Improving **Wind Power Forecasts** by considering the **Spatio-Temporal structure** of Wind Power Forecast Errors

DTU Informatics
Department of Informatics and Mathematical Modelling

Olivier Corradi

1 Introduction and motivation

Denmark counted 24% of installed wind power capacity in 2008 [1] and aims at 50% in 2020 [2]. With the ongoing increase in installed wind power and due to the variability of this resource, forecasting production is becoming a critical process. So far, production forecasts have usually been generated individually for a given site of interest (single wind farms or a group of wind farms), without taking into consideration spatial or temporal correlations between those.
Because the production forecast is based on a meteorological forecast, errors in the meteorological forecast are transfered to the production forecast. Analysis show that patterns in the production forecast errors can be observed throughout multiple sites within a time window of a couple of hours [3].
This project aims at identifying those patterns, and develop a model capable of correcting the one hour ahead production forecast by considering the previous production forecast errors together with the meteorological forecast. This is investigated on the simple case of western Denmark, divided in 3 zones. The model is then tested with a division in 15 zones.
Part of this work was presented and developed under the collaborative project SafeWind, funded by the European Commission under the 7th Framework Program.

2 Observations and analysis

The dataset studied consists of weather forecast and wind power production data over the DK1 network (Jutland and Fynen) from 1 July 2006 to 25 October 2007 with a temporal resolution of 1 hour. As the production is obtained for each Energinet.dk transformer, data is aggregated according to zones (Figure 1) by using the positions of the wind mills connected to the transformer. Running the wind production forecasting model and comparing it to the actual production produces time series of **wind prediction errors** which will aim to minimize.

