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 Dc-6.10
"Overall assessment of WP-6 contributions"

DOCUMENT TYPE	Deliverable
DOCUMENT NAME:	swind.deliverable_Dc-6.10.pdf
VERSION:	V2.0
DATE:	2012.09.15
CLASSIFICATION:	R0: General public
STATUS:	Approved

Abstract: This work package of the SafeWind project concentrated on predicting and facing extremes situations in the short to medium term (up to 2-3 days ahead) with purely statistical and probabilistic methods accounting for the evolution of meteorological variables, possibly indicating different weather regimes. A large component here related to the question of estimating and communicating prediction uncertainty, as it is a crucial aspect of wind power forecasting and for warning forecast users about potential forthcoming extremes. This report gives an overview of the work performed throughout the project, giving some of the main methodological elements, showing sample evaluation results, while giving appropriate links to more extensive descriptions published in the form of deliverable project reports and of peer-reviewed articles in international scientific journals.

AUTHORS[1], REVIEWERS				
MAIN AUTHOR/EDITOR:	P. Pinson			
AFFILIATION:	Technical University of Denmark, DTU Informatics			
ADDRESS:	Asmussens Allee 305, 2800 Kgs Lyngby, Denmark			
TEL.:	+45 4525 3428			
EMAIL:	pp@imm.dtu.dk			
FURTHER AUTHORS:	Miguel Bermejo (UC3M), Lueder von Bremen (Oldenburg), Robin Girard (Mines ParisTech), Jouyoung Jeon (Oxford), Patrick McSharry (Oxford), George Sideratos (NTUA)			
PEER REVIEWERS:	Robin Girard, Georges Kariniotakis (Mines ParisTech)			
REVIEW APPROVAL:	Approved :	X	Rejected (improve as indicated below) :	
SUGGESTED IMPROVEMENTS:	*For a long list of remarks make reference to another document*			
APPROVER:	Pierre Pinson (DTU)			

STATUS, CONFIDENTIALITY, ACCESSIBILITY										
STATUS:				CONFIDENTIALITY:				ACCESSIBILITY:		
S0	Approved/Released	**X**		**R0**	General public	**X**		Private web site	**X**
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 comment**s**			**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.**

Contents

1 Introduction

Massive deployment of renewable energy sources, with the leading role of wind energy historically, calls for the development of new integrated approaches for the optimal management of power systems operations, most often in a market environment. This in turn requires the use of forecasts as input to monitoring and decision-making, with forecasts that may be of different forms, for different lead times and with varied space-time resolution (Jones 2011). Forecasts are especially valuable in a market environment like those existing in Europe or in the US (Botterud *et al.* 2010) among others. An extensive and recent review of the state of the art in wind power forecasting was compiled in Giebel *et al.* (2011).

In the frame of the EU-FP7 project SafeWind, the role and main emphasis of WP-6 was to develop and assess new methodologies for forecasting wind power generation using statistical concepts, which would answer the new varied needs of forecast users with focus on extreme events. Extreme events were indeed identified as comprising one of the most pressing challenges when it comes to an optimal integration of wind energy into existing power systems and electricity markets. They translate to severe threats to systems adequacy and operations safety, while yielding extreme costs for the actors of the power systems and electricity markets: power producers may have to deal with additional costs linked to additional maintenance and repair or market penalties, while system operators have to guarantee a proper system operation at all times. These so-called *extreme events* are perceived in a fairly different manner depending on the type of forecasts user. In parallel the meaning of “extreme” in the case of the wind power application is seen differently by meteorologists, forecasters, power systems engineers, etc. (Pinson *et al.* 2010).

The original objectives of this work package related to lead times in the early-medium range (up to 2-3 days) while works in WP-4 and WP-5 were somewhat more focused on the short-term (a few hours ahead) and medium range (more than 2 days ahead), respectively. A number of different routes were taken for the proposal of forecasting approaches relevant for different types of extreme events. Since extreme events ought be rare by nature, and their potential occurrence uncertain at the time of issuing predictions, a particular emphasis was given to probabilistic approaches. These may provide forecast users with probabilities of occurrence of pre-defined *categorical* events e.g. cut-offs and high-variability periods, or alternatively with a full *continuous* description of all potential levels of power generation with their likelihood. The general methodological aspects related to probabilistic forecasting are reviewed in Section 2. This will include probabilistic forecasting in parametric and nonparametric setups in Section 2. Subsequently, regime-switching modelling and forecast combination will be discussed in Section 3. The proposal of an event-based view of wind power forecasting is introduced in in Section 4, with events being e.g. cut-offs, ramps and high-variability periods. Finally, the specifics and importance of the communication of forecast information (mainly uncertainty) will be further developed in Section 5. The report ends with conclusions in Section 6.

2 Probabilistic forecasting of wind power generation

By probabilistic forecasting is meant here the prediction of the power generation (as a continuous variable) from a wind farm or wind power portfolio for a lead time of interest. Other forms of probabilistic

forecasting will be discussed in further parts of this report, for instance for the case of events.

The work performed included a part of methodological developments before its application to the various test cases considered. Some of these developments generally contributed to advancing the science of forecasting while some others are mainly driven by the wind energy application specifically. It should be possible to transpose them in a fairly straightforward manner to the forecasting of other renewable energy sources e.g. solar and wave energy. These methodological developments were mainly directed towards the proposal of new parametric distributions, new approaches to quantile modeling and forecasting, as well as forecast verification.

2.1 Parametric vs. nonparametric approaches to probabilistic forecasting

Probabilistic forecasts are there to inform about ranges of potential wind power generation with given probabilities, hence giving the possibility to extract single-valued forecasts that are commonly used by forecast users as input to decision-making, while providing a valuable information about the uncertainty of such forecasts. They may take various forms, i.e. quantiles, intervals and full predictive densities, depending on the forecasting methodology employed and the optimal forecast products that is needed. In the most general case, focus is given to predictive densities, which are discussed further in the following. They may be generated in a parametric or nonparametric framework, while using various types of data and information as input. By writing y_t the power production for a wind farm at time t and k the lead time, a predictive density issued at time t for time $t+k$ is denoted by $\hat{f}_{t+k|t}$, with $\hat{F}_{t+k|t}$ the corresponding cumulative distribution function. The forecaster at time t states that

$$y_{t+k} \sim \hat{F}_{t+k|t} \tag{1}$$

where for simplicity y_{t+k} also denotes a random variable at time $t+k$. The above equation simply tells that power generation is a random variable that is fully characterized by the probabilistic forecast $\hat{F}_{t+k|t}$.

2.1.1 Nonparametric approaches

In a nonparametric setup, no assumption is made regarding the shape of $\hat{F}_{t+k|t}$. Therefore, in order to fully described that distribution, it is necessary to summarize it with a set of quantiles,

$$\hat{F}_{t+k|t} = \{\hat{q}^{(\alpha_i)}_{t+k|t} \,;\, 0 \leq \alpha_1 < \ldots < \alpha_i < \ldots < \alpha_m \leq 1\}, \quad i = 1, \ldots, m, \tag{2}$$

with chosen nominal proportions spread over the unit interval. A model is then to be proposed for each of the m quantiles permitting to define nonparametric distribution. The fact that different models may be setup for each and every quantiles may render the task cumbersome and computationally expensive. Generally, for a good representation of predictive densities, it is accepted that 21 quantiles may be necessary, with nominal proportions $\alpha_i \in \{0, 0.05, 0.1, \ldots, 0.9, 0.95, 1\}$. In a normalized setup where power generation is within [0,1], the extreme quantiles are naturally given by these bounds since it would be very difficult to restrict the absolute range of potential power generation with complete certainty. In the case where power values are not normalized, power may then range between 0 and the nominal capacity of the wind farm or portfolio of interest.

Figure 1 depicts an example set of nonparametric probabilistic forecasts issued on the 3.4.2007 at 16:00 for 5 aggregated zones over Western Denmark (Energinet.dk dataset) for a total capacity of 2.5 GW. Roughly, zone 1 is of the North tip of the Jutland peninsula, zone 2 for the Western part, zone 3 the Eastern one, zone 4 for the South-West corner finally zone 5 for the South-East corner and the island of Funen. Corresponding measurements are also shown. The predictive densities are represented by 20 quantile forecasts since the median is omitted. Both forecasts and measurements are normalized by the nominal capacity in each of the zones, so as to take values in [0,1].

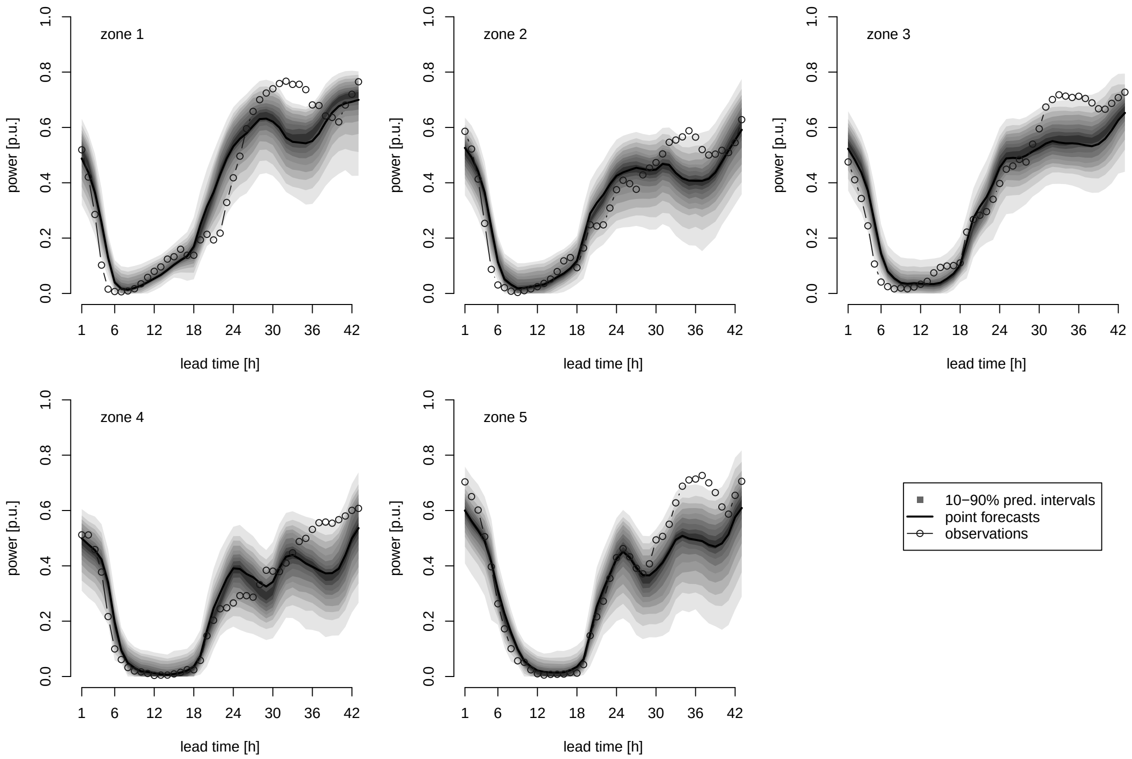

FIGURE 1: *Nonparametric probabilistic forecasts of wind power generation issued on the 3.4.2007 at 16:00 for 5 aggregated zones defined over Western Denmark, represented as a river-of-blood fan chart. Predictive densities are defined by a set of quantile forecasts with nominal proportions* $\alpha \in \{0.05, 0.1, \ldots, 0.45, 0.55, \ldots, 0.9, 0.95\}$*, yielding prediction intervals with nominal coverage rates between 10 and 90%. These forecasts are normalized and hence take value in [0,1].*

The flexibility of nonparametric approaches clearly is appealing for a nonlinear and complex process like wind power generation. A number of approaches existed before the beginning of the SafeWind project, like the quantile regression approach of Bremnes (2006), the time-adaptive quantile regression of Møller *et al.* (2008), the adaptive resampling approach of Pinson and Kariniotakis (2010) and the ensemble-based method of Pinson *et al.* (2009). However, various investigations have shown that the estimation and predictions of quantiles with nominal proportions close to the bounds of the unit interval (say, 0.01 and 0.99 for instance) could lack robustness. This may hence be problematic when being interested in extreme events.

Novel approaches to nonparametric forecasting of wind power generation were developed in the frame of the SafeWind project. As an example, the work of McSharry and Anastasiades (2011) investigated the possibility of improving the prediction of quantiles, which form the basis of predictive densities,

using a new type of information in the form of variability indices. These variability indices are there as explanatory variables to inform about the recent volatility of wind power time-series, therefore potentially resolving among situations with various levels of uncertainty. It is foreseen that adding such type of explanatory variables to quantile forecasting models could help improving robustness of quantile forecasts with nominal proportions close to the bounds of the unit interval. In a similar vain, Sideratos and Hatziargyriou (2012b) proposed a new artificial-intelligence based approach that can additionally account for weather uncertainty information as quantified by the standard deviation of wind speed and direction forecasts used as input, in the area around a wind farm of interest. Application to the case of Greek wind farms demonstrated the interest of the proposal.

