Clustering Analysis to Improve the Reliability and Maintainability of Wind Turbines with Self-Organizing Map Neural Network

ZAFAR HAMEED* and KESHENG WANG
Norwegian University of Science and Technology (NTNU), Trondheim, NORWAY
(Received on May 03, 2012, revised on September 03, 2012)

Abstract: Reliability and maintainability of wind turbines are posing new challenges and issues due to advancements in new and sophisticated technologies. This has necessitated development of novel, efficient, and cost-effective strategies for enhancing the availability for power output and operational life. There are certain ways to achieve such objectives where every approach has certain pros and cons. One possible technique is to use the wind speed and power output data for exploring the behavioral similarities of different wind turbines. Based on the similarity measures, a group of turbines may appear together in the form of clusters. This is accomplished by working with the vast piles of data which are pre-processed by using statistical time domain features to provide input to a self-organizing map (SOM) neural network. Based on the clustering results, operational and maintenance strategies are planned for a group of wind turbines in contrast to doing the same work for individual ones. A case study is presented where it has been shown how the information obtained from the clustering analysis would be used for predicting the power output and then developing the optimal operational and maintenance strategies in an integrated manner.

Keywords: Wind Speed, power output, features, cluster, self-organizing map (SOM)

1. Introduction

Reliability and maintainability of wind turbines are of vital importance to attract potential investors in order to secure further growth and development. A higher level of reliability and maintainability will mean maximum availability of production with the least possible downtime due to failure. The enhancement of production availability with minimum downtime is a considerable task especially for the offshore wind industry. There are numerous challenges and issues that can hinder high levels of availability for production. Few of them are: suitability of the selected wind turbine for the designated site; what kind of condition monitoring system (CMS) is installed; how the information coming from the sensors is analyzed to understand the operational state in real time; what type of measures have been undertaken to address the causes of failure; and the level of priority that is to be allocated to develop optimal inspection and repair strategies. All these aspects are important so that they can be analyzed and treated in a holistic way to enhance the life of the asset which is the wind turbine.

In the last two decades work has been focused on the development of state-of-the-art CMS and operational & maintenance (O&M) strategies. Attempts have been made to address different aspects of wind turbines to increase their reliability and maintainability levels. It has been proposed by [1] how to increase the reliability of wind turbines by using event sequence analysis, fault tree analysis and other structural reliability methods to improve their design. Different aspects regarding the optimization of O&M strategies for wind farms including access and energy output were covered by [2-5]. The impact of installation on the overall performance of wind turbines was discussed in detail by [6] where different aspects were highlighted. Quantification and optimization of maintenance
were described and principles for undertaking these processes were elaborated by [7]. The concept of self-maintenance machines [8] was put forth by introducing redundancy based on failure analysis to avoid the system having a sudden shutdown. Different aspects of condition monitoring system, its architecture, and fault detection algorithms for wind turbines assemblies were summarized and discussed in detail by [9,10]. Also [7, 11] studied how to use the total productive maintenance, risk-based inspection, and reliability-centered maintenance (RCM) and suggested how to conduct the maintenance optimization of the wind turbines. A number of ways were suggested by [12, 13] to lower the O&M costs in an optimal way and put forth ideas about how to standardize the failure and maintenance data. Failure mode and effect analysis were discussed by [14] where different failure causes were identified and then levels of criticality were outlined. Challenges in the reliability data collection, condition based maintenance strategies and grouping strategies were addressed by [15-17].

