RESEARCH ARTICLE

Modelling the effects of environmental conditions on wind turbine failures

Maik Reder | Julio J. Melero

CIRCE - Universidad Zaragoza, Campus Río Ebro, C/ Mariano Esquillor Gómez 15, Zaragoza 50018, Spain

Correspondence
Julio J. Melero, CIRCE - Universidad Zaragoza, Campus Río Ebro, C/ Mariano Esquillor Gómez 15, Zaragoza 50018, Spain.
Email: melero@unizar.es

Abstract
Operation and maintenance is one of the main cost drivers of modern wind farms and has become an emerging field of research over the past years. Understanding the failure behaviour of wind turbines (WTs) can significantly enhance operation and maintenance processes and is essential for developing reliability and strategic maintenance models. Previous research has shown that especially the environmental conditions, to which the turbines are exposed to, affect their reliability drastically. This paper compares several advanced modelling techniques and proposes a novel approach to model WT system and component failures based on the site-specific weather conditions. Furthermore, to avoid common problems in failure modelling, procedures for variable selection and complexity reduction are discussed and incorporated. This is applied to a big failure database comprised of 11 wind farms and 383 turbines. The results show that the model performs very well in several situations such as modelling general WT failures as well as failures of specific components. The latter is exemplified using gearbox failures.

KEYWORDS
failure, operation and maintenance, reliability, weather, wind turbine, zero-inflated models

1 | INTRODUCTION

Wind energy generation has evolved to one of the most important renewable energy sources worldwide. To compete with conventional energy sources, however, the overall cost of wind energy must yet be lowered significantly. Being a main cost driver of modern wind farms (WFs), operation and maintenance (O&M) has become an emerging field of research. Operators can profit substantially by shifting from corrective towards predictive maintenance strategies.

Understanding the failure behaviour of wind turbines (WTs) and their components plays an important role in this transition. Sophisticated reliability models are used to anticipate failures and to establish cost-effective O&M strategies.

Reliability models usually take into account the operational life of the system or component using so-called lifetime distributions. These are probability density functions defined over a usage parameter, such as time, distance, and cycles.

Weibull distributions are the most widely used lifetime distributions in reliability engineering; see, eg, O'Connor and Kleyner. When modelling the failure rates over time, these are resulting in the famous bathtub curve.

The latter is divided into 3 phases—early, useful life, and ageing phase. The early life is characterized by a decreasing failure rate over time and modelled with a Weibull parameter $\beta < 1$. Ageing or wear-out is represented by an increasing failure rate with $\beta > 1$. During the whole useful-life period, the failure rates are assumed to be constant ($\beta = 1$).

For WF operators, the failure rates on relatively short time intervals, eg, on a monthly basis, are of high interest when setting up their maintenance strategies. Assuming constant failure rates (eg, during the entire useful-life period) can lead to wrong conclusions and higher O&M expenses. Furthermore, lifetime models are meant for situations where the operational life is the only measurable driver of the subject’s reliability degradation, and they do not account directly for the effect of the surrounding meteorological conditions. Wind turbine systems are highly exposed to weather phenomena, and as stated by Kuik et al., apart from the system’s operational life, reliability degradation of WT components is often related to complex combinations and sequences of operational and environmental conditions. This can result in highly varying component degradation processes throughout the year, which still need to be fully understood. Much research effort has been dedicated to analysing failure data of WTs and their
components; examples can be found in literature.\textsuperscript{3-7} Also, the environmental conditions before WT failures have been analysed in previous research.\textsuperscript{8-14} These studies have shown that not only the turbine age but also certain combinations of weather conditions can affect their lifetime negatively.

Fewer studies have been conducted on actually modelling the failure behaviour with respect to the environmental conditions the WTs are exposed to. A work by Wilson et al\textsuperscript{15} uses artificial neural networks with the aim of analysing the effects of rainfall, pressure, relative humidity (RH), temperature, wind direction, wind speed, and gust speed at ground level on different WT components. Wilson and McMillan\textsuperscript{16} subsequently present a nonparametric mixture model to compare the distributions of weather conditions, including relative humidity, temperature, and wind speed, in the presence of failures to normal site conditions. Further, Wilson and McMillan\textsuperscript{17} establish a Markov chain Monte Carlo model to analyse the effects of wind speed on O&M cost. A work presented by Faulstich et al\textsuperscript{18} proposes an additive Weibull failure rate model for WT rotor blades based on their age, as well as the hours of full load and wind speed (overload). A recently published study by Slimacek and Lindqvist\textsuperscript{19} uses a Poisson-gamma model for modelling the rate of occurrence of failures (ROCOF) with a time-constant ROCOF as base function. A prox-i-covariate indicating the number of stops caused by external natural factors was used to represent the harshness of the environmental conditions. They conclude that the latter is the most significant factor for modelling WT reliability and should be considered in further studies. Nonetheless, their database did not permit to include environmental covariates directly.

Previous studies have successfully shown that correlations between environmental parameters and WT failures exist and models describing their effects on the component degradation were established. However, in most cases, the environmental parameters were not included directly as covariates. In other cases, the influence of different weather variables on the WT failure behaviour is only modelled separately for each weather variable.

This paper proposes a novel approach to model monthly WT failures directly incorporating complex combinations of 6 on-site environmental parameters, as well as 5 turbine-specific and operational variables. The performance of several different multivariate regression models is compared using a large failure database including various turbine technologies. With increasing model complexity, more sophisticated parameter estimation and variable selection (PEVS) techniques have to be considered in order to account for common problems related to, eg, overdispersed failure data, heterogeneity, optimal covariate subsets, and multicollinearity. The best combination of regression model, parameter estimation, and variable selection techniques is determined and proposed for further use in the field.

