Technical University of Denmark

TradeWind Deliverable 2.2: Forecast error of aggregated wind power

Giebel, Gregor; Sørensen, Poul Ejnar; Holttinen, Hannele

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Giebel, G., Sørensen, P. E., & Holttinen, H. (2007). TradeWind Deliverable 2.2: Forecast error of aggregated wind power. (Risø-I, Vol. 2567).

Further Developing Europe's Power Market
for Large Scale Integration of Wind Power

Forecast error of aggregated wind power

Gregor Giebel, Poul Sørensen, Hannele Holttinen
Risø National Laboratory

April 2007

Agreement n.:	EIE/06/022/SI2.442659
Duration	November 2006 – October 2008
Co-ordinator:	European Wind Energy Association
Supported by:	Intelligent Energy Europe

Document information

Document Name:	Forecast error of aggregated wind power
Document Number:	Risø-I-2567(EN)
Author:	Gregor Giebel, Poul Sørensen, Hannele Holttinen
Date:	April 2007
WP:	WP2
Task:	Task 3
Revision:	
Approved:	

Diffusion list

All TradeWind Consortium Partners

Contents

Preface

This report is written in fulfilment of Task 2.3 in the TradeWind project (EU sponsored, under the Intelligent Energy Europe initiative): Wind Power Integration and Exchange in the Trans-European Power Market. The Task description is as follows:

Task 2.3: Forecast error of aggregated wind power

Estimates of forecast error of aggregated production for time horizons of intraday and dayahead markets in future will be produced. This will be done by reference to published studies of forecasting for wind generation, and from internal knowledge of WP2 participants. Modelling of wind power fluctuations for aggregated wind generation capacity.

Introduction

Short-term forecasting of wind power is an indispensable tool for every system operator and market actor in places where wind power penetration surpasses some 5 or 10% of the total demand. Therefore, already in 1990 the first tools had to be developed in Denmark. For a thorough overview of most literature before 2003, see Giebel et.al. [1].

In short, typical wind power forecasting systems for a horizon larger than a few hours ahead are based on Numerical Weather Prediction (NWP), i.e. results from the runs of full-blown weather models at operational meteorological centres. This is because any time-series analysis technique is typically worse than the accuracy achievable with NWP after 4-6 hours (in some cases even after two hours, as in the figure below). An example of performance of forecast errors dependency on forecast length is illustrated in Figure 1.

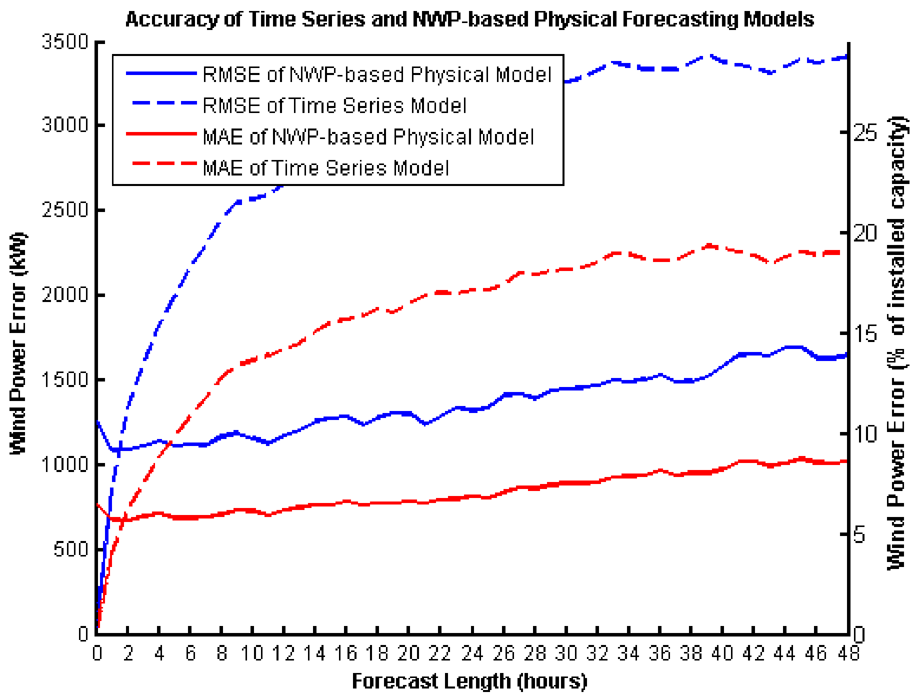

Figure 1. Performance of NWP-based physical prediction method compared to time series method.

Time-series analysis techniques are typically used for the first few hours though, since NWP model results tend to have an error even for the very short horizons. This is due to the fact that the data acquisition is not instantaneous (often it takes up to two hours until all data is available for the model run), does not cover the full area of the NWP model (eg over the oceans), and is too sparse to allow a perfect description of the atmospheres initial state. After the data acquisition cut-off, when most data is available, the model, typically set up to cover the whole globe or at least a quarter of it, runs for another two hours or so, which means that the newest and most accurate results of the NWP model are based on a four hour old snapshot of the atmosphere. Also, often the NWP model needs to "relax" into a self-consistent state after it has been forced by the new data, which takes up to six hours. So the accuracy of a NWP model is typically best after 3-6 hours, when it is comparable to the accuracy of the persistence model. The persistence model just states that the forecasted value (usually wind power) is the same as the value at the last measurement. For short horizons, this model does quite well, but since the only constant thing in the atmosphere is change, it gets successively worse for longer horizons. There are more advanced time series analysis models like ARMA, Neural Networks, Kalman Filters etc. However, they are only slightly better than persistence, often around 10% [1], due to an in-built averaging component, which drags them back to the mean generation. While this improves the statistical scores, it is not always useful for the client. That is why state-of-the-art systems for short-term forecasting have both a NWP based component and an autoregressive part based on time-series analysis techniques.

