submitted to Wind Energy - (February 16, 2010)

Spatio-temporal analysis and modeling of short-term wind power forecast errors

Julija Tastu[1], **Pierre Pinson**[1,*], **Ewelina Kotwa**[1], **Henrik Madsen**[1]
Henrik Aa. Nielsen[1,2]

[1] DTU Informatics, Technical University of Denmark, Kgs. Lyngby, Denmark
[2] ENFOR A/S, Hørsholm, Denmark

Abstract

Forecasts of wind power production are increasingly being used in various management tasks. So far, such forecasts and related uncertainty information have usually been generated individually for a given site of interest (either a wind farm or a group of wind farms), without properly accounting for the spatio-temporal dependencies observed in the wind generation field. However, it is intuitively expected that, owing to the inertia of meteorological forecasting systems, a forecast error made at a given point in space and time will be related to forecast errors at other points in space in the following period. The existence of such underlying correlation patterns is demonstrated and analyzed in this paper, considering the case-study of western Denmark. The effects of prevailing wind speed and direction on autocorrelation and cross-correlation patterns are thoroughly described. For a flat terrain region of small size like western Denmark, significant correlation between the various zones is observed for time delays up to five hours. Wind direction is shown to play a crucial role, while the effect of wind speed is more complex. Nonlinear models permitting capture of the interdependence structure of wind power forecast errors are proposed, and their ability to mimic this structure is discussed. The best performing model is shown to explain 54% of the variations of the forecast errors observed for the individual forecasts used today. Even though focus is on one-hour-ahead forecast errors and on western Denmark only, the methodology proposed may be similarly tested on the cases of further look-ahead times, larger areas, or more complex topographies. Such generalization may not be straightforward. While the results presented here comprise a first step only, the revealed error propagation principles may be seen as a basis for future related work.

Keywords: wind power prediction, forecast errors, correlation analysis, spatio-temporal modeling, non-linear regime-switching modeling

* *Corresponding author:*
P. Pinson, DTU Informatics, Technical University of Denmark,
Richard Petersens Plads (bg. 305 - 212), DK-2800 Kgs. Lyngby, Denmark.
Tel: +45 4525 3428, fax: +45 4588 2673, email: pp@imm.dtu.dk, webpage: www.imm.dtu.dk/∼pp

Introduction

The optimal integration of wind energy into power systems requires high-quality wind power forecasts, preferably accompanied with reliable estimates of the forecast uncertainty. So far, state-of-the-art prediction systems typically provide forecasts for a single wind turbine, for a wind farm, or over a region with significant installed wind power capacities [1, 2]. Even if forecasting methodologies are developed for different spatial resolutions, see e.g. Siebert [3], the spatio-temporal interdependence structure in the wind generation field[1] is seldom considered, since it is assumed to be fully captured by meteorological predictions used as input. Recently however, some research works have concentrated on wind speed prediction using spatio-temporal correlation with application to wind power prediction (for short look-ahead times, typically up to two-hours ahead), thus showing potential interest in accounting for these aspects [4, 5, 6]. These works mainly deal with cases where wind behavior between sites is easier to model, owing to terrain topology, or wind climatology. For instance, Larson & Westrick [4] consider the test case of a potential site for a wind farm located at the exit to the Columbia River Gorge, while meteorological observations are available from the entrance of this same Gorge. Another example relates to the work of Damousis *et al.* [5], for which information is available upstream of the location considered in the Thessaloniki area, and with quite steady prevailing thermal winds. In a more general setup, Gneiting *et al.* [6] have recently proposed several regime-switching models which account for two dominant wind directions while predicting wind speed up to two-hours ahead, with an interesting extension to probabilistic forecasting. The results of all the works mentioned above show significant improvements compared to benchmark prediction methods e.g. persistence. But, if these methods were to be applied to other types of case-studies, for which wind behavior is more complex and where no channeling effect is present, or for larger areas, one should then not expect similar performance of the models in terms of forecast quality. More advanced models may be needed for such cases, as discussed by Hering & Genton [7] for instance, potentially requiring significant expertise for identification of their structure or estimation of their parameters.

In operational conditions, state-of-the-art forecasting methods of wind power generation are commonly optimized with a focus on the wind farm (or aggregation of wind farms) of interest. So far, they do not account for potential information from neighboring sites, for example other wind farms or meteorological stations. Having a broader view of the forecasting problem, one could account for the possibility that, even though forecasting systems are optimized for local conditions, the inertia in meteorological systems might have the effect that a wind power forecast error at a certain point in space and time could propagate to other locations during the following period. Therefore, in view of the significant installed capacities of wind power installed all over Europe today (see current status and expected developments at `www.ewea.org`), analysis and understanding of the spatio-

[1]by wind generation field is meant here a complete description of the wind power generation characteristics over a domain of interest.

temporal characteristics of wind power forecast errors are of major importance. Indeed, errors in meteorological forecasts might translate to fronts of imbalances, i.e. taking the form of a band of forecast errors propagating across entire regions. Studies on the spatio-temporal characteristics of wind fields have already been deemed as highly informative for judging the adequacy of available generation and potential reserves in the UK for instance [8]. Regarding wind power forecasting errors, a relevant analysis of the spatial smoothing effect (thus related to the analysis of the correlation of forecast errors at the spatial level only) has been performed by Focken *et al.* [9] for the specific case of Germany. However, such an analysis does not provide information on how spatial patterns in forecast errors (or of smaller/larger forecast uncertainty) may evolve in space and time. Potential benefits of spatio-temporal analysis and associated modeling of forecast errors include global corrections of wind power forecasts, associated increased knowledge on the interdependence structure of forecast uncertainty, and correspondingly improved decision-making from the provided forecasts. This may concern both wind power producers with a geographically spread portfolio, and Transmission System Operators (TSOs) managing a grid with significant wind penetration. Better understanding of spatio-temporal dependencies may also be beneficial at the planning stage, for the optimal dispatch of wind farms in order to improve the predictability of wind generation at the regional level.

One of the main goals of this paper is to make a first step in analyzing spatio-temporal propagation of the wind power forecast errors. Therefore the first objective is to demonstrate that such a spatio-temporal interdependence structure of wind power prediction errors exists. Another objective is to show how some explanatory variables, more precisely wind speed and wind direction, may affect the nature and strength of this interdependence structure, in view of the geographical layout of the wind farms. A complementary objective is to propose models that have the ability to capture such effects. The case study considered relates to the western Denmark area, for which both hourly measurements of wind power and corresponding forecasts are available over a period of several months in 2004. Forecasts of wind speed and direction used as input to the wind power prediction method used in the analysis. A detailed description of this case study and available data is given in a first part of the paper. Subsequently, classical time-series analysis tools are employed in order to highlight the spatio-temporal characteristics of the wind power forecast errors. Based on the results of this analysis, a set of models and methods is proposed with the aim of capturing the revealed non-linear behavior of forecast errors. Three types of statistical models are considered. Firstly a linear model based on observed forecast errors for various groups of wind farms is presented. It is followed by a regime-switching approach permitting to switch between different linear models, depending upon the forecasted wind direction. Finally, the effect of wind speed forecasts is accounted for by generalizing the models considered, then taking the form of conditional parametric models. Some possible directions for future research are presented and discussed in a final part of the paper.

Case Study and Available Data

Owing to its already significant share of wind generation in the electricity mix as well as very ambitious objectives in the medium term, focus is given to the test case of Denmark. This country has set the goal of having 50% of the electricity demand met by wind energy in 2025 [10], which will clearly result in challenges related to the management of the grid. More precisely, the case study of this paper relates to the western Denmark area, including the Jutland peninsula and Funen island, which is connected to the UCTE (Union for the Co-ordination of Transmission of Electricity) system and has around 70% of the entire wind power capacity installed in Denmark. Another reason for the choice of this test case is that operational developments and application of wind power forecasting systems started around 1994 in Denmark [11], and it is thus common practice today to have forecasts of wind power production at different spatial resolutions and at a state-of-the-art level of accuracy. One more reason for choosing this area for the analysis is due to the orographical particularity of the territory. Denmark has a very smooth and flat terrain, while there is in general only one prevailing weather front dominating in the whole territory at any given moment. As a consequence, our analysis of spatio-temporal dependencies in forecast errors does not require for any particular orographic or vegetation particularities to be taken into account. Note that orographic effects at the very local scale are smoothed by the grouping process.

The data selected for this work comes from 22 wind farms of different nominal capacities (details can be found in [3]), and spread throughout the area considered. For all these wind farms, measurements of wind power production with an hourly resolution are available, along with wind power forecasts provided by the Wind Power Prediction Tool (WPPT). WPPT is a state-of-the-art forecasting system. Methods included in this forecasting system are described in [12, 13] (and references therein), while application results may be found in e.g. [14]. For the present case, it provides forecasts of wind power generation for each of the wind farms with a temporal resolution of one hour up to a 48-hour lead time. Forecasts are generated every hour. The inputs for WPPT are historical power measurements at the level of the wind farm considered, along with meteorological forecasts of wind speed and direction. An initial so-called power curve model permits the nonlinear conversion of wind speed and direction forecasts to power. In a second stage, a dynamical model permits recalibration of the power curve model output to correct for potential diurnal cycles not captured by the meteorological forecasts and to adapt to local conditions by accounting for the local dynamics of the wind farm considered. Adaptive estimation of the model parameters permits accommodation of long-term variations in the wind generation process characteristics because of e.g. seasonality, or ageing of the turbines. In the present case, the meteorological forecasts used as input are provided by the HIRLAM model of the Danish Meteorological Institute (see `http://www.dmi.dk/eng/index/research_and_development/dmi-hirlam-2009.htm`). The forecasts are available over a 40×42-nodes grid (horizontal resolution is 3 km) covering Denmark

