Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Modelling Wind Turbine Failures based on Weather Conditions

To cite this article: Maik Reder and Julio J. Melero 2017 *J. Phys.: Conf. Ser.* **926** 012012

View the article online for updates and enhancements.

Related content

- Wind Turbine Failures - Tackling current Problems in Failure Data Analysis
  M D Reder, E Gonzalez and J J Melero
- Advances in Constitutive and Failure Models for Sheet Forming Simulation
  Jeong Whan Yoon and Thomas B Stoughton
- Condition monitoring of a wind turbine doubly-fed induction generator through current signature analysis
  Estefania Artigao, Andres Honrubia-Escribano and Emilio Gomez-Lazaro

This content was downloaded from IP address 85.251.108.81 on 01/12/2017 at 18:39

Modelling Wind Turbine Failures based on Weather Conditions

Maik Reder and Julio J. Melero

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

E-mail: `mreder@fcirce.es`

Abstract. A large proportion of the overall costs of a wind farm is directly related to operation and maintenance (O&M) tasks. By applying predictive O&M strategies rather than corrective approaches these costs can be decreased significantly. Here, especially wind turbine (WT) failure models can help to understand the components' degradation processes and enable the operators to anticipate upcoming failures. Usually, these models are based on the age of the systems or components. However, latest research shows that the on-site weather conditions also affect the turbine failure behaviour significantly. This study presents a novel approach to model WT failures based on the environmental conditions to which they are exposed to. The results focus on general WT failures, as well as on four main components: gearbox, generator, pitch and yaw system. A penalised likelihood estimation is used in order to avoid problems due to for example highly correlated input covariates. The relative importance of the model covariates is assessed in order to analyse the effect of each weather parameter on the model output.

1. Introduction

Over the past years wind turbine (WT) operation and maintenance (O&M) has become an emerging field of research. Wind farm operators can profit substantially from shifting towards predictive maintenance strategies rather than employing corrective tasks. Understanding the wind turbine degradation in operating conditions plays an important role in this transition. Predictive maintenance models take into account age-dependent component reliability degradation. Modelling the accessibility of wind farms is usually carried out including the weather conditions at site. However, the impact of meteorological conditions on the component lifetime is mainly neglected. Yet, taking into account the environmental conditions that are more likely to provoke component failures, could enhance predictive maintenance models significantly.

Several studies have been carried out investigating the failure behaviour of wind turbines. Many different data bases have been analysed with the aim of quantifying failure rates and downtime related to the wind turbine systems and their components. Examples are given in [1–6]. The change in failure rate over time is frequently modelled using Weibull distributions, deriving the so-called bathtub curve. The latter is divided into three phases - early, random (useful life) and ageing phase. Early life failures are characterised by a decreasing rate over time, while ageing or wear-out is represented by an increasing failure rate. During the useful life period the failure rates are assumed to be constant.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

Previous studies have shown that not only the turbine age, but also certain combinations of weather conditions can affect their life-time negatively. During the useful life-time, the assumption of constant failure rates does not always hold true, especially when considering shorter time-intervals, such as for example the failure occurrences on a monthly basis. There are significant variations in failure rates throughout the year that need to be taken into account by the wind farm operators. With this they can react properly to upcoming failure by initiating preventive or opportunistic maintenance actions.

Much research effort has been dedicated to identifying the critical meteorological parameters that influence the WT failure behaviour negatively. One of the first extensive analysis on the effects of weather on WT reliability was presented by Hahn et al. [7], showing increased failure rates of certain WT components with rising average daily wind speeds. In Tavner et al. [8] an annual periodicity in failure rates due to seasonal variation in weather conditions is reported, by analysing the correlation between monthly averaged wind speed conditions and component failures. The effects of wind speed on WT downtimes with regards to the turbines' energy- and time-based availability is examined by Faulstich et al. [9]. Wilson et al. [10] use artificial neural networks to show that especially the gearbox and generator are very likely to fail in changeable wind conditions. McMillan et al. [11] use Markov Chain models to calculate the seasonal component failure rates and wind farm availability based on weather conditions. In Wilkinson et al. [12] a strong relationship between high wind turbine downtime and temperature and wind speed was found. Tavner et al. [13] cross-correlate component failures with average monthly maximum and mean wind speed, maximum and minimum air temperature and average daily mean relative humidity. The studies presented by Wilson et al. in [14, 15] use Markov Chains and Monte Carlos simulations to model the WT failure behaviour with respect to certain weather conditions. Recently, Faulstich et al. [16] included wind speed in an additive Weibull failure rate model in addition to the three age-dependent phases of the bathtub curve. A study published by Slimacek et al. [17] indirectly considers the environmental conditions for modelling the rate of occurrence of wind turbine failures by taking into account the number of stops due to harsh weather conditions. They conclude that the number of stops due to harsh environmental conditions are the most important drivers of their model.