The NWP model results come as gridded values, with a horizontal resolution of some kilometres to few tens of kilometres, and with 30-40 values describing the vertical extent of the atmosphere, having a denser spacing close to the surface. This means that one has to interpolate or "downscale" the wind speed values coming as NWP forecasts to the actual wind power site. There are two principal approaches to that: the physical and the statistical approach. The statistical approach uses advanced statistical models, usually self-adapting ones like neural networks, support vector machines, Kalman filters or adaptive recursive least squares estimation with exponential forgetting factors, in order to directly estimate a power curve from the NWP wind speed and direction and the measured power. In the physical approach, one tries to use wind flow modelling with high-resolution models (meso-scale models like MM5 or WRF or micro-scale models like CFD or WAsP) to calculate the wind field in the area surrounding the wind farm. As those models only calculate the wind field, the physical models need an extra step to convert the wind speed to power. This can be done using the manufacturer's power curve, or one could employ the full-blown statistical models to do the task. As the NWP wind speed error is fairly independent of the forecasted wind speed, there is a filtering happening through the power curve, increasing error sensitivity between ca 4 and 11 m/s, and decreasing error sensitivity over 12 m/s up to the area of shutdown, or under about 3 m/s. Just imagine a 1 m/s error: in the flat parts of the power curve (zero power or full power), that error is irrelevant, while in the steep parts of the power curve, it gets amplified.

This yields a forecast for a single wind farm. Typical accuracies for those, measured as percentage of the installed capacity, are 7-20 % of RMS error over one year. For an overview of applicable error measures for short-term forecasting of wind power, see Madsen et al [2]. In short, not only the yearly average Root Mean Square (RMS) error or Mean Absolute Error (MAE) are applicable (often normalised with the installed capacity to yield the NMAE), but also the error correlation and the coefficient of determination R^2. The error depends often on the complexity of the terrain the wind farm is located in – more complex terrain is generally harder to predict by the NWP models [3].

1 Benchmark results for single wind farms

Typical forecast accuracies for single wind farms can vary quite dramatically. For the EU Anemos project, a comparison of 11 state-of-the-art tools was made for 6 sites in Europe [4], and the comparison shows that the differences between the wind farms, but also between the forecasting models are quite large.

Figure 2 shows the NMAE variation for each sites. The ALA test site is characterized as highly complex, SOT and GOL as complex, KLI and WUS as flat, and TUNO as offshore. The forecast errors are generally higher for more complex terrain, and the difference between the tools is also most significant for most complex terrain.

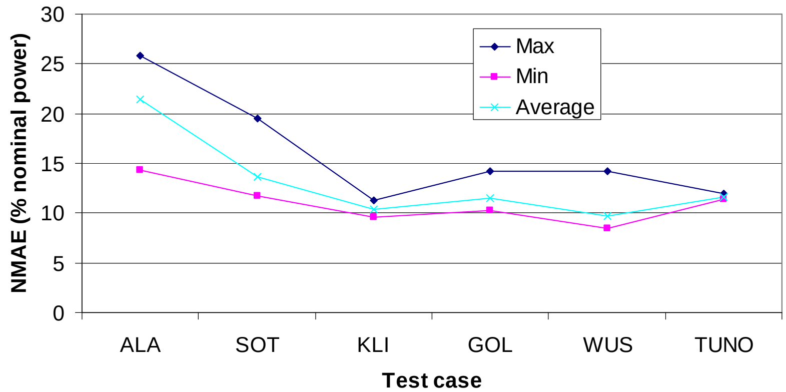

Figure 2. NMAE variation for each test case. 12 hours forecast horizon. Qualitative comparison.

For the best wind farms and for a large region like Germany, the average Mean Absolute Error for the day-ahead forecast can get down to about 5% of installed capacity.

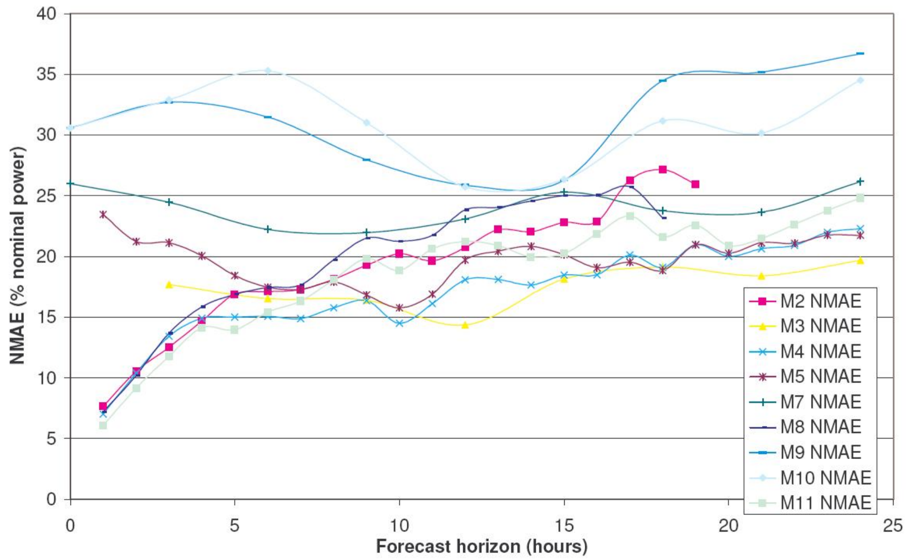

Figure 3. NMAE vs. forecast horizon with 10 different forecast systems for Alaiz wind farm in Spain characterized by highly complex terrain.