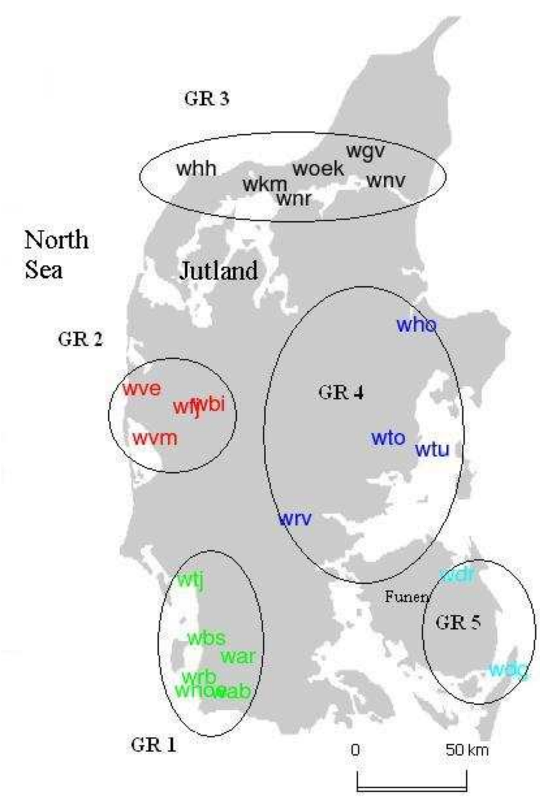


FIGURE 1: *Geographical locations of the 22 wind farms and selected groups of wind farms.*

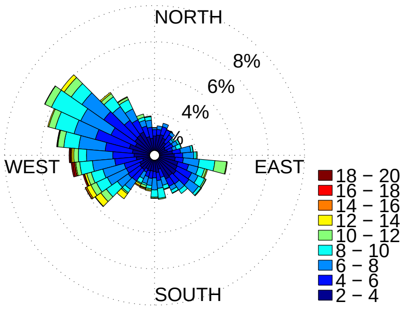
and surroundings, including a large part of the North Sea. To provide weather forecasts for each wind farm, data available at the HIRLAM nodes is respectively sub-sampled and interpolated (performed at DMI). Meteorological forecasts are delivered every six hours with an hourly temporal resolution up to 48 hours ahead. Wind forecasts are available at different vertical levels. Only wind forecasts at 10 meters a.g.l. (above ground level) have been considered here, though, due to the fact that WPPT also uses this particular level as input. The period for which both measurements and predictions have been made available for this study is from the fall of 2003 until July 2004. Since a new version of WPPT was installed in the fall of 2003, some time is needed for the model parameters to settle. Therefore, it was decided to disregard data originating from the last few months of 2003. The final dataset includes data from the first seven months of 2004.

Forecast errors are defined as the difference between power predictions and corresponding measurements, subsequently normalized by the installed wind power, following the framework described in [15]. Only one-hour ahead forecast errors are considered. The random variable corresponding to the forecast error at time t is denoted by x_t.

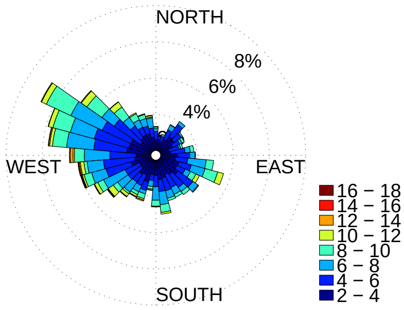
It has been chosen to study and model errors for groups of wind farms instead of concentrating on errors for each separate wind farm. This approach is preferred since spatial smoothing reduces the dependency on the local behavior and permits to focus on more on global phenomena affecting the spatio-temporal movements of forecast errors. In the first step consisting of grouping the data, both the geographical layout of the wind farms and the extensive study performed by [3] have been accounted for. Based on clustering analysis, [3] proposed to form 3 groups of wind farms. We chose to further split data from 3 groups into 5 in order to have more flexibility while accounting for different wind directions. This splitting has been performed mainly considering the geographical layout of the wind farms. An additional correlation analysis was performed. It did not play a crucial role in our decision, as it was difficult to interpret. Since the core objective of the present paper is to check whether spatio-temporal error propagation can be modeled and used for forecast improvement, not much effort has been made to optimize the grouping of wind farms. Possibly the grouping technique could be the focus of further work, and result in additional improvement of forecast performance.

The obtained groups of wind farms, along with the location of various wind farms, are depicted in Figure 1. In the following, particular attention will be given to Group number 5 (corresponding to Funen island), as a large part of weather fronts propagation over Denmark come from the North Sea (mainly from W-NW, see Figure 2 demonstrating wind-rose plots for the forecasted wind at different groups). It is then expected that the most significant spatio-temporal characteristics of forecast errors will be observed if forecast errors at Group 5 are considered as the response variable to errors observed in the other groups of the Jutland peninsula.

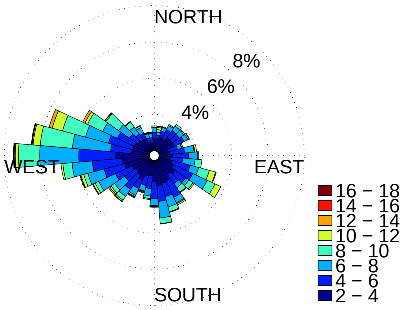
The group errors have been calculated as an average of the errors within the groups. In parallel, since the aim is to study the effects of wind speed and direction on the spatio-temporal characteristics of forecast errors, a procedure was defined for obtaining representative wind speed and direction forecasts for each of the groups of wind farms considered. Here, instead of employing a vectorial approach that would involve adding wind vectors, and then deriving average wind speed and direction from the norm and orientation of resulting vector, a more geometrical approach was used. The representative wind speed is given as the average of various wind speed values for the wind farms of the group. In parallel, wind direction at each of the farms defines the orientation of a set of unit vectors. The resultant vector then represents the wind direction for the group of wind farms. Note that these wind speed and direction data are forecasts, more precisely one-hour ahead forecasts, since focus is one-hour ahead forecast errors of wind power. To insist on this aspect, the notations $\hat{u}_t$ and $\hat{\theta}_t$ will be used for wind speed and direction forecasts, respectively. Since meteorological forecasts are updated every six hours only, these one hour forecasts are obtained by using the last relevant available meteorological forecast series.



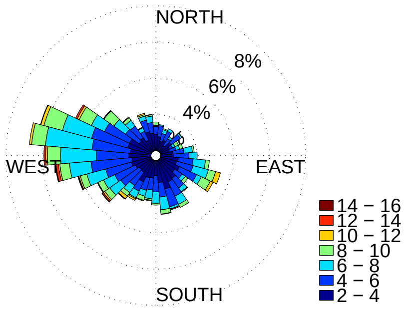
(a) Group 1



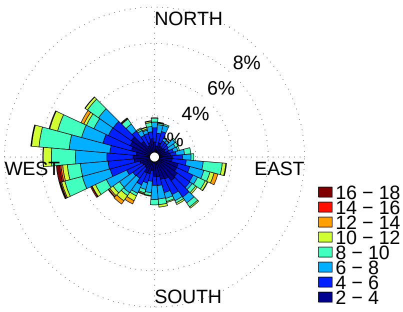
(b) Group 2



(c) Group 3



(d) Group 4



(e) Group 5

FIGURE 2: *Wind-rose plots showing the occurrences of forecasted wind speeds (m/s) and directions for each group.*

Highlighting some Spatio-temporal Characteristics of Wind Power Forecast Errors

An analysis of the available data is performed in order to reveal some of the spatio-temporal characteristics of wind power forecast errors. Such an analysis is crucial for understanding the underlying processes and for proposing a set of relevant models that would permit capture and reproduction of the various process characteristics. More precisely, the analysis performed aims at answering the following two questions:

- Is there a significant linear dependency within and between the groups, possibly with some time lag?
- Can the forecast variables, wind direction and wind speed, be used to contribute to revealing a stronger dependency?

The interest in answering the first of these two questions lies in the fact that, if linear dependency within and between groups (possibly with some time lag) is observed, it will then be straightforward to build linear models to capture such an effect. In parallel, a possible (nonlinear) relationship with some explanatory variables such as wind speed and direction forecasts could also be integrated in the proposed models with various approaches. It may appear as more relevant to study the dependency on the measured wind speed and direction, but since in real-world application such information will obviously not be available for the few following hours, forecasted values are preferred.

In the statistical literature there exists a set of standard tools that can be employed for analyzing these types of linear dependencies in datasets (for more details see e.g. [17, 18]). In further works, nonlinear dependencies could be considered as well, with methods described in [19] for instance. For the questions raised in this paper the set of necessary tools includes Auto-Correlation Function (ACF) and Cross-Correlation Function (CCF). Each of these functions describes different types of dependencies and will be briefly introduced below. The analysis is structured as follows: *(i)* firstly the dependencies within each group of wind farms are examined; *(ii)* secondly the dependencies between groups are characterized; *(iii)* finally, the effects of wind speed and wind direction forecasts on both types of dependencies are studied.

Dependency within the groups

This section investigates the effects that previous values of time-series of forecast errors (for each group) have on the current state of the group. The time-series of forecast errors for the group of wind farms j is denoted by $\{x_{j,t}\}$ where t is the time index. The following analysis is based on the ACF of the time-series of forecast errors. An assumption for its use concerns the stationarity of the process considered, which in general terms translates to the idea that process characteristics do not change with time. For more information on (strictly) stationary stochastic processes see [18]. However conclusions on significant dependencies at various time lags can be formulated even though wind power forecast errors are not strictly stationary.

The ACF in lag k for the group of wind farms j, denoted by $\rho_j(k)$, is given by

$$\rho_j(k) \; = \; \rho[x_{j,t}, x_{j,t-k}] \; = \; \frac{\mathbb{E}\left[(x_{j,t} - \mu_j)(x_{j,t-k} - \mu_j)\right]}{\sigma_j^2} \tag{1}$$

where μ_j is the mean of the time series $\{x_{j,t}\}$ and σ_j is its standard deviation. The ACF gives the correlation between the two lagged time-series $\{x_{j,t}\}$ and $\{x_{j,t-k}\}$. Therefore ρ_j takes values in $[-1, 1]$: 1 indicates a perfect positive linear dependency, -1 a perfect negative linear dependency, while 0 stands for no linear dependency at all. It is obvious that for $k = 0$ we have $\rho_j(0) = 1, \forall j$.

