IOP Conf. Series: Journal of Physics: Conf. Series **1037** (2018) 062003 doi:10.1088/1742-6596/1037/6/062003

A Bayesian Approach for Predicting Wind Turbine Failures based on Meteorological Conditions

M Reder, J J Melero

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

E-mail: `mreder@fcirce.es, melero@unizar.es`

Abstract. With the growing wind energy sector, the need for advanced operation and maintenance (O&M) strategies has emerged. So far, mainly corrective or preventive O&M actions are applied. Predictive modelling, however, is expected to significantly enhance existing O&M practice. Here, anticipating wind turbine component failures can enable operators to lower the O&M cost and is particularly useful for wind farms located in remote areas or offshore locations. Previous research has shown that the failure behaviour of wind turbines and their components is highly influenced by the meteorological conditions under which the turbines operate. Hence, there is a significant need for robust models for failure prediction taking into consideration these conditions. Furthermore, solutions need to be found in order to determine the most suitable input variables for enhancing their prediction accuracy.

This study uses failure data obtained from 984 wind turbines during 87 operational WT years. Bayesian belief networks (BBN) are trained based on failure records, technology specific covariates, as well as measurements of the environmental and operational conditions at site. Subsequently, the failure events in a wind farm during a period of 36 months are predicted with the BNN, whereas the failure events of six main components are predicted separately. Furthermore, an extensive sensitivity study is carried out to find the model with the highest prediction accuracy for each component. The influence of each meteorological, operational or technical covariate are discussed in detail. The models achieved a very good accuracy and were able to predict the majority of the component failures over the prediction period.

1. Introduction and Problem Statement

Over the past years wind farm operation and maintenance (O&M) has become an emerging field of research. In the wind industry, a very large share of the levelised cost of energy is directly related to O&M activities. Operators need sophisticated models to anticipate wind turbine (WT) failures and to estimate future maintenance actions.

Here, one core element are failure and reliability models of wind turbine systems and components. These are used by operators to indicate the probability of having a WT failure in the near future and to adjust the maintenance strategies accordingly. Current maintenance strategies are almost exclusively time-based. In the context of failure or reliability models, this implies that the system age is the only driver for the reliability degradation. One of the most widely used reliability models is based on a Weibull distribution and results in the famous bathtub curve. The latter assumes a constant failure rate throughout a very big part of the WTs' life time, the so called 'useful life'. However, these models were designed for machinery working under mostly constant operating conditions and not exposed to changing weather conditions, as wind turbines are. Thus, the assumption of having a constant failure rate throughout the useful life time of wind turbines and their components, does not hold true. It has been successfully shown in several previous studies, see e.g. [1–8], that the failure behaviour of wind turbines and their components is highly related to the meteorological conditions to which the WTs are exposed to. Hence, advanced WT reliability models should include these meteorological and operational conditions in order to account for the variable surroundings, [9]. This significantly enhances the failure prediction and can eventually result in much lower O&M costs.

This paper uses a Bayesian approach to estimate the conditional probabilities of having one or more wind turbine failures in the presence of complex combinations of several weather variables. The latter include wind speed (WS), wind gusts (MaxWS), precipitation (Rain), relative humidity (RH) and ambient temperature (Temp). For this, Bayesian networks are trained using information on monthly wind turbine failures and the related weather conditions, as well as turbine technology specific attributes. The trained models are then employed to predict the failure events of specific components in a wind farm during a prediction period of 36 months. Besides the failure events related to the whole turbine system, also component failures of the blades, gearbox, generator, main bearing, pitch and yaw system are modelled. In total, the data used in this study are comprised of records obtained during 87 operational WT years. An extensive sensitivity analysis is carried out in order to determine the most suitable input variables for each model of six main WT components.

To the authors' knowledge, the combination of predicting binary outcomes such as the occurrence of one or more failures in a wind farm based on multivariate probabilistic models, has not been subject to previous studies in the field of wind energy, yet. The results of this study are expected to significantly contribute to research in O&M and enable operators to predict and understand the conditional probabilities of having a failure under given meteorological conditions.