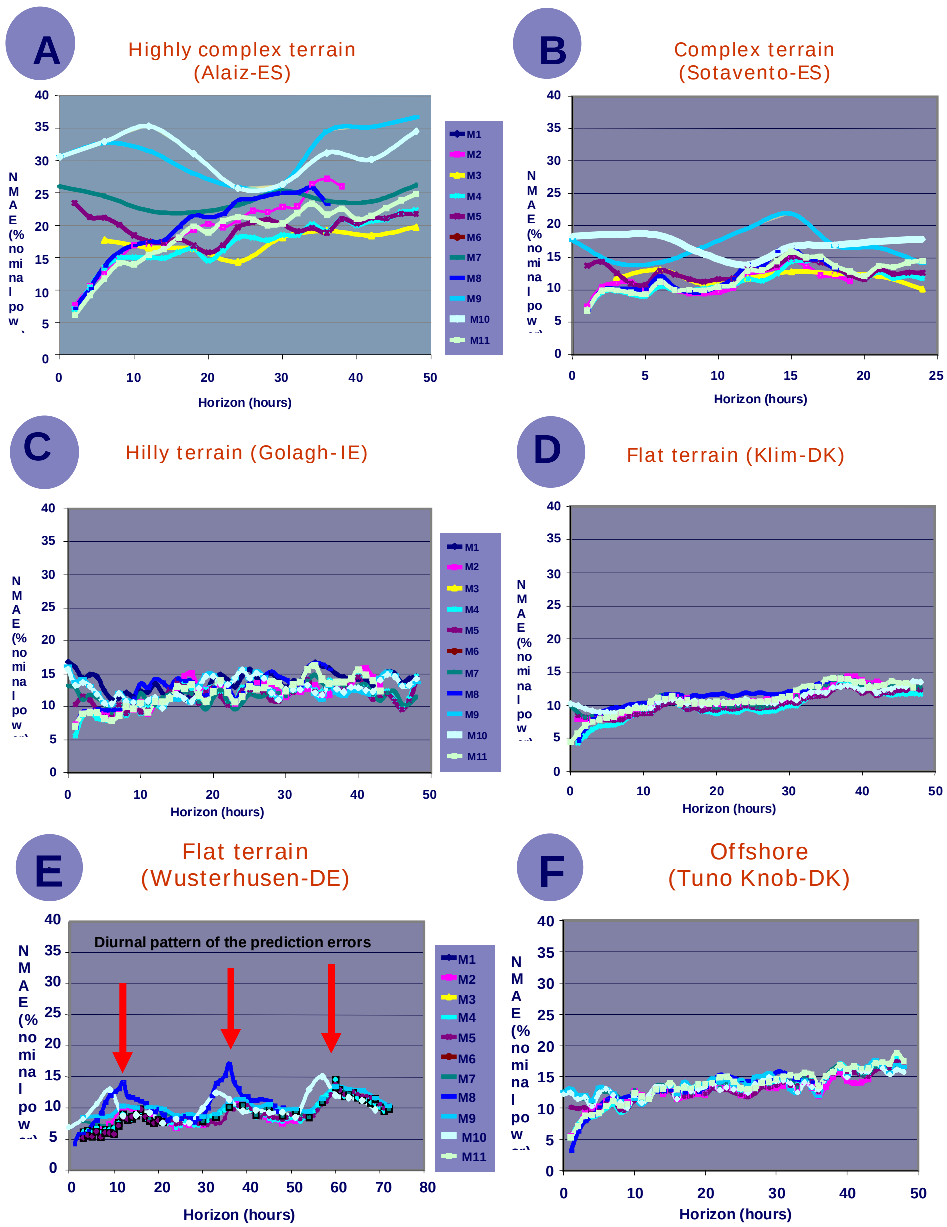

Figure 4. NMAE vs forecast horizon for 6 different wind farms from 10 different forecast systems.

Figure 3 shows the NMAE for the most complex site, Alaiz (ALA) in Spain. It is seen that the forecast errors are quite high, in some cases above 35 %. Also, it is clear that the forecast tools perform very different on this site. On other, less complex sites, comparisons showed less errors and more even performance of the different tools.

Figure 4 shows the NMAEs for 10 different forecast systems for 6 sites. The NWP input was the same for all forecasting models at a particular site. One can see various things in these plots. First of all, the performance of short-term forecasting in general is quite site-specific. Easy terrain is predicted quite well by the NWP model, and the quality of the short-term prediction model itself is not so determining for the result. Secondly, it is not always the same model which is best across horizons and across sites. Thirdly, some models contain autoregressive parts dependent on online data, and are therefore better for the very short horizons – see e.g. the case of Tunø Knob. Furthermore, some short-term forecasting systems model the daily variation in error explicitly, and therefore can get rid of the extreme diurnal pattern in Wusterhusen.

2 Upscaling and spatial smoothing

The results as described above are for single wind farms. However, many larger clients are more interested in the result for a region, be it an electrically defined region as for Transmission System Operators (TSOs) or a market region as for traders. In only very few cases, typically where wind power only took off in the last few years, there is online data available for all turbines in a region. In many cases though, like in Denmark, the production data for most wind turbines is only available from the accounting system for payments for the wind turbine owners, with a delay of up to a month. This means that for the purposes of an online forecast, it is useless. Therefore, a correlation has to be found between a few wind farms delivering online data within a region, and the much later determined total regional production. This approach is called *upscaling*.

The upscaling approach is illustrated for the WPPT forecasting system (Wind Power Prediction Tool, developed at the Technical University of Denmark and now sold by Enfor) in an application for an owner of large wind farms in Figure 5. According to [5], this configuration is used by a large wind farm owner in Denmark, and the installation has the following characteristics:

- A reasonable (less than 20) number of wind farms
- Online power production data is available for a number of wind farms.
- Offline production data with a resolution of 15 min. is available for almost all wind turbines. These offline data is released with a delay of 3-5 weeks.

As illustrated in the figure, the upscaling combines the present online production data with the historical offline data to predict the production.

Since not all wind farms in a region see the same wind speed at the same time, and since the error made by the NWP is temporally and spatially distributed, the error for forecasting a region is smaller than the error for a single wind farm. In this context it is interesting to investigate the spatial correlations between both the wind power generation and the wind power forecasting errors, as it is the uncorrelated part of the error which generates the error improvement due to spatial smoothing.

The variability of an averaged time series, e.g. expressed as the relative standard deviation of this time series, depends on the variabilities of the single time series, and on the correlation between the various series. For wind power forecasting, there are two effects which reduce the forecast error for a region in comparison to the one of a single wind farm: the generation as such is already smoother for a region due to the uncorrelated frequencies of the single wind farm generation profiles, making it thereby more easily predictable, and the forecast errors are uncorrelated on an even smaller length scale. For the former issue, refer to the literature overview given by Giebel [6]. In most studies, the generation correlation vanishes on a length scale of about 750 km.

