Reliability Engineering and System Safety 169 (2018) 554–569

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Data-driven learning framework for associating weather conditions and wind turbine failures

Maik Reder[1], Nurseda Y. Yürüşen[1], Julio J. Melero*

CIRCE - Universidad de Zaragoza, C/ Mariano Esquillor 15, 50018, Zaragoza, Spain

ARTICLE INFO

Keywords:
Wind turbine
Failure
Weather
Big data
Association rule mining
k-means clustering
Data mining
Machine learning
Operation & maintenance

ABSTRACT

The need for cost effective operation and maintenance (O&M) strategies in wind farms has risen significantly with the growing wind energy sector. In order to decrease costs, current practice in wind farm O&M is switching from corrective and preventive strategies to rather predictive ones. Anticipating wind turbine (WT) failures requires sophisticated models to understand the complex WT component degradation processes and to facilitate maintenance decision making. Environmental conditions and their impact on WT reliability play a significant role in these processes and need to be investigated profoundly. This paper is presenting a framework to assess and correlate weather conditions and their effects on WT component failures. Two approaches, using (a) supervised and (b) unsupervised data mining techniques are applied to pre-process the weather and failure data. An apriori rule mining algorithm is employed subsequently, in order to obtain logical interconnections between the failure occurrences and the environmental data, for both approaches. The framework is tested using a large historical failure database of modern wind turbines. The results show the relation between environmental parameters such as relative humidity, ambient temperature, wind speed and the failures of five major WT components: gearbox, generator, frequency converter, pitch and yaw system. Additionally, the performance of each technique, associating weather conditions and WT component failures, is assessed.

1. Introduction

Throughout the past years the wind energy sector has been growing rapidly. New wind turbine (WT) technologies and higher installed capacity are resulting in increased cost and complexity of operation and maintenance (O&M) tasks. Current maintenance strategies are mainly focusing on corrective and preventive actions, yet, operators could benefit extremely from applying predictive maintenance approaches.

This issue applies to many industrial fields and the solutions proposed in previous research are mainly based on the system age. For this purpose, e.g. Martorell et al. [1] studied age-dependent reliability models such as the proportional age set back (PAS) and proportional age reduction (PAR) frameworks. Both of them take into account maintenance efficiency, environmental and operational conditions as generic imperfect maintenance functions. In their study, three different environmental setups are considered indicating good, normal and bad conditions, which were found to highly influence the age and reliability functions. In Martorell et al. [2] these maintenance models have been extended by including the systems failure causes. It was shown that failures can be related to certain root causes, which need to be taken into consideration. Carlos et al. [3] applied the PAS models in wind energy, with the aim of minimizing the O&M cost of wind farms while maximising their revenue considering the maintenance frequency as decision variable. However, they neglected operational and environmental conditions in the maintenance scheduling framework. A similar situation can be seen also for age reduction models. Ding and Tian [4] studied an opportunistic maintenance approach for wind turbines by using an imperfect maintenance model, which depends on an age reduction ratio. However, they also were not including the environmental conditions. In general, environmental conditions are being neglected because of their complexity although having a very high impact on maintenance actions, as demonstrated by XiaoFei and Min [5]. They model the effect of system age and environmental conditions on the hazard rate function by introducing a stochastic process with two states, indicating normal and severe environment. They conclude that the system reliability degrades more rapidly when exposed to severe environmental conditions. Furthermore, the history of the environmental events is said to affect the hazard rate. Albeit, their study did not include the specific meteorological parameters themselves.

* Corresponding author.
E-mail addresses: mreder@fcirce.es (M. Reder), nyildirim@fcirce.es (N.Y. Yürüşen), melero@unizar.es (J.J. Melero).
[1] Shared first authorship - the authors contributed equally to this work.

https://doi.org/10.1016/j.ress.2017.10.004
Received 18 January 2017; Received in revised form 5 October 2017; Accepted 7 October 2017
Available online 9 October 2017
0951-8320/

Thus, previous research shows that for developing effective maintenance decision models, a deep understanding of the component degradation processes including the highly variable environmental conditions is required. Complex environmental and operational conditions must be defined and studied in detail to generate realistic ageing and reliability functions. Especially for systems that are interacting strongly with their surroundings. This paper shows how to complement these models by associating failure information with significant environmental variables by providing a tool that enables the interpretation of failures based on dynamic meteorological conditions. With this, the presented framework aims at supplementing existing maintenance and reliability models and enhancing the outcome of maintenance decision models. The framework will be applied to wind energy but can easily be adapted for other areas.

1.1. Background in wind energy

As stated in Van Kuik et al. [6], environmental conditions and their complicated combinations still have to be fully understood and methodologies have to be developed in order to analyse and correlate them to WT failures. Especially short term environmental variations, such as for example high-speed wind gusts, can cause a severe impact on components and need to be analysed extensively. This knowledge is essential for further use in WT reliability models and complex maintenance decision making tools.

