SafeWind


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

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

Grant Agreement N°: 213740

Deliverable Dp-4.1

"Spatio-temporal modelling and forecasting of wind power prediction errors"

DOCUMENT TYPE	Deliverable
DOCUMENT NAME:	swind.deliverable_Dp-4.1.pdf
VERSION:	V2.0(*)
DATE:	2011.11.01
CLASSIFICATION:	R0: General public
STATUS:	Approved

Abstract: This report concentrates on the possibility of modelling the spatio-temporal propagation of wind power forecast errors with the final objectives of improving short-term prediction of power, in relation with the Task 4.1 of the EU project SafeWind. It concentrates on 2 different approaches, either based on the conditional parametric modelling of errors for power generation at the aggregated level (e.g. for control zones), or based on the detailed modelling of the forecast errors for all locations at which wind power generation is recorded. The methodological developments result in forecast correction, communication of new probabilistic forecasts, and further understanding of the properties of the spatio-temporal propagation of forecast errors.

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:	R. Girard (Armines), S.C. Thomsen (DTU), O. Corradi (DTU), J. Tastu (DTU), H. Madsen (DTU)			
PEER REVIEWERS:	P. McSharry (University of Oxford)			
REVIEW APPROVAL:	Approved :	X	Rejected (improve as indicated below) :	
SUGGESTED IMPROVEMENTS:	*For a long list of remarks make reference to another document*			
APPROVER:	L. von Bremen (ForWind)			

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 comments			**R3**	Restricted to WP members + PL				
S4	Under preparation			**R4**	Restricted to Task members +WPL+PL				

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

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

Contents

Publication acknowledgements **3**

1 Introduction **4**

2 Data for an analysis of aggregated power generation **5**

3 Characterization of the spatio-temporal dynamics of aggregated of forecast errors **5**

3.1 Cross-correlation between groups 6

3.2 Dependency on metereological conditions 6

4 Spatio-temporal correction of forecasts for aggregated wind power generation **7**

4.1 The CP-VAR model 8

4.2 Assessment of the point forecasts 9

4.3 Global averaged wind versus local averaged wind 9

5 Probabilistic Forecasts **11**

5.1 Truncated multivarite Normal distributions 11

5.2 Assessment of the probabilistic forecasts 11

6 Data, objectives and foreseen applications for the detailed spatio-temporal models **13**

7 Characterization of the dependence structure at two distant wind farms **16**

7.1 Joint distribution 16

7.2 Conditional expectation 17

7.3 Variance decomposition 19

8 Characterization of spatio-temporal correlation of prediction errors accross Denmark **20**

8.1 Propagation time 20

8.2 Propagation velocity and direction 21

8.3 Main velocity and direction 22

9 An alternative definition of propagation **23**

9.1 Propagating pattern 24

9.2 Estimation of speed 24

9.3 Results 25

10 Spatio-temporal propagation of errors and wind **25**

11 Conclusions **28**

Publication acknowledgements

Part of the work described in this report was published in the form of journal articles. Acknowledgements for the reproduction of this material are in order. The publications concerned are:

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

Girard R, Allard D. Spatio-temporal propagation of wind power prediction errors.*Wind Energy* 2011, submitted.

1 Introduction

Several European countries witness substantially increasing wind power integration in their energy mix. For some of the power systems, such as that of Denmark, wind power generation was seen to reach more than 100% of total power consumption for some hours over the year. Some other systems may encounter this type of situation in the coming years since wind energy is the fastest growing renewable energy source in the world: the total installed capacity has increased from 10 GW in 1998 to 158 GW in 2009[1].

The variable and hardly-predictable nature of the wind resource causes several difficulties in the operation and management of the power systems. Short-term forecasts of wind generation for the next minutes and up to 2 or 3 days ahead can facilitate the management of power systems by operators. Wind power forecasts are useful as input to various management tasks, like the dynamic quantification of reserves [1] or the optimization of combined wind-hydro power plant scheduling [2] for instance. They also turn out to be valuable when incorporated in bidding strategies for participating at electricity markets, see e.g. [3, 4]. Representative extensive reviews of the state of the art in wind power forecasting are given in [5–7].

The increasingly large number of wind farms in a country such as Denmark is challenging for the network management. However, it also yields a dense network of real time observations of the state of the atmosphere at low altitude. A single wind farm power measurement is far from being as relevant as a good calibrated meteorological station, but the 15 minutes resolution of the SCADA system and the number of wind power measurement in a country such as Denmark, contains substantially more information than the few meteorological stations that exist with 3 hourly data. One should recognize that the absence of meteorological information at this temporal and spatial scale is also related to the complexity that would arise from trying to resolve the whole physical process at this temporal scale. Also, any treatment of this large amount of data for explanatory or predictive purpose can only be a simplified approach of statistical nature.

In operational conditions, state-of-the-art forecasting methods of wind power generation are commonly optimized with 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. With 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, analysis and understanding of the spatio-temporal characteristics of wind power forecast errors are of major importance. Indeed, errors in meteorological forecasts might translate to fronts of imbalances, 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 (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 modelling of forecast errors include global corrections of wind power forecasts, associated increased knowledge of the interdependence structure of forecast uncertainty, and correspondingly improved decision-making from the forecasts available. 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 regional level.

This report can be seen as inlcuding two major parts, the first ones placing emphasis on the modeling and forecasting of spatio-temporal effects for wind power generation aggregated in various connected zones (Sections 2-5),

[1] http://www.ewea.org

while the second concentrates on the the case of considering all detailed locations at which wind power generation is observed (Sections 6-10). In both cases, the importance of meteorological conditions in the displacement of forecast errors is underlined, even though the modeling approaches and underlying analyses are fairly different. The report ends with conclusions in Section 11, as well as perspectives for future work.

2 Data for an analysis of aggregated power generation

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. Denmark has set the goal to meet 50% of electricity demand with wind energy in 2025 [10], and this will clearly result in challenges related to the management of the grid. More precisely, the case study for the first part of this report relates to western Denmark, including the Jutland peninsula and the island of Funen, which is connected to the UCTE (Union for the Co-ordination of Transmission of Electricity) system and stands for 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 here in around 1994 [11], and it is common practice today to have forecasts of wind power production at different spatial resolutions and at a state-of-the-art level of accuracy.

The work covered in the following few sections is based on power forecasts for groups of wind turbines spread throughout the western part of Denmark. Two data sets have been considered:

- 5 group dataset. First seven month of 2004. Used for the first study documented in [12] and described in Section 3. These groups are for limited numbers of turbines over Western Denmark;
- 15 group dataset. 1. January, 2006 to the 24. October, 2007. Used for the Safewind study documented in [13, 14] and described in Section 4. These groups are for all turbines in Western Denmark, except for offshore wind capacities. This is illustrated in Figure 1.

The power forecasts were generated based on a model similar to the Wind Power Prediction Tool (WPPT) [15, 16]. The forecasts have a temporal resolution of one hour and up to a 48-hour lead time.

The forecasts are based on power measurements at the wind farms, as well as meteorological forecasts. These are provided by the Danish Meteorological Institute and issued every six hours. They have an hourly temporal resolution up to 48 hours ahead. Only wind forecasts at 10 m above ground level were considered here since they are the input used by WPPT. The meteorological forecast are also considered as input to some of the spatio-temporal prediction models.

It is not the power forecasts themselve that are studied but their corresponding errors. Forecast errors are commonly defined as the difference between power measurements and forecasts, subsequently normalized by the installed wind power, following the framework described in [17]. Only one-hour-ahead forecast errors are considered in the study which are deemed suitable for the scale of Denmark.

3 Characterization of the spatio-temporal dynamics of aggregated of forecast errors

This section reviews the findings of Tastu *et al.* [12]. The work were not directly performed in the frame of Safewind, though they serve as the basis for the spatio-temporal prediction model described in the next sections. The analysis in [12] clearly ndicates that there is something to be gained utilizing the spatio-temporal structure and metereological dependencies in the forecast error. In [12] Western Denmark was grouped into five groups for simplicity.

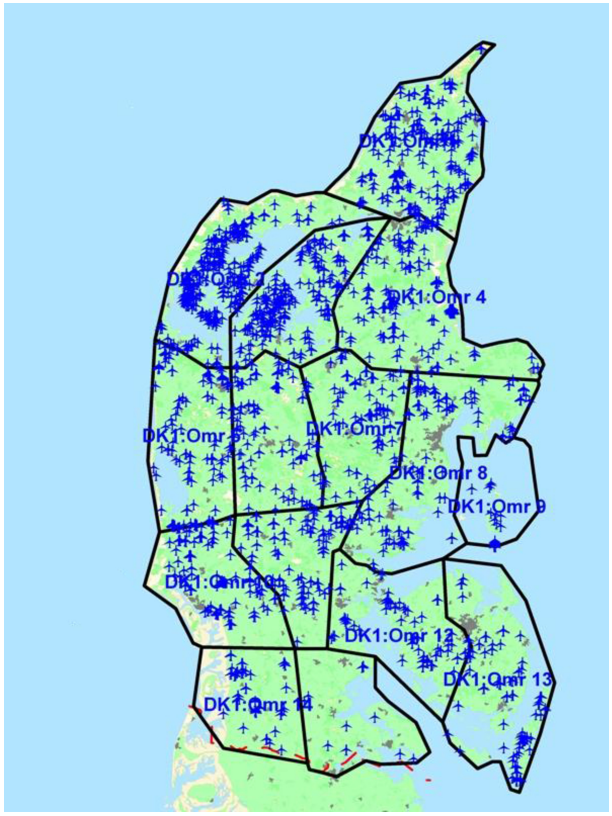

FIGURE 1: *Map of Western Denmark, with the division of all wind power capacities into 15 groups.*

3.1 Cross-correlation between groups

Figure 2 shows the cross-correlation function (CCF) for the groups denoted 5 and 1 in the study. Group 5 corresponds to Funen (center of Denmark) and group 1 south western Jutland (Western Denmark). The CCF shows that the forecast errors observed for group 5 are correlated with past forecast errors witnessed for group 1.

This is the general picture seen for various combination of groups. This is a clear indication that the forecasts can be improved by taking advantage of the cross-correlation structure.

3.2 Dependency on metereological conditions

It is natural to assume that some metereological conditions give rise to higher cross-correlation between some groups than for some others. In [12] the dependency on wind speed and wind direction was analyzed. Table 1 shows how the correlation between group 5 and 1 depends on the wind direction. The wind direction is grouped into 4 intervals. It is seen that the highest correlation is experienced with westerly winds which is the most dominating wind condition in Denmark.

Table 3 shows how the CCF of group 5 and 1 depends on the wind speed. The wind speed is divided into 5 intervals. Although a clear dependency is seen in the plot is should be noted that some of the intervals contain rather limited data points. The plot indicates that the cross-correlation between the groups increases with the wind speed.

