SafeWind


Collaborative project funded by the European Commission under the 7th Framework Program, Theme 2007-2.3.2: Energy

"Multi-scale data assimilation, advanced wind modelling & forecasting with emphasis to extreme weather situations for a safe large-scale wind power integration"

Grant Agreement N°: 213740

Deliverable Dp-6.5

"Statistical forecasting methods with focus on cut-off event"

DOCUMENT TYPE	Deliverable
DOCUMENT NAME:	swind.deliverable_Dp-6.5_v1
VERSION:	V1.0
DATE:	12.12.2012
CLASSIFICATION:	R0: Genera Public
STATUS:	S0 : Approved

Abstract: This Deliverable of SafeWind project presents some approaches for the prediction of cut-off events. Prediction of these events is important to wind farm owners, since it has a significant impact on power production. A sudden strong wind may shut down and entire wind farm, meaning that the wind farm will go from full to no production in a short period of time. This prediction is, however, very complex, since those events, altought important, are rarely observed.

AUTHORS[1], REVIEWERS				
MAIN AUTHOR/EDITOR:	I. Sánchez			
AFFILIATION:	Universidad Carlos III de Madrid (UC3M)			
ADDRESS:	Avda. Universidad, 30, 28911, Leganés, Madrid, Spain			
TEL.:	+34 91.624.9179			
EMAIL:	Ismael.sanchez@uc3m.es			
FURTHER AUTHORS:	Stig Mortensen (ENFOR), Philipp Brandt (Overspeed)			
PEER REVIEWERS:	Ismael Sánchez (UC3M), Stig Mortensen (ENFOR), Philipp Brandt (Overspeed)			
REVIEW APPROVAL:	Approved :		Rejected (improve as indicated below) :	
SUGGESTED IMPROVEMENTS:				
APPROVER:				

VERSION HISTORY			
VERSION[2]:	DATE:	COMMENTS, CHANGES, STATUS:	PERSON(S):
1.0	25.08.2012	Combination of reports from UC3M and ENFOR	I. Sánchez
2.0	19.10.2012	Combination of reports from Overspeed	I. Sánchez
3.0	1.12.2012	Updated with final revision of contributors	I. Sánchez

STATUS, CONFIDENTIALITY, ACCESSIBILITY									
STATUS:				CONFIDENTIALITY:				ACCESSIBILITY:	
S0	Approved/Released	**X**		**R0**	General public	**X**		Private web site	
S1	Reviewed			**R1**	Restricted to project members			Public web site	**X**
S2	Pending for review			**R2**	Restricted to European Commission			Paper copy	
S3	Draft for comments			**R3**	Restricted to WP members + PL				
S4	Under preparation			**R4**	Restricted to Task members +WPL+PL				

PL: *Project leader* **WPL:** *Work package leader* **TL:** *Task leader*

[1] **The authors of this document are solely responsible for its content, which does not represent the opinion of the European Community and the European Community is not responsible for any use that might be made of data appearing therein.**

[2] **VERSION NAMING : V0.x** draft before peer-review approval, **V1.0** at the approval, **V1.x** minor revisions, **V2.0** major revision

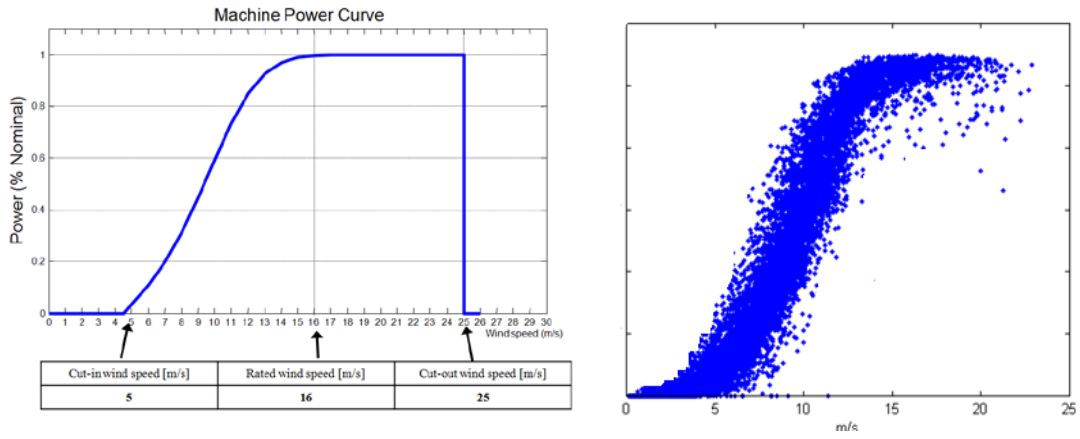
Contents

1. Statistical Forecasting of Unobserved Cut-off Events in Wind Power Generation, by UC3M

1.1 Introduction

The power curve of a wind turbine represents the relationship between the electrical power output of the turbine and the wind speed. Its main wind speed values are:

- Cut-in speed: Minimum wind speed at which the wind turbine will generate power.
- Rated wind speed: Minimum wind speed at which the wind turbine will generate its designated rated power.
- Cut-off wind speed: Wind speed at which shut down occurs. Having a cut-out speed is a safety feature which protects the wind turbine from damage. Beyond this cut-off wind speed the rotor does not turn. The operation is resumed when the wind speed decreases and passes a wind speed that is slightly smaller than the cut-off wind speed. The magnitudes of these wind speeds and their difference will in general depend on the wind turbine in question. Kristensen et al (2004) showed that in order to evaluate the loss of power due to cut-off events, this difference can be ignored. Therefore, we will only consider a single value of wind speed to characterize cut-off events.



(a) (b)

Figure 1: Example of power curves. (a) Machine (theoretical) power curve and main speed values. (b) A real power curve.

Figure 1(a) shows the so-called machine power curve, where the relationship between wind speed and output power is deterministic. This curve is usually computed from controlled experiments and under some standardized conditions (Hunter el al (2001)). Conversely, real power curves are stochastic, as seen in Figure 1(b). In a real power curve it is more difficult to define the above mentioned main wind speed values. Another important aspect in real power curves is that wind speeds close or above cut-off are rarely observed.

The goal of this work is to propose a statistical methodology for the prediction of cut-off events due to extreme wind speed in wind power forecasting using observed data. This prediction is important both at a wind farm level and at a regional level. For a wind farm, a cut-off event may translate to the rapid loss of a very significant share of the power supply, since we will pass from a production at rated power to no production at all. If such an extreme weather condition is observed in a large region, the fast loss of power can be dangerous for the reliability of the system.

The prediction of a cut-off event is not easy, because these events are rarely observed. Consequently, it is not easy to solve it with an empirical model. If we would have enough registered cut-off situation, we could incorporate that information into a model. For instance, a non-linear model could be fitted using non-parametric

techniques. Also, a binary response model could be built. In these models, the response is a dummy variable where the cut-off situation is labeled as 1 and the remaining situations are labeled as 0. Then some logistic regression (parametric or non-parametric) could be estimated using some explanatory variables, like predicted wind speed as input (see, for instance Hosmer and Lemeshow (2000)).

In this work we will pursue a very different approach, which we consider more realistic. We will consider that all the observed data is below cut-off. That is, our goal is to build a methodology to predict the unobserved event of a cut-off.

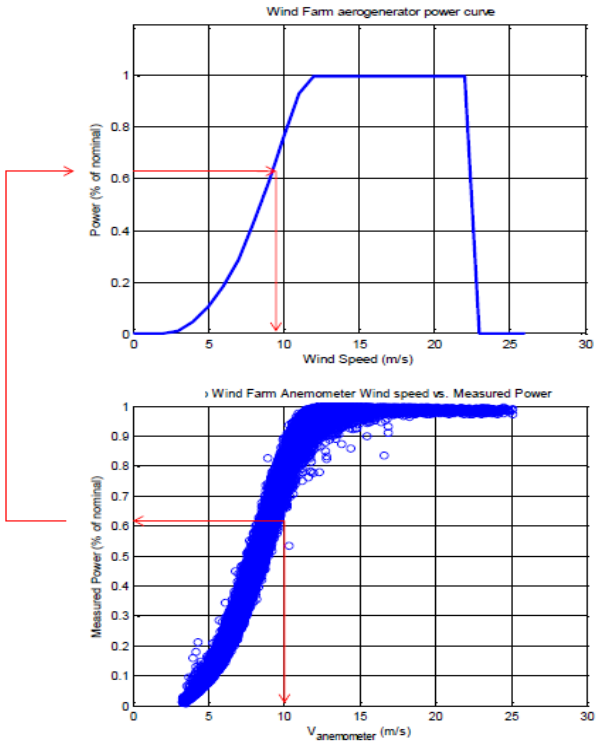


Figure 2: Alternative wind speeds for a given measured power.

The procedure can be summarized as follows. For a given measured power P_t we can collect two alternative wind speeds (see Figure 2): the "machine wind speed" that corresponds to the ideal machine power curve, hereinafter referred to as M_t, and the registered wind speed, denoted as R_t. Depending on the application, the registered wind speed can be the one measured by an anemometer or the predicted by a NWP method. The dataset will then consist on a set of vectors (P_t, M_t, R_t) $t=1,...,T$.

Let us denote as V_{cut} the cut-off wind speed according to the machine power curve. The wind turbine is designed to stop if . However, M_t is not observed. The only wind speed we observe is R_t. For a given R_t it can be interpreted that M_t is a random variable that needs to be predicted. Once the conditional distribution of M_t is characterized, we can compute the probability of a cut-off

.

This probability can be used, for instance to correct a, say, traditional wind power prediction where the cut-off event is not taken into account. The goal then is to build a model to predict M_t using R_t as input. The steps to build this model are the following. First, the best possible model describing the relationship between and must be found, comparing alternative approaches and deciding through the analysis of the errors of the residuals which is the best one. This model will be used to predict a conditional mean. Residual analysis is performed in this step with the aim of obtaining a conditional normal distribution. If conditional normality is attained, the forecasting exercise needs to be completed by the definition of its conditional variance. This conditional variance will be calculated using also alternative approaches. The results will be evaluated using some real data from the Klaus storm.

1.2 Models for M_t

In this section, real data measurements from a wind farm located in Spain is used to illustrate the methodology. Here Rt is the wind speed from an anemometer mast and Pt is the output power of a single wind turbine. Table 1 contains its most representative speed values. For the whole wind farm, a wind farm power curve can be computed by adding the power curve of all the wind turbines.

Cut-in wind speed (m/s)	Rated wind speed (m/s)	Cut-off wind speed (m/s)
3	12	23

Table 1. Speed specifications of the wind turbine

The nominal power is 1500kw, but hereinafter power will be represented as % of nominal in the graphs, as it is the most common and useful notation. The algorithm developed starts facing in an x-y diagram, for a given power, the values of M_t and R_t as seen in Figure 2. The result is shown in Figure 3. Note that Mt wind speed data larger than the rated wind speed and lower than cut-in wind speed cannot be identified.

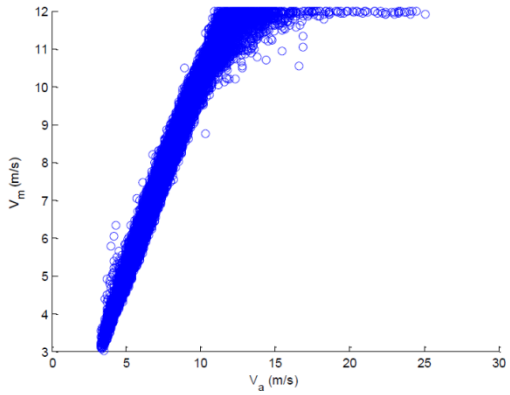


Figure 3. Wind farm anemometer wind speed vs. ideal wind speed

Also, in order to avoid the influence of the change of regime in the output power (from stop to run) the data in the vicinity of cut-in and rated wind speeds will also be discarded. Figure 4 shows the selected range of data that is used to build the model. Shaded areas are discarded. The model can be written as

)+

Several models for) are proposed. The evaluation has been made by splitting the sample in two parts. The model is fitted to the first part of the whole set of points. The second part of the data is used to evaluate the quality of the forecasts of each model. Figure 5 shows the predictions made by each of the models fitted to this set of points, and the models are shown in Table 2.

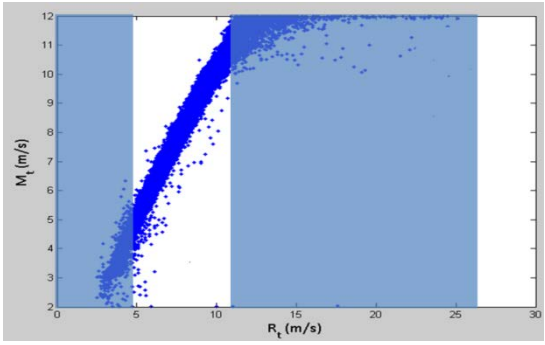


Figure 4. Sets of points used to estimate the models and forecast the errors

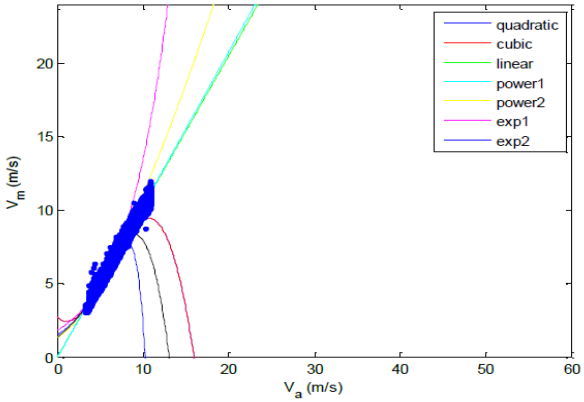


Figure5. Fitting of different pattern models to the set of Real vs. Ideal speeds

Fitted model	Equation
Linear	
Sqrt	
Quadratic	
Cubic	
Local regression	
Power 1	
Power 2	
Exp 1	
Exp 2	

Table 2. Equations of the models fitted to the set of points

Models such as cubic, exp2 or quadratic are completely useless according to the result of the fitting. Also the exp1 model could be discarded. Local regression has also a poor performance. Besides, local regression is of doubtful utility here, given that our purpose is to extrapolate outside the range of observed data. The error measurements used to evaluate the quality of the fittings will be the MAE (Mean Absolute Error) and the RMSE (Root Mean Squared Error), defined as:

Figure 6 shows the MAE of the best subset of models. RMSE shows similar pattern and is omitted.

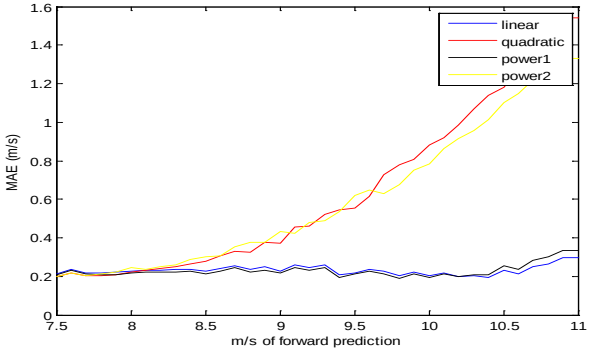
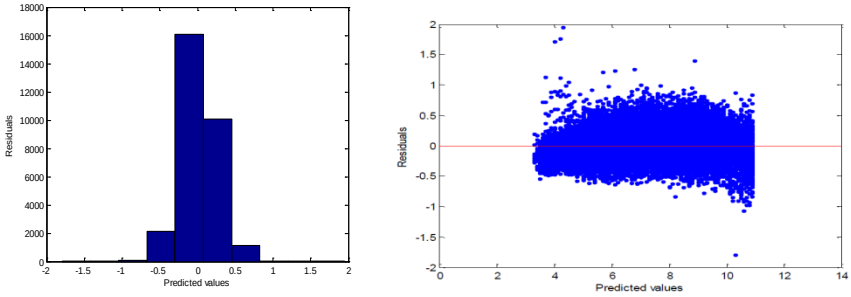


Figure 6 .MAE of some selected models

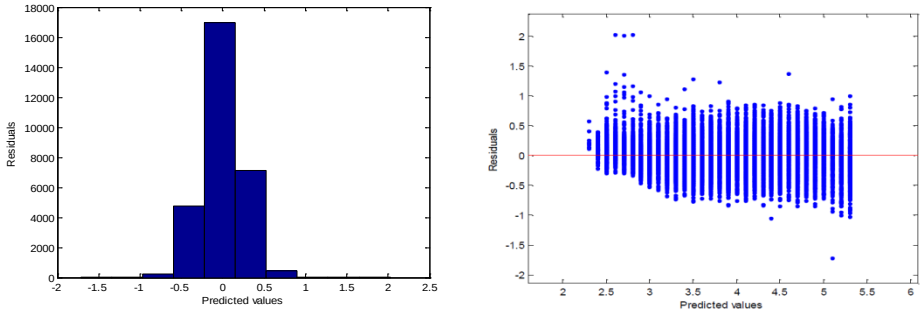
The conclusion is that both the linear and the model based on a power transformation (Power 2 in Table 2) show the best results. In order to understand these models we are going to analyse the residuals from the linear model. They are shown in Figure 7. Figure 7(b) shows the plot of residuals versus predicted values. This plot shows a lack of linearity and also heteroskedasticity, which makes the linear model inadequate. These results suggest that a power transformation should be used instead.



