Intelligent wind power prediction systems
– final report –

Henrik Aalborg Nielsen (han(at)imm.dtu.dk)
Pierre Pinson (pp(at)imm.dtu.dk)
Torben Skov Nielsen (tsn(at)imm.dtu.dk)
Lasse Engbo Christiansen (lec(at)imm.dtu.dk)
Henrik Madsen (hm(at)imm.dtu.dk)
Informatics and Mathematical Modelling
Technical University of Denmark
Lyngby

Gregor Giebel (gregor.giebel(at)risoe.dk)
Jake Badger (jake.badger(at)risoe.dk)
Xiaoli Guo Larsen (xiaoli.guo.larsen(at)risoe.dk)
Wind Energy Department
Risø National Laboratory
Technical University of Denmark
Roskilde

Hans Ravn (hansravn(at)aeblevangen.dk)
RAM-løse edb
Smørum

John Tøfting (johto(at)dongenergy.dk)
Lars Voulund (larvo(at)dongenergy.dk)
DONG Energy A/S
Skærbæk

June 29, 2007

PSO Project number:	6503 (FU 4101)
Ens. journal number:	79029-0001
Project title:	Intelligent wind power prediction systems

Contents

Preface

This report summarizes the results of the project *Intelligent wind power prediction systems* which has been conducted over the period May 2004 to April 2007. The project was financially supported by the Danish utilities PSO fund (FU-4101) and with Informatics and Mathematical Modelling, Technical University of Denmark as the entity responsible for the project. The further participants in the project were Risø National Laboratory (part of the Technical University of Denmark as of January 1, 2007), RAM-løse edb, Elsam Kraft A/S and Energi E2 A/S (now DONG Energy A/S). Furthermore, Vattenfall participated in some of the last project meetings. The transmission companies, first Elkraft System and then Energinet.dk, followed the progress and attended meetings. With the aim of fertilizing international cooperation, parts of the project have been conducted in collaboration with sub-contractors. Specifically, CENER in Spain and VTT in Finland. The project budget was DKK 3.37 million, of which DKK 2.07 million was granted under the PSO rules.

Measurements of power production were provided by the project participants. Furthermore power data have been provided by ACCIONA in Spain. Meteorological forecasts have been provided by the Danish and Spanish meteorological institutes and the National Centers for Environmental Prediction (USA). The forecasts from Deutscher Wetterdienst were obtained commercially via financial support by the Danish PSO fund under contract 101295 (FU-2101). Furthermore, CENER in Spain provided Alaiz wind farm data and forecasts based on MM5 modelling and LocalPred for Spain and for the meteorological forecasts from the National Centers for Environmental Prediction for Denmark. Data on power prices were provided by Nord Pool.

The first experiments using mesoscale modelling for the Aliaz wind farm were conducted under the EU funded ANEMOS project (ENK5-CT-2002-00665), see Giebel et al. (2006). The Intelligent Prognosis project has benefitted from the experience gained in that earlier project. Karlsruhe University/Research Center Karlsruhe is acknowledged for permission to use KAMM. As part of the collaborative partnerships developing the mesoscale model WRF, the following institutions in the U.S.A. are acknowledged; the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (the National Centers for Environmental Prediction (NCEP) and the Forecast Systems Laboratory (FSL), the Air Force Weather Agency (AFWA), the Naval Research Laboratory, Oklahoma University, and the Federal Aviation Administration (FAA).

The project and the results were presented at the following conferences:

- European Wind Energy Conference & Exhibition 2006, Athens, Greece, 2006 (poster) (Nielsen et al., 2006b).
- Global Windpower 2006, Adelaide, South Australia (poster) (Nielsen et al., 2006b).

- European Wind Energy Conference & Exhibition 2007. Milan, Italy (talk and corresponding paper in the Scientific Proceedings) (Nielsen et al., 2007b).

At the conference in 2007, most of the main results of the project were presented.