The mentioned research has proved that the failure behaviour of wind turbines and their components is strongly influenced by the meteorological conditions the systems are exposed to. Nonetheless, no models have been developed yet, to directly describe the WT failure behaviour based on combinations of external covariates. Thus, this work presents a novel approach to model the failures of WTs during the useful life including the effect of environmental conditions. The model will be applied to a 2 MW case study wind farm and the failures of the whole turbine system as well as four main components are modelled. In this context, failures are defined as events that can be associated to a component breakdown, which causes a WT stop and needs intervention such as replacement or repair.

2. Methodology

A regression model based on a generalised linear model (GLM) is applied to the data. The model is set up with a Poisson response distribution and a logarithmic link function. Subsequently, a *ridge* regression is employed to estimate the model parameters.

The *ridge* regression, see e.g. [18] and [19], is a penalised parameter estimation technique, which is frequently used to estimate the parameters of a regression problem with a high number of covariates. It introduces a penalty term on the squared ℓ_2-norm of the coefficient vector.

The penalisation helps to avoid over-fitting due to the high number of covariates. Although *ridge* regression is known to introduce a bias by shrinking the model coefficients towards zero, it is capable of significantly reducing the variance, in comparison to for example ordinary least square methods.

The Poisson regression model is given by:

$$\hat{y} = \exp(\beta_0 + \sum_{i=1}^{N} \beta_i x_i^T) \quad , \tag{1}$$

where $\hat{y}$ is the failure rate, x are the external covariates, N the number of observations and β are model parameters. Eq. 1 is fitted by maximising the penalised likelihood:

$$max_{\beta,\beta_0} \quad \frac{1}{N} \sum_{i=1}^{N} \left(y_i (x_i^T \beta + \beta_0) - e^{x_i^T \beta + \beta_0} \right) - \lambda \left(\frac{1}{2} \sum_{i=1}^{N} \beta_i^2 \right) \quad , \tag{2}$$

with the number of observations N and a penalisation parameter λ. The latter dictates the degree of shrinkage. Thus, larger values for λ result in more shrinkage and higher bias. At the same time, however, with increasing λ values the variance decreases. Choosing a value of the penalisation parameter λ is realised by fitting a model with maximum regularization and subsequently decreasing it until the model becomes too complex and overfitting occurs. To ensure computationally efficient parameter estimation the *L-BFGS* solver, see e.g. [20], is used.

2.1. *Case Study*

The model is applied to a case study including 30 turbines located in a wind farm in Spain. The WTs are geared, three bladed and pitch-regulated machines, with a rated capacity of $2MW$ each. The terrain complexity of the site is 4, which was classified using the normative IEC 61400-12-2, [21]. The data set consists of failure logbooks, SCADA and meteorological tower (met-mast) data collected during three years of operation. At the start of the data collection the turbines were five years old. An average wind farm year is modelled, whereas the observation period is introduced into the model by means of an exposure variable called model offset. Thus, the model outcome can be considered as the rate of failure occurrence in an average operational year.

2.2. *Data*

In this study, the monthly failures of an average year will be modelled. In the results section the months January to December are indicated with the numbers 1 to 12. The model covariates include the monthly average wind speed (WS) and turbulence intensity (TI), and the monthly maximum wind speed (MaxWS), measured at a height of 45 meters at the wind farm met-mast. Additionally, the monthly mean relative humidity (RH), precipitation (Rain) and ambient temperature (Temp) taken from closely located weather station are included. Table 1 shows the minimum, median, mean, maximum and standard deviation values for each meteorological parameter of the input data. Additionally, as an indicator for the time operating at full capacity the monthly mean power production (PWR) taken from the turbines SCADA systems divided by the rated capacity, was chosen. In order to avoid problems due to the different covariate magnitudes, the latter are centred to a mean of 0 and standardised by dividing by their standard deviation.