(a) (b)

Figure 7. Residuals for the linear model

The power transformation is of the type , where is optimized by searching for the highest correlation between and Mt. Some other estimation methods can be used for (see, for instance, Box and Cox (1964) and Carrol and Rupert (1981)). For the example in Figure 7, the chosen value is The residual analysis of this transformed model can be seen in Figure 8. Now the diagnosis of the residuals reveals that the linearity problem is solved. However, some heteroskedasticity is still present, but of a much lower magnitude. It seems that for very low wind speed the variance is lower. Since our purpose is to predict for large values of wind speed, we can ignore this effect.



(a) (b)

Figure 8. Distribution of the residuals of the lineal model after the transformation

The next task is to look for a distribution for the prediction error of the power-transformed model. This distribution will allow for the computation of the cut-off probability. The first distribution tested for the residuals is the normal one. The hypothesis of normality has been rejected in the analysed cases, as well as other similar distributions like the logistic one. However, what we need is to model the distribution of the residual conditioned to a value of the input variable. Figure 8 suggests that, it might be possible that the distribution of the whole set of residuals, that is, the unconditional distribution of the residuals, could just be the mixture of different normal distributions, but with different variance, which all together do not fit to a normal one. This effect would be a consequence of the heteroskedasticity of the whole set of residuals. Taking for example the residuals for a predicted value of 4,3 m/s, and doing the Kolmogorov-Smirnov test to check its normality we obtain p-values large enough in all of the cases. Figure 9 illustrates this conditional normality of the residuals. Similar analysis has been performed of more values of the predicted speed, with similar conclusions. Therefore we can count on the normality of the residuals of the power-transformed model.

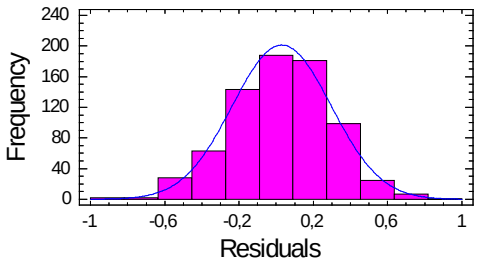


Figure 9. Normal distribution of residuals for 4.3 m/s

The model for M_t is then

with

1.3 Models for the variance of e_t

In the case of homoskedasticy of the error term of a regression, the variance of the prediction of M_i for a given input is, according to classical regression theory (Draper & Smith, 1998)

Where T is the sample size, is the sampling average of and is the variance of the error term, estimated with the residual variance. However, since we are going to use the model to predict outside the range of measured values, we will build a model where this extrapolation exercise is, somehow, simulated. To this aim, the variance will be estimated using an increasing window of data, sorted according to . That is, the sample will be divided into several nested subgroups, where each subgroup contains the data of the previous subgroup plus new data. For example, the second set (B) includes all the observations of the first one (A) plus new data. The third set (C) includes those of the second (B) and the first one (A) plus new data, and so on. Figure 9 shows this increasing window scheme used in the estimation of .

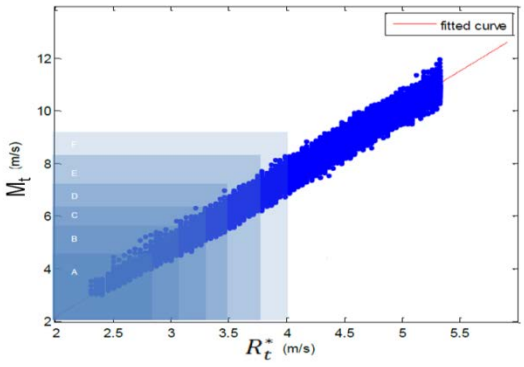


Figure 9. Set of increasing windows

For each window, we estimate the power-transformed model in the estimation subsample, and predict the remaining points of the forecasting subsample. Each out of sample prediction is classified in bins according to the prediction range, in a similar fashion as the prediction horizon in time series, as shown in Figure 10. For instance, if the estimation subsample covers the values of up to 4 m/s and we predict the observation corresponding to =4.5 m/s the prediction range is 0.5 m/s. At the end, we will have a large number of prediction errors, each one associated with a prediction range (or prediction horizon, measured in m/s). The goal then is to make a model that relates the variance of the prediction errors with the prediction range. The variance of the prediction error is computed by averaging the squared of the prediction errors in each horizon.

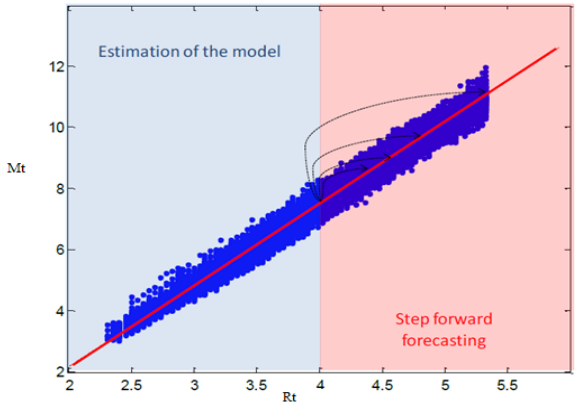


Figure 10. Forecast of wind speed at different ranges of forecasting

As a result, a discrete number of points defining the variance as a function of the range of forward speed forecasting will be available (see Figure 11). Then, the only thing left is to fit these points to a polynomial model in order to estimate the variance of the conditional distribution of M_t given R_t.

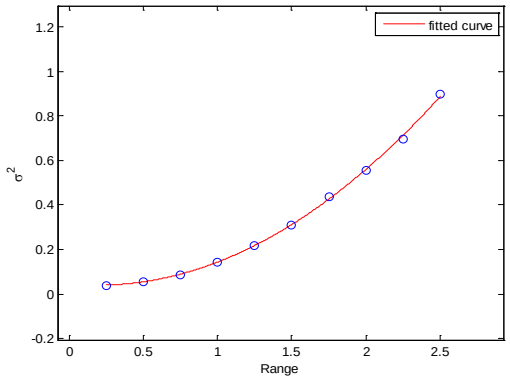


Figure 11. Fitting of a quadratic model to the variance of the residuals

Several polynomials have been used to fit this curve. In all the analysed cases, the best fit is obtained with a quadratic model.

1.4 Summary of the procedure

The proposed methodology for the prediction of a cut-off event can be summarized in the following steps:

Step 1: Collect historical data on generated power P_t and predicted wind speed R_t.

Step 2: Build a wind farm machine power curve by adding the machine power curves of each wind turbine.

Step 3: Compute the machine wind speed M_t corresponding to each value of P_t by using the wind farm machine power curve.

Step 4: Select the data (R_t ,M_t) in the range of ($M_{cut\text{-}in}$ +1) and (M_{rated}-1), where $M_{cut\text{-}in}$ is the theoretical cut-in wind speed and M_{rated} is the theoretical rated wind speed according to the machine power curves. This selection can be more restrictive in order to assure that the selected data is free from transition effects to the different regimes of the wind farm.

Step 5: With the selected data (R_t ,M_t) fit a power transformed model by ordinary least squares

with , where the power is the one that maximizes the correlation with M_t. More complex estimation procedures could be used like fitting a Box-Cox transformation by maximum likelihood.

For a given wind speed the machine wind speed is a random variable that is assumed normally distributed with mean

.

Step 6: With the residuals of the previous model fit a polynomial model for the heteroskedasticity. The model is

For a given wind speed the machine wind speed is a random variable that is assumed normally distributed with variance

Step 7: For each R_t , compute the probability of cut-off by

where M_{cut} is the cut-off wind speed according to the machine power curve and
,).

Step 8: Use this probability to correct your wind power prediction for time t, with the risk of having a cut-off

Alternatively, in order to increase robustness, some threshold can be applied to the probability to be used as a correction of the predicted power. That, is, the corrected power would be

where is some threshold probability below which the correction is not applied, and is the indicator function.

1.5 The Klaus storm

The Klaus Storm was a windstorm which made landfall over several areas of central and southern France, Spain, Andorra and also some parts of Italy on the 23rd and 24th of January 2009. This storm caused serious damages, especially in northern Spain and caused up to twenty-six fatalities, as well as problems for public transport and power supplies, with approximately 1.7 million homes in southwest France and tens of thousands of homes in Spain under power cuts. Severe damage to property and major forest damage was caused as well. During the storm, peak gusts were over 200 km/h in Spain, being some of the highest wind speeds ever registered in this country. Therefore, this storm represents a great opportunity to test the forecasting algorithm. We have collected data from different wind farms located northern Spain. For the sake of conciseness in this work we only include one wind farm. Figure 12 shows the wind farm machine power curve, built as the aggregation of the individual machine power curves. The figure also shows he observed data on predicted wind speed (from a NWP provider) and observed output power. All data are hourly mean values. The cut-off speed of these turbines is . Once the whole model of the wind farm is defined, the forecasts will be compared with the real value of generated power at the moment of the cut-off.

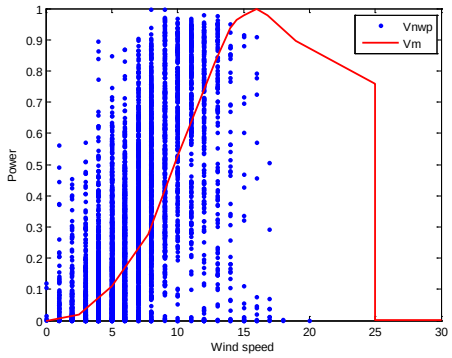


Figure 12. Wind farm machine power curve and actual data

In Figure 13 the pattern between R_t and M_t can be seen. The range of data to build a linear model between both wind speeds is . We have used as threshold probability . Table 3 shows the prediction of the mean values of and the probability of cut-off.

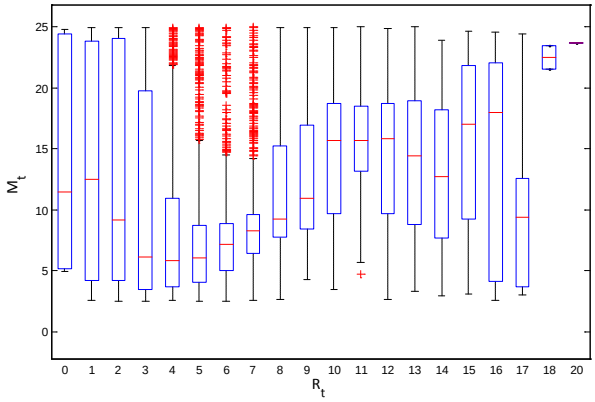


Figure 13. Predicted wind (R_t) versus machine wind speed (M_t)

Table 4 shows the main hours where the cut-off took place are displayed. The table includes the date (hh.mm.ss), the measured wind power production (in the [0,1] interval), the wind power predictions made by a forecasting tool that does not use the information of the cut-off (Pred), the wind speed prediction R_t, the predicted machine wind speed using the proposed methodology , the probability of cut off and the corrected power production.

Date dd.hh.mm	Production	Pred	**New Pred**	V_{NWP}		P(cut-off)
23.16.00	0.9096	0.2615	0.2615	15	24.2571	0.4436
23.17.00	**0.3389**	**0.3695**	**0.1547**	**16**	**26.0731**	***0.5812***
23.18.00	**0.0134**	**0.1785**	**0.0520**	**17**	**27.8800**	***0.7088***
23.19.00	**0.0388**	**0.2149**	**0.0626**	**17**	**27.8800**	***0.7088***
23.20.00	**0.0706**	**0.1727**	**0.0503**	**17**	**27.8800**	***0.7088***
23.21.00	**0.0131**	**0.2245**	**0.0417**	**18**	**29.6784**	***0.8142***
23.22.00	**0**	**0.1916**	**0.0356**	**18**	**29.6784**	***0.8142***
23.23.00	**0.0013**	**0.1872**	**0.0348**	**18**	**29.6784**	***0.8142***
24.00.00	**0**	**0.3707**	**0.0001**	**26**	**43.8113**	***0.9998***
24.01.00	**0.0012**	**0.3616**	**0.0208**	**20**	**33.2517**	***0.9425***
24.02.00	**0**	**0.0905**	**0.0379**	**16**	**26.0731**	***0.5812***
24.03.00	0.0064	0.1162	0.1162	14	22.4314	0.3119
24.04.00	0	0.0416	0.0416	14	22.4314	0.3119
24.05.00	0.0019	0	0	14	22.4314	0.3119
24.06.00	0.0009	0.0888	0.0888	14	22.4314	0.3119

Table 4. Results of the methodology to the Klaus Storm in a wind farm

The results of Table 4 reveals that the RMSE of the wind power prediction without the correction for cut-off risk is RMSE= 0.2453 and MAE=0.1849, whereas the RMSE of the corrected predictions is RMSE=0.1794 and MAE=0.0874. The reduction in error is very large since the large deviations due to the cut-off event are avoided.

Bibliography

- Box, G.E.P., and Cox, D.R. (1964). "An analysis of transformations". Journal of the Royal Statistical Society, Series B 26 (2): 211–252.
- Carroll, RJ and Ruppert, D. On prediction and the power transformation family. Biometrika 68: 609–615.
- Draper, N.R., and Smith, H. (1998). Applied Regression Analysis, John Wiley and Sons. Canada.
- Hosmer, D.W., and Lemeshow, S. (2000). Applied Logistic Regression, 2nd ed.. New York; Chichester, Wiley.
- Hunter, R., Dunbabin, P., Pedersen, T.F., Antoniou, I., Frandsen, S., Klug, H., Albers, A., Lee, W.K. (2001). “Measurement Method to Verify Wind Turbine Performance Characteristics” . Deliverable of Task 1 of EU project “European Wind Turbine Testing Procedure Developments”, contract no. SMT4-CT96-2116. Risø National Laboratory, Roskilde.
- Kristensen, L., Højstrup, J., and Rathmann, O. (2004). “Power Loss in the Upper Cut-out Wind Speed Interval”, Wind Energy, 7, 37-46.
- Hosmer, D.W., and Lemeshow, S. (2000). Applied Logistic Regression, 2nd ed.. New York; Chichester,

2. Cut-off prediction module by ENFOR

2.1 Introduction

Modern wind turbines are equipped with a safety feature which ensures that the wind turbine shuts down to prevent damage during periods with very high wind speeds. Each turbine measures the current wind speed at the nacelle, and when this wind speeds become too high it will cease power production and shut down. Typically 'too high' means that the current wind speeds exceeds a threshold, known as the cut-off wind speed. The cut-off wind speed is normally somewhere in the range from 20 − 25 m/s. After a while when the wind speed has dropped the wind turbine will resume production again, see Figure 1.

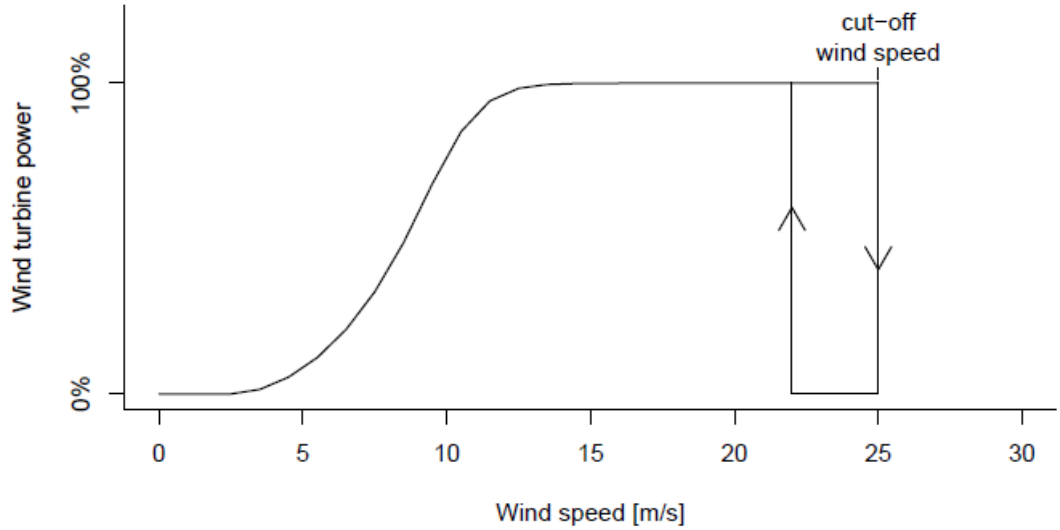


Figure 1: Wind turbine power curve with wind speed cut-off limit

Prediction of cut-off events is important to wind farm owners, since it has a significant impact on power production. A sudden strong wind may shut down and entire wind farm, meaning that the wind farm will go from full to no production in a short period of time.

