A decision Tree Based Approach for Data Analysis and Fault Prediction For Wind Turbine Using SCADA dataset

[1]Roheela Amin, [2]Sushma Gupta, [3]Muheet Ahmed Butt, [4]Majid Zaman
[1]M.Tech. Student, [2]Assistant Professor, [3,4]Scientist D
[1,2]Swami Devi Dayal Institute of Engineering and Technology, Haryana, India,
[3,4]University of Kashmir, Srinagar, J&K

Abstract - The main objective of Supervisory Control and Data Acquisition (SCADA) is to give a means to the human operator to control and to command a highly automated process. Supervisory control and data acquisition systems (SCADA) are widely used in industries for supervisory control and data acquisition of industrial processes. It has also revolutionized the field of data sciences in a much inevitable manner. The communication between the control centre of SCADA and the remote station takes place by various communication channels such as optical fibre cables, microwave technology and power line carrier communication. The presence of such an efficient system give rise to certain faults results in malfunction of such systems. As the demand for wind energy continues to grow at exponential rate, reducing operation and maintenance costs and improving reliability have become top priorities in wind turbine maintenance strategies. Prediction of wind turbine fault before they reach a catastrophic stage is critical to reduce the operational and maintenance cost due to unnecessary scheduled maintenance. To this end, it is important to be able to perform maintenance before it's needed. Instead, by performing complex analysis of existing data from turbine's supervisory control and data acquisition system (SCADA) system, valuable insights into turbine performance can be obtained at a much lower cost. This paper proposes a methodology of fault prediction for wind turbine based on stored SCADA data set using decision tree approach. Fault analysis and diagnosis of faults help in protecting both hardware and software of devices employed in SCADA based systems. On the basis of information received, the SCADA operator is in a position to take important decision related to smooth and faultless generation, transmission and distribution of power. The proposed research provides insight analysis for relationship ships of various attributes like rotor speed, turbine speed, wind speed, power and output of the dataset and also provides fault analysis using Decision Tree Data Mining approach on the said dataset care of various parameters.

Keywords: SCADA system, Wind Turbine, Decision tree algorithm.

1. Introduction

SCADA systems are installed in different substations to facilitate the day to day operations. They are also being installed in oil pipelines to detect leakage in pipelines. In fact in plants such as nuclear power plant where there is high risk to carry out the condition monitoring of the different auxiliary equipment by the help of manpower, SCADA systems automatically carry out this job by employing the use of field devices that continuously sense the operating states of different critical inaccessible equipment and pass on the information to the home station. The operator sitting at the home station can now see the status of all equipment on the display provided to him on operator console.

Supervisory control and data acquisition (SCADA) Systems are frequently used in water and waste water systems to monitor and control tank levels, remote well pumps, lift station pumps, high service pumps, valves, and chemical pumps. SCADA systems range from a simple to large configurations. Most of the SCADA applications use a human machine interface (HMI) software that permits users to interface with machine to control the devices. HMI is thus connected to motors, valves and many more devices. The software of SCADA receives the information from programmable logic controllers (PLCs) or remote terminal units (RTUs), which in turn receives their information from sensors or inputted values which we have given manually. SCADA is used in a power system to collect, analyse and to monitor the data effectively, which will reduce the waste potentially and will improve the efficiency of the entire system by saving money and time.

Among the challenges, it has been noted that 20% wind energy by 2030, are improvement of wind turbine performance and reduction in operating and maintenance costs. Maintenance costs can be reduced through the continuous, automated monitoring of wind turbines. Wind turbines often operate in severe, remote environments and require frequent scheduled maintenance. Unscheduled maintenance due to unexpected failures can be costly, not only for maintenance support but also for lost prediction time. In addition, as wind turbines age, parts fail, and power production performance degrades, maintenance costs increase as a power production performance degrades maintenance costs increase as a percentage of production. Monitoring and data analysis enables condition based rather than time-interval-based maintenance and performance time-ups. Experience from other industries shows that condition monitoring detects failures before they reach a catastrophic or secondary-damage stage, extends assets life, keeps assets working at initial capacity factors, enables better maintenance planning and can reduce routine maintenance.

In this research paper we describe our exploration of existing wind turbine SCADA data for development of fault detection and diagnostic techniques. Our ultimate goal is to be able to use SCADA-recorded data to provide advance warning of failures or performance issues. A number of measurements from turbine are used to develop anomaly detection algorithms. Classification techniques such as clustering and principal components analysis were investigated for capturing fault signatures. The developed algorithms were tested with data from a failed gearbox incident. Here we then describe the wind turbine and data used for this work, followed by a description of the algorithm approaches used, results, discussion and concluding remarks. A monitoring system that provides detection coverage for multiple top failure modes will be valuable in a maintenance system. With this objective, we explore the use of SCADA data for fault detection and diagnostics.