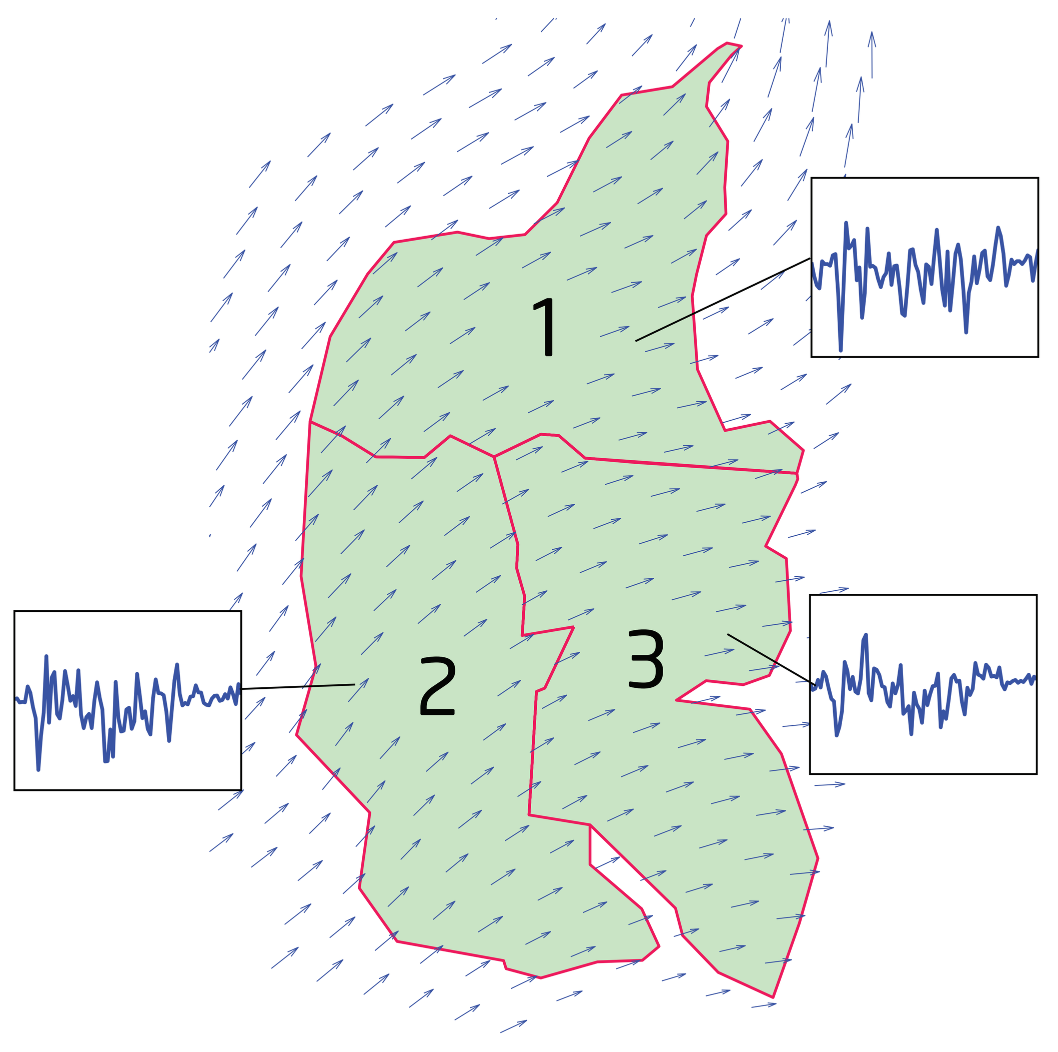

FIGURE 1: Division of the DK1 network into 3 zones. A time series of forecasting errors for each zone is presented together with a snapshot of the wind field at 10m height.

However, as the wind production forecasting model uses as input a weather forecast, it is intuitively expected that the forecasting errors propagate in time and space (Figure 2).

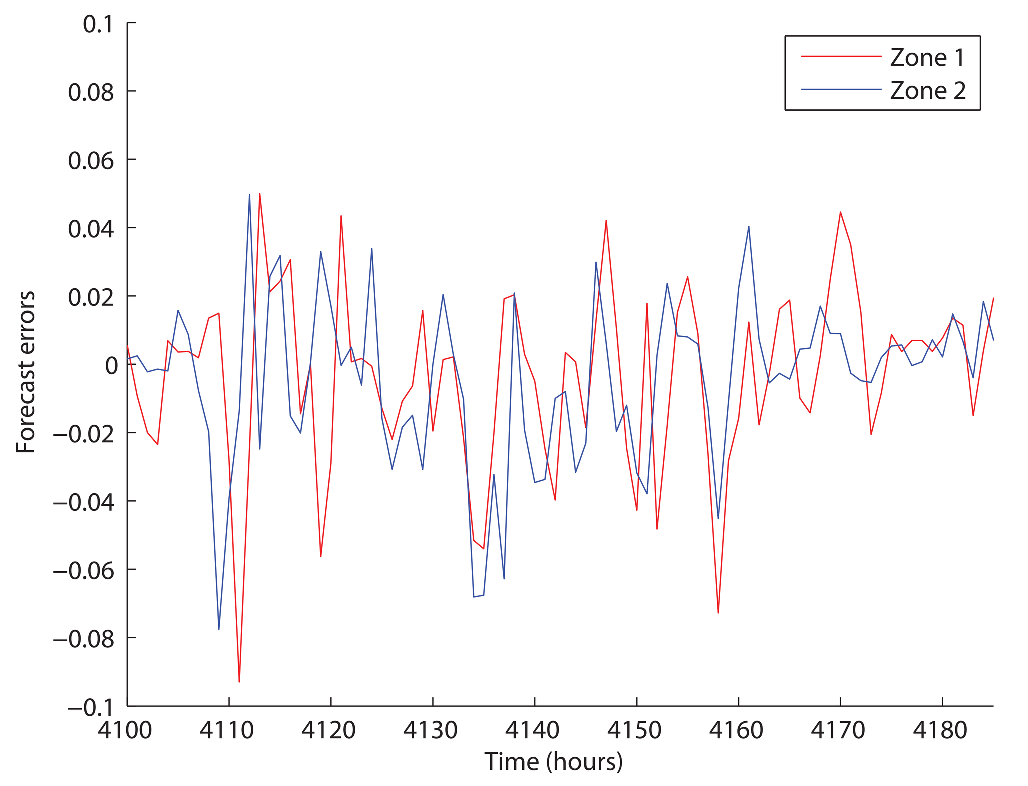

FIGURE 2: Forecasting errors seem to propagate from zone 2 to zone 1.