In short, there have been a number of approaches and methods proposed to address the reliability and maintenance issues for wind turbines. However it is worth mentioning that most of the ideas were based on a single wind turbine which might be extended to the group of installed wind turbines. The reason behind this rationale is that wind turbines are installed in the form of arrays to exploit maximum wind potential for a given site. There is a small possibility that land-based wind turbines might be installed alone but in the case of offshore locations, such likelihood of this happening is remote. Thus the situation demands that the analysis of wind turbines should be carried out at the wind farm level in order to address the reliability and maintainability challenges and issues on a broader scale. Consequently, in the current work an attempt has been made to work with the wind turbines in the form of groups which are part of a wind farm. The performance of wind turbines was investigated based on their exposure to the wind speed and the resultant power output. It is expected that working in this fashion will yield certain patterns which could provide a sound basis to cluster the various wind turbines together. Based on such patterns, a group of wind turbines is to be demarcated to plan and implement their O&M strategies in an integrated and coherent way rather than working on single wind turbines in isolation.

The need for the current work is discussed in Section 2. The objectives are outlined in Section 3 where the implications and relevance of the proposed idea are specified. Section 4 discusses modeling aspects of statistical time domain features and the SOM neural network. The outline of the proposed method is illustrated in Section 5. Results and discussions are given in Section 6. Finally conclusions are drawn in Section 7.

2. Need

Wind turbines are installed in the form of an array which could be of a regular or irregular shape as shown in Figure 1. Based on the given layout of wind turbines, there are possibilities that certain units would behave similarly due to the exposure to comparable wind loads, failure rates, design margins, controller type, and by the manufacturer. For example, as wind speed and direction show similar patterns for a given geographic location (latitude and longitude), it is expected that wind turbines installed in that region might behave differently compared to others in a nearby site. Moreover it is also possible that for a given grid of latitude and longitude, certain patterns would emerge for defining a particular cluster of wind turbines. Examples of clusters have been shown in Figure 1. These factors motivate one to investigate wind turbines on a macro scale at the wind farm level.
Operational and maintenance costs are high compared to the total investment costs. Therefore any improvement at the operational level will enhance the likelihood of profits as well as making the wind power industry lucrative for investors. When wind turbines are installed at a given site, the question arises about how to plan and execute the necessary inspection and repair strategies. The answer to this is not easy because it is a demanding task to decide whether to conduct inspection on the wind turbines or respond only when failure occurs. One of the possible alternatives might be to collect the necessary information relevant to the wind turbines like the wind speed data, power output data and if possible failure data. Then it is necessary to analyze the collected information to search for certain patterns in the data which may tell us which turbines are behaving similarly. It is worth mentioning that a wind farm may have turbines from different manufacturers, design types, and controller types. It might not be a good option to simply divide the wind turbines based on these types and plan the operational activities on this basis. For example there are chances that wind turbines at the outer edges of the farm might behave differently than the turbines which are right in the middle. This is because the central wind turbines are more affected by the wake effect than those peripheral turbines. Such facts indicate the need for collection of real-time information from every wind turbine, analysis of the data, followed by decisions about potential clusters. Moreover, the demarcation of wind turbines based on the wake effect need computational fluid dynamic analysis which is not discussed in detail here due to scope of the current work.

3. Objectives

There are number of objectives needed to carry out the current work which include:

- addressing the reliability and maintainability issues in a holistic way
- suggesting ways of how to pre-process the vast amounts of data being used as input
- demonstrating the use of neural networks in the clustering analysis and power predictions
- understanding the behaviour of wind turbines at the wind farm level
- analyzing how reliability changes with geographic locations
- planning and optimizing the inspection and repair strategies in an optimal fashion

Based on the clustering analysis, challenges and issues pertaining to the O&M can be addressed in an optimal way. The question arises as to how? When the turbines are clustered together based on the real time data, how is it possible to group certain repair and inspection activities together by sharing the set-up costs. For more details on this activity see [16-18]. When the groups of turbines have been determined, then there are
possibilities to address the external factors like access, logistics and transportation in the development and implementation of inspection and repair strategies. When the work is to be planned for a cluster of wind turbines, it is expected to reduce the frequency of sudden failures significantly. The reason behind this claim is the rationale that the power output of the given wind turbines reveals the real ongoing state of operational health. If the power output is not at a satisfactory level, then it means that the system has started to move into an abnormal state. Based on this indication, proper corrective measures should be undertaken so that the system will be brought back in a normal state. Moreover while planning and conducting work in the form of clusters, it will be necessary to address the safety and risk issues which are present while optimizing the maintenance activities, see [19-20].

