International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 0925

WIND TURBINE PITCH SYSTEM FAILURE PREDICTION BASED ON SOM AND RBF NEURAL NETWORK

Dan Xu, John Soraghan, Alasdair McDonald

Department of Electronic and Electrical Engineering,

University of Strathclyde,

Glasgow, UK

ABSTRACT

Wind turbine pitch systems play an important role in wind energy extraction. Such systems are known to suffer high failure rates resulting in long downtime maintenance. Consequently, industries wish to utilize failure prognostic techniques to schedule effective maintenance in order to reduce losses. A cost-effective method that analyses existing data collected from the built-in sensors in the wind turbine (WT), such as the supervisory control and data acquisition (SCADA) data is presented in this paper. Information such as wind speed, power generation and subsystem measurements are collected for every 1 second resulting in large datasets. By combining a SOM clustering technique and radial basis function neural network (RBFNN), patterns are revealed while reducing redundant information. The performance of the WT failure prediction system was evaluated and tested using 37 sensors from SCADA data for a single WT in Levenmouth, Scotland. A failure state classification accuracy of between 96%-99% was observed.

Keywords: wind energy, self-organizing map, k-means, radial basis function neural network, clustering, classification, etc.

1. INTRODUCTION

Unscheduled maintenance of unexpected failures usually has a great impact on the operation and maintenance (O&M) cost [1]. To minimize the cost, efficient failure prediction is desired. As wind turbines (WTs) have expanded to the offshore because of the persistent wind source the operational condition is more severe than on the land. Salt particles contained in the wind and seawater may accelerate corrosion and aging problems [2]. The accessibility is also limited by the distance between site and shore due to the sea wave is largely influenced by weather conditions. For safety reasons, human intervention should be managed effectively.

An effective pitch system is important in energy production performance. By adjusting the angle between blades and wind, energy from wind can be extracted with optimum efficiency. By surveying several wind farms in Europe with an 8-year period, work [3] shows that the pitch system also occupying the highest failure rate, although there was no requirement of major replacement such as generator or gearbox.

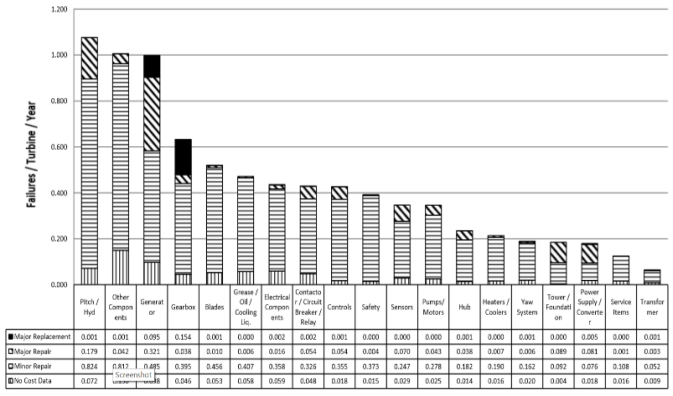

Fig 1 Sub-system Components Failure Rates and Their Cost Categories [3]

There are varied Non-Destructive Techniques (NDT) in the condition monitoring (CM) category, such as visual

Selection and peer-review under responsibility of the scientific committee of the 11th Int. Conf. on Applied Energy (ICAE2019).

inspection and vibration analysis [4]. However, the former requires in situ examination, which consumes time. The latter may require additional sensors and the signals may be affected by the noisy background. A model-based approach [5] was developed, which took into account different motor behaviours and processed the current signals in the frequency domain. Work [6] has developed an algorithm-based method for pitch system diagnosis. By combining regression and classification models, an online failure detection technique based on Supervisory Control and Data Acquisition (SCADA) data was developed in [7].

According to its economical characteristic, the sensor information in the SCADA data was considered. The SCADA is a data collection system that records the signals through pre-built sensors in the WT. Comprehensive information such as subsystem measurements, weather parameters, health states and alarms are usually recorded as 10-min averaged value for CM purposes. In our research, a higher frequency data (sampled at 1 Hz) was utilized. Intuitively, higher frequency data is more easily affected by noise and the noise can be mitigated by an averaging process. However, this process may lead to information loss. Study [8] shows that by using higher frequency SCADA data, a reliable power curve can be estimated from a small number of data for WT monitoring because of an increase of information. Study [9] provides evidence that accurate power curve model and improved detection abilities can be achieved through higher sampling data. This is because it provides a deeper understanding of WT dynamic behaviours. Although wind power curve can indicate various failures by analysing different deviations to the normal values, the modelling process can be affected by inconsistent data and data quality [20].