There are a number of inherent difficulties in predicting the occurrence of cut-offs. The wind turbine reacts to short lived wind conditions, such a gust winds with a duration of seconds and short wind speed averages of the past few minutes. Such accuracy is not available in day-ahead wind speed forecasts, and statistical methods are therefore needed to predict the cut-off risk. Furthermore events are binary, either it happens or it does not. For this reason the cut-off prediction is not well represented in the point power prediction, but is better given as an addition risk probability forecast along with the point prediction.

This report will focus on the treatment of cut-off events as binary, meaning that a cut-off event probability is the proper type of prediction.

It should be noted that some wind turbines have a more gradual decline in power production for high wind speeds, similar to the increase seen for wind speeds between cut-in and rated wind speed. Such wind turbines do not have a hard cut-off wind speed and are thus not the focus here. For these types of turbines the power production for high wind speeds is better captured in the normal point prediction.

2.1.1 Definition of cut-off events

For a single turbine measured on a short time resolution, the cut-off event is well defined. A cut-off event occurs when high wind speeds forces the turbine to do a safety shut-down, resulting in production going from rated power to zero in a short period of time.

For wind farms, a cut-off may happen for only a fraction of the turbines, for example as strong winds are moving across the wind farm area. As this is generally during a transition period (i.e. when a strong wind is moving across the wind farm area), it is preferably to still treat the event as binary, such that a wind farm cut-off is defined as when a certain fraction of turbines has shut down.

For wind farms measured on a longer time scale such as one hour steps, there is the added complexity that a cut-off may have occurred for only part of the time step. Again it is desirable to be have a binary classification. If predictions are given on the one hour time scale but measurements are available on a shorter, say 10 minute, time scale, a one hour cut-off event may be defined as an hour with at least one 10 minute period with a large fraction of turbines shut down due to cut-off.

In general, the definition of event is important and should be determined based on discussions with the user, as it determines the problem which is solved by the cut-off prediction.

2.1.2 Classification of cut-off events

Any choice of cut-off event definition must be supported by the available data for the wind farm. The best data is a direct measurement of e.g. the number of wind turbines shut down due to cut-off at any given time step or a similar type of measurement. Should than not be available, it is necessary to do a manual classification based on measured wind speed and power production. In lack of a measured wind speed the short lead times of a predicted wind speed can be used instead.

In Figure 2 an example is shown with 24 hourly measurements of production together with an hourly classification of cut-offs based on a measured wind speed. In total there are 10 time steps classified as having a cut-off, but there are only two event periods, where an event period is defined as a period of sequentially classified cut-off time-steps.

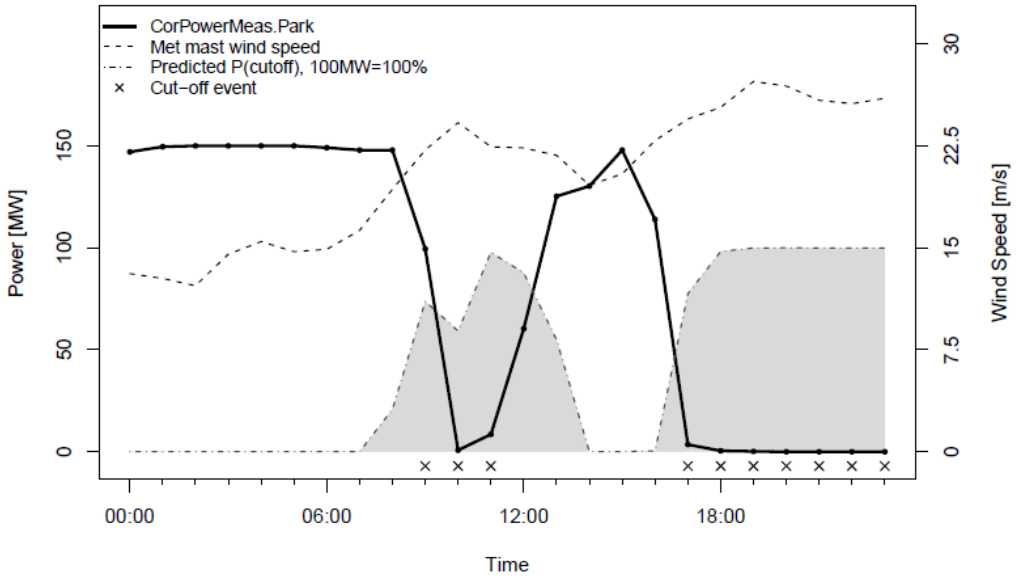


Figure 2: Example classification and prediction during a day with cut-off.

2.1.3 Evaluation of cut-off predictions

In the following a number of methods for evaluation of a cut-off probability prediction will be considered. To illustrate the methods an example data set is used where cut-off has been

N = 2665	Observed	99.9% CI	Model prediction
Cut-off events, rate	2.62%	[1.72%, 3.81%]	2.08%
Cut-off events, count	70	[45.9, 101.4]	55.3

Table 1: Observed and predicted event rate/count.

classified on a one hour time scale. The data is compared to a cut-off prediction with 1-6 hour lead time updated every 6 hour, meaning that all observations are used once in the evaluation. Only time points where data was available such that a possible cut-off could have been positively identified has been used, giving a data set with 2665 hourly observations. During this period 70 cut-offs where identified, this gives an observed cut-off rate of 2.7%.

Overall reliability of predicted probability

The overall reliability of the predicted cut-off risk can be determined by comparing the predicted event rate to the observed event rate. The predicted event rate is given by

– (1)

If classified cut-off events are assumed to be independent, we can compare the model prediction with a confidence interval of the observed event rate. This comparison is found in Table 1.

It is seen that for this case the predicted event rate matches well with the observed.

Contingency table based on Alert threshold

It may be difficult for an operator to properly use a cut-off probability prediction. The probability may therefore be reduced to an alert/no-alert forecast for each time step based on a threshold.

In Table 2 a contingency table is shown based on a threshold of 0.05. The sensitivity is found to be 81.4% and the specificity is found to be 98.4%.

Since single time-step cut-off events are correlated as illustrated in Figure 2, an evaluation of single time steps will over-emphasise performance during long cut-off periods. If we instead count cut-off periods, and count a period as alerted when 50% of the period is above the threshold, we get the table shown in Table 3. No summary is made for columns, as the counted

	Alerted	Not-alerted			
Event time-steps	57	13	70	81.4%	(sensitivity)
No-event time-steps	40	2555	2595	98.4%	(specificity)
	97	2568	2665	3.6%	
	58.8%	99.5%	2.6%		

Table 2: Contingency table of event and alerts (alert threshold is p(cut-off) > 0.05).

	Alerted	Not-alerted			
Event periods	9	4	13	69.2%	(sensitivity)
No-event time-steps	40	2555	2595	98.4%	(specificity)

Table 3: Contingency table of event and alert periods (alert threshold is p(cut-off) > 0.05).

alerts for event periods and no-event time-steps are not of equal duration and thus not directly comparable.

Receiver operating characteristic curve

For a continuous predictor of a binary event, it is possible to derive a receiver operating characteristic (ROC) curve. This describes the predictors performance at various alert thresholds by comparing the true positive rate (TPR) against the false positive rate (FPR). In relation to the contingency tables above, the TPR equals the sensitivity and FPR is 1 - specificity. A ROC curve is shown in Figure 3. From the figure it is seen that the threshold of 0.05 chosen for the contingency tables above gives a reasonable balance between the true positive rate and the false positive rate.

Reliability of cut-off probabilities, binned

The evaluation in Section 2.3 focuses on the overall reliability of the predicted probabilities. However, it is also of interest to see how the predicted probabilities performs for smaller ranges.

Figure 4 shows the observed event rates split by the predicted event rates in groups of 0%−10%,10% − 20%, etc. As before, confidence intervals are based on an assumption of independent observations. If a diagonal line passes through most of the confidence intervals this indicates that the predicted probabilities are reliable.

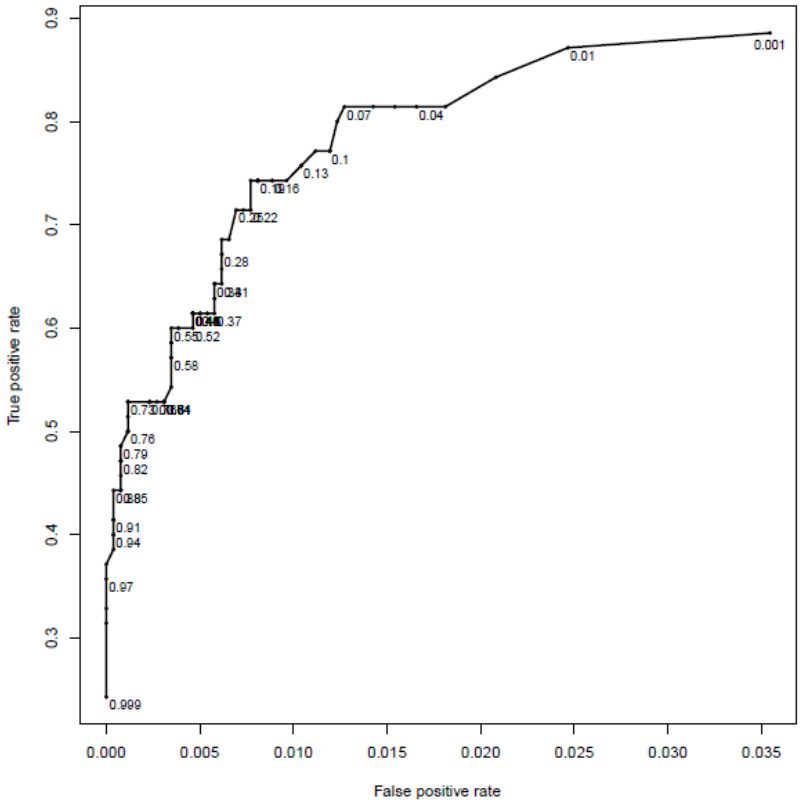


Figure 3: Receiver operating characteristic (ROC) curve for a range of alert thresholds

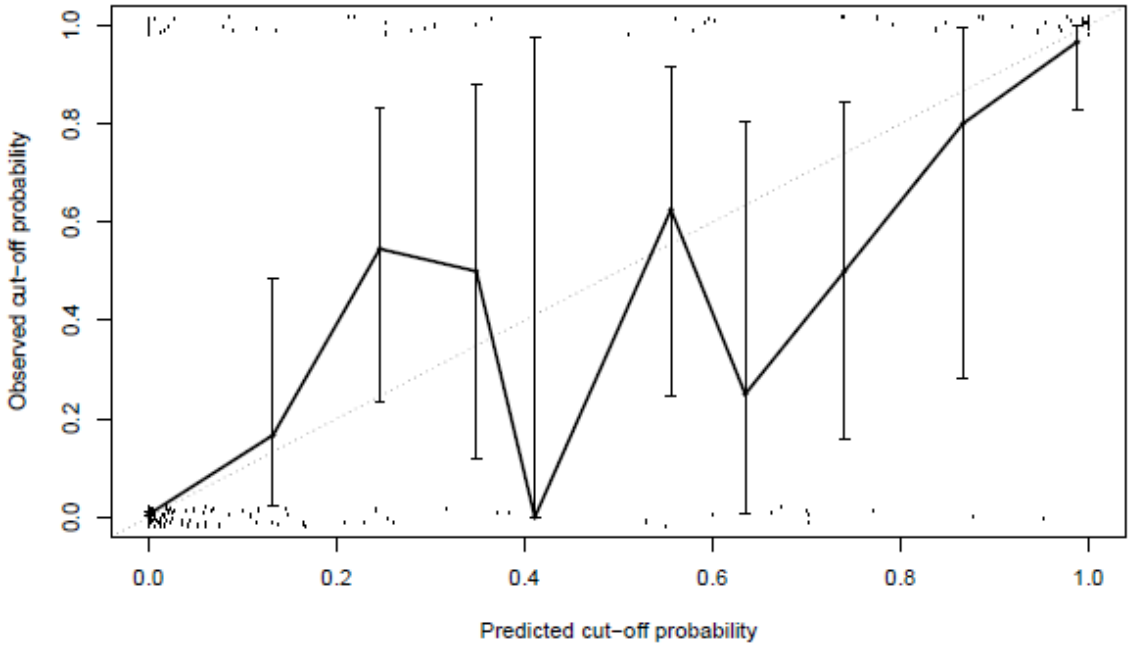


Figure 4: Reliability of predicted cut-off probabilities. Bars are 95% confidence intervals. Dots are event/no-event (1/0) versus predicted cut-off probability.

2.2 ENFOR cut-off module

The ENFOR cut-off module provides a probability prediction of the cut-off risk for wind farms consisting of wind turbines with a sharp wind speed cut-off limit. Predictions are based on latest available NWP update, and follows the NWP in update frequency and maximal lead time. The time resolution may vary independent of the time resolution in the NWP update.

The module requires a wind farm power curve and offline or online measured power and avail- ability for the wind farm. The wind farm power curve can be wind direction dependent, but does not need to be. A sufficient wind farm power curve can be derived by adding the default manufactures free wind power curve for all wind turbines in the wind farm, an example of such a power curve is seen in Figure 1. The module is designed such that it does not rely on any form of internal automatic classification of cut-off events to be used for training, as these events are normally rare and therefore not suitable for training of a statistical model.

The module estimates a model which calibrates the NWP wind speed forecast to the wind speed used in the wind farm power curve. For a non-wind direction dependent power curve this corresponds to a calibration from NWP wind speed to the actual free wind speed experience by the wind turbines. A measurement of this free wind speed experienced by the wind farm as a whole is accurately given by the power measurement corrected for availability. As an example, if a wind farm is producing at 70% of installed capacity this may correspond to an average free wind speed of 9.5 m/s for the wind farm. Along with a calibration model for the NWP wind speed, the module also estimates the uncertainty in the NWP wind speed relative to the derived measured free wind speed. Given the calibrated NWP wind speed and its uncertainty, the module calculates the probability of actual realised wind speed being above the wind farm cut-off limit. As mentioned, using this method the cut-off probability is derived without classification of prior cut-off events which may or may not be available in the available training data.

The model used for calibrating the NWP wind speed is given in equation below

$$\hat{w}_t = a_0 + a_1(\theta_t^{\mathrm{NWP}}) w_t^{\mathrm{NWP}} + \epsilon_t, \quad \epsilon_t \in N(0, \sigma^2) \tag{2}$$

where is the estimated actual wind speed and e_t is the deviation between the actual and forecasted wind speed with variance σ^2. The offset a_0 adjusts for overall bias in the NWP wind speed, and scale parameter a_1 adjusts for scaling differences between NWP wind speed and the free wind experienced by the wind turbines. The scale parameter is estimated with a wind direction dependence to account for shadow effects from other turbines, terrain or other factors which affects the relation between the NWP wind speed and the wind speed experienced by the wind farm. The wind direction dependence is modelled using a lower order harmonic expansion.

The cut-off probability is found as the probability of the actual wind speed being above the cut-off threshold, wcut, i.e. the probability is defined as Pr(> wcut).

2.2.1 Estimation and evaluation

The estimation procedure is designed to give unbiased estimates of the parameters in the calibration model in Eq. 2. As the model is calibrated based on wind speed derived from the power measurement, the best estimates of wind speed is given when the wind farm is producing somewhere along the monotonely increasing part of the power curve typically found approximately in the range 3 − 17 m/s, but still information of actual wind speed when the park is producing at either 100% or 0% is naturally important and also used in the calibration. The estimation of the model parameters is lead time dependent to account for decreasing forecast accuracy for longer lead times and also varying NWP model dynamics for different lead times.

2.2.2 Simulation results

The properties of the estimation procedure have been evaluated based on a simulation study. The aim is to verify that the estimation procedure is able to give unbiased estimates of a_0, a_1 and σ from Eq. 2.

The simulation is based on 4 years of NWP wind speed data for an offshore wind farm with a cut-off wind speed of 25 m/s. The NWP data is updated every 6 hours. For each NWP a sequence of 6 hourly steps in the forecasts is extracted, giving a single complete time series of forecasts for one lead time group. For this wind speed forecast time series (wNWP) a number of simulated actual wind farm wind speeds (w_t) have been simulated by varying a_0, a_1 and σ in Eq. 2. The variance σ has been used with values 1, 2, and 3 m/s to simulated forecast accuracy for short, medium and long lead times. The offset a_0 has been used with values 0 and -1 to simulated NWPs with correct low wind speeds and NWPs with a tendency to forecast to high wind speeds in low wind situations (some NWP models add "artificial" wind to avoid numerical problems in low wind situations). The slope parameter a_1 has been used with values 0.9, 1.0, and 1.3 to simulate NWPs with positive and negative bias. The simulated actual wind speed is used to make a simulated power measurement, and the simulated power measurement is then finally used with the estimation procedure to estimate a_0, a_1 and σ and then to calculate a cut-off risk for each time step in the 4 year period.