The remainder of this paper is organized as follows. Section 2 describes in detail the data set and model covariates. Section 3 discusses the challenges of recent regression modelling and parameter estimation techniques. Additionally, the objectives and the novelty of this study in the context of the discussed literature are presented. Section 4 gives the mathematical formulations of the used models. The results and discussions are presented Section 5. Finally, the conclusions and future objectives are given in Section 6.

2 \hspace{1em} DATA

The data used in this work are composed of historical failure logbooks, WT SCADA data, and the WFs’ measurement tower (met mast) data. Additionally, measurements obtained at meteorological stations close to the respective WFs are used. Figure 1 presents the composition of the used data regarding the failed main components. It is shown that the gearbox contributed to a very high number of failures and will be used to test the model performance on component level. The gearbox has been identified as one of the most critical WT components.\textsuperscript{7,20} Furthermore, although being embedded within the nacelle, it is highly affected by environmental conditions.\textsuperscript{15,21,22}

2.1 \hspace{1em} WF specifications and historical failure data

Table 1 summarizes the WF characteristics and failure data. The observation period is 3 years (January 2013 to December 2015), resulting in 1149 operational WT years for 11 WFs and, in total, 383 turbines. All farms are located on-shore in different areas of Spain and operate 3 bladed,
Table 1 shows the number of WT system and gearbox failures registered during the observation period, as well as the failures per turbine and month. Figure 2A,B displays histograms for the monthly WT system and gearbox failures, respectively.

### 2.2 Weather and operational data

This work considers 6 environmental parameters. These were identified in literature as the most critical ones influencing the WT failure behaviour negatively. The monthly mean wind speed (WS), turbulence intensity (TI), and monthly maximum wind speed (MaxWS) were measured at a met-mast height of 45 m. The monthly mean RH, ambient temperature (Temp), and total monthly precipitation (Rain) are taken from closely located weather station. Figures 3A to 3F show histograms of the measured environmental covariates. The turbulence intensity is only available for WF-A and WF-C and does not enter in the model for all WT technologies. But the influence of TI on the outcome is discussed separately for WF-C in section 5.5.

As operational indicator, the average monthly active power (taken from the SCADA system) in percent of the turbine’s rated capacity, is introduced as the covariate PWR. This covariate indicates indirectly how long and with how much capacity the turbines are operating on average during each month. The hub-height, rated capacity, and turbine diameter also enter the models to distinguish between the different WT technologies. To avoid problems due to distinct covariate magnitudes, all inputs are standardized: centred to a mean of 0 and divided by their standard deviation (scaled). This allows the relative comparison of the variable importance within the context of the respective model.

Figure 4 shows the correlation between the input variables. TI is excluded, as it was only available for 2 WFs. Red indicates negative correlation, and blue positive correlation. Although only pairwise correlation is displayed, it can be seen that many environmental variables are correlated to each other.

### 2.2.1 Model memory

Previous studies have shown that the weather conditions can have a delayed and/or cumulative effect on the WT components. Thus, the environmental covariates of the previous month will be included in the model. This is further referred to as model memory, and the additional covariates...
are indicated with a suffix .mem. With this, the total number of covariates rises from 11 to 17. The effects of this extension are discussed in section 5.3.

3 | BACKGROUND AND PROBLEM STATEMENT

Modelling failures based on the environmental conditions can be realized in several ways. For continuous responses such as the time to failure of nonrepairable systems with externally induced stresses, accelerated life models\textsuperscript{25} or proportional hazard models\textsuperscript{26} may be alternatives to the Weibull models. These, however, are not appropriate when modelling discrete responses, such as the ROCOF. The latter are often modelled using Poisson regression, which is very similar to applying the proportional hazard model. Nevertheless, Poisson models have shown to perform poorly in certain situations causing several problems. Section 3.1 discusses the problems that come along with the Poisson regression and introduces 3 advanced alternatives. The latter result in increased model complexity, and more sophisticated PEVS methods must be considered, as shown in section 3.2. The presented models in connection with the estimation techniques will be applied to the data, and the most appropriate combination will be determined, as discussed in section 3.3.
3.1 | Failure models

The Poisson distribution dictates that its mean is equal to its variance. Serious problems can arise with overdispersed count data, having a larger variance than the Poisson distribution permits. Overdispersion can occur because of a variety of reasons, such as the presence of unobserved effects (heterogeneity) and/or an abundance of zero counts in the data. In particular, when modelling the actual monthly failure observations (not the average of a typical year), the distributions of the latter are usually highly right skewed. This was shown in Figure 2A,B of Section 2, where the number of zeros is significantly larger than the non-zero counts. In these cases, Poisson models are likely to overestimate the number of failures.

The Poisson model is often extended to the negative-binomial (NegBin) regression model, a mixture of the Poisson and a gamma distribution accounting for the unobserved effects. In WT reliability modelling, Slimacek and Lindqvist19 use a Poisson-gamma model and maximum likelihood estimation (MLE) for estimating the model coefficients.

However, this could also be handled using more advanced modelling techniques, for example, by considering 2 separate processes to model the failure counts and the zeros. One process generates the structural zeros using a binary distribution; the other one is governed by a regular count process, such as the Poisson or, even better, the NegBin distribution that generates counts including occasional zeros and takes into account heterogeneity. These types of models are often referred to as zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) and have successfully been applied in many different research areas.27 Having 2 separate processes, these models tend to have a very high number of regressors and are sometimes avoided for this reason. In combination with a suitable parameter estimation other than MLE and advanced variable selection methods, however, they can lead to much better results than other modelling techniques.