2. Data

In this study failure data of 29 different wind farms obtained during 1045 operational months are used. The wind farms operate in total 984 wind turbines, of which 638 are stall and 349 are pitch regulated. All turbines are three bladed and geared drive machines aged between one and 16 years with rated capacities of 300 kW to 2000 kW. Figure 1 summarises the meteorological and turbine specific model covariates.

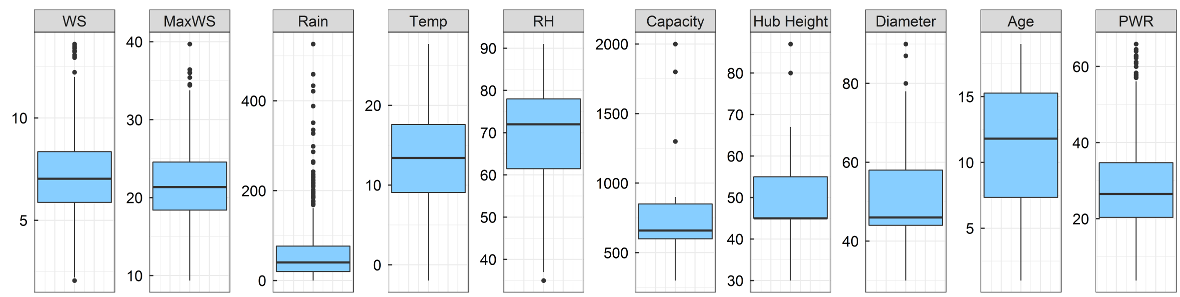


Figure 1: Summary of the environmental and turbine specific input data: WS (m/s), MaxWS (m/s), Rain (mm); Temp (°C), RH (%), Rated Capacity (kW), Hub Height (m), Diameter (m), Age (years), PWR(%)

.

The monthly mean measurements for the meteorological variables relative humidity (RH) and ambient temperature (Temp), as well as the total monthly precipitation (Rain) were taken from close-by located weather stations or if available from the wind farms met mast. The monthly mean wind speed (WS) and maximum wind gust (MaxWS) measurements were taken directly from the wind turbines' SCADA systems. Furthermore, the measured monthly mean active power production in percent of the rated capacity (PWR) taken from each turbine's SCADA system, is included into the model in order to account for the time in operation per month.

The failures of six main WT components are analysed. These have been defined in [10] as the most critical components in terms of failure rate and downtime. Furthermore, the failures of the whole wind turbine system without distinguishing between the failed component will be analysed. This will be referred to as 'all Failures' in the further. Table 1 displays the number of failure events per component registered over the whole observation period.

Table 1: Number of failure events per component in the database used for this study.

Component	Blades	Gearbox	Generator	Main Bearing	Pitch	Yaw	**Total**
Failure events	232	230	104	14	6	18	**609**

In the context of this paper only failures that resulted in a wind turbine stop, which required manual interaction such as repair or replacement, are considered. Thus, (bi-)annual services and inspections are excluded.

3. Methodology

In this study a naive Bayesian classifier, [11, 12], will be used, which is a special form of a Bayesian Belief Network (BBN). It is particularly suited for high dimensional input data and applies Bayes' theorem with a strong independence assumption among the features. Despite the fact that in reality this assumption is rarely true, it has shown to perform surprisingly well in real world classification problems. It allows to simplify the multidimensional classification task by breaking it down into several one-dimensional ones, [11]. The conditional probability using Bayes' theorem is defined as:

$$p(c_j|\mathbf{x}) = \frac{p(c_j)p(\mathbf{x}|c_j)}{p(\mathbf{x})} \quad , \tag{1}$$

where $\mathbf{x}$ is a vector of variables for which the posteriori probability $p(c_j|\mathbf{x})$ that $\mathbf{x}$ belongs to class c_j is determined. The class prior probability is given by $p(c_j)$, the predictor prior probability by $p(\mathbf{x})$ and the the likelihood $p(\mathbf{x}|c_j)$. Under the 'naive' conditional independence assumption the naive Bayes classifier model results in:

$$\hat{y} = \underset{j}{\operatorname{argmax}} \left\{ p(c_j) \prod_{i=1}^{n} p(x_i|c_j) \right\} . \tag{2}$$