The results from the simulation study is shown in Table 4 and 5. In the first table the estimation is done while fixing a_0 to 0 and only estimating a_1 and σ, in the latter all three parameters have been estimated.

The first three columns a_0, a_1, and σ show the values used in simulation. The following three Columns a_0, a_1, σ shows the estimates found based on the simulated power measurements. The column column Ncut-off counts the true number of cut-off defined as the number of time steps where the simulated actual wind speed was above 25 m/s. Column E($N_{cut\text{-}off}$) shows the expected number of cut-off events based on the predicted cut-off risk, this is defined as E($N_{cut\text{-}off}$) = and the column CI 95% gives a 95% confidence interval for the average.

a_0	a_1	σ	$\hat{a_0}$	$\hat{a_1}$	$\hat{\sigma}$	$N_{\text{cut-off}}$	CI 95%	E($N_{\text{cut-off}}$)		TPR	FPR
0	0.9	1	0.000	0.904	0.967	6	1.54;15.50	6.2	✓	1.00	0.00
0	1.0	1	0.000	1.006	0.961	23	12.6;38.2	27.8	✓	1.00	0.00
0	1.3	1	0.000	1.314	0.943	390	342.9;441.0	440.7	✓	0.99	0.01
0	0.9	2	0.000	0.905	1.975	10	3.72;21.39	10.5	✓	0.90	0.00
0	1.0	2	0.000	1.008	1.964	29	17.0;45.9	37.8	✓	1.00	0.00
0	1.3	2	0.000	1.319	1.940	473	419.4;531.2	539.6	+	0.96	0.02
0	0.9	3	0.000	0.906	2.992	13	5.58;25.49	22.0	✓	0.92	0.00
0	1.0	3	0.000	1.008	2.985	55	37.8;77.1	66.7	✓	0.80	0.01
0	1.3	3	0.000	1.325	2.959	621	559.2;687.5	704.3	+	0.96	0.05
-1	0.9	1	0.000	0.805	1.030	3	0.339;10.734	0.1	-	0.33	0.00
-1	1.0	1	0.000	0.903	1.031	16	7.61;29.18	6.2	-	0.81	0.00
-1	1.3	1	0.000	1.197	1.025	259	220.7;301.5	163.9	-	0.93	0.00
-1	0.9	2	0.000	0.806	1.979	5	1.08;14.14	1.4	✓	0.60	0.00
-1	1.0	2	0.000	0.905	1.977	21	11.1;35.9	10.5	-	0.81	0.00
-1	1.3	2	0.000	1.204	1.970	330	285.4;379.3	233.0	-	0.93	0.01
-1	0.9	3	0.000	0.808	2.964	9	3.13;20.00	5.3	✓	0.78	0.00
-1	1.0	3	0.000	0.907	2.965	32	19.3;49.7	21.6	✓	0.72	0.00
-1	1.3	3	0.000	1.210	2.970	429	377.8;484.9	350.6	-	0.88	0.03

Table 4: Estimation of only a1 and σ, a0 assumed to be zero

a_0	a_1	σ	$\hat{a}_0$	$\hat{a}_1$	$\hat{\sigma}$	$N_{\text{cut-off}}$	CI 95%	$E(N_{\text{cut-off}})$		TPR	FPR
0	0.9	1	-0.179	0.922	0.966	6	1.54;15.51	8.0	✓	1.00	0.00
0	1.0	1	-0.211	1.029	0.960	23	12.6;38.2	32.2	✓	1.00	0.00
0	1.3	1	-0.271	1.347	0.940	390	342.9;441.1	507.3	+	1.00	0.01
0	0.9	2	0.090	0.896	1.973	10	3.72;21.39	9.6	✓	0.90	0.00
0	1.0	2	0.025	1.005	1.964	29	17.0;45.9	37.2	✓	1.00	0.00
0	1.3	2	-0.157	1.338	1.941	473	419.4;531.2	575.7	+	0.97	0.03
0	0.9	3	0.381	0.868	2.975	13	5.58;25.49	16.3	✓	0.77	0.00
0	1.0	3	0.257	0.981	2.975	55	37.8;77.1	57.2	✓	0.80	0.00
0	1.3	3	-0.058	1.332	2.960	621	559.2;687.5	718.2	+	0.96	0.05
-1	0.9	1	-1.195	0.921	0.964	3	0.339;10.816	3.1	✓	1.00	0.00
-1	1.0	1	-1.209	1.025	0.962	16	7.6;29.2	20.4	✓	1.00	0.00
-1	1.3	1	-1.313	1.348	0.943	259	220.7;301.5	330.8	+	0.99	0.01
-1	0.9	2	-0.809	0.886	1.973	5	1.08;14.14	4.3	✓	0.80	0.00
-1	1.0	2	-0.874	0.993	1.968	21	11.1;35.9	21.4	✓	0.86	0.00
-1	1.3	2	-1.126	1.331	1.947	330	285.4;379.2	387.4	+	0.97	0.02
-1	0.9	3	-0.385	0.846	2.974	9	3.13;20.00	7.6	✓	0.78	0.00
-1	1.0	3	-0.471	0.954	2.973	32	19.3;49.7	29.7	✓	0.78	0.00
-1	1.3	3	-0.950	1.317	2.973	429	377.8;484.9	495.0	+	0.94	0.04

Table 5: Estimation of all model parameters a_0, a_1, and σ.

observed cut-offs. The marks -, √, and + indicates if the expected number of cut-off based on the model predictions is below, inside or above the confidence interval. Lastly, the columns TPR and FPR gives the true-positive-rate and false-positive-rate based on a threshold of 0.05 for the predictions.

The purpose of trying to estimate with $a_0 = 0$ is to see 1) does the estimates improve in situations where there is in fact no offset, and 2) if there is an offset, is the estimate of a_1 and σ strongly affected by assuming no offset. With respect to the first point, it is seen from Table 4 that we get good estimates of a_1, and σ when the both the simulation and estimation is without offset (first half of the table). However, when an offset is introduced in the simulation but not the estimation then a general tendency to underestimate a_1 is observed. The scaling factor for the estimate of the actual wind speed is then to small, and this results in too low estimates of the cut-off risk, seen by many $E(N_{\text{cut-off}})$ being below the CI. When the offset is included in the estimation in Table 5, it is seen that we still get good estimates for a_1, and σ when the true offset is 0, and when there in fact is an offset we see that estimating it more or less removes the bias in estimating a_1. Based on this it is decided to include the offset in the estimation model.

The estimate of σ is important, as this parameter determines the cut-off risk for a given estimate of the actual wind speed for the wind farm. Table 5 shows that it is accurately estimated in all situations.

The general performance of the method is good when evaluated by the expected number of cut-offs event based on the predictions (the column E(Ncut-off) in Table 5), although there is a small tendency to give to high cut-off risks if the NWP is generally to low (a1 = 1.3). The TPR shows that the method in general catches about 80%-100% of the actual cut-off events, and gives false warnings in about 0%-5% of the time periods where the is not an actual cut-off event. Overall the simulation shows that the methods works reliably and with satisfactory accuracy.

3. Detection and prediction of cut-off events, by Overspeed

3.1 Introduction

This deliverable evaluates the predictability of cut-off events. The first two sections describe cut-off events (section 3.2) and illustrate the need for their prediction (section 3.3). The following section, section 3.4, describes the data sets used for this evaluation. Section 3.5 describes an approach for the identification of cut-off events based on time series data of a wind farm and its evaluation. In Section 3.6 an approach for the prediction of cut-off events is presented and different results for its application are presented. The results of this evaluation are summarised in Section 3.7.

3.2 Description of cut-off events

A Cut-off event describes the emergency shut down of a wind turbine at high wind speeds. Beyond a certain wind speed the forces exposed to the blades by the extraction endanger the security of the turbine. When a turbine cuts out its power production drops from nominal power to zero within seconds.

The controller of the turbine follows the wind speed measurements of the anemometer mounted at the top of the turbine and shuts down, when the cut-out wind speed u1 is exceeded (see Figure 1). The exact value of u1 can be different for different turbine types but usually is about 25 m/s as it is for the shown example turbine. To prevent frequent starting and stopping of the turbine when the wind speed is around u1 hysteresis is applied. I.e. the turbine resumes it production not before the measured wind speed has fallen below the cut-in wind speed u2.

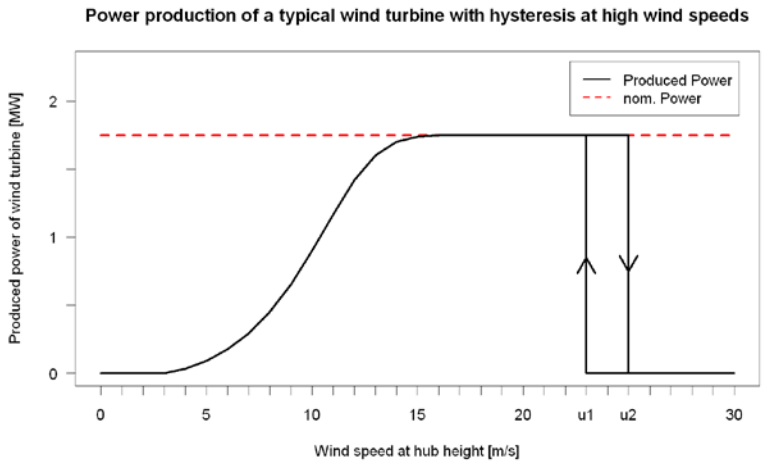


Figure 1: : Power curve of a typical wind turbine with hysteresis at high wind speeds. The turbine cuts out, when the wind speed exceeds u2, and resumes its production, when the wind speed falls below u1.

When cut-off events are analysed on an wind farm level, full and partial cut-offs can occur. Full cut-offs describe a shut down of all turbines of the wind farm as shown in the example time series of the left graph of Figure 2. In this case the production of the whole wind farm drops to zero at half past nine p.m.. As the wind is slowed down while passing through a wind farm, it can also happen that only a sub set of the turbines cuts out. This also causes the production of the wind farm to drop, but only proportional to the number of turbines that have been shut down. The right example of Figure 2 also shows a full shut down between 19 and 20 o'clock. The events before and after that are partial shut-downs.

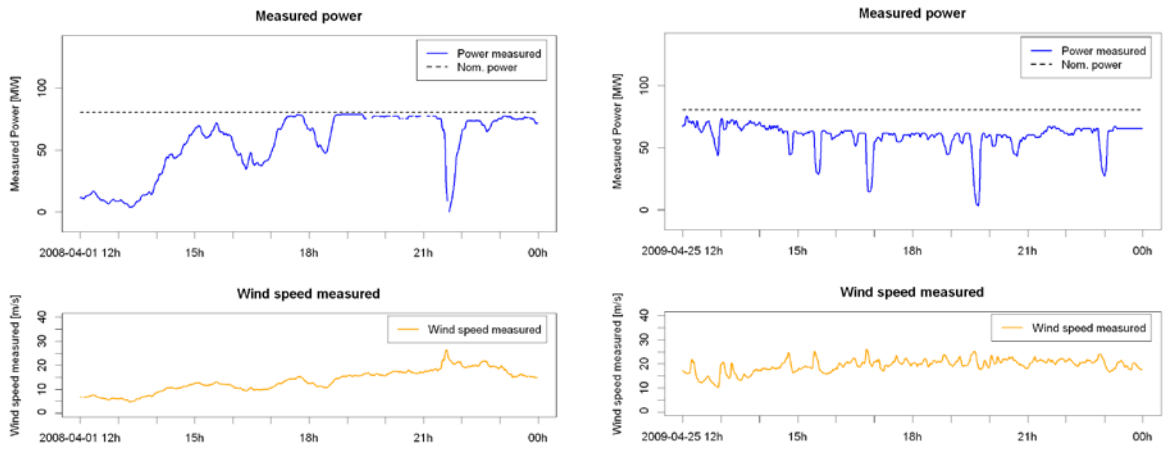


Figure 2: Examples for a full (left) and multiple partial (right) cut offs. Left: The power production of the whole farm drops to zero within seconds. Right: Several times some of turbines of the farm cut out for some minutes.

In the following evaluation full and partial cut-offs are both regarded as cut-offs events and treated equally, but only cut-off events with a drop of more than 30% of its nominal power are considered.

3.3 Motivation for cut-off prediction

Cut-off events cause large ramps in the power production: falling, when the turbines shut down, and rising, when they resume their production. They severely endanger the grid stability due to these rapid and large changes in the power production. Much regulation capacity is needed to handle these events. Today's wind power prediction models focus on the prediction of the average wind power production for the coming hours, mostly day ahead. To minimise the needed balancing energy they are optimised to minimise the average prediction error, which is mostly expressed in the root mean square or mean absolute error. This approach does not address the problem of cut-off prediction.

The following two figures show the predicted power of a physical and a statistical wpp model during two example cut-offs for the Australian farm (left) and one of the Irish farms (right). The physical model uses the numerical weather data as input. The statistical model additionally uses the latest SCADA data available for its predictions. In both cases the predicted power of both models does not indicate cut-off events. This leads to intermediate errors almost as big as the nominal power. Like in these two example situations most of the cut-off events are not recognised by state-of-the-art wind power prediction models.

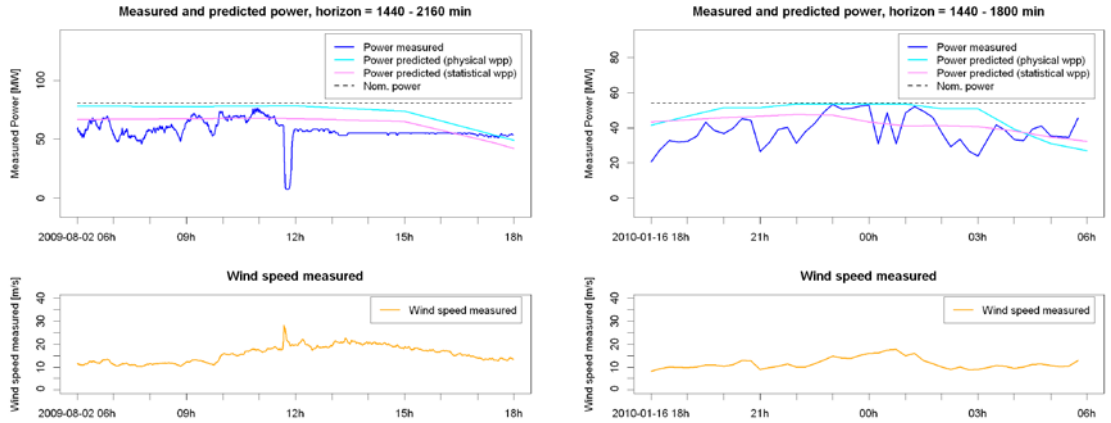


Figure 3: Examples of cut-off events, which are not recognised by state of the art wind power prediction models. In both cases neither the statistical (pink line) nor the physical (turquoise line) predict the cut-off.

Figure 4 shows two situations. The prediction of the statistical model (pink line) even corresponds to quite low wind speed around 12 m/s, whereas the physical model predicts full production, but no cut-off events. But taking the prediction error into account, it is questionable if the predictions are helpful in these cases. As the cut-offs mostly last for less than half an hour and many of them are only partial cut-offs the prediction error of the physical model is huge. Ignoring the indication of cut-off events as done by the statistical model would even lead to a much lower error for this period. This indicates that wind power prediction and cut-off prediction have different aims of optimisation which contradict like for the prediction of ramp events. In addition it raises the question how events, which last shorter than the resolution of the nwp models can be predicted based on nwp data.

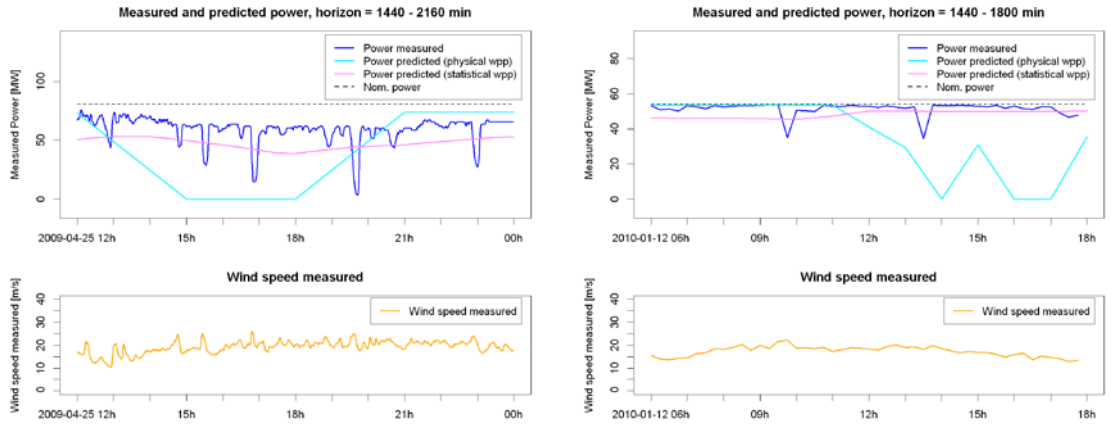


Figure 4: Examples of cut-off events, which are recognised by the prediction models. In both cases the physical prediction model (turquoise line) indicates cut-off events, leading to an even higher prediction error though. Both events are not predicted by the statistical model.