In parallel, Jeon and Taylor (2012) introduced a new type of power curve modelled with conditional kernel density estimation, allowing to convert probabilistic forecasts of wind into probabilistic forecasts of power, by account for uncertainties in both input wind and observed power data. An example of such power curve for the test case of the Aeolos wind farm in Greece in shown in Figure 2.

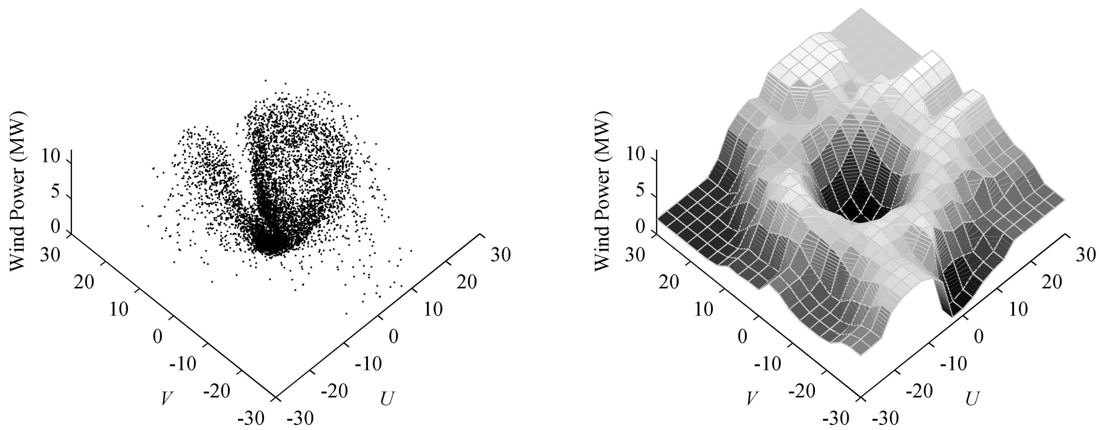

FIGURE 2: *Power curve as wind speed and direction versus power generation (left) and a smooth surface fitted using a Nadaraya-Watson estimator (right), for the Aeolos wind farm. (figure taken from Jeon and Taylor (2012))*

In all cases, the authors showed non-negligible improvements from their innovative proposals if compared with the existing. More precisely, considering variability indices as in McSharry and Anastasiades (2011) and weather predictability as in Sideratos and Hatziargyriou (2012b) allowed improving sharpness and resolution of the probabilistic forecasts, that is, their inherent ability to inform about situations with various level of forecast uncertainty. And, Jeon and Taylor (2012) showed that forecasts issued based on their proposal bivariate wind forecasts (i.e. , its u and v components) and power curve using conditional kernel density estimation, had a better calibration than benchmark approaches, where calibration is seen as the probabilistic correctness of the forecasts. More work is to be done in order to improve the robustness and skill of quantile forecasts for extreme events, but generally, it may be that if one is interested in extreme and rare events, parametric approaches should be preferred anyway.

2.1.2 Parametric approaches

A known drawback of nonparametric approaches is that they may not be very robust in the tails of distributions, which is the part representing extreme events in terms of forecast errors. Therefore, it may

be of particular relevance to employ parametric approaches instead. The interest of these is to rely on parametric assumptions, which can be seen as pre-defined shapes, for predictive densities. Based on a chosen assumption, the distribution of y_{t+k} as predicted at time t is

$$y_{t+k} \sim F(y; \hat{\boldsymbol{\theta}}_{t+k}), \tag{3}$$

where F is the distribution of choice. The set of parameters $\hat{\boldsymbol{\theta}}_{t+k} = \left[\hat{\theta}_{1,t+k}, \hat{\theta}_{2,t+k}, \ldots, \hat{\theta}_{m,t+k}\right]^\top$ fully determines the predictive density. The choice for a distribution or the other, e.g. Gaussian, censored Gaussian, etc., provides different information on the range of potential outcomes and their probability. Note that the Gaussian one (and its truncated and censored versions) is the most employed in practice. However, it may be that such distributions do not appropriately describe the asymmetry of forecast uncertainty close to the power generation bounds, or the increased peakedness of the distributions. Other possibilities were therefore investigated during the project.

Example works in that direction include that of Bermejo and Sánchez (2012) who explored the possibility of modelling conditional moments of predictive densities, to then be used in order to issue predictive distributions based on parametric assumptions (e.g. Gaussian and Beta) and based on information-theoretical concepts, following the maximum entropy principle. An example result for a Greek wind farm in shown in Figure 3, where the evolution of moments as a function of the level of power predicted by point forecasts (used as input here to issue predictive densities) follow some pattern already empirically observed in previous work. The interest of the moment-based method is that it provides the flexibility to fit and issue different types of predictive densities based on fairly simple models for these moments.

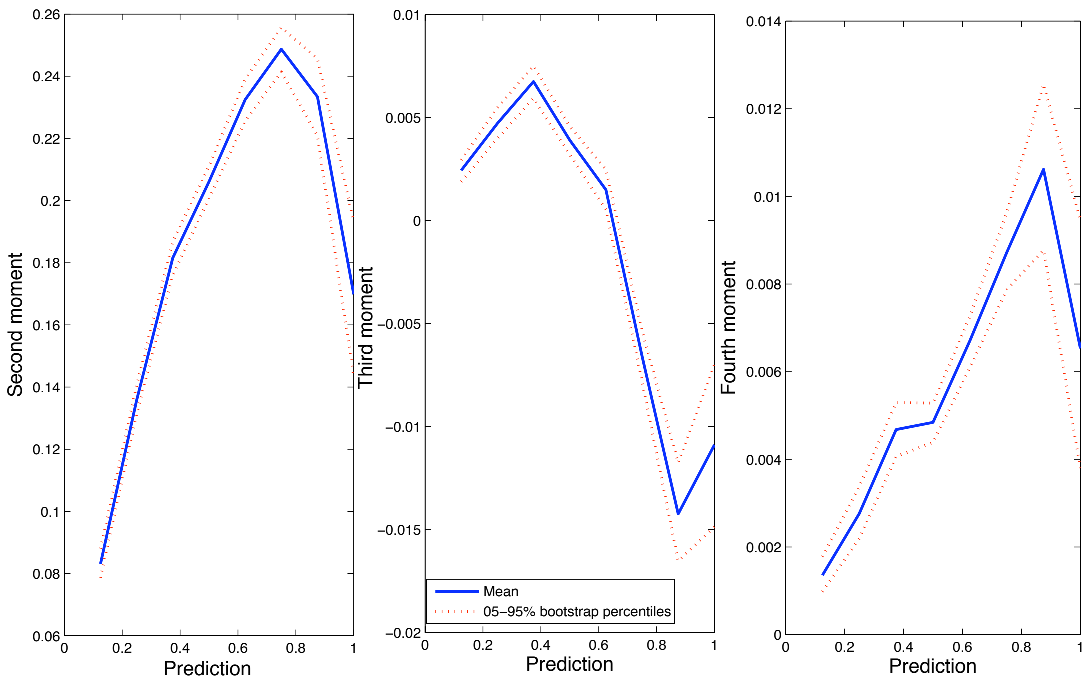

FIGURE 3: *Modelling of conditional moments following the method of Bermejo and Sánchez (2012) as a basis for issuing predictive densities of wind power generation following parametric assumptions or information-theoretical concepts. The moments are here conditional to the level of power predicted by input point forecasts.*

In parallel, Pinson (2012) proposed a new type of predictive density using Generalized Logit-Normal

(GL-Normal) distributions for the modelling and forecasting of short-term fluctuations, i.e., for lead times of a few minutes to a few hours ahead. It is expected that such distributions may be generically employed for a wider range of lead times since permitting to respect some of the inherent characteristics of wind power forecast uncertainty stemming from its nonlinear and double-bounded nature. Such characteristics may for instance be observed in Figure 3, where variance, skewness and kurtosis (related to the moments of order 2, 3 and 4) exhibit a typical pattern of the effect of the power curve on forecast uncertainty. That is, the uncertainty is increased for mid-level power generation, with a noticeable asymmetry (moment of order 2), while the skewness of distributions goes from positive to negative as power increases (moment of order 3).

The proposal of Pinson (2012) was evaluated on the test case of an offshore wind farm for a period of several months in 2005-2006, by issuing probabilistic forecasts with lead times of 10 minutes. The forecasting results in terms of predictive skill (quantified with the Continuous Ranked Probability Score - CRPS) are collated in Table 1, and compared to those for other types of predictive densities, namely Normal and Beta, plus two benchmark approaches that are climatology and persistence. Similar dynamical models for the parameters of predictive densities were employed in all cases so that it is mainly the shape of densities that is evaluated. It is clear that GL-Normal predictive densities allow for better probabilistic forecasts. This can be explained by sharper forecasts close to the power generation bounds, but also to a better probabilistic calibration of these distributions.

TABLE 1: *Monthly results for the evaluation of density forecasts with a CRPS criterion (in % of the nominal capacity P_n). Best score values are highlighted using bold fonts.*

Month	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	All
Climatology	18.67	19.75	20.07	18.61	19.79	24.65	21.81	21.63	20.47
Persistence	2.30	2.06	2.03	2.06	1.96	2.35	2.29	2.02	2.12
Normal AR	1.85	2.01	1.99	1.98	1.87	2.31	2.23	1.97	2.01
Beta AR	1.81	1.91	1.96	1.93	1.78	2.08	2.08	1.90	1.92
GL-Normal AR	1.76	1.80	1.86	1.86	1.81	**1.93**	1.92	2.03	1.86
GL-Normal CP-AR	**1.73**	**1.78**	**1.82**	**1.81**	**1.71**	1.94	**1.88**	**1.79**	**1.80**

On the parametric side overall, the introduction of new approaches to the modelling of conditional moments, comparison of various predictive densities, as well as proposal of new probability distributions, allowed for a number of interesting advances. While a few years ago emphasis was mainly placed on the development of nonparametric approaches for wind power probabilistic forecasting, it was shown during the project that parametric (and semi-parametric) approaches could actually be advantageous in terms of simplicity of modelling and better robustness when it comes to the tails of predictive densities, which are the part of densities that cover the extreme events.

2.2 Forecast verification

Assessing forecast quality is of utmost importance for appraising the relevance of models and forecasting approaches developed throughout the project period. The verification protocols proposed by Madsen *et al.* (2005) and Pinson *et al.* (2007) for the case of deterministic and probabilistic predictions, respectively, were used as a basis. However, in view of the proposal of new forecasting products (event-based

predictions), while placing more focus on probabilistic forecasts and the tails of predictive densities, it was chosen to review recent advances in forecast verification in statistics and economics e.g. Clements (2005), as well as in meteorology and climate e.g. Stephenson and Jolliffe (2012).

It would be too lengthy to get into detail with aspects of forecast verification here, while it was not the primary focus of this work package of the SafeWind project. A updated overview of relevant approaches to forecast verification when it comes to wind energy and to the types of forecasting products considered was compiled in McSharry *et al.* (2009). More targeted work was carried out on the assessment of the probabilistic correctness (i.e., their calibration) of probabilistic forecasts, by focusing on the potential dramatic effect of serial correlation on reliability assessment, which should be taken into account (Pinson *et al.* 2010). This was demonstrated based on the test case of the Danish Energinet.dk dataset for all onshore capacities of Western Denmark. Approaches permitting to cope with this correlation issue and also permitting to visualize the potential uncertainty in reliability assessment owing to both sampling and correlation effects were then introduced and validated.

In parallel, emphasis was placed on the verification of some of the new type of forecasting products considered in the SafeWind project, that is, scenarios of short-term wind power generation. Figure 4 depicts an example set of space-time scenarios of short-term wind power generation for the same episode as in Figure 1, i.e. issued on the 3.4.2007 at 16:00 for the 5 aggregated zones. Corresponding measurements are also shown. The number of scenarios displayed is of 12. By construction, these scenarios respect the marginal predictive densities shown in Figure 1 while adding information on the expected spatio-temporal dependencies. Similarly if generated a sufficiently high number of scenarios, the average of these scenarios would be equivalent to the point forecasts commonly used by most forecast users.

For the verification of scenarios of short-term wind power generation, the proposal made by Pinson and Girard (2012) in the frame of this work package was to generalize the existing framework for the verification of univariate probabilistic forecasts to the case of multivariate ones, based on recent methodological proposals available in the literature. This extended framework includes a number of diagnostic tools and scores that may allow discriminating among rival approaches to scenario generation, based on their forecast quality. This verification framework was illustrated and demonstrated based on test cases with data form French wind farms, as provided by Électricité de France (EDF). Interestingly, the work in Pinson and Girard (2012) hinted at the fact that defining events (like ramps for instance) based on such scenarios may actually ease the verification process since reducing dimensionality.