Furthermore, the correlations between zones seem to vary with the wind direction (Figure 3).

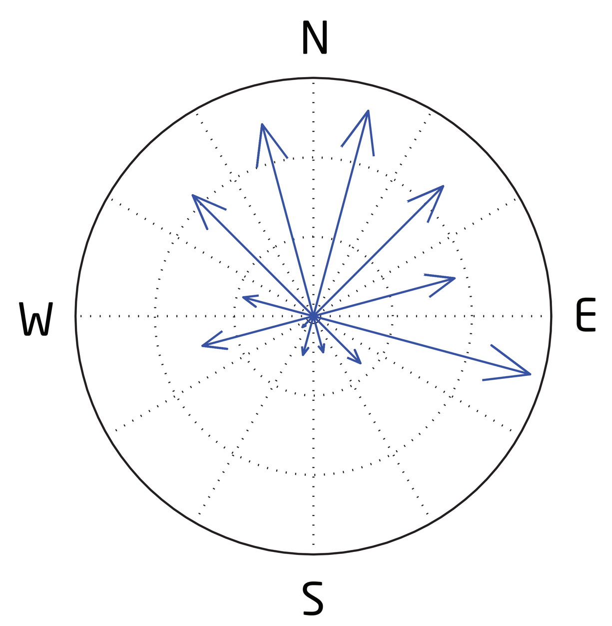

FIGURE 3: Significant wind directions correlating zone 1 and zone 2 with 1 hour delay. For each wind direction, the length of the arrow is the amount of correlation.

A model capturing this information and aiming at improving the wind power forecasts is then built.

3 Model building

Consider the column vector $\mathbf{x}_t$ consisting of the forecast errors of each zone at a certain time t. A first simple approach is to fit a linear model relating the errors in all the zones to themselves one hour ago:

$$\mathbf{x}_{t+1|t} = \mathbf{A}\mathbf{x}_t + \epsilon_t \tag{1}$$

where ϵ_t is multivariate Gaussian with zero mean. Such a model is called a Vector Auto-Regressive (VAR) model with lag time 1 hour.
However, the matrix $\mathbf{A}$ is constant and does not depend on meteorological conditions. Introducing an input signal w_t representing the wind direction averaged over the whole Denmark at time t, the model can be extended to a Conditional Parametric Vector Auto-Regressive (CPVAR) model by having the matrix $\mathbf{A}$ depend on a global wind direction w_t:

$$\mathbf{x}_{t+1|t} = \mathbf{A}(w_t)\mathbf{x}_t + \epsilon_t \tag{2}$$

This model will be labeled as CPVAR Global, as it uses as input a wind direction averaged over the whole Denmark. Note that the hypothesis that the wind field over Denmark is homogeneous has been made, meaning that the global wind field is a good approximation to the local wind field. A model that would use a locally averaged wind would have each element a_{ij} of the matrix $\mathbf{A}$ depending on a wind direction w_{ij} averaged over zones i and j:

$$\mathbf{x}_{t+1|t} = \begin{pmatrix} a_{11}(w_{11}) & \dots & a_{1n}(w_{1n}) \\ \vdots & \ddots & \vdots \\ a_{n1}(w_{n1}) & \dots & a_{nn}(w_{nn}) \end{pmatrix} \mathbf{x}_t + \epsilon_t \tag{3}$$

This model will be labeled as CPVAR Local, as a locally averaged wind direction has been used.
In the three models, the estimations are done by approximating the functions $a_{ij}(\cdot)$ of $\mathbf{A}$ with a polynomial function. The regression problem is then solved by using the Least Squares (LS) estimator.