3.2 | Parameter estimation and variable selection approaches

When dealing with high-dimensional regression problems, two of the most serious issues are overfitting (e.g., a model with too many input variables) and strongly correlated covariates (multicollinearity). Variable selection and regularization are methods applied during the parameter estimation with the aim of reducing the number of model covariates to the most relevant ones. They often provide remedy for the mentioned issues.

Parameter estimation is usually conducted via MLE, a popular procedure of finding the set of model parameters that maximizes a known log-likelihood function. Maximum likelihood estimation, however, does not provide any variable selection criteria and can lead to overfitting and high variance of the coefficient estimates when including more model covariates.

Also, for small sample sizes and in the presence of high multicollinearity, MLE can be heavily biased. To avoid overfitting and to select the best subset of input variables, several manual, forward, backward, and stepwise model selection approaches have been developed. These techniques, however, received much criticism and are said to be quite unstable. Furthermore, multicollinearity is still a remaining issue.28,29 Penalized likelihood estimation such as LASSO $\ell_1$ (least absolute shrinkage and selection operator)29 and ridge $\ell_2$ (Tikhonov)30 regularization accomplish the same goal in an automatic, more stable, and computationally efficient manner. Additionally, they account for multicollinearity, and LASSO is able to perform variable selection.

Nonetheless, besides other excellent properties, LASSO has shown to be a highly biased estimate.31 Thus, to overcome bias issues, further penalized regression techniques have been developed. Among these are the smoothly clipped absolute deviation (SCAD) penalty31 and the minimax concave penalty (MCP).32 Both, SCAD and MCP initially apply the same penalization rate as LASSO but relax it towards zero with increasing coefficient values. Hence, they apply less shrinkage to the non-zero coefficients and, with this, reduce the bias. Still, LASSO, SCAD, and MCP have certain limitations in the presence of collinearity, as they assume the independence between penalty and correlation among predictors.

The ridge $\ell_2$ regularization often performs better in the presence of highly correlated inputs.33 But, since it does not perform variable selection, Zou and Addendum34 proposed the elastic net (Enet) method combining the $\ell_1$ and $\ell_2$ penalties. With a similar objective, a combination of the MCP and $\ell_2$ penalties called Mnet was proposed by Huang et al.34 Both have the ability of eliminating correlated covariates as a group. Thus, there are very sophisticated alternative parameter estimation techniques to MLE, such as Enet and Mnet, which help to avoid bias in the regression coefficients, reduce multicollinearity and overfitting, and guarantee an effective variable selection.

3.3 | Methodology and objectives

The objective of this paper is to extend existing research by developing models that directly consider the environmental conditions and are able to capture their combined effect on the failure behaviour of WT systems and their components. As the weather significantly changes throughout the year, the failures are modelled on a monthly basis. Previously used regression techniques and parameter estimation methods have not been found to entirely accomplish the objectives. Advanced alternatives need to be applied. Because of the very high numbers of (correlated) input covariates, the latter, however, might require PEVS techniques other than MLE.

To find the most suitable modelling technique, 4 regression models, namely, Poisson, NegBin, ZIP and ZINB in combination with 3 PEVS methods such as MLE, Enet, and Mnet, are compared. To the authors’ knowledge, these types of models and variable selection methods have not been applied in the field of WT failure modelling yet.

To cover the most relevant aspects, the following steps are performed and will be presented in this order in the results section:

1. **Evaluating the model performance with MLE**: WT system failures and gearbox failures are modelled via Poisson, NegBin, ZIP, and ZINB models in combination with MLE. The latter is a standard procedure for parameter estimation, does not delete input covariates, and serves to understand
the model performance when including all input variables. The gearbox was chosen to test the performance on component level, as it is a critical subsystem and one of the most expensive and highly maintenance intensive components causing the highest downtime.4,7

2. Effect of the model memory: The benefits of including the environmental covariates of the previous month are shown.

3. Performance of the PEVS methods: This is evaluated by combining each of the 4 regression models with either Enet or Mnet regularization and comparing the results with the ones of step 1. Figure 5 visualizes this process. Because of limited space in this paper, this is conducted only for the gearbox failure data. Wind turbine components have shown to react differently to certain weather conditions, and in some situations, establishing separate models for each component is more useful for operators than modelling the whole WT system without further distinguishing between the failed components.

4. Model performance on data of a single WF: Operators often need separate models for each WF to plan O&M actions properly. The combinations of environmental conditions can vary strongly at each site and affect the turbines differently. Thus, the capability of the proposed models for establishing separate WF models will be assessed in a final step.

4 | MATHEMATICAL FORMULATION OF THE MODELS

In Section 3, different models have been explained. Problems of current WT failure models including overfitting, overdispersion, excess zeros in the response variable, variable selection, heterogeneity, and multicollinearity have been discussed. Possible solutions to these problems were presented. Their mathematical formulations are introduced in this section. To account for the different numbers of turbines per WF, the failure counts are modelled with an offset of the number of turbines. This is essentially the same as modelling the failures per turbine over a certain time interval (ROCOF).