In a first step the model is applied to the whole data base, without further distinguishing between the failed components. Subsequently, the failure data of four main components: the gearbox, generator, pitch and yaw system are extracted from the same set and the model is applied again. In order to analyse the importance of each input variable, the standardised model coefficients are compared. This is commonly done to interpret which of the covariates contributes the most to modelling the output and helps to see which weather conditions are important for modelling the failures.

Table 1. Statistical properties for each meteorological parameter .

Variable	Minimum	Median	Mean	Maximum	Standard Deviation
WS (m/s)	4.238	6.535	6.675	10.87	2.408
TI (-)	0.105	0.118	0.119	0.134	0.012
MaxWS (m/s)	16.33	23.29	23.13	30.89	5.250
Temp (°C)	4.414	12.18	12.59	20.92	6.020
RH (-)	0.634	0.657	0.673	0.719	0.035
Rain (mm)	5.067	37.20	39.78	94.00	30.02

2.3. Model Evaluation

The model performance is evaluated using the metrics root mean square error (RMSE), mean absolute error (MAE) and the coefficient of determination R^2. Due to confidentiality reasons regarding the failure data and in order to enable the comparison of the five data input cases, the first two metrics are displayed in a normalised form indicated by NRMSE and NMAE. The metrics are defined as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2}, \qquad NRMSE = \frac{RMSE}{y_{max} - y_{min}}; \tag{3}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \sqrt{(\hat{y}_i - y_i)^2}, \qquad NMAE = \frac{MAE}{y_{max} - y_{min}}; \tag{4}$$

$$R^2 = 1 - \frac{\sigma(\hat{y} - y)^2}{\sigma(y)^2}; \tag{5}$$

where $\hat{y}$ is the modelled variable, y the original data value and σ denotes the standard deviation. The values y_{max} and y_{min} are the maximum and minimum of the original data.

3. Results and Discussion

A summary of the model performance metrics for the five different cases is shown in Table 2. The R^2 value ranges from 0 to 1 and indicates how shows the goodness of fit of the model to the data, where 1 indicates the best fit. It can be seen that in general the model performs well for all five failure classes.

The models for the generator and pitch system, however, showed lower R^2 and higher error values than the other ones. This indicates that for these two components additional covariates, which were not included in the model, could be of importance. This will be assessed in further studies.

Figure 1 shows the original failures (black) and the modelled data (green) for wind turbine system failures without distinguishing between their components. The values are normalised to the maximum number of failures for confidentiality reasons. As an example for a separately

Table 2. Model metrics for the different components.

Measure	All failures	Gearbox	Pitch System	Generator	Yaw System
R^2	0.7982	0.8139	0.5492	0.5623	0.8837
NMAE	0.1087	0.1034	0.2456	0.1697	0.0694
NRMSE	0.1218	0.1280	0.2894	0.2079	0.0884

modelled WT component, Figure 2 displays the original and the modelled data for yaw system failures.