2. SCADA System For Wind Energy Generation

Wind power is an important source of renewable energy for which the UK has a excellent resource. Wind turbines are generally grouped together to form wind farm. Here the author describes how SCADA system play key role in distributed control system used to operate and maintain wind farm. SCADA is for remote supervision and control of a wind turbines and wind parks The SCADA system from wind power technology offers full remote control and supervision of entire wind park and the individual wind turbines. The SCADA system can run on a computer in the control room of the wind park or it can run on any internet connected computer accessing the wind park using TCI/IP.
Here are the main features of SCADA system in wind energy generation:
Park overview: The park overview of the SCADA system provides a graphical overview of the wind park indicating the status of each individual turbine.
Park control: The park overview makes it possible to start/stop entire wind park, clusters of turbines or individual wind turbines. Furthermore the park control can be used for setting production limits for wind park.
Turbine Overview: The turbine overview of SCADA system gives a full overview of all relevant parameters of wind turbine, for instant temperatures, pitch angle, electrical parameters, rotor speed, yaw system etc.
Log Viewed: The SCADA system features a flexible browsing of the log data of wind turbine. all relevant log data are accessible and can be sorted by different parameters.
Report Generator: The report generator of SCADA system makes it possible to make all relevant reports based on the log data. The reports can be graphically presented to provide the best possible overview.
The wind energy generation based on SCADA system is shown in figure 1. Below:

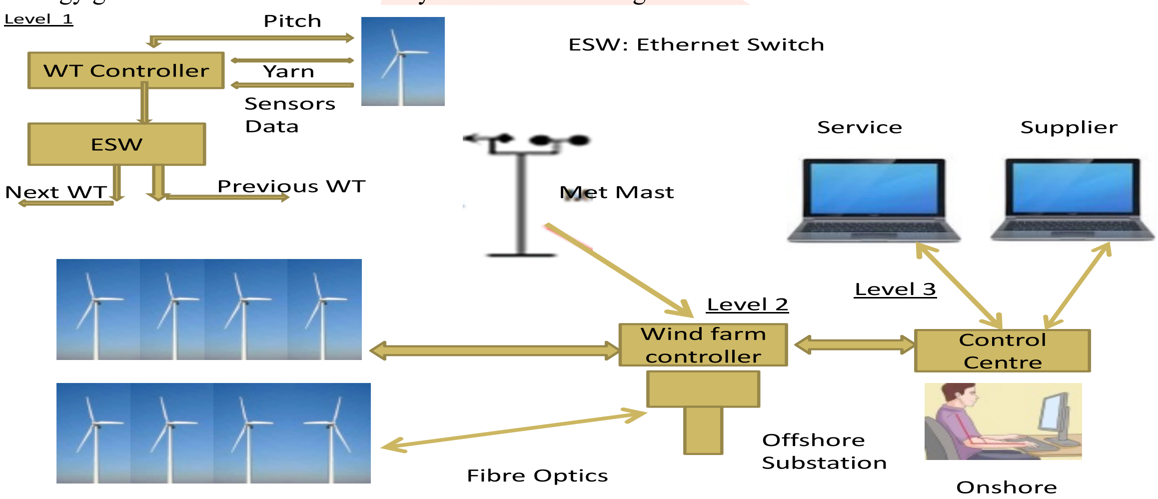

Figure 1: SCADA based wind energy generation

The SCADA system can be used for a larger wind farms. A wind farm is a group of wind turbines in the same location used to produce electricity. A large wind farm may consist of several hundred individual wind turbines and cover an extended area of hundreds of square miles. For such a larger area, it is not possible to visit each and every turbine to collect all the data like, daily generation, warning, status, message, faults etc. So in that case SCADA system is utilised to monitor the turbine from a single platform and keeping the record of generations (KWH) and faults. In SCADA system all the turbines are connected with single server either by optical fibre or by GSM network.
SCADA can be used for small wind turbines that are being used for stand alone application but it will not be financially viable as it will add extra cost to turbine in case to visit the turbine to collect the data and attend the faults. SCADA really just refers to all the data collected by various sensors and supervisory control aspects. Supervisory control is things like requesting (from the central park controller) the turbines to start or stop due to electricity stop price signals or to take advantage of low wind speeds to perform a maintenance routine. Critical control like high speed shut down, emergency brake programs, pitching, speed control/active power regulation etc are performed within the turbine controller and are not considered SCADA (although the turbine controller produces SCADA data.

3. Wind turbine components degradation and consequences

The main objectives in wind turbine control systems are avoidance of excessive mechanical loads, maximisation of energy capture, and the provision of appropriate power quality. Maximum power extraction can be achieved using bigger diameter of wind turbine blades. Blades are made by either glass fibre with epoxy composites or high density wood. Blades are one of the weakest components in a wind energy conversion system and some faults might happen directly on the blades such as hub and blade corrosion, crack, serious aero-elastic deflection, and rotor imbalance. Furthermore, high force on the blades causes high vibration and fatigue in the main body of the turbine. it is important to note that failures in the pitch controlled system will cause the blades breakdown. Tower of the wind turbine is basically designed to maintain vibrations within win speed changes as well as to hold the nacelle. A tower of the wind turbine might fail due to a storm, fire, or earthquake. A gearbox in the drive train increases the rotor speed to values more than the acceptable range of wind turbine speed (i.e., from 20-50 rpm to 1000-1500 rpm). The failure of the gearbox might be due to the shaft imbalance, shaft damage, shaft misalignment, gear damage, bearing damage, broken shaft, high oil temperature, and leaking oil. Failures in the wind turbine will cause several issues such as abnormal noises and excessive vibration in the generator. Generator fault can be occurred due to either internal or external failures. External failures happen due to the short circuits in the AC grid which will lead to generator overheating. Internal failures are usually because of the poor insulation or mechanical failures such as slip ring and bearing. The yaw system failure can be occurred by failures in its components, typically control system and drive motor failure. Failure in the yaw system may cause some common failures such as losing the track of the wind direction, results the output power decline. Regarding the failures in the brake system, the wind turbine must be stopped. Otherwise it will cause serious damages to the mechanical units, main structure and blades.