4.1 | Regression models

This study uses 4 regression models: Poisson, NegBin, ZIP, and ZINB, which are based on different probability distributions. The probability distribution for the Poisson model is given by

$$Pr(y_i|x_i) = \frac{\mu_i^y}{y_i!} e^{-\mu_i},$$

where $y$ is the response variable of nonnegative integer values. The mean and variance are $E(y_i|x_i) = \text{Var}(y_i|x_i) = \mu_i = \exp(x_i \beta_i)$, with the estimation coefficients $\beta_i$. The covariates $x_i$ for each observation $i$ include the variables described in section 2.2:

$$x_i = \begin{bmatrix} \text{age} \\ \text{MaxWS} \\ \vdots \\ \text{Rain.mem} \\ \text{RH.mem} \end{bmatrix}.$$ (2)

As previously discussed, the Poisson model can be extended to the NegBin model with the probability distribution:

$$Pr(y_i|x_i, \theta) = \frac{\Gamma(y_i + \theta)}{y_i! \Gamma(\theta)} \left( \frac{\theta \mu_i}{\theta + \mu_i} \right)^\theta \left( 1 - \frac{\mu_i}{\theta + \mu_i} \right)^{y_i}.$$ (3)
where \( \mu = \exp(x_i \beta) \) and \( \theta \) is a dispersion parameter adjusting the model regarding the degree of overdispersion. The NegBin model is a Poisson mixture model, with mean \( E(y_i|z_i) = \mu_i = \exp(x_i \beta + z_i) \) and variance \( \operatorname{Var}(y_i|z_i) = \mu_i + \mu_i^2 / \theta \), taking into account the unobserved effects \( z_i = \exp(\epsilon_i) \), which follow a gamma distribution.

As stated in Section 3, zero-inflated models use one process governed by a binomial distribution and a second one governed by a count distribution such as Poisson or NegBin. The probability distribution for the ZIP model is

\[
\Pr(y_i|x_i) = \begin{cases} 
\sigma_i + (1 - \sigma_i) e^{-\lambda_i} & \text{for } y_i = 0 \\
(1 - \sigma_i) e^{-\lambda_i} & \text{for } y_i \geq 1
\end{cases}
\]

with mean \( E(y_i|x_i, z_i) = \mu_i (1 - \sigma_i) \), variance \( \operatorname{Var}(y_i|x_i, z_i) = E(y_i|x_i, z_i)(1 + \mu_i \sigma_i) \), and \( \mu_i = \exp(x_i \beta) \). The term \( \sigma_i \) is the zero-inflation probability and is given by the logistic link function \( \sigma_i = \exp(y_i z_i) / (1 + \exp(y_i z_i)) \), with the zero-inflated regressors \( z_i \) and parameters \( \gamma_i \).

The probability distribution for the ZINB model is denoted by

\[
\Pr(y_i|x_i, \theta) = \begin{cases} 
\sigma_i + (1 - \sigma_i) \left(1 + \frac{\gamma_i}{\theta} \right)^{-\gamma_i} & \text{for } y_i = 0 \\
(1 - \sigma_i) \left(1 + \frac{\gamma_i}{\theta} \right)^{-\gamma_i} & \text{for } y_i \geq 1
\end{cases}
\]

with the logistic link function given above. The mean is \( E(y_i|x_i, z_i) = \mu_i (1 - \sigma_i) \) and the variance \( \operatorname{Var}(y_i|x_i, z_i) = E(y_i|x_i, z_i)(1 + \mu_i (\sigma_i + \theta^{-1})) \), with \( \chi_i \) as defined in Equation 2.

### 4.2 Parameter estimation and variable selection

The PEVS techniques MLE, Enet, and Mnet are used in this paper, as explained in the following.

Maximum likelihood estimation aims at finding the values for \( \beta \) that maximize the log-likelihood function \( L(\beta|x_i) = \ln(L(\beta|x_i)) \) with the covariates shown in Equation 2. As maximizing the log-likelihood is equivalent to minimizing the negative log-likelihood (loss function), the MLE estimators are given by

\[
\hat{\beta}_{\text{MLE}} = \arg\max_{\beta} \{ L(\beta|x_i) \} = \arg\min_{\beta} \{-L(\beta|x_i)\}.
\]

In penalized regression, instead of minimizing the loss function, the objective function \( M(\beta) \) is minimized:

\[
\hat{\beta}_{\text{pen}} = \arg\min_{\beta} \{ M(\beta) \} = \arg\min_{\beta} \{-L(\beta|X) + \lambda P(\beta)\},
\]

where \( P(\beta) \) is the penalty function and \( \lambda \) is a parameter that controls the trade-off between the penalty and the fit. Subtracting the penalty introduces sparsity and shrinks the coefficients. The penalty function for the regularization depends on the used technique. The Enet regularization penalty\(^{33}\) is given as

\[
\lambda P(\beta, \alpha) \propto \lambda \left( a ||\beta||_1 + \frac{1}{2} (1 - a) ||\beta||_2^2 \right) \quad \text{with} \quad \lambda \geq 0,
\]

where \( a \in [0, 1] \) controls the mixture between the \( \ell_1 \) and \( \ell_2 \) penalties. For \( a = 1 \), a pure LASSO regression is obtained, penalizing the sum of the absolute values of the model coefficients \( ||\beta||_1 = \sum_{k=1}^{p} |\beta_k| \), where \( p \) is the number of model covariates. For \( a = 0 \), the ridge penalty for model coefficients \( ||\beta||_2^2 = \sum_{k=1}^{p} \beta_k^2 \) is applied. In this study, an Enet with equal share of \( \ell_1 \) and \( \ell_2 \) penalty is used (\( a = 0.5 \)). The Enet estimates are defined as

\[
\hat{\beta}_{\text{Enet}} = \arg\min_{\beta} \{ M(\beta; \lambda, \alpha) \}.
\]

In contrast to Enet, which uses the \( \ell_1 \) penalty in the first term, Mnet uses the MCP, which is given by\(^ {32}\)

\[
\lambda P(\beta, \gamma) = \sum_{p=1}^{p} \rho(\beta_p; \lambda_1, \gamma) + \frac{1}{2} \lambda_2 ||\beta||_2^2 \quad \text{with} \quad \lambda = (\lambda_1 \geq 0, \lambda_2 \geq 0),
\]

with a regularization parameter \( \gamma \) and

\[
\rho(\beta; \lambda_1, \gamma) = \begin{cases} 
\lambda_1 \beta - \frac{\beta^2}{2} & \text{for } \beta \leq \gamma \lambda_1 \\
\frac{1}{2} \gamma^2 \lambda_1^2 & \text{for } \beta > \gamma \lambda_1
\end{cases}
\]