Higher sampling rate will result in a large dataset. To reduce the computational complexity, the information of data is compressed and clustered by PCA and SOM. With learning the clustering properties, the pitch faulty state should be well predicted through the RBFNN. The structure of this paper is organized as follows. Explanations of the specification of the involved WT are summarized in section 2. Theory and methodology are included in Section 3. Results with analysis are shown in Section 4. The summary of work is in Section 5.

2. THE LEVENMOUTH WIND TURBINE

2.1 Wind turbine specifications

In our research, the data used for simulation was collected from the Levenmouth (Scotland) Demonstration Turbine and provided by the Offshore Renewable Energy (ORE) Catapult company. The turbine is 7 MW with 574 SCADA signals. According to the 2-year downtime analysis provided by the company, the pitch system was found to generate frequent failure alarms. With both long downtime hours and high alarm counts, failure prediction needs to be emphasized for the pitch system. Unlike commercial WTs, this turbine is a prototype and is used for academic testing purposes. There are several limitations should be considered due to its unique design. Firstly, the training and benchmark data are lacked for simulation. Secondly, for the reason of its design defect, noise is produced when the WT is facing a certain direction. This will lead to an artificial shutdown in preventing noise pollution to the nearby residents. Research activities and maintenance are also executed frequently, which results in a discontinuous operation mode. Since human intervention may trigger false alarms, the quality of SCADA data then becomes an issue for analysis purpose [10].

2.2 SCADA dataset and the pitch states

Before pre-processing, there are 51 variables from pitch system and general parameters such as power, wind speed and temperatures. However, the sources are uncleared for some variables. Such variables and variables with logic units are removed together with missing data, and 38 variable left. The pitch states and the number of training samples are shown in Table 1. Our target is to be able to identify the failures (state 3).

Current pitch state	Code	Candidates
PITCH_STATE_STANDBY	0	18941
PITCH_STATE_ENABLED	1	28517
PITCH_STATE_BATTERY_TEST	2	0
PITCH_STATE_FAULTED	3	29405
PITCH_STATE_OTHER	4	699

Table 1 Pitch States

3. WT PITCH SYSTEM FAILURE PREDICTION SYSTEM

The system is consisting of two stages: information compression stage and neural network prediction stage. The first stage is combined with PCA, SOM and k-means. PCA is used for dimensionality reduction, SOM is applied for compressing data in its codebook and clustering. Utilizing k-means is for clustering refinement, and also important for the RBF neuron centres initialization and spread constant determination. In the second stage, the RBFNN is trained and tested for classification. The system procedure is illustrated in Fig 2.

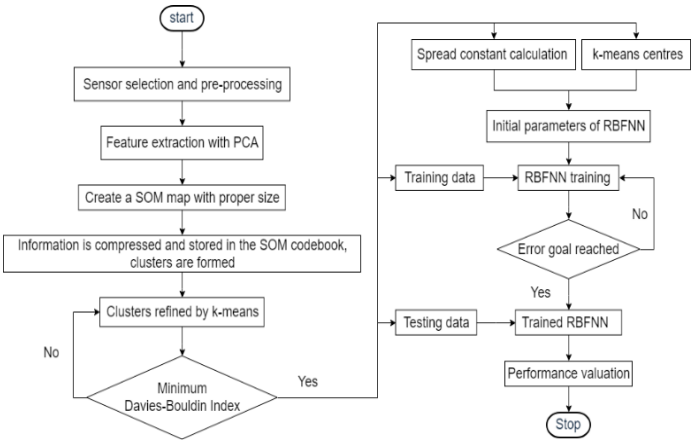

Fig 2 System Flowchart

3.1 Stage 1: Information Compression

3.1.1 PCA and SOM

Since the dataset is large, a time range from 10 minutes before to 10 minutes after the failure occurrence was selected. The reason for choosing a time window is that data integrity may be affected by human activities and some failures only last a few seconds, also the occurring time was not always consistent with the record. The selected data contains three normal operation states (state 0, state 1, state 4) and a faulty state (state 3). There are no values for state 2 and a small amount of state 4 data within the selected data, which becomes one reason for poor classification in the results.