In this paper the boolean model response (class) is defined as the event of having one or more failures in a wind farm during one month. The model considers in total 17 input covariates, including the environmental parameters presented in Section 2. As the meteorological conditions often have a delayed or even accumulative impact on the failure behaviour, the meteorological measurements taken throughout the previous month are also taken into consideration and are indicated with the suffix '*.b*', which stands for '*.before*'. Furthermore, it is supposed that distinct WT technologies are affected differently by certain combinations of environmental parameters. Thus, in order to distinguish between the WT technologies, additionally, the turbine age, rotor diameter, hub height, (pitch and stall) regulation and rated capacity are added as model covariates. The BBN with all input variables is visualised in Figure 2. This is an example of how the covariates could be connected, however, the interconnections may vary for each component model.

The values for each climatological input variable are grouped into four equally-populated categories via quantile discretisation. The data set is split randomly into training and testing data, so that a testing period of 36 months is obtained, for which the predictions will be carried out. During training, the model derives the conditional probabilities of having a failure event in the presence of the respective categories of each covariate. Subsequently, the trained naive Bayes classifier is used to predict failure occurrences during 36 months of operation.

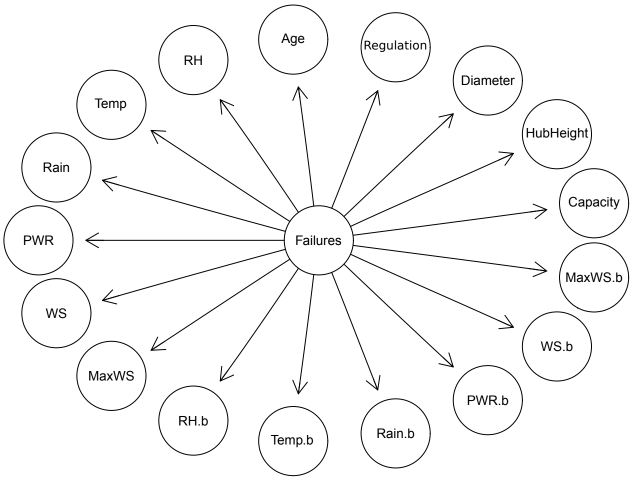


Figure 2: Example of a Bayesian Belief Network including all model covariates of this study.

3.1. Sensitivity Analysis

The distinct combinations of environmental conditions are not influencing the failure behaviour of all WT components equally, [6]. Hence, for each component a sensitivity study is carried out in order to find the most suitable model, compromising between prediction accuracy and model complexity. This, furthermore, helps to understand which combination of covariates is the most effective for predicting the specific component failure.

For this, separate models using all possible combinations of input covariates are trained and predictions are carried out on the test data set. Each combination can involve between 1 and 17 covariates (members). Thus, given 17 types of different variables in the data set, and $n = \{1...17\}$ members possibly involved in each combination, a total of 131071 models are trained for each component. In order to evaluate the performance of the different classification models, each one of them is used to predict the classes of the response variable in the test data set. Then, the confusion matrices of all predictions are compared. These contain information on the true positive (TP), true negative (TN) as well as the false positive (FP) and false negative (FN) predictions. The best model will be chosen based on four quality measures of the binary prediction. The sensitivity (true positive rate TPR) and specificity (true negative rate TNR) as well as the prediction accuracy (ACC) are stated as:

$$TPR = \frac{TP}{TP + FN}; \qquad TNR = \frac{TN}{TN + FP}; \qquad ACC = \frac{TP + TN}{TP + TN + FP + FN}. \tag{3}$$

Further, the Mathews correlation coefficient (MCC) is given by:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP(TN + FN))}}. \tag{4}$$

IOP Conf. Series: Journal of Physics: Conf. Series **1037** (2018) 062003 doi:10.1088/1742-6596/1037/6/062003

The TPR, TNR and ACC, take values on an interval $\{0, 1\}$, with 1 being the best case. MCC reaches from -1 to 1, where -1 represents a complete disagreement with the observed data, 0 indicates a random and 1 a perfect prediction. This measure is often considered a slightly more robust metric than the TPR, TNR and ACC, which are often found to introduce a certain bias to the model evaluation process.