Therefore new approaches accompanying today's daily wind power forecast are needed for the prediction of cut-off events.

3.4 Description of data sets

The data used for this evaluation comes from four wind farms in two different countries. The Lake Bonny wind farm with 80.5 MW installed capacity is located in the south western part of Australia near the shore. The SCADA data of this wind farm provides power, wind speed as well as wind direction measurements in one minute resolution. Three years of data from January 2008 till December 2010 have been available for the evaluation. The power data has a very good quality with an availability of more than 97% for the time frame. The wind speed and direction measurements data in contrast have some significant gaps for approx. 10% and 20% of the time frame. The numerical weather data used for this wind farm comes from the ECMWF with 3 hour horizon steps, an update frequency of 12 hours and a horizontal resolution of 0.25° and ~16km respectively.

The other three wind farms Slieve Rushen 2 (nominal power = 54 MW), Garves (nominal power = 15 MW) and Altahullion 2 (nominal power = 11.2 MW) are located in North Ireland. Slieve Rushen 2 is located in the south west close to the border of Ireland, Garves and Altahullion 2 in the northern part but still 10-15 km away from the sea. For the Irish wind farms almost two years of data has been available for the evaluation from December 2009 till October 2011. The available SCADA data has a 15 minutes resolution and provides power and wind speed measurements. For all three wind farms the wind speed data has a very high availability with almost 100%, whereas the power data is available for ~92% of the time frame. The numerical weather data used for these wind farms comes from the BMO with 1 hour horizon steps, an update frequency of 6 hours and a horizontal resolution of 12 km).

3.5 Detection of cut-off events

In this section two approaches for the identification of cut-off events are presented. Section 3.5.1 describes manual inspection of time series data. In section 3.5.2 filter for their detection are presented and evaluated.

3.5.1 Manual inspection

The prerequisite for the development and evaluation of methods for the prediction of cut-off events is to know about their occurrence. When and where did they occur for how long and how many turbines have been affected? Having this information about the events is needed, but unfortunately mostly not available. As a consequence the only way to get the needed information about cut-off events is an inspection of available measurements of the wind farms. Other type of events like ramp events have a definition, e.g. power change of 40% of the nominal power within 3 hours, whose application to the measured power time series of the wind farm leads to their unambiguous identification. There is no such definition for cut-off events.

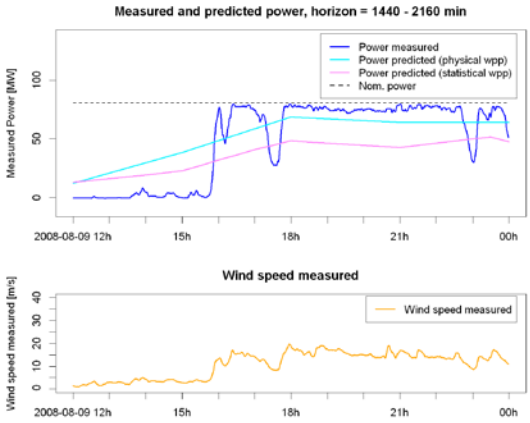


Figure 5: Example for a time series of measured power that could be classified as a cut-off event. A look at the measured wind speed time series shows it is not.

As illustrated by the example of Figure 2 to Figure 4 searching for a steep drop within the power production followed by an increase of the same magnitude is a helpful indicator for the identification of cut-off events. The power time series of Figure 5 could also be easily identified as a full shut down by this characteristic. But only relying on the power production can be misleading. The power time series of Figure 5 also shows characteristics of a partial cut-off at half past seven p.m. and around 11 o'clock p.m.. But taking also a look at the wind speed data reveals that the power change is not caused by cut-off events. The steep changes of the power production are caused by wind speed changes at the very steep part of the power curve between 5 and 15 m/s (see also Figure 1). This highlights the importance of the wind speed measurements for the detection of cut-off events. E.g. the left time series of Figure 4 from the Australian wind farm shows even very small spikes in the production, whereas this detailed characteristics are not shown in the time series of the Irish wind farm in the right graph.

Unfortunately the identification of cut-off events even with power and wind speed measurements is not always as easy as it might be expected. An important aspect for their identification is the resolution of the measured data. The one minute averages of the measurements of Lake Bonny very precisely shows the course of the measurements. Even changes for only a few minutes can be clearly identified. This is not the case for the data of the Irish wind farms, for which only the average of 15 minutes is available. The lower resolution smoothens the characteristics of the event, which makes the identification of cut-off events difficult. A full cut-off, which lasts for a couple of minutes will only cause a drop of 50% of the nominal power within this data. Especially hard is the identification of partial cut-offs as a small drop of the power production can also be caused by variations in the wind speed during the averaging period.

In addition the identification is made difficult due to interference of the TSO. Down regulation of some turbines (see Figure 6, left) or the whole farm as well as shut-downs of turbines for maintenance can show similar characteristics as cut-offs in the power production. In combination with a low resolution of the measured data the separation of cut-off events and interference of the TSO becomes very hard.

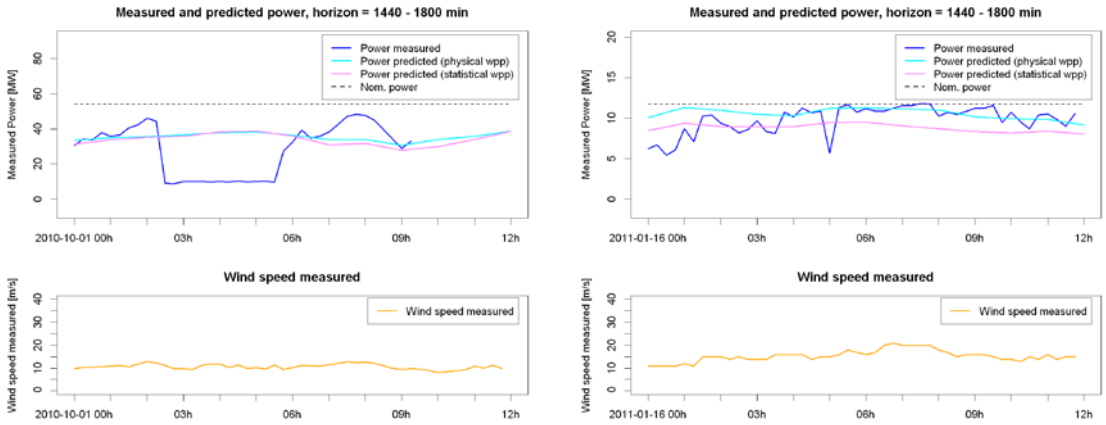


Figure 6: Examples, which illustrate some difficulties of the classification of cut-off events by time series data. Left: The data may show cut-off-like characteristics due to down regulation of the TSO. Right: The time series shows an ambiguous characteristic.

Figure 6 (right) shows an event that is very ambiguous. When the power production drops rapidly the wind speed stays constant and is at a quite low level. When the power production starts to rise again after the drop of the power production the wind speed even increases. During subsequent hours much higher wind speeds do not seem to cause any turbines to cut out. The power drop could be caused by a strong fluctuation of the wind speed causing a full cut-off for some minutes, but can also be caused by down regulation for a very short time. A higher resolution of the wind speed time series would help to resolve this.

For the inspection the minimum power change for partial cut-offs has been set to 30% of the nominal power. Events with a smaller power change have not been regarded as significant as the focus of this work is on extreme events. For the Australian wind farm the manual inspection lead to a number of 43 cut-off events with an average of almost 15 events per year. Due to the high resolution of the data, these results are very reliable. For Irish wind farms Slieve Rushen Phase 2 and Garves 16 and 9 events have been found with an average of 9 and 5 per year. The relatively low number of events that could be clearly identified for the Irish farms is probably due to the low resolution of the data. There have been some events, for which typing was very hard. To minimise the risk of wrong typing these events have been categorised as no cut-offs. In addition for the Irish farms full down regulation has been identified during some periods with very strong wind speed. As Cut-offs during these periods would have been very likely this action has probably been taken to prevent the grid from multiple cut-offs.

The data inspection for the wind farm AL2 has shown a lot of cut-off like events, which could not be explained. Within the power time series 24 out of 45 cut-off like events show a drop of mostly 40-50% of the nominal power in the production between 4:45 and 5:00 in the morning. This raised many doubts about the reliability of the dataset, which has therefore not been discharged for further evaluation.

3.5.2 Automatic inspection

To avoid the effort of a manual inspection of the whole time series two approaches have been evaluated to automatically detect cut-off events in historic data. Filter have been used for the detection using the change in the power production, the wind speed and the correlation between the measured wind speed and the power production as parameters. As cut-off events are caused by high wind speeds the measured wind speed is a very obvious parameter for their detection. Another characteristic of cut-off events is the drop in the power production due to the cut out of the turbines. As

the cut-off is caused by an increase of the wind speed and causes a drop in the power production a negative correlation between these time series is expected. For other situations only positive value for the correlation are expected. The three filter criteria for these parameters are defined as follows:

A. Maximum wind speed. The maximum wind speed u_{max} within x is $\geq u_{cutoff}$.
B. Power change. The maximum falling power ramp Δp within x is larger than Δp_{cutoff}.
C. Correlation. The correlation coefficient r of the measured wind speed and the measured power production of the time frame x and the previous hour (a-2:00 to a:00) is less than r_{cutoff}.

x is a one hour time frame x ranging from a-1:00 to a:00 with $a \in \{1\text{-}23\}$. The filter F_1 classifies all x as cut-off events, for which A and B are true. The second filter F_2 also utilises C. They are defined as follows:

$$F_1: x_{cutoff} = \{ x \mid A \wedge B \} \quad \text{and} \quad F_2: x_{cutoff} = \{ x \mid A \wedge B \wedge C \}$$

As all cut-off events with a decrease in the power production of more than 30% should be considered at least 0.3 is used for Δp_{cutoff}. The exact value of Δp_{cutoff} as well as u_{cutoff} and r_{cutoff} have to be defined for each site individually. They are strongly influenced by e.g. the location and height of the wind speed measuring device, the turbine type as well as the resolution of the power and wind speed time series.

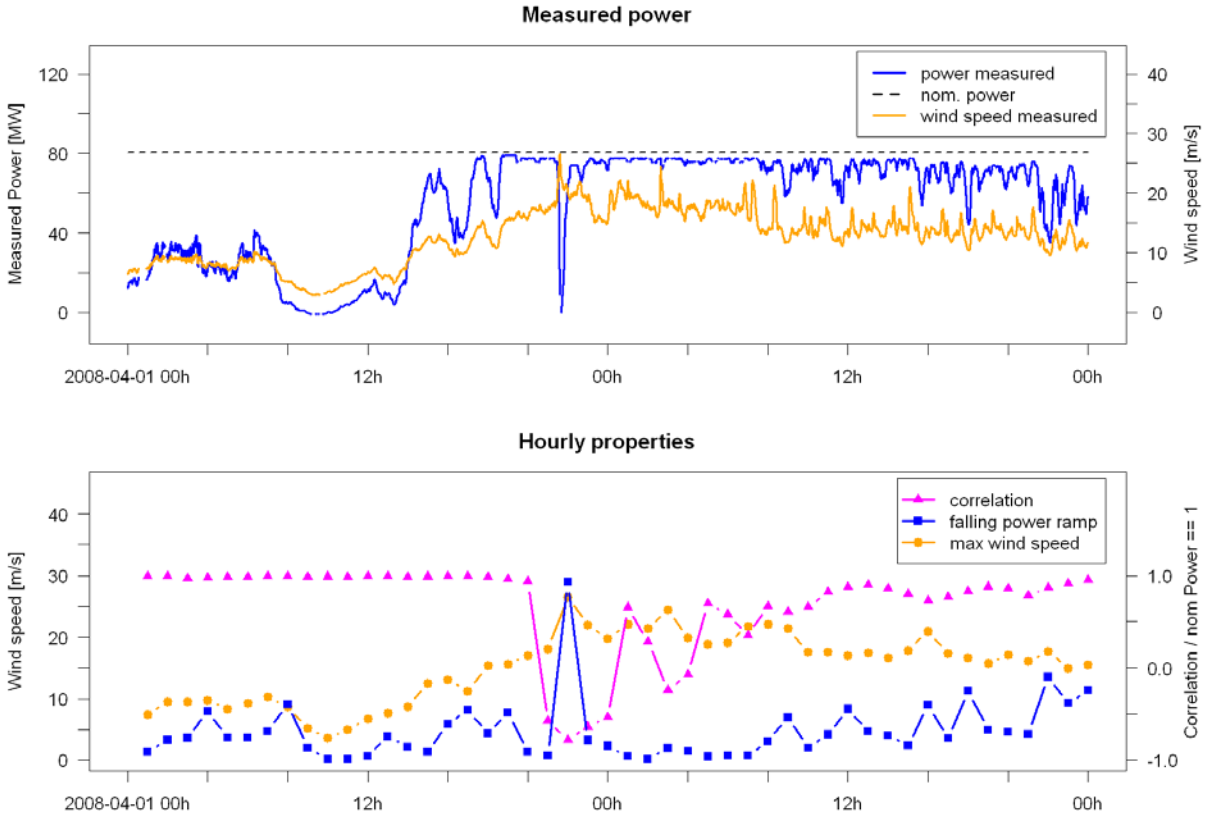


Figure 7: Illustration of the three indicators (power change, wind speed and correlation of measured power and wind speed) used for an automatic detection of cut-off events. During a cut-off event the first two show high and the latter one low values.

Figure 7 illustrates the three indicators A, B and C for the cut-off detection. The blue curve showing the maximum falling power ramp for each hour has a spike during the cut-off event. The hourly maximum wind speed is shown by the orange curve. It also has its maximum value for the cut-off event and shows also high values for some of the following hours. The correlation of the measured wind speed and the power is almost exactly equal to one for the period before the cut-off event. During this period the power production ranges from the cut-in of the turbine to the rated power and changes wind speed cause also changes in the production (see also Figure 1). During the cut-off event the

correlation becomes almost –1 because the wind speed increases while the power drops significantly. As the wind farm production has reached its nominal power for the 10 hours following the cut-off event the correlation also shows values significantly below 1. During this period variations in the wind speed do not affect the power production. For this example time series e.g. Δp_{cutoff} = 40%, u_{cutoff} = 20 m/s and r_{cutoff} = 0 look like reasonable values for the cut-off detection. Figure 8 shows the three properties for an example with multiple cut-off.