PCA is a popular technique used for removing redundant information. There are 7 principal components (PCs) containing more than 95% of most useful information are obtained over this step. They are then fed into SOM (Fig 3).

SOM is an unsupervised neural network based on the competitive learning principle. The advantages of applying SOM is that is has simple interpretation and visualisation [19]. The learning process can be thought of as a net of similar neurons spanning among the input data. The map size is determined as :

$$M \approx 5\sqrt{S} \quad (1)$$

where M is the total number of neurons, and S is the number of total training samples. This is slightly different from the formula proposed in work [11], where S is referring to the number of rows in the training data matrix.

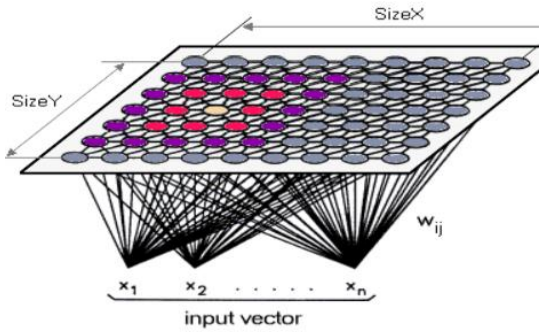

Fig 3 Illustration of SOM [12]

Depending on different randomly distributed parameters, neurons will respond to the input data with similar characteristics. The neuron that receives the largest input is the "winner" and will reach its maximum value. While others within the cluster will be set to minimized values. There is a certain amount of weight of each winning neuron, during the learning process, the system will shift the weight position to reach an equilibrium state. Finally, similar input data will be gathered together and remain efficient distance to separate different patterns [13]. Those winning neurons are also called the Best Matching Unit (BMU). Since the quantity of BMUs is sufficiently smaller than inputs but has learned most "structural" features, information is then compressed.

3.1.2 RBF Parameters Determination

One of the widely used RBF is the Gaussian function:

$$H = e^{-||x-\mu||^2/2\sigma^2} \quad (2)$$

where x is represented as input vectors, μ and σ^2 are the centre and variance (spread constant) of the hidden units. Improper selection of spread constant may require more neurons for fitting either a fast-changing or a smooth function and further increase the computational complexity [14]. In order to simplify computation, the clustering properties are used to calculate the spread constant.

The k-means algorithm is applied after the SOM, for clustering improvement. Its centroids are used as the Gaussian neuron centres. The integer k is treated as an unknown variable in our research. By setting the maximum number of 10 iterations, the optimum number of clusters is obtained according to the Davies-Boulding Index (DBI) [15]. The DBI computes the ratio between intra-cluster similarity and inter-cluster difference, and the best classification result can be said to be done with the lowest value. The spread constant is then calculated according to equation (9):

$$\sigma_{\text{spread}} = d_{max}/\sqrt{k} \quad (3)$$

where d_{max} is the maximum distance between two centres, k is the number of clusters.

3.2 Stage 2: RBFNN Training

The RBFNN is a three-layer feedforward neural network, it consists of input layer, hidden layer and output layer. There is one hidden layer that contains nonlinear units, who locally respond to the input data with a minimum distance. It is one of the favoured techniques for function approximation because of its robust tolerance to input noise [16]. Since there is only one single hidden layer, an RBFNN can be trained with less complexity compared to multilayer neural networks.

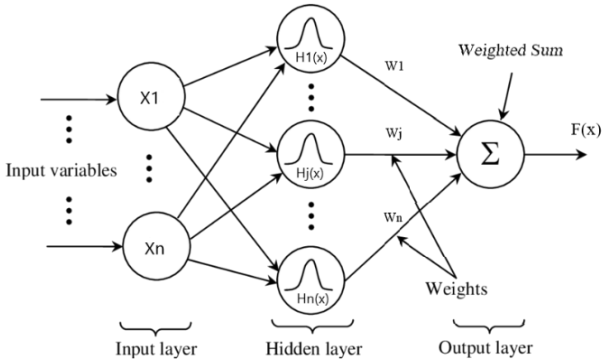

Fig 4 Radial Basis Function Neural Network [17]

The units with unknown "weights" connected between the hidden layer and output layer are linear. Compared with backpropagation networks, the learning speed is faster since the unknown weights can be calculated from solving a linear equation.