3 Regime-based modelling and forecast combination

At the beginning of the project work, it was observed and concluded that most wind power forecasting models aimed at producing the best forecasts for all types of forecast conditions, in the sense of minimizing a quadratic (Root Mean Square Error - RMSE) or linear (Mean Absolute Error - MAE) error criterion, without any specific consideration for extreme events. Such events may actually occur in certain types of weather and power production regimes. It may therefore be beneficial to apply regime-based approaches to the modelling and forecasting of wind power generation, where the forecasts in various regimes may be combined eventually. Different strategies have been explored, based on local measurements only, using both local and offsite measurements, as well as based on artificial intelligence techniques with

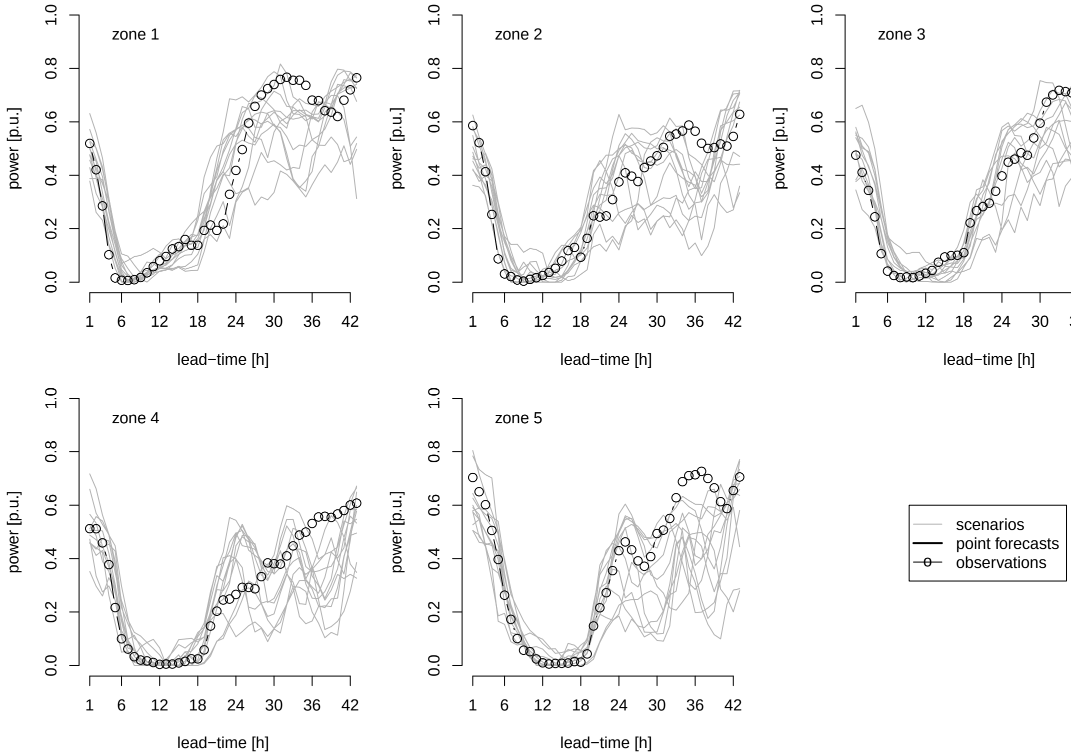

FIGURE 4: *Spatio-temporal scenarios of wind power generation issued on the 3.4.2007 at 16:00 for 5 aggregated zones over Western Denmark.*

measurement and meteorological forecasts as input.

3.1 Time-series approaches based on local historical measurements

Time-series approaches based on local historical measurements are better suited to cases where lead times up to few hours ahead (say, 6 to 8) are of most interest. This forecast range is relevant for participation in the intra-day energy market and for the management of operations at wind farms.

Emphasis was placed on a class of regime-switching models, the Self-Exciting Threshold Autoregressive Models (SETAR), in this part of the work. SETAR models have the benefit of simplicity since allowing using linear models in each of the regimes. Also they are more easily interpretable since the changes among the different regimes are governed by an observed variable e.g. past power values, the so-called threshold variable. A common SETAR structure was proposed to obtain better forecasts than if using classical univariate time-series models. The identification of the structure of SETAR models is one of the major problems with their use today, due to the complexity of efficiently estimating thresholds, that is, the points in which the process changes among regimes.

In the present case the proposal of Bermejo *et al.* (2011) has been to identify the SETAR structure by using a novel automatic procedure that was called Aut-ARLS. This procedure is based on the recursive and time-varying estimation of the parameters. It only needs to define a set of possible threshold variables as input. When applied to the case of wind power prediction, the existence of different regimes when

wind power generation was increasing or decreasing was investigated and shown. This means that the dynamics of wind power generation are somewhat different when wind power output tends to go up or down. It could be potentially explained by the physical characteristics of wind turbines, since this type of behaviour was never observed for wind itself. Subsequently, the existence of different regimes depending on the levels of wind power is studied, with the regime sequence being determined by the previous wind power measurements. Known results are confirmed there, which are that 3 different regimes should be considered, for high, medium and low levels of wind power. These three regimes were obtained for both cases of power generation increasing and decreasing.

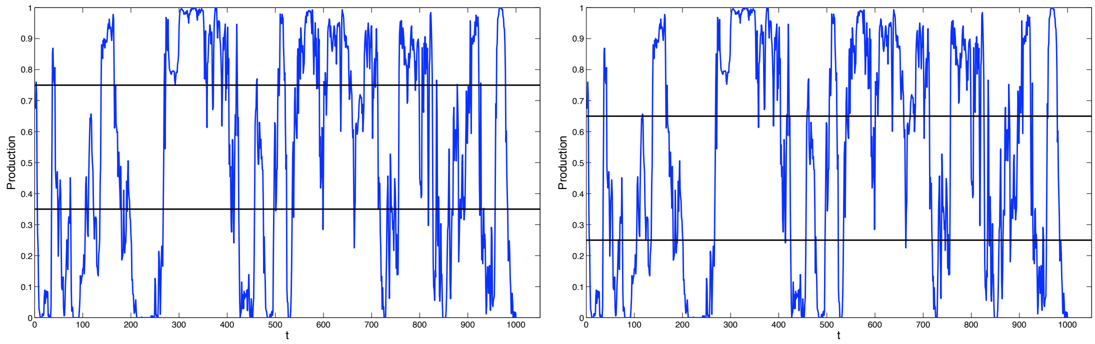

FIGURE 5: *SETAR structure identified when the threshold value is the first lag of the process, and when the wind power is increasing (a) or decreasing (b).*

Once the SETAR structure has been decided upon, forecasts are issued based on Monte Carlo simulations of future regimes. Subsequently, different trajectories are generated, whose marginal distributions follow some estimated predictive density (in the parametric or nonparametric framework discussed in the above). Different parametric distributions are considered: Truncated Normal (TN), Censored Normal (CN), Beta and Maximum Entropy Principle (MEP), following the work of Bermejo and Sánchez (2012). The estimation of the sequence of predictive densities is based on the method of conditional moments mentioned in Section 2.1.2 with input the point forecasts already employed by forecast users.

This SETAR model with conditional moments is referred to as SETAR-TVCD. Forecasts were generated and compared for a set of competing dynamical models and strategies for generating predictive densities:

ARMA: The best linear model is chosen by using AIC,

TV-AR: Several AR(1) model with time-varying parameters are estimated and the best model is chosen,

SETAR: The classical SETAR models using sample moments,

SETAR-TVCD: Conditional moments are used.

This was carried out based on the test cases of Greek wind farms with several years of data. An important conclusion of that work was that the same SETAR structure was found for all available Greek wind farms. This can be seen from the box-plot of the estimated threshold values for all the analyzed wind farms in Figure 6.

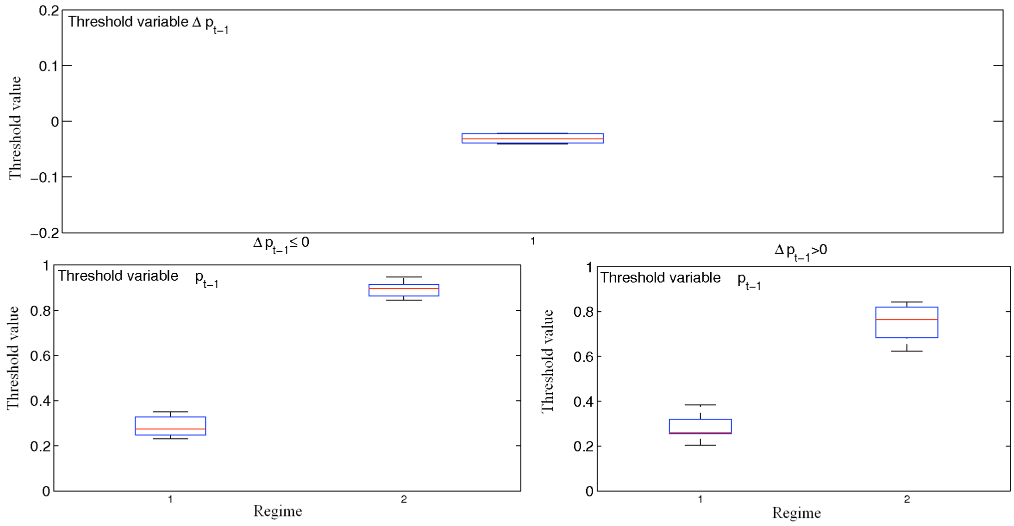

FIGURE 6: *Box-plot of the estimated threshold values for all available wind farms.*

Finally, Table 2 gathers all the results from the empirical analysis work, based on a RMSE criterion for forecast evaluation. It confirms that employing SETAR-TVCD dynamics with MEP distributions may be the best approach to take.

TABLE 2: *Summary of the results for all available wind farms based on a RMSE criterion.*

Mean of the Relative Root Mean Square Error in 6 wind farms										
			SETAR				SETAR-TVCD			
h	ARMA	TV-AR	TN	CN	BETA	MEP	TN	CN	BETA	MEP
1	1,18	1,08	1	1	1	1	1	1	1	1
2	1,26	1,12	1,18	1,18	1,19	1,09	1,15	1,11	1,09	1
3	1,42	1,17	1,29	1,22	1,26	1,10	1,21	1,17	1,13	1
4	1,65	1,22	1,35	1,33	1,34	1,11	1,26	1,21	1,16	1
5	1,87	1,25	1,57	1,44	1,45	1,12	1,35	1,27	1,26	1
6	1,96	1,28	1,66	1,56	1,50	1,17	1,41	1,30	1,32	1

3.2 Time-series approaches also using offsite measurements

With the large scale deployment of wind power capacities, it is often the case that wind farms are not that far from each other, with distances typically in the order of tens of kilometers. At such a spatial scale, information from neighboring wind farms in the form of current and past observations (meteorological, power) may be beneficial in order to improve short-term forecasts for lead times in the range of minutes to a few hours (Tastu *et al.* 2011). Especially in the context of regime switching and forecast combination, early detection of a change of regime at a neighboring wind farm would directly translate to a better prediction of a change of regime at a wind farm of interest.

In order to analyse this problem, focus was given to the specific test case of 3 Irish wind farms represented in Figure 7, at the sites of Carnsore, Richfield and Ballywater, where time-series of power measurements

were available with a resolution of 15 minutes. Forecasts were generated for a lead time of 15 minutes. This was for the sake of example only, and further lead times could be similarly considered. The potential forecast improvements stemming from considering offsite information are expected to decrease with lead times though. An overview of this modelling work and analysis is available in Trombe *et al.* (2012).

As a base example, Figure 8 shows the parallel evolution of normalized wind power generation at these 3 wind farms over a period of 3-4 days. It shows that Carnsore and Richfield appear to have very similar wind power dynamics and to be fairly in phase (over that period at least), while power generation at Ballywater, even though looking very different from the others, seem to give some advance information about trends in the evolution of power generation. For instance, power production is picking up a few hours before at the Ballywater wind farm on the 3.10.2008, then followed by the other two. Their geographical proximity and the fact that these wind farms witness the same wind regimes certainly explain such observations. As a consequence, it was investigated whether or not one could take advantage of this set-up for better modelling and predicting regime changes in power production, in a probabilistic forecasting framework.

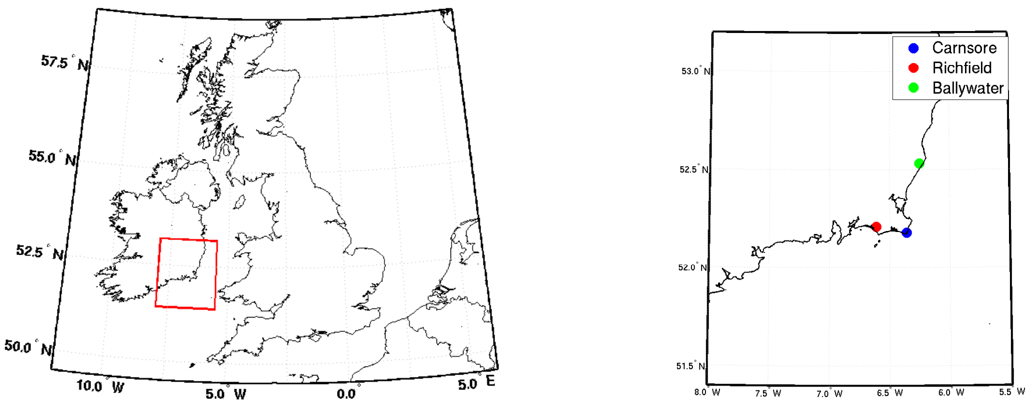

FIGURE 7: *The Carnsore, Richfield and Ballywater wind farms are located in the South-East of Ireland. Carnsore and Richfield are separated by an approximate distance of 17 km, and the distance between Richfield and Ballywater is of 40 km.*

For that purpose, Threshold AutoRegressive (TAR) models were first considered as a basis. TAR models comprise the general family of piecewise linear regime-switching models for which regime changes are governed by an observable regime sequence. SETAR models, as considered in the above, are for the specific case where this regime sequence is given by the past observed value of the time series of interest, i.e. , the last power measurement before a forecast is to be issued. In the present case TAR models were set up by considering that it is information at offsite locations that should govern the regime changes. Alternatively, TARX models were used, for which the regime sequence is determined based on local information only, but where the dynamical models in each regime use offsite measurements as explanatory variables. AR and ARX univariate models were also identified and estimated so as to provide a basis for comparison.