4. Literature Survey

The literature survey is conducted on different energy management systems fetching the data from the generation to consumption by using different means and approaches. Based on different papers the smart energy architecture is proposed and a smart energy data warehouse is implemented for smart grid. With the wide use of wind energy resources the traditional energy resources have been adjusted and modulated. In the recent years, supervisory control and data acquisition (SCADA) system has been applied in power system substation automation and becomes a focus of electrical utility. A brief literature survey related to wind energy and its electricity generation based on SCADA system is given below:
Xingjio Yao, Shu Liu and Xiaodong Wang, "Research on Maximum Wind Energy Capture Control Strategy",2010, presented a paper which proposes a optimum control strategy to utilize the wind energy and improve efficiency of wind generation system. Maximum wind energy capturing is one of the most important technology for improving the conversion efficiency of wind power generation system and achieving maximum power point tracking. In this paper analysis of wind turbines operation characteristics, the principle and process of capture the maximum wind energy were analysed.[1]

Preben Maegaard, "Wind Energy Development and Application Prospects of non – grid – connected Wind Power",2009 has proposed a paper in which renewable energy is driving element of word economic development. The faster development of wind power has been making more contribution for the development of renewable energy. In this paper the non-grid connected wind power is a new concept in renewable energy field. Non-grid-connected wind power system directly link wind energy into other various high energy consumption industries. Theoretical researches and applications of large-scale non-grid-connected wind power, will open up the new areas of wind power diversified applications.[2]

Simon Gill, Edward Barbour, IA Grant Wilson, David Infield, "Maximising revenue for non-firm distributed wind generation with energy storage in an active management scheme",2013 has given a paper in which connection of high penetrations of renewable generation such as wind to distribution networks requires new active management techniques. In this paper the renewable subsidies are considered as a potential third revenue stream. The result clearly shows that storage using both operating modes increase revenue over either mode individually.[3]

Linas Gelazanskas, Kelum A.A garmage, "Managing Renewable Intermittency in Smart Grid: Use of residential hot water heaters as a form of energy storage",2016, has presented a paper which discusses a novel wind generation balancing technique to improve renewable energy integration to the system. the novel individual hot water heater controllers were modelled with the ability to forecast and look ahead the required energy, while responding to the electricity grid imbalance. The method developed in this research are not limited to wind power balancing and can be used with any other type of renewable generation source. Artificial intelligence and machine learning techniques were used to learn and predict energy usage [4]

Qui Yingning Sun Juan, CAO Mengnan FENG Yanhui, " Model based wind turbine gearbox detection fault detection on SCADA Data", 2014 has presented a paper in which typical wind turbine gearbox condition monitoring is based on vibration signals, which is effective to detect failure with high frequency signal range. To systematic understand wind turbine systems, this paper presents research results of model based wind turbine gearbox fault detection. The result obtained in this work is useful for wind turbine gearbox design and effective algorithm development of fault detection.[5]

D Gopi Krishna, "Preventive maintenance of wind turbines using remote instrument monitoring system", 2012 has given a paper in which effective and efficient usage of renewable energies like wind, solar, hydro and tidal power should be the vital issue of the decade to meet the current energy requirements. This paper proposes remote instrument monitoring system (RIMS) which constitute of wireless sensor network based supervisory control and data acquisition (SCADA) system, with the multiple sets of sensors distributed wide across each turbine in wind farm. [6]

N Chen, R Yu, Y Chen, H Xie "A Hierarchical method for wind turbine prognosis using SCADA data", 2016, gave a paper in which a hierarchical method based on GP(Gaussian processes) and PCA (Principal Component Analysis) is proposed for turbine prognosis using SCADA data. The method includes two levels of prognosis: 1) detect which wind turbine behaves abnormally and has potential defect. 2) determine the defective components in abnormal turbine. The field dataset including 24 failed turbines is used to validate the proposed hierarchical method. Thus the validation results show that proposed method can achieve wind turbine prognosis with 79% detection rate on turbine level and 75% detection on component level.[7]

Yusuf DEBBAG, Ercan Nurcan YILMAZ, "Internet based monitoring and control of a wind turbine via PLC",2015 presents a paper in which policy challenges associated with global warming, the prospect of increasingly expensive fossil fuels and serious concerns about safety of nuclear power are encouraging many country of world economics to develop smart grids. All information and data which belong to wind turbine, must be gathered in one central unit. If we entre central station IP address and port number of SCADA program into address line of web browser, we can monitor the system via any web browser of an internet connected computer.[8]

B Chen, YN Qui, Y Feng, PJ Tavner, WW Song, "A wind turbine SCADA alarm pattern recognition", 2011, gave a paper which develops a wind turbine supervisory control and data acquisition (SCADA) system that contains a alarm signals providing significant important information. This paper presents the feasibility study of SCADA ALARM and diagnosis method using an artificial neural network (ANN). Based on this study, we have found that the general mapping capability of ANN help to identify those likely wind turbine faults from SCADA alarm signals, but a wide range of representative alarm pattern are necessary for supervisory training.[9]

Se-Yoon Kim, In-Ho Ra, Sung-Ho Kim, "Design of wind turbine fault detection system based on performance curve", 2012, has presented a paper in which wind turbine fault detection system based on wind verses power performance curve obtained from SCADA is studied. SCADA data obtained from 850 KW wind turbine system installed in Kunsan Korea are used and various simulation studies were carried out.[10]

R. Palma-Behnke, D. Ortiz, L. Reyes, G. Jimenez-Estevez, N. Garrido, "A social SCADA approach for renewable based micro-grid – The Huatacondo project", 2011 has presented a paper which proposes a novel SCADA approach for renewable based micro-grid. The optimization scheme provides online set points for each generation unit, operation modes for a water supply system, and signals for consumers based on demand side management mechanism. This concept is applied to a smart micro-grid composed of photovoltaic panels, a wind turbine, a diesel generator, a battery bank and a water supply system. The result shows the economics of energy management system and impact of community participation.[11]