4. Results and Discussion

In this section, firstly, the results of the sensitivity study are analysed in order to find the minimal number of members in the combination of covariates, which results in the best model for each WT component. Then, the component models will be explained in detail including the conditional probabilities of the input variables. This helps to understand which meteorological variables affected most the respective component model.

4.1. Sensitivity analysis

Figure 3 shows the sensitivity versus specificity of the predictions for each possible combination of input covariates for the failure model of the whole WT system ('all failures'). The graph is sorted by the number of possible members in each combination n. It can be seen that the best compromises between specificity and sensitivity are obtained for mid-range n. The model using all input variable did not show the best performance out of all models. Nonetheless, with these graphs it remains a difficult task to determine the best performing model.

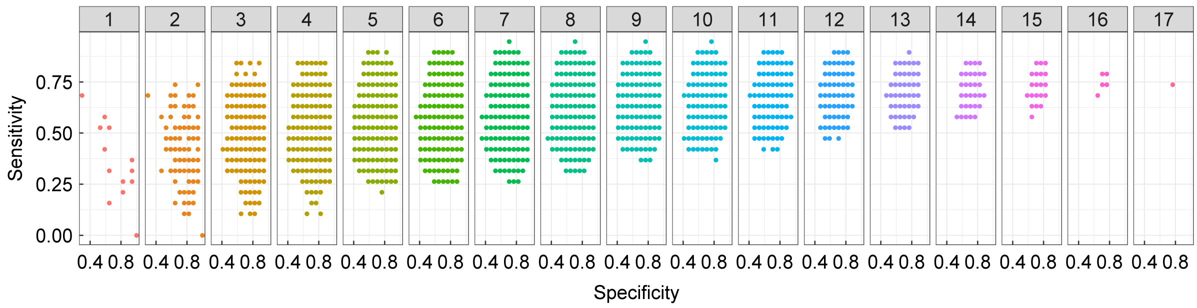
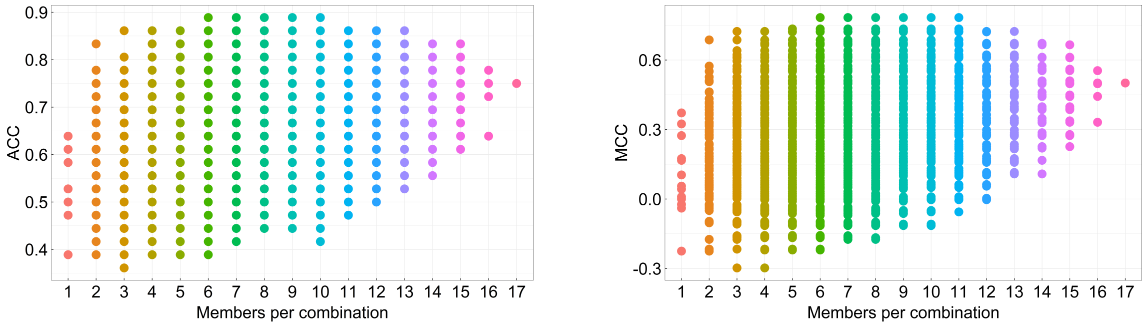


Figure 3: Sensitivity versus specificity for failures of the whole wind turbine system.

This becomes clearer, when looking at Figure 4, which shows the ACC and MCC for each n. One can see that both, ACC and MCC, show their maximum values for n between 6 and 11 members.

In Figure 5 the MCC of the predictions for all components and all numbers of members per combination are displayed. It can be seen that the model for all WT failures obtains its maximum value for MCC with six model covariates. The model for the blades needs a minimum of $n = 9$ members in a combination of covariates, while the gearbox needs $n = 6$, the generator $n = 5$ and the main bearing $n = 3$ members. The models for yaw and pitch system did not perform well in the predictions. Both models were no better than a random model. This, could be due to the limited amount of failure events in the used data set for these components. Furthermore, this could also imply that the failure behaviour of these components is influenced more by other variables, which were not included into the models. The models for the remaining components, however, showed very high values for the MCC, which confirms that these models serve as very good predictors.