In the work of [12] a first step study on various prediction models were examined. This comprises linear models, regime-switching and conditional parametric models. In general it was seen that performance of the forecasts were

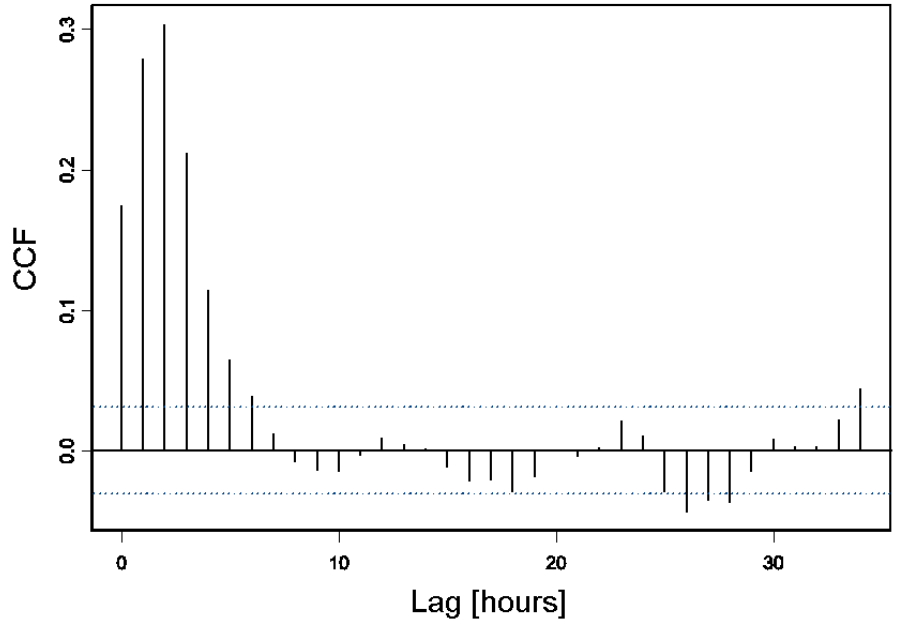

FIGURE 2: *CCF for the Groups 5 and 1. Dotted lines show 95% confidence intervals under the assumption of independence. Values outside of such intervals can be considered as significant correlation. .*

TABLE 1: *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

improved (measured in RMSE) by accounting for the meteorological conditions such as direction and wind speed.

4 Spatio-temporal correction of forecasts for aggregated wind power generation

In the scope of the Safewind project it has been the aim to derive a prediction model that could be made operational. More specifically it should capture the dominant spatio-temporal characteristics as seen in the analysis and deal with an arbitrary grouping of Denmark in a multi-dimensional approach. The CP-VAR model fulfills these criteria with a level of complexity making it suitable for implementation.

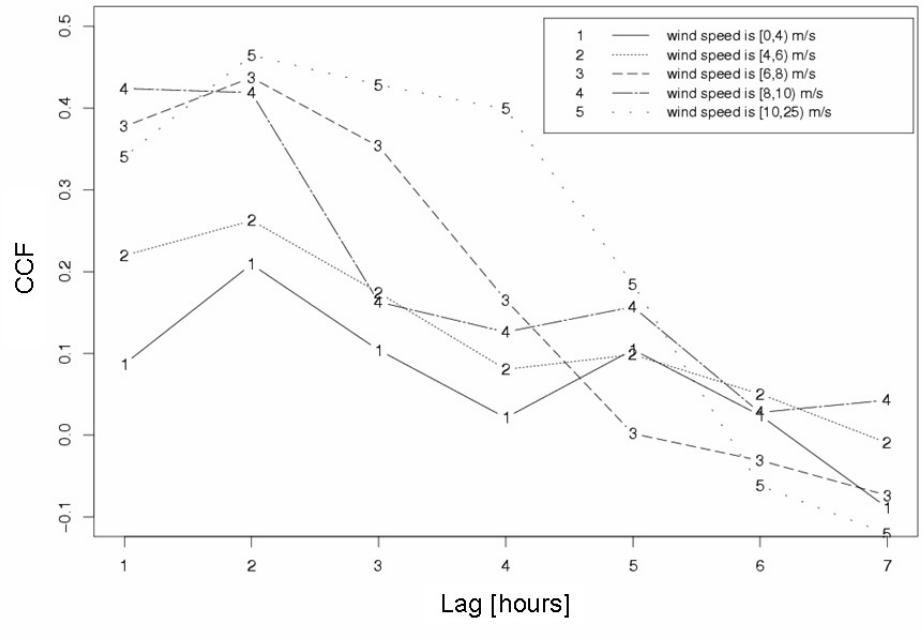

FIGURE 3: *Cross-correlation between forecast errors for Groups 5 and 1 and for the wind sector (180,270]. Cross correlation is given for different wind speed levels, and as a function of the lag.*

4.1 The CP-VAR model

Conditional parametric autoregressive models comprise a class of models with a linear structure (like a Vector AutoRegressive model (VAR), for instance), but for which the coefficients are replaced by smooth functions of other variables, i.e.

$$\mathbf{w}_t = \sum_{i=1}^{p} \mathbf{A}_i(\mathbf{z}_t)\mathbf{w}_{t-i} + \boldsymbol{\epsilon}_t. \tag{1}$$

If $\mathbf{A}$ does not depend on any parameters the model simplifies to a standard VAR model. It can be fitted to data in a straightforward manner, with $\mathbf{w}_t$ being a vector of dimension $[m \times 1]$ showing wind power forecast errors at m groups of wind farms ($m = 15$ when considering the laregr dataset) obtained for time t, $t = 1, \ldots, N$, $\mathbf{z}_t$ is a vector of dimension $[1 \times l]$ representing a signal obtained at time t which conditions model coefficients $\mathbf{A}$. If assumed that $\mathbf{A}_i$ depends on average forecasted wind direction for time t (wd_t), then $l = 1$ and $\mathbf{z}_t = wd_t$. If $\mathbf{A}_i$ is conditioned on both average wind speed and direction, then $l = 2$ and $\mathbf{z}_t = [wd_t, ws_t]$, where ws_t denotes an average forecasted wind speed for m groups of wind farms.

The results from CP-VAR models with respect to both wind speed and direction forecasts will not be presented and discussed in this work, as they did not show any improvement in terms of forecast RMSE compared to a model conditioned on the forecasted wind direction only. For similar reasons it was decided to present only the results for the CP-VAR model with $p = 1$. This model will be referred to as CP-VAR in the following. If using model (1) for a forecasting application, the one-step-ahead forecast at time t (denoted as $\hat{\mathbf{w}}_{t+1|t}$) will be given by

$$\hat{\mathbf{w}}_{t+1|t} = \sum_{i=0}^{p-1} \hat{\mathbf{A}}_{i,t}(\mathbf{z}_t)\mathbf{w}_{t-i} \tag{2}$$

where $\hat{\mathbf{A}}_{i,t}$ is the estimate of $\mathbf{A}_i$ evaluated at time t. Details of estimation approaches for these models are described in [13] and references therein.

As mentioned above a spatially averaged wind is considered for describing the metereological conditions. By spatial average we understand a global average of the wind field over Denmark. Using the global average simplifies the problem and has shown to be a robust approach. This is discussed further in Section 4.3.

4.2 Assessment of the point forecasts

Since after installation of the models some time is needed for parameter values to settle, in this work it was decided to disregard the first 5000 data points (approximately 30% of the data) in the evaluation step for both point forecast and probabilistic forecast assessments. The objective of this section is to illustrate and analyze the ability of the presented models to capture the spatio-temporal characteristics of wind power forecast errors, as well as the effects of forecasted wind direction on those characteristics. The forecasted errors are added to the original WPPT predictions to get the adjusted forecast based on spatio-temporal dependencies. The accuracy of such corrected forecasts is compared to that of original WPPT forecasts based on the RMSE criterion.

Comparison is made between:

1. The RMSE of the original WPPT forecast, i.e. estimate of the errors from the state-of-the-art wind power forecasting system, without accounting for the spatio-temporal characteristics.
2. The RMSE of VAR model based predictions, i.e. estimate of the errors resulting from the corrected forecast based on linear spatio-temporal patterns without considering the effects of meteorological conditions.
3. The RMSE of the CP-VAR model (1) based predictions, i.e. estimate of the errors resulting from the corrected forecast based on spatio-temporal patterns with consideration of the effects of the average forecasted wind direction.

Results are presented in Table 2. For each of the models, RMSE estimates are given for all the 15 groups together with the estimates of the error reduction in terms of the RMSE. The reduction in RMSE (denoted as ΔRMSE) is given as a percentage decrease in RMSE in comparison to the RMSE of the WPPT forecast for each group. It is seen that accounting for the spatio-temporal characteristics (VAR model) results in a reduction in RMSE for all the groups. The CP-VAR model outperforms VAR and this proves that the forecasted wind direction influences the spatio-temporal patterns and taking it into consideration in the model permits more accurate predictions.

Figure 4 shows the distribution of ΔRMSE resulting from the CP-VAR model through the considered geographical area. One can note that the larger improvements correspond to the eastern part of the region. This is in line with the fact that in Denmark the prevailing wind direction is westerly, so the easterly located groups are usually situated "down-wind" from the rest of the region. Therefore the spatio-temporal models show better predictive performance on the eastern part of the region as the information propagates following the wind direction. An interesting point to mention is that for Group 9 the observed improvement in the RMSE (4.08%) is not as large as for the other surrounding groups. This could be explained by the fact that Group 9, in contrast to the rest of the groups, is situated off the mainland. Therefore it is very likely that the dynamics of Group 9 are different from the rest of the groups.

4.3 Global averaged wind versus local averaged wind

In the CP-VAR model (equation 1) the global averaged wind direction (wind direction averaged over Denmark) was used as the conditional parameter. This was done under the hypothesis that the wind field over Denmark is

Group	WPPT	VAR		CP-VAR	
	RMSE [%]	RMSE [%]	Δ RMSE [%]	RMSE[%]	Δ RMSE [%]
1	3.32	3.11	6.40	3.08	7.32
2	2.98	2.88	3.13	2.81	5.62
3	3.39	2.99	11.77	2.87	15.22
4	3.29	2.83	13.98	2.76	16.13
5	3.15	3.06	2.65	3.01	4.39
6	3.27	2.92	10.82	2.83	13.31
7	3.53	3.01	14.69	2.92	17.09
8	2.93	2.47	15.45	2.39	18.46
9	3.34	3.22	3.46	3.20	4.08
10	3.58	3.45	3.60	3.39	5.31
11	3.29	2.83	14.18	2.72	17.26
12	3.21	2.75	14.38	2.66	17.22
13	2.97	2.63	11.63	2.57	13.41
14	3.77	3.64	3.50	3.58	4.94
15	3.49	3.11	10.76	3.01	13.67

TABLE 2: *Evaluation of the forecast performance of the various models in terms of the RMSE.*

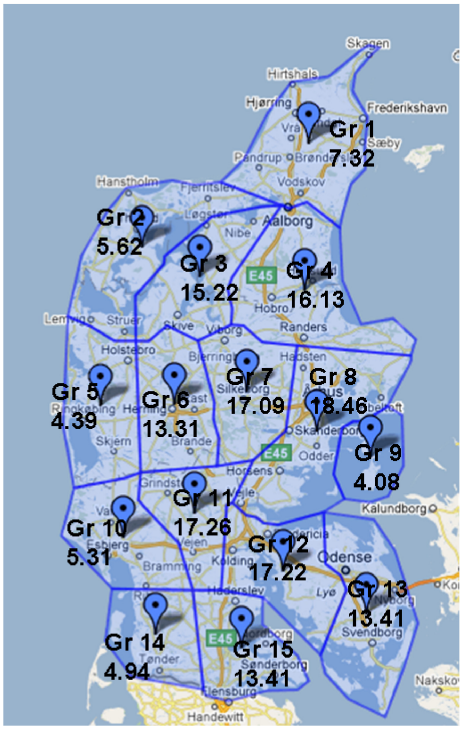

FIGURE 4: *Predictive performance of the CP-VAR model in terms of a percentage reduction in the RMSE (ΔRMSE) of the forecast errors. (Produced using* `http://maps.google.dk/`*)*

sufficiently homogenious with good results. In the work described in [14] it was examined if the prediction error could be reduced even more by considering locally averaged wind directions. By locally averaged wind direction we understand the average wind direction between pairs of zones. The resulting model has the structure:

$$\mathbf{w}_t = \Sigma_{i=1}^{p} \begin{pmatrix} a_{i,11}(\mathbf{z}_{11}) & \dots & a_{i,1m}(\mathbf{z}_{mm}) \\ \vdots & \ddots & \vdots \\ a_{i,m1}(\mathbf{z}_{n1}) & \dots & a_{i,mm}(\mathbf{z}_{mm}) \end{pmatrix} \mathbf{w}_{t-1} + \epsilon_t \tag{3}$$

As illustrated every coefficient of the prediction model is condition to a "local" parameter. In this work $\mathbf{z}_{ij} = wd_{ij}$ are the locally averaged wind directions. For simplicity the order of the model was set to $p = 1$.