4. Modeling Framework

There are number of sensors installed on wind turbines to carry out different types of measurements like wind speed, wind direction, temperature, pressure and condition monitoring of the components. It is of vital importance to decide what type of information is needed and how this information is to be extracted from the considerable amounts of data. It is important to define the input features based on the available information. Normally the raw data is available in hundreds of megabytes which might contain white noise and other external disturbances. It often becomes imperative to extract the valuable features by reducing the dimensionality of the data. The task of feature extraction in data manipulation provides the basis of reliable, efficient, and accurate results.

In the current work, the vast amounts of data were at hand and the processing was posing a great challenge to extract valuable information. To overcome this problem, it was decided to use statistical time domain features to preprocess the available information. For this purpose four statistical moments were selected, namely root mean square (RMS), variance, skewness, and kurtosis. Suppose we have \( X \) random samples with a sample size of \( N \) and \( E[X] \) is the expected value, then we can define the selected statistical moments as:

\[
RMS(X) = \sqrt{\frac{\sum X^2}{N}}
\]

(1)

\[
Variance(X) = E[X^2] - (E[X])^2
\]

(2)

\[
Skewness(X) = \frac{E[(X-E[X])^3]}{(\sigma[X])^3}
\]

(3)

\[
Kurtosis(X) = \frac{E[(X-E[X])^4]}{(\sigma[X])^4}
\]

(4)

In Equations (3) and (4), \( \sigma \) is the standard deviation, which is the square root of variance \( \sigma[X] = \sqrt{\text{Variance}(X)} \) given in Equation (2). These four moments are used as preprocessors to extract the valuable information in the form of features from each bin of the data. Consequently this preprocessed data is to be used as input to the self-organizing map (SOM) neural network for clustering analysis.

The SOM neural network defines an ordered mapping, a kind of projection from a set of given data items onto a regular, usually two-dimensional grid which could be rectangular, circular or any other shape. According to [21], a data item is mapped into the node whose model is most similar to the data item. For this purpose, we assume if \( X_j \) is representing the input data, (after preprocessing as in Equations 1 to 4), there are \( m \) neurons in the competitive layer, \( U_i \) is the \( i^{th} \) neuron, \( W_{ij} \) is the weight value based on the connection of competitive neuron \( U_i \) with the input data, the output value \( D_i \) of the neuron \( U_i \) is computed by as an example by using Euclidean distance [22]:

In Equations (3) and (4), \( \sigma \) is the standard deviation, which is the square root of variance \( \sigma[X] = \sqrt{\text{Variance}(X)} \) given in Equation (2). These four moments are used as preprocessors to extract the valuable information in the form of features from each bin of the data. Consequently this preprocessed data is to be used as input to the self-organizing map (SOM) neural network for clustering analysis.

The SOM neural network defines an ordered mapping, a kind of projection from a set of given data items onto a regular, usually two-dimensional grid which could be rectangular, circular or any other shape. According to [21], a data item is mapped into the node whose model is most similar to the data item. For this purpose, we assume if \( X_j \) is representing the input data, (after preprocessing as in Equations 1 to 4), there are \( m \) neurons in the competitive layer, \( U_i \) is the \( i^{th} \) neuron, \( W_{ij} \) is the weight value based on the connection of competitive neuron \( U_i \) with the input data, the output value \( D_i \) of the neuron \( U_i \) is computed by as an example by using Euclidean distance [22]:
$D_i = \|X_j - W_i\| = \sqrt{\sum_{j=1}^{n}(x_j - w_{ij})^2}$  \hspace{1cm} (5)