Since existing empirical analysis work looking at data with similar temporal resolution showed that observable regime sequences may not be optimal, alternative approaches based on unobservable regime

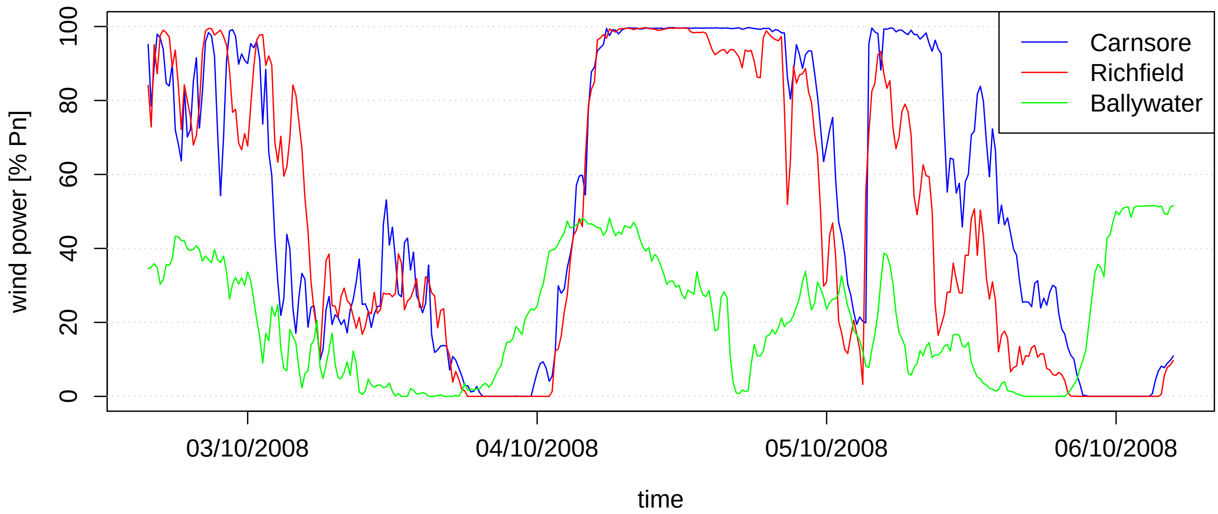

FIGURE 8: *Normalized wind power generation at Carnsore, Richfield and Ballywater.*

sequences were also considered. This translated to identifying and estimating Markov-Switching AutoRegressive (MSAR) models. Similarly to the case of TAR models, MSARX models were also used, so as to directly incorporate offsite measurements within the dynamical models.

3.3 Artificial intelligence based techniques with measurements and meteorological forecasts as input

A new regime switching method based on artificial intelligence was proposed to produce better wind power forecasts in the case of extreme events, compared to the more common methods of the state of the art that consider all types of forecast conditions indifferently. The methodology was extensively described in Sideratos and Hatziargyriou (2012). Firstly, potential extreme events were grouped into six different regimes. For the identification of extreme events, two different power curve models are used together with a continuous wavelet analysis. The under-estimation or over-estimation of wind power by the power curve models indicates intensity errors and phase errors related mainly to ramp events. Finally, setting a threshold on the coefficients of the continuous wavelet analysis allowed dividing wind power time-series in periods with high variability and low variability. By doing so, the three regimes mentioned previously were extended to six regimes depending upon wind power variability.

Subsequently, the Bayes rule and the radial basis function models were integrated in a ARTMAP network forming the novel so-called RBF-pARTMAP network. The RBF-pARTMAP can be used to estimate the probability occurrence of each regime. The final prediction is then obtained from the combination of the regimes probabilities with the predictions of the six RBFNNs. For on-line operation, a novel adaptive learning algorithm was designed and applied, enhancing the RBFNNs performance using new observations.

The overall performance of that approach was evaluated based on a real-world case study with data for a period of one-year. The performance was observed to be better than that of persistence (that is, a naive predictor stating that power generation at time $t+k$ will be equal to the last available measurement at time

t) for all lead times except for 1-hour ahead. This can be seen from Figure 9 which shows the evaluation of the NMAE (Normalized Mean Absolute Error) as a function of lead time for the newly proposed regime-switching approach and for the considered benchmark methods. In parallel to the comparison with respect to persistence forecasting, a 12.5% global improvement (i.e. , if considering all lead times, indifferently) compared to the state-of-the-art approach of Sideratos and Hatziargyriou (2007), was observed.

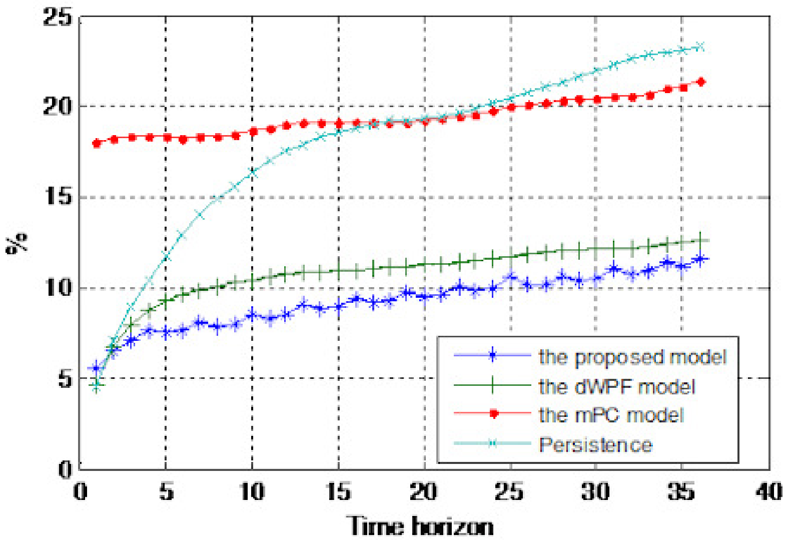

FIGURE 9: *The NMAE of the regime switching model, compared with that for the state-of-art dWPF model (Sideratos and Hatziargyriou 2007) and for persistence, as a function of lead time.*

Figure 10 illustrates reasons from improvement in forecasts accuracy from the application of this regime-switching approach, for a number of events of particular importance for power systems operations. These events show a better fit of forecasts and measurements overall, in terms of both magnitude and phase.

4 Event-based view of the forecasting problem

4.1 Event-based definition of the prediction problem

When considering and discussing extremes and wind power forecasting, it appeared that extremes corresponded to specific types of events about which the forecast users wanted to be informed in advance. These specific events were seen to have substantial consequences on their assets' safety and on their revenues. In view of these exchanges with forecast users and the existing knowledge in meteorological forecasting (and forecasting science more generally), it was proposed to introduce the concept of *event-based* forecasting for the renewable energy application. Examples events of interests are cut-offs, ramps and periods with high power fluctuations.

Event-based prediction is actually fairly common in meteorology and climate, where typical extreme events are heat waves, hurricanes, etc. They are defined by setting a threshold on the value of a defining

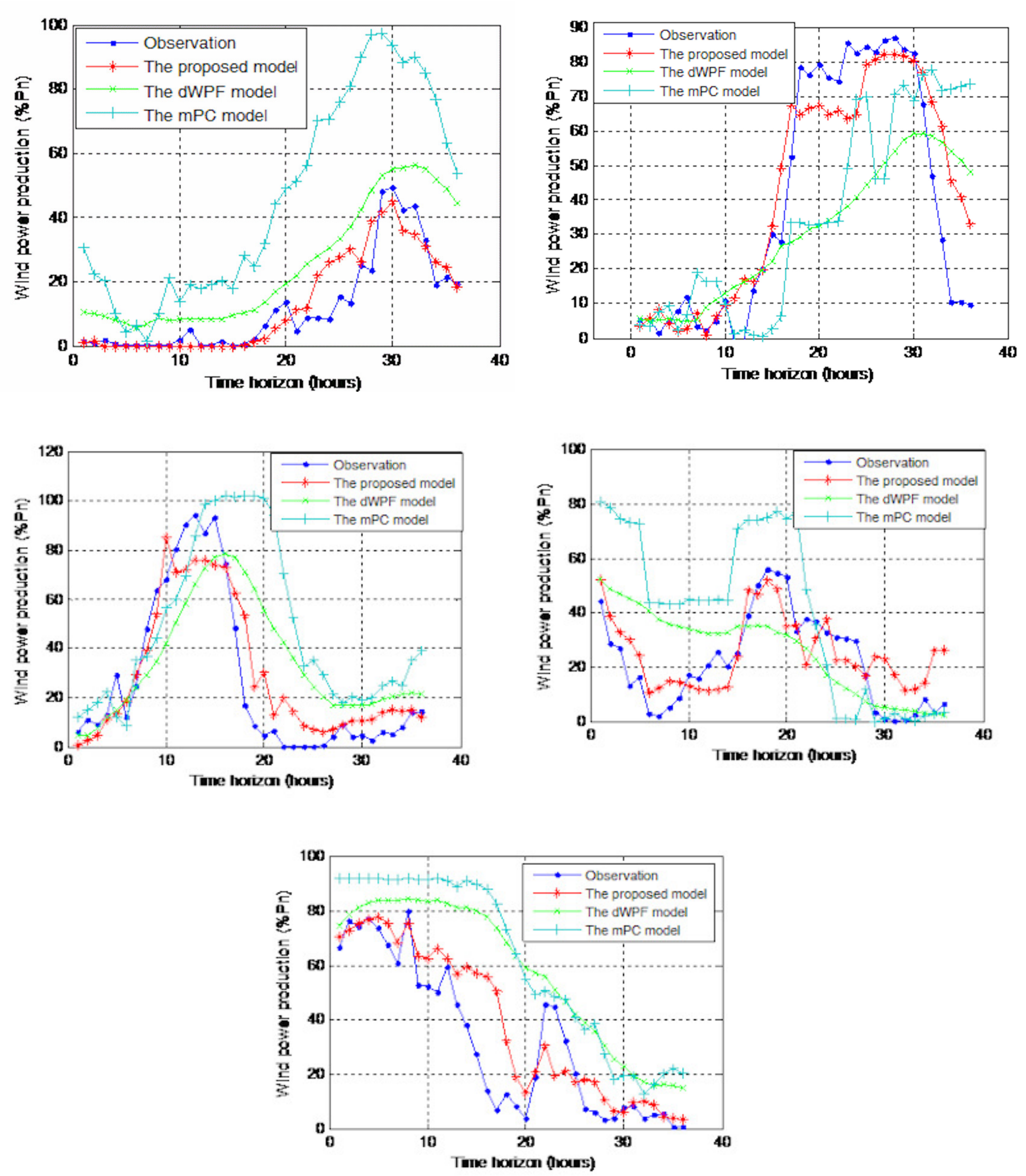

FIGURE 10: *The predictions of the regime switching model, of the state-of-art dWPF model (Sideratos and Hatziargyriou 2007) and of a direct power curve model during several extreme power system events: a) on up-ramp event, b) on intensity error event, c) on a phase error event, d) on a down-ramp event and e) on a high wind power fluctuation event.*

variable over a period of interest. For instance a heat wave may be defined as the observed temperature being above 35 degrees for more than 3 days in a row. The interest of working with events is that instead of considering the full range of potential values for the defining variable, observations and forecasts consist of 0s and 1s only — the heat wave is there or not. If applied to the case of wind energy and some of the extreme events to be studied, extreme events can be defined based on a similar concept. As an example, a ramp roughly correspond to a gradient of power production over a certain period of time

being higher than a certain threshold.

In a generic manner, the event is defined as a binary random variable, with observation $x_t \in \{0, 1\}$ at a given time t. We use the same notation for the corresponding random variable. Subsequently, an event-based prediction $\hat{x}_{t+k|t}$ issued at time t for time $t + k$ informs of the probability of x_t occurring at that time,

$$\hat{x}_{t+k|t} = \mathrm{P}\left[x_{t+k} = 1\right] \tag{4}$$

given the information available at time t. In the fully probabilistic case, $\hat{x}_{t+k|t} \in [0, 1]$ while in the deterministic case, $\hat{x}_{t+k|t}$ can only be equal to 0 and 1. Event-based forecasts can be directly based on statistical models such as Generalized Linear Models (GLMs). For a review of GLMs, see Madsen and Thyregord (2011).

An example of event based forecasts in the form of ramp predictions is depicted in Figure 11, as issued on the 4$^{\text{th}}$ April 2007 at 00:00UTC for the whole onshore capacity of Western Denmark (for a nominal capacity of 2.515 GW on that day), along with related ramp probability forecasts. Ramps are defined as a change of more than 500MW in power generation within a 6-hour time window. It shows forecast probabilities of observing ramp events at different lead times in the future, where ramps events are defined as changes of more than 500 MW in power generation within a 6-hour time window. Lead times therefore correspond to the centre of these time windows.