The RBF method was originally used for solving interpolation problems [18]. Given N inputs $\{x_i \in R^n | \, i = 1, \dots, N\}$ and N targets $\{\mathrm{y_i} \in \mathrm{R^n} | \, \mathrm{i} = 1, \dots, \mathrm{N}\}$, a function F is seeking to satisfy the interpolation conditions:

$$F(x_i) = y_i \; i = 1, \dots, N \quad (4)$$

and the problem can be solved by the RBF expansion:

$$F(x) = \sum_{i=1}^{N} \alpha_i h\left(||x - x_i||\right) \quad (5)$$

where α_i is an unknown coefficient, x represents the input vectors and x_i is the RBF centre. Define:

$$H = \; h\left(\left||x_j - x_i|\right|\right) \quad j = 1, \dots, N \quad (6)$$

Equation (3) is the response function of the ith locally-tuned unit. h is a positive definite function and normally chosen as the unit normalized Gaussian function. Substituting the interpolation condition (4) and the response function (6) to the RBF expansion (5), it becomes:

$$y_i = \sum_{i=1}^{N} \alpha_i H \quad i = 1, \dots, N \quad (7)$$

According to equation (7), the calculation of coefficient α_i depends highly on the size of matrix H. The computational complexity grows approximately to the third power of the size of the matrix (N^3). When N is large, there will be a higher probability of ill-conditioning to the equation. Therefore, a small number of neurons for transformation are desired. Equation (5) becomes:

$$F(x) = \sum_{j=1}^{K} \alpha_j h\left(\left||x - x_j|\right|\right) \quad (8)$$

The value of K is smaller than N and the system becomes overconstrained. By least-squares approach, coefficients can be solved by:

$$\alpha = H^{+} y \quad (9)$$

where H^+ refers to the pseudoinverse of H with a dimension of $(K \times N)$ [18].

4. RESULTS AND DISCUSSION

4.1 Stage 1: K-means Clustering Results

The K-means clustering results are shown in Fig. 5. Fig 5(a) shows one result of DBI calculation, the lowest value around 0.45 suggesting the optimal number of clusters is 4. The spread constant is then computed as 9.455. Fig 5(b) is the colour code embedded in the SOM-Toolbox. From the clustering results in Fig. 5(c) high dimensional data is represented in a 2-D grid and data with similar behaviour are formed as clusters. It can be seen that two colour regions on the top are fused together. Participants in each cluster are summarized in Table 2. The green region contains the most participants of state 0 and state 1, which are normal operational states. There is also a small amount of state 3 and state 4. The blue and yellow regions consist of single participants, which are state 0 and state 3 respectively. The participants in the dark blue region are state 1, state 3 and state 4. The reason for the existence of multi-states within one cluster is inferred as the fusion process or false alarms. Since SOM needs a sufficient amount of data to create meaningful clusters. Lack of data leads to an increase of randomness in the grouping process. Further clusters analysis is required in the future.

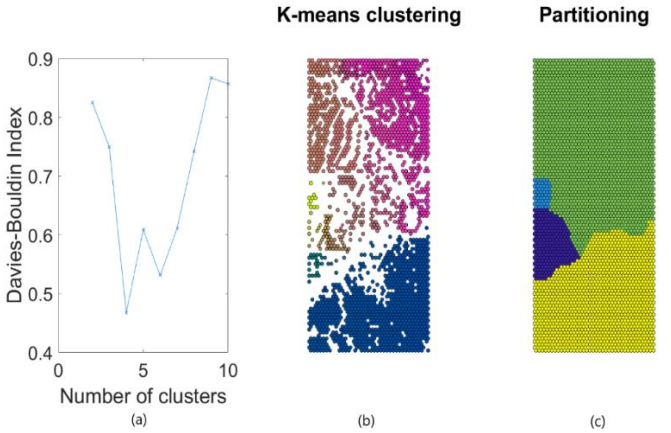

Fig 5 K-means Clustering (a) Davies-Bouldin Index. (b) Built-in olour code in the SOM-Toolbox (c) Partition

Colour Code	Composition
Green	State 0, State 1, state 3, state 4
Blue	State 0
Dark blue	State 1, State 3, State 4
Yellow	State 3

Table 2 Participants in Each Cluster

4.2 Stage 2: RBFNN Results

Fig 6 illustrates one of the testing results from the RBFNN. The red line represents the actual pitch state and the blue line is the testing result from the neural network. The result is relatively good, with most of the faults being predicted. Although some states are not correctly predicted, such as between sampling points 30 and 40. With simple rounding and filtering techniques, values are made integer and negative and extreme values can be filtered.