The Mnet estimator is defined as\(^ {34}\)

\[
\hat{\beta}_{\text{Mnet}}(\lambda, \gamma) = \arg\min_{\beta} \{ M(\beta; \lambda, \gamma) \}.
\]
For more detailed explanations of the 3 regularization techniques, reference is made to the respective literature. In all cases, a K-fold cross validation (with $K = 10$) is conducted before fitting the models in order to find the appropriate shrinkage parameter $\lambda$ that minimizes the mean squared error.

### 4.3 Model evaluation metrics

The proposed models will be fitted to the data and compared via the evaluation metrics: mean absolute error (MAE), root mean square error (RMSE), and the Akaike information criterion (AIC). The MAE and RMSE are given by

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2},$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_i - y_i|,$$

where $\hat{y}$ is the modelled and $y$ the original data value. Both metrics express the average model error reaching from 0 to $\infty$. However, the RMSE is more sensitive to occasional large errors than the MAE. Furthermore, the width of the confidence intervals is proportional to the RMSE.

Additionally, the models will be compared using the AIC, a measure of the relative quality of statistical models that have been applied to the exact same data set. It is defined as

$$\text{AIC} = 2k - 2\ln(\hat{\mathcal{L}}),$$

with the maximum value of the likelihood function $\hat{\mathcal{L}}$ and the number of estimated parameters $k$. For small data samples, the AIC has been refined with a correction factor as

$$\text{AICc} = \text{AIC} + \frac{2k(k + 1)}{n - k - 1},$$

with the sample size $n$. The AIC and AICc account for the complexity of the models and increase with a higher number of model parameters.

According to literature, the AIC and AICc have several advantages over the $F$ test and are supposed to find the most suitable model more reliably.

### 5 RESULTS AND DISCUSSIONS

In this section, the results of the evaluation steps explained in section 3.3 are presented. To facilitate the model interpretation, some comments will be made firstly.

#### 5.1 Comments on the model interpretation

Several things need to be considered when analysing the effect of each regressor on the dependent variable. The latter can only be investigated when simultaneously assuming the other covariates to be constant. This, however, is very unlikely considering the dynamics of meteorological parameters. One has to keep in mind that due to the naturally high correlation between many weather parameters, the covariates that are deleted by Enet and Mnet estimations still need to be considered in the interpretation. This can be done by looking at the correlations between the different variables in Figure 6. This graph shows the pairwise Pearson correlation values on the upper right triangle, the histograms and density functions of the data in the diagonal, and the data scatter plots with smooth lines (LOESS) on the lower left triangle. The following observations can be made:

- With increasing temperatures, the WS and MaxWS decrease. Thus, these variables are highly negatively correlated.
- Temperature and RH are also negatively correlated.
- WS and RH are positively correlated.
- PWR is highly positively correlated with wind speed and maximum wind speed; and negatively correlated to temperature.
- Precipitation is not significantly correlated with the other input parameters.
- As RH is defined as the percentage of moisture held in the air in relation to the possible maximum moisture content at a given temperature, lower RH can be caused by both, higher temperatures and/or less precipitation. However, as shown in Figure 6, there is nearly no significant correlation between RH and precipitation, whereas the temperature seems to be more important for the RH changes.

#### 5.2 Performance of the regression models

Table 2 shows the model evaluation metrics for the 4 modelling techniques with MLE applied to the WT system failures. Table 3 displays the results for the gearbox failures. In both cases, the ZINB models have a substantially better fit to the data indicated by the lowest values for all evaluation metrics. Furthermore, when comparing these results with those of other studies, the error metrics for both models are significantly lower.
**FIGURE 6** Correlation plots of the meteorological inputs. MaxWS, maximum wind speed; RH, relative humidity; WS, wind speed [Colour figure can be viewed at wileyonlinelibrary.com]

**TABLE 2** Evaluation metrics for the wind turbine failure models with MLE

<table>
<thead>
<tr>
<th>Measure</th>
<th>Poisson</th>
<th>NegBin</th>
<th>ZIP</th>
<th>ZINB</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>1144.73</td>
<td>1004.02</td>
<td>1056.88</td>
<td>991.45</td>
</tr>
<tr>
<td>AICc</td>
<td>1146.92</td>
<td>1006.21</td>
<td>1065.94</td>
<td>1000.51</td>
</tr>
<tr>
<td>MAE</td>
<td>1.226</td>
<td>1.206</td>
<td>1.242</td>
<td>1.137</td>
</tr>
<tr>
<td>RMSE</td>
<td>1.903</td>
<td>1.848</td>
<td>1.883</td>
<td>1.812</td>
</tr>
</tbody>
</table>

Abbreviations: AIC, Akaike information criterion; MAE, mean absolute error; MLE, maximum likelihood estimation; NegBin, negative binomial; RMSE, root mean square error; ZINB, zero-inflated negative binomial; ZIP, zero-inflated Poisson.