As focus is mainly given to Group 5, Figure 3 illustrates the ACF of the corresponding time-series of forecast errors. Qualitatively similar results have been found for the other groups, and are not discussed here. Figure 3 gives the value of $\rho_5(k)$ as a function of k, along with 95% confidence intervals under the assumption of independence for a Gaussian process. Please note, that the data actually is not absolutely Gaussian, even though it has some of its properties. Therefore the 95% intervals shown are preliminary and used only for highlighting data characteristics, but can not be fully trusted for building models. In a hypothesis testing framework, one may then reject the hypothesis of independence $\{x_{5,t}\}$ and $\{x_{5,t-k}\}$ if the value of $\rho_5(k)$ lies outside of this interval. In practice, if the value of ρ for a given lag k is clearly outside this interval, one often concludes on a significant autocorrelation for that lag.

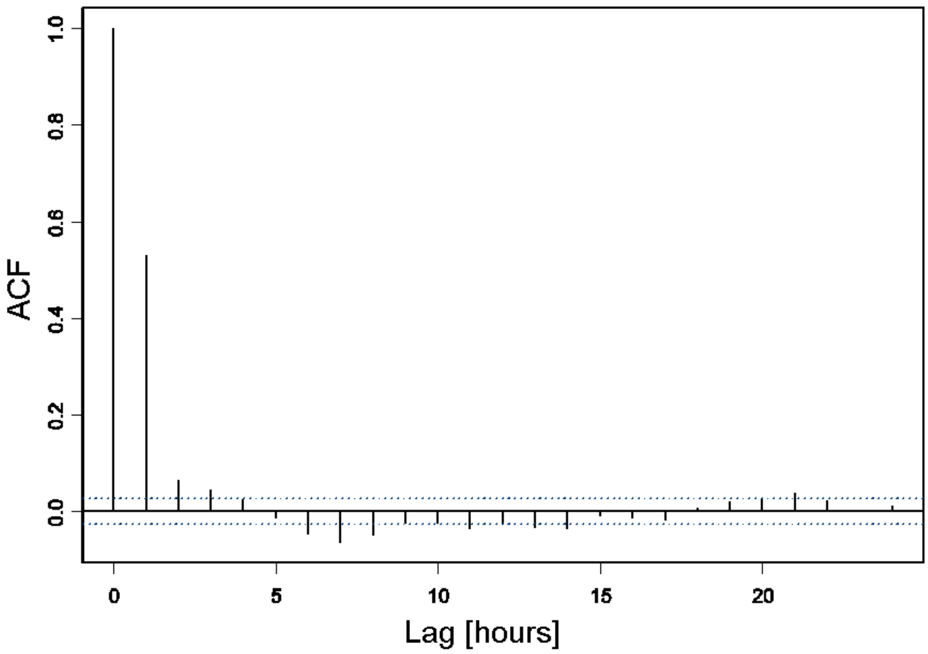


FIGURE 3: *ACF for Group 5, including a 95% confidence interval under the assumption of independence (dotted line). Values outside of this interval can be considered as significant correlation.*

From Figure 3, it can be seen that the ACF is a rapidly dampened exponential function, with a dominant autocorrelation in lag 1. The periodic waves for further lags are difficult to interpret. Non-negligible ACF values indicate that a fitted wind power prediction model (WPPT in this case) was not an ideal one, since the errors are not totally random. The better the fitted model is, the smaller the ACF values that would be observed. They would be 0 for all lags (starting from the lag 1) with an ideal prediction tool. Here, from looking at the ACF results, it is clear, that there is still room to improve prediction accuracy. In a general manner the $\{x_{5,t}\}$ time-series can be appropriately modeled with an Auto Regressive Moving Average (ARMA) model. This would translate to saying that there are two layers of dynamics in this time-series of forecast errors: a long-term inertia defining the MA part, and short-term dynamics making the AR part. It may be concluded from the Figure that there clearly are dependencies between forecast errors at different lags within a group of wind farms. However, some external signals (i.e. forecast errors for other groups of wind farms) may be related to such dependencies and this might better explain the observed behavior. This calls for further analysis.

Spatio-temporal dependencies between the groups

Once the autocorrelation pattern of forecast errors within the groups has been discussed, one can then proceed with the investigation of cross-dependencies between the groups. Information about potential cross-dependencies at certain time lags would be of great importance, since this provides crucial information for model structure identification. Demonstration of the existence of such a pattern would also translate to showing that there is spatio-temporal propagation of forecast errors between the groups.

For the case of cross-dependencies (possibly with some time-lag) between the time-series of forecast errors for the various groups of wind farms, the standard tool to consider is the CCF. The CCF between the time-series of forecast errors for Groups i and j, denoted by $\rho_{ij}(k)$, is given by

$$\rho_{ij}(k) \;=\; \rho[x_{i,t}, x_{j,t-k}] \;=\; \frac{\mathbb{E}\left[(x_{i,t}-\mu_i)(x_{j,t-k}-\mu_j)\right]}{\sigma_i \sigma_j} \tag{2}$$

where μ_i and μ_j are the mean of the time-series $\{x_{i,t}\}$ and $\{x_{j,t-k}\}$, respectively, while σ_i and σ_j are their corresponding standard deviations.

As previously described, particular focus is given to Group 5, since its geographical location and the meteorological characteristics of western Denmark make it the most interesting group to study. Group 5 is located downwind of the other groups when the wind direction is from W-NW (which is dominant for that part of Denmark). Table 1 summarizes the CCF evaluation (with respect to all other groups, and for lags between 0 and 5 hours) as well as the ACF evaluation performed above.

The cross-correlation values at lag 0 are significantly different from 0 for all groups, and this indicates that wind power forecasting errors for Group 5 have a tendency to be positively correlated with all the other groups. Furthermore, this correlation is typically higher for groups with a closer geographical location. Depending on the lag considered, the same group of wind farms does not always exhibit the highest correlation. For a time lag of one hour Group 4 shows the highest correlation, while for a time lag of two hours, Group 1 has the highest (as highlighted by the bold numbers in the Table). The forecast errors in the two other groups also have some correlation with forecast errors in Group 5 for the various time lags, though of minor magnitude. It is intuitively expected that this is due to the geographical layout of the various groups of wind farms and meteorological particularities of the area (prevailing W-NW wind), Groups 1 and 4 being the most strongly related to Group 5 (see Figure 1).

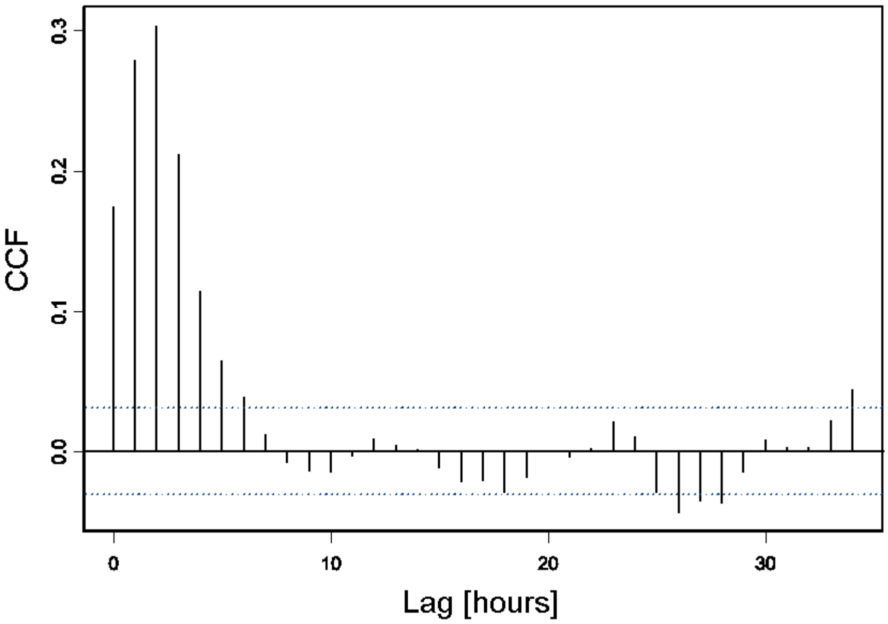
TABLE 1: *Cross- and auto-correlation for Group 5.*

		lag					
	Group	0	1	2	3	4	5
cross-correlation	1	0.175	0.282	**0.307**	0.217	0.119	0.070
	2	0.191	0.187	0.169	0.163	0.138	0.079
	3	0.1578	0.148	0.114	0.081	0.074	0.060
	4	0.2893	**0.320**	0.260	0.139	0.050	0.018
auto-correlation	5	1.000	0.527	0.059	0.040	0.019	-0.016

Owing to the dominance of Groups 1 and 4 in the observed cross-correlation patterns, it was decided to further study their dependency on Group 5. Since a visual inspection of the CCF may be more informative, the corresponding CCFs are depicted in Figure 4. The large cross-correlation values for the small lags denote the dependency on the lagged forecast error values for Groups 1 and 4 on the current forecast error at Group 5. More precisely, one retrieves the fact that for Groups 5 and 1 the highest cross-correlation is observed in lag 2, whereas for Groups 5 and 4 this peak is at lag 1. The reason is most likely due to the geographical layout, especially the distance between the groups. A closer look at Figure 4 reveals periodic oscillations in the CCF for both groups for lags larger than 6-7 hours. In line with our comment about dependencies within a group, such oscillations indicate some long-term dynamics in the forecast error process.