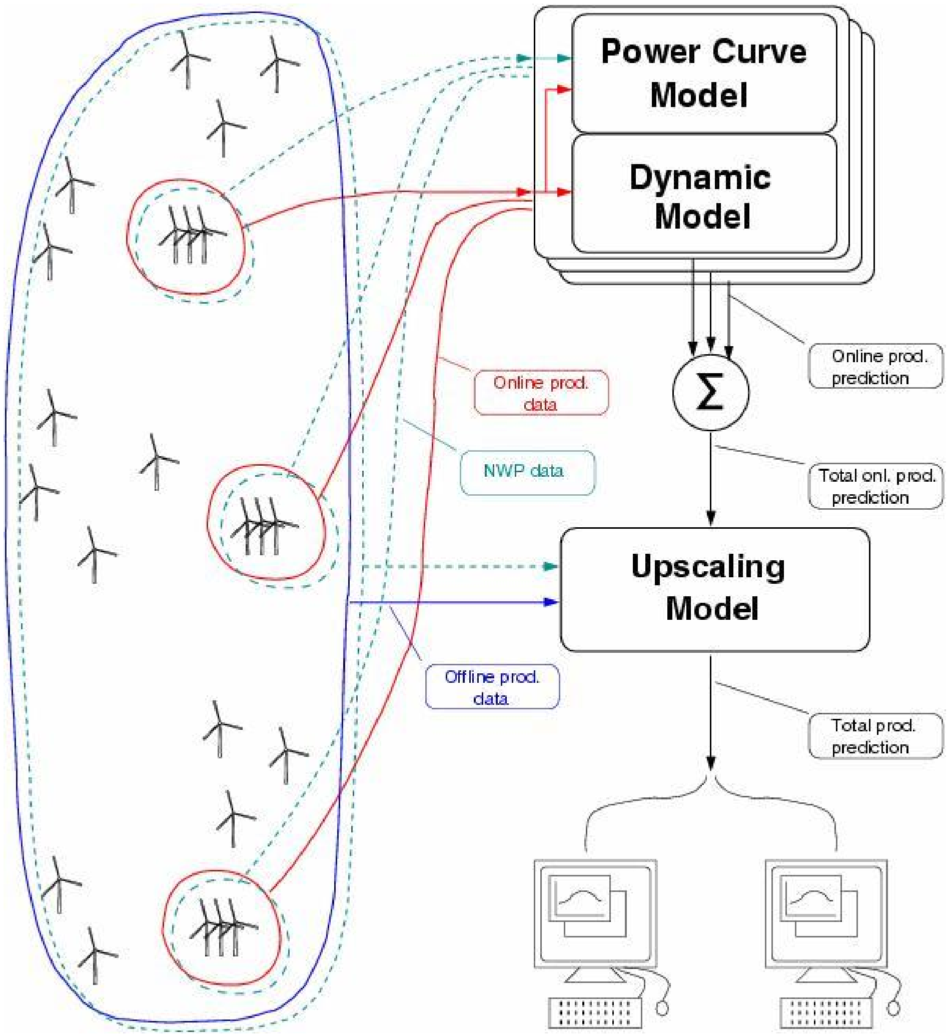

Figure 5. The WPPT forecasting system in an application where online production data is available only for some wind farms, and offline production data is available for most wind farms. Figure by Henrik Madsen, IMM/DTU.

Since the correlation between forecast errors is getting weaker with distance, the forecasts for a region are much more accurate than the forecast for single wind farms. This error reduction scales with the size of the region in question. Within this region, only a certain number of wind farms is needed to predict the power production in a region quite accurately. For regions, the error autocorrelation is also stronger on a time scale of days than for single wind farms.

There are only few published results to the reduction of the error due to spatial smoothing effects. Söder [7] tried to develop a simple simulation model for the error of distributed wind power forecasts for power system modelling purposes, but as his model errors were based solely on persistence forecasting (and not on NWP results), his conclusions have to be treated with caution.

Boone [8] studied wind speed forecast errors and used a simple ARMA(1,1) time series model to simulate the wind speed forecast errors for single wind speed, and studied cross-correlation coefficients between forecast errors on two wind speeds, to support models with multiple wind speeds.