The competitive unit with the shortest distance is that which is closest to the current input pattern and therefore it is representative and called the winning node. After the identification of the winning node, the weights of the neighboring neurons are corrected. This means that the closer the neurons are to the winning node, the larger the change in their weights. The corrections in weights are scaled down which is called the Topology Dependent Function:

$$a(*) = a(d_c - d_i) = a(d_c - d_i)$$  \hspace{1cm} (6)

Here, $(d_c - d_i)$ is the topological distance between the central neuron “c” (winning neuron) and the current neuron “i”, while the extent of the simulation depends on the function $a(*)$. In order to improve the learning procedure, the topological dependent function is multiplied by another monotonically decreasing function, being given in Equation (7).

$$\eta(t) = (a_{\text{max}} - a_{\text{min}}) \times \left(\frac{t_{\text{max}} - t}{t_{\text{max}} - t_{\text{min}}}\right) + a_{\text{min}}$$  \hspace{1cm} (7)

where, $t_{\text{max}}$ is the maximum number of epochs, the two constants $a_{\text{max}}$ and $a_{\text{min}}$ define the upper and lower limits between which the correction $\eta(t)$ is decreasing from the beginning to the end of the training. The correction of weights of the $i^{th}$ neurons depends on the criterion used to select the central neuron “c”. The weights of the SOM are updated as:

$$\Delta w_{ij} = \eta(t).a(d_c - d_i)(x_j - w_{ij})$$  \hspace{1cm} (8)

$$w_{ij}^{\text{new}} = w_{ij}^{\text{old}} + \Delta w_{ij}$$

where, $x_j$ is the component of input vector $X_j$. More in-depth knowledge about the SOM neural network along with its training algorithm can be found in [21, 22].

5. Method

An outline of the proposed method is shown in Figure 2. Sensor data was collected and compiled before the pre-processing. In the pre-processing phase, wind speed and power output samples from the selected wind turbines were evaluated. This process was considered important to check the presence of any outliers in the data. If the data were replete with unrealistic information, then even a small fraction of such information might have a detrimental impact on the final results. For example, if the maximum power output of the selected wind turbine is 1MW (Mega Watt) but the data samples were available with {1.1, 1.2 …}MW, then it means the samples were not truly representative for carrying out the analysis.

After removal of the outliers from the data samples, the next stage is to bin the data as in Equations (1 to 4), before they are used as input to the SOM neural network. Each feature of wind speed and corresponding power output from the selected wind turbines was given as input to train the map of the SOM neural network. For example: RMS for a given turbine $k$ was compiled with the RMS of other $k+1$ wind turbines as an input. The same process was repeated for other features i.e., variance, skewness and kurtosis. The clustering pattern of each feature was evaluated separately and then the map with best boundaries among the wind turbines was selected.

After determining the size of clusters, the next task is to develop the operational strategy for a group of wind turbines based on the clustering analysis and the expected power output. An integrated approach regarding how to develop this kind of operational strategy has been outlined in Figure 3. Information from the clustering analysis is used to predict the power output of the clustered wind turbines. Moreover the components from the clustered wind turbines will be selected to optimize their inspection and maintenance intervals for accomplishing the maintenance optimization process.
The maintenance optimization process will be completed based on the failure data. The components with a higher expected rate of failures would be prioritized. Then necessary preventive and corrective measures would be adopted to cope with their higher failure rates. In the current approach, the maintenance optimization process, its planning and implementation would go side by side with the failure analysis, clustering analysis and power prediction. When the necessary optimization strategy has been selected to optimize the inspection and maintenance intervals of the components, then the necessary resource allocation process will be carried out before its implementation. The planning and implementation phases will be completed bearing in mind the root cause analysis of the failures and the likely impact of certain causes of failures on the overall performance of the system. The approach presented in Figure 3 is called the integrated operational strategy because it is dealing with clustering analysis, power prediction, maintenance optimization, and failure analysis all at the same time.