**TABLE 3** Evaluation metrics for the gearbox failure models with MLE

<table>
<thead>
<tr>
<th>Measure</th>
<th>Poisson</th>
<th>NegBin</th>
<th>ZIP</th>
<th>ZINB</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>557.18</td>
<td>532.34</td>
<td>542.74</td>
<td>523.52</td>
</tr>
<tr>
<td>AICc</td>
<td>559.37</td>
<td>534.53</td>
<td>551.80</td>
<td>532.58</td>
</tr>
<tr>
<td>MAE</td>
<td>0.6005</td>
<td>0.6126</td>
<td>0.6430</td>
<td>0.5733</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.7749</td>
<td>0.7827</td>
<td>0.7830</td>
<td>0.7572</td>
</tr>
</tbody>
</table>

Abbreviations: AIC, Akaike information criterion; MAE, mean absolute error; MLE, maximum likelihood estimation; NegBin, negative binomial; RMSE, root mean square error; ZINB, zero-inflated negative binomial; ZIP, zero-inflated Poisson.

Figures 7 and 8 display hanging rootograms of the models (with MLE) for WT and gearbox failures, respectively. Rootograms are useful methods to show the model fit to the data and are able to display issues such as overdispersion and problems with excess zeros. The distance between the bars and the reference line highlights the dissimilarity between expected and observed frequencies. Figures 7A and 8A indicate that the Poisson model has problems with high overdispersion and the zero counts. Figures 7C and 8C display the ZIP models, which still show some degree
of overdispersion, but a significantly better fit for the zeros. Figures 7b and 7d and 8b and 8d present almost identical results: NegBin and ZINB models handle overdispersion and excess zeros significantly better than Poisson and ZIP.

Consequently, the ZINB models show the best overall performance for both, the whole WT system and the gearbox failures. Comparing Tables 2 and 3, one can observe that the gearbox models have much lower errors. The WT system data contain any failed component during the observation period, and their failure behaviour is substantially harder to capture by models. Hence, as each component is affected differently by weather conditions, the components should be modelled separately. Therefore, in the following sections, only the results for the gearbox failure models will be displayed.
5.3 Model memory

As the ZINB model has shown the best results, it will be used to display the effects of including a model memory. For the present case, MLE will be used in order to emphasize the effects of the model memory without eliminating covariates, as Mnet and Enet would do.

It can be seen in Table 4 that including a model memory significantly decreases the MAE and RMSE about approximately 9%. Although the model with memory has more covariates, the AIC and AICc show lower values. This leads to the assumption that certain input covariates of the model memory are important and need to be included in the failure model. As introducing a memory results in a higher number of model covariates, previously discussed variable selection techniques should be considered for reducing model complexity by selecting the relevant covariates, as discussed in the following.

5.4 Variable selection method

In this section, the ZINB model in combination with the Enet and Mnet penalty will be applied to the gearbox failure data. The resulting evaluation metrics are presented in Table 5. It can be obtained that the zero-inflated models outperform the Poisson and NegBin models in both cases. ZINB again showed the best results. There is only approximately a 1% difference in MAE and RMSE for ZINB-Enet and ZINB-Mnet. However, the values for the AIC and AICc are in favour of the ZINB-Mnet model. Comparing Table 3 and Table 5, one can see that Mnet also outperforms MLE.

Tables 6 and 7 show the standardized model coefficients, standard errors, and 95% confidence intervals for ZINB-Mnet and ZINB-Enet models, respectively. As discussed in Section 4, is the model specific dispersion parameter.

The coefficients that were eliminated by the penalized regressions are not displayed, as they are not included in the resulting models.

The Enet model has 13 coefficients; the Mnet model has only 6, while having a very similar fit (RMSE and MAE). Additionally, the Mnet model shows lower AIC and AICc values.

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>Comparing ZINB models with and without memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
<td>MLE ZINB without memory</td>
</tr>
<tr>
<td>AIC</td>
<td>542.63</td>
</tr>
<tr>
<td>AICc</td>
<td>544.82</td>
</tr>
<tr>
<td>MAE</td>
<td>0.6702</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.7936</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 5</th>
<th>Evaluation metrics for the models with Enet and Mnet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
<td>Enet Poisson NegBin ZIP ZINB Mnet Poisson NegBin ZIP ZINB</td>
</tr>
<tr>
<td>AIC</td>
<td>554.19 530.85 512.36 508.32 552.43 527.39 503.27 494.45</td>
</tr>
<tr>
<td>AICc</td>
<td>555.34 532.18 513.88 510.80 553.76 528.72 504.42 495.43</td>
</tr>
<tr>
<td>MAE</td>
<td>0.6047 0.6103 0.6035 0.5468 0.6012 0.6090 0.5936 0.5503</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.7776 0.7812 0.7768 0.7394 0.7754 0.7804 0.7704 0.7419</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 6</th>
<th>Results of the estimation of the ZINB models with Mnet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Coefficient (β)</td>
</tr>
<tr>
<td>Intercept</td>
<td>−0.321</td>
</tr>
<tr>
<td>Rated Capacity</td>
<td>−0.327</td>
</tr>
<tr>
<td>Age</td>
<td>−0.430</td>
</tr>
<tr>
<td>Temp</td>
<td>−0.288</td>
</tr>
<tr>
<td>Rain</td>
<td>0.209</td>
</tr>
<tr>
<td>RH.mem</td>
<td>−0.929</td>
</tr>
<tr>
<td>Log(β)</td>
<td>1.139</td>
</tr>
</tbody>
</table>

Abbreviations: RH, relative humidity; ZINB, zero-inflated negative binomial.
TABLE 7  Results of the estimation of the ZINB models with Enet