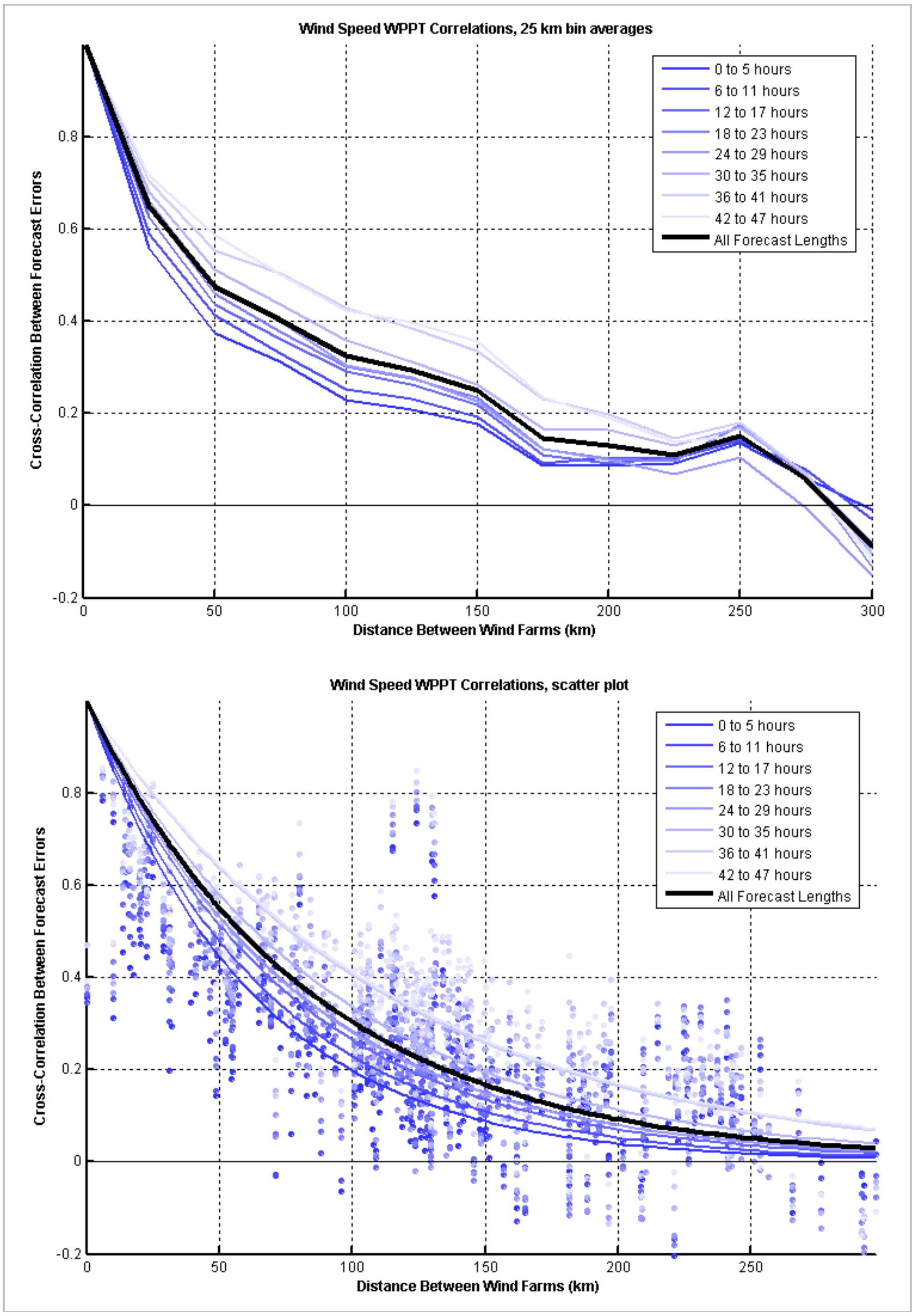

Figure 6. Correlations between wind speed forecast errors recorded in 2003 for 23 wind farms in western Denmark. In the upper plot, the correlations have been averaged over 25 km bins, while in the lower plot, each correlation is shown along with exponential fits. Darker shades refer to shorter forecast lengths. Figure is from Boone [8].

As a result of the cross-coherence study, Boone shows the following plots of cross-correlations and their decrease with distance. He used two different operational forecasting systems, WPPT and Prediktor, with two different NWP systems (the Danish HIRLAM model from the Danish Meteorological Institute and the German Lokalmodell of the Deutscher Wetterdienst) to develop an error simulation module for the Wilmar power system modelling tool [9]. Of those results, only the combination of DMI-HIRLAM and WPPT is shown here in Figure 6.

A least squares fit of the correlations for each set of forecast lengths has been made according to the exponential function

$$r = \exp(-d/\lambda) \tag{1}$$

where r is the cross-correlation, d is the distance between the wind farms, and λ is giving the relevant length scale. The estimated length scale increases with the forecast horizon from 62 km for the 0-5 h horizon to 113 km for the 42-47 h horizon, with an average of 81 km.

The methodically most relevant study on the subject was made by Lange [10] and Focken [11]. They applied power measurements on 30 wind farms in Germany to study the accuracy of the aggregated power output of wind farms distributed over given regions.

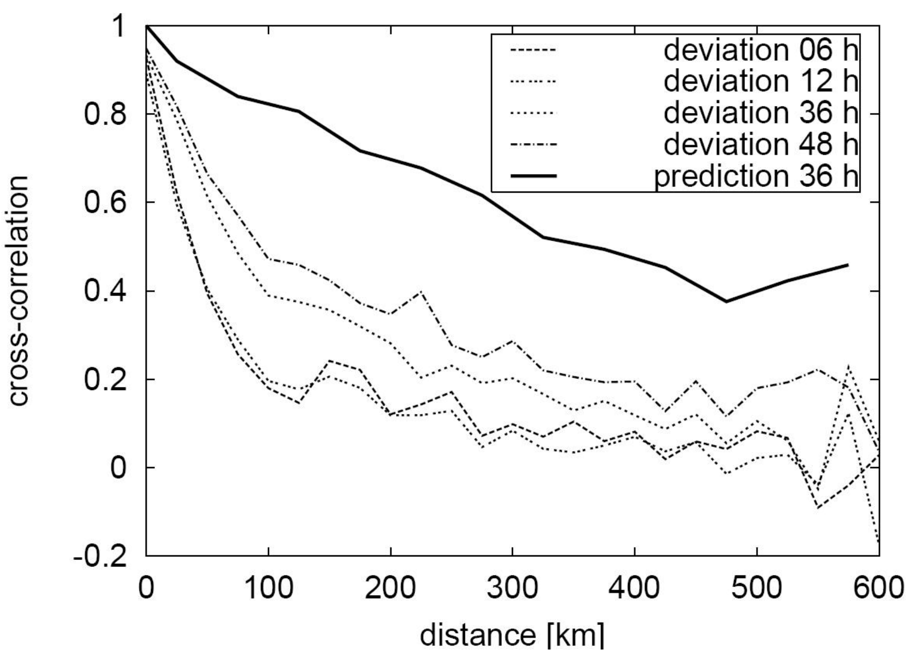

Figure 7: Spatial cross-correlation of prediction deviations for various prediction times based on German data for the years 1996–1999. For comparison the cross-correlation coefficients of the prediction (36 h) are also shown. All cross-correlation coefficients have been averaged over 25km bins. The figure is provided by M. Lange, energy & meteo systems GmbH.

One of the results of Lange and Focken's studies is the calculated cross-correlations shown in Figure 7, using a prediction method based on NWP results. The German results exhibit significantly longer distances than the Danish results in Figure 6. Comparing the Danish and German results, they agree quite well for distances less than 100 km. For distances above 100 km, the Danish results show less cross-correlation than the German, but the cross-correlations are relatively low at those distances, especially for the shorter forecast horizons (6-12h). For longer forecast horizons (36-48h), the cross correlation decays slower with distance, especially for the German results.