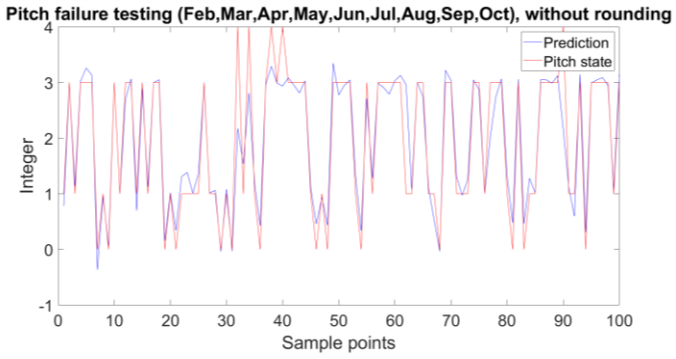

Fig 6 RBFNN Testing Result for One Simulation

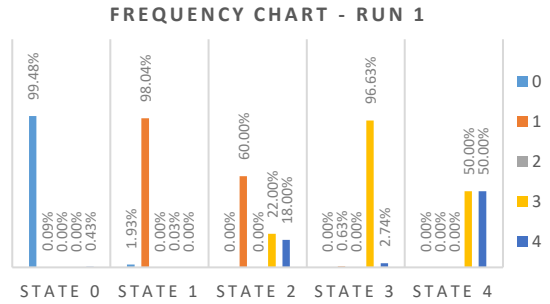

Fig 7 Frequency Chart Example

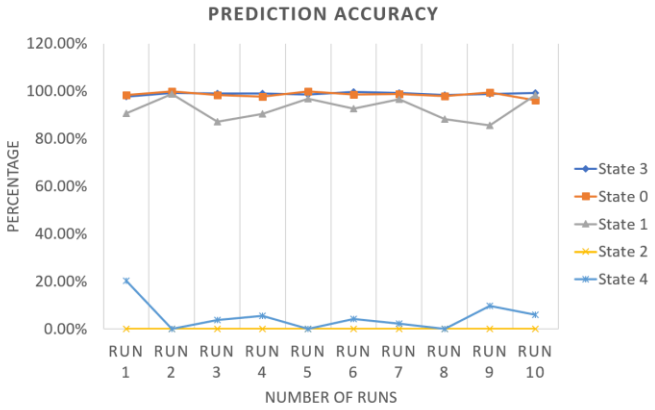

Fig 8 Prediction Accuracy of Each Pitch State

Fig 7 shows the frequency chart for one experiment. The x-axis labels are the predicted results from the RBFNN, and the colour bars represents real values in the dataset. Most values of state 0, state 1 and state 3 have been accurately classified. While poor performance for state 4 due to lack of training samples, so that the neural network could not learn their behaviours for prediction. The presence of state 2 in the predicted result was due to rounding process, so its performance will not be considered. By mapping the predicted results to the original dataset, the prediction accuracy is calculated. In Fig 8, the accuracy of identifying failures (state 3) has achieved to 96% - 99%. The performance for identifying states 0 and 1 is also high.

In general, the main advantage of our approach is that it requires less prior knowledge of WTs. The system design is also simplified by using a shallow neural network such as RBFNN, which is also tolerant to input noise. However, the weakness of this approach is the lack of analysing WT dynamic behaviours provided by the higher frequency SCADA data.

5. CONCLUSIONS

Due to the high failure rate and subsequent downtime, accurate failure prognostic techniques are desirable for the WT pitch system. Analysing WT SCADA

data for WT failure prediction represents a cost effective method, requiring no additional sensors. A data sampling frequency of 1 Hz was utilized in our study. In order to alleviate the computation burden of the larger dataset, a compressed method combining SOM and RBFNN was proposed in this paper. From the experimental results using SCADA data from the Levenmouth WT, an accuracy of identifying pitch faults of 96% - 99% has been achieved. Future work will be focus on classification methods of the causes for alarms and multi-states in the clusters.