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient (β)</th>
<th>Standard error</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−0.395</td>
<td>0.189</td>
<td>(−0.706 to −0.085)</td>
</tr>
<tr>
<td>Rated Capacity</td>
<td>−0.336</td>
<td>0.158</td>
<td>(−0.596 to −0.077)</td>
</tr>
<tr>
<td>Age</td>
<td>−0.314</td>
<td>0.173</td>
<td>(−0.599 to −0.029)</td>
</tr>
<tr>
<td>WS</td>
<td>0.176</td>
<td>0.234</td>
<td>(−0.208 to 0.560)</td>
</tr>
<tr>
<td>Temp</td>
<td>−0.512</td>
<td>0.255</td>
<td>(−0.932 to −0.092)</td>
</tr>
<tr>
<td>PWR</td>
<td>−0.112</td>
<td>0.250</td>
<td>(−0.523 to 0.300)</td>
</tr>
<tr>
<td>RH</td>
<td>−0.372</td>
<td>0.263</td>
<td>(−0.805 to 0.060)</td>
</tr>
<tr>
<td>Rain</td>
<td>0.286</td>
<td>0.120</td>
<td>(0.089 to 0.484)</td>
</tr>
<tr>
<td>WS.mem</td>
<td>−0.187</td>
<td>0.259</td>
<td>(−0.614 to 0.239)</td>
</tr>
<tr>
<td>RH.mem</td>
<td>−0.704</td>
<td>0.288</td>
<td>(−1.177 to −0.230)</td>
</tr>
<tr>
<td>Rain.mem</td>
<td>−0.097</td>
<td>0.138</td>
<td>(−0.323 to 0.130)</td>
</tr>
<tr>
<td>PWR.mem</td>
<td>0.214</td>
<td>0.233</td>
<td>(−0.170 to 0.598)</td>
</tr>
<tr>
<td>MaxWS.mem</td>
<td>−0.240</td>
<td>0.197</td>
<td>(−0.564 to 0.084)</td>
</tr>
<tr>
<td>Log(θ)</td>
<td>1.294</td>
<td>0.895</td>
<td>...</td>
</tr>
</tbody>
</table>

Abbreviations: MaxWS, maximum wind speed; RH, relative humidity; WS, wind speed; ZINB, zero-inflated negative binomial.

As most input variables are highly correlated (see Figure 6), Mnet successfully eliminated as many of the latter as possible in order to prevent collinearity. The Enet model tends to include most of them. Thus, the Mnet model is able to describe the data very well by only using very few of the input variables.

Furthermore, as the model with less complexity shall be preferred, the Mnet model results will be discussed further.

5.4.1  Interpreting the model covariates

By analysing the estimated coefficients of the ZINB-Mnet model given in Table 6, one can interpret the effect of each covariate on the model response. The following observations were made for non-environmental covariates:

• **Age**: The negative model coefficient for the covariate age suggests that the gearbox failures occur mainly in younger WT systems. This is consistent with literature, where premature gearbox failures are often considered a principal problem of WT systems. Although being designed for a lifetime of 20 years, literature states that gearboxes frequently suffer from major damages within the first 2 to 11 years and commonly fail at least once within the first 5 years of their lifetime. The most common cause of premature gearbox failures is problems with the gear bearings.

• **Rated capacity**: The negative coefficient for the model covariate rated capacity suggests that turbines with lower rated capacity suffer from more gearbox failures. This, however, could be also because the database itself includes more WT systems with lower rated capacity.

The coefficients of the meteorological covariates show the following behaviour:

• **Rain**: The precipitation covariate has a positive coefficient, indicating that with more precipitation, the gearboxes fail more frequently.

• **Temperature (Temp)**: The model states that more gearbox failures occur for lower monthly mean temperatures. This is consistent with earlier findings. Furthermore, during months with colder mean temperatures (winter), usually the mean wind speeds are higher. This causes increased wear on the gearbox.

• **Relative humidity (RH.mem)**: The negative coefficient for the model memory covariate RH.mem suggests that for lower mean relative humidities during the month prior to the gearbox failures, the latter occur more often.

As discussed in section 5.1, some of the model covariates are highly correlated. For this reason, combinations of the latter need to be interpreted while taking into consideration the failure modes of the component.

Failure mode

Wind turbine gearboxes can fail in a variety of different ways. However, according to the National Renewable Energy Laboratory gearbox failure database, more than 60% of all gearbox failures are directly related to the gear bearings. The principal reasons for WT gearbox bearing failures are oil degradation and contamination. Additionally, temperature-related changes in oil viscosity can affect the gearbox. Oil contamination is caused by moisture, particles, and entrained air (foam), which can result in high vibrations and wear. These contaminations can enter the gearboxes in a variety of ways. They could be introduced during manufacturing or maintenance or generated internally. Additionally, they can be ingested through air exchange with the ambient air. The latter occurs often during warmer months or because of diurnal temperature variations, which cause air to be sucked into the gearbox through the seals and "breathers." The typical breather systems in WT gearbox housings are usually not sufficiently...
preventing the contaminants from entering the system. The model suggests that the gearbox failures occur in the presence of these temperature variations, as discussed in the next 2 paragraphs.

The month before failure

The model indicates that the month before the failure is characterized by lower RH and thus higher temperatures, lower mean wind speeds, and less PWR. Although the RH might be lower at higher temperature, the increased air exchange contributes to a higher risk of contamination inside the gearbox. As the effect of these contaminations usually occurs time delayed, the component might only fail after a certain period or when the operational conditions change because of higher wind speeds and/or increased operational time.

The month of failure

The month of failure is defined by lower temperatures. This is consistent with earlier studies conducted by the authors. As the previous month was rather characterized by higher temperatures, this is likely to indicate a transition month from warmer to colder seasons. In particular, the daily temperature swings during these months can cause wear due to oil viscosity changes, which are resulting in less oil flow. Additionally, temperature variations due to heavy rain facilitate further air exchange between the ambient air and the interior of the gearboxes through the breathers. Along with lower temperatures, usually higher mean wind speeds are registered, causing more WT shutdown and start-up events and higher times in operation. Under these conditions, the gearboxes are mechanically challenged, and possible damages due to previously entered oil contaminations can lead to a component breakdown. So a combination of degraded and contaminated lubricant due to previous air exchange with the surroundings and problems with oil viscosity and higher loads during the failure month are affecting the gearbox lifetime behaviour negatively.