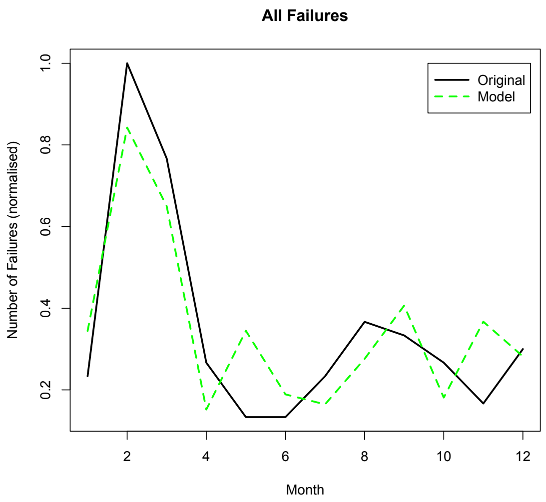

Figure 1. Model vs. data: all failures

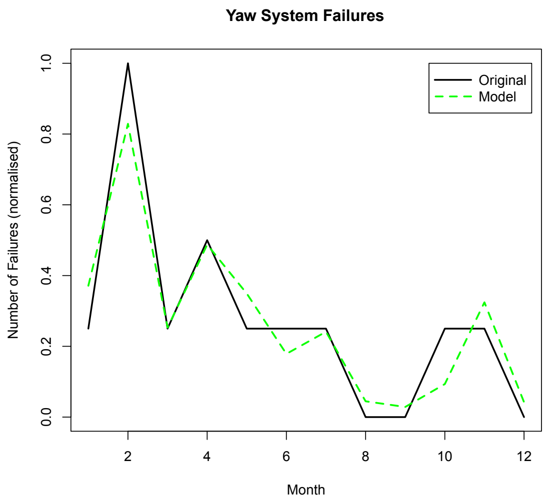

Figure 2. Model vs. data: yaw system

It can be seen that in both cases the highest failure occurrences are recorded within the second to fourth month of the year. This is consistent with previous literatures, where it is stated that failures mainly occur during the winter months, e.g. [11] and/or the transition periods between seasons, see e.g. [22].

Figures 3 to 7 show the standardised coefficient magnitudes. Positive coefficients (blue) indicate that with increasing covariates the output increases as well, while negative coefficients (orange) respectively describe the opposite behaviour. As the input data were centred and scaled, these coefficients can be interpreted as the influence of every model covariate on the dependent variable. For an easier interpretation the plots show the magnitudes as fraction of the most important regressor, which is scaled to 1.

As a multivariate model is used, the coefficients have to be interpreted with care though. The isolated effect of one regressor can theoretically only be interpreted, if the other regressors were kept constant. As this is very unlikely for environmental conditions, which in many cases are correlated, the effect should be analysed considering combinations of the meteorological parameter. Furthermore, only relative effects can be investigated.

As shown in Figure 3 the input parameters TI, $Temp$ and WS have the highest importance for modelling the data including failures of any WT component. High monthly mean turbulence intensity, high mean wind speed and low temperatures play a significant role. This is consistent

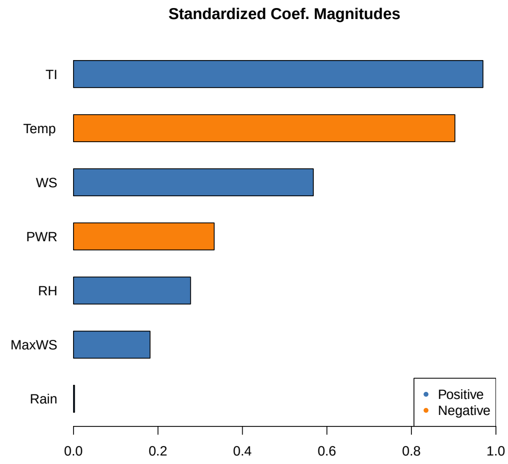

Figure 3. Variable importance: all failures

with the aforementioned literature, which stated that higher mean wind speeds and low temperatures can be correlated to higher number of WT failures. Low PWR values seem to influence the failure behaviour as well, but not very dominantly. As the active power output is usually positively correlated to wind speed, having positive coefficient magnitudes for wind speed and negative ones for the PWR variable this might seem contradictory. However, under faulty conditions the wind turbine is often performing below the expected capacity and underperformance can be seen as indicator for component failures.

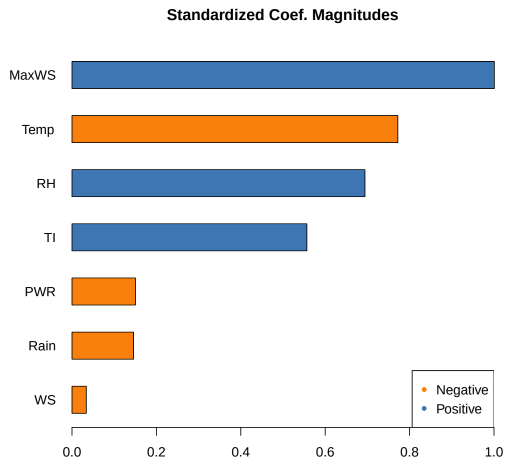

Figure 4. Variable importance: pitch system

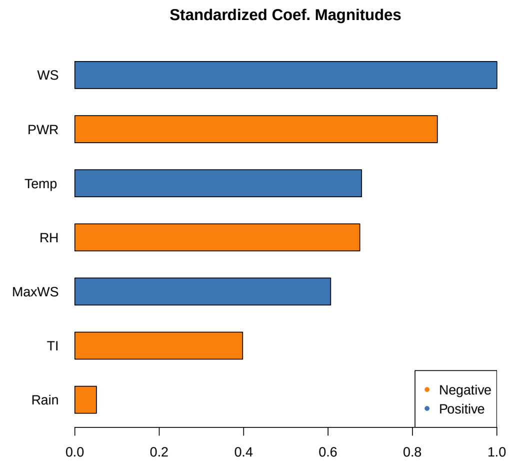

Figure 5. Variable importance: yaw system

In literature, e.g. [6] and [23], it has been shown that the different components react differently to certain combinations of environmental conditions. Thus, these should be analysed separately in order to obtain more meaningful results.