In the following emphasis is placed on 3 different types of events that were deemed of importance for forecast users: *(i)* ramps, *(ii)* cut-offs, and *(iii)* periods with high variability of power generation.

4.2 Ramp events

Since issuing probabilistic forecasts in the form of predictive densities for every locations and lead times, individually, may not be suitable for describing some complex events like rapid changes in power production (the so-called ramps), it was proposed to forecast ramps on the basis of a specific time at which they may occur, complemented by the probability of observing a ramp within a set of time intervals. These sets of time intervals, also referred to as prediction intervals (temporal ones, not to be mistaken with the classical interval forecasts for continuous variables), are centered on that lead time at which a ramp is expected to occur, also covering some lead times before and after that. Thus, this approach also provides an uncertainty estimate about the timing of ramp events. An illustration of that concept is given in Figure 12 below.

The proposed methodology for issuing ramp forecasts and timing uncertainty relies on numerical weather prediction ensembles, as provided by ECMWF, then converted to wind power ensembles. Advanced filtering techniques are then applied for converting the information given by the scenarios of future wind power generation into a meaningful signal corresponding to potential rapid changes in power generation levels. An extensive description of that proposal is available in Bossavy *et al.* (2012). It was implemented in the ANEMOS platform and is running operationally at several transmission system operators (e.g. PPC, RTE).

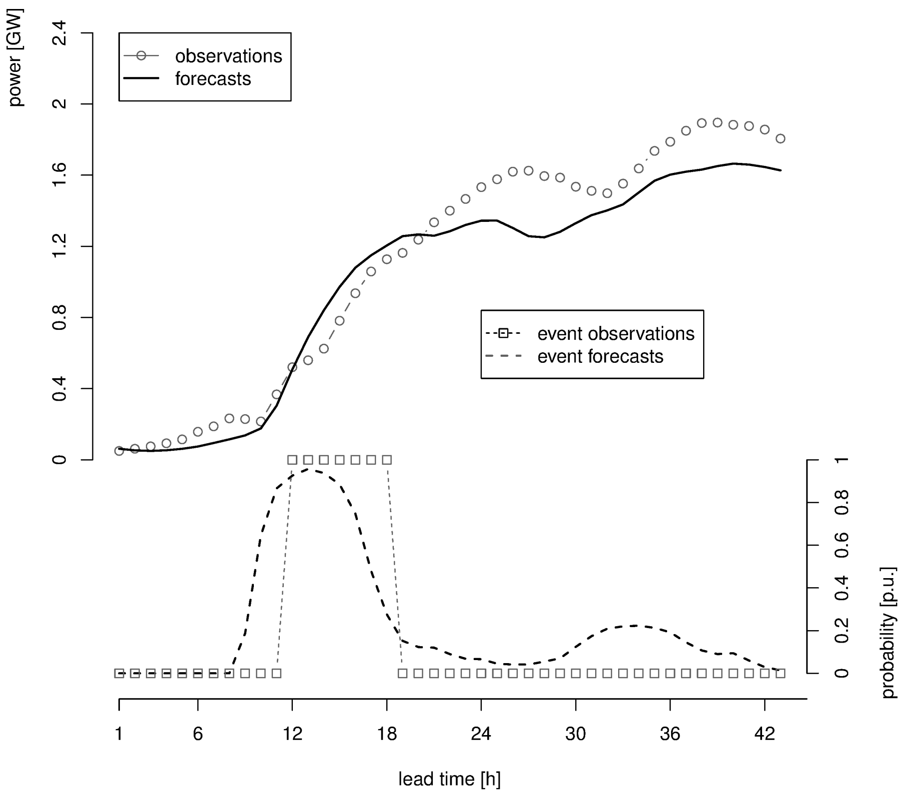

FIGURE 11: *Point forecasts of wind power generation issued on the 4th April 2007 at 00:00UTC for the whole onshore capacity of Western Denmark (for a nominal capacity of 2.515 GW on that day), along with related ramp probability forecasts. Ramps are defined as a change of more than 500MW in power generation within a 6-hour time window. Ramp forecasts are filtered from 1000 scenarios of short-term wind generation. (figure taken from Morales et al. (2013))*

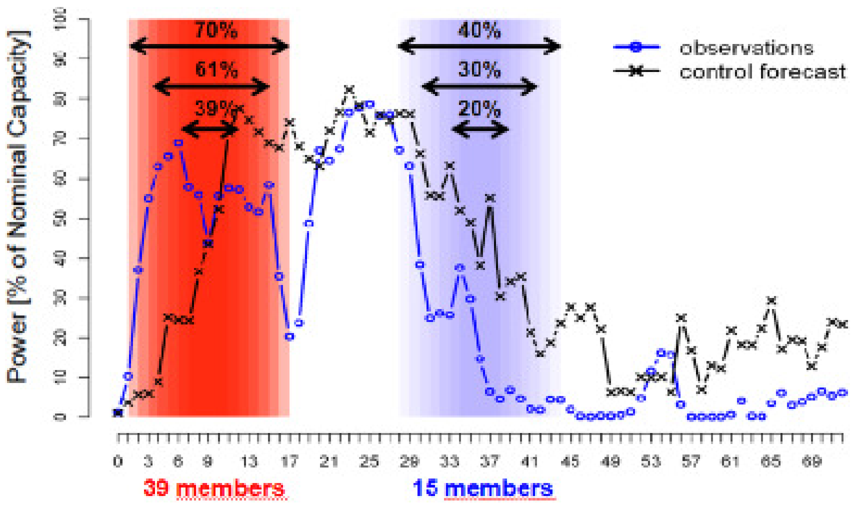

FIGURE 12: *Example of a ramp forecast complemented with a prediction interval informing about the ramp timing uncertainty. The coloring indicates an increased risk of observing a ramp event, while the width of the prediction interval tells about the time window over which such an event may be observed.*

4.3 Cut-offs

Wind turbines are equipped with a safety feature which ensures that they shut down to prevent damage during periods with very high wind speeds. More precisely when wind speed exceeds a threshold, known as the cut-off wind speed, turbines simply stop producing electricity. Here a statistical methodology for the prediction of cut-off events due to extreme wind speed in wind power forecasting using observed data was proposed. This prediction is important both at a wind farm and regional levels. For a wind farm, a cut-off event may translate to the rapid loss of a very significant share of the power supply, since going 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. It will occur in a smoother manner though, that is, the decrease in production may happen over a longer period of time thanks the geographical smoothing effect.

The prediction of a cut-off event is not easy, since these events are rarely observed. The goal is then to build a methodology to predict a cut-off for a wind farm that has never experienced that event.

For a given measured power P_t two alternative wind speeds are collected, as illustrated in Figure 13. The first one is the "machine wind speed" M_t that can be read using the machine power curve for the measured power P_t. The second one is the "registered" wind speed R_t. Depending on the application, the registered wind speed can be that measured by an anemometer or that predicted by a NWP method for time t. The dataset will then consist on a set of vectors (Pt, Mt, Rt), $t = 1, \ldots, T$.

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 $M_t > V_{\text{cut}}$. 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, $\mathrm{P}\left[M_t > V_{\text{cut}}\right]$.

This probability can be used, for instance to correct a wind power prediction $\hat{p}_{t|t-k}$ where the cut-off event is not taken into account. The proposed methodology for the prediction of a cut-off event that was developed and evaluated in the SafeWind project 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 $(M_{\text{cut-in}} + 1)$ and $(M_{\text{rated}} - 1)$, where $M_{\text{cut-in}}$ is the theoretical cut-in wind speed and $(M_{\text{rated}}$ 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,

$$M_t = \alpha_0 + \alpha_1 R_t^{\delta} + \epsilon_t \tag{5}$$

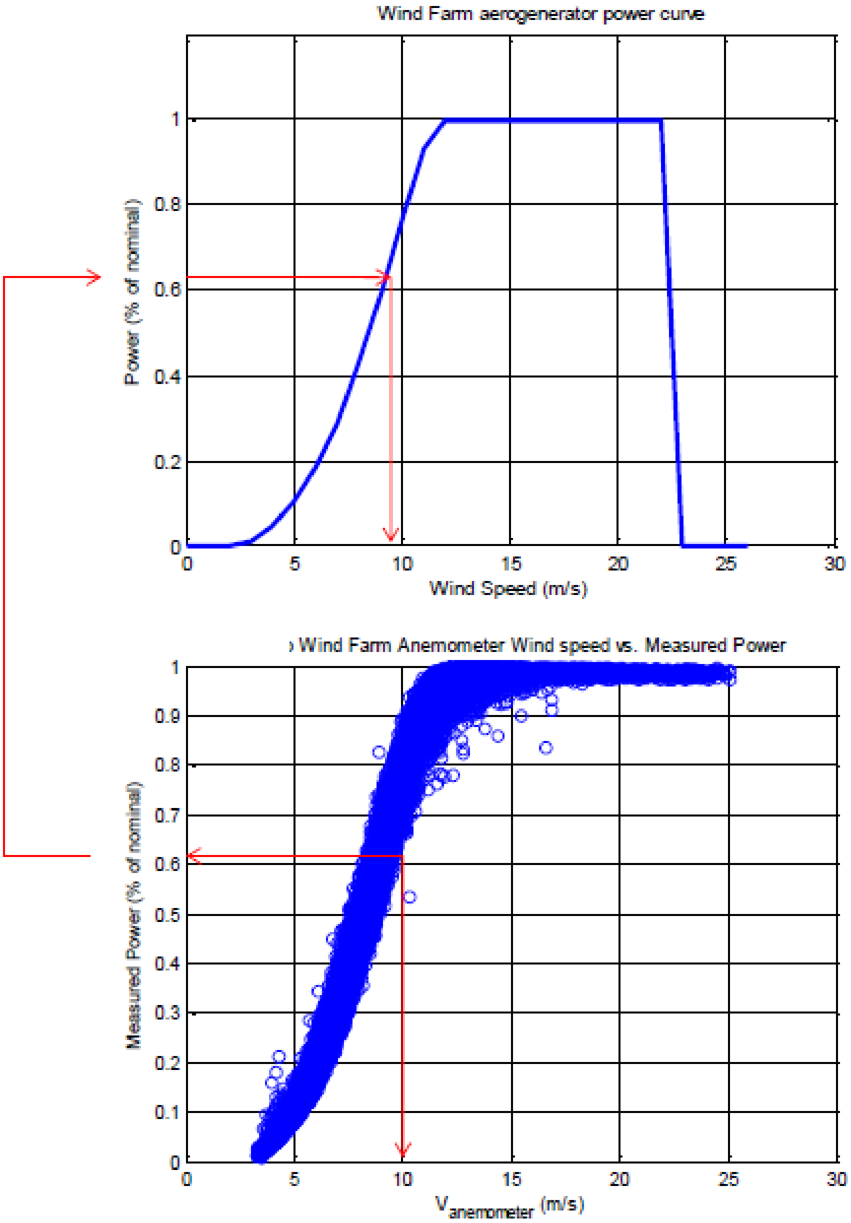

FIGURE 13: *Alternative wind speeds for a given measured power. The upper plot shows the 'machine wind speed', whereas the lower plot shows the 'registered' wind speed.*

with the power exponent in R_t^δ chosen so as that which 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 R_t the machine wind speed is a random variable that is assumed normally distributed with mean

$$\hat{M}_t = \hat{\alpha}_0 + \hat{\alpha}_1 R_t^\delta. \tag{6}$$

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

The model is

$$\epsilon_t^2 = \beta_0 + \beta_1 R_t^\delta + \beta_2 \left(R_t^\delta\right)^2 + \nu_t. \tag{7}$$

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

$$\hat{\sigma}_t^2 = \hat{\beta}_0 + \hat{\beta}_1 R_t^\delta + \hat{\beta}_2 \left(R_t^\delta\right)^2 \tag{8}$$

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

$$\pi_{\text{cut}} = \text{P}\left[X > M_{\text{cut}}\right], \tag{9}$$

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

$$X \sim \mathcal{N}\left(\hat{\alpha}_0 + \hat{\alpha}_1 R_t^\delta, \hat{\beta}_0 + \hat{\beta}_1 R_t^\delta + \hat{\beta}_2 \left(R_t^\delta\right)^2\right). \tag{10}$$

Step 8 : Use this probability to correct your wind power prediction $\hat{p}_{t|t-k}$ for time t, with the risk of having a cut-off

$$\hat{p}^c_{t|t-k} = \hat{p}_{t|t-k}[1 - \pi_{\text{cut}}]. \tag{11}$$

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

$$\hat{p}^c_{t|t-k} = \hat{p}_{t|t-k}[1 - \mathbf{1}\{\pi_{\text{cut}} > \pi^0_{\text{cut}}\}\pi_{\text{cut}}]. \tag{12}$$

where π^0_{cut} is some threshold probability below which the correction is not applied, and is the indicator function.