6. Results and Discussions

The data of minimum, mean, and maximum wind speed and corresponding power output were collected so that every sample was gathered after 15 minutes around the clock. In this way a vast amount of data was collected and one of the sampled pieces of raw data is shown at the top of Figure 4. One can see that it is very hard to use this raw information as it is, an alternative way is to pre-process the information. The data were processed with statistical time domain features with bin sizes of n=25 and n=50. One processed piece of information for RMS, variance, skewness, and kurtosis for a wind turbine with bin size of 50 is illustrated below.

The data were processed in two bins of size 25 and 50 to evaluate their influence on the clusters. Moreover it was decided to use power output separately as input and then
combine it with wind speed to see how the cluster pattern emerges in both scenarios. In this way, four possibilities of potential clusters were generated.

Eight different wind turbines at zero sea level (represented by coloured hexagons in Table 1 were selected for clustering purposes. The clusters were obtained by using Matlab software. All four features (RMS, variance, skewness, and kurtosis) were tested with “only power output with bin size of 25 and 50” and “with wind speed and power output with bin size of 25 and 50” and the results are shown in Figure 5. Evaluating Figure 5(a) to (c) with features RMS, variance and skewness, certain patterns were not emerging which might tell us about some similarity among a given number of wind turbines. It is interesting to note that in Figure 5(d) with the kurtosis feature, some clear patterns started to emerge. In the first subfigure 5(d) with “only power output with 25 bins”, turbines of E32B, E33A, E40, V25 and WW2700 were clustered together with some chances of possible mixing with other turbines. In the same subfigure, turbines E17 and E32A were mixing with each other. It means that there were some similarities. In the second subfigure 5(d), when the bin size was doubled i.e., from 25 to 50 with “only power output”, the patterns were not emerging as clearly except for turbine E32A. In the third subfigure 5(d), with “wind speed and power output with bin size 25”, turbines E32A, E32B, E33A, E40, V25, and WW2700 were clearly making clusters while the others were mixing with each other. In the last subfigure 5(d), there was a strong overlapping between turbines E32A and E40 while E33A was doing the same with E32A. Again in this figure, turbines V25 and WW2700 were not overlapping with other turbines.

Figure 4: Raw and Processed Data for Minimum, Mean and Maximum Wind Speed and Power Output for One Wind Turbine
Table 1: Description of Features and Wind Turbines in Clusters with Designated Hexagons

<table>
<thead>
<tr>
<th>Features</th>
<th>Bin Size (n)</th>
<th>Detail of Turbines in Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. Wind Speed/Power Output</td>
<td>25/50</td>
<td>E17, E32A, E32B, E33A</td>
</tr>
<tr>
<td>Mean Wind Speed/Power Output</td>
<td>25/50</td>
<td>E33B, E40, V25, WW2700</td>
</tr>
<tr>
<td>Max. Wind Speed/Power Output</td>
<td>25/50</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5: Clustering of Wind Turbines through Chosen Features (a) RMS (b) Variance (c) Skewness (d) Kurtosis

Based on the patterns that emerged in Figure 5(d), it can be concluded that the kurtosis feature of “only power output” and “wind speed with power output”, with bin size of 25, was useful to explore the possible similarities between the eight selected wind turbines. When the bin size was doubled, the boundaries started overlapping with each other. Moreover kurtosis (being the fourth moment) was useful to capture the existing patterns among the wind turbines in contrast to the first three moments (RMS, variance, skewness).