In [14] is was seen that no improvement measured in Root Mean Squared Error (RMSE) was gained using the local averaged wind. In fact the CP-VAR model with global averaged wind direction seems more robust resulting in a slightly better prediction of the forecast errors than. This indicates that the globally averaged wind is sufficient for the scale of Denmark and there is no need to introduce the additional complexity of locally averaged wind. This could be explained by the fact that the wind field is very homogeneous in Denmark (80% of the wind dataset is located under a variance of 0.2.

5 Probabilistic Forecasts

Besides providing improved point forecasts of the groups focus has also been on providing probabilistic knowledge of the improved point forecasts. In order to build such probabilistic forecasts for the prediction errors, a parametric approach employing a truncated MultiVariate Normal distribution (MVN) is used [18].

5.1 Truncated multivarite Normal distributions

Briefly, the assumption is that:

$$\mathbf{w}_t \sim N_{\mathbf{a_t}}^{\mathbf{b_t}} \left(\hat{\mathbf{w}}_{t|t-1}, \mathbf{\Sigma}(\mathbf{z}_t) \right) \tag{4}$$

where $\hat{\mathbf{w}}_{t|t-1}$ denotes the mean of the distribution, which is assumed to be equal to the point forecast of the wind power forecast errors obtained from (2). The parameters $\mathbf{b}_t$ and $\mathbf{a}_t$ are vectors of the dimension $[m \times 1]$ denoting upper and lower truncation limits of the distribution. The need to truncate the distribution arises from the fact that standardized wind power predictions issued by WPPT for time t (denoted by $\mathbf{pp_t}$) lie between 0 and 1, since one cannot obtain negative power as well as a quantity larger than a nominal capacity. Therefore the errors of the power predictions are also bounded:

$$\mathbf{b_t} = 1 - \mathbf{pp}_t \quad \mathbf{a_t} = -\mathbf{pp}_t \tag{5}$$

$\mathbf{\Sigma}(\mathbf{z}_t)$ is a covariance matrix of the distribution which is conditional on the external signal $\mathbf{z}_t$ which in this case equals an average forecasted wind direction at time t. Estimation is performed in a recursive adaptive way, similar to the framework of estimation in CP-VAR models. Further details can be found in [13].

5.2 Assessment of the probabilistic forecasts

A primary requirement for probabilistic forecasts relates to their calibration, which corresponds to the statistical consistency between the probabilistic forecasts and the observations [19]. In the univariate case, calibration can be verified using the Probability Integral Transform (PIT). In the ideal situation, i.e. if the observations were drawn from the predictive distribution, the PIT would have a uniform distribution on the unit interval $[0, 1]$ [19]. Therefore, in order to assess calibration for a univariate case, one can plot a PIT histogram and check for its uniformity. Assessing probabilistic forecasts of multivariate quantities is more complex. Some of the tools are presented in [20]. Since in this work 15-dimensional data having a truncated distribution with varying parameters is analyzed, the implementation and evaluation of genuinely multivariate approaches presented in [20] become troublesome. Instead, as the first step, the adequacy was checked on the individual marginals of the estimated multivariate density. After checking PIT histograms for each of the 15 groups, it was observed that the results for

all groups look very similar. Figure 5 shows as an example the PIT for Groups 8. The histogram is relatively close to uniform which is the general picture for the various groups.

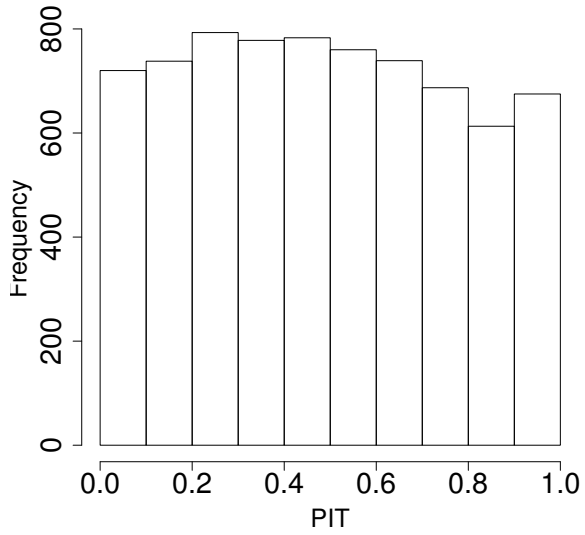

FIGURE 5: *PIT histogram for the reliability assessment of the univariate probabilistic forecasts for Group 8*

For illustration purposes, a probabilistic forecast for Group 8 is depicted in Figure 6 for a one-week period beginning on the 3rd of August 2007.

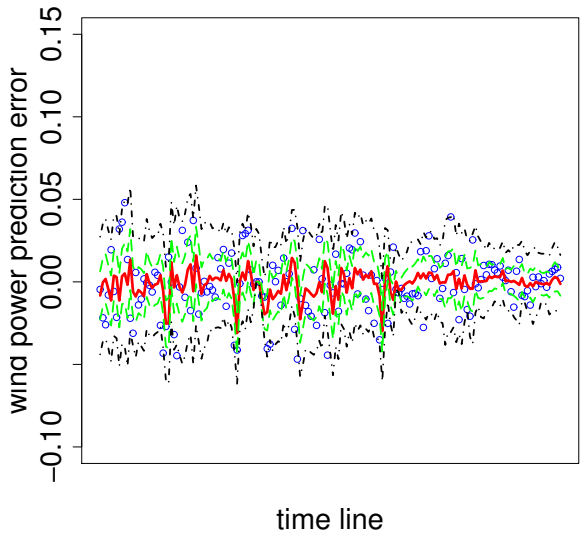

FIGURE 6: *Probabilistic forecast for Group 8. Blue circles denote the observed WPPT errors, red solid lines show the corresponding point forecasts produced by the CP-VAR model. Green and black broken lines correspond to upper and lower limits of the 50% and 90% forecast intervals (based on the quantiles of marginal MVN distributions).*

6 Data, objectives and foreseen applications for the detailed spatio-temporal models

Usual short term wind power prediction model are statistical models that estimate the link between the forecast or analysed meteorological situation, the observed wind power production, and the future wind power production (see for example [6]). When forecasting wind power at look ahead times smaller than 4 hours it is known that the forecast meteorological situation that is provided by usual numerical weather prediction is not as relevant as the in situ wind power measurement. At this temporal scale, the unsused information that is contained in the spatio-temporal field of wind power measurement at the scale of a country might be more relevant.

The main purpose of this work is to analyse the spatio-temporal stochastic structure of wind power prediction errors for look ahead times smaller than 4 hours and with a 15 minutes temporal resolution. At this resolution, we use a simple $AR(4)$ model to forecast wind power and generate prediction errors. One AR model is fitted per wind farm. This can be considered as a convenient pre-whitening transformation that permits the spatio-temporal analysis of wind power production. We have performed the same analysis with persistence forecast and with AR model of higher order and the results remain roughly the same. The corresponding transformation is also convenient for understanding the potential of spatio-temporal information as a complement to in situ measurement in wind power forecast. We concentrate on a test case in Denmark (domain $DK1$) which contains 212 transformation stations whose location can be seen in Figure 7. For each of these spatial locations, a 2 year power production is observed at a 15 minute resolution. The first year of data (2007) is used to estimate the parameters of the AR models and the second year is used for the analysis performed in this paper. In the rest of the report, these transformation stations are assimilated as wind Farms.

FIGURE 7: *Transformation stations to be used in the study (All points). Red points are associated to number that are reused in the report.*

The quantity of interest will now be the prediction errors themselves. For a wind farm at position $p = (x, y)$ we will denote by $\epsilon^p_{t+h|t}$ the standarized prediction error (forecast value minus observed value normalized by nominal capacity) of the forecast obtained at time t for wind farm p and look ahead time h. The horizons are given in minutes and $t + 15$ refers to the time 15 minutes after time t which is also the first horizon after time t. The density of prediction errors for one representative wind farm and different look ahead times are shown Figure 8. These

errors are unbiased, not gaussian, have negative skewness that shift the mode of the distribution toward negatives values. This shift increases continuously with the horizon.

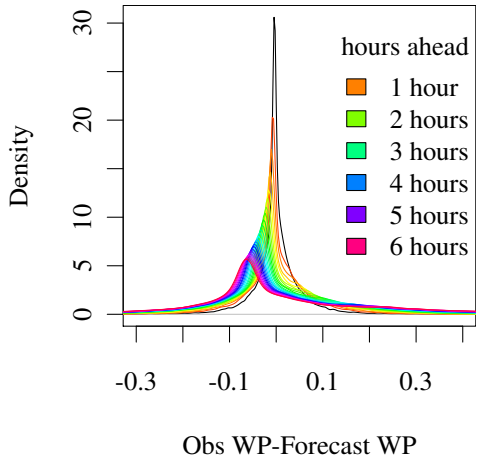

FIGURE 8: *Density of prediction errors for wind farm* 7.

We will denote by H the largest possible look ahead time and use the notations $\mathcal{E} = (\epsilon^p_{t+h|t})_{t=1,\dots,T,h=1\dots,H,p=p_1,\dots,p_n}$ for the whole process. For a given run time $t \in \{1,\dots,T\}$ (i.e. a time for which a forecast trajectory is generated) we will denote by $\mathcal{E}_t = (\epsilon^p_{t+h|t})_{h=1,\dots,p=p_1,\dots,p_n}$ the corresponding spatio-temporal field of wind power prediction errors. In this report we consider that $\mathcal{E}_1,\dots,\mathcal{E}_T$ are (possibly dependent) relalizations of random variables identically distributed with the same distribution as a random vector denoted $\mathcal{E}^+$ (which can be considered as a spatio-temporal process). This said, the objective of this report can be rephrased into the analysis of the stochastic structure of $\mathcal{E}^+$ in a spatio-temporal perspective and conditionnally to meteorological variables.

This work may be further motivated by other possible applications such as the followings.

Correction is the task of predicting at time t for wind farm p_0 the vector of errors

$$(\epsilon^{p_0}_{t+15|t},\dots,\epsilon^{p_0}_{t+H|t})$$

using as explanatory variable all the errors measured until time t among all wind farms:

$$\begin{pmatrix} \epsilon^p_{t|t-15} & \epsilon^p_{t|t-30} & \cdots & \cdots \\ & \epsilon^p_{t-15|t-30} & \cdots & \cdots \\ & & \epsilon^p_{t-30|t-45} & \cdots \\ & & & \cdots \end{pmatrix}_{p\in\{p_1,\dots,p_n\}}$$

Let us denote by Φ_t the vector containing this whole set of information. If one can specify the distribution of $\mathcal{E}$ and if we observe $\Phi_t = x$ the correction task is done by evaluating

$$\mathbb{E}\left[\epsilon^p_{t+15|t},\dots,\epsilon^p_{t+H|t}|\Phi_t = x\right]$$

If $\mathcal{E}$ is a Gaussian process, this conditional expectation is fully specified by the knowledge of the covariance matrix of $\mathcal{E}$ and the application of the correlation Theorem.