According to Focken et.al. [11], the increased cross-correlation for increased forecast horizons might be due to the growing systematic errors for increasing forecast horizon which give rise to higher spatial correlations. For comparison the cross-correlation coefficients for the 36 h power prediction have been calculated in the same way and are shown in [11] as well.

Lange and Focken have also analyzed normalized standard deviations of forecast errors. The standard deviations are normalized with the rated power of the corresponding wind power. If an ensemble consists of a number N of wind farms, then the relative standard deviation σ_{ensemble} of the ensemble forecast error can be calculated according to

$$\sigma_{\text{ensemble}} = \sqrt{\frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1}^{N} r_{xy}\sigma_x\sigma_y} \qquad (2)$$

where σ_x is the relative standard deviation of the forecast error of wind farm x power, and r_{xy} is the cross correlation coefficient between forecast errors on wind farms x and y.

Lange and Focken used the corresponding standard deviation of a single wind farm defined as the average standard deviation, i.e.

$$\sigma_{\text{single}} = \frac{1}{N}\sum_{x=1}^{N}\sigma_x$$

and then the standard deviation ratio σ_{ensemble} / σ_{single} is a measure for the reduction of the relative forecast error.

Lange and Focken have calculated the standard deviation ratio for different prediction horizon times as shown in Figure 8. The three curves represent different sizes of regions, with diameters 140, 350 and 730 km. It is seen that the deviation ratio depends quite little on the prediction horizon, while it depends significantly on the region size.

Lange and Focken also analyzed the relation between forecast error standard deviation ratio and the region size as given in Figure 9. The result indicates that the standard deviation ratio decreases exponentially with the region size.

For three wind farms in the UK with a maximum separation of 450 km, Parkes et al [13] report a portfolio effect of a 5% reduction in NMAE, from about 15% for the day-ahead forecast for a single wind farm to about 10% for the prediction for all three wind farms. This led to a potential saving of £3/MWh. The portfolio effect of three wind farms in Spain with maximum separation of 600 km also yielded a reduction in NMAE of about 5%, this time from about 20% for a 20-hour horizon down to about 15%.

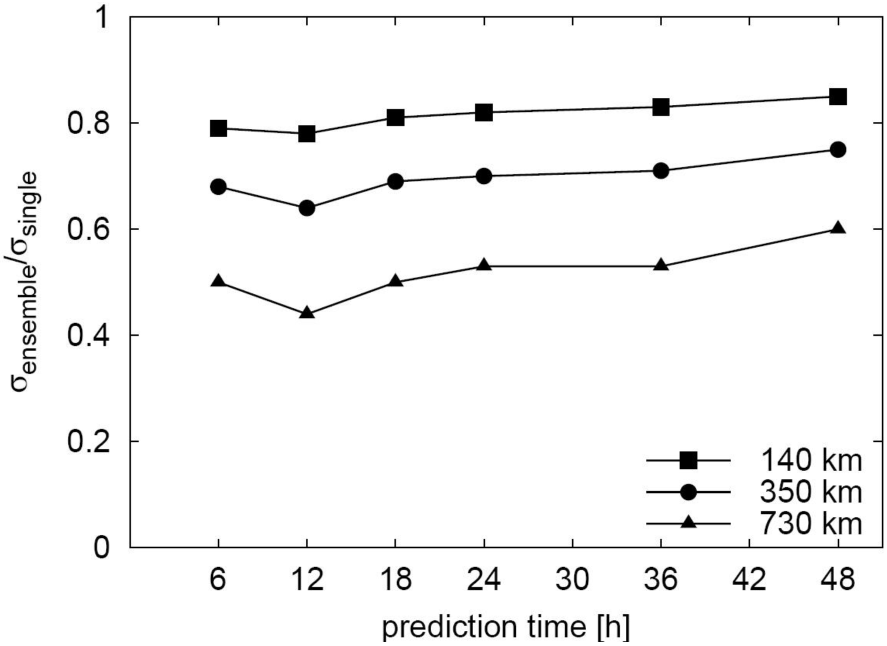

Figure 8. Forecast error standard deviation ratio versus prediction horizon for regions with diameters 140 km, 350 km and 730 km respectively. The figure is provided by M. Lange, energy & meteo systems GmbH.

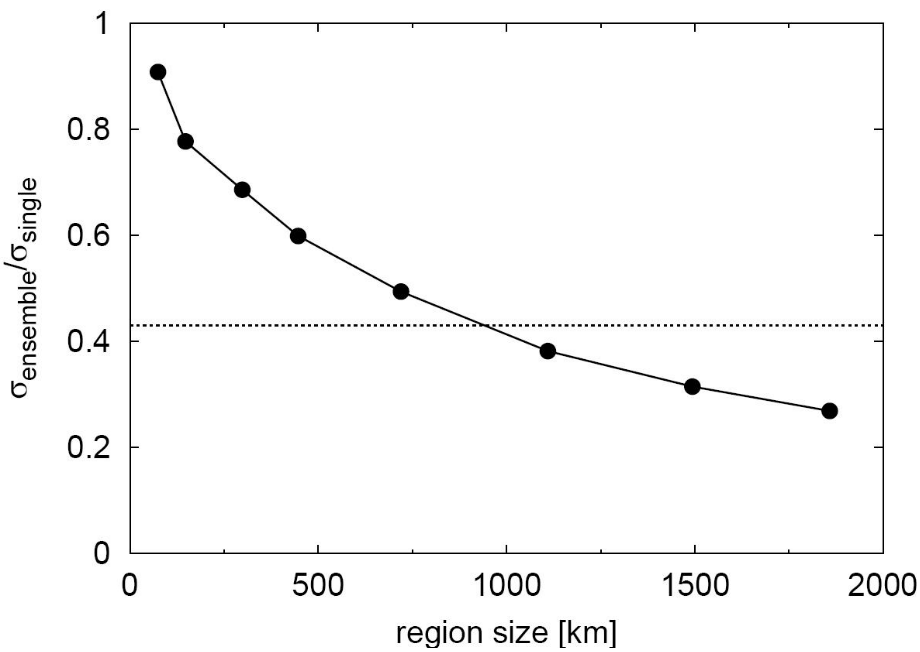

Figure 9. Forecast error standard deviation ratio versus region size quantified by the region diameter. The horizontal line gives the expected error reduction for an area the size of Germany. The figure is provided by M. Lange, energy & meteo systems GmbH.