The selected eight turbines were drawn on the map with their given latitude and longitude, as shown in Figure 6. Turbines E32B and E33A were too close to each other, so that these appeared as one dot on the map. Turbines V25 and E40 were close to each other but they were not clustering together, see Figure 5(d). This means that their behaviour was different owing to their different manufacturers in spite of close vicinity. The same is true for WW2700 which was making a separate cluster although it was slightly mixing with turbine E32A. It might be possible that these two turbines were exposed to the same wind loads when the direction of wind was shown with arrow of “wind direction 1”. Some behavioural similarity was noticed between turbines V25 and E33A which might have developed over the time due to “wind direction 2” when both units were exposed to the wind loads from the same direction. Turbines E17, E32A and E33A were overlapping with each other and interestingly these three machines were
located in close proximity. The manufacturers of turbines WW 2700 and V25 were
different from the E type machines. Turbine V25 did not show any similarity with the E
type turbines although WW 2700 was correlated with E32A to some extent.

Figure 6: Location of Wind Turbines with Possible Wind Loads and Barriers

It is important to know how to cross the barrier (shown in Figure 6) for planning the
operational strategies of V25 and E40 turbines with other machines. From the subfigure
5(d), it was clear that turbines E32A and E40 were behaving similarly. This means that in
one way the barrier could be crossed to group the inspection and maintenance activities of
these two turbines together. Moreover turbine V25 showed the same tendency as E32B
and E33A. All three turbines could also be grouped together for planning their reliability
and maintenance related tasks. In short, there are a number of other possibilities which
could be based on the clustering information for developing optimal operational strategies.
Based on Figure 5 and 6, some of the potential clusters of wind turbines are summarized
in Table 2.

Table 2: Summary for Possible Clusters of Wind Turbines

<table>
<thead>
<tr>
<th>Cluster #</th>
<th>Potential clusters of wind turbines</th>
<th>Behavioral similarly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>E17, E32A, E33A</td>
<td>Strong</td>
</tr>
<tr>
<td>2</td>
<td>E32B, E33A, V25</td>
<td>Moderate</td>
</tr>
<tr>
<td>3</td>
<td>E32A, E40</td>
<td>Moderate</td>
</tr>
<tr>
<td>4</td>
<td>E32A, WW2700</td>
<td>Fair</td>
</tr>
</tbody>
</table>

From the clustering analysis, it was found out that the turbines E17, E32A and
E33A were strongly behaving similarly as shown in Table 2. This finding implies that
there are high chances for those wind turbines to be clustered which are located in the
close proximity to each other. There is another interesting point regarding the location of
turbines E32B and E33A, they appear to be too close to each other in that they appear as
one dot in the map as shown in Figure 6. These two turbines along with V25 were
showing similar patterns as shown at cluster # 2 in Table 2. One possible explanation for
such kind of behaviour would be that the locations as well as wind direction are important factors in determining the clusters of wind turbines. The turbines E32B and E33A are from the same manufacturer but that of V25 is from another manufacturer but all three are showing behavioural similarity. The inclusion of V25 in cluster #2 might be attributed to the fact that all three wind turbines are exposed to wind direction 2 which influenced these machines. It is equally important to assess how the wind turbines react to the wind loads in producing the power output. This means that although turbines E32B and E33A are too close to each other and failure rates are nearly same in Figure 8(c), their reaction to wind loads is quite different in terms of producing the power output. The inclusion of turbines E32A and E40 at cluster #3 in Table 2 shows that manufacturer’s type has influence in determining the clustering patterns to some extent even though turbines are installed little far from each other. On the other hand, cluster # 4 in Table 2 shows the influence of close proximity in determining the clusters even though the manufacturers are of different types in case of turbines E32A and WW2700. One fact is evident; it is relatively hard to define the clear boundaries among the clusters of wind turbines. The reason behind the rationale is that the correlation of one wind turbine with others varies from strong to fair as shown in Table 2. This means that the clustering pattern of turbines may change over the period of time due to a number of factors like wind loads, wind direction, local topographical variations, and intrinsic reliability of the wind turbines.