Much research effort has been dedicated to identifying weather dependent failure rates and downtimes of WT components. Hahn [7] was the first to carry out an extensive analysis on the effects of weather on WT reliability. They showed that with rising average daily wind speed the failure rates of certain components increased as well. Tavner et al. [8] analysed the effect of monthly averaged wind speed conditions on component failure rates, using the Danish reliability database WindStats. Assuming the monthly mean wind speed across Denmark to be representative for the considered wind farms, they concluded that there is an annual periodicity in failure occurrences due to seasonal variation in weather conditions. Faulstich et al. [9] carried out an analysis on the effects of wind speed on WT downtimes, taking into account energy- and time-based availability.

Along with high wind speeds, literature also frequently addresses temperature and humidity as critical parameters. Especially humidity can cause corrosion and other highly dangerous degradation processes. Costa at al. [10] stated that correlating relative humidity and failures using classical analysis approaches, such as the analysis of variance (ANOVA) technique, is one of the most challenging tasks. Tavner et al. [11] cross-correlated component failures with average monthly maximum and mean wind speed, maximum and minimum air temperature and average daily mean relative humidity. Their study refers to three specific wind farms using yearly and monthly mean weather conditions taken from closely located meteorological stations. However, neither the actual conditions at the time of failure occurrence nor short-time weather events, e.g. high-speed gusts, were considered. Also, only old wind turbine technologies with a rated power of 300 *kW* were analysed. They suggested that future studies should ensure to include modern WTs and short-time weather variations, along with details on maintenance strategies. Wilkinson et al. [12] showed the impact of environmental conditions on WT reliability using an availability database and WT downtimes. The environmental data were taken from the *Modern-Era Retrospective Analysis for Research and Applications* (MERRA) database and the results showed that there is a strong relationship between wind speed, temperature and high downtime. Wilson and McMillan [13] modelled the relationship between failures and weather conditions using Markov Chains, based on data for one wind farm over a relatively short period of time.

Recently, the possibilities of using Supervisory Control and Data Acquisition (SCADA) data for these purposes have been discovered. SCADA systems are among the standard equipment of modern wind turbines and can provide a huge amount of information on many operational WT parameters, conditions of its components, as well as external weather parameters. Such extensive data require sophisticated analysis techniques and high computational effort. Thus, a number of research projects have been trying to apply data mining and machine learning techniques using WT SCADA data. For example Kusiak and Verma [14] used apriori rule data mining to find frequent item sets in SCADA data in order to identify WT pitch system faults. Machine learning can reveal very interesting results on component degradation, however, they strongly depend on the availability of a considerable amount of input data.

Previous studies evaluating weather effects on component failures are mostly limited to a very low number of analysed turbines and/or outdated WT technologies. The environmental data considered are mostly comprised of yearly or monthly average values obtained at close-by located weather stations, and short-term changes in these parameters are not taken into account. However, the turbines' SCADA systems provide a rich data source and can give a realistic representation of the conditions at the WTs. In an earlier study by Reder and Melero [15], the wind speed effects on WT component reliability have been assessed using 10-minute mean SCADA data taken from failed WTs. It has been found, that short term wind changes need to be analysed further, in order to obtain a clear image of the conditions leading to component failures. Also, modern WTs and further environmental parameters should be taken into account. Hence, there is a significant need for extending previous studies, such as [11,12]. Certainly, failures are often caused by cumulative stress over a large period of time. However, especially short-time weather events including high-speed gusts and ambient temperature changes, can cause the final impact on a component and lead to its failure and are an important factor in weather dependent failure analysis.

Considering the fact that this will be based on a significantly higher amount of data than previous analyses, methodologies capable of handling big data and deriving meaningful results have to be developed. Thus, in this study an effective framework for processing and correlating the different inputs is introduced. This applied to a historical failure database of modern wind turbines located in Spain as well as environmental data taken from the affected wind turbines' SCADA system and weather stations close to the wind farm. Five components are considered: gearbox, generator, pitch system, yaw system and frequency converter. These were chosen according to the results of previous studies, e.g. Reder et al. [16], where the components with the highest share of the total WT failure rates and downtime were identified.

A summary of the relevant literature and the problem statement were presented in this section. The remainder of this paper is structured as follows. The input data for a case study, to which the framework will be applied to, are described in Section 2, comparing the available operational years, sample sizes and WT technology types to previous studies. The established framework including the data pre-processing and the methodology behind the supervised and unsupervised learning algorithms are described in Section 3. Here, also an example for the calculation of the evaluation metrics for interpreting the results is given. The results of applying the established framework to the case study are presented in Section 4. Conclusions are given in Section 5, where the key findings and the possible application areas are emphasized.

2. Case study - data set description

The framework was tested and verified using the data sets presented in this section. These are comprised of historical failure data, SCADA data, historical weather data and expert judgements.

2.1. Historical failure data

As shown in Table 1, the historical failure database contains a significantly higher number of turbines compared to previous studies. It represents modern three bladed, pitch regulated WTs with rated power between 660 kW and 2000 kW, produced by the same WT manufacturer.