Huuan Long,Long Wang, Zijun Zhang, Zhe Song, Jia Xu, "Data-driven wind turbine power generation performance monitoring, 2015 has given a paper which investigates the wind turbine power generation performance monitoring based on SCADA data. The proposed approach identifies turbines with weekend power generation performance through assessing the wind power curve profiles. Two blind industrial studies are conducted to validate the effectiveness of the proposed monitoring approach, and the results demonstrate high accuracy in detecting the abnormal power curve profiles of wind turbines and their associated time intervals.[12]

Meik Schlechtingen, IImar Ferreira Santos, Sofiane Achiche, "Using data mining approaches for wind turbine power curve monitoring: a comparative study", 2013 has presented a paper in which four data mining approaches for wind turbine power curve monitoring are compared. In this research, cluster centre fuzzy logic, neural network an d k- nearest neighbour models are built and their performance compared against literature. The comparison shows the decrease of error rates and of ANFIS models when taking into account the two additional inputs and the ability to detect faults earlier.[13]

Blind Chen, Peter C Matthews, Peter J Tavner, "Automated on-line fault prognosis for wind turbine pitch systems using supervisory control and data acquisition", 2015 has proposed a paper in which current wind turbine studies focus on improving their reliability and reducing the cost of energy, particularly when wind turbines are operated offshore. Ideally, a wind turbines health condition or state of the components can be deduced through rigorous analysis of SCADA data. This study proposed a new method for analysing wind turbine SCADA data by using an a priori knowledge based adaptive neuro-fuzzy inference system with the aim to achieve automated detection of significant pitch fault.[14]

Yanjun Yan, James Zhang, "Using edge-detector to model wake effects on wind turbines", 2014 has presented a paper in which a healthy wind turbine is essential for efficient wind energy generation, and fault monitoring. One major phenomenon that is not a fault but will cause power reduction is wake effect. This paper proposes to use the SCADA data to learn the wake pattern for each turbine. the accurate wake pattern generated by this approach is helpful to separate wakes from true faults, and to understand the vulnerability of the turbines.[15]

Michael Wilkinson, Brian Darnell, Thomas Van Delft, Keir Harman, "Comparison of methods for wind turbine condition monitoring with SCADA data", 2014 has presented a paper which proposes that wind turbine operational can be reduced by monitoring the condition of major components in the drive train. In this paper three SCADA based monitoring methods were reviewed: signal trending, self-organising maps and physical model. The physical model was identified as being the most reliable as predicting impending component failure. An advance detection period of between 1 month and 2 years was achieved by the method.[16]

Ran Bi, Chengke Zhou, Donald M Hepburn,"Applying instantaneous SCADA data to artificial intelligence based power curve monitoring and WTG fault forecasting", 2016 has given a paper in which power curve (PC) monitoring can be applied to evaluate the wind turbine generator (WTG) power output and detect deviations between the expected and measured value, often a precursor of unexpected faults. In this research, the instantaneous SCADA data is used to show the fault forecast ability of artificial intelligence (AI) based PC monitoring of a pitch regulated WTG. A case study illustrated that AI models, using wind speed, generator speed and pitch angle inputs, would have successfully detected a pitch fault due to the slip ring malfunction nearly 5hours earlier than the existing fault detection mechanisms.[17]

Pramod Banglore, Lina Bertling Tjernberg, "An artificial neural network approach for early fault detection of gearbox bearings", 2015 has proposed a paper in which it is proven that gearbox is to be a major contributor towards downtime in wind turbines. The majority of failures in the gearbox originate from the gearbox bearings. This paper introduces a self-evolving maintenance scheduler framework for maintenance management of wind turbines. Furthermore, an artificial neural network (ANN) based condition monitoring approach using data from supervisory control and data acquisition system is proposed. The results demonstrate that the proposed ANN based condition monitoring approach is capable of indicating severe damage in component being monitored in advance.[18]

Yingying Zhao, Dongsheng Li, Ao Dong, Li Shang, Dahai Kang, Jaijai Lin, "Fault Prognosis of Wind Turbine Generator Using SCADA data", has presented a paper which proposes that accurate prognosis of wind turbine generator failures is essential for reducing operation and maintenance costs in wind farms. This paper presents a prognosis method to predict the remaining useful life (RUL) of generators, which requires no additional hardware support beyond widely adopted SCADA system. This paper presents a data-driven wind turbine anomaly detection method and time series analysis method to predict the wind turbine generator remaining useful life (RUL).[19]

W.A.W. Ghoneim, A.A. Helal, M.G. Abdel Wahab, "Renewable Energy Resource and Recovery Opportunities in Waste Water Treatment Plants", 2016, has presented a paper which describes the different methodologies for energy recovery and renewable energy utilization techniques in waste water treatment plants which are one of the largest consumable of electric power. To reach self-sustainability the plant can get thus the rest of its energy demand through convectional renewable resources such as solar, wind, hydro and tidal power.[20]

5. Decision Tree Methodology

Decision tree is the main technology mainly used for classification and prediction using a desired dataset. Decision tree learning is a typical inductive algorithm based on instance, which focuses on classification rules displaying as decision trees incidental from a group of disorder and irregular instance [21]. In top-down recursive way, it compares attributes between internal nodes of decision tree, judges the downward branches according to different attribute of the node, and draws a conclusion from leaf nodes in the decision tree. So from a root to a leaf node corresponds to a conjunctive rule, and the entire tree corresponds to a group of disjunctive expression rules. Take the decision tree as a Boolean function. The input of the function is the object or all property of situation, and the output is the "yes" or "no" decision value. In the decision tree, each tree node corresponds to a property test, each leaf node corresponds to a Boolean value, and each branch represents one of the possible values of testing attribute. The most typical decision tree learning system is ID3, which originated in the concept learning system CLS, and finally evolved into C4.5 (C5.0), which can deal with continuous attributes. There are multiple decision tree methods, such as ID3, C4.5, PUBLIC, CART, CN2, SLIQ, SPRINT etc. Decision tree classification algorithm is usually divided into two steps: 1. To construct decision tree and 2. To prune Decision Trees.