The relationship among the clustering patterns and failure rates of the wind turbine was also investigated. For this purpose, failure rates along with root causes of the selected wind turbines are shown in Figure 7. Bearing in mind the failure rates of the wind turbines, it was found that wear was one of the dominant cause of failure as shown in Figure 7(a) and (b). A comparison for the total failure rate of all the wind turbines was carried out in Figure 7(c). Turbines E32A had the highest failure rate while that of turbines E40 and V25 were at the lowest ebb. From the clustering analysis, it was found out that the turbines E17, E32A and E33A had strong behavioral similarity at cluster # 1 in Table 2 but their failure rates were different from each other as shown in Figure 7(c). This means that their close proximity, wind loads, and wind direction had played an important role in clustering them. At cluster #2 in Table 2, wind turbines E32B and E33A had close proximity and similar failure rates but the third wind turbine V25 was located little far apart with different failure rates compared to the other two. It is worth mentioning that the turbines at cluster # 2 are exposed to same wind direction (wind direction 2 in Figure 6) and it might have played a pivotal role in determining the clusters even though manufacturer of V25 turbine is different from the other two in the same group. Turbines E32A and WW2700 were showing some similarity at cluster # 4 in spite of a different manufacturer. While analysing the clustering pattern, it is interesting to note that turbine E32A is appearing at cluster #1, 3 and 4 with varying degrees of behavioral similarity. It might be attributed to the fact that it is surrounded by number of wind turbines as shown in Figure 6.

After analyzing Figures 6 and 7 and Table 2, it might be concluded that there are a number of factors which are influencing the clustering patterns over the period of time. Some of the important ones are highlighted in Figure 8.

It is important to show how the information obtained from clustering analysis would provide underpinning to develop optimal O&M strategies for the clustered wind turbines. For this purpose, the turbines E17, E32A (type E32), and E33A (type E33) at cluster # 1 from Table 2 were selected. The power prediction of the clustered wind turbines was carried out by using a self-back propagation (SBP) neural network within Matlab software. For more knowledge about SBP neural network and its applications, see [22].
The network was trained with a huge amount of wind power data. Based on the trained network, the power was predicted for the next 480 days. The predicted power output and percentage error (in predicted and actual power output) are shown at the top and bottom in Figure 9 respectively. The percentage error for turbines E17 and E32 was less than 1% and the same was less than 4 % in turbine E33. The percentage error in predicting the power output was fairly reasonable in all three clustered wind turbines.
Figure 8: Important Influencing Factors on Determining the Wind Turbine Clusters

It has been illustrated in Figure 10 how the proposed integrated operational strategy, being proposed in Figure 3, has to be implemented. Information from the clustered turbines, E17, E32A, and E33A were used to predict the power output shown in Figure 9. The profile of predicted power output was used to identify the time periods when its expected value would be low to plan and carry out the inspection and maintenance tasks. For accomplishing the maintenance optimization process, the dynamic grouping strategy was chosen as one means to develop optimal operational strategy. Further details regarding the different aspects of grouping strategies can be found in [16-18, 23-26].

In the current dynamic grouping strategy, the components were selected based on their age and condition based models for evaluating their likelihood to be a part of a certain group of activities. The prime objective in grouping inspection and/or maintenance is to share the set-up costs with as many activities as possible to lower the operational costs. It has been proposed in [27] how to group different inspection and maintenance activities together from three wind turbines where eight components (related with gearbox, its cooling system, main bearing, and couplings) from each of wind turbine were selected.

Groups 1 to 4 had 11, 15, 17 and 16 activities which had to be conducted at the optimal interval of each group at 3.54, 6.64, 9.93, and 13.20 months. In Figure 10, all these four groups of activities were shown in pink circles with their respective group number. The optimal intervals were converted into days as the predicted power output was also in days. The expected power output of the clustered wind turbines were mapped on the planned execution time of the four groups of activities which were 107, 201, 302, and 401 days respectively.