Such an approach was applied and evaluated on the number of test cases. One of the main conclusions is that owing to the rare occurrence of such events, more test cases would be necessary in order to draw significant conclusions. Another important conclusion is that it is required in the future that all turbines may be able to provide information on wind, power, but also its state (i.e., specifying whether or not a cut-off has really occurred as triggered by the machine).

4.4 Periods with high wind power variability

4.4.1 Characterization of strong wind power fluctuations

In contrast to single ramp events, periods with high wind power variability, i.e. series of ramps with a time scale of minutes, were investigated. The knowledge about fluctuating wind power becomes highly interesting for very large-scale offshore wind farms when the range of short-term fluctuations reaches several 100 MW and put risk to the distribution and/or transmission grid (and requires a lot of balancing

power). For instance, fluctuations of the size of 100MW have been observed for Horns Rev (Akhmatov *et al.* 2007). As soon as many spatially distributed wind parks or wind turbines are aggregated fluctuations on the time scale of minutes are smoothed out and the desired effect to study weather-related wind power fluctuations disappears. Since the future wind power density is highest for offshore sites, though offshore data where not available for that particular work, emphasis was placed on near-shore sites instead, where the weather conditions resemble most the offshore ones. The selection of such a site had to fulfil the following criteria of *(i)* rather concentrated wind power capacity, *(ii)* near-shore conditions and *(iii)* data resolution of at least 15 minutes. Thus, the chosen test case is the transformer stations of Tjaereborg where approximately turbines with 20MW installed power are connected. The site is close to Esbjerg and can be considered as near-shore. The data made available by Energinet.dk for the years 2006-2010 was used.

The aim was to study under which atmospheric conditions strong wind power fluctuations occur and which variables of NWP models can be used for characterization and forecasting. It is out of scope for the next couple of years to forecast wind fluctuations precisely in amplitude, frequency and phase for any site. Even the forecast with short lead times cannot be successful, as it will be impossible to gather (or say, measure) all initial conditions accordingly. Furthermore the atmospheric models will still exhibit model errors in the dynamical core and deficiencies in the physical representation of atmospheric processes that will amplify very quickly in the model and render any very high spatial and temporal resolved predictions impossible.

The distinct forecast of frequency and amplitude of wind (power) fluctuations is not only impossible but actually also not the information that transmission system operators (TSO) ask for. To quantify the amount of secondary reserve in a certain time period to level the fluctuating feed-in of a specific wind park might be more important. In the coming years, high offshore installation densities combined with little spatial distribution will urge the transmission operator to keep high shares of balancing power and to spend high share of regulation energy in case a large offshore wind farm enters a period of high wind power fluctuations.

The analysis of short-term wind power fluctuations requires a proper metric to quantify fluctuations. The metric should account for wind power gradients that occur during a certain time interval/period. However, not all wind power gradients are relevant to consider as wind power fluctuations that might harm the integration of wind power. When analysing the distribution of temporal wind power gradients at Tjaereborg it was found that 4% of all 15 minute gradients exceed the size of 10% of rated power (20MW) (von Bremen 2012). von Bremen and Bush-Saleck (2010) have proposed a simple metric to quantify the strength and the occurrence of wind power fluctuations with gradients higher than 10% of rated wind power capacity by adding up the absolute value of gradients, e.g. in a 6h time period with

$$totalfluc_{0.1-1} = \sum_{i \in T_{6h}} |P_i - P_{i-1}|, \quad |P_i - P_{i-1}| > 0.1 \tag{13}$$

Figure 14 shows the time series of $totalfluc_{0.1-1}$ (right) and also $totalfluc_{0-0.1}$ (left) that summarizes only the gradients $<$10% of rated power. It can be seen that fluctuations with gradients $<$10% of rated power are distributed quite equally during the seasons and no exceptionally high values occur. Conclusively, gradient $<$10% of rated power are basically happening all the time and do not require special attention. This is different for $totalfluc_{0.1-1}$. which is largest in autumn and beginning of winter, but is

smallest in summer.

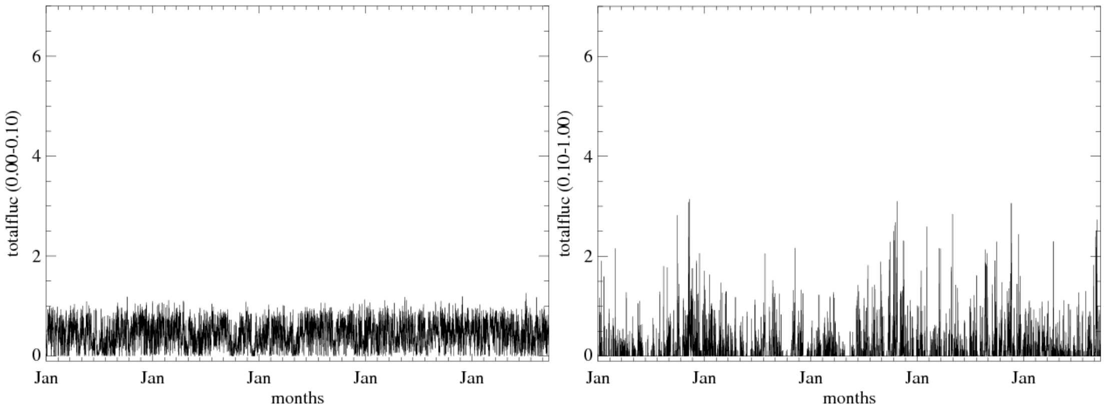

FIGURE 14: *Time-series of $totalfluc_{0-0.1}$ (left) and $totalfluc_{0.1-1}$ (right) at transformer station TJAe (Tjaereborg) 2006-2010.*

4.4.2 Wind power fluctuations related to atmospheric conditions

In von Bremen *et al.* (2010), Vincent and Hahmann (2011) and Vincent *et al.* (2010), it was showed that the highest wind power fluctuations occur in north-western flow conditions. North-western flow is often (not always) connected with cold-air outbreaks from the polar ice caps. As the air mass is very cold, the stratification is unstable. For the choice to describe atmospheric stability it is important to consider that the parameters are available from NWP models. Unfortunately, parameters like Monin-Obhukov length or Richardson number that are normally used to determine atmospheric stability are not available from NWP. However, it is not the stability in the atmospheric surface layer that leads to convection and the development of clouds. Moreover the temperature difference between the underground (sea surface) and the entire air mass is responsible to allow or suppress lifting of air parcels. In particular, cold-air outbreaks over the open North Sea leading to organized convection (Noteboom 2006), namely open cellular convection (OCC) can be considered as a source of wind fluctuations due to local wind conditions that develop in the down and updraft of the clouds. In this study the air temperature difference between the lowest atmospheric model level from ECMWF (10m) and the (surface) skin temperature (SKT) is used as a proxy to characterize if the stratification is rather stable or unstable. As Tjaereborg can be considered near-shore the closest sea surface temperature has been used.

Figure 15 shows the relation between wind power fluctuation ($totalfluc_{0.1-1}$) and 70m wind speed and the temperature difference T91-SKT. All meteorological parameters are taken from 6-hourly ECMWF analyses. The black dots indicate that no wind power fluctuations with gradients $>10\%$ occur for low wind speeds. The highest fluctuations occur for wind speeds between 8-12 m.s^{-1} and negative differences of air temperature and skin temperature. In case the sea surface skin temperature is higher than the air temperature an unstable stratification exists. It is obvious that the characterization of the meteorological condition is not strong enough to determine perfectly if convection and related wind fluctuations will occur. Sometimes high fluctuations (>1.9) also occur in stable conditions and occasionally almost no fluctuations (blue points) occur in very unstable conditions and wind speeds of around 10m/s. It can be concluded that definitely also other dependencies exists. However, the general relationship of

$totalfluc_{0.1-1}$ between wind speed and stability remains and can be used to quantify the conditional risk for severe wind power fluctuations, i.e. for given (forecasted) weather conditions fluctuations exceeds a certain threshold.

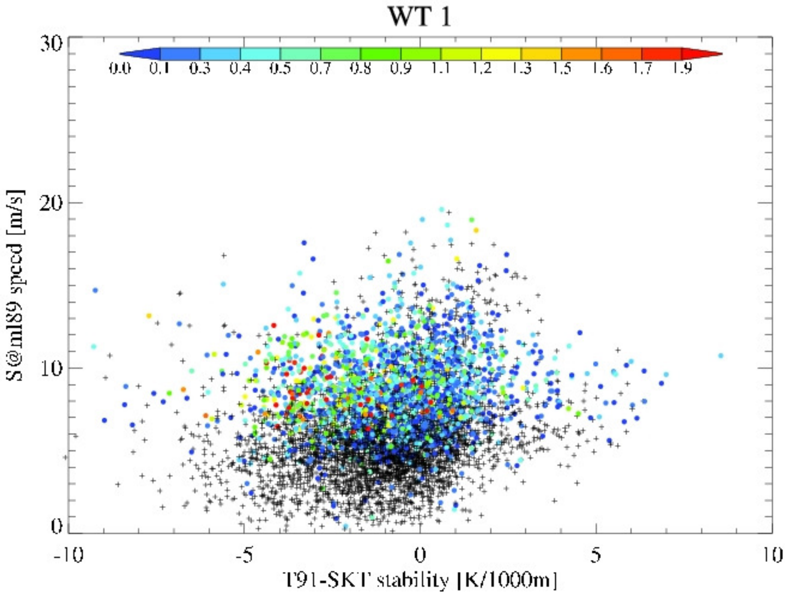

FIGURE 15: *Wind power fluctuations expressed as $totalfluc_{0.1-1}$ over 70m wind speed (ECMWF analysis) and thermal stability of the atmosphere (T91-SKT) for the years 2006-2009 at transformer station TJAe (Tjaereborg).*

Figure 16 shows the 50%, 20%, 10% and 5% quantiles of $totalfluc_{0.1-1}$. conditioned on 70m wind speed and stability (T91-SKT). The highest fluctuations occur for wind speeds of 8-12m/s and temperature differences between air and surface skin temperature of -2 to -4K. With a probability of 10% $totalfluc_{0.1-1}$ is larger than 1.5 and with a probability of 5% $totalfluc_{0.1-1}$ exceeds even 1.9. As can be seen from Figure 1 $totalfluc_{0.1-1} = 1.9$ is quite rare and can be regarded as very strong fluctuations. An example of time series with wind power fluctuations of 2.58 is shown in the SafeWind Deliverable Report Dp-6.7 (von Bremen 2012).

5 Communication of forecast information with focus on uncertainty

5.1 General considerations

The SafeWind project has focused extensively on developing methods for constructing, evaluating and communicating probabilistic forecasts of wind power generation, with a particular emphasis on the tails of predictive distributions. The underlying processes that govern wind power generation are influenced by several meteorological variables and modelling is complicated by the nonlinearity of the transformation between wind speed and wind power. With the complexification of forecasting approaches and resulting products, the communication and appraisal of forecast uncertainty information may become cumbersome for most forecast users. This is the reason why efforts were spent on considering simple and meaningful ways to communicate such an information.

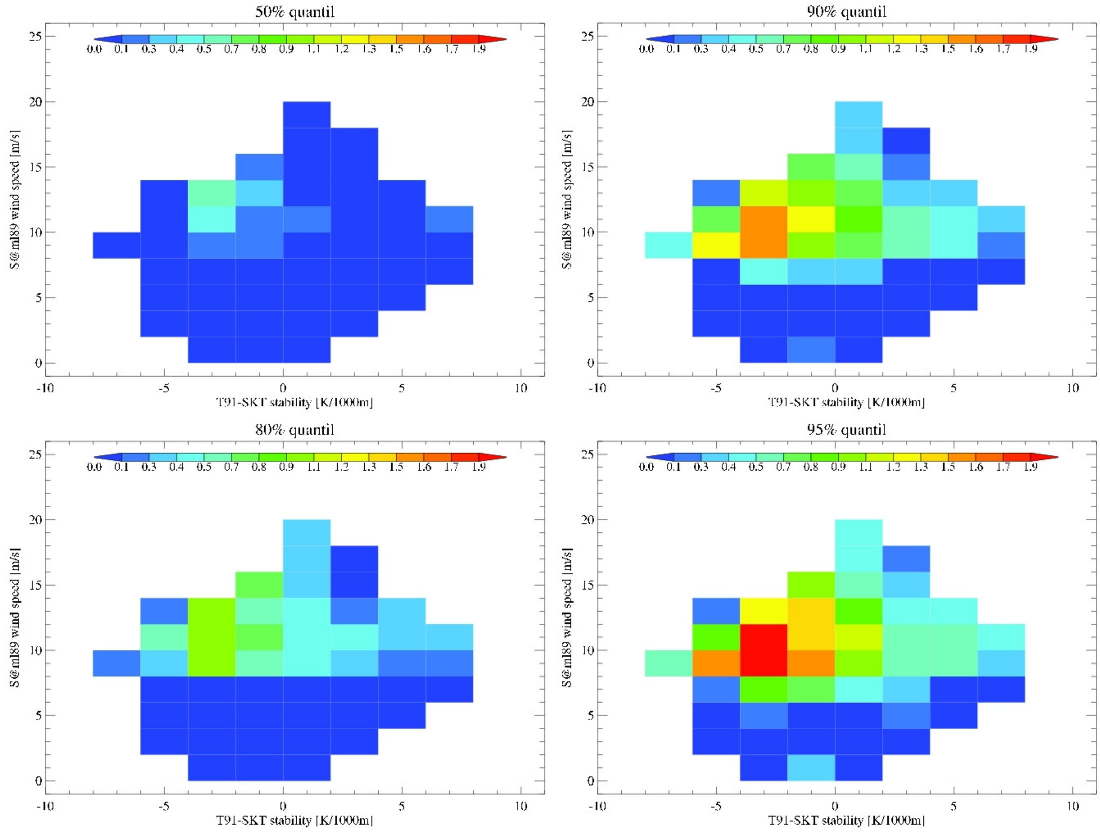

FIGURE 16: *Quantiles of $total fluc_{0.1-1}$ depending on 70m wind speed (ECMWF analysis) and thermal stability of the atmosphere (T91-SKT) for the years 2006-2009 at transformer station TJAe (Tjaereborg).*

A review of the approaches available for assessing predictability of extremes, accounting for uncertainty and communicating this information in a coherent format to end-users was performed, with particular emphasis placed on weather-related processes. Describing uncertainties relating to extremes is not trivial: in contrast to extreme value analysis that offers a static assessment of extremes, the approach developed here should provide a dynamical assessment of risk. As an example wind power forecast ensembles derived from ensembles of numerical weather predictions offer a means of assessing predictability in a dynamic manner. A practical example of communicating the spatial profile of risk levels will be given in the following for the particular extreme events associated with the storm Xynthia.