Dependency on wind direction

The following analysis consists of assessing how wind direction forecasts can further characterize the spatio-temporal dependencies highlighted above. As it is known that wind direction clearly affects spatio-temporal dependencies in wind power production, similar effects are intuitively expected for the propagation of forecast errors. Group 5 is chosen here again as the group of focus, while the forecast errors from the other groups play the role of explanatory variables. Note that sim-



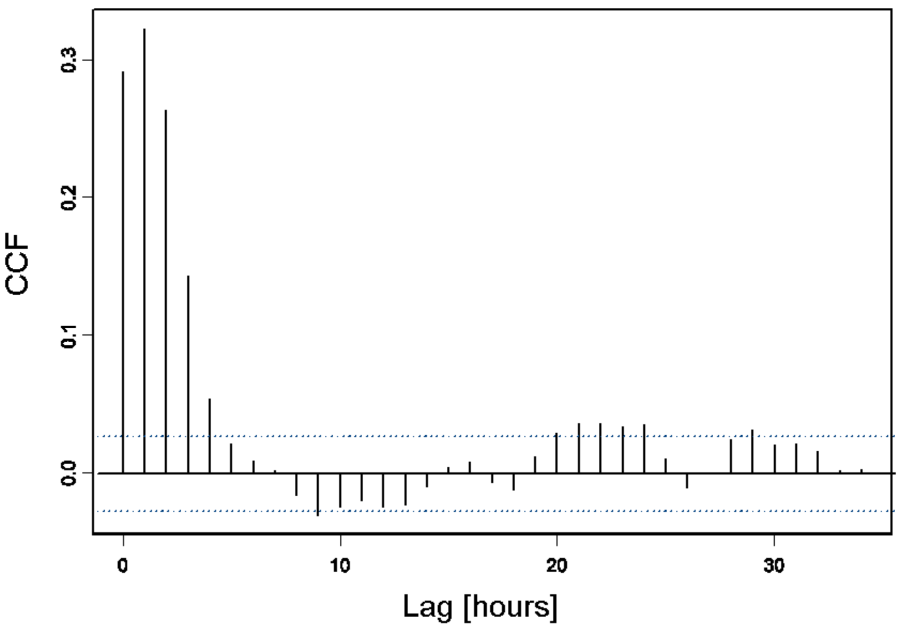


FIGURE 4: *CCF for the Groups 5 and 1 (top) and Groups 5 and 4 (bottom). Dotted lines show 95% confidence intervals under the assumption of independence. Values outside of such intervals can be considered as significant correlation.*

ilar results could be obtained from considering forecast errors in any other group as the response, potentially explained by forecast errors in the remaining ones. They would not be as significant as for Group 5, as this group is ideally situated downwind from most of the other groups.

In order to examine whether wind direction has any effect on the observed spatio-temporal dependencies between forecast errors for the various groups, the available dataset of forecast errors is divided according to the forecast wind direction in Group 5. The division is performed by constructing four intervals for potential wind directions: (0-90], (90-180], (180-270] and (270-360]. Therefore each interval corresponds to a specific sector, i.e. (0-90] the sector between North and East, (90-180] that between East and South, etc. For each of these sectors, a correlation analysis is performed between forecast errors at Group 5 and those of the other groups. This then translates to performing some kind of regime-based analysis of the spatio-temporal dependencies, the regime being defined by wind direction only. Owing to the fact that the correlation structure of forecast errors is studied conditional on the wind direction, such correlation is referred to as *directional correlation* in the following. As it is shown in the previous section, that the strongest correlation structures are between Groups 5 and 1, and between Groups 5 and 4, only corresponding results are given here. The cross-correlation values for lags ranging between 0 and 5 hours are given in Tables 2 and 3.

TABLE 2: *Directional correlation for Groups 5 and 1, for lags ranging from 0 to 5 hours.*

lag	regime			
	(0-90]	(90-180]	(180-270]	(270-360]
0	0.0457	0.1472	0.2240	0.1580
1	0.0499	0.2856	0.3597	0.2361
2	0.0672	0.3103	**0.4213**	0.2219
3	0.0358	0.1810	0.3218	0.1542
4	-0.0166	0.0985	0.2193	0.0519
5	0.0115	0.1130	0.1347	-0.0099

TABLE 3: *Directional correlation for Groups 5 and 4, for lags ranging from 0 to 5 hours.*

lag	regime			
	(0-90]	(90-180]	(180-270]	(270-360]
0	0.1390	0.3200	0.2615	0.3460
1	0.2212	0.2691	0.2570	**0.4514**
2	0.1788	0.2049	0.2075	0.3762
3	0.1288	0.1555	0.0978	0.1831
4	0.1014	0.0965	0.0158	0.0485
5	0.0252	0.0735	0.0102	-0.0157

Recall that the analysis of the spatio-temporal dependencies in Table 1 revealed that there was a maximum correlation at lag 1 between forecast errors for Groups 5 and 4, and at lag 2 between forecast errors for Groups 5 and 1. This can be seen again in Tables 2 and 3. Focusing on the correlation pattern between Groups 5 and 1, reveals that the correlation at lag 2 is at its maximum when the forecast wind direction is in the (180-270] sector (between South and West). Even for the other lags, the maximum correlation value between forecast errors for Groups 5 and 1 is attained in this wind direction regime. It then seems that the wind power forecast errors have a tendency to propagate following the wind direction. Note that there is also significant correlation at various lags for the two adjacent sectors, i.e. for wind directions originating from the sectors (90-180] and (270-360], although these are of lower magnitude. This is in line with the idea that errors in wind power forecasts are directly linked to errors in weather forecasts. In case the input meteorological forecasts are wrong, they are likely to be wrong over a part of the region considered, if not the whole region, thus leading to a non-negligible correlation of wind power forecasts among the groups. For the case of Groups 5 and 1 (Table 2) note that the values in column 2 are very similar to the ones of column 4. This could be explained by the fact that, for North-West and South-East wind, both Group 5 and Group 1 meet the weather condition at approximately the same time. None of these two groups is clearly up-wind in these two sectors. Correlation values are finally much lower for the remaining sector (wind directions in (0-90]) and this for all lags. In parallel, for the case of the correlation pattern of forecast errors for Groups 5 and 4, it is clear that in general correlation values are higher than for the case of Groups 5 and 1. This may certainly be explained by the fact that Groups 5 and 4 are geographically closer than Groups 5 and 1. Then, similar to the above, it seems that forecast errors tend to propagate following the wind direction, since the maximum correlation between Groups 5 and 4 (for a lag of one hour) is observed for wind sector (270-360], which is consistent with the geographical layout of the groups of wind farms. For this wind sector, Group 4 is located upwind of Group 5. The fact that the lag for which the maximum correlation is reached is shorter for Group 4 than Group 1 confirms the importance of distance between groups. In conclusion, it appears that the impact of wind direction on forecast errors is quite straightforward: they seem to be transported by the wind and thus propagate along the prevailing wind direction.

Dependency on wind speed

Since wind appears to be a driving force for the propagation of wind power forecast errors, another potential explanatory variable to be examined is the wind speed forecast. Indeed, as the distance between groups of wind farms seems to play a significant role, wind speed should also make the propagation of forecast errors slower or faster. This holds even though the speed of the error propagation is not necessarily the same than the forecasted wind speed, as the speed/direction of atmospheric features might be different with the surface wind speed/direction. In order to study the potential effect of wind speed, our strategy is to divide the dataset depending on wind speed, and to analyze the correlation pattern of forecast errors (as a function of the lag). This is done for each wind sector individually, as the impact of wind speed may be more significant for the wind sector that exhibits the clearest interdependence of forecast errors. The propagation of forecast errors may then be seen as a finite impulse response conditioned by wind speed. As an example, focus is given here to the correlation pattern between Groups 5 and 1, and for wind sector (180-270], for which Group 1 is located directly upwind of Group 5. The dataset is divided into five smaller datasets depending on the (forecasted) wind speed in Group 5. The five wind speed intervals considered are (in m/s): 1 - [0,4), 2 - [4,6), 3 - [6,8), 4 - [8,10), 5 - [10,25). The CCF is then calculated for each of these wind speed intervals, and for lags between one and seven hours. Figure 5 then illustrates how the correlation pattern varies depending on the wind speed. Note that this analysis has some restrictions, as the number of observations is not the same among the intervals. Each of them contains 200-400 data points, which makes the results significant, but not straightforward to compare, since the level of significance for the estimates differs from interval to interval. However, it may still allow us to observe some general features that would be explained by the wind speed level.

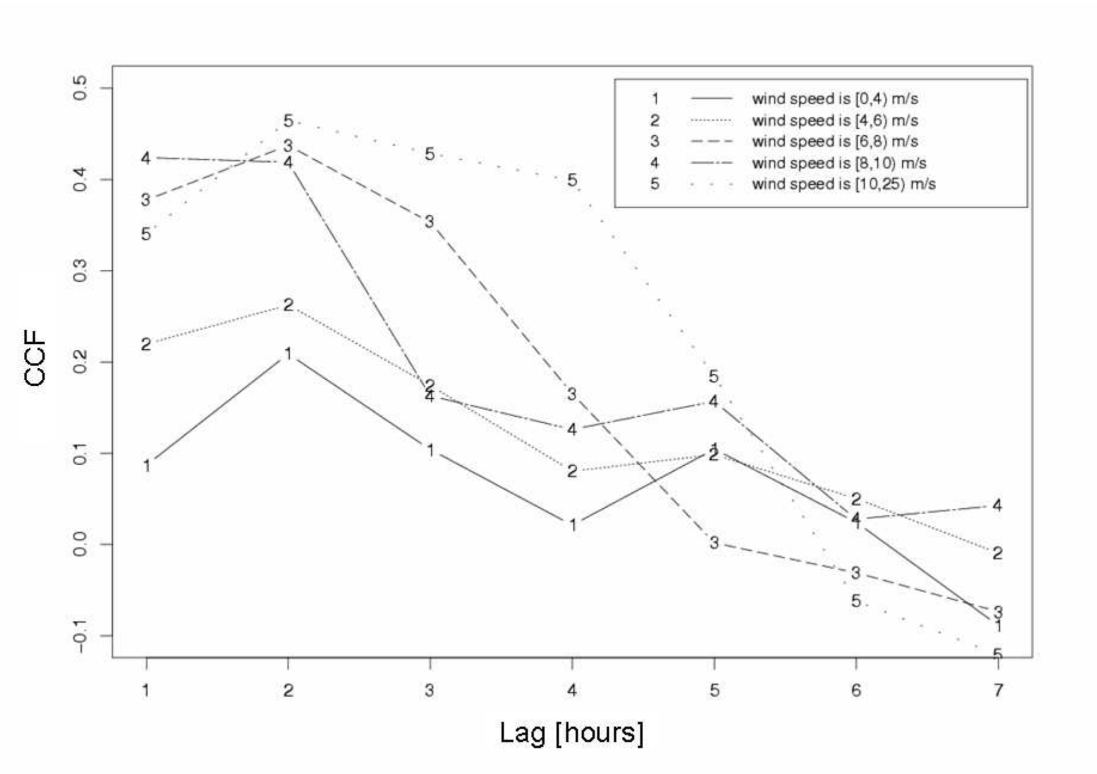


FIGURE 5: *Cross-correlation between forecast errors for Groups 5 and 1 and for the wind sector (180,270]. Cross correlation is given for different wind speed levels, and as a function of the lag.*