5.5 Application of the model to a single WF

In sections 5.2 to 5.4, it was shown that the ZINB models with Mnet penalty perform significantly better than other combinations of modelling techniques. The latter has lead to very satisfying results when applied to a large data set containing several different WFs and turbine technologies. As distinct technologies can react differently to combinations of environmental conditions, operators might prefer modelling each technology separately. To test the model performance on a data set containing only 1 technology, a single WF taken from the data set presented in Section 2 is tested. As WF-C contained information for the turbulence intensity and with an age of 8 years represents a WF that is well into its useful life without being very old, the latter was chosen for the analysis. The results are expected to show slightly different model coefficients than for the whole data set. Figure 9A,B displays the modelled versus the original monthly failures for ZINB-Mnet with and without the covariate TI. Furthermore, the constant failure rate that is usually assumed during the useful lifetime is displayed. Figure 10 displays the kernel density plots for the 2 set-ups. It is shown that the ZINB-Mnet model including TI performs best. Additionally, one can see that assuming a constant failure rate throughout the year leads to quite wrong assumptions and can delay the repair and maintenance processes for months.

As shown in Table 8, including TI substantially increases the model performance. Significantly lower AIC, MAE, and RMSE values were recorded when TI was included. As the model has 1 additional covariate, the difference between AIC and AICc for the model with TI is slightly higher. Nonetheless, its value is lower than for the model that does not include TI. When comparing the evaluation metrics with the results presented in Table 5 of section 5.4, much lower errors were recorded when modelling a WT technology separately. This was an expected result, as the failure behaviour of distinct technologies and different sites varies.

In section 5.4.1, the terrain class was eliminated by both variable selection algorithms for modelling the gearbox failure data of all WT technologies. However, as the terrain complexity affects wind conditions such as the turbulence intensity, which has lastly shown to play quite a big role when modelling gearbox failures, the terrain complexity is entering indirectly into the model.

![Figure 9](https://wileyonlinelibrary.com)  WF-C gearbox failure model with and without turbulence intensity (TI): A, failures per month; B, cumulative failures. ZINB, zero-inflated negative binomial [Colour figure can be viewed at wileyonlinelibrary.com]
Table 9 displays the results for the estimation process of the ZINB-Mnet model with TI. Six environmental parameters were selected by the Mnet algorithm. Again, the combination of low temperatures and high precipitation raises the failure frequency. Additionally, increasing TI and MaxWS during the month of failure leads to more failures. The Mnet algorithm did not eliminate the TI covariate, showing that it is an important parameter for modelling the gearbox failures in this WF.

As the data set for WF-C was rather small, the coefficients show wider confidence intervals than the results of section 5.4. Bigger databases (longer observation periods) would be needed to obtain lower confidence intervals, which is likely to slightly affect the selection of the model coefficients. Nonetheless, it has been shown that for modelling one WT technology separately, the ZINB-Mnet model also performs very well, having even lower errors than when modelling several turbine technologies at the same time.

6 CONCLUSIONS AND OUTLOOK

This paper presents a novel approach to model WT failures based on the meteorological conditions the turbines are exposed to. Several regression models have been tested on the data. Suitable PEVS techniques to reduce issues with high-dimensional regression problems have been implemented and evaluated. To the authors’ knowledge, this is the first work of this type in the context of modelling WT failures based on environmental conditions.

The use of ZINB models with Mnet penalty is proposed. The model takes into account unobserved effects (heterogeneity) with a gamma distribution. It handles the excess numbers of zeros by using a separate process that generates the structural zeros, which is governed by a
binomial distribution. The Mnet penalized regression has shown to be capable of selecting the most important input covariates very efficiently. Using the proposed model in combination with the penalized regression helps to prevent several problems such as overdispersion, overfitting, and multicollinearity. Including a model memory with the meteorological covariates of the previous month enhances the model performance significantly.

The proposed model has shown to perform well for modelling WT system failures of different turbine technologies without specifying the failed component. When considering the different components separately instead of modelling the whole system, the models performed even better, as the environmental conditions affect the components’ reliability differently. This has been shown using a subset of the data containing only gearbox failures, which is one of the most critical WT components. The model reveals that low temperatures, high maximum wind speeds, and precipitation affect the gearbox failure behaviour negatively. The findings related to wind speed are consistent with earlier studies.\textsuperscript{15,21} However, this study extends the latter by developing a more thorough model with further environmental variables and including a model memory.

Furthermore, the model performance was tested on the failure data of a single WF. Different WT technologies behave differently when exposed to certain weather conditions. Consequently, operators frequently use separate models for each turbine technology. Modelling the different technologies and their components separately has shown to lead to the best results and is recommended for future studies in this context. However, to analyse WFs separately, further work shall consider bigger databases. Model improvements could also take into account noise in the input variables by using, eg, error-in-variable models. Additionally, the authors will extend this study by modelling other WT components and including further operational SCADA data.

ACKNOWLEDGEMENTS

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

ORCID

Maik Reder \url{http://orcid.org/0000-0002-3830-9308}
Julio J. Melero \url{http://orcid.org/0000-0003-2360-0845}

REFERENCES


How to cite this article:  Reder M, Melero JJ. Modelling the effects of environmental conditions on wind turbine failures. Wind Energy. 2018;1–16.  https://doi.org/10.1002/we.2201