The ensemble members represent different scenarios and the spread among the members is expected to provide a measure of predictability. This spread may be translated into risk indices with the purpose to give a simple and meaningful information about the expected forecast error, particularly the risk for large errors. As a starting point, a previously proposed definition of a risk index (Pinson *et al.* 2009), based on the weighted standard deviation of the ensemble members, was evaluated. The evaluation was made on three French wind farms on different temporal and spatial scales. Furthermore, a number of alternative definitions of risk indices were proposed and examined to assess whether they could be more suitable. Finally an investigation of how risk indices can be presented to end-users was performed, including suggestions and evaluations of a number of alternative methods.

Results also demonstrated how that the risk indices were able to discriminate among situations with various levels of forecast error observed a posteriori. The choice of prediction model, particularly with regard to how the ensemble spread is affected when meteorological ensembles are converted to power ensembles and eventually to risk indices, was found to be important. Although the alternative risk indices do not always outperform the previously used definition, some of them give similar performance while being easier to calculate and interpret. Finally, it was found that the way in which risk indices are presented to end-users is important. Some of the proposed alternative methods were found to outperform those used in previous studies.

5.2 Specifics of an example contribution

As a more specific contribution illustrating the general considerations covered in the above, let us consider the example case of an approach to communicate forecast uncertainty for regions, e.g. control zones that are characterized by an inhomogeneous distribution of wind power capacities. Much work has been done in the past to estimate forecast uncertainty and communicate prediction risk indices for either single sites (Pinson *et al.* 2009) or dynamic prediction intervals at the national level (Dobschinski *et al.* 2010). But even if the aggregated forecast uncertainty is known, end-users like transmission system operators for instance might be interested in how the forecast uncertainty is distributed geographically. This knowledge can be used to discriminate regions with lower and higher forecast uncertainty levels. Thus, it may for instance be possible in the future to allocate preferably reserve power in areas that exhibit a high risk that wind power forecast are imperfect. Or, in case wind power production and demand are displaced geographically, reserve power can be obtained closer to the sites of demand. In future power systems with Smart Grid technology, i.e. distributed storage, PV and biogas systems and demand side management the uncertainty in wind power production will be important information to develop economic balancing strategies.

The work here was based on the ensemble wind power forecasts produced in some other tasks of the SafeWind project. The details on the wind power forecast model are given in Deliverable Dp-5.10 (von Bremen *et al.* 2012). The ensemble wind power forecast utilize the Ensemble Prediction System (EPS) of ECMWF and the new 100m winds that became operational on 26 January 2010. The results presented here are for the 50Hertz control zone in Eastern Germany. It is worthwhile to mention that real wind power production data with a spatial resolution similar to the EPS (0.25) does not exist. Thus, 100m winds from the ECMWF analysis are utilized to compute the spatially resolved wind power production. The configuration of the wind power forecasting model is identical to the forecasting mode.

As an indicator of forecast uncertainty prediction risk directly based on the ensemble spread, i.e. standard deviation of the wind power ensemble, is used. Testing has shown that the 70% inner quantile range of the ensemble distribution is for higher lead times too sensitive with respect to the average forecast error. Figure 17 shows the RMSE of the ensemble mean wind power forecast (solid line) with forecast lead time. The average ensemble spread of wind power (dashed line) is slightly smaller which indicates too little spread (Leutbecher and Palmer 2008). However, the two lines have the same slope and do not diverge. In contrast to this, the 70% inner quantile range of the ensemble distribution increases steeper with lead time compared to the RMSE of the forecast. Conclusively, the 70% inner quantile can be regarded to be too sensitive for higher lead times.

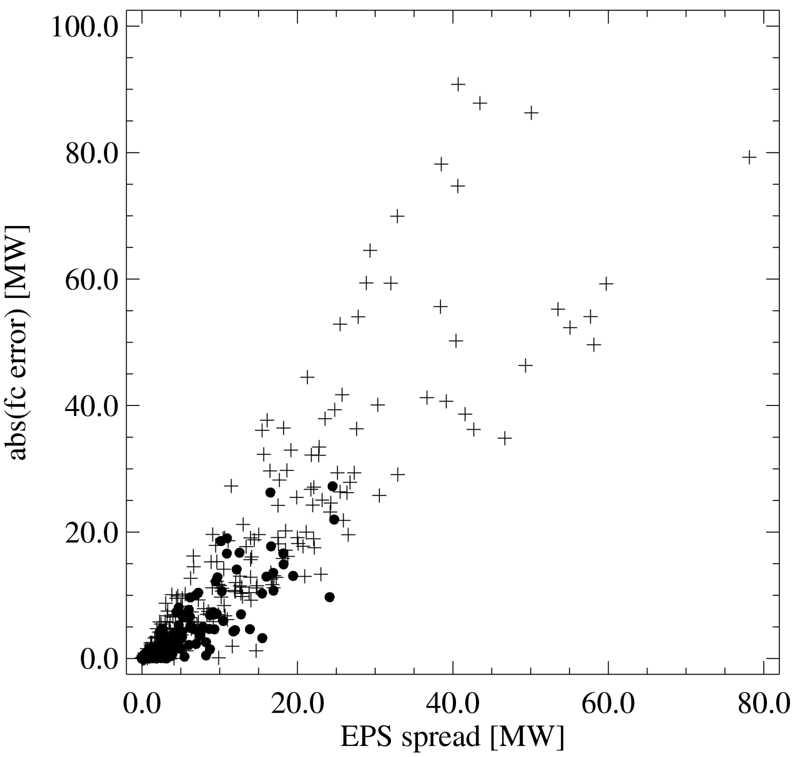

FIGURE 17: *RMSE of ensemble mean (solid), ensemble spread (dashed) and inner quantile (p=70%) of the distribution of ensemble wind power forecasts for 50Hertz in normalized units (to installed wind power) for February and March 2010.*

For the occasion of winter storm Xynthia that hit Germany between 29 February and 1 March 2010 the spatial distribution of wind power uncertainty in the 50Hertz control zone is shown and discussed here. The powergrams in Figure 18 show that there are clear indications of Xynthia 4 days in advance for the 50Hertz control zone. The 48h wind power forecast of Xynthia (Figure 18, on the right side) is very good in every sense, i.e. the amplitude and the timing is correct for the deterministic and the ensemble forecast, while ensemble forecasts indicate a low level of forecast uncertainty.

The spatial distribution of the poor 96h wind power forecast of Xynthia can be seen in Figure 19 (left). Large absolute forecast errors of more than 90 MW per grid point occur (Figure 19, bottom left). However, the ensemble spread (Figure 19, middle left) in those grid points is very large. Conclusively, a high uncertainty was indicated and enough time was given to take actions like for instance to increase the amount of regulating reserves to balance deviations from the day-ahead forecast. But as seen in Figure 19 (right bottom) the 48h forecast is far better and the maximal forecast error is less than 30MW per model point. This good forecast skill is anticipated from the low forecast uncertainty, i.e. ensemble spread in Figure 19 (middle right).

The scatter plot between the absolute forecast error and the ensemble spread at the level of model points in the 50Hertz control zone is shown in Figure 20 for the +96h (plus signs) and the +48h (bullets) forecast horizon. For both forecast horizons the relation between forecast error and EPS spread is very similar. The occurring values of the EPS spread are much smaller for the +48h forecast and thus it is anticipated that the forecast error is considerably smaller.

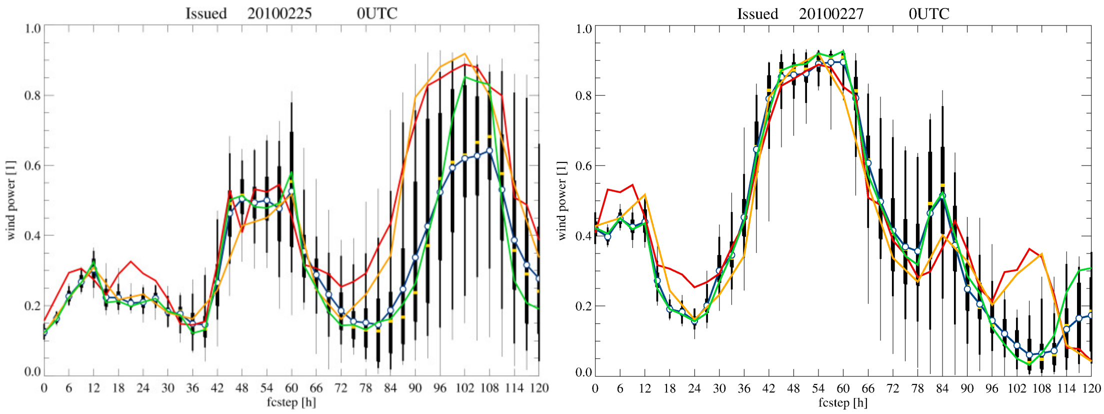

FIGURE 18: *Powergram of probabilistic wind power forecast of storm Xynthia at 1 March 2010 (0UTC) in the 50Hertz control zone. The forecast uncertainty is extremely reduced in the newer forecast run (right) issued at 27 February (0UTC) compared to the medium-term forecast (left). Measured wind power in red, ensemble mean in blue, control (deterministic) forecast in green and simulated wind power in orange. The vertical boxes represent the 50% and 90% inner quantiles while the minimal and maximal value of the ensemble is indicated by the tip of the vertical thin line.*

6 Conclusions and perspectives

The work in the WP-6 of the SafeWind project placed emphasis on novel methods for wind power forecasting and extremes, where the central lead times of interest where those relevant for operations management and market participation, i.e. up to 48-72 hours ahead. Shorter lead times (in the order of 10 minutes to 6 hours) were considered for some of the investigations related to new predictive densities and regime-switching modelling for instance since it is where the highest benefits from these new approaches could be shown. Also when looking at the use of meteorological ensemble forecasts as provided by ECMWF as input, longer lead times (up to 5-7 days ahead) were studied since an acceptable level of predictability could still be reached. Whatever the forecast length and temporal resolution, a main objective was to propose, develop and evaluate forecasting methodologies that could allow issuing

- *more accurate point forecasts*, since comprising the forecasting product that this the most employed by forecast users in practice,
- *more reliable and skilled probabilistic forecasts* (in various forms), since such types of forecasts may bring the highest value as input to decision-making,
- *more meaningful information about forecasting uncertainty*, since too complex forecast products are difficult to appraise by most forecast users, and may then provide then with misleading signals.

As a combination of these three points, particular attention was given to the case of extreme events, which may be of meteorological nature e.g. the winter storm Xynthia , or of particular importance to market participation and power systems operations e.g. very large forecasting errors.