Note from Figure 5 that there is a general trend such that the average correlation of forecast errors between Groups 5 and 1 (and for the wind sector considered) increases as the wind speed gets larger. Indeed for low wind speeds, forecast errors may be mainly due to local phenomena, and thus do not propagate at all to the neighboring groups of wind farms. For higher wind speeds, one retrieves the finite impulse behavior mentioned earlier, with wind speed directly influencing the magnitude of the correlation between forecast errors, as well as the lag for which this correlation is maximum. This particular lag is of 1-2 hours. Note that the high correlation for lags up to 5 hours in the case of high wind speeds (i.e. here between 10 and 25 m/s) may be due to very large and long-lasting discrepancies e.g. phase shifts between forecasts and measurements. Such phase shifts correspond to timing (or phase) errors in the forecasts and directly translate to clusters of errors of significant magnitude with the same sign, thus increasing their observed autocorrelation. This phenomenon is more common for higher levels of wind speed, in relation to meteorological fronts crossing the area, and to significant ramping in wind power generation. In a general the effect of wind speed on the propagation of wind power forecast errors appears to be more difficult to perceive than that of wind direction and is clearly nonlinear (Figure 5). If the dependence was linear, the shape of the CCF for different wind speed intervals would be the same with a possible shift in the dominant lag or potential linear deformation of the CCF. This nonlinear effect can also be seen for other wind sectors and other groups (though not shown and commented on here), even though this effect is also conditioned by the geographical layout of the groups of wind farms, mainly their respective positioning and the distance between them.

Proposal of Relevant Models

In the above analysis, it has been demonstrated that wind power forecast errors indeed have some spatio-temporal characteristics at the level of western Denmark, and that this propagation of forecast errors is also affected by wind speed and direction. Our objective in this section is then to propose a set of relevant models that may be used to capture and reproduce the observed behavior of forecast errors. Remember that focus is here is on one-hour ahead forecast errors here but that an analogous methodology could be applied for the modeling of forecast errors related to further look-ahead times. Since the most significant correlation patterns observed over the whole data analysis are those obtained when concentrating on forecast errors at Group 5, it is decided to concentrate on this case. The overall methodology and set of models are thus introduced for this specific case, though they could be similarly derived if considering other groups of wind farms. The one-hour ahead forecast errors at Group 5 are seen as the response variable and denoted by y_t (instead of $x_{5,t}$), while the explanatory variables, which consist of the one-step ahead forecast errors in Groups 1 to 4, are denoted by $x_{1,t}, \ldots, x_{4,t}$, t being the time index. The notation for the errors in Group 5 is here changed in order to make it easier to see the difference between response and explanatory variables.

Three types of models appear to be relevant for modeling the observed spatio-temporal characteristics of the wind power forecast errors. Firstly since there is some significant linear correlation between forecast errors at different time lags and for different groups, a straightforward starting point is to use Auto-Regressive models with eXogenous input (referred to ARX models in the following). Such models also comprise a natural benchmark against which more complex models should be evaluated. Indeed, for capturing the effect of wind direction on the spatio-temporal characteristics of forecast errors, it is proposed in a second stage to use a regime-switching approach. Such a regime-switching approach will permit switching between different ARX models, depending on the forecast wind direction. Finally, the more complex effect of wind speed on these spatio-temporal characteristics is accounted for by upgrading ARX models to conditional parametric models in each wind direction regime, thus making the coefficients of the model a nonparametric function of wind speed. In all cases, it is assumed that the time-series considered have stationary properties. This assumption may be relaxed in the future, and model coefficients may be adaptively estimated in an estimation framework including exponential forgetting.

Linear models

As the most simple linear model to be employed for the modeling of one-hour ahead forecast errors, one may think of a simple AutoRegressive (AR) model. However, since our aim here is to consider the spatio-temporal effects highlighted above, it appeared more relevant to also account for some explanatory variables, namely the one-hour ahead forecast errors observed in the other groups and for different points in time in the past. This then led to the building of an ARX model. For more information related to the theory behind the building of ARX models, we refer to [17, 18, 20]. The general structure of an ARX model is given by

$$y_t = \beta_0 + \sum_{l=1}^{p} \beta_l y_{t-l} + \sum_{i=1}^{n} \sum_{j=1}^{k_i} \beta_{i,j} x_{i,t-j} + \epsilon_t \tag{3}$$

where the response variable y_t is linearly explained by its p previous values in the auto-regressive part, and by n external input variables, each up to lag k_i ($i = 1, \ldots, n$). ϵ_t is a purely random variable with zero mean and finite variance, which represents the noise that cannot be explained by the model.

The estimation of ARX model parameters can be straightforward performed with Least Squares (LS) estimation methods. Again, extensive details and discussion on this topic can be found in [17, 18, 20]. The procedure employed for selecting the input explanatory variables and their lags is detailed in [16]. Finally, the model structure obtained is the following

$$y_t = \beta_0 + \sum_{l=1}^{7} \beta_l y_{t-l} + \sum_{i=1}^{3} \beta_{1,i} x_{1,t-i} + \sum_{j=1}^{2} \beta_{4,j} x_{4,t-j} + \epsilon_t \tag{4}$$

meaning that the current one-hour ahead forecast error in Group 5 can be explained by a linear combination of its last 7 values, in addition to last 3 forecast errors made for Group 1 and the last 2 forecast errors made for Group 4. This is consistent with the results from the analysis of the spatio-temporal characteristics of forecast errors performed in the previous section. Note that number of lags used in the model is different from the number suggested by Figures 3 and 4. Information from those figures is only used as a first step towards understanding and highlighting data characteristics. Final model structure is decided on the basis of Akaike's Information Criterion (AIC) and optimization of determination coefficients (see [16] for exact details) in order to achieve the best possible performance of the model.

Regime-switching models based on wind direction

Here we recall the idea of what we defined above as directional correlation, which was used in the data analysis performed above. The main purpose of defining such directional correlation is to analyze and model the effect of wind direction on the spatio-temporal dependencies of forecast errors. We claim that if the wind direction is compatible with the direction of the vector having its beginning in a given group of wind farms and ending in another group of wind farms, then the dependency between errors for these two groups (possibly with some lag) should be higher than in case of different directions.

Regime-switching models extend the idea of linear models by having a set of linear models, each of them being active in a certain regime. The switch between regimes can be governed by previous values of the response variable, external signals or unobservable stochastic processes. Here, focus is on the second type of regime-switching models as the regime switches will be governed by the wind direction forecast in Group 5. Regimes are defined by threshold values for the wind direction variable $\hat{\theta}_t$. These thresholds correspond to the upper bounds of the intervals in which the given 'sub-model' is active. The corresponding models employed may then be referred to as Threshold AutoRegressive with eXternal input (TARX) models. This type of regime-switching model has initially been introduced in [21], and extensively described in [22]. For the specific case of the wind power application, basic concepts of regime-switching modeling may be found in [23].

The potential range of values for the wind direction variable $\hat{\theta}_t$ is $\mathcal{I} = (0, 360]$. Define intervals $R_1 \cup \ldots \cup R_k = \mathcal{I}$ such that $R_i \cap R_j = \emptyset, i \neq j$. Each interval is given by $R_i = (r_{i-1}, r_i]$. The values $r_0, ..., r_k$ are the so-called threshold values which define switches between regimes. The threshold values are in general to be estimated from the data. However here, we consider the case when the values are known in advance, since they have been derived from an analysis of the data similar to that performed above. The motivation for such an assumption is that we analyze the case for which the regimes are governed by wind direction. From physical knowledge and intuition about the process characteristics, the choice of regimes may be fairly straightforward. The general form of the models examined further is

$$y_t = \beta_0^{(s_t)} + \sum_{l \in L_y^{(s_t)}} \beta_l^{(s_t)} y_{t-l} + \sum_{i=1}^{4} \sum_{j \in L_{x_i}^{(s_t)}} \beta_{i,j}^{(s_t)} x_{i,t-j} + \epsilon_t \tag{5}$$

where

$$s_t = \begin{cases} 1, & \text{if } \hat{\theta}_t \in R_1 \\ 2, & \text{if } \hat{\theta}_t \in R_2 \\ \vdots & \\ k, & \text{if } \hat{\theta}_t \in R_k \end{cases} \tag{6}$$

In the above, $\hat{\theta}_t$ serves as the external signal which determines regime switching, t being the time index. In parallel, y_t is the response variable i.e. the one-hour ahead forecast errors at Group 5, the $x_{i,t-j}$ are the forecast errors for Group i and at lag j, and $\{\epsilon_t\}$ is zero mean white noise. $L_y^{(s_t)}$ and $L_{x_i}^{(s_t)}$ are sets of non-negative integers defining the auto-regressive and input lags (for Group i) of the model. The superscript (s_t) indicates that these sets of integers may be different for each of the regimes, i.e. along for different model structures depending on wind direction. The $\beta_{j,i}^{(s_t)}$ coefficients are the linear coefficients to be estimated in each regime s_t. Since the thresholds are known, the estimation problem for TARX models is solved by fitting different linear models to the data in each of the regimes. The estimation method to be employed is described in detail in [23].