As displayed in Figure 5, many components showed a wide span of possible n, for which the highest MCC values were obtained. It shall be mentioned that, as we strive for the best compromise between model complexity and prediction accuracy, only the combinations



(a) (b)

Figure 4: Metrics: (a) ACC and (b) MCC for failures of the whole wind turbine system.

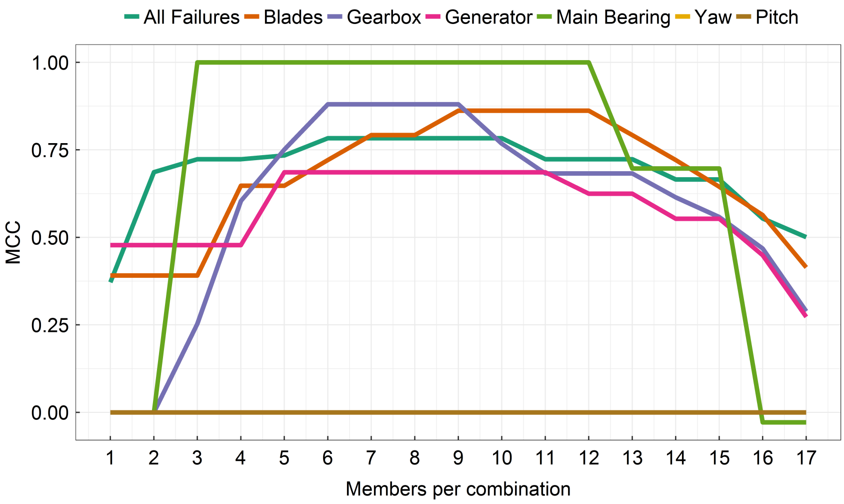


Figure 5: MMC of the predictions made for all components and all possible numbers of members per combination.

of covariates with the minimum number of members for which the highest MCC value was obtained, will be considered in the further. However, more covariates could possibly be added to the respective model, while having the same prediction performance.

Table 2: Results of the predictions for all component models with n_{best}.

Component	n_{best}	MCC	ACC	Sensitivity	Specificity	TP	FP	TN	FN
All Failures	6	0.783	0.889	0.842	0.941	16	1	16	3
Blades	9	0.862	0.944	0.800	1.000	8	0	26	2
Main Bearing	3	1.000	1.000	1.000	1.000	1	0	35	0
Gearbox	6	0.880	0.972	0.800	1.000	4	0	31	1
Generator	5	0.686	0.944	0.500	1.000	2	0	32	2
Pitch System	-	0	0.973	0	1.000	0	0	36	1
Yaw System	-	0	0.973	0	1.000	0	0	36	1

Table 2 summarises the evaluation metrics of the predictions for all component models with the number of members per combination n_{best}, which results in the best model. It is remarkable, that only one false positive was obtained for all predictions. In general the models for the blades, gearbox, generator, main bearing and the whole WT system performed very well, having high sensitivities and specificities, as well as good detection rates. As expected, the number of failure events in the available data set for each component influences the accuracy of the predictions, as the models need a certain number of positive events during the training in order to learn the structure of the data. The models for the pitch and yaw system will be excluded from further discussion, as they did not perform well. Also, the main bearing failure records only contained one event during the testing period. This was successfully predicted, however, the model covariates discussed in the next section might be biased with this restriction. As shown in Table 3, in some cases there was little to no difference between the best and the second-best performing component model obtained for n_{best}. The models for the main bearing and gearbox, showed the same performance for several combinations of input covariates. Due to the limited space in this paper, however, only one combination for each component can be analysed in detail.

Table 3: Results of the predictions for all component models with n_{best}.