Literature on spatiotemporal correction includes [21, 22] [12] [23] (spatial ARMAX) Kalman Filter [24].

Scenario generation. Probabilistic predictions for several successive lead times are most often generated and presented as marginal distributions for each lead time, or as some of their characteristics e.g. quantiles or central prediction intervals. The most prominent example certainly consists of the Bank of England's fan charts discussed and analyzed by [25]. They can be sufficient if the decisions to be made for a given lead time are independent of the others, as for the wind energy trading problem of [4]. In a general manner, however, making optimal decisions may require knowledge of the interdependence in the stochastic process over various lead times — possibly also for a number of locations and many different variables. It is the case for instance with the optimal integration of renewable energies into existing power systems and electricity markets [26]. This then requires the forecasting of the characteristics of the joint distribution of the process for the set of lead times of interest. Since modeling, estimation and communication of complex multivariate densities may be intractable, this type of forecasts takes the form of simultaneous prediction intervals [27–30] or of time trajectories as for example in [31] or [32], also referred to as ensemble forecasts in the meteorological community [33].

Formally trajectory generation is the task of generating, at time t, K spatio-temporal trajectories

$$(e^{p,k}_{t+15|t}, \ldots, e^{p,k}_{t+H|t})_{p=p_1,\ldots,p_n,k=1,\ldots,K}$$

that can be assimilated as realizations of K independent random variables that have the same distribution as $(\epsilon^{p}_{t+15|t}, \ldots, \epsilon^{p}_{t+H|t})_{p=p_1,\ldots,p_n}$.

Ressource assessment from a predictability point of view. In resource assessment, the aim is to take optimal decisions on the location of new wind farms. As penetration increases predictability and forecasting tools become of paramount importance. Nowadays, with the development of very large scale wind farms (i.e. offshore) and also with the development of electricity markets, wind predictability may become an issue at the level of site selection and design of a new wind farm. There is now an interest in developing research on the new concept of "resource assessment versus predictability". The predictability of a site and especially the issue of extreme events can be considered when taking decisions for the installation of a new wind farm. This naturally involves understanding the relation (correlation or non linear relations) in time and in space between the forecast errors at a given site that already exists in a portfolio of wind farms and a given number of potential new sites were no wind farm exist and hence no measurement is made. In this case it is necessary to develop a parametric (or semi-parametric) approach modeling the link between the interdependency structure, the time and the geographical position.

Storage system: dimensioning, optimized use. Energy Storage Systems (ESS) constitutes an important family of solutions for smoothing renewable power. The combination of an ESS with a renewable generator provides a hybrid system capable of providing deterministic power of the required quality at the scheduled time. The two dual class of problems that arise from the introduction of a storage system in a set of geographically distributed renewable power plant are

- Dimensioning and positioning (with a minimal cost).
- Optimal use (for maintaining a specified imbalance or for participating into the electricity market)

One should notice that an answer to one of these problems should be associated to an answer to the other problem. In the simple case when one has a wind farm and wants to design a storage to upper bound the probability to have, for example, 15 minutes of produced power under a given threshold within a year, it is required to know the temporal correlation of the power production. Similarly, when the problem comes to dimensioning a storage system for a large number of wind farms geographically distributed, with a more complex goal related to imbalance of wind power production, a deep understanding of the spatio-temporal interdependence structure of the wind power forecast errors is required. Also, a modeling adapted to extreme events is required.

Outlier detection. Outlier detection is a classical statistic thematic. When applied to spatio-temporal field of wind power forecast errors, the basic ideas for outliers detection suggest that the spatio-temporal coherence of the underlying process should be used. Indeed, it is clear that part of a wind power forecast error, for a group of wind farms, show something abnormal when part of the group for a given period of time is in contradiction with the assumed spatio-temporal stochastic interdependence structure. Considering different time and spatial scales to detect abnormal errors is a relatively new idea that requires the knowledge of the spatio-temporal interdependence structure of wind power forecast errors.

The description of the stochastic structure of errors can take different forms depending on the nature of the stochastic relations that exist between forecast errors at different time and wind farms. In Section 7 we restrict ourselves to the study of two distant wind farms to understand the nature of the above mentioned relation. This allows us to show that wind power forecast errors at two different wind farms have maximum correlation for a given lead time. After showing suitability of correlation between errors to describe spatio-temporal propagation from one wind farm to another, the corresponding propagation is analyzed in more details in Section 8 (using the 212 transformation stations assimilated as wind farms). The propagation we describe with time averaged statistical indices (i.e. correlation) and in Section 9 we give an alternative definition of propagation based on a planar wave modeling. In these two analysis, we identify two advections characteristics that are related to displacement of meteorological system of cold/warm air advection and is also related to wind direction. In Section 10 we analyze the spatio-temporal correlation of errors conditionally on wind. Since it is not obvious whether wind at 1000 hPa is giving the best direction and speed for analyzing displacement of errors in space and in time, we also use wind at higher heights.

7 Characterization of the dependence structure at two distant wind farms

Understanding a spatio-temporal process over a country, such as the wind power production or it prediction error involves four dimensions, two spatial dimensions and two temporal dimensions: the time t and the prediction look ahead, denoted h. A first step in understanding such a complex process is to consider few spatial locations and to analyze the corresponding multidimensional temporal process. In this section we reduce the study to a subset of 2 wind farms (Farms 23, and 25) whose locations are marked in Figure 7. This permits us to describe the nature of the nonlinear relations that exist between errors at different locations and look ahead time. We wil show that this relation is nonlinear but that linear correlation is nevertheless suitable for describing and quantifying the spatio-temporal propagation of wind power prediction errors.

7.1 Joint distribution

From Figure 8 we have already seen that the prediction errors have smooth unimodal distribution. In Figure 9 we have represented the density of the joint distribution $(X, Y) = (\epsilon^{23}_{t+60|t}, \epsilon^{25}_{t+15|t})$, as estimated with the bivariate kernel density estimation implemented in the Package `KernSmooth` [34],[35] of statistical analysis software `R`. We used a Gaussian kernel and an automatic window bandwidth selection based on the univariate variance estimation. It is represented for a reduced range of values around its mode in Figure 9. Clearly, the the contours of the joint distribution are not elliptical, showing that the joint distribution $(\epsilon^{25}_{t+15|t}, \epsilon^{23}_{t+60|t})$ is not bi-Gaussian . In agreement with Figure 8 the marginal distributions of $\epsilon^{23}_{t+h|t}$ present a negative skewness: while being centered, the mode of the error vector $(\epsilon^{25}_{t+15|t}, \epsilon^{23}_{t+60|t})$ found around $(-0.005, -0.015)$, is shifted toward the negative values. A significant part of the mass is concentrated in a very small area around this mode.

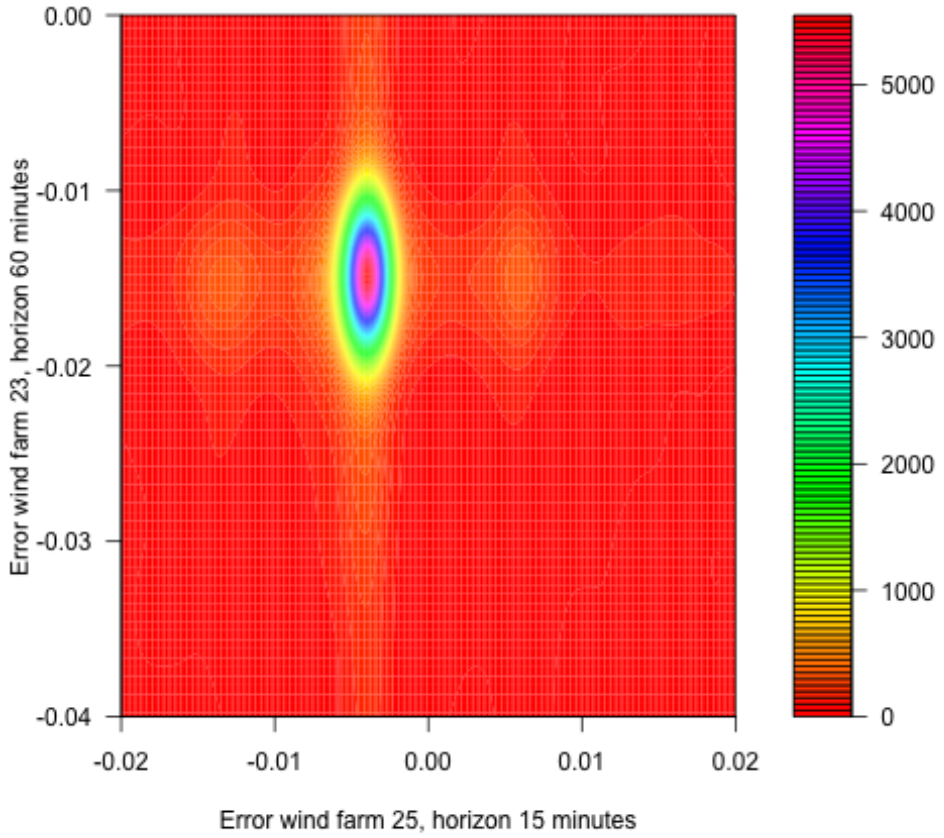

FIGURE 9: *Kernel density estimation of the joint distribution of* $(X, Y) = (\epsilon^{25}_{t+15|t}, \epsilon^{23}_{t+60|t})$.

7.2 Conditional expectation

There are many different ways of characterizing the dependence between two random variables X and Y. The simplest and most common one is clearly the linear correlation, but it only characterizes linear relationships. Its interest is therefore questionable since we have seen in the previous subsection that the relationship between the errors is non-linear.

The relationship between wind power prediction errors at different locations and for different look ahead times is further explored in Figure 10 in which is represented local polynomial estimates of the conditional expectation $y(x) = E[\epsilon^{23}_{t+h|t} | \epsilon^{25}_{t+15|t} = x]$ on a reduced set of values for x: $|x| < 0.1$. The estimation was done with the function `loess` in `R`. Outside the interval $[-0.1, 0.1]$ the number of samples is too small to give any conclusion without further modeling. These local polynomial estimations are shown with their 99% confidence intervals under the assumption that the estimation bias that results from using local polynomials is null. The conditional expectations shown in Figure 10 are non linear (except for the first values of h).