For the test considered in the present paper, after analysis of the data in order to split it into various wind direction regimes, and then in each regime in order to identify the structure of the linear models (for more details, see [16]), the following general structure of the TARX model was obtained. First of all, the regimes are defined as follows:

$$s_t = \begin{cases} 1, & \text{if } \hat{\theta}_t \in (0, 90] & \text{(North-East sector)} \\ 2, & \text{if } \hat{\theta}_t \in (90, 180] & \text{(East-South sector)} \\ 3, & \text{if } \hat{\theta}_t \in (180, 270] & \text{(South-West sector)} \\ 4, & \text{if } \hat{\theta}_t \in (270, 360] & \text{(West-North sector)} \end{cases} \tag{7}$$

Originally, the choice for these regimes was dictated by easiness of interpreting the effect of wind direction forecasts which in this case is compatible with geographical cardinal directions. In fact, other divisions of the range of wind direction values were studied, and the improvement in model fit was considered insignificant or none. In a second stage, focus is on the linear models to be fitted in each of the regimes. The optimal number of lags is selected separately for each of the regimes. Table 4 describes the structure of the resulting TARX model. While building the models, AIC was used to decide on the final number of lags used. The choice of the variables seems to be reasonable if the position of the groups of wind farms is taken into account (see Figure 1). For example, the sector

(270,360] corresponds to situations with the wind direction forecast from the North-West sector and in this case the effect on Groups 1 and 4 is seen to be most significant. Also, the maximum lags taken for Groups 1 and 4 conform with the directional distance from Group 5 in this regime. By the directional distance in this case we consider a projection of the distance between the corresponding groups on the axis following the middle wind direction of the current regime (e.g. equal to 315 for regime 4).

TABLE 4: *Threshold model structure: number of lags in the autoregressive part of the model, and selected lags for each of the other groups.*

s_t	AR	Group 1	Group 2	Group 3	Group 4
1	10	-	-	4th	1st
2	5	1st	-	-	1st
3	6	1st, 2nd, 4th	-	-	1st
4	6	1st and 3rd	-	-	1st

Conditional parametric models with regime-switching

It is now aimed at upgrading the previous regime-switching model by integrating the complex non-linear effect of wind speed on the spatio-temporal characteristics of forecast errors. The underlying idea is that the time delay for the propagation of errors is directly linked to wind speed. For the purpose of accounting for such an influence, it is proposed here to transform the linear models in each of the regimes of the TARX model described above by conditional parametric models. Conditional parametric models comprise a class of models with a linear structure (like an ARX model), but for which the linear coefficients are replaced by smooth functions of other variables. For an extensive description of conditional parametric models, we refer to [24, 25].

More specifically, it is chosen to employ conditional parametric ARX models in each of the regimes in order to obtain a conditional parametric regime-switching approach. A conditional parametric ARX model with the model coefficient being smooth functions of wind speed, and with regime switches based on wind direction, can be written as

$$y_t \;=\; \beta_0^{(s_t)}(\tilde{u}_t) + \sum_{l \in L_y^{(s_t)}} \beta_l^{(s_t)}(\tilde{u}_t) y_{t-l} + \sum_{i=1}^{4} \sum_{j \in L_{x_i}^{(s_t)}} \beta_{i,j}^{(s_t)}(\tilde{u}_t) x_{i,t-j} + \epsilon_t \tag{8}$$

where the regime switches with respect to wind direction forecast are governed by (6). $\{\epsilon_t\}$ is a white noise sequence, i.e. a sequence of independent and identically distributed random variables with zero mean and finite variance. In addition, as in the case for the simpler TARX models introduced above, $L_y^{(s_t)}$ and $L_{x_i}^{(s_t)}$ define the model structure (i.e. the lags to be considered), with the superscript

(s_t) indicating that these sets of integers may be different for each of the regimes. Note that for simplification and for direct comparison with the results that will be obtained with TARX models, the structure of the conditional parametric regime-switching model is defined similarly to the TARX model described above, that is, by (7) for the regime switches, and by Table 4 for the model structure. In the following, conditional parametric models with regime-switching will be abbreviated as CP-TARX models.

Then, in contrast to the TARX models, the $\beta_{j,i}^{(s_t)}$ coefficients are smooth functions of a representative wind speed $\tilde{u}_t$ (discussed below). Since the thresholds on wind direction are known, the estimation problem simplifies to the independent estimation of a conditional parametric model in each of the regimes. For this purpose, the LFLM (Local Fitting of Linear Model) software developed at the Technical University of Denmark [26] is employed. For an extensive description of the estimation methods involved, we refer to [13]. Coefficient functions have been locally approximated with first-order polynomials, for a number of 150 fitting points uniformly spread over the range of potential wind speed values. Tricube kernels have been chosen, with a nearest-neighbor bandwidth covering the 40% wind speed data closest to each fitting point, allowing smooth local estimates of the coefficient functions.

The variable $\tilde{u}_t$ in equation (8) is a filtered wind speed at time t, which is representative of the wind field potentially affecting forecast errors at the group of wind farms considered. Firstly, it was decided to take wind speed in Group 5 ($\hat{u}_{5,t}$) at time t as the representative wind speed $\tilde{u}_t$. Of course, in this case the information about wind speeds in other groups was lost and not accounted for by the model. It was therefore decided that $\tilde{u}_t$ should be a summary of wind speed information at all the groups included in the model, based on a filter employing weighted linear regression. The weights were selected according to the corresponding coefficients of a linear regression of y_t on the errors from the other groups included in the model. For instance, assume that we want to explain y_t using $x_{1,t-2}$ and $x_{2,t}$. Then, in order to obtain the representative wind speed, the linear model $\tilde{u}_t = a\hat{u}_{1,t-2} + b\hat{u}_{2,t}$ is employed, where $\hat{u}_{1,t-2}$ and $\hat{u}_{2,t}$ denote wind speeds in Group 1 at time $t-2$ and in Group 2 at time t, respectively. The coefficients a and b in this model are the weight coefficients estimated for the model $y_t = ax_{1,t-2} + bx_{2,t}$. Such representation of a forecasted wind speed showed a better model performance in terms of R^2, therefore was chosen for the further analysis. Note that filtered wind speed values certainly are different from the forecasted wind speeds. The values for the filtered wind speed range between 0 and 5 m/s.

Application Results

The objective of this section is to illustrate and analyze the ability of the various models presented above to capture the spatio-temporal characteristics of wind power forecast errors, as well as the effects of both wind speed and direction on those characteristics. The modeled errors are subtracted from the original forecasts issued by WPPT in order to get the forecasts adjusted after consideration of spatio-temporal dependencies. The accuracy of these adjusted forecasts is compared to that of original WPPT forecasts based on two different criteria. Here, models are fitted on the dataset considered, which has a limited size (seven months). Ideally, one year or more of data would be preferred. In addition, since regime-switching uses different models for each regime, this further reduces the amount of data used for estimation of model parameters. In order to optimally use this limited dataset, the approach employed is firstly to fit the various models to the whole dataset (seven months from the year 2004), with the aim of evaluating their ability to capture the effects highlighted in the previous section. In a second stage, a cross-validation exercise allows us to comment on the generalization ability of the models, i.e. on their potential ability to reproduce observed and modeled effects if trained and used on different data.

Model fitting

Comparison is made between the linear ARX model of equation (4), the regime-switching TARX model of equation (5), and the regime-switching conditional parametric CP-TARX model of equation (8). The structure of the last two models is detailed in Table 4. Remember that the linear ARX model only accounts for autoregressive effects and linear effects from neighboring groups of wind farms, while the TARX model additionally accounts for the dependency on wind direction and the CP-TARX model aims at capturing the dependency on both wind speed and direction.

For evaluation of the fit of the various models, two criteria are employed. On the one hand, the coefficient of determination R^2 tells how much of the variations in the wind power forecast errors at Group 5 are explained by the models. Its value is between 0 and 1, 1 being a perfect power of explanation. It may then conveniently be expressed in percentage units. On the other hand, it is chosen to employ the Root Mean Square Error (RMSE) criterion. The RMSE is a quadratic error measure, thus giving more weight to large residuals and being in line with idea of LS fitting of the models. It is given here as a percentage of the installed capacity of Group 5. For more details on evaluation of statistical model fitting, we refer to [20, 27], and also to [15] for the specific case of the wind power application. Table 5 gathers the corresponding results. Note that RMSE values have been calculated for the same data set on which the model parameters have been estimated, thus informing about the quality of the fit of the models in a LS sense. It may therefore be that the higher ability of some of the models to better explain the errors come from some form of overfitting. This will be discussed in more details and accounted for in the following subsection, when performing a cross-validation exercise.

The linear ARX model already has a certain ability to explain variations in wind power forecast

TABLE 5: *Evaluation of the fitting of the various models over the whole dataset.This evaluation is based on the coefficient of determination R^2 and on the error criterion RMSE. The regime is determined by the wind direction forecast.*

Regime	ARX model		TARX model		CP-ARX model	
	R^2 [%]	RMSE [%]	R^2 [%]	RMSE [%]	R^2 [%]	RMSE [%]
1	-	-	38.4	7.5	48.5	5.3
2	-	-	46.1	4.4	48.3	4.3
3	-	-	49.3	6.8	55.6	6.4
4	-	-	54.9	5.6	57.7	5.4
Overall	47.8	5.8	49.9	5.7	54.2	5.4

errors at Group 5, since it has an R^2 of 47.8%. However in a general manner, this ability is increased by accounting for the effects of wind speed and direction. Indeed, TARX and CP-TARX exhibit higher values for the coefficient of determination, reaching 49.9% and 54.2%, respectively. The overall RMSE values for these two models are also lower than for the linear ARX model, with a non-negligible advantage for the more complex CP-TARX model. For comparison, the RMSE for one-hour ahead forecasts for this group of wind farms is 11.67% of nominal capacity before application of the various models studied here. Note that the fairly high level of original prediction error may be explained by the fact the nominal capacity for Group 5 is small. Such reduction in the RMSE criterion means that whatever the type of model chosen, the most reduction in forecast errors actually comes from the initial idea of accounting for spatio-temporal effects, while going for complex models, including wind speed and direction, mainly allows for better performance in certain meteorological conditions. Indeed, going into more detail, one notices differences in the values of evaluation criteria among the various regimes. These differences may be due to the more or less appropriate structures of the (sub)models in each regime, or due to different amounts of data used for model fitting, as well as different inherent predictability levels in various meteorological conditions. Since wind primarily blows from western directions over this region, a large share of the data available corresponds to regimes 3 and 4. Differences among regimes are of higher magnitude for the TARX model, with the CP-TARX model always having higher R^2 values as well as lower RMSE. The most significant improvements are observed for regimes 1 and 3, corresponding to the North-East and South-West direction, and for which the model structure is quite different. In the former case, the model mainly has an autoregressive pattern, while for the latter case the model has a lighter autoregressive pattern and relies more on past forecast errors at Group 1. This confirms a general interest of having the model coefficients as a function of wind speed.