For the pitch system model (Figure 4) low temperatures and high monthly maximum wind speeds are significant. In addition to that, high relative humidity and turbulence intensity play a role when modelling the pitch system failures. These are the conditions, where the pitch system is mostly active in order to regulate the rotor speed and thus is subject to higher stresses and possible damages.

Figure 5 shows the standardised coefficient magnitudes for yaw system failures. The differences between the coefficient magnitudes are not as large as they are for the other WT components. Many meteorological factor seem to play a role in this model. A clear under-performance of the turbine can be seen, as rising wind speeds and falling PWR values lead to increased numbers of yaw system failures. This leads to the assumption that despite the higher mean wind speeds, the wind direction changed frequently and the yaw system had to be constantly searching for the best wind direction. Thus, a possible yaw-misalignment resulted in higher wear and under-performance. However, this should be investigated in more detail by including the wind direction in further studies.

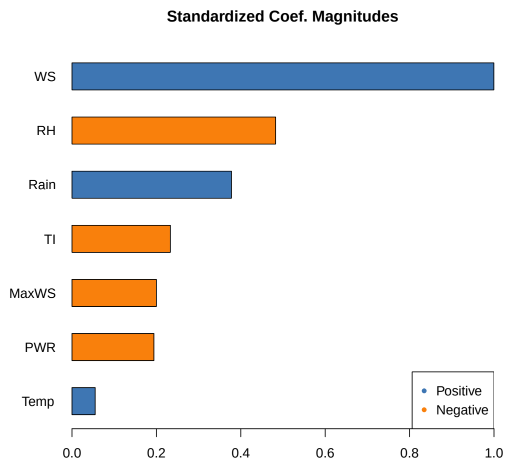

Figure 6. Variable importance: gearbox

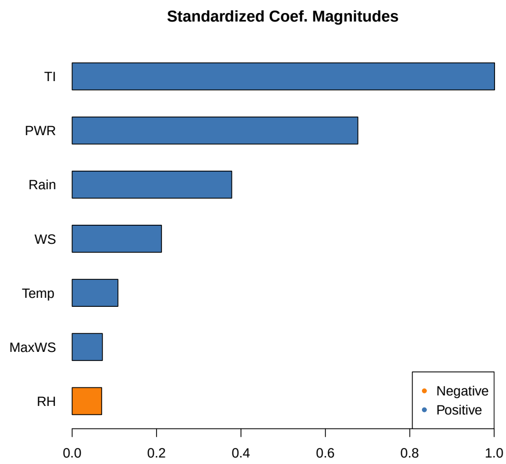

Figure 7. Variable importance: generator

Modelling the number of gearbox failures is influenced mostly by high wind speeds, as displayed in Figure 6. With higher mean wind speeds, the load on the gearbox is increasing and the component is more likely to fail.

The generator failure model (Figure 7) is mostly driven by increasing turbulence intensity and power output. Elevated TI introduces higher loads on the generator that has to adapt to these varying input speeds. Positive coefficients for the variables PWR and WS indicates that no under-performance was recorded before the failures. Additionally, it states that with higher power production, the generators are more likely to fail. Generator failures usually occur abruptly due to sudden changes in turbulence intensity and highly varying wind conditions, see e.g. [24]. Furthermore, the amount of precipitation plays a significant role, as water intrusion highly affects electronic equipment.

4. Conclusion

A methodology to model wind turbine failures based on the environmental conditions to which the turbines are exposed to has been introduced and applied to five failure classes. It was shown that the models work well and that it is possible to derive the relative importance of each input variable in order to analyse the effect of different covariates on the modelling output. This was used to assess which combinations of environmental conditions affected the respective component the most. Furthermore, the models have shown, that mechanical components, especially the yaw system, come along with a significant turbine under-performance. Thus, these failures could be anticipated by immediately detecting under-performance in combination with certain environmental conditions. Hence, the herein presented models serve to identify which environmental parameters influence the failure behaviour of certain wind turbine components. This information can help to anticipate failures and significantly enhance predictive maintenance models.