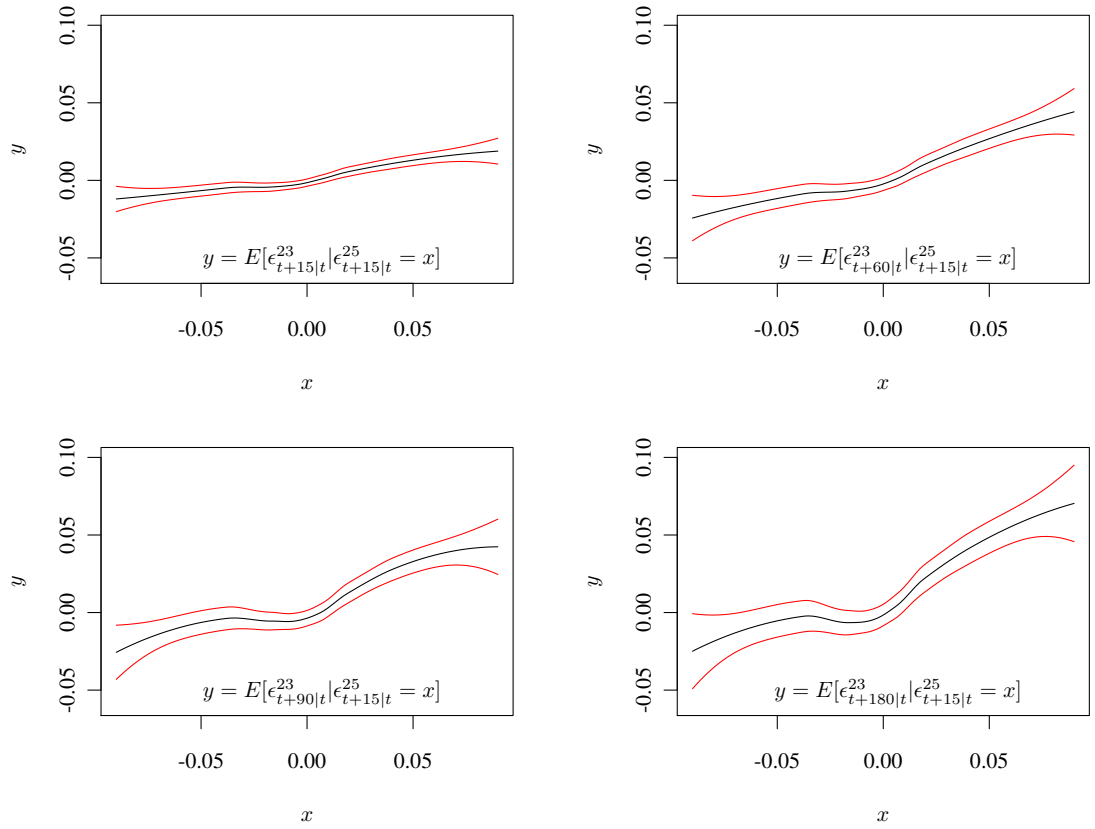

FIGURE 10: *Local polynomial fitting* $\hat{y}(x)$ *of* $y(x) = E[\epsilon^{23}_{t+h|t}|\epsilon^{25}_{t+15|t} = x]$ *(*$h \in \{15min, 60min, 90min, 180min\}$*) and confidence intervals at* 99% *(red lines) for this estimation (obtained under the assumption that* y *can be accurately described with local polynomials).*

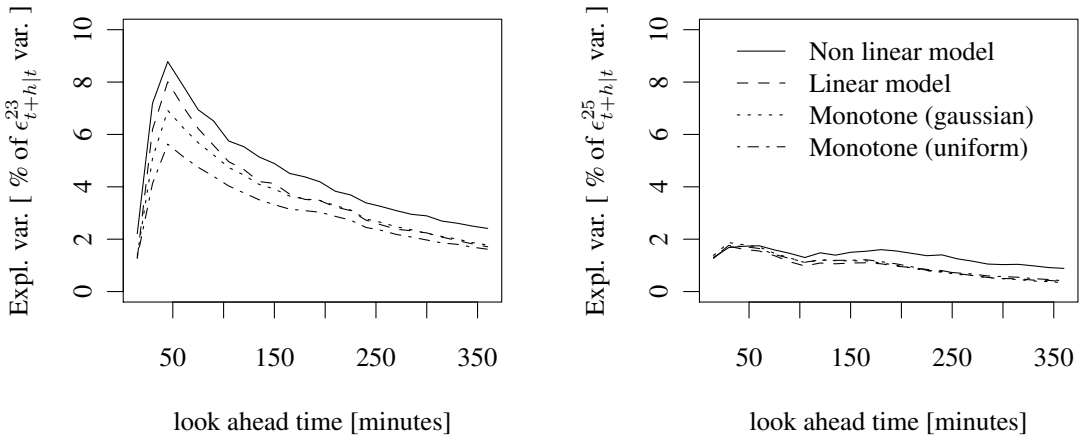

FIGURE 11: *Left hand side: interdependence between $\epsilon^{25}_{t+15|t}$ and $\epsilon^{23}_{t+h|t}$ as a function of h. Expressed as a proportion of variance of $\epsilon^{23}_{t+h|t}$ explained by $\epsilon^{25}_{t+15|t}$. Right hand side: interdependence between $\epsilon^{23}_{t+15|t}$ and $\epsilon^{25}_{t+h|t}$ as a function of h. Expressed as a proportion of variance in $\epsilon^{25}_{t+h|t}$ explained by $\epsilon^{23}_{t+15|t}$*

7.3 Variance decomposition

Measuring the strength of the non linear relationship $Y = f(X)$ can be done by decomposing the variance of Y, using that $E[Y|X]$ is an orthogonal projector:

$$Var(Y) = Var\left(E\left[Y|X\right]\right) + Var\left(Y - E\left[Y|X\right]\right). \tag{6}$$

The first term represents the variance of Y explained by X. If $E\left[Y|X = x\right]$ is linear, then the ratio $Var\left(E\left[Y|X\right]\right)/Var(Y)$ is the squared correlation between observations of X and Y. If this relation is non linear, there are different alternatives. One can operate monotonic transforms of X and Y to a predefined distribution before computing the squared correlation coefficient between the transformed variables. Obviously, this can be efficient when the relationship between Y and X is monotonic. When the predefined distribution is uniform, it leads exactly to the Spearman correlation coefficient. In the general case, one can estimate $E\left[Y|X\right]$ with a non parametric procedure and compute the associated variance directly or with Monte Carlo simulations. The corresponding fraction of variance is then called Sobol indices [36].

In Figure 11 the strength of the relationship between $Y = \epsilon^{23}_{t+h|t}$ and $X = \epsilon^{25}_{t+15|t}$ is analyzed on the left panel, while the relationship between $Y = \epsilon^{25}_{t+h|t}$ and $X = \epsilon^{23}_{t+15|t}$ is analyzed on the right panel. This is done through the computation of different correlation coefficients as a function of the look ahead time h: the squared correlation between $Y = \epsilon^{23}_{t+h|t}$ and $X = \epsilon^{25}_{t+15|t}$ (respectively $Y = \epsilon^{25}_{t+h|t}$ and $X = \epsilon^{23}_{t+15|t}$), the squared Spearman correlation, the squared correlation obtained after transforming the variables into normal random variables, and an estimation of $Var\left(E\left[Y|X\right]\right)/Var(Y)$ obtained with non-parametric local polynomials estimate of $E\left[Y|X\right]$ and Monte Carlo simulations.

It is clear from Figure 11 that for any of the correlation coefficient considered the relationship between $\epsilon^{23}_{t+h|t}$ and $\epsilon^{25}_{t+15|t}$ is stronger than the relationship between $Y = \epsilon^{25}_{t+h|t}$ and $X = \epsilon^{23}_{t+15|t}$. We can also observe that the influence of $\epsilon^{25}_{t+15|t}$ on $\epsilon^{23}_{t+h|t}$ increases with h until a maximum of influence is reached and then decreases. The amount of variance explained by the relationship depends on the correlation coefficient (the largest is for the general non-linear relationship; then it is larger for a linear then for monotonic transformations). However, whe considered as functions of the look ahead time, all these variance analysis show a maximum of influence at the

same look ahead time.

8 Characterization of spatio-temporal correlation of prediction errors across Denmark

Moving toward a spatio-temporal description of the prediction errors requires indentifying spatial and temporal scales. In this paper, we identify these scales through the analysis of the temporal and spatial lags that maximize the dependence between errors.

We now study the spatio-temporal dependence structure of prediction errors at a larger spatial scale. The problem is reduced to that of analyzing correlation between different locations and look ahead times. The use of correlation for this purpose has been justified in the preceding Section. Although linear correlation does not perfectly capture the relationship between forecast errors at two different wind farms it captures the evolution of this relation along look ahead times which is the quantity of interest in this section. It is thus a suitable tool to describe the temporal and spatial scales that are involved in the average propagation of errors.

Formally, we analyze the spatio-temporal correlation structure of $\mathcal{E}^+$ defined at the end of Section 6. As already mentioned, the empirical correlation between errors $\epsilon^{p_1}_{t+h_1|t}$ and $\epsilon^{p_2}_{t+h_2|t}$ at two different spatial positions p_1 and p_2 and for two different look ahead times h_1 and h_2, denoted $C(h_1, h_2, p_1, p_2)$ is a complex quantity depending on 2 temporal variables and 4 spatial coordinates. However, our aim is to use this correlation function to characterize the propagation observed in the preceding section and ultimately to analyze the speed and direction of the propagation of wind power forecast errors in Denmark. Built upon the last remark of the preceding section, we will first define the temporal scale associated to the propagation; then we will characterize its direction and speed. We will end this section with a the definition of simple statistic that quantifies the main propagation direction and speed.

8.1 Propagation time

The correlation between wind power forecast errors at two distant wind farms is a usually a smooth function of look ahead times h_1 and h_2. It conveys important information regarding the propagation of errors between pairs of wind farms from which we can derive propagation times and propagation speeds. Two important observations were consistently made on the data. Firstly, for a given look ahead time, one either observes that

$$\max_{h_2} C(h, h2, p1, p2) \gg \max_{h_2} C(h, h2, p2, p1)$$

or

$$\max_{h_2} C(h, h2, p1, p2) \ll \max_{h_2} C(h, h2, p2, p1).$$

Furthermore, this property does not depend on h and we can fix $h = 15$. Secondly, and this is seen on Figure 11, when one of the look ahead time is kept fixed, say $h = 15$, the function $C(h, h2, p1, p2)$ of h_2 shows a single maximum. There is thus a particular look ahead time h_2 at which the error $\epsilon^{p_2}_{t+h_2|t}$ has the highest correlation with the error $\epsilon^{p_1}_{t+1|t}$. As a consequence, between any two wind farms (p_1, p_2) there is a prefered direction of correlation. If $\max_{h_2} C(15, h2, p1, p2) > \max_{h_2} C(15, h2, p2, p1)$ we say that p_1 is upstream to p_2, which will be denoted $p_1 \prec p_2$, and vice-versa.