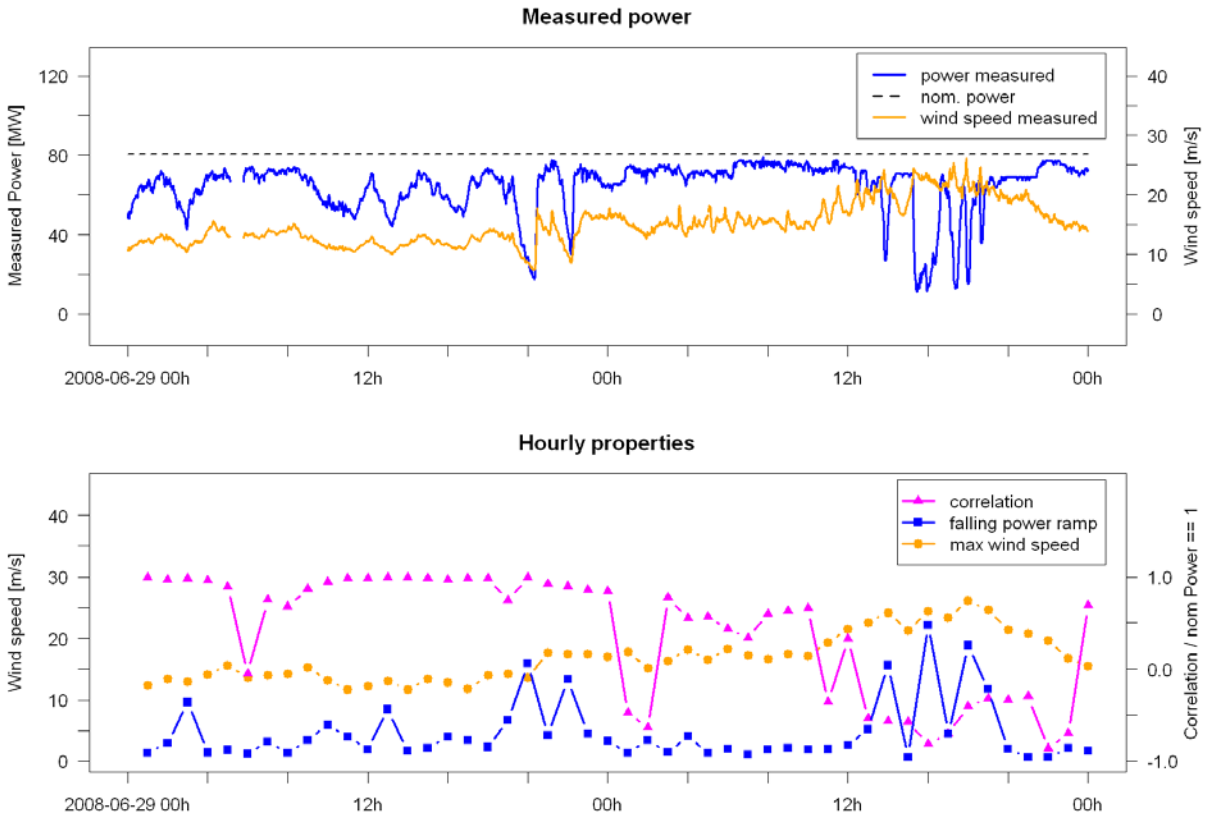


Figure 8: Illustration of the three indicators (power change, wind speed and correlation of measured power and wind speed) for a multiple cut-off event.

To evaluate the use of the three properties the two filters F_1 and F_2 have been applied to the available wind farm data. To separate the tuning and evaluation of the filter the data of the wind farms has been divided into two periods. Due to the very low number of cut-off 9 events for the wind farm Garves, the evaluation has been limited to the wind farms Slieve Rushen Phase 2 and Lake Bonny with 16 and 43 events. For the assessment of the results it has to be noted that 16 events is still quite low number of events for this type of evaluation. But for reasons of comparability the evaluation has also been done for Slieve Rushen Phase 2. In addition to the quite high number of detected cut-off events for Lake Bonny the data set is quite long with three years. To analyse the influence of the length of the training period three training periods with a different length have been evaluated for Lake Bonny (see Table 1). One cut-off event of the training period of III-LKB is caused by a sudden gust and lasts only for about three minutes. As its values u_{max} and r for this x_{cutoff} are differing quite strongly from the other cut-off events a second tuning for this time frame is made (III-LKB*), where this event is ignored.

The cut-offs events of the tuning period, which have been manually detected, have been used to determine the values for Δp_{cutoff}, u_{cutoff} and r_{cutoff}. For this evaluation a rather simple approach has been used for their identification. For Δp_{cutoff} and u_{cutoff} the maximum and for r_{cutoff} the minimum value that still classifies all of the training events as cut-offs has been chosen. This approach gives more weight to detection accuracy. The tuning values are than used for the application of the filters to the evaluation period.

The accuracy of the automatic detection has than been evaluated against the manually detected events. The accuracy is measured with the detection accuracy and the event capture, which are defined as follows:

detection accuracy = no. correct forecasted events / no. forecasted events and

cut-off capture = no. of correct forecasted events / no. cut-off events .

In the following sections the application of F_1 and F_2 will be evaluated.

3.5.3 Cut-off detection based on wind speed and power change

Table 1 shows the information of the different training periods for Lake Bonny and Slieve Rushen Phase 2. For each period the determined tuning parameters for F_1 and the results of their application to the tuning period is shown. The increasing number of events for the different training periods of Lake Bonny lowers the wind speed threshold from 24 m/s to 19 m/s, whereas the threshold of Δp_{cutoff} does not change. For Slieve Rushen Phase 2 the wind speed threshold is significantly lower, as the measured wind speed is much lower in general for this wind farm. This can be due to the averaging and also to different measurement heights. As only 19 of the 20 available cut-off events are used for the training for III-LKB* it has a lower event capture of 95%. The high values for the detection accuracy and the cut-off capture show a very good separation of the cut-off events non-cut-off events for the tuning period. In the following these tuning parameters for F_1 will be used for the detection of events in the evaluation period. Only for III-LKB and I-SL2 3 out of 23 and 3 out of 11 are wrong forecasts of cut-offs.

Data set	Period [months]	No. events	Δp_{cutoff}	u_{max}	Cut-off capture [%]	Detection Accuracy [%]
I-LKB	6	6	- 39 %	24,22	100 %	100 %
II-LKB	8,5	13	- 39 %	24,02	100 %	100 %
III-LKB	12	20	- 39 %	19,23	100 %	87 %
III-LKB*	12	20	- 39 %	22,70	95 %	100 %
I-SL2	10	8	- 33 %	15,68	100 %	72,7 %

Table 1: The threshold for power change (Δp_{cutoff}) and wind speed (u_{max}) have been determined for each training period. The cut-off capture and detection accuracy columns show the accuracy for their application to the training data.

The application of Δp_{cutoff} and u_{max} to the evaluation shows a 100% detection accuracy for all Lake Bonny data sets except for III-LKB (see Table 2). Both periods with 12 month of training show the best results for the event capture with 87%. The very good results for these data sets indicate that a less strict setting of the wind speed threshold would probably also have improved the detection for I-LKB and II-LKB. A look into the data showed that each of the 8 (I-LKB) and 10 (II-LKB) events that were missed but detected in III-LKB, would have been detected with a slightly lower wind speed threshold. For Slieve Rushen Phase 2 the accuracy is much lower than for LKBONNY with only 6 out of 12 forecast being correct.

Data set	Period [months]	No. events	No. events detected	Hits	Misses	Cut-off capture [%]	Detection Accuracy [%]
I-LKB	30	37	24	24	13	64,9	100,0
II-LKB	27,5	30	19	19	11	63,3	100,0
III-LKB	24	23	27	20	3	87,0	74,1
III-LKB*	24	23	20	20	3	87,0	100,0
I-SL2	11,5	8	12	6	2	75,0	50,0

Table 2: The evaluation results of the threshold values of Table 1 are shown. To estimate the accuracy hits, misses, cut-off capture and detection accuracy is given for each time frame.

3.5.4 Cut-off detection based on wind speed, power change and correlation

For the second evaluation with F_2 the correlation (r_{cutoff}) is added as a filter parameter for the cut-off event detection. Table 3 shows that the tuning parameters for r_{cutoff} are quite robust for Lake Bonny, except for III-LKB* due to the outlier. The values for Δp_{cutoff} and u_{max} are the same as determined for F_1. The accuracy results for the application of F_2 to the training data sets shows that adding r_{cutoff} even slightly improves the separation of the cut-off and non-cut-off events for III-LKB and I-SL2 (see Table 1).

Data set	Period [months]	No. events	Δp_{cutoff}	u_{max}	r_{cutoff}	Event capture	Detection Accuracy [%]
I-LKB	6	6	- 39 %	24,22	- 0,31	100 %	100 %
II-LKB	8,5	13	- 39 %	24,02	- 0,30	100 %	100 %
III-LKB	12	20	- 39 %	19,23	0,64	100 %	100 %
III-LKB*	12	20	- 39 %	22,70	- 0,30	95 %	100 %
I-SL2	10	8	- 33 %	15,68	0,45	100 %	80 %

Table 3: The threshold for power change (Δp_{cutoff}), wind speed (u_{max}) and correlation (r_{cutoff}) have been determined for each training period. The cut-off capture and detection accuracy columns show the accuracy for their application to the training period.

The results of the application of F_2 to the evaluation data sets in Table 4 show a decrease of the cut-off capture and an increase of the detection accuracy in some cases compared with the results of F_1. Due to the additional filter criteria fewer events are detected for all cases. For I-LKB, II-LKB and III-LKB* this lowers the event capture of approximate 20% due to a fewer number of hits and therefore an increases of the number of misses. For III-LKB and I-SL2 the detection accuracy has improved.

Data set	Period [months]	No. events	No. events detected	Hits	Misses	Event capture	Detection Accuracy [%]
I-LKB	30	37	18	18	19	48,6	100,0
II-LKB	27,5	30	14	14	16	46,7	100,0
III-LKB	24	23	22	20	3	87,0	90,9
III-LKB*	24	23	15	15	8	65,2	100,0
I-SL2	11,5	8	11	6	2	75,0	54,5

Table 4: The evaluation results of the threshold values of Table 3 are shown. To estimate the accuracy hits, misses, cut-off capture and detection accuracy is given for each data set.

3.5.5 Conclusion of automatic detection of cut-off events

The results of the evaluation confirm that the parameter Δp_{cutoff}, u_{max} and r_{cutoff} are very good indicators for the separation of cut-off and non-cut-off events. Both filter methods have shown very good results for the cut-off detection for Lake Bonny especially for the periods III-LKB + III-LKB* with 12 months of training. The accuracy for Slieve Rushen Phase 2 is significantly lower as for Lake Bonny, which is probably due to the lower resolution of the data.

The large difference of the tuning parameters of III-LKB + III-LKB* and the impact on the detection quality indicate that it would be promising to evaluate different tuning approaches to improve the chosen and rather simple tuning approach. As the results improve with the length of the tuning period and in general only few data is available for tuning a model developing an adaptive tuning approach would be useful.

Nevertheless there will probably be no tuning procedure for F_1 and F_2 to get a configuration to detect all cut-off events with 100% event capture and detection accuracy. An analysis of the whole three years of Lake Bonny data showed that there is no such configuration of Δp_{cutoff}, u_{max} and r_{cutoff}. A chance to overcome these draw backs is only seen by making use of other detection procedures.

3.5.6 Conclusion on cut-off detection

The identification of cut-off events for a wind farm is a challenging task. Mostly there is no data available to researchers about their occurrence but this information is needed to develop approaches for their prediction.

By the first approach presented the events were identified manually by analysing the wind speed measurements and power production time series. The main characteristic that has been utilized is the sharp drop of the power production a very high wind speeds. Cut-off events have been identified for one wind farm in Australia and two in Ireland. This approach has shown that the resolution of the data is of great importance for the identification of cut-off events. The events within the Australian data set with one minute resolution have been identified with a good reliability. The data of the Irish wind farms on the other hand is averaged to a 15 minutes resolution, whereby much of the details of the time series is lost. This makes a reliable classification of some events impossible. The averaging of the data probably also cause misses of some events. But as the detection results can no be verified this can not be proven.

The second approach was an automatic detection of the cut-off events by filtering the time series with two filter utilizing the following three indicators: the power change of the falling power ramp, the maximum wind speed and the correlation between the wind speed and the power production. By a training with 20 events a detection accuracy with an event capture and detection accuracy of 90 has been achieved for a time frame with 23 cut-off events for the high resolution data. To identify all cut-off events with this approach a manual inspection of the time series will probably always be needed.

3.6 Prediction of cut-off events

In this section an approach for the prediction of cut-off events is presented an evaluated. Section 6.1 describes the tuning of this method. In the following two sections evaluation results are presented. The results of section 6.2 are based on the predicted wind speed as it is commonly used in the wind power prediction domain. It is the average prediction of the four corner points of the grid cell of the site. The evaluation of section 6.3 is based on wind speed prediction of only the closest grid point.

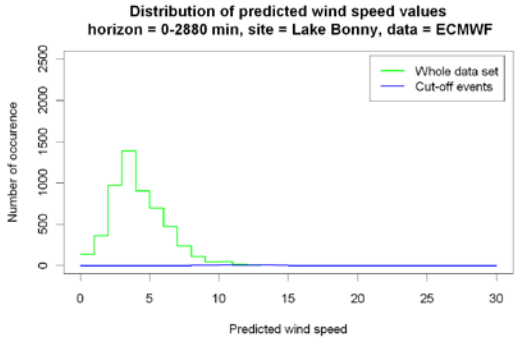
3.6.1 Determination of wind speed thresholds

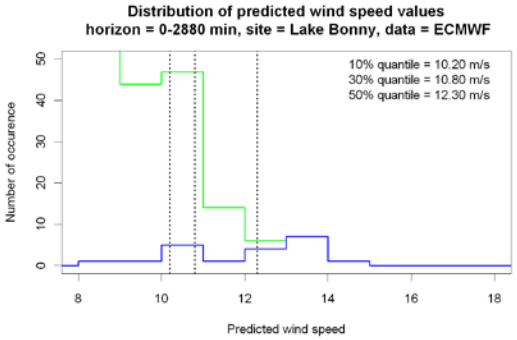
To determine a suitable threshold value for predicting cut-off events, the wind speed predictions of the training periods of section 5.2 have been analyzed. Figure 9 shows the distribution of the predicted wind speed for each training period (green line). In addition these figures show the distribution of the predicted wind speeds of the time stamps closest to each cut-off event (blue line). The left figures show the overall distribution, whereas the right figures are a zoom in to the tail of the distribution at high wind speeds to illustrate the congruence of both distributions at high wind speeds. The nwp data is filtered with a prediction horizon of $t < 2$ days, therefore four predictions are available for each time stamp for both distributions. For each period the 10, 30 and 50% quantile of the cut-off event wind speed predictions has been determined.

The different shape of the distributions for the whole period for the two farms might be caused by the different numerical weather models. The predictions for the Slieve Rushen Phase 2 come from the British Met Office and have a higher temporal and special resolution than the prediction for Lake Bonny, which come from the ECMWF. Due to the higher resolution less temporal and spatial effects do not have to be averaged by the predictions for Slieve Rushen Phase 2. Extreme values are not levelled out and the distribution is spread over a wider range.

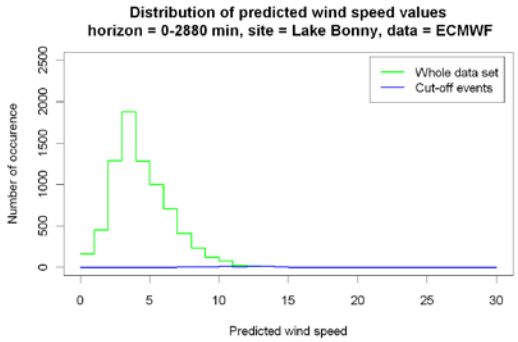
Distribution of predicted wind speeds for the whole data set (green) and for cut-off events (blue) for Lake Bonny and Slieve Rushen Phase 2 for different training periods:

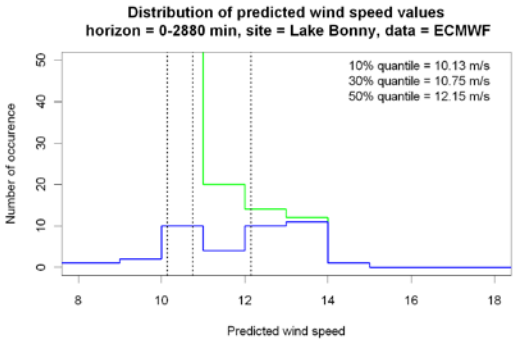
I-LKB:



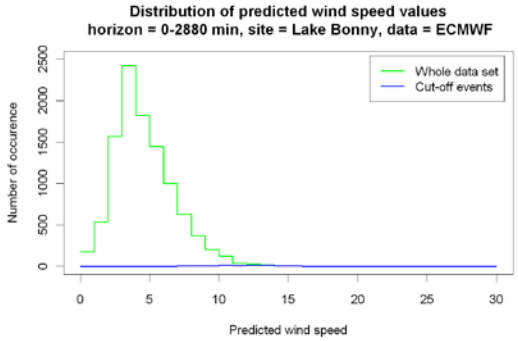


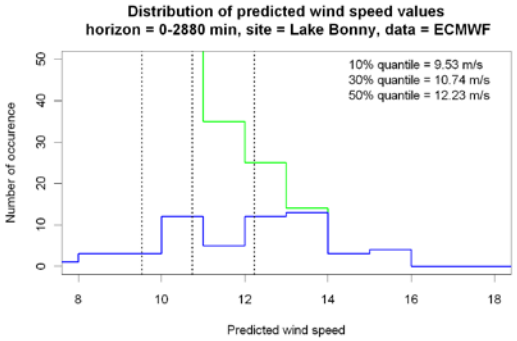
II-LKB:



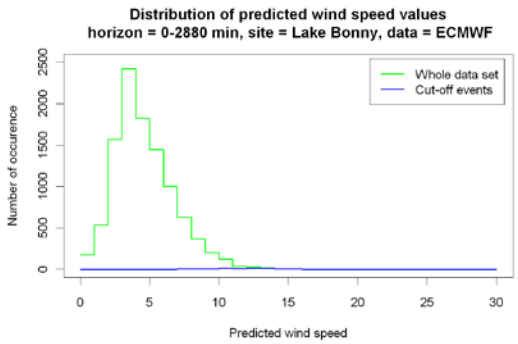


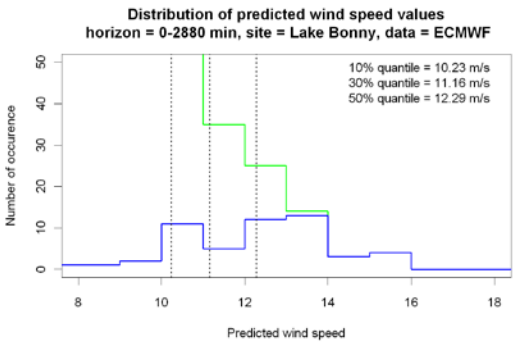
III-LKB:





III-LKB*:





I-SL2:

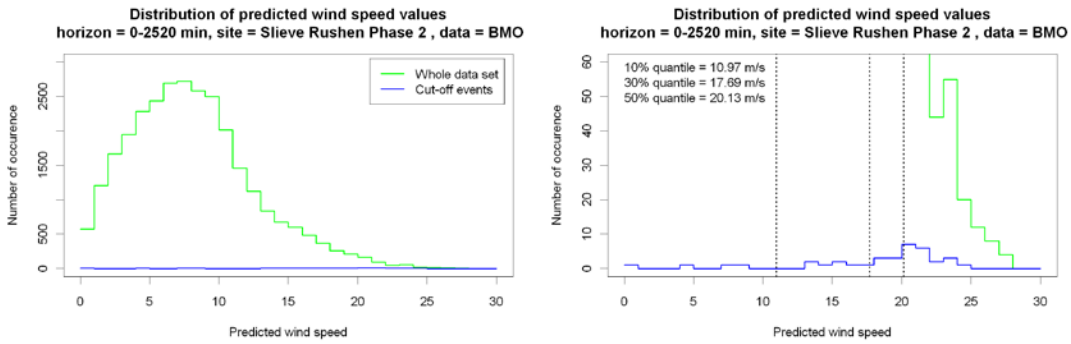


Figure 9: Distribution of predicted wind speed values of the whole (green line) and during the cut-off events (blue) for different training periods. The right graph is a zoom in of the left graph. Threshold values of the cut-off wind speeds are determined by the 10, 30 and 50% quantile of the cut-off distribution.