The following publications were produced as part of the project:

Badger et al. (2007): Badger J., Giebel G., Larsen X.G., Nielsen T.S., Nielsen H.Aa., Madsen H., and Tøfting J. Report on the use of stability parameters and mesoscale modelling in short-term prediction. Technical Report Risø-R-1614, Risø National Laboratory, Technical University of Denmark, Frederiksborgvej 399, DK-4000 Roskilde, 2007.
`http://www.risoe.dk/rispubl/VEA/ris-r-1614.pdf`

Christiansen et al. (2007): Christiansen L.E., Nielsen H.Aa., Nielsen T.S., and Madsen H. Automatic selection of tuning parameters in wind power prediction. Technical Report IMM-2007-12, Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2007.
`http://www.imm.dtu.dk/pubdb/views/publication_details.php?id=5316`

Holttinen and Ikäheimo (2007): Holttinen H. and Ikäheimo J. Wind prediction and bids in Denmark. VTT Energia, 2007.
`http://www.imm.dtu.dk/~han/pub/FU4101_DKvind_bids_and_balance_VTT_2006.pdf`

Nielsen et al. (2005): Nielsen H.Aa., Nielsen T.S., Madsen H., and Giebel G. Candidate prediction models and methods. Informatics and Mathematical Modelling, Technical University of Denmark, 2005.
`http://www.imm.dtu.dk/~han/pub/FU4101_Candidate_pred_models.pdf`

Nielsen et al. (2007a): Nielsen H.Aa., Nielsen T.S., Madsen H., San Isidro M.J., and Marti I. Optimal combination of wind power forecasts. *Wind Energy*, 2007. Accepted.
`http://www.imm.dtu.dk/~han/pub/FU4101_WEpaperCFCaccepted.pdf`

Nielsen et al. (2007b): Nielsen H.Aa., Pinson P., Christiansen L.E., Nielsen T.S., Madsen H., Badger J., Giebel G., and Ravn H.F. Improvement and automation of tools for short term wind power forecasting. In *Scientific Proceedings of the European Wind Energy Conference & Exhibition*. Milan, Italy, 2007.
`http://www.ewec2007proceedings.info/allfiles2/475_Ewec2007fullpaper.pdf`

Nielsen et al. (2007c): Nielsen T.S., Nielsen H.Aa., and Giebel G. Initialisation of power curve models using à priori information. Technical Report IMM-2007-13, Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2007.
`http://www.imm.dtu.dk/pubdb/views/publication_details.php?id=5329`

Nielsen et al. (2006b): Nielsen T.S., Nielsen H.Aa., Madsen H., San Isidro M.J., and Marti I. Optimal wind power forecasting by model combination. Poster presented at the European Wind Energy Conference & Exhibition 2006, Athens, Greece, 2006.
http://www.imm.dtu.dk/~han/pub/FU4101_EWEC06poster.pdf

Nielsen et al. (2006c): Nielsen T.S., Nielsen H.Aa., Pinson P., Madsen H., Giebel G., and Badger J. Improvement and automation of tools for short term wind power forecasting. Poster presented at Global Windpower 2006, Adelaide, South Australia, 2006.
http://www.imm.dtu.dk/~han/pub/FU4101_GW06poster.pdf

Nolsøe (2006): Nolsøe K. Estimating the power curve. Informatics and Mathematical Modelling, 2006.
http://www.imm.dtu.dk/~han/pub/FU4101_EstimatingThePowerCurve.pdf

Pinson et al. (2007c): Pinson P., Nielsen H.Aa., and Madsen H. Robust estimation of time-varying coefficient functions – application to the modeling of wind power production. Technical Report IMM–2007-09, Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2007.
http://www.imm.dtu.dk/pubdb/views/publication_details.php?id=5275

Ravn (2006): Ravn H.F. Short term wind power prognosis with different success criteria. In *Proceedings of the 9th International Conference on Probabilistic Methods Applied to Power Systems*. 2006.
http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?isnumber=4202205&arnumber=4202303&count=196&index=97

Furthermore, the following master thesis project was closely linked to the project:

Linnet (2005): Linnet U. *Tools supporting wind energy trade in deregulated markets*. Master's thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2005.
http://www.imm.dtu.dk/pubdb/views/publication_details.php?id=3969

Summary

This report summarizes the results of the Danish PSO-project *Intelligent wind power prediction systems*. During the project further improvement and automation of tools for short term wind power forecasting, typically with horizons up to 48 hours, have been studied. Also, the use of such forecasts in bidding on the spot marked have been considered.