After verifying that model residuals are not correlated, a bootstrapping technique (following the framework introduced in [30] and more specifically the functions described in [26]) is applied to check the level of uncertainty associated with the estimates of model coefficients. As an example,

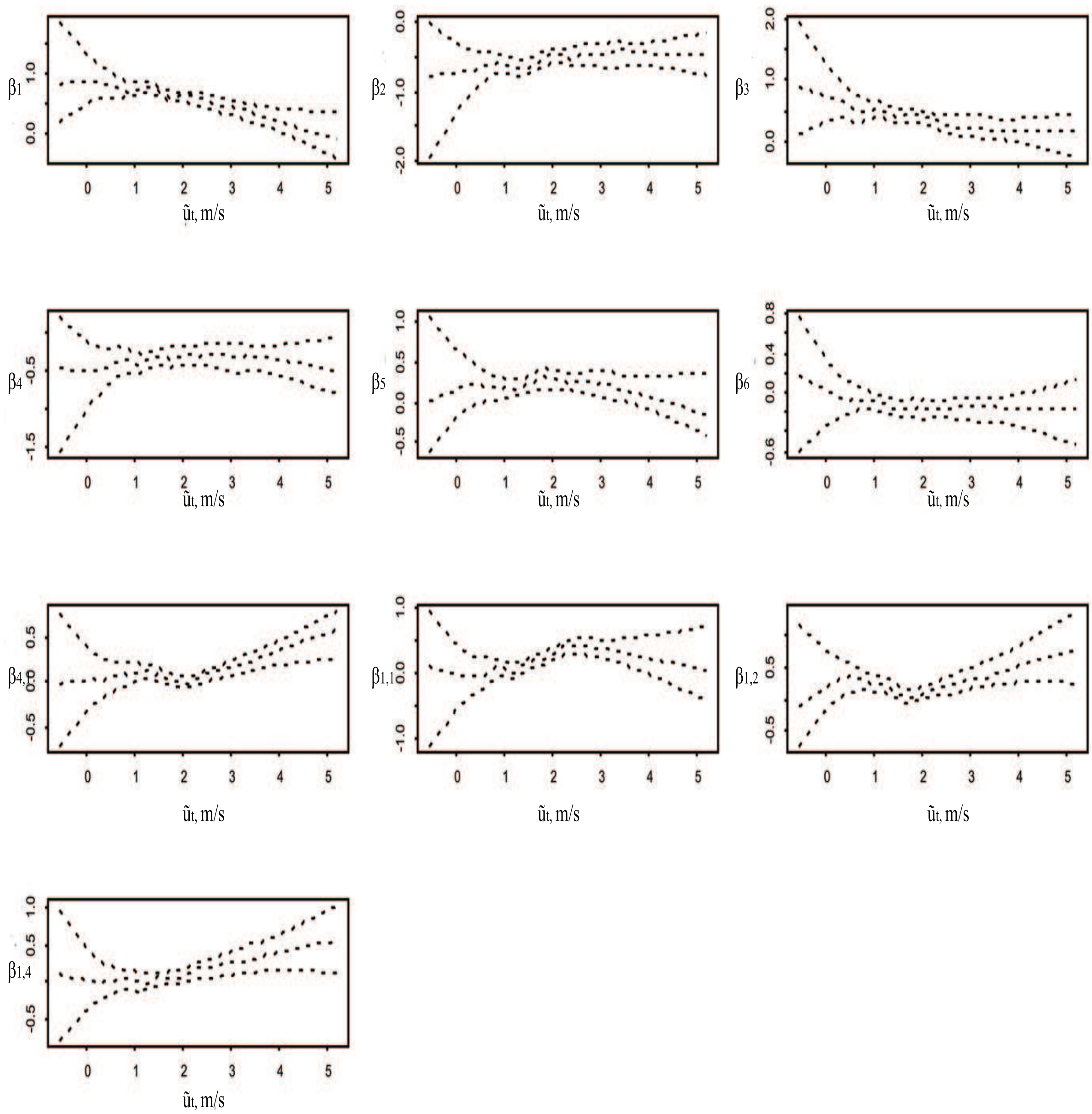


FIGURE 6: *Coefficients of the CP-TARX model (see equation* (8)*) fitted in regime 3, along with 95% confidence intervals based on 200 bootstrap replicates.*

the results for regime 3 are shown in Figure 6. Results for other regimes are qualitatively similar, and not discussed here. They are extensively commented on in [16]. A first interesting point with Figure 6 is the noticeable evolution of the model coefficients as a function of the wind speed level. One sees for instance that as the wind speed level increases, there is a general trend that the autoregressive coefficients get closer to zero, while the coefficients values for the different lags in forecast errors at Groups 1 and 4 globally increase. This observation is mainly based on the results from the area where the bootstrap confidence intervals are narrow enough to make it possible to conclude on the behavior of coefficients, i.e. where the wind speed level is between 1 and 3 m/s.

In the areas corresponding to very low or high wind speed levels, due to the lack of data, the confidence intervals are broad, prohibiting determination of coefficient behavior. This translates to saying that the higher the wind speed, the larger the effect of upstream information (from Groups 1 and 4 in regime 3, see map in Figure 1) and the less significant is the autoregressive pattern. Such a behavior may actually be fairly intuitive: as a wind front is stronger and moves faster, it possibly could transport forecast errors and dominate over local effects, which in contrast may be the main source of forecast errors for calm periods (thus corresponding to low wind speeds). In parallel, the impact of the distribution of representative wind speed values on the uncertainty of model coefficients is visible: as filtered wind speed values are more concentrated between 1 and 4 m.s^{-1} the 95% bootstrap confidence intervals are fairly tight, while they get wider for representative wind speed values outside of this range. This uncertainty in the value of the model coefficients directly relates to an insufficient amount of data available for those intervals. As may be noticed from the Figure, coefficient functions actually prolong for representative wind speed values below zero. This is due to the estimation method employed, and does not mean that representative wind speed values below zero may be encountered.

Cross-Validation and generalization ability

In the model-fitting exercise carried out above, both R^2 and RMSE measures have been calculated for the same data for which the model parameters have been estimated. Therefore, the higher ability of some of the models to better explain the variations of prediction errors may come from a numerical artifact, namely the so-called over-fitting. As a consequence, in order to verify if the models would perform similarly if applied to new (unseen) data, a cross validation procedure is employed. The idea of cross-validation is to use a subset of data for estimation of the model parameters, while the other subset is employed for model evaluation. More precisely, 3-fold cross validation is applied (as described in [17, 20]). The data in each regime is divided into three equal subsets. Two of the constructed subsets are used for parameter estimation and the third subset is used for checking the model performance. By repeating the procedure three times, one obtains three different estimations of the model parameters, with corresponding evaluation on independent subsets. The results are presented in Table 6, with a focus on regimes 3 and 4 only, since these are deemed as more interesting in the above analysis, owing to the higher amount of data available, and better performance of the fitted models. Also, emphasis is on the TARX and CP-TARX models, since the effect of wind speed on the spatio-temporal characteristics of forecast errors is more complex, and the way conditional parametric models permit (or not) to capture them should be verified.

Cross validation results show that TARX models seem to have better generalization ability than the more complex CP-TARX models. R^2 values for the three subsets for TARX models are fairly stable and at an almost similar level as for fitting performed on the whole dataset in the previous section

TABLE 6: *3-fold cross validation results for both TARX and CP-TARX models in regimes 3 and 4.*

Model (regime n^o)	subset1		subset2		subset3	
	R^2 [%]	RMSE [%]	R^2 [%]	RMSE [%]	R^2 [%]	RMSE [%]
CP-TARX model (regime 3)	47.8	5.89	40.4	7.78	49.2	8.03
TARX model (regime 3)	48.1	5.93	47.8	7.20	47.5	7.96
CP-TARX model (regime 4)	52.5	5.58	51.6	5.81	57.0	6.89
TARX model (regime 4)	53.4	5.43	52.9	5.56	55.3	6.60

(see Table 5). RMSE values exhibit higher differences though. In parallel, the cross validation exercise for CP-TARX models yields more significant differences in both RMSE and R^2 from one evaluation subset to the other, with a significant decrease in R^2 if compared to the model-fitting results of Table 5. And, in a general manner forecast accuracy for the TARX models is slightly better than that of the CP-TARX models, while this was not the case for using the whole data set above. Such results may be interpreted as a higher generalization ability of TARX models in comparison to the CP-TARX models. However, it is important to note the limited amount of data used in the present study. As already mentioned, the available data covers a period of seven months only. The number of observations in regimes 3 and 4 for this period is 1536 and 1640, respectively. When it comes to the cross validation exercise, each of the constructed subsets includes data from 2.33 months period only, making the number of observations available for the estimation step drop to around 1000 (which is 2 subsets or 4.66 months) for the specified regime. Taking into account that each set also has to be divided according to wind direction and that different wind speed levels have to be considered, it is likely that this seven-month period is actually not sufficient to draw final conclusions on the ability of CP-TARX models to capture the spatio-temporal characteristics of forecast errors accounting for the effects of wind speed and direction. Results obtained for the TARX models may appear as more trustworthy as they are based on more data for each wind regime considered (since no division according a wind speed is needed), but this may not be still the case if extending this study to longer periods.

Considering the data in regimes 3 and 4 only, the RMSE values obtained if averaging over the 3 cross-validation subsets are of 6.71% and 6.49% for the CP-TARX and TARX models, respectively. For comparison, the RMSE value calculated on this same data subset for regimes 3 and 4, but before applying any of the spatio-temporal models, is of 8.99%. This shows that both of the presented models can significantly reduce the forecasting errors of the state-of-the-art prediction tool.