Rohrig presents the German experience from the day-ahead forecast (24h to 48h ahead regarding the start of forecast model at the weather service): Single Wind Farm: 10 % to 20 % (RMSE % of nominal capacity) - Single Control Area: 7.5 % to 10 % - All Control Areas (whole Germany): 5% to 6.5%. Further reductions can be expected from combining different forecasting models: The first results from Germany show the best model performing at 5.1 % RMSE, a "simple" combination 4.2 % and "intelligent" combination 3.9 %.

Table 3. Level of accuracy of wind power predictions in Germany (NRMSE = normalized root mean square error, % of installed wind capacity). Source: Rohrig [12].

NRMSE [%]	Germany (all 4 control zones) ~1000 km	1 control zone ~ 350 km
day-ahead	5.7	6.8
4h ahead	3.6	4.7
2h ahead	2.6	3.5

Likewise, for Finland Holttinen et al. [14] present a reduction in forecasting error from up to 16% for the single site 24-h ahead forecast down to about 10% for the total error of four wind farms with a maximum spacing of about 380 km. They write: *The Mean Absolute Error (MAE) normalized by installed capacity is between 11 –15 % for 12 hours ahead for one site. Assuming the same installed wind power capacity in all 4 sites this drops the forecast error to 9%. For 36 hours ahead, one site errors are 13 –18 % and aggregated error drops to 11%.*

Figure 10 shows the forecast errors for one and four sites respectively, versus the forecast horizon.

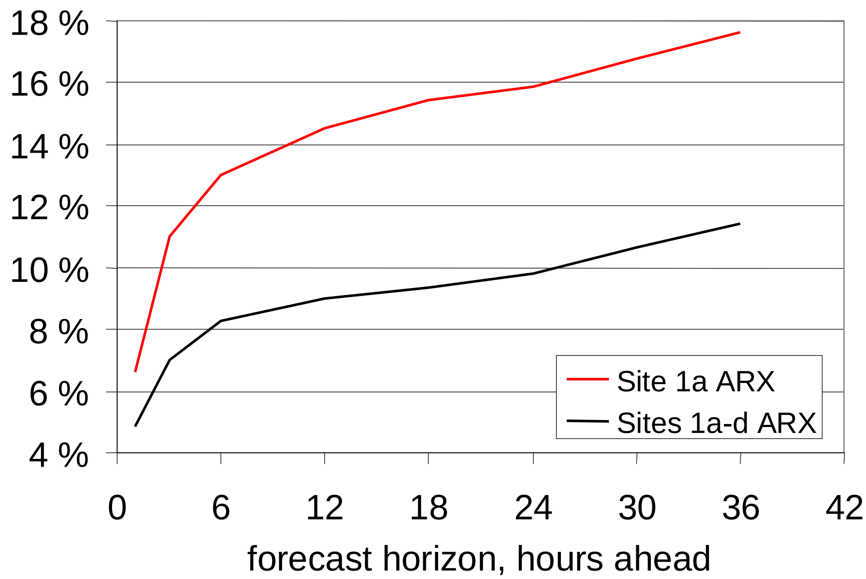

Figure 10. Mean absolute forecast error in % of capacity – year 2004 Finland. The graph is from Holttinen et al. [14]

Upscaling has also been a topic in the Anemos project [15]. For Jutland, a reduction down to 6.2% NMAE is reported, while for Ireland, the error only reduces to 11.6%. A larger report on the topic is upcoming.

Within that framework, Siebert and Kariniotakis [16] have looked into the optimal number of reference wind farms for the Jutland/Fyn area. Out of a total of 23 available wind farms, the optimal number of reference wind farms was shown to be only 5. This surprisingly low number is a combination of the sufficient coverage of those 5 farms of the main meteorological regions in the area, plus the very good data quality those 5 could offer. More wind farms would have led to more noise in the input signal for the upscaling algorithm.

3 Advances in Short-term Forecasting

Recently, a few papers have been published on the increasing quality of short-term prediction services during the last years.

The ISET (Institut für Solare Energieversorgungstechnik e.V., Kassel, Germany) was the first short-term forecasting provider for transmission system operators in Germany. In a paper for the EWEC 2006, Lange et. al. [17] presented the following plot for the accuracy in the E.On control zone.

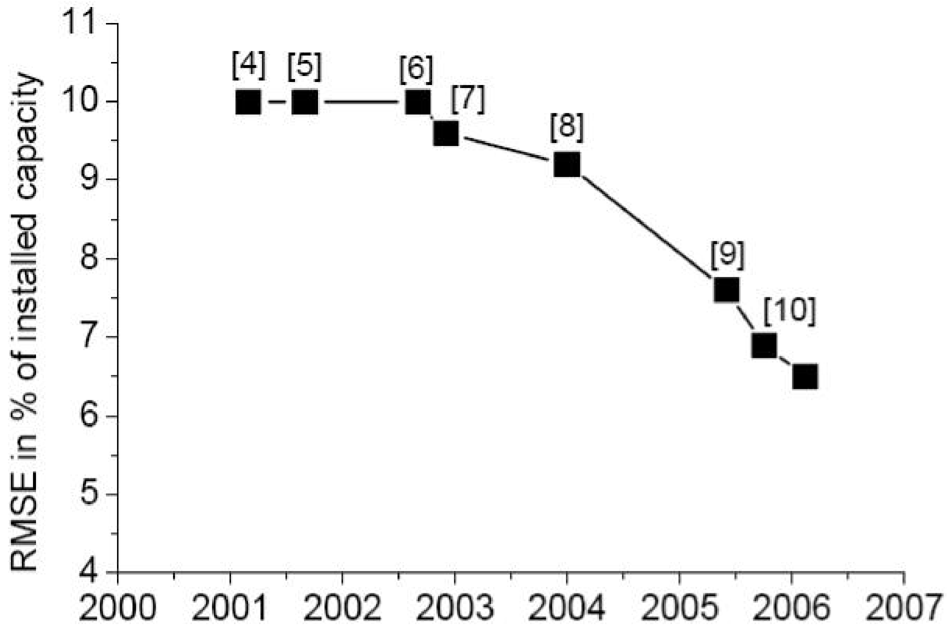

Figure 11: The development of the forecast error during the last years in the E.On Netz area. The numbers in square brackets are references from Lange et.al.