5.1. Construction of Decision Tree: The input of decision tree construction algorithm is a set of classic labelled examples. The result of structure is a binary tree or ternary tree. The internal nodes of a binary tree (non-leaf nodes) are usually represented as a logical judgment, such as in the form of (AI=VI) logical judgment, AI is the attribute, of which VI is a value. Tree's edge is the branch outcome of logic. The internal node of ternary tree is an attribute, of which edges are all values. Where there are several attribute values, there are several sides. Tree leaf nodes are category tag. Method of constructing decision tree is a top-down recursive structure. With the ternary tree as an example, its structural idea is starting to establish decision tree with single node represented the training sample. If the samples are all in the same class, it can be leaf nodes, and contents of nodes are the category tags.

5.2. Pruning of Decision Tree: The object of Data mining is real world data, which are generally not perfect. Maybe there are some missing values in attribute field; or lack essential data resulting to incomplete data; or data are inaccurate even wrongly, or containing noise, so it is necessary to discuss the problem of noise. The basic decision tree construction algorithm does not consider the noise, so the generated decision tree fits completely with training examples, which will lead to excessive fitting and will destruct predictive performance. Pruning is a technique to overcome noise, at the same time it also can make the tree simplified and easy to understand. The Two pruning strategies are forward pruning and Post pruning.

5.2.1. *Forward-Pruning:* is pruning before the decision tree's growth process is completed, when decide to continue dividing the impure training subset or shutdown.

5.2.2. Post-Pruning: is pruning after the decision tree growth process is completed. It is a Fitting- and-simplifying of the two stage method. First generate a decision tree fitting completely with training example, and then trim the leaves of the tree from the bottom to the top, gradually to the root. When Pruning, it is used a test data set.

6. SCADA Data Set For Wind Turbine Description

A one year data for a wind turbine was collected which amounted to 60745 records. The dataset composed of Record Id, Rotor Speed in Kilometre per hour, Turbine Speed in RPM, Wind Speed, Power, Output of Turbine as shown in table below. The output had two outcome values NORMAL and ERROR. The experiment and the related analysis which was carried out is shown in next section.

Record ID	Rotor Speed - KPH	Turbine Speed - RPM	Wind Speed	Power	Output
954	49.18768967	515.16	14.31	1000	NORMAL
955	49.55814994	414	11.5	401.9138756	NORMAL
956	40.4945132	297	8.25	246.4114833	NORMAL
957	50.94920888	1532.37	51.97	1771.689197	ERROR
958	42.60087531	360.72	10.02	331.1004785	NORMAL
959	45.04538961	393.12	10.92	374.1626794	NORMAL
960	48.08164649	460.8	12.8	1000	NORMAL
961	42.17361707	305.28	8.48	257.4162679	NORMAL
962	41.80015591	288.36	8.01	234.9282297	NORMAL
….	…….	…….	……..	……	…….

7. Experimental Analysis For SCADA Based Turbine Data

The primary parameters of the data set are

1. Rotor Speed
2. Turbine Speed
3. Wind Speed
4. Power
5. Output

In order to understand the working of the system, we first need to determine the relation between various parameters, besides predict as and when the system will stop besides determine those factors which cause the system to stop.

7.1. Rotor Speed

The relation of Rotor Speed with other parameters is analysed and shown below.