From the predicted power output, it was found out that the execution time of groups 1 and 2 occurred when the expected power output would be high. On the other hand the execution time of groups 3 and 4 occurred at a time when the expected power output would be relatively low. In such a scenario, what should the operational strategy be? To address this issue, it is proposed to conduct the root and cause analysis of the failures to evaluate their likely impact on the overall availability of the system. One such study was conducted by [28] where the causes of failures and their impact on the system were analysed as shown at the bottom of Figure 10.
There are a number of possibilities regarding how to plan and implement the four groups of activities. In the first scenario, the execution of the first group of activities was due but expected power was high. In such a case it should be investigated what was the likely root cause of the failures of components which were part of a group 1. If the root cause had a high chance to shut down the system, then it is important to implement group 1 irrespective of the high expected power output. In the second scenario, the activities from group 1 and the cause of failure of components did not play a role in shutting down the system but these causes had “other consequences”. It was then necessary to delay the execution of group 1 maintenance, to wait for the periods when expected power output would be low. The same is the case with group 2. The execution of groups 3 and 4 occurred when the expected power output was relatively less, so their planning and implementation was not as complex as for groups 1 and 2. By integrating the expected power output and the execution time of the maintenance-related activities, proposed cost-effective operational strategies can be planned to reduce the per unit energy price being generated from wind turbines.

7. Conclusion

An approach was presented to cluster wind turbines in a wind farm based on an SOM neural network. It has been demonstrated how to pre-process the data by using statistical time domain features as input to the selected neural network.

Four moments namely RMS, variance, skewness and kurtosis were selected. For the example provided it was found that kurtosis had given satisfactory results in determining the cluster boundaries. It was also found that kurtosis results with bin size 25 was quite reasonable, which showed that the bin size should be not be too large.

The turbines which were from the same manufacturer behaved similarly in most of the cases. It was found that the turbines from the same manufacturer, which were installed in close vicinity, had clustered together. This was an interesting result which showed that in future wind turbines could be installed in this fashion to exploit this similarity measure. Moreover, the clustering of wind turbines from the same manufacturer was not a hard and fast rule. Some similarity was found among the wind turbines which were made from different manufacturers due to wind loads, wind direction, and site location.
It was demonstrated how the operational and maintenance tasks could be scheduled based on the information obtained from the cluster analysis. Power prediction was carried out by using the SBP neural network for clustered wind turbines. An integrated approach was presented to develop an integrated operational and maintenance strategy to enhance the reliability and availability of the wind turbines.

Future work includes using higher moments and comparing their influence on the cluster boundaries. Moreover it is planned to employ other standard pattern search techniques.
approaches and compare them with the current approach for evaluating the viability and accuracy.

References


Zafar Hameed was born in Punjab, Pakistan. He obtained his B.Sc. and M.Sc. degrees in Mechanical Engineering from Pakistan and South Korea respectively. Currently he is pursuing his Ph.D. degree in RAMS (Reliability, Availability, Maintainability and Safety) analysis for offshore wind turbines from Norwegian University of Science and Technology (NTNU) Trondheim Norway. His main areas of interests are RAMS Engineering, CAD Modeling, FEM Analysis and Dynamic Simulations of mechanical systems.

Kesheng Wang holds a Ph.D. in production engineering from the Norwegian University of Science and Technology (NTNU), Norway. In 1993 he was appointed Professor at the Department of Production and Quality Engineering (IPK), NTNU. He is a director of the Knowledge Discovery Laboratory (KDL) at IPK, NTNU at present. He was elected member of the Norwegian Academy of Technological Sciences in 2006. He is also an active researcher and serves as a technical adviser in SINTEF. He has published 15 books and over 180 technical peer reviewed papers in international journals and conferences. Prof. Wang’s current areas of interest are intelligent manufacturing systems, applied computational intelligence, data mining and knowledge discovery, swarm intelligence, condition-based maintenance, structured light systems for 3D measurements, RFID for supply chain management. E-mail: Kesheng.wang@ntnu.no