5. Outlook

In future studies the models will be refined and the performance of different modelling techniques and input variables will be investigated. These should be capable of handling high numbers of inputs and reduce collinearity. Further environmental variables will be tested, in order to determine the ones that mostly affect the components' failure behaviour. Instead of relative humidity, absolute humidity is expected to lead to better interpretation of the results in terms of the effect of humidity on the components. Especially yaw system failures should be modelled taking into account wind direction and wind shear.

The model will be tested on larger data sets, including more wind farms and different turbine technologies. The objective is to establish generic models describing the influence of environmental conditions on WT component failures. These can then be used for failure prediction and for assessing the environmental conditions that have the highest impact on WT components. Furthermore, these models are expected to enhance predictive O&M strategies and contribute to decreasing the overall cost of wind farms.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 642108.

References

[1] Hahn B, Durstewitz M and Rohrig K 2007 *Wind Energy* 329–332
[2] Bard J, Faulstich S and Lynding P 2011 January (Fraunhofer Inst itute for Wind Energy and Energy System Technology (IWES))
[3] Wilkinson M, Harman K, Hendriks B, Spinato F and Van Delft T 2011 *Proceedings of the European Wind Energy Conference* (Brussels, Belgium: European Wind Energy Association) pp 1 – 8
[4] Carroll J, McDonald A and McMillan D 2015 *Wind Energy* **19** 1107–1119
[5] Reder M, Gonzalez E and Melero J J 2016 *Journal of Physics: Conference Series* **753** 1–10
[6] Reder M and Melero J J 2016 *Wind Europe Conference* (Hamburg, Germany)
[7] Hahn B 1997 *FGW-Workshop Einfluss der Witterung auf Windenergieanlagen,*
[8] Tavner P, Edwards C, Brinkman A and Spinato F 2006 *Wind Engineering* **30** 55–72
[9] Faulstich S, Lyding P and Tavner P 2011 *European Wind Energy Conference* (Hamburg, Germany)
[10] Wilson G, Mcmillan D and Ault G 2012 *International Conference on Sustainable Power Generation and Supply (SUPERGEN 2012)* (Hangzhou, China)
[11] McMillan D and Wilson G 2013 *EWEA Offshore 2013* (Frankfurt, Germany)
[12] Wilkinson M, Van Delft T and Harman K 2012 *European Wind Energy Conference* April (Copenhagen, Denmark)
[13] Tavner P, Greenwood D M, Whittle M W G, Gindele R, Faulstich S and Hahn B 2013 *Wind Energy* **16** 175–187
[14] Wilson G and McMillan D 2014 *Journal of Physics: Conference Series* **524** 1–10
[15] Wilson G and McMillan D 2014 *Safety, Reliability and Risk Analysis: Beyond the Horizon - Proceedings of the European Safety and Reliability Conference, ESREL 2013* 801–809
[16] Faulstich S, Berkhout V, Mayer J and Siebenlist D 2016 *Journal of Physics: Conference Series* **749** 1–11
[17] Slimacek V and Lindqvist B H 2016 *Wind Energy* **19** 1991–2002
[18] Hoerl A E and Kennard R W 1970 *Technometrics* **12** 55–67
[19] Antoniadis A, Gijbels I and Nikolova M 2011 *Annals of the Institute of Statistical Mathematics* **63** 585–615
[20] Byrd R H, Nocedal J and Schnabel R B 1994 *Mathematical Programming* **63** 129–156

[21] IEC 2013 International Standard IEC 61400: Wind turbines Part 12-2; Ed. 1, 2013/03, Power performance of electricity-producing wind turbines based on nacelle anemometry Tech. rep. Geneva, Switzerland
[22] Guo P and Bai N 2011 *Energies* **4** 2077–2093
[23] Reder M, Yürüşen N Y and Melero J J 2018 *Reliability Engineering & System Safety* **169** 554–569
[24] T Wei 2010 *Wind Power Generation and Wind Turbine Design* (Southampton, Boston: WIT Press)