Component	n_{best}	**MCC - best model**	**MCC - 2^{nd} best model**	**Comment**
All Failures	6	0.783	0.734	–
Blades	9	0.862	0.792	–
Main Bearing	3	1.000	1.000	11 models showed the same MCC
Gearbox	6	0.880	0.880	7 models showed the same MCC
Generator	5	0.686	0.533	–

4.2. Interpretation of the Resulting Models

This section discusses the predictions made with the model that was determined as the best performing in the sensitivity study for the respective component, being displayed in Table 2. Figure 6 shows the observed and predicted failure events for each component during the prediction period of 36 months. Generally, a very good accuracy is obtained and most component failures were found.

Figure 6: Monthly observed ans predicted failures (boolean).

Table 4 gives the levels to which the covariates were assigned to during the quantile discretisation. The conditional probabilities for having one or more failures of a certain component within the whole wind farm under the given values of the meteorological variable are shown in Figures 7 to 9 as well as Table 5.

Table 4: Assigned levels for the model covariates by quantile discretisation.

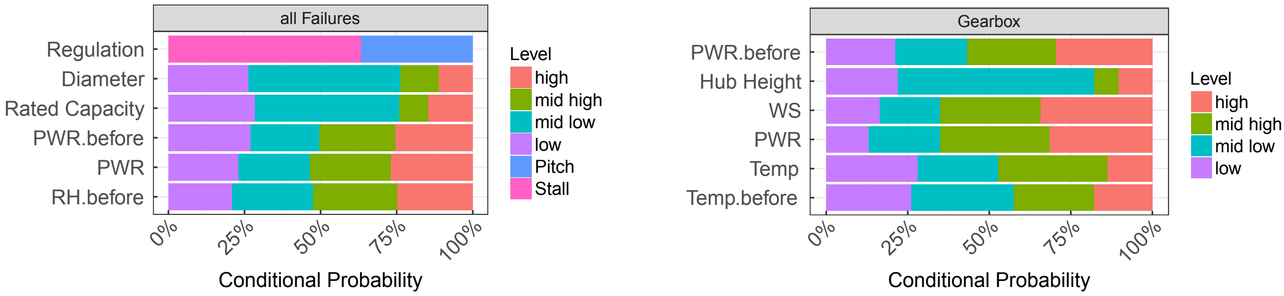
Level	**WS**	**MaxWS**	**Rain**	**Temp**	**RH**	**Capacity**	**Hub Height**	**Diameter**	**Age**	**PWR**
SI-unit	m/s	m/s	mm	°C	%	kW	m	m	years	%
low	2.1-5.7	9.4-18.4	0-20.7	(-2) - 8.6	35-59.9	300-600	30-35	30-42	1-6.9	3.8-19.2
mid low	5.8-6.8	18.5-21.6	20.8-35.5	8.7-13.6	60-69.9	660 -850	43-55	44-58	7-11.6	19.3-25.7
mid high	6.9-8.1	21.6-25.3	35.6-70.3	13.7-17.9	70-76.9	900-1300	57-67	59-78	11.7-15.1	25.8-33.6
high	8.1-13.6	25.3-40.5	70.4-526	18.27.7	77-91	1800-2000	80-87	80-90	15.2-19	33.7-64.5

The figures display how each covariate level contributed to the overall conditional probability of the respective covariate. As we use the minimum n for which the highest values of MCC were obtained, these figures do not include all the model covariates that could possibly affect the failure behaviour.

In the following each covariate will be analysed separately for each component model. However, the reader should keep in mind that the failure events occur under the condition of a combination of all of these covariates at specific levels.