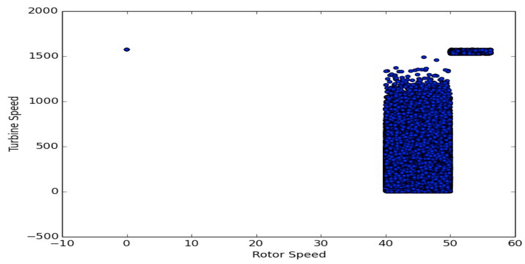

Figure 2: Rotor Speed & Turbine Speed

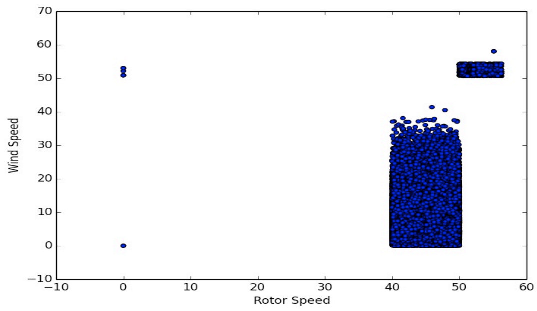

Figure 3: Rotor Speed & Wind Speed

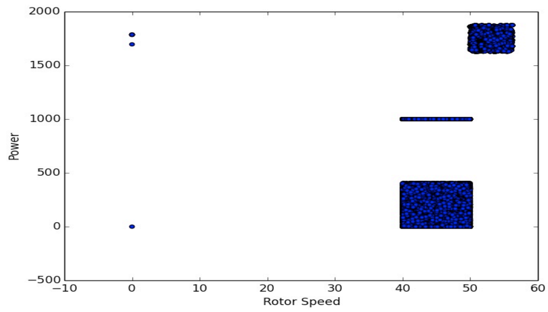

Figure 4: Rotor Speed & Power

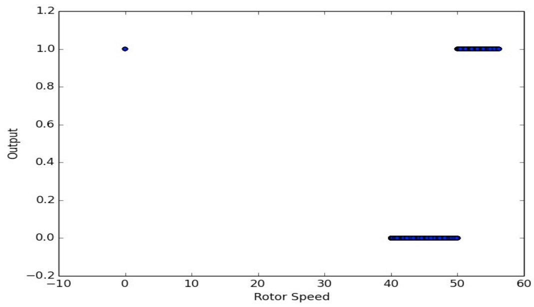

Figure 5: Rotor Speed & Output

From the above analysis it is very evident that there is not much effect of Rotor speed on overall functioning of the system. Rotor Speed varies from 40-50 and change in Rotor Speed does not essentially change the functioning of other parameters neither Output. However after in depth analysis it can be determined that whenever system is in stop state the rotor speed is above 50 however reverse is not always true and thus Rotor Speed cannot solely be factor which determines the overall working of the system. The basic co-relation between Rotor Speed and other parameters can be established that whenever Rotor speed is above 50 other system is likely to be above normal values.

7.2. Turbine Speed

The relation of Turbine Speed with rest of the parameters is analysed and shown below:

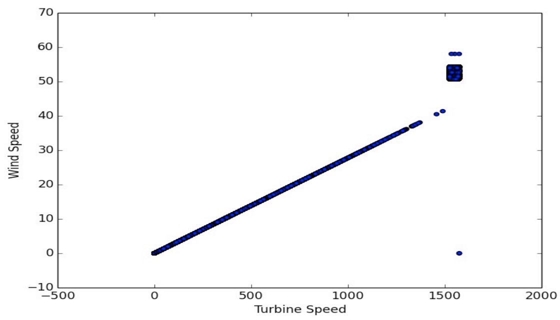

Figure 6: Turbine Speed & Wind

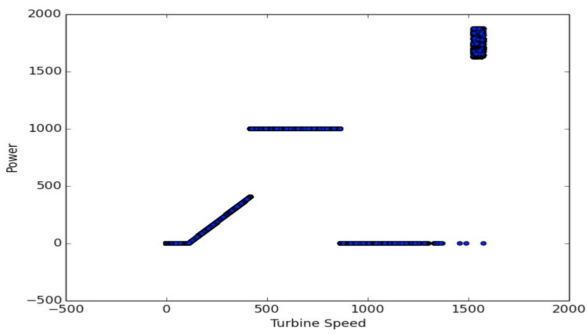

Figure 7: Turbine Speed & Power

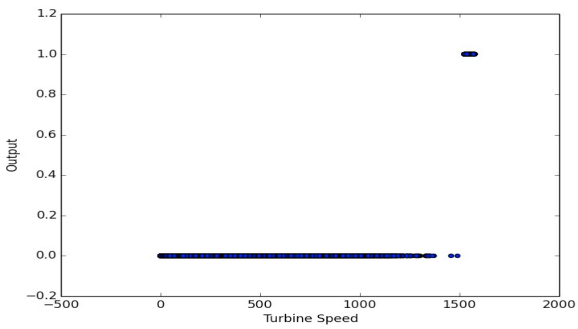

Figure 8:Turbine Speed & Output

From the above analysis it is very evident that with the rise in Wind speed there is rise in turbine speed up to the certain value when there is no increase in turbine speed no matter what the wind speed is, this is the value when the power is maximum and results in shutdown of the system, thus it can be concluded that wind speed determines turbine speed and overall functioning of the system, however resultant turbine speed of 1500+ causes system to shut down.

7.3. Wind Speed, Power & Output

The relation of rest of the parameters is analysed and is mentioned below

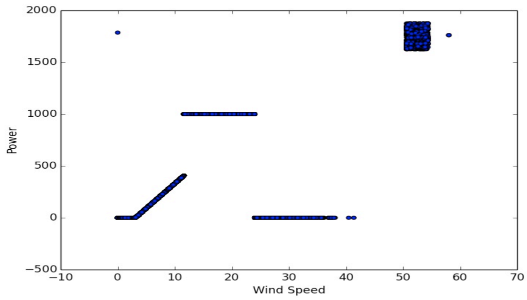

Figure 9: Wind Speed and Power

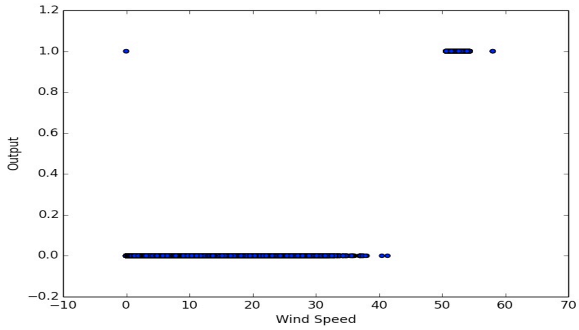

Figure 10: Wind Speed & Output

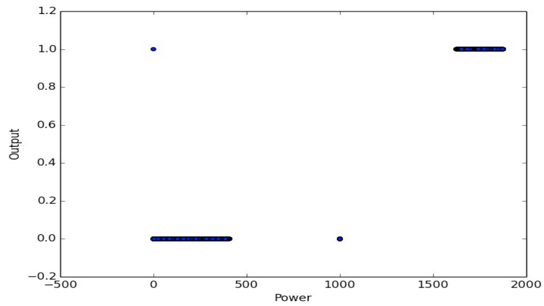

Figure 11: Power & Output

From the above analysis it is clear that wind speed of 50 and above results in power of 1500 + which causes system to shutdown, thus system is predicted by the wind power which cause turbine speed to increase which results and high power(1500+) and thus causing the system to shut down. Further amount of energy generated by SCADA is dependent upon wind however in order to ensure system

8. Proposed Framework for Fault and Performance Prediction

A Decision Tree algorithm is used to predict the system outcome, given below is the decision tree algorithm for the prediction of system credibility.