It has been demonstrated that it is indeed possible to achieve further automation and robustness by modification of the methods underlying tools such as the Wind Power Prediction Tool. Initialisation of a statistically based power curve model using a physical model results in significant performance improvements for high wind speeds during the first months of operation. Presumably, performance for events which are even more rare, as e.g. cut-outs, can be improved generally or at least until sufficient data from the particular type of event are available.

For a complex site it has been shown that stability indices can be used to improve the accuracy of wind speed predictions based on a simple statistical model. The stability indices are derived from the same operational forecasting model that provides the usual inputs for wind speed predictions. This improvement indicates the advantages of using more of the data available within the operation forecast models. The physical basis for the improvement given by the stability indices is revealed by mesoscale modelling. Mesoscale modelling is then used to provide corrections for wind speed predictions based on wind speed, direction and stability. These corrections improved the prediction accuracy. The mesoscale model results are verified using two different models.

The adaptive non-parametric methods underlying modern forecasting tools require a small number of tuning parameters to be selected. Methods for continuous adjustment of these parameters have been developed. The criteria used for adapting these are directly related to the prediction performance. Also a robust version of the adaptive non-parametric estimation method has been developed. For data containing outliers, this method results in markedly more accurate estimation of the underlying power curve. Operational robustness is easily achieved using meteorological forecasts from two or three different suppliers. When more than one forecast is available, combined forecasting can be used. The improvement in performance which can be achieved depends on the variance and correlation of the meteorological forecast errors. For realistic values of these, significant improvements are observed.

Finally, the type of criterion used during estimation or when forming a bid on the spot marked may have economic consequences for wind power producers. Also, it has been shown to have consequences at the level of the electricity system as whole in that it affects the requirement for regulating power.

1 Introduction

With the increasing focus on renewable energy the share of wind power production is rapidly increasing worldwide. As a consequence of this an increasing focus on state-of-the-art forecasting tools is observed. These systems must be operating for wind farms or regions which can be quite different in terms of location, layout, quality of meteorological forecasts, etc. This report summarize the results obtained as part of the project *Intelligent wind power prediction systems*. During the project, further improvement and automation of tools for short term wind power forecasting, typically with horizons up to 48 hours, have been studied. Also, the use of such forecasts in bidding on the spot marked has been considered.

Based on the requirements outlined above the methods underlying the tools must be self-calibrating, also called adaptive (Nielsen et al., 2002a,b, 2006a). In this project the Wind Power Prediction Tool (WPPT) (Madsen et al., 2005a; Nielsen et al., 2005) is used as an example of such a system and thus forms the technological basis of the project. In Section 2 WPPT is briefly described. Since WPPT is based on statistical modelling, the system must learn the relation between the meteorological forecast and the observed power production before reasonable forecasts can be produced. In Section 3 it is investigated how this can be (partly) avoided by the use of initial models generated using physical modelling and software such as WAsP[1]. The section also describes the benefits from such an approach. Section 4 considers improvement of the underlying physical models. The section sets out by considering various stability measures, and the benefit of including these in a linear regression model relating the forecasted wind speed to the observed wind speed is investigated. The stability measures are based on the data from the operational weather forecast model providing the usual basic wind speed inputs for wind farm prediction systems. This part of the study aims to find out what benefits may come about by using more of the stability related information within operational weather forecasting models. Thereafter, mesoscale modelling of a wind farm placed in complex terrain is described. The section also includes a suggestion for using these extended models together with WPPT. For further automation of adaptive forecasting tools Section 5 describes approaches to continuous adjustment of tuning parameters (forgetting factor and bandwidth). Section 6 considers estimation criteria. First, a general introduction to economically based estimation criteria and their relation to quantile regression and robust estimation is given. Thereafter a robust version of the estimation method currently used in WPPT Nielsen et al. (2000a) is presented. Finally, in Section 6.3, the system-wide consequences of bidding on the spot marked using economical estimation criteria are considered. Operational robustness can be improved by having access to several meteorological forecasts. However, most of the time, all meteorological forecasts will be available. Section 7 considers the improvement in performance which can be achieved by using several meteorological forecasts. Finally, in Section 8 the main conclusions of the project are outlined.