A similar plot, though constrained to the last two years, was shown by Krauss et.al. [18] for the EnBW TSO area. They show the monthly accuracy of three different forecasting systems for the aggregate error, and conclude that there are significant changes in forecast accuracy from month to month, and that the ranking of the three models changes from month to month as well.

4 Conclusions

Short-term forecasting of wind power has errors depending on the horizon, on the complexity of the site, the level of the predicted wind speed and on the quality of the NWP model and the short-term prediction model, to name just a few.

Substantial Benchmark tests in the EU ANEMOS project of different prediction tools on different sites have shown that the prediction tools perform very different on a highly complex terrain, while the performance is more equal on less complex terrain. Over the last years, prediction accuracy has improved steadily.

The cross correlation between forecast errors on different sites fit reasonably to an exponential decay. This correlation decay implies that the larger the region, the larger the smoothing of the error. This is theoretically founded, and found in real life, too. Results from actual portfolios of wind power show reductions of about 5% NMAE for wind farms distributed over some 400 km. The error reduction is larger for the shorter horizons.

Acknowledgements

We thank M. Lange for providing his figures in printable resolution.

References

[1] Giebel, G., G. Kariniotakis, R. Brownsword: *The State-of-the-Art in Short-Term Forecasting of Wind Power - A Literature Overview.* Position paper for the Anemos project, to be downloaded from anemos.cma.fr. 2003

[2] Madsen, H., P. Pinson, G. Kariniotakis, H.Aa. Nielsen, T.S. Nielsen: *Standardizing the Performance Evaluation of Short-term Wind Power Prediction Models*. Wind Engineering **29**(6), pp. 475-489, 2005

[3] G. Kariniotakis, I. Martí, D. Casas, P. Pinson, T. S. Nielsen, H. Madsen, G. Giebel, J. Usaola, I. Sanchez, A.M. Palomares, R. Brownsword, U. Focken, M. Lange, J. Tambke, P. Louka, G. Kallos, C. Lac, G. Descombes, G. Sideratos: *What performance can be expected by short-term wind power prediction models depending on site characteristics?* Paper for the EWEC Conference London (UK), Nov. 2004

[4] I. Martí, G. Kariniotakis, P. Pinson, I. Sanchez, T. S. Nielsen, H. Madsen, G. Giebel, J. Usaola, A.M., Palomares, R. Brownsword, J. Tambke, U. Focken, M. Lange, G. Sideratos, G. Descombes. *Evaluation of Advanced Wind Power Forecasting Models – Results of the Anemos Project.* EWEC 2006, Athens (GR).

[5] http://www.enfor.eu/wind_power_prediction_tool_wppt.php

[6] Giebel, G.: *Wind Power has a Capacity Credit -A Catalogue of 50+ supporting Studies.* WindEng EJournal, windeng.net, 2005

[7] Söder, L., 2004, "Simulation of Wind Speed Forecast Errors for Operation Planning of Multi-Area Power Systems", 8th International Conference on Probabilistic Methods Applied to Power Systems, 12 - 16 September 2004, Iowa State University, United States

[8] Boone, A.: Simulation of Short-term Wind Speed Forecast Errors using a Multi-variate ARMA(1,1) Time-series Model. Masters thesis at the Kungliga Tekniska Högskolan, Stockholm, 95 p., 2005

[9] See e.g. www.wilmar.risoe.dk and the references linked from there

[10] Lange, M., U. Focken: *Physical Approach to Short-Term Wind Power Prediction.* Springer, Berlin / Heidelberg 2006, ISBN 3-650-25662-8

[11] Focken, U., M. Lange, K. Mönnich, H.-P. Waldl, H.G. Beyer, A. Luig: *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.* J. Wind Eng. Ind. Aerodyn. **90**(3), pp. 139-249 (March 2002), DOI 10.1016/S0167-6105(01)00222-7

[12] Rohrig, K, (ed) 2005. Entwicklung eines Rechenmodells zur Windleistungsprognose für das Gebiet des deutschen Verbundnetzes,

Abschlussbericht Forschungsvorhaben Nr. 0329915A, gefördert durch Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit (BMU). Kassel, Germany, 2005.

[13] Parkes, J., J. Wasey, A. Tindal, and L. Munoz: *Wind Energy Trading Benefits Through Short Term Forecasting.* European Wind Energy Conference and Exhibition, Athens (GR), 27.2.-2.3. 2006

[14] Holttinen, H., P. Saarikivi, S. Repo, J. Ikäheimo, and G. Koreneff: *Prediction Errors and Balancing Costs for Wind Power Production in Finland.* Global Wind Power Conference, Adelaide, 2006

[15] Nielsen, T.S., H. Madsen, H. Aa. Nielsen, P. Pinson, G. Kariniotakis, N. Siebert, I. Marti, M. Lange, U. Focken, L. von Bremen, P. Louka, G. Kallos, G. Galanis: *Short-term Wind Power Forecasting Using Advanced Statistical Methods*. European Wind Energy Conference and Exhibition, Athens (GR), 27.2.-2.3. 2006

[16] Siebert, N., and G. Kariniotakis: *Reference wind farm selection for regional wind power prediction models.* European Wind Energy Conference and Exhibition, Athens (GR), 27.2.-2.3. 2006

[17] Lange, B., K. Rohrig, B. Ernst, F. Schlögl, Ü. Cali, R. Jursa, and J. Moradi: *Wind power prediction in Germany – Recent advances and future challenges.* European Wind Energy Conference and Exhibition, Athens (GR), 27.2.-2.3. 2006

[18] Krauss, C, B. Graeber, M. Lange, U. Focken: *Integration of 18GW Wind Energy into the Energy Market – Practical Experiences in Germany*. Workshop on the Best Practice in Short-term Forecasting, Delft (NL), 25 October 2006.