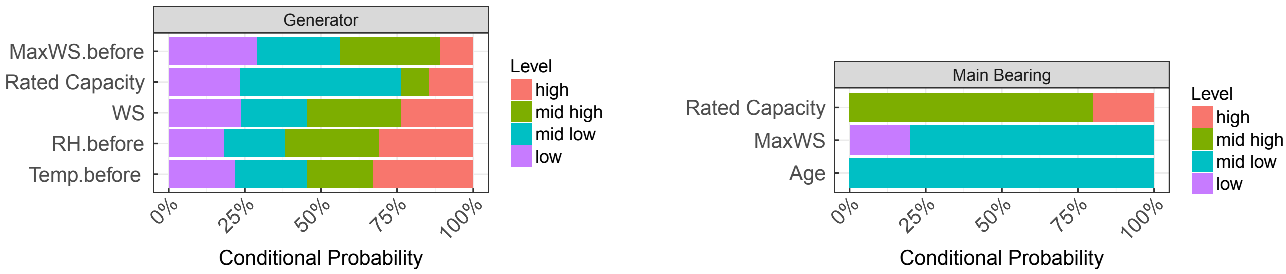
- **All failures:** Figure 7a shows that the conditional probabilities for having a WT failure (without distinguishing between the failed components), are much higher for stall regulated turbines than they are for pitch regulated ones. Furthermore, WTs with lower rated capacities and smaller diameters are affected more often. Slightly elevated power production and relative humidity also contributed to higher probabilities of failure.
- **Gearbox:** Figure 7b shows that colder temperatures throughout the month before the failure as well as during the failure month, higher power production and higher mean wind speeds increase the probability of having a failure of this component. Furthermore, lower hub-heights showed to be affected more often.
- **Generator:** The conditional probabilities of the generator model indicated that the failures are likely to occur under higher mean wind speeds and high relative humidity. Turbines of lower rated capacity were affected more often. During the month before the failure slightly higher temperatures and lower maximum wind speeds resulted in a higher probability of failure. In literature, e.g. [13], it is stated that generator failures often occur under highly variable environmental conditions, as for instance during transition periods from summer to winter. This might explain the difference of temperature and wind speed indicated by the model.
- **Main Bearing:** The compromise of best performing model and prediction accuracy for the main bearing failures only considered 3 covariates. According to these, younger turbines with higher rated capacity operating under wind conditions with low gust speeds, showed higher failure probabilities.
- **Blades:** High precipitation, high relative humidity throughout the month before failure, slightly higher wind speeds during the failure month as well as the month prior to the failure resulted in higher probabilities of blade failures. Furthermore, marginally older turbines were failing more often.

IOP Conf. Series: Journal of Physics: Conf. Series **1037** (2018) 062003 doi:10.1088/1742-6596/1037/6/062003



(a) (b)

Figure 7: Model covariates and conditional probabilities for (a) all failures, (b) gearbox.



(a) (b)

Figure 8: Model covariates and conditional probabilities for (a) generator, (b) main bearing failures.

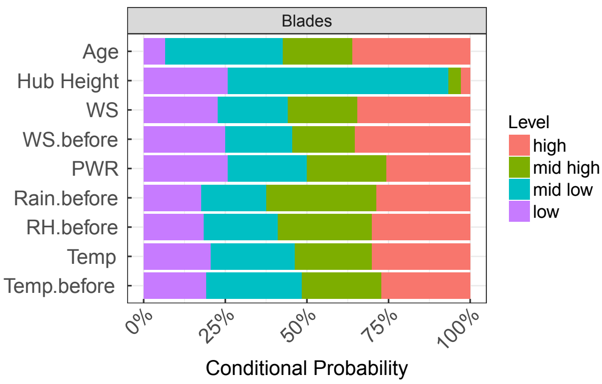


Figure 9: Model covariates and conditional probabilities for blade failures.

5. Conclusions

The failure events of six main WT components were predicted with a Bayesian belief network based on meteorological and operational conditions as well as WT technology specific covariates. It was shown that the herein presented techniques are capable of reliably predicting failure events during 36 operational WT months. An extensive sensitivity analysis was carried out, in order to find the best performing model for each of the six main components, compromising between model complexity and prediction accuracy. It could be seen that each component model is driven by different input covariates. Furthermore, even with a certain degree of multicollinearity among the environmental input variables, as well as the independence assumption of the naive Bayes classifier, the BBN performed very well.