ACKNOWLEDGEMENT

I would like to show my gratitude to the Offshore Renewable Energy (ORE) Company for providing information and data to support the simulation in this research.

REFERENCE

[1] Kim K, Parthasarathy G, Uluyol O, Foslien W, Sheng S, Fleming P. Use of SCADA Data for Failure Detection in Wind Turbines. ASME 2011 5th International Conference on Energy Sustainability, Parts A, B, and C. 2011;.

[2] Price S, Figueira R. Corrosion Protection Systems and Fatigue Corrosion in Offshore Wind Structures: Current Status and Future Perspectives. Coatings. 2017;7(2):25.

[3] Carroll J, McDonald A, McMillan D. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. Wind Energy. 2015;19(6):1107-1119.

[4] García Márquez F, Tobias A, Pinar Pérez J, Papaelias M. Condition monitoring of wind turbines: Techniques and methods. Renewable Energy. 2012;46:169-178.

[5] Kandukuri S, Huynh V, Karimi H, Robbersmyr K. Fault Diagnostics for Electrically Operated Pitch Systems in Offshore Wind Turbines. Journal of Physics: Conference Series. 2016;753:052005.

[6] Wu D, Liu W, Zhai Y, Shen Y. Fault Diagnosis for the Pitch System of Wind Turbines Using the Observer-Based Multi-innovation Stochastic Gradient Algorithm. Theory, Methodology, Tools and Applications for Modeling and Simulation of Complex Systems Communications in Computer and Information Science. 2016;526–38.

[7] Guo H. A Novel Degradation Identification Method for Wind Turbine Pitch System. IOP Conference Series: Earth and Environmental Science. 2018;133:012011.

[8] Mücke T, Wächter M, Milan P, Peinke J. Langevin power curve analysis for numerical wind energy converter models with new insights on high frequency power performance. Wind Energy. 2014;18(11):1953-1971.

[9] Gonzalez E, Stephen B, Infield D, Melero J. On the use of high-frequency SCADA data for improved wind turbine performance monitoring. Journal of Physics: Conference Series. 2017;926:012009.

[10] Schlechtingen M. A global condition monitoring system for wind turbines. DTU Mechanical Engineering DCAMM Special Report. 2013;No. S150.

[11] Vesanto J. [Internet]. Cda.psych.uiuc.edu. 2000 [cited 14 May 2019]. Available from: http://cda.psych.uiuc.edu/martinez/edatoolbox/Docs/toolmet2000.pdf

[12] Lohninger, H. (2012). *Kohonen Network - Background Information*. [online] Lohninger.com. Available at: http://www.lohninger.com/helpcsuite/kohonen_network_-_background_information.htm [Accessed 15 May 2019].

[13] Rumelhart D, Zipser D. Feature Discovery by Competitive Learning*. Cognitive Science. 1985;9(1):75-112.

[14] Uk.mathworks.com. (2019). *Design radial basis network - MATLAB newrb- MathWorks United Kingdom*. [online] Available at: https://uk.mathworks.com/help/deeplearning/ref/newrb.html [Accessed 15 May 2019].

[15] Davies D, Bouldin D. A Cluster Separation Measure. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1979;PAMI-1(2):224-227.

[16] Hashemi Fath, A., Madanifar, F. and Abbasi, M. (2018). Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks to predict solution gas-oil ratio of crude oil systems. *Petroleum*.

[17] Broomhead D. S, Lowe D. Multivariable functional interpolation and adaptive networks. Complex Systems. 1988;2:321-355.

[18] Poggio T, Girosi F. A Theory of Networks for Approximation and Learning. Technical Report Massachusetts Institute of Technology, Cambridge, MA, USA. 1989;.

[19] Pang, K. (2003). *Self-organizing Maps*. [online] Cs.hmc.edu. Available at: https://www.cs.hmc.edu/~kpang/nn/som.html [Accessed 18 Jun. 2019].

[20] Wang, Y., Hu, Q., Srinivasan, D. and Wang, Z. (2019). Wind Power Curve Modeling and Wind Power Forecasting With Inconsistent Data. *IEEE Transactions on Sustainable Energy*, 10(1), pp.16-25.