Some of the modelling proposals and investigations permitted to gain new insight on the dynamics and predictability of wind power generation, for instance through the regime-switching and forecast combina-

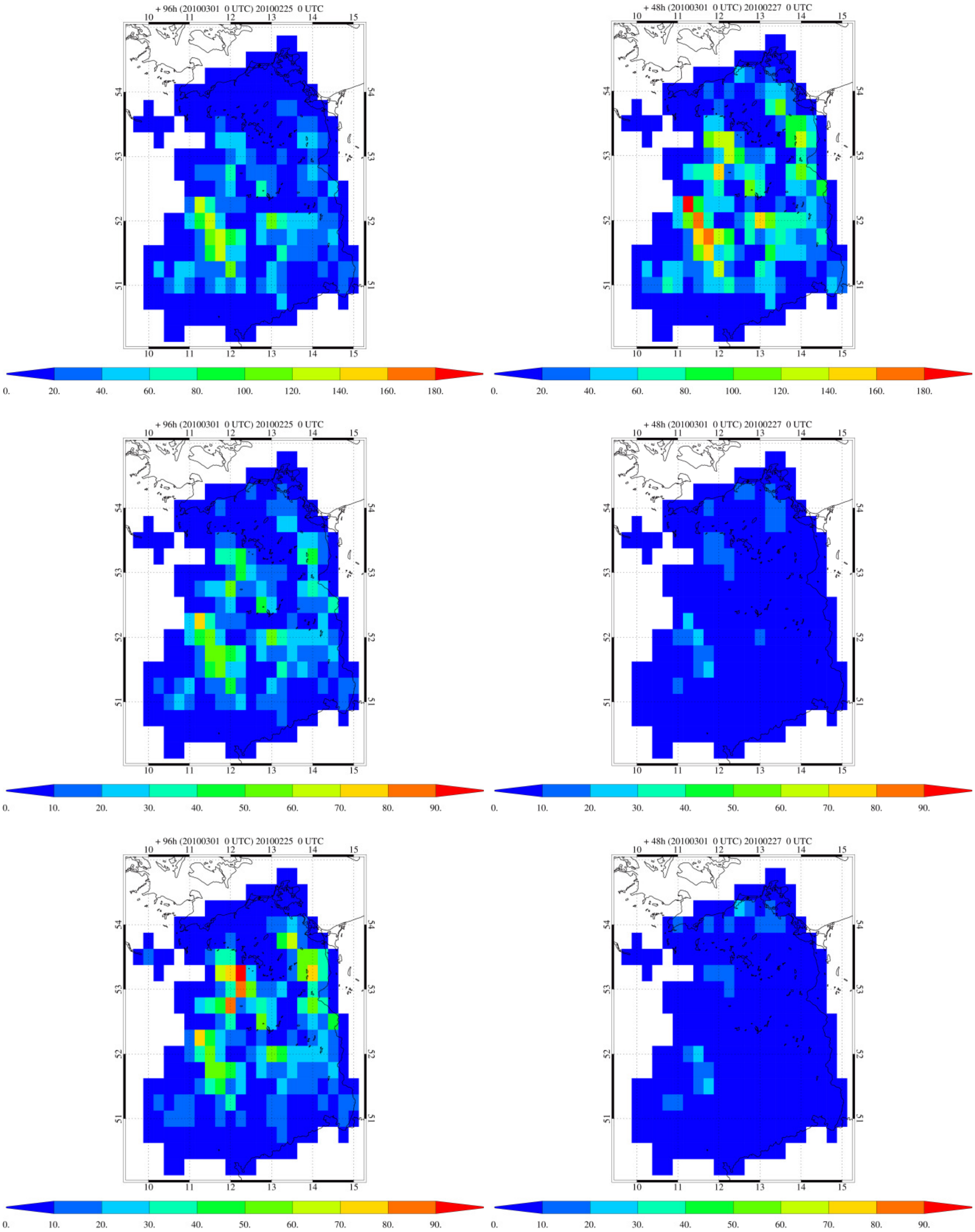

FIGURE 19: *96h (left) and 48h (right) wind power forecast of storm Xynthia for the 50Hertz control zone valid for 01 March 2010, 0UTC. Ensemble mean (top), forecast uncertainty expressed as ensemble spread (middle) and absolute forecast error (bottom) in MegaWatt.*

tion tasks telling us how offsite observations may inform of regime switches and how local measurements may already provide valuable information on power generation regimes. These investigations came as a natural extension of previous works on wind power forecasting. In contrast, other work highly benefited from a collaboration between statisticians, energy forecasters, meteorologists and forecast users,

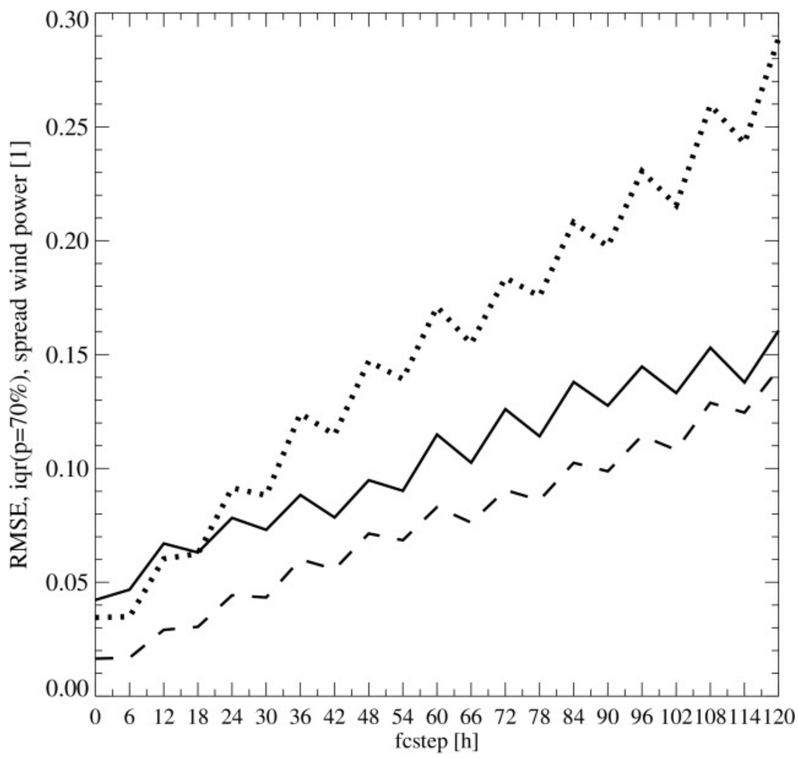

FIGURE 20: *Relation between forecast error and ensemble spread for each model point in the 50Hertz control zone for a +48h (bullets) and +96h (crosses) forecast valid at 1 March 2010 0UTC.*

for instance for the case of the optimal communication of forecast uncertainty and for the definition of event-based prediction problems. Finally, new ways to approach the wind power forecasting problem were though off through the work in this work package. As an illustrative example, the concepts of scenarios and event-based prediction open the door to the development of new forecasting methodologies and of forecasting products that may be of utmost importance to a wide range of forecast users.

Acknowldegements

Acknowledgements are due to all participants of SafeWind WP-6 for their contribution to a challenging part of the project with varied tasks and a wealth of available test case data to be used for evaluation. In addition the authors are grateful to all those who provided contributions of some sort to this report and who have not been cited in the list of authors. Finally, the authors are indebted to all data providers for the time and efforts they spent on making these data available.

References

Akhmatov V, Rasmussen C, Eriksen PB, Pedersen J (2007). Technical aspects of status and expected future trends for wind power in Denmark. *Wind Energy* **10**: 31–49.

Bermejo MA, Peña D, Sánchez I (2011). Identification of TAR models using recursive estimation. *Journal of Forecasting* **30**(1): 31–50.

Bermejo MA, Sánchez I (2012). Adaptive densitive estimation of wind power predictions. Technical Report, Universidad Carlos III Madrid, Dpt. of Statistics, 2012.

von Bremen L (2012). Methodology to forecast fluctuating wind and wind power. Deliverable Report Dp-6.7, EU-FP7 project SafeWind [Available at:`www.safewind.eu`].

von Bremen L, and co-authors (2012). Impact of various ensemble configurations for wind power predictions. Deliverable Report Dp-5.10, EU-FP7 project SafeWind [Available at:`www.safewind.eu`].

von Bremen L, Busch-Saleck N (2010). Minimizing the risk in offshore wind power integration induced by severe wind power fluctuations. In *Proc. of the third conference on "The Science of making Torque from Wind"*, Heraklion, Crete.

von Bremen L, Busch-Saleck N, Heinemann D (2010). Estimate severe offshore wind power fluctuations for better grid integration. In *Proc. of German Wind Energy Conference 2010 (DEWEK)*, Bremen, Germany.

Bremnes JB (2006). A comparison of a few statistical models for making quantile wind power forecasts. *Wind Energy* **9**: 3-11.

Bossavy A, Girard R, Kariniotakis G (2012). Forecasting ramps of wind power production with numerical weather prediction ensembles. *Wind Energy*, available online.

Botterud A, Wang J, Miranda V, Bessa RJ (2010). Wind power forecasting in U.S. electricity markets. *Electricity Journal* **23**(3): 71–82.

Clements MP (2005). *Evaluating Econometric Forecasts of Economic and Financial Variables*. Palgrave-Macmillan: London.

Dobschinski J, Lange B, Rohrig K, Saint-Drenan Y, Wessel A, von Bremen L (2009). The potential of advanced shortest-term forecasts and dynamic prediction intervals for reducing the wind power induced reserve requirements. In *Proc. of EWEC 2010 (European Wind Energy Conference)*, Warsaw, Poland.

Giebel G, Brownsword R, Kariniotakis G, Denhard M, Draxl C (2011). The state of the art in short-term prediction of wind power — A literature overview, 2nd Edition. Deliverable Report Dp-1.4, EU-FP7 project SafeWind [Available at:`www.safewind.eu`].

Jeon J, Taylor JW (2012). Using conditional kernel density estimation for wind power density forecasting. *Journal of the American Statistical Association* **107**(497): 66–79.

Jones LE (2011). Strategies and decision support systems for integrating variable energy resources in control centers for reliable grid operations. Technical Report, Alstom Grid, Washington DC (USA). [Available at uwig.org/doe_wind_integration_report.pdf]

Leutbecher M, Palmer T (2008), Ensemble forecasting. *Journal of Computational Physics* **227**(7): 3515-3539.

Madsen H, Pinson P, Nielsen TS, Nielsen HAa, Kariniotakis G (2005). Standardizing the performance evaluation of short-term wind power prediction models. *Wind Engineering* **29**: 475–489.

Madsen H, Thyregord P (2011). *General and Generalized Linear Models*. Chapman & Hall/CRC: New York.

McSharry P, Pinson P, Girard R (2009). Methodology for the evaluation of probabilistic forecasts. Deliverable Report Dp-6.2, EU-FP7 project SafeWind [Available at:`www.safewind.eu`].

McSharry P, Anastasiades G (2011). Quantile forecasting of wind power using variability indices. Technical report, University of Oxford, Smith School for Enterprise and the Environment.

Morales JM, Conejo AJ, Madsen H, Pinson P, Zugno M (2013). *Renewable Energy in electricity Markets: Operational Problems*. Springer Verlag: New York (in preparation).

Møller JK, Nielsen HAa, Madsen H (2008). Time-adaptive quantile regression. *Computational Statistics and Data Analysis* **52**(3): 1292–1303.

Noteboom S (2006). Open cell convection and closed cell convection. Chapter in: *SatManu. Stageverslag, KNMI, De Bilt*.

Pinson P (2010). Catalogue of complex to extreme situations. Deliverable Report Dc-1.2, EU-FP7 project SafeWind [Available at:`www.safewind.eu`].

Pinson P (2012). Very short-term probabilistic forecasting of wind power with generalized logit-Normal distributions. *Journal of the Royal Statistical Society, Series C* **61**(4): 555-576.

Pinson P, Girard R (2012). Evaluating the quality of scenarios of short-term wind power generation. *Applied Energy* **96**: 12–20.

Pinson P, Kariniotakis G (2010). Conditional prediction intervals of wind power generation. *IEEE Transactions on Power Systems* **25**(4): 1845-1856.

Pinson P, Madsen H (2009). Ensemble-based probabilistic forecasting at Horns Rev. *Wind Energy* **12**(2): 137–155.

Pinson P, McSharry P, Madsen H (2010). Reliability diagrams for nonparametric density forecasts of continuous variables: accounting for serial correlation. *Quarterly Journal of the Royal Meteorological Society* **136**(646): 77–90.

Pinson P, Nielsen HAa, Madsen H, Kariniotakis G (2009). Skill forecasting from ensemble predictions of wind power. *Applied Energy* **86**(7-8): 1326–1334.

Pinson P, Nielsen HAa, Møller JK, Madsen H, Kariniotakis G (2007). Nonparametric probabilistic forecasts of wind power: required properties and evaluation. *Wind Energy* **10**(6): 497–516.

Sideratos G, Hatziargyriou N (2007). An advanced statistical method for wind power forecasting. *IEEE Transactions on Power System* **22**(1): 258–265.

Sideratos G, Hatziargyriou N (2012). Wind power forecasting focused on extreme power system events. *IEEE Transactions on Sustainable Energy* **3**(3): 445-454.

Sideratos G, Hatziargyriou N (2012b). Probabilistic wind Power forecasting using radial basis function neural networks. *IEEE Transactions on Sustainable Energy*, available online.

Stephenson DB, Jolliffe IT (2012). *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 2nd Edition. Wiley & Sons: New York.

Tastu J, Pinson P, Kotwa E, Nielsen HAa, Madsen H (2011). Spatio-temporal analysis and modeling of wind power forecast errors. *Wind Energy* **14**(1): 43–60.

Trombe P-J, Pinson P, Madsen H (2012). Regime-switching modelling and offsite observations in short-term wind power forecasting. Technical report, Technical University of Denmark, Dpt. of Informatics and Mathematical Modelling.

Vincent C, Hahmann AN (2011). Hour-scale wind fluctuations over the North Sea. In *Proc. of EWEA 2011 conference (European Wind Energy Association)*, Brussels, Belgium.

Vincent C, Pinson P, Giebel G (2010). Wind fluctuations over the North Sea. *International Journal of Climatology* **31**(11): 1584–1595.