In future work, the influence of different data pre-processing for the BBN should be investigated, as the form of input discretisation can affect the model outcome. Additionally, other meteorological variables should be included. Especially, the turbulence intensity, as well as wake effects are expected to have a significant impact on the failure behaviour of certain components. Also, data should be used that contain more failure information for the main bearing, pitch and yaw system. These were represented with a low number of failures in the data set of this study, and thus, it was not possible to generate reliable prediction models for these components. Furthermore, the model can be easily extended by including WT technologies (e.g. direct drive WTs) that were not considered in this study. Finally, future studies could focus on different failure modes of each component, as it is supposed that each failure mode is provoked by different combinations of meteorological and operational conditions.

Table 5: Conditional Probabilities for all models and input variables

		low	mid low	mid high	high
all Failures	Diameter	0.263	0.500	0.127	0.111
	Rated Capacity	0.285	0.475	0.095	0.146
	PWR	0.231	0.234	0.266	0.269
	PWR.before	0.269	0.228	0.250	0.253
	RH.before	0.209	0.266	0.278	0.247
	Regulation	Stall = 0.633		Pitch = 0.367	

		low	mid low	mid high	high
Gearbox	Hub Height	0.219	0.603	0.075	0.103
	WS	0.164	0.185	0.308	0.342
	PWR	0.130	0.219	0.336	0.315
	PWR.before	0.212	0.219	0.274	0.295
	Temp	0.281	0.247	0.336	0.137
	Temp.before	0.260	0.315	0.247	0.178

		low	mid low	mid high	high
Generator	WS	0.236	0.218	0.309	0.236
	MaxWS.before	0.291	0.273	0.327	0.109
	RH.before	0.182	0.200	0.309	0.309
	Temp.before	0.218	0.236	0.218	0.327
	Rated Capacity	0.236	0.527	0.091	0.145

		low	mid low	mid high	high
MB	Age	0.000	1.000	0.000	0.000
	MaxWS	0.000	1.000	0.000	0.000
	Rated Capacity	0.236	0.527	0.091	0.145

		low	mid low	mid high	high
Blades	Age	0.066	0.360	0.213	0.360
	WS	0.228	0.213	0.213	0.346
	WS.before	0.250	0.26	0.191	0.353
	PWR	0.257	0.243	0.243	0.257
	Rain.before	0.176	0.199	0.338	0.287
	RH.before	0.184	0.228	0.287	0.301
	Temp	0.206	0.257	0.235	0.301
	Temp.before	0.191	0.294	0.243	0.272
	Hub Height	0.257	0.676	0.037	0.029

References

[1] Faulstich S, Lyding P and Tavner P 2011 *European Wind Energy Conference*
[2] Wilkinson M, Van Delft T and Harman K 2012 *European Wind Energy Conference*
[3] 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
[4] Wilson G and McMillan D 2014 *Journal of Physics: Conference Series* **524** 1–10
[5] Tavner P, Greenwood D M, Whittle M W G, Gindele R, Faulstich S and Hahn B 2013 *Wind Energy* **16** 175–187
[6] Reder M and Melero J J 2017 *Journal of Physics: Conference Series* **926** 012012
[7] Reder M and Melero J J 2017 Time series data mining for analysing the effects of wind speed on wind turbine reliability *Safety and Reliability Theory and Applications* (CRC Press Taylor & Francis Group: CRC Press) p 109
[8] Reder M, Yürüsen N Y and Melero J J 2018 *Reliability Engineering & System Safety* **169** 554–569
[9] Kuik G A M V, Peinke J, Nijssen R, Lekou D, Mann J, Ferreira C, Wingerden J W V, Schlipf D, Gebraad P, Polinder H, Abrahamsen A, Bussel G J W V, Tavner P, Bottasso C L, Muskulus M, Matha D, Lindeboom H J, Degraer S, Kramer O, Lehnhoff S, Sonnenschein M, Morthorst P E and Skytte K 2016 *Wind Energy Science* **1** 1–39
[10] Reder M D, Gonzalez E and Melero J J 2016 *Journal of Physics: Conference Series* **753** 072027
[11] Hand D J and Yu K 2001 *International Statistical Review* **69** 385–398
[12] Zhang H 2004 *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference FLAIRS 2004* **1** 1 – 6
[13] T Wei 2010 *Wind Power Generation and Wind Turbine Design* (Southampton, Boston: WIT Press)