From these, when p_1 and p_2 are two wind farms such that p_1 is upstream (i.e. $p1 \prec p_2$) we define a propagation time $T(p_1, p_2)$ between the two wind farms:

$$T(p_1, p_2) = \arg\max_{h_2} C(15, h_2, p_1, p_2) - 15, \tag{7}$$

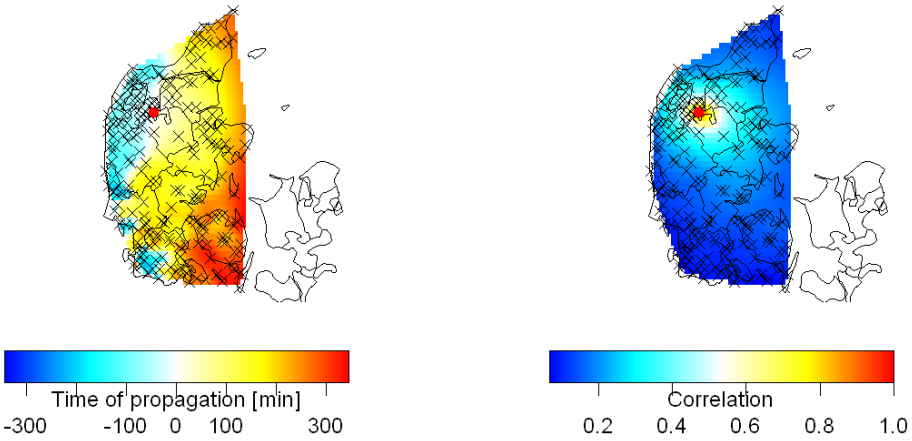

FIGURE 12: *Velocity of propagation of maximum correlation from a particular wind farm (Wind farm 59). On the left, propagation time from wind farm 59: the value of $p \to T(p_{59}, p)$ where T is defined by Equation 7, interpolated to be represented as an image. On the right, the associated value of maximum correlation is represented also using interpolation.*

and denote by $\rho_{max}(p_1, p_2)$ the associated maximum correlation:

$$\rho_{max}(p_1, p_2) = C\left(15, T(p_1, p_2) + 15, p_1, p_2\right). \tag{8}$$

The left panel of Figure 12 represents the propagation time from wind farm 59 to all other wind farms. For visual conciseness, the propagation time from wind farms located upstream is also represented with negative values. The main observation that can be made is that wind farms p located west from wind farm 59 are usually upstream wind farms (i.e. with negative values of the propagation time) while those located east from wind farm 59 are usually downstream. This shows that, at least for wind farm 59, the propagation is westerly with a slight south tendency.

The right panel of Figure 12 is a spatial interpolation of the correlation $\rho_{max}(59, p)$ when p is located downstream to the wind farm 59. For wind farms located upstream, the quantity depitced is $\rho_{max}(p, 59)$ associated to the propagation time $T(p, 59)$. Figure 13 represents the average of this correlation as a function of distance between p_1 and p_2. We can see that the correlation decreases rapidly down to below 0.2 for distances up to 50 km. For larger distances, it decreases much more slowly.

8.2 Propagation velocity and direction

Figure 13 shows how the maximum correlation $\rho_{max}(p_1, p_2)$ decreases with distance $p_2 - p_1$. Furthermore, Figure 14 clearly indicates that the east-west direction is a privileged direction: for a given distance, ρ_{mas} is maximum in this direction and the contour lines of $T(p_1, p_2)$ are almost perpendicular to this direction.

For two wind farms p_1 and p_2 such that $p_1 \prec p_2$, the knowledge of the propagation time from p_1 to p_2 permits the definition of a propagation velocity from a wind farm p_1 to a wind farm p_2

$$v(p_1, p_2) = \frac{d(p_1, p_2)}{T(p_1, p_2)} \text{ if } p_1 \prec p_2, \tag{9}$$

where $d(\cdot, \cdot)$ is the Euclidian distance. In a given direction $\theta \in (-\pi, \pi)$, we can further quantify the propagation speed by defining a directional average propagation velocity $V(\theta)$. Let us denote $\angle(p_i, p_j)$ the angle in $(-\pi, \pi)$

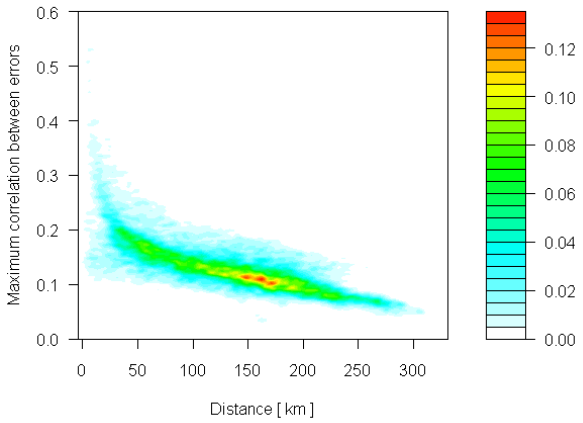

FIGURE 13: *Maximum correlation (ρ_{max} as defined by Equation 8) between wind farms as a function of distance. The color level gives the density of points.*

made by the vector $p_i p_j$ with the X axis. Then we define

$$V(\theta) = \frac{\sum_{i,j} v(p_i, p_j)\chi(p_i, p_j)}{\sum_{i,j} \chi(p_i, p_j)}, \tag{10}$$

where

$$\chi(p_i, p_j) = 1 \text{ if } p_i \prec p_j,\ \angle(p_i, p_j) \in (\theta - \delta_\theta, \theta + \delta_\theta),$$

and $\chi(p_i, p_j) = 0$ otherwise.

This directional average propagation velocity is defined for all directions $\theta \in (-\pi, \pi)$. The drawback of $V(\theta)$ for quantifying the general propagation direction is that, even in directions with only very few couples of (upstream, downstream) wind farms, this quantity will be defined and non null. To get a more realistic picture, we would need to average out propagations in a given direction θ with those in the opposite direction $\theta + \pi$. We thus also define a signed average of the velocity $V^s(\theta)$ for the directions $\theta \in (-\pi/2, \pi/2)$ by considering the weights

$$\chi(p_i, p_j) = \begin{cases} 1 & \text{if } p_i \prec p_j \text{ and} \\ & \angle(p_i, p_j) \in (\theta - \delta_\theta, \theta + \delta_\theta) \\ -1 & \text{if } p_i \prec p_j \text{ and} \\ & \angle(p_i, p_j) + \pi \in (\theta - \delta_\theta, \theta + \delta_\theta) \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Note that the condition $p_i \prec p_j$ and $\angle(p_i, p_j) + \pi \in (\theta - \delta_\theta, \theta + \delta_\theta)$ is equivalent to $p_j \prec p_i$ and $\angle(p_i, p_j) \in (\theta - \delta_\theta, \theta + \delta_\theta)$

8.3 Main velocity and direction

The two velocities $V(\theta)$ and $V^s(\theta)$ are represented in Figure 14, on which the proportion of positive weights is also represented as a function of θ. We see that for east west direction, this proportion is null while it is around 100% for West-East direction. In other words, there are no couples of wind farms aligned in the East-West direction such that the wind farm located east is upstream. In the north direction, the proportion is around 40 %: farms located North can be downstream to farms located South but the opposite is more frequent. This confirms what was observed in Figure 12: the main propagation direction is East, with a slight inclination toward South.

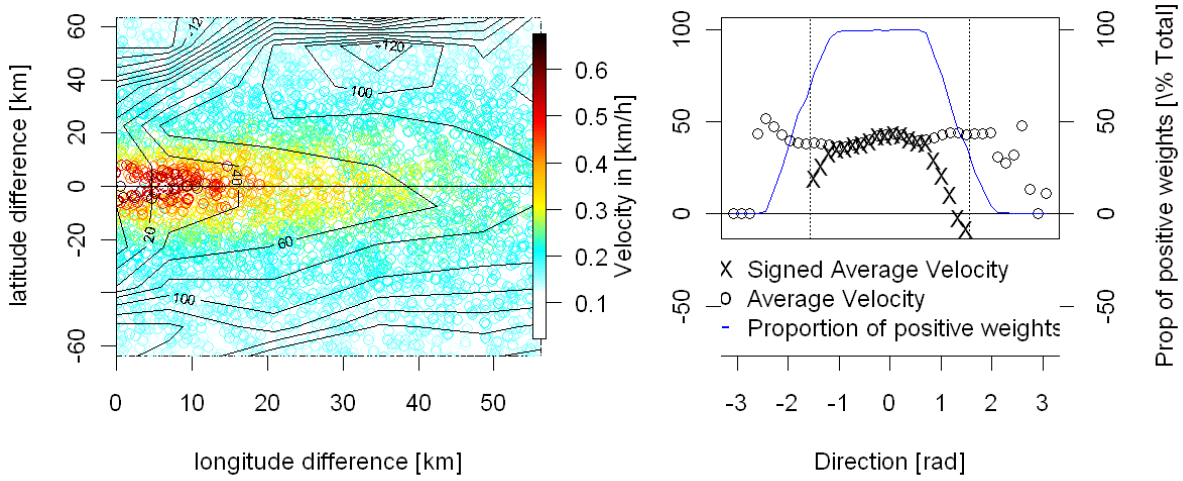

FIGURE 14: *Velocity and direction of propagation of maximum correlation, from all wind farms. On the left the maximum correlation $\rho_{max}(p_1, p_2)$ between to locations p_1 and p_2 is represented with a color scale as a function of the vector $p_1 - p_2$, the corresponding propagation time $T(p_1, p_2)$ is represented on the same graph in minutes with level set contour lines as a function of the vector $p_1 - p_2$. The Figure on the right gives the estimated directional average propagation velocity $V(\theta)$ calculated in different directions as defined by Equation 10, and the propagation velocity $V^s(\theta)$ obtained with the signed weights defined by 11*

The value of $V(\theta)$ varies slightly between 30 km/h and 50 km/h for a large range of directions. It is equal to $V^s(\theta)$ when 100% of the wind farms couples (p_1, p_2) with the given orientation satisfy $p_1 \prec p_2$ (i.e. when the proportion of positive weights is 100%). In the intermediate cases of $|\theta| \approx \pi/2$ propagations in opposite directions cancel out, leading to lower value of $V^s(\theta)$.

The agreement in the east directed propagations is accounted by V^s. The signed directional velocity can thus be used to define a main stream propagation direction and velocity:

$$\vec{V} = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} V^s(\theta) \vec{u}_\theta d\theta \tag{12}$$

where $\vec{u}_\theta$ is the unit vector in direction θ. In our case the obtained direction is around $-\pi/16$ and the associated speed is 23.66 km/h.

9 An alternative definition of propagation

The propagation analyzed in the preceding section and summarized by the main velocity given (Equation 12) is meant to summarize an average propagation of errors. However, it is not directly computed from any observed propagation of errors. In this section, we evaluate the propagation speed in the east west direction through a planar wave model for each run time t and compute the associated average propagation speed.

9.1 Propagating pattern

Let us consider here that the space-time function of the errors can be represented, at least locally in time, as a spatial pattern propagating for all position $p \in \mathcal{P}$ according to a translation at constant speed $\vec{v}_t$ along the look ahead time. Then, for $h = 1, \ldots, H$,

$$\epsilon^p_{t+h|t} = f_t(p - \vec{v}_t(h-15)) + \eta^p_{t+h|t} \tag{13}$$

where $f_t(\cdot)$ is the shape function of the propagating pattern, $\vec{v}_t$ the speed (vector of propagation and $\eta^p_{t+h|t}$ a centered noise. For a fixed speed vector $\vec{v}_t$, the shape function is defined to reach the minimum in the optimization problem:

$$\min_{f_t \in \mathcal{F}} \sum_{h=1}^{H} \sum_{p \in \mathcal{P}} \left(\epsilon^p_{t+h|t} - f_t(p - \vec{v}_t(h-15)) \right)^2 \tag{14}$$

where $\mathcal{F}$ is a class of smooth functions. We will denote $\hat{f}(\vec{v}_t)$ the result of this minimization, to emphasize the dependence of the shape function on the propagation speed $\vec{v}_t$. In this work, we used the bivariate local polynomials implemented in `loess` function in `R`. This type of propagation model, based on a planar wave description of the errors goes back to [37, 38]. Since then few work have built upon this. Among them are [39] in astrophysics or [40] in meteorology. Extension of the corresponding model have been considered, see e.g. [21] [22] and the references therein.