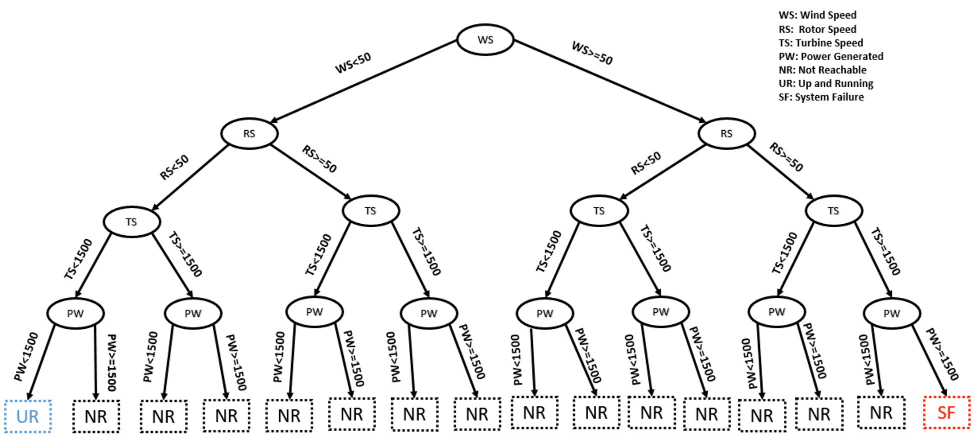

Figure 12: Decision Tree for the Proposed Framework

From the analysis of data and execution of decision tree on above data set results shown above are derived in the table below:

Case No.	Wind Speed	Rotor Speed	Turbine Speed	Power	Output	Remarks
Case 1	>50 (4145)	>50 (4139)	>1500	>1500	1	System

			(4147)	(4146)	(4147)	crashes
Case 2	<50 (46220)	<50 (46226)	<1500 (46218)	<1500 (46219)	0 (46218)	System is up (Working mode)
Total cases	50365	50365	50365	50365	50365	

The transformation of the above data for Decision Tree Algorithm is shown below in a .csv format.

WS	RS	TS	PW	Output
LOW	LOW	LOW	LOW	UR
LOW	LOW	LOW	LOW	UR
LOW	LOW	LOW	LOW	UR
LOW	LOW	LOW	LOW	UR
LOW	LOW	LOW	LOW	UR
HIGH	HIGH	HIGH	HIGH	SF
LOW	LOW	LOW	LOW	UR
LOW	LOW	LOW	LOW	UR
HIGH	HIGH	HIGH	HIGH	SF
LOW	LOW	LOW	LOW	UR
LOW	LOW	LOW	LOW	UR
LOW	LOW	LOW	LOW	UR

Where WS is Wind Speed, RS is Rotor Speed, TS Turbine Speed, PW Power and Output is the final outcome of the record.

9. Proposed Algorithm

The Proposed algorithm for the proposed dataset is shown below:

Step 1: Place the best attribute of the dataset of WIND SCADA at the root of the tree. Here the best attribute is the Wind Speed.

Step 2: Split the training set into subsets. Subsets should be made in such a way that each subset contains data with the same value for an attribute. In our case the subsets include Rotor Speed, Turbine Speed and Power.

Repeat step 1 and step 2 on each subset until you find leaf nodes in all the branches of the tree.

Interpretation: In the proposed a decision tree algorithm, for determining various determine predictability possible in the proposed system. We compare the values of the root attribute with other record's attribute present in the SCADA Dataset. On the basis of comparison, we follow the branch corresponding to that value and jump to the next node. We continue comparing our record's attribute values with other internal nodes of the tree until we reach a leaf node with predicted class value. In our case these values are UR, NR and SF representing up and Running, Not Reachable and System Failure respectively. This approach modeled decision in a manner that faults could easily be predicted based on a desired record values.

From the above decision tree algorithm following is deduced

- ✓ Wind Speed varies, however the moment it crosses 50 it creates ripple effect on rest of the parameters.
- ✓ Rise in Wind Speed above 50 causes rise in Rotor Speed up and above 50, this is directly proportionate relationship.
- ✓ Rise in Rotor Speed above 50 causes rise in Turbine Speed up and above 1500.
- ✓ Rise in Turbine Speed above 1500 cause power generation of 1500+ which cause system to shut down.
- ✓ As long as wind speed is less than 50, system is up, however generation of power is dependent upon wind speed but the moment it crosses 50 it will result is system shutdown.
- ✓ Prediction System demands that wind speed above 50 and system should automatically shut down.