A comparison of the different training periods of Lake Bonny show a quite good correlation of very high predicted wind speeds and the cut-off events. Almost all timestamps with a predicted wind speed above 13 m/s lead to a cut off event. Nevertheless the gap of the two curves starts to increase rapidly for lower wind speeds. There is only quite little change of the blue curve between the training periods II-LKB and III-LKB. As some of the additional events of the longer period III-LKB are multiple cut-offs covering sequent hours they referrer to the same wind speed prediction. These values are not counted multiple times. The threshold values for the different quantiles is quite stable over the different training periods except of III-LKB. The outliner lowers the 10% quantile value quite strong as it would be expected.

For the Irish farm the distribution for the whole period looks different due to the different characteristics of the wind farm. In contrast to Lake Bonny the distribution of predicted wind speeds close to cut-off events for Slieve Rushen Phase 2 does not look promising for cut-off prediction. The highest wind speeds predicted do not lead to cut-off events and the number of predicted wind speeds leading to cut-off events is only very low compared with the overall distribution.

This analysis already shows that a clear separation of cut-off events and “non cut-offs” based on the predicted wind speed is not possible. Nevertheless for Lake Bonny the results look promising.

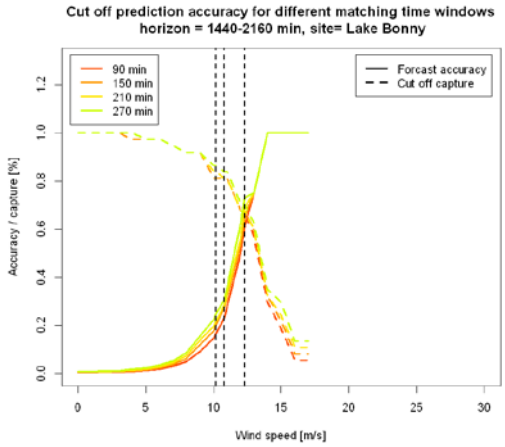
3.6.2 Prediction of cut-off events

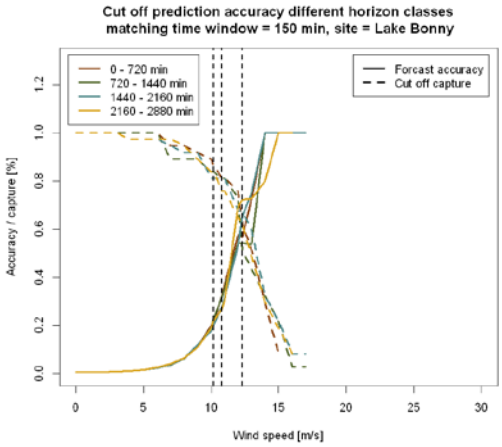
In this section the predictability of cut-off events based on the predicted wind speed is evaluated. The three different data sets cover 24, 27,5 and 30 months with 23, 30 and 37 cut-off events for Lake Bonny and 11,5 months with 8 events for Slieve Rushen Phase 2 (see also Table 1). Various wind speed thresholds have been applied to the evaluation period of each data set. Each of the data sets has been separated according to the wind speed threshold into probable cut-off events and non cut-offs, classifying each wind speed prediction above the threshold as a cut-off event. For these predictions the forecast accuracy and cut-off capture was calculated the same way is in section 5.2.

A cut-off predictions for a time stamp t has been classified as a hit, if a cut off event x_{cutoff} has been detected within a certain tolerance b ($\exists x_{cutoff} : |t - x_{cutoff}| < b$). Various values for b have been evaluated. The lowest possible value for a is dependent on the resolution of the nwp and is therefore 90 min for Lake Bonny and 30 min for Slieve Rushen Phase 2 as their weather predictions have a resolution of 180 and 60 min. The nwp data has also been filtered with different horizon classes, which have been chosen to have unique time stamps in each horizon class. As the predictions for Slieve Rushen Phase 2 are updated twice that frequently the horizon classes differ slightly.

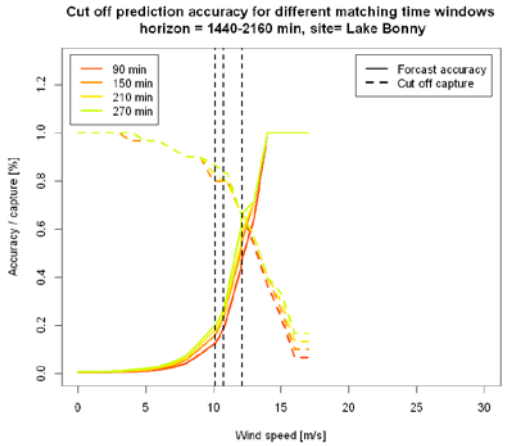
The following graphs show the evaluation for each data set for different horizon classes and matching tolerances:

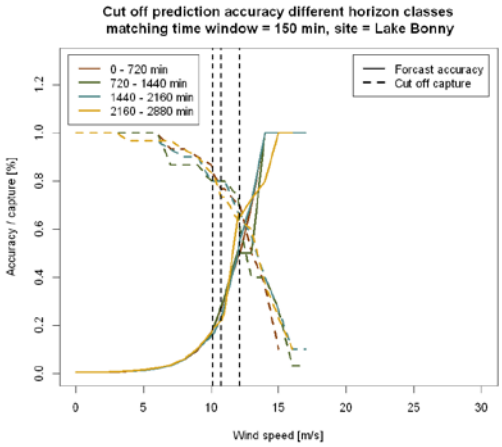
I-LKB:



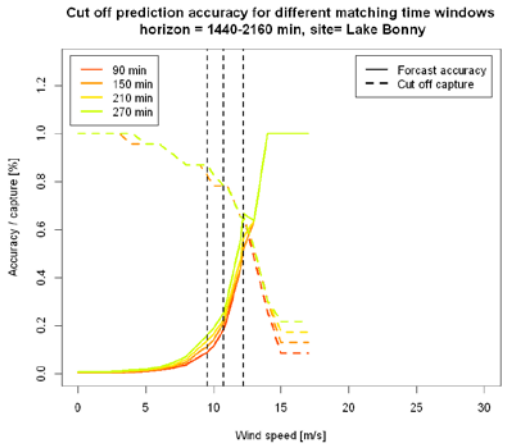


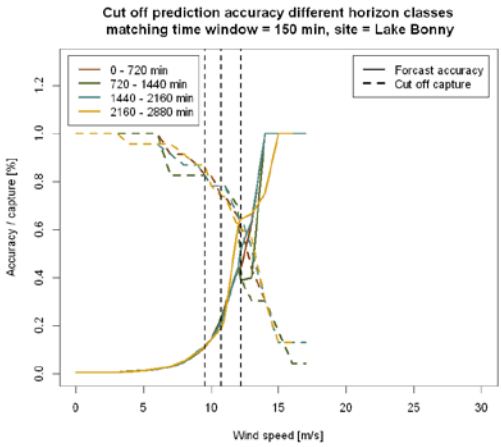
II-LKB:





III-LKB:





III-LKB*:

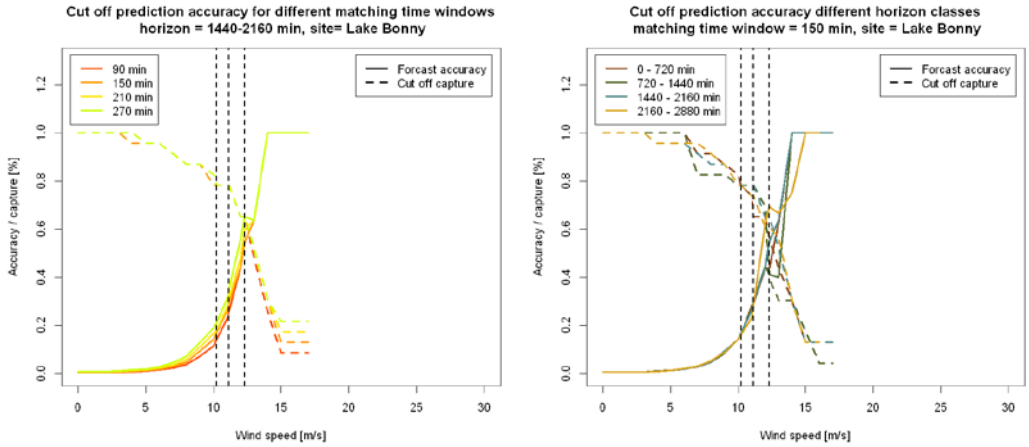


Figure 10: Accuracy evaluation of cut-off prediction based on wind speed thresholds determined in 6.1 for the different evaluation periods of Lake Bonny. The forecast accuracy and cut-off capture is evaluated for different matching time windows (left) and prediction horizons (right).

The evaluation results for the Lake Bonny data sets in Figure 10 show quite similar results. The highest sensitivity of the accuracy values is between 10 and 15 m/s. Within this interval the cut-off capture drops from around 0.8 to 0.2, whereas the forecast accuracy increases from around 0.2 to 1 independent from the prediction horizon or the matching tolerance.

If the intersection of both curves would be regarded as the optimum trade-off of both accuracy measured 0,6 – 0,7 would be achieved for values between 12 and 13 m/s. If the forecast accuracy would be the most important measure 1 could be achieved capturing every third event or more for all cases.

A comparison of the left side graphs show that the variation of the matching tolerance has almost no influence. So the timing of predicted high wind speed events seem to be very precise. This leads to the assumption that the intensity of the remaining events is probably not that well captured within the predicted wind speed.

The prediction horizon also does not seem to have a big influence on the accuracy. For all horizon classes the cut-off capture and the forecast accuracy are very close together for most of the threshold values. The small deviations are probably negligible regarding the relatively small number of events. The only significant deviation of the curves is between the values 12 and 13. Inspection of the data has shown that for these threshold values two periods with 4 cut-off events within 4 and 6 hours are captured differently by the different forecasts horizons having a big impact on the accuracy.

It could be argued that the chosen evaluation method gives too much weight to these events as each period of multiple cut-off events is still caused by one meteorological event. But on the other hand is the impact on the grid for such a period much bigger, which is captured by this evaluation method. Nevertheless as the dependency of the horizon does not show a consistency and shows the best accuracy for the highest prediction horizon it could be doubted that this effect could be generalised.

The vertical dotted lines of the evaluation graphs of Figure 10 are the three chosen quantiles that have been determined in the previous section. It shows that the 50% quantile would be the best choice for this site. For all three periods a wind speed of the 50% quantile hits the intersection of the forecast accuracy and the cut-off capture quite well. The cut-off capture values for the chosen quantiles are slightly higher than the expected 50, 70 and 90%. The accuracy plots also show that the determination of the threshold is a very crucial point. A value that is slightly lower or higher than the optimal value decreases the prediction accuracy significantly.

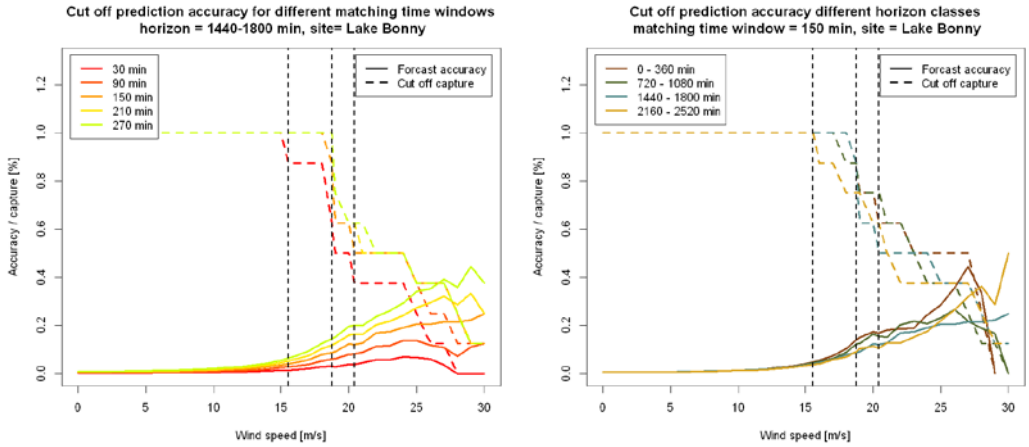


Figure 11: Accuracy evaluation of cut-off prediction based on wind speed thresholds determined in 6.1 for Slieve Rushen Phase 2. The forecast accuracy and cut-off capture is evaluated for different matching time windows (left) and prediction horizons (right).

As it was already assumed the results for the Irish farm in Figure 11 look very different and have a much lower accuracy as for Lake Bonny. Due to the very low number of only 8 events the results of the evaluation should be assessed carefully. The interval of high sensitivity of the two curves lies between 15 and 30 m/s. The forecast accuracy shows a strong dependency on the matching tolerance with a maximum of up to 0.4 for a tolerance of 270 min and values < 0.1 for a tolerance of 30 min. These low accuracy values indicate that there is no threshold for the predicted wind speed, which implies cut-off events with a high reliability. At the intersection of the forecast accuracy and the cut-off capture values around 0.4 are reached for a matching tolerance of >= 150 min. As well as for Lake Bonny the accuracy does not show a strong dependency on the prediction horizon.

The chosen quantiles lead to a cut-off capture around the expected values of 0.9, 0.7 and 0.5, but with only a very low forecast accuracy. Aiming at quantiles that are located at the intersection of each of the two curves, higher quantiles would have had to be chosen..

The big difference of the evaluation for the two sites raises the question about reasons for that. The most likely explanation for the low accuracy is the low number of detected events. As it was mentioned in section 5 it can be doubted that only 16 cut-off events have taken place within the 22,5 months. These events have therefore not been detected due to the low resolution of the measured data, but the high wind speed causing these events might still be represented within the prediction data.

3.6.3 Cut-off prediction based on “raw” grid point predictions

In this section the approach for cut-off prediction presented in the previous section is evaluated using data from the nearest grid point of the Lake Bonny wind farm. In 3.6.3.1 the data of the four grid points of the cell, in which the wind farm is located. The prediction based on the data of one of the grid points is evaluated in 3.6.3.2.