9.2 Estimation of speed

In Equation 13, the speed is to be fixed or specified, but of course it is unknown and must be estimated simultaneously to the estimation of $f(\vec{v}_t)$. Following [40], we will seek the speed which maximizes the fraction of variance of $\mathcal{E}_t = (\epsilon^p_{t+h|t})_{p \in \mathcal{P},\ h=1,\ldots,H}$ explained by the planar wave $\hat{f}(\vec{v}_t)$. We will thus maximize

$$R^2(\vec{v}_t) = 1 - \frac{\sum_{h=1}^{H} \sum_{p \in \mathcal{P}} [\epsilon^p_{t+h|t} - \hat{f}_t(p - \vec{v}_t(h-15))]^2}{\sum_{h,p} (\epsilon^p_{t+h|t} - \bar{\epsilon}_{t+\cdot|t})^2}, \tag{15}$$

where $\bar{\epsilon}_{t+\cdot|t}$ is the average on p and h of the error at time t.

In order to simplify the computations, the whole panel of possible directions is not considered here since our main interest is to compare the speed obtained by this approach to the propagatio velocity $\vec{V}$ defined in Equation 12, for which the direction is known to be westerly. Considering more directions would increase the size of the parameter space and require more complex optimization procedure without providing new insight on our dataset.

We thus imposed an East-West propagation direction, with a speed bounded in the interval $[-v_{\max}, v_{\max}]$, with $v_{max} = 250$ km/h. Note that negative speed would imply a propagation going from East to Weast. The optimization is done by simple grid search on this interval, with a 1 km/h resolution. For each speed considered, the optimal shape function $\hat{f}(\vec{v}_t)$ is found according to Equation 14. The propagation speed maximizing Equation 15, denoted $\hat{v}_t$, is then the optimal propagation speed at time t. For each time t considered, the result of this optimization procedure is a propagation speed $\hat{v}_t$, shape function $\hat{f}(\hat{v}_t)$ and a fraction of variance $R^2(\hat{v}_t)$.

9.3 Results

We performed this estimation procedure for the two years of prediction and 4 different look ahead horizons: $H \in \{90, 120, 150, 180\}$. The fraction of variance explained by the model in Equation 13 is variable and is not always very large. Figure 15 represents the mean and standard deviation of the estimated propagation speed as a function of the fraction of explained variance. For doing so, we first removed all cases for which the optimal speed was at the border of the search grid (i.e. $|\hat{v}_t| = 250$ km/h). Such estimated speeds were interpreted as a lack of fit of the model in Equation 13. They represented a total of 20% of the cases. On the remaining 80% cases, called the estimable cases, the average and the standard deviation of the estimated speeds corresponding to time t for which the associated fraction of variance $R^2(\hat{v}_t)$ was larger than a value δ were computed. For the representation of Figure 15, the threshold δ itself varied from 0% to 50% with a 1% increment. On this Figure is also represented the proportion of considered observations, for a given δ.

We can first note that the proportion of considered obervations decreases with the fraction of explained variance. There are far less cases in which the propagation model explains more 40% of the variance than cases in which this fraction is larger than, say 10%. This proportions decreases also with the horizon H. For example, the proportion of explained variance exceeds 20% for 49.9% of the cases for $H = 90$, and 34.3% for $H = 180$. These results show that this model coud be of interest for predictive purpose, in particular with the shorter horizon.

For $\delta = 0$, the i.e. when considering all estimable instants t the average propagation speed is around 30 km/h, in agreement with the value of $\vec{V}$ obtained in the previous section. As δ increases, the average speed increases and the standard deviation incresases also, but slowly. This indicates that situations with higher propagation speed are on average better described by the model in Equation 13 than situations with lower propagation speed. They are also slightly more variable.

In conclusion, the simple mode in Equation 13 is able to partly describe the spatio-temporal propagation of errors.

10 Spatio-temporal propagation of errors and wind

In the previous section, it was shown that there exists specific temporal lags and directions of higher correlation between errors. Whilst this cannot constitute a physical description of any weather situation, the spatio-temporal dependence structures exhibited in this work do suggest some sort of propagation. Predicting the exact evolution of such pattern is beyond the scope of this work. In this section we will analyze the main average propagation speed and direction as defined by Equation 12 in relation with wind data at a larger scale.

The used meteorological data are wind speed and direction obtained from the analysis of a mesoscale model at 3 different pressure levels: 500hPa, 750hPa, and 1000hPa on a grid point located in the center of $DK1$. Since these data are 3 hourly, we extrapolated them linearly to get a time step similar to the one power prediction errors (i.e. 15 minutes). This approximation is deemed sufficient for our purpose, which is exploring the role of large scale wind characteristics in the spatio-temporal correlation of prediction error. Wind roses are shown in Figure 16. They show that the wind is mainly oriented from East to West, and that wind speed is larger in this direction than in any other direction.

The propagation analysis in the correlation structure performed section 8 is now done conditionnaly to wind speed and direction. We get an average propagation speed and direction, as defined by Equation 12, conditionally on wind the variable defined above. In order to evaluate the relationship between wind and propagation, using wind at different model levels and even wind shear, we represent in Figure 17 the main average propagation direction conditionally on wind direction and main average propagation speed conditionally on wind speed.

The main average propagation direction is almost linearly explained by wind direction at 1000 hPa. Although not being linear at higher pressure levels, it is still strictly increasing. Figure 17 can be used as a tool to model the

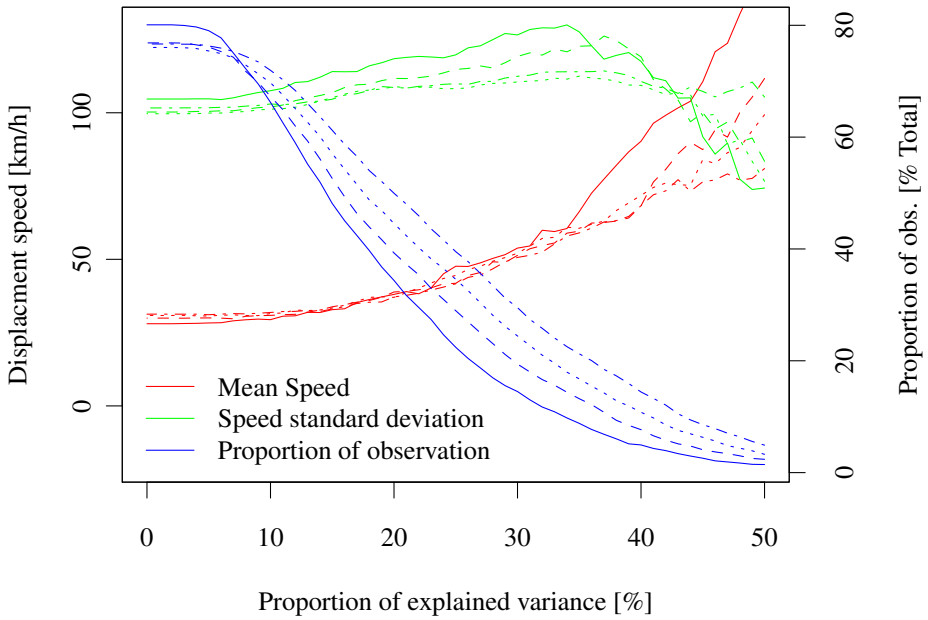

FIGURE 15: *Average Displacement speeds with* 4 *different maximum look ahead times represented with dashed-dotted lines for* $H = 90$ *minutes, dotted lines for* $H = 120$ *minutes, dashed lines for* $H = 150$ *minutes and solid lines for* $H = 180$ *minutes. Displacement speed maximizes the fit in the westerly propagation model given by Equation 13*

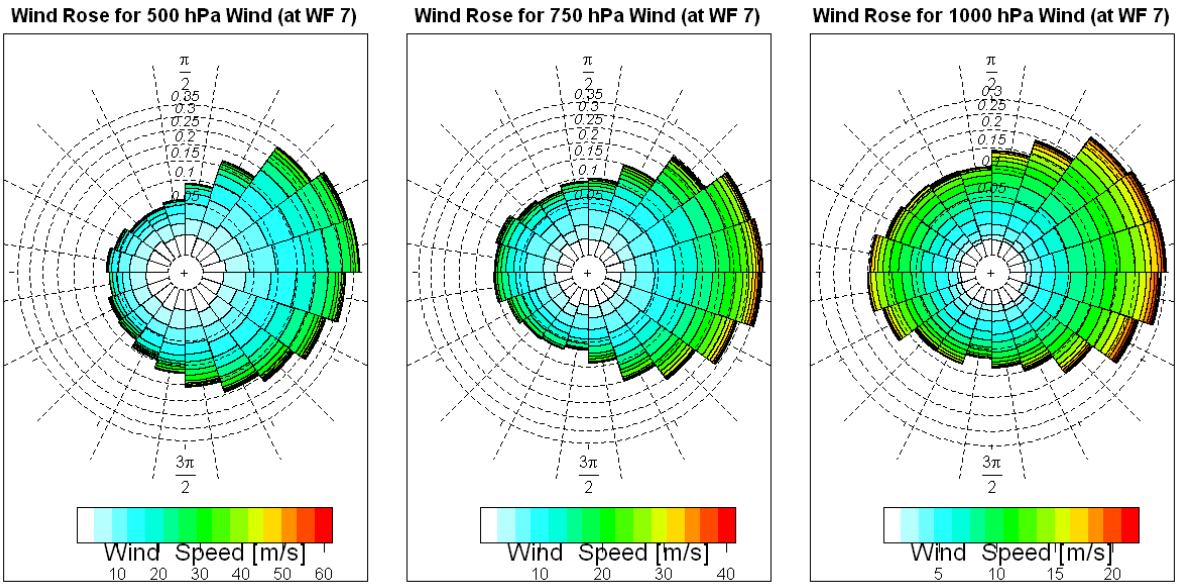

FIGURE 16: *Wind roses at 3 different model levels, colored area give the direction of the wind. These wind are obtained from a mesoscale model analysis. We have used the convention that a colored area around zero corresponds to wind blowing from the west. A colored area around* $\pi/2$ *corresponds to wind blowing from south.*

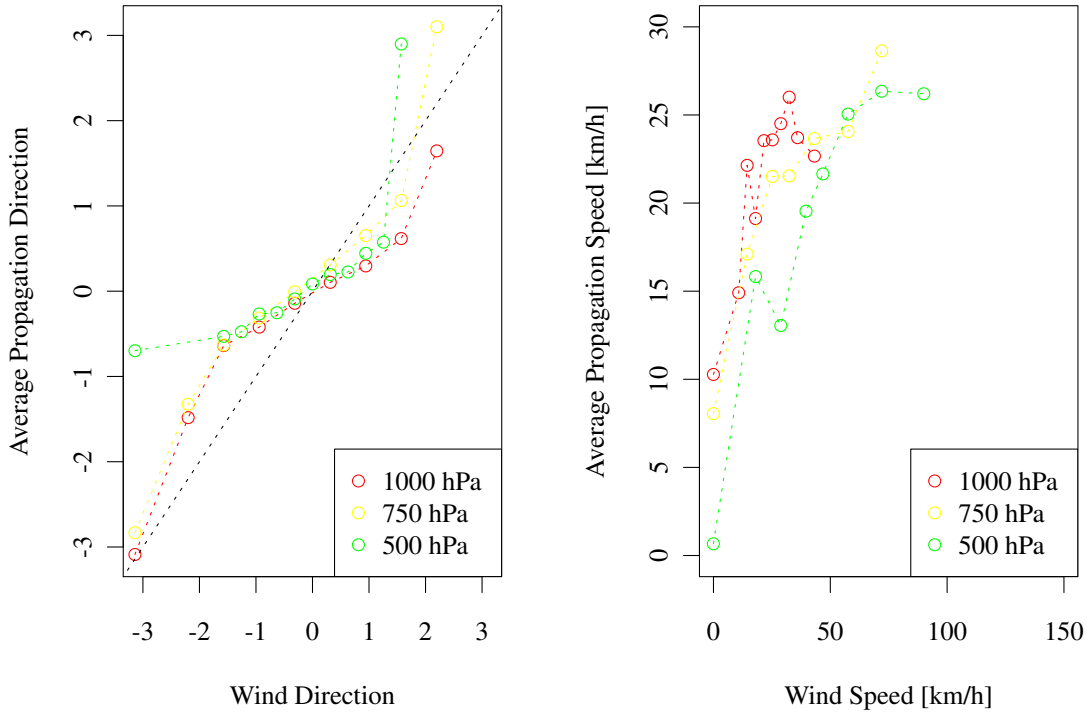

FIGURE 17: *Average characteristics of propagation of prediction errors conditionally on wind characteristics.*

relation between wind direction and the main average propagation direction. The relation between wind speed and main average propagation speed is also of interest. There is a direct, almost linear relationship between the average propagation speed and wind speed, up to 30 km/h for 1000 hPa wind, and 50 km/h for the 500 hPa wind. At higher wind speed, the propagation speed of the errors is more or less constant. This could be interpreted as a saturation effect of wind power predictions.