Conclusions and Perspectives

The present paper can be seen as the first step towards understanding and capturing the complex nature of spatio-temporal propagation of wind power forecast errors. The test case of the western Denmark area is of particular relevance, in view of the significant installed wind power capacities spread over this region, and of the resulting management challenges for the TSO or for power producers with a geographically spread wind portfolio. A thorough analysis of the available forecast and measurement data has permitted formulation of a set of important conclusions. Such conclusions go along the line of our main objective, which is to show that there clearly exists some spatio-temporal patterns in the characteristics of wind power prediction errors. First of all, there exists in general a significant cross-correlation between forecast errors for neighboring areas with lags of a few hours. For the present case study, lags with significant dependency are up to five hours, while the lags with most effect are the one and two-hour lags. This cross-correlation pattern is clearly conditioned by the prevailing weather situation, mainly characterized by wind speed and direction. Wind direction is shown to play a crucial role, while the effect of wind speed is more complex. Prevailing wind speed affects the dependency in the following way: the higher the wind speed the stronger the dependency on more remote places; while in case of lower wind speeds, more influence comes from a local origin (thus exhibiting an autoregressive pattern).

In terms of modeling, this means that the dependency on wind direction may be easily accounted for by state-of-the-art regime-switching approaches, while dependency on wind speed should be captured by more complex models. This has been performed here by embedding conditional parametric models in the regime-switching approach. The superiority of such a proposal for capturing the complex effect of wind speed has not been demonstrated, possibly because of the limited size of the available dataset (only data from a seven-month period was available). The best spatio-temporal model proposed has been shown to explain up to 54% of one-hour ahead wind power forecast errors in terms of R^2. Finally when applied to new, “unseen” data, the regime-switching models permitted to reduce the forecast error level from the initial 8.99% to 6.49%, in terms of the RMSE criterion.

Note that owing to the choice of such a short look-ahead (1 hour ahead), forecast errors may be due to large ramps in wind power generation, which are difficult to predict when a strong weight is given to the past few power measurements (as is done by a state-of-the-art model like WPPT for forecasts up to ca. six-hours ahead). The various potential origins of the forecast errors do not alter the interest of the proposed approach, since they involve statistically characterizing spatio-temporal patterns in forecast errors, and subsequently taking advantage of this knowledge for forecast correction. The proposed analysis and methodology could also be extended to the case of errors for further look-ahead times (up to several hours ahead) if working on the same terrain as Denmark. In order to make models valid for data coming from a larger region or from a region with a more complex terrain than Denmark, some adjustments would have to be done in the modeling approach due to the fact that it is not always possible to use one prevailing wind speed or direction as a representative of the situation in the whole region. The methodology presented in this paper could be evaluated for such more complicated cases. Possibly if working with a larger region and small time-lags, the region could be divided into sub-regions and each sub-region could be analyzed separately. If considering further look-ahead times (more than several hours ahead), then data from a larger region should be considered along with a thorough examination of the weather forecasts, in order to evaluate how weather fronts normally move along the region and which parts of the entire region may affect each other at the time scales considered. Such a generalization of the proposed methodology might not be straightforward. We believe, however, that the principles introduced for highlighting spatio-temporal characteristics of forecast errors, model building and estimation, can be seen as generic in future related work.

For a small area like western Denmark, which is the first to be touched by fronts coming from North-West, the use of online measurements from the United Kingdom, or from measurement devices in the North Sea, might lead to highly significant improvements on a longer time horizon. In parallel, considering the number of turbines spread over western Denmark, it appears crucial to propose a modeling approach that would allow for dynamic evolution of the overall wind installations. Indeed new wind farms should be easily accounted for in the model, without having to re-estimate all coefficients and/or change the structure of the existing models. A potential solution could be to employ a lattice approach, for which a data assimilation step would permit accommodation of all online measurements before modeling the spatio-temporal dynamical process. Then, in order to make the general approach more generic, and potentially applicable to larger regions (potentially with various local wind climatologies), methodology adjustments should account for the fact that it may not be possible to consider a unique prevailing wind speed and direction as being representative of the weather regime over the whole area. For such more complicated cases it may be needed to switch from conditional parametric models to varying-coefficient models.

Acknowledgments

The authors would like to thank Energi E2 A/S (now part of DONG Energy) for providing the wind power data, the Danish Meteorological Institute (DMI) for the meteorological forecasts, as well as Torben S. Nielsen and the WPPT team for the wind power forecasts. This work has been partly supported by the Danish Public Service Obligation (PSO) fund, through the 'Improved Wind Power Prediction' project (under contract PSO-5766), which is also acknowledged. Acknowledgements are finally due to two reviewers whose comments allowed to enhance the paper.

References

1. Giebel G, Kariniotakis G, Brownsword R. The state-of-the-art in short-term prediction of wind power - A literature overview. EU project Anemos, Deliverable Report D1.1, 2003.
2. Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitona E. A review on the young history of wind power short-term prediction. *Renewable & Sustainable Energy Reviews* 2008; **12**:1725-1744.
3. Siebert N. Development of methods for regional wind power forecasting. Ph.D. Dissertation, Ecole des Mines de Paris, Paris, France, 2008.
4. Larson K, Westrick K. Short-term wind forecasting using off-site observations, *Wind energy* 2006; **9**:55–62.
5. Damousis GI, Alexiadis MC, Theocharis JB, Dokopoulos PS. A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. *IEEE Transactions on Energy Conversion* 2004; **19**:352–36 .
6. Gneiting T, Larson K, Westrick K, Genton, MG, Aldrich E. Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space-time method. *Journal of the American Statistical Association* 2006; **101**: 968–979.
7. Hering A, Genton MG. Powering-up with space-time wind forecasting. *Journal of the American Statistical Association*, to appear.
8. Miranda M, Dunn, RW. Spatially correlated wind speed modeling for generation adequacy studies in the UK. in *Proceedings of the 2007 IEEE Power Engineering Society General Meeting*, Tampa, Florida, USA, 2007.
9. Focken U, Lange M, Monnich M, Waldl H-P, Beyer H-G, Luig A. Short-term prediction of the aggregated power output of wind farms - A statistical analysis of the reduction of the prediction error by spatial smoothing effects. *Journal of Wind Engineering and Industrial Aerodynamics* 2002; **90**:231–246.
10. Ea Energy Analyses. 50% wind power in Denmark by 2025 - English summary. Technical Report, Ea Energy Analyses, Copenhagen, Denmark, 2007. [Available at: www.windpower.org]
11. Madsen H (ed.). Models and methods for predicting wind power. Technical Report, ELSAM/IMM, Technical University of Denmark, Informatics and Mathematical Modeling, 1996. (ISBN 87-87090-29-5)
12. Nielsen TS, Madsen H, Nielsen HAa. Prediction of wind power using time-varying coefficient functions. Proc. IFAC 2002, 15th World Congress on Automatic Control, Barcelona, Spain, 2002.
13. Nielsen HAa, Nielsen TS, Joensen AK, Madsen H, Holst J. Tracking time-varying-coefficient functions. *International Journal of Adaptive Control and Signal Processing* 2000; **14**:813–828.
14. Cutler N, Kay M, Jacka K, Nielsen TS. Detecting, categorizing and forecasting large ramps in wind farm power output using meteorological observations and WPPT. *Wind Energy* 2007; **10**:453–470.
15. Madsen H, Pinson P, Nielsen TS, Nielsen HAa, Kariniotakis G. Standardizing the performance evaluation of short-term wind power prediction models. *Wind Engineering* 2005; **29**:475–489.
16. Kotwa E, Vlasova J. Spatio-temporal modeling of short-term wind power prediction errors. M.Sc. Thesis Dissertation, Lund University, Lund, Sweden, 2007.
17. Chatfield C. *The Analysis of Time Series: An Introduction, 6th edition*, Chapman & Hall/CRC: New York, 2003.
18. Madsen H. *Time Series Analysis*, Chapman & Hall/CRC: London, 2007.
19. Nielsen HAa, Madsen H. A generalization of some classical time series tools. *Computational Statistics and Data Analysis* 2001; **37**:13–31.
20. Madsen H, Holst J. *Modeling Nonlinear and Nonstationary Time Series*. Lecture Notes, Technical University of Denmark, Dpt. of Informatics and Mathematical Modeling, Kgs. Lyngby, Denmark, 2000.
21. Tong H. *Threshold Models in Non-linear Time Series Analysis*. Lecture Notes in Statistics (no. 21), Springer Verlag: New York, 1983.
22. Tong H. *Nonlinear Time Series: A Dynamical System Approach*. Oxford University Press: Oxford, UK, 1990.
23. Pinson P, Christensen LE, Madsen H, Sørensen PE, Donovan MH, Jensen LE. Regime-switching modeling of the fluctuations of offshore wind generation, *Journal of Wind Engineering and Industrial Aerodynamics* 2008; **96**: 2327–2347.
24. Chambers JM, Hastie TJ. *Statistical models in S*. Wadsworth&Brooks/Cole: Pacific Grove, California, 1992.
25. Cleveland WS, Devlin SJ. Locally weighted regression: An approach to regression analysis by local fitting, *Journal of American Statistical Association* 1988; **83**:596–610.
26. Nielsen HA. An S-PLUS/ R library for locally weighted fitting of linear models. Technical Report, Technical University of Denmark, Informatics and Mathematical Modeling, IMM report 1997-22, Kgs. Lyngby, Denmark, 1997.
27. Pierce DA. R^2 measures for time-series. *Journal of the American Statistical Association* 1979; **74**: 901–909.
28. Lange M. On the uncertainty of wind power predictions - Analysis of the forecast accuracy and statistical distribution of errors. *Journal of Solar Energy Engineering* 2005; **127**: 177–184.
29. Pinson P. Estimation of the uncertainty in wind power forecasting. Ph.D. Dissertation, Ecole des Mines de Paris, Paris, France, 2006.
30. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. Chapman & Hall/CRC: New York, 1993.