3.6.3.1 Analysis of grid points

To obtain wind speed predictions for a wind farm the mean value of the prediction for the grid points of the four corners of the grid cell, in which the wind farm is located, are calculated. This common approach has also been used to get the wind speed predictions used in the previous sections. The map in Figure 12 shows the location of the wind farm Lake Bonny and its nearest grid points. With a grid size of 0,5° the horizontal distance between the grid points is in this case ~ 20km and ~ 30km vertically. The map shows that the location of the wind farm is quite different of the mean location with a distance of 5-15km(?). It is even closer to the south west grid point then to the centre. As the wind farm is located near the shore the surface of the four grid points is very different: The south west grid point is located in the sea, whereas the other three are on shore. Therefore it is expected that the

prediction for the grid points differ and an equal weighted treatment might not be appropriate. The data for the following analysis covers the two year period from 01-2009 to 12-2010.

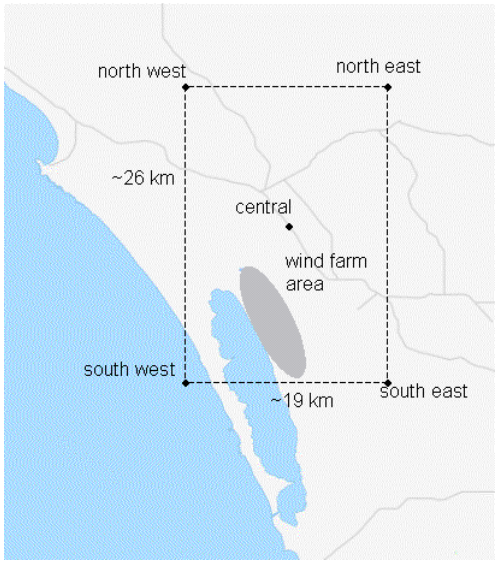


Figure 12: Location of wind farm Lake Bonny (grey area) and its closest grid points

In the following two figures the characteristics of the four grid points is compared. Figure 13 shows the distribution of the predicted wind speed and direction of the four grid points. Their deviation from the mean value is shown in Figure 14. The two wind direction plots on the left side show quite evenly distributed values with a dominance of the southerly directions. The differences from the mean are only minor with most of them staying below 10° and rarely exceeding 20°.

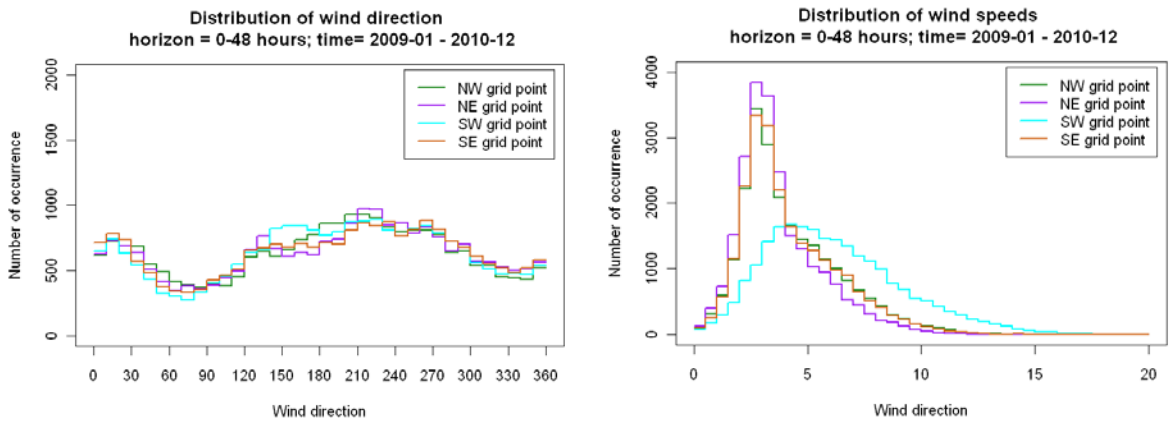


Figure 13: Distribution of predicted wind speed and direction for the four grid points closest to Lake Bonny wind farm.

But as it would have been expected the wind speed is increasing with more closeness to the sea due to the lower roughness over the water. The south west grid point (turquoise line) has got by far the highest wind speeds. The north east grid point (purple) has slightly lower wind speeds and the north west and south east grid point (green and brown) show a similar distribution. Differences of more than 4 m/s to the mean and therefore an even higher difference to the other grid points occur. But the predictions are not just biased. Figure (wspd distr.) shows different characteristics of the Weibull distributions. Especially the scale parameter is different for the four distributions.

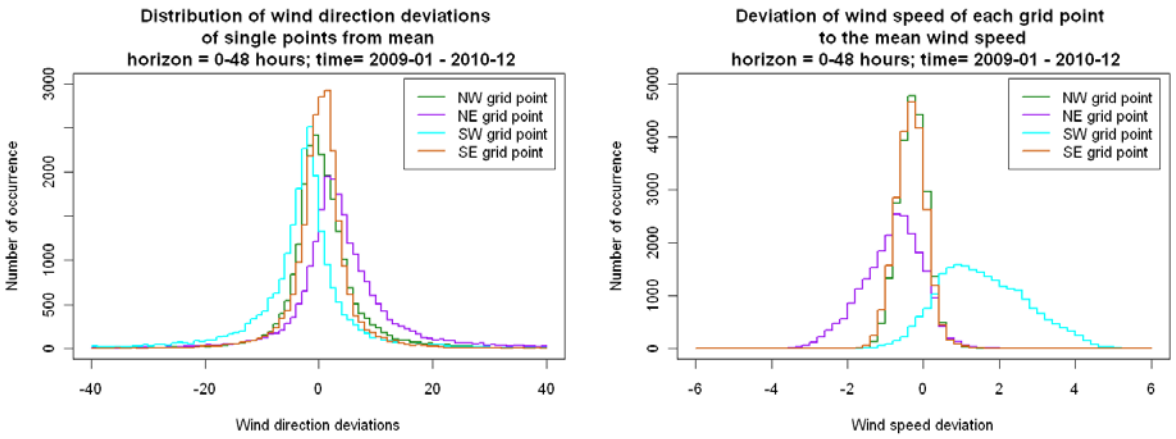


Figure 14: Deviations of the predicted wind speed and direction of each of the four grid points from their mean.

In Figure 15 the scatter of the wind speed predictions for each grid point and the averaged wind speed, the regression line and the correlation is shown. All four plots show a high correlation, but taking into account their distance of only about 25 km, the values for the south west and north east grid point show only a quite weak correlation with ~ 0.94 .

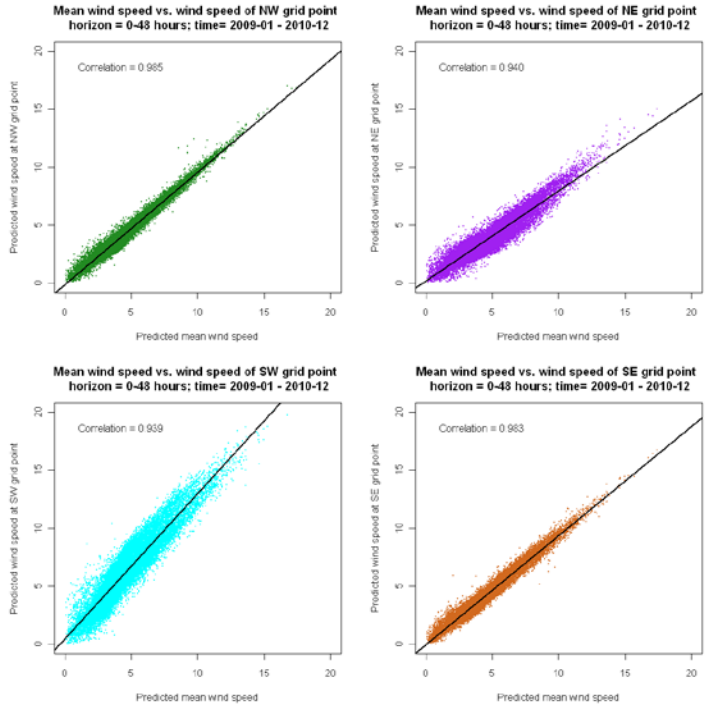


Figure 15: Correlation of the predicted wind speed for each of the four grid points with thei mean

The described differences of the wind speed predictions of the four grid points are levelled out by using the mean value. This approach has proven good results for wind power prediction, but can be a draw back for the prediction of extreme events like cut-offs as the characteristics of the single grid points can not used be utilised. The exploitation of this information might therefore improve the prediction of cut-off events.

This presumption is emphasised by the analysis of Figure 16. It compares the ratio of the wind direction for the whole time series (black line) with the wind direction during all detected cut-off events +/- 1 hour (multiple colours). It shows that for the whole time frame the wind direction is quite evenly distributed with a slightly higher share of the southerly winds. For the cut-off events with ~ 80% south

west (180 – 270°) is clearly the dominant wind direction. Therefore it can be expected that during most cut-off events the wind at the wind farm shows similar characteristics as the south west grid point. The utilisation of predictions for this grid point for cut-off predictions might give better results than the average of all grid points. This will be evaluated in the next section.

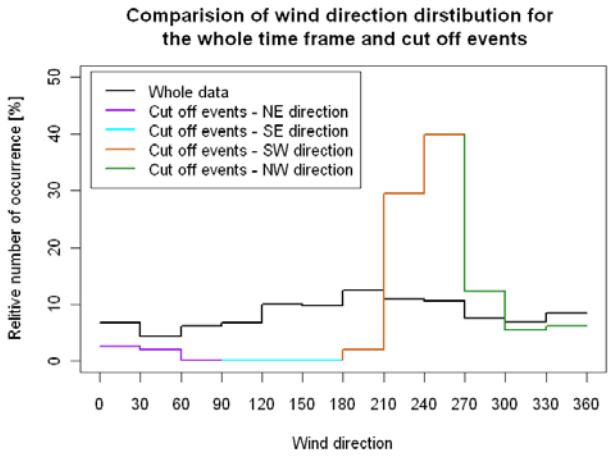


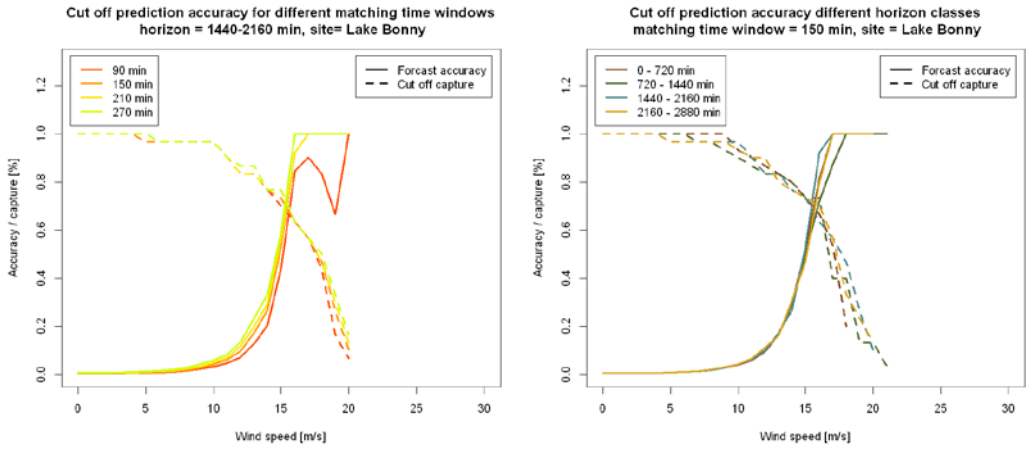
Figure 16: Comparison of the relative distribution of the predicted wind direction for the whole time frame and during cut-off events

3.6.3.2 Cut-off prediction based on “raw” grid point predictions

In this section the use of the wind speed predictions of the south west grid point for cut-off prediction is evaluated. As the grid data is only available from 08-2008 – 12-2010 not all evaluation periods of the previous section are covered. To allow a comparison the time frame of the evaluation periods of II-LKB and III-LKB used for the evaluation without a tuning period.

The following four graphs show the accuracy results for the time frame II-LKB with 30 (Figure 17 top row) and time frame III-LKB with 23 cut-off events (Figure 17 bottom row). They are similar to the evaluation of the mean wind speed in Figure 17. Due to the in average higher wind speed the wind speeds with the highest sensitivity are slightly shifted to the right for the south west grid point. In this evaluation also neither the size of the matching time window nor the prediction horizon show a big influence on the prediction accuracy. Nevertheless with the intersection of the curves at about 70% higher accuracy is achieved as for the mean value.

II-LKB:



III-LKB:

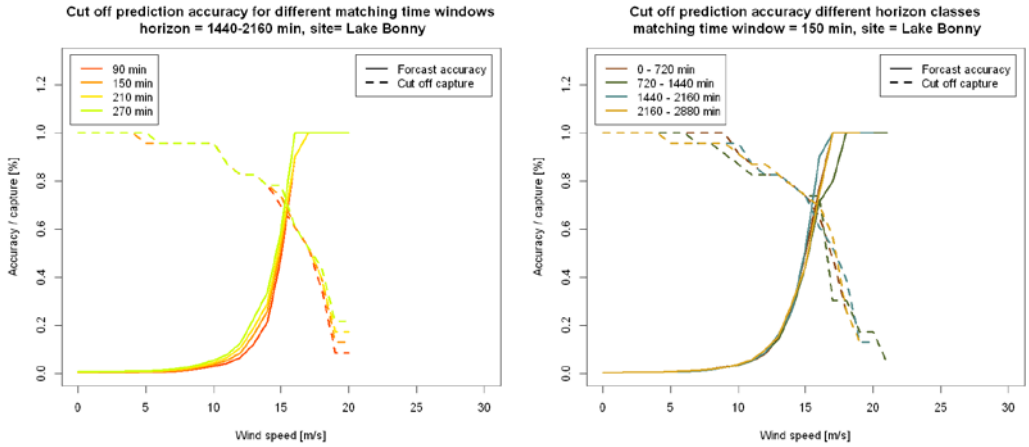


Figure 17: Accuracy evaluation of cut-off prediction based on various wind speed thresholds for Lake Bonny. The forecast accuracy and cut-off capture is evaluated for different matching time windows (left) and prediction horizons (right).

3.7 Conclusion on cut-off prediction

A methodology for cut-off prediction based on the predicted wind speed has been presented and evaluated. To determine an appropriate threshold value the predicted wind speed for detected cut-off events has been analysed. The use of the 10, 30 and 50 % quantile has been analysed, in which the 50% quantile has shown almost optimal results for Lake Bonny. The choice of the threshold is very sensitive to the prediction accuracy.

For the evaluation of the wind speed threshold the cut-off capture and the forecast accuracy have been analysed for different data sets. The evaluation has shown that neither the size of the matching time window nor the prediction horizon has a big influence on the prediction accuracy. For Lake Bonny wind farm for wind speeds of 13 m/s more than 60% for the cut-off capture and the forecast accuracy have been achieved.

For the wind farm Slieve Rushen Phase 2 the results were not promising. It is supposed that many cut-off events of the site have not been detected due to the low resolution of the input data and frequent down regulation.

The analysis of single grid points has shown that by averaging the wind speed predictions of the four corner points of the grid cell characteristics of these predictions are levelled out. Almost all the prediction for the south west grid point are higher than the mean with deviations of up to 4 m/s at 10 meters height. An analysis of the wind direction during cut-off events showed a domination of the south west direction with ~80% of the events, which is not reflected in the general distribution. Using the south west grid point for cut-off prediction showed an improvement of ~10% in comparison to the prediction based on the mean wind speed. Nevertheless this effect is only expected if the four grid points show different characteristics. This especially applies to wind farms located near the shore due to the differences of the roughness length between sea and land.

3.8 Results and Conclusion

In this evaluation a detection and prediction approach for cut-off has been presented.

For the detection of cut-off events two filter approaches utilising the power change, wind speed and their correlation have been presented and evaluated. The accuracy of these approaches is much dependent on the resolution of the data. For one minute resolution data very reliable identification of cut-off events was possible, whereas the accuracy decreases significantly for 15 minutes data.

Furthermore a cut-off prediction approach based on the predicted wind speed has been presented. Due to low data quality and only few detected events, the evaluation has only been done for two sites. The results for one site still had to be assessed very carefully. An accuracy of more than 60% of cut-off capture and forecast accuracy has been achieved for a prediction horizon of 24-36 hours. The tolerance of the predicted timing of the events for these results have been 90 minutes and six events were used for tuning. The evaluation of different tolerances for the timing of the predicted cut-off events as well as the prediction horizon did not show a big impact. An extension of the training period did not lead to significant improvements of the accuracy either.

These prediction results have been improved by using the wind speed predictions of the nearest grid point as predictor instead of the mean. This method is probably only promising if the four grid points for the calculation of the mean show different characteristics.

To improve the cut-off detection approach an evaluation of different methods for the determination of the filter parameter or adding adaptation seems promising. The evaluation of other different variables as predictor for cut-off events like the gust factor might improve the accuracy as well as the application of other prediction approaches, e.g. from the data mining domain. As a comparison of the impact of different temporal and spatial resolutions of numerical weather predictions has not been possible, this could also be part of future research.