The same analysis was performed on thermal wind, defined and computed as the difference between wind values at different levels. The relation between propagation and thermal wind (not presented here) were much less strong and difficult to interpret. They did not present an interesting scope for additional modelling of the prediction errors.

11 Conclusions

Due to the fact that wind power prediction tools generate forecasts for individual sites only, without properly accounting for the information coming from the neighboring territories, there is potential for improving the forecast quality by including the spatio-temporal aspects into the models. In this report a CP-VAR model was first proposed for this purpose, ideally for the case of power generation aggregated by zones (i.e. control areas for instance). The model coefficients being recursively estimated smooth functions of the forecasted average wind direction permit to consider the influence of meteorological data on the spatio-temporal inter-dependence among the sites and also to account for the long-term variations in the process characteristics. It was shown that correspondingly corrected forecasts when evaluated on the test case of western Denmark result in a reduction of prediction errors up to 18.46% in terms of RMSE. The coefficients are based on a spatial average of the wind speed rather than a local wind speed for each group considered. This simplifies the problem and for a relatively small region like Denmark it is seen to be of no consequence.

The adjusted forecasts can be communicated along with estimates of their associated uncertainty. Predictive densities are modelled as truncated multivariate normal distribution. The evaluation exercise carried out showed that the proposed method results in reliable univariate probabilistic forecasts for each individual group. However, additional research is needed before concluding if the multidimensional forecast is calibrated with the estimated multivariate density. This is one of the points of interest for future work related to probabilistic forecasts issued with CP-VAR models.

The prediction approach and spatio-temporal concepts were analysed and validated for the western Denmark area, and by considering one-hour ahead predictions. The proposed methodology could be extended to the case of further look ahead times (up to several hours ahead) if working on the same terrain as Denmark. Having in mind that western Denmark is the first to be touched by fronts coming from North-West, the use of on-line measurements from the United Kingdom or from the measurement devices in the North Sea, might lead to significant improvements in making predictions for longer time horizons. In order to make models valid for data coming from a region with a more complex wind climatology than Denmark, some adjustments would have to be done in the modelling approach due to the fact that it is not always possible to use average wind direction as a representative of the situation in the whole region. For such more complicated cases it may be needed to switch from conditional parametric models to varying-coefficient models.

Since the CP-VAR model structure and complexity may be highly dependent upon the number of sites considered, it is of particular importance in the future to propose a modelling approach which would permit to easily include or exclude a new wind installation into the model without having to change the whole structure and to re-estimate the coefficients. A potential solution for this could be to employ a lattice approach.

The analysis of the spatio-temporal stochastic structure of wind power prediction errors in Denmark goes into that direction. We showed that errors at two distant wind farms for different look ahead times are in relation that is close to be monotone and that a linear model does not give a satisfactory modeling but permits to capture the main characteristic of the dependence structure along the look ahead times. For example, the correlation between two

distant wind farms is a smooth function f of the look ahead times. We noticed that f increases rapidly and that after a maximum correlation is reached it decreases slowly. The time for which the maximum correlation is obtained was studied in detail through its relation with distance and direction between the considered wind farms. This permits us to conclude on a privileged direction of propagation of errors which is westerly. We noticed that this direction is that of the average wind in the country. We then adopted a dual approach to characterize propagation of wind power forecast errors through the use of constant speed planar wave modeling. The average propagation speed obtained with this dual method appears to be similar to that obtained through the correlation analysis. Finally we analyzed the correlation of wind power prediction errors conditionally on wind and thermal wind respectively. We showed that wind direction and speed are connected to main propagation direction and speed. Wind shear seems to have less interest in explaining propagation direction and speed. Finally, this analysis permitted us to understand that the joint distribution of forecast errors at different location and for different look ahead time is not Gaussian, even if spatio-temporal modeling of Gaussian process has already witness some advanced contributions [41]. However, errors might be well modeled by a mixture of a Gaussian and a heavy-tailed distribution. We also observed that the interdependence of the forecast errors is well structured with respect to the temporal and spatial dimensions. These observations will be used as a starting point for further statistical spatio-temporal modeling of wind power prediction errors.

References

[1] Doherty R, O'Malley M. A new approach to quantify reserve demand in systems with significant installed wind capacity. *IEEE Transactions on Power Systems* 2005; **20**: 587-595.

[2] Castronuovo ED, Pecas Lopes JA. On the optimization of the daily operation of a wind-hydro power plant. *IEEE Transactions on Power Systems* 2004; **19**: 1599-1606.

[3] Angarita JM, Usaola J. Combining hydro-generation and wind energy: Biddings and operation on electricity spot markets. *Electric Power Systems Research* 2007; **77**: 393-400.

[4] Pinson P, Chevallier C, Kariniotakis G. Trading wind generation from short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems* 2007; **22**:

[5] Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitosa E. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews* 2008; **12**: 1725-1744.

[6] Giebel G, Kariniotakis G, Brownsword R, Denhardt M, Draxl C. The state of the art on short-term wind power prediction - A literature overview. 2nd Edition. Technical report, SafeWind project, Deliverable Report D1.4, 2011. [Available online: http://www.safewind.eu/].

[7] Monteiro C, Bessa R, Miranda V, Botterud A, Wang J, Conzelmann G. Wind power forecasting: State-of-the-art 2009. Technical Report, Argonne National Laboratory, Chicago, Illinois (USA), 2009.

[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 Analysis. 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] Tastu J, Pinson P, Kotwa E, Madsen H, Nielsen HAa. Spatio-temporal analysis and modeling of short-term wind power forecast errors. *Wind Energy* 2011; **14**: 43-60.

[13] Tastu J, Pinson P, Madsen H. Multivariate conditional parametric models for a spatio-temporal analysis of short-term wind power forecast errors. in *Proceedings of the EWEC'2010, European Wind Energy Conference*, Warsaw, Poland, 2010.

[14] Corradi O. Improving wind power forecasts by considering the spatio-temporal structure of wind power forecasts. *Poster presentation, GrønDyst Conference, Copenhagen, Denmark*, 2010.

[15] Nielsen TS, Madsen H, Nielsen HAa. Prediction of wind power using time-varying coefficient functions. in *Proceedings of the IFAC 2002, 15th World Congress on Automatic Control*, Barcelona, Spain, 2002.

[16] 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.

[17] 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.

[18] Horrace WC. Some results on the multivariate truncated normal distribution *Journal of Multivariate Analysis* 2005; **94**: 209-221.

[19] Gneiting T, Balabdaoui F, Raftery AE. Probabilistic forecasts, calibration and sharpness *Journal of the Royal Statistical Society, Series B* 2007; **69**: 243-268.

[20] Gneiting T, Stanberry LI, Grimit EP, Held L, Johnson NA. Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble predictions of surface winds *Test* 2008; **17**: 211-235.

[21] Ailliot P, Baxevani A, Cuzol A, Monbet V, Raillard N. Space-time models for moving fields with an application to significant wave height fields. *Environmetrics* 2011; **22**: 354–369.

[22] Ailliot P, Monbet V, Prevosto M. An autoregressive model with time-varying coefficients for wind fields. *Environmetrics* 2006; **17**: 107-117.

[23] Stoffer D. Estimation and identification of space-time ARMAX models in the presence of missing data. *Journal of the American Statistical Association* 1986; **81**: 762-772.

[24] Huang H-C, Cressie N. Spatio-temporal prediction of snow water equivalent using the Kalman filter. *Computational Statistics and Data Analysis* 1996; **22**: 159-175.

[25] Wallis K. Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts. *International Journal of Forecasting* 2003; **19**: 165-175.

[26] Morales JM, Minguez R, Conejo AJ. A methodology to generate statistically dependent wind speed scenarios. *Applied Energy* 2010; **87**: 843-855.

[27] Chew V. Simultaneous prediction intervals. *Technometrics* 1968; **10**: 323-330.

[28] Ravishanker N, Hochberg Y, Melnick EL. Approximate simultaneous prediction intervals for multiple forecasts. *Technometrics* 1987; **29**: 371-376.

[29] Ravishanker N, Wu LSY, Glaz J. Multiple prediction intervals for time-series: comparison of simultaneous and marginal intervals. *Journal of Forecasting* 1991; **10**: 445-463.

[30] Kolsrud D. Time-simultaneous prediction intervals for a time-series. *Journal of Forecasting* 2007; **26**: 171-188.

[31] Pinson P, Madsen H, Nielsen HAa, Papaefthymiou G, Klöckl B. From probabilistic forecasts to statistical scenarios of short-term wind power production. *Wind Energy* 2009; **12**: 51-62.

[32] Jordà Ò, Marcellino M. Path forecast evaluation. *Journal of Applied Econometrics* 2010; **25**: 635-662.

[33] Leutbecher M, Palmer TN. Ensemble forecasting. *Journal of Computational Physics* 2008; **227**: 3515-3539.

[34] Wand MP, Ripley B. Functions for kernel smoothing. R package version 2.23-4, 1995 [url = http://CRAN.R-project.org/package=KernSmooth].

[35] Wand MP, Jones MC. *Kernel smoothing*. Monographs on statistics and applied probability, Chapman and Hall: London, 1995.

[36] Sobol I. Sensitivity analysis for non-linear mathematical models. *Mathematical Modeling and Computational Experiment (Engl. Transl.)* 1993; **1**: 407-414.

[37] Briggs BH. On the analysis of moving patterns in geophysics–I. Correlation analysis. *Journal of Atmospheric and Terrestrial Physics* 1968; **30**: 1777-1788.

[38] Briggs BH. On the analysis of moving patterns in geophysics–II. Dispersion analysis. *Journal of Atmospheric and Terrestrial Physics* 1968; **30**: 1789-1794.

[39] Little LT, Ekers RD. A Method for analysing drifting random patterns in astronomy and geophysics. *Astronomy and Astrophysics* 1971; **10**: 306-309.

[40] Brillinger DR. *An application of statistics to meteorology: estimation of motion*. in Festschrift for Lucien Le Cam: research papers in probability and statistics, 1998.

[41] Gneiting T. Nonseparable, stationary covariance functions for space-time data. *Journal of the Americam Statistical Association* 2002; **97**: 590-600.