10. Conclusion

SCADA is the solution for industrial automation. It enables you to achieve monitoring and controlling your plant operation remotely. Field instruments and devices will be linked to SCADA engine and PLC, therefore you can simply trigger via SCADA screens (HMI) to control your instruments, as well as simply monitor any live data and status via the screens. This benefits both operators and supervisors in operating the process, as it is able to reduce much time and efforts. Most of the countries globally are opting for natural energy generation for example solar, wind etc. While tapping of solar energy is more predictable, it is generation of wind energy which can cause unexpected effects. In this paper the structure of typical SCADA system used for operation and maintenance of turbines has been presented and extended with a fault detection system, which is based on the assumption that wind turbines near each other in a wind farms should be under the same operating conditions. Wind causes turbines to propel which in turn generates energy but it can also lead to catastrophic effect in certain cases especially if wind is blowing above certain threshold, in this case 50, this threshold needs to be determined in order to ensure safety of the entire system and its surrounding. The proposed research not only provides various relationships between Dataset attributes but also determine predictability of the system using Decision Tree algorithm for appropriate decision that can be taken at real time Sand so that to avoid any failure. The failure conditions that are not trapped properly can cause lot damage to the wind turbine infrastructure.

11. REFERENCES

[1] Xingjia Yao, Shu Liu, Xiadong Wang, Zuoxia Xing, Research on maximum wind energy capture control strategy, Optics photonics and energy engineering (OPEE), 2010 International conference on 1, 34-37, 2010.

[2] Preben Maegaard, Wind energy development and application prospects of non-grid-connected wind power, proceedings of, 1-3, 2009.

[3] S Gill, E Barbour, D infield, Maximising revenue for non-firm distributed wind generation wind energy storage in an active management scheme, IET Renewable Power Generation, 2013 – IET.

[4] Linas Gelazanskas, Kelum AA Gamage, Managing renewable intermittency in smart grid: Use of residential hot water heaters as a form of energy storage, Electrical apparatus and technologies,2016, 19th international symposium on, 1-4, 2016.

[5] Q Yingning, S Juan, C Mengnan, W Hao, Model based wind turbine gearbox fault detection on SCADA data, 2014 – IET.

[6] DG Krishna, Preventive maintenance of wind turbines using remote instrument monitoring system, Power India Conference, 2012 IEEE fifth, 2012 – ieeexplore.ieee.irg.

[7] N Chen, R Yu, Y Chen, H Xie, Hierarchical method for wind turbine prognosis using SCADA data, IET Renewable power generation, 2016-IET.

[8] Y Debbag, EN Yilmaz, Internet based monitoring and control of a wind turbine via PLC, smart grid congress an fair, 2015 – ieeexplore. Ieee.org.

[9] Bindi Chen, YN Qui, PJ Tavner, WW. Song, A Wind turbine SCADA alarm pattern recognition, IET Digital Library,2011.

[10] Se – Yoon Kim, In-Ho Ra, Sung-Ho Kim, Design of wind turbine fault detection system based on performance curve, Soft computing and intelligent systems and 13th international symposium on advanced intelligent systems, 2012 joint 6th international conference on, 2033-2036, 2012.

[11] R Palma-Behnke, D Ortiz, L Reyes, A social SCADA approach for arenewable based micro grid – The Huatacondo – project, Power and energy society general meeting, 2011 IEEE, 1- 7, 2011

[12] Huan Long, Long Wang, Data-driver wind turbine power generation performance monitoring, IEEE Transactions on industrial electronics 62 (10), 6627-6635, 2015.

[13] Meik Schlechtingen, IInmar Ferreira Santos, Sofiane Achichew, Using data- mining approaches for wind turbine power curve monitoring: a comparative study, IEEE Transactions on Sustainable Energy 4 (3) 671 – 679, 2013.

[14] Bindi Chen, Peter C Matthews, Peter J Tavner, Automated on-line fault prognosis foe wind turbine pitch systems using supervisory control aqnd data acquisition, IET Renewable Power Generation 9 (5), 503-513, 2015.

[15] Y Yan, J Zhang, Using edge-detector to model wake effects on wind turbines, health management (PHM), 2014,IEEE, 2014 – ieeexplore.ieee.org.

[16] M Wilkinson, B Darnell, T van Delft, K Harman, Comparison of methods for wind turbine condition monitoring with SCADA data, IET Renewable Power, 2014 – IET.

[17] Ran Bi, Chengke Zhou, Donald M Hepburn, Applying instantaneous SCADA data to artificial intelligence based power curve monitoring and WTG fault forecasting, Smart grid and clean energy technologies (ICSGCE), 2016 International Conference on, 176 – 181, 2016.

[18] P Banglore, LB Tjernberg, An artificial neural network approach for early fault detection of gearbox bearings, IEEE Transactions on smart grid, 2015, ieeexplore.ieee.org.

[19] WAM Ghoneim, AA Helal, MG Abdel Wahab, Renewable energy resources and recovery opportunities in waste water treatment plants, Renewable energies for developing countries (REDEC), 2016 3rd international conference on 1-6, 2016.

[20] Y Zhao, D Li, A Dong, J Lin, D Kang, Fault prognosis of wind turbine generator using SCADA data, North America Power, 2016 – ieeexplore.ieee.

[21] www.sciencedirect.com>pii>pdf, SCADA based software for renewable energy management.

[22] http://www.researchgate.net> publication, SCADA data for wind turbine.

[23] http://www.google.com> patents, wind farm control system (SCADA).

[24] http://www.wind energy -the-fact.org>, SCADA data analysis

[25] Qing-yun Dai, 2Chun-ping Zhang and Hao Wu, " Research of Decision Tree Classification Algorithm in Data Mining", International Journal of Database Theory and Application Vol.9, No.5 (2016), pp.1-8