4 Results

The Root Mean Squared Error (RMSE) criterion is used to assess the efficiency of a given model. The improvements are measured as the percentage of reduction of the RMSE in single-zone one hour ahead forecasts.

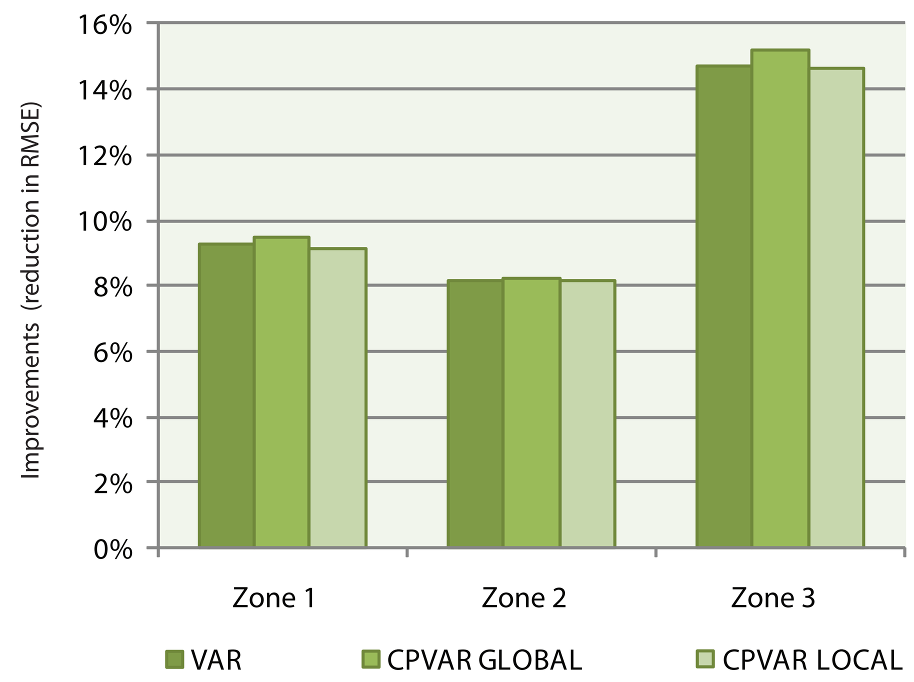

FIGURE 4: Major improvements are seen in zone 3 due to the dominance of wind blowing East/North-East in Denmark.

The CPVAR Global model (using a wind averaged over the whole country) yields the best results. 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). Applying this model to the DK1 network divided in 15 zones yields improvements up to 15% with the CPVAR Global model, whereas the other models are less robust and provide smaller improvements (up to 13%).

5 Conclusion and outlook

Combining the production errors and the forecasted wind direction in a statistical model permits to reduce the errors made on the wind production forecasts up to 15% depending on the zone. The model used can easily be implemented as an extra layer on existing forecasting systems, and an adaptive version of the CPVAR can even be used for real-time updating of the model coefficients.
The quality of the results encourages further developments like extending the model to include neighboring countries, integrating different time scales (based on the distance between zones for example) or investigating other meteorological variables like pressure, temperature or higher winds at 100m height. Those further development might give even more significant improvements in the wind power forecast.

References

[1] Danish Energy Authority, *Danish Annual Energy Statistics 2008*. September 2009.

[2] Energinet.DK, *The Danish Wind Case Fast Facts DK version 2008*. 2008.

[3] Julija Tastu, Pierre Pinson, Henrik Madsen, *Multivariate Conditional Parametric models for a spatio-temporal analysis of short-term wind power forecast errors*. 2010.