[1] www.wasp.dk

2 The Wind Power Prediction Tool

The technological basis of much of the work presented here is the Wind Power Prediction Tool (WPPT) (Madsen et al., 2005a; Nielsen et al., 2002a, 2000b). Therefore, for the sake of clarity, the system is briefly described below. A more complete description can be found in (Nielsen et al., 2005). Furthermore, configuration examples are shown in (Madsen et al., 2005a).

WPPT has run at all major Danish utilities for more than a decade. Since version 4 the direction dependent power curve modelling of WPPT has been based on conditional parametric models (Cleveland, 1994). In such models the the response y_t at time t is modeled using two groups of explanatory variables. One group of variables $\mathbf{x}_t$ enters globally through coefficients depending on the other group of variables $\mathbf{u}_t$, i.e.

$$y_t = \mathbf{x}_t^T \boldsymbol{\theta}(\mathbf{u}_t) + e_t, \tag{1}$$

where $\boldsymbol{\theta}(\cdot)$ is a vector of coefficient-functions to be estimated and e_t is the noise term.

Depending on how the software is set up, the power curve directly models the relation between the meteorological forecast and the power production in a wind farm or in a region. The rationale behind this is explained in e.g. (Nielsen et al., 2002b).

To account for auto correlation and diurnal variation, WPPT uses the output from the conditional parametric model as input to an ARX-model (Ljung, 1987) which adjusts for these effects. These models can be written in the form

$$y_t = \mathbf{x}_t^T \boldsymbol{\theta} + e_t \ , \tag{2}$$

where t is the time index, y_t is the output, $\mathbf{x}_t$ is a vector containing inputs and possibly lagged values of the output, $\boldsymbol{\theta}$ is a vector containing the coefficients of the model, and e_t is the error term.

All models in WPPT are automatically re-calibrated as new information becomes available. For ARX-models adaptive recursive least squares (Ljung, 1987) are used, whereas for conditional parametric models the method described in (Nielsen et al., 2000a) is used. Further information about WPPT can be obtained via `www.risoe.dk/zephyr/`.

3 Initial estimates

The adaptive estimation methods need to learn the underlying dependencies from data, i.e. sets of meteorological forecasts and actual power productions. At startup, initial estimates must be supplied and over time the influence of the initial estimates will vanish. Given historic data, the software can be run in off-line mode and hence the estimates will be appropriate when going into on-line mode. However, if historic information is not available, the initial estimates will influence the forecasts generated until the effect of the initial estimates has vanished. Depending on the geographical location, this may take several months for combinations of wind speeds and directions which only occur rarely.

Here we focus on the power curve model of WPPT and investigate ways in which initial estimates can be accommodated by the adaptive estimation method. The initial estimates are calculated using the software WAsP (`www.wasp.dk`) Park from Risø National Lab also Appendix A in Nielsen et al. (2005).

3.1 Initialization

The power curve model in WPPT is estimated recursively and adaptively using a procedure as described in (Nielsen et al., 2000c) based on locally linear regression. The coefficients, θ, and correlation matrix, $\mathbf{R}$, of the local linear models (see also section 5) must be initialized prior to the estimation. This section compares the traditional initialization scheme for the power curve model of WPPT with a new simulation driven initialization scheme.

In recursive linear estimation, the coefficients and correlation matrix are traditionally initialized with $\theta = 0$ and $\mathbf{R} = \alpha\mathbf{I}$, where $\mathbf{I}$ is the identity matrix and $\alpha << 1$. This is the initialization scheme used in the current operational version of WPPT.

In the simulation driven scheme a simulated data series covering a period prior to commencement of the observations is generated using a physical derived power curve model for the wind farm. A set of simulated wind speed and wind direction series are generated and used as input to the physical model to calculate a simulated power series from the wind farm.

A simulated data set has been calculated for the Klim wind farm in Northern Jutland, Denmark, covering a period of one year prior to the observed data set used in the evaluation (period: 29/11/2002-20/07/2004). Here the wind speed and wind direction series have been generated as two sets of independent uniformly distributed variables. The deterministic wind farm model used is the WAsP model developed by Risø and the deterministic power curve used in the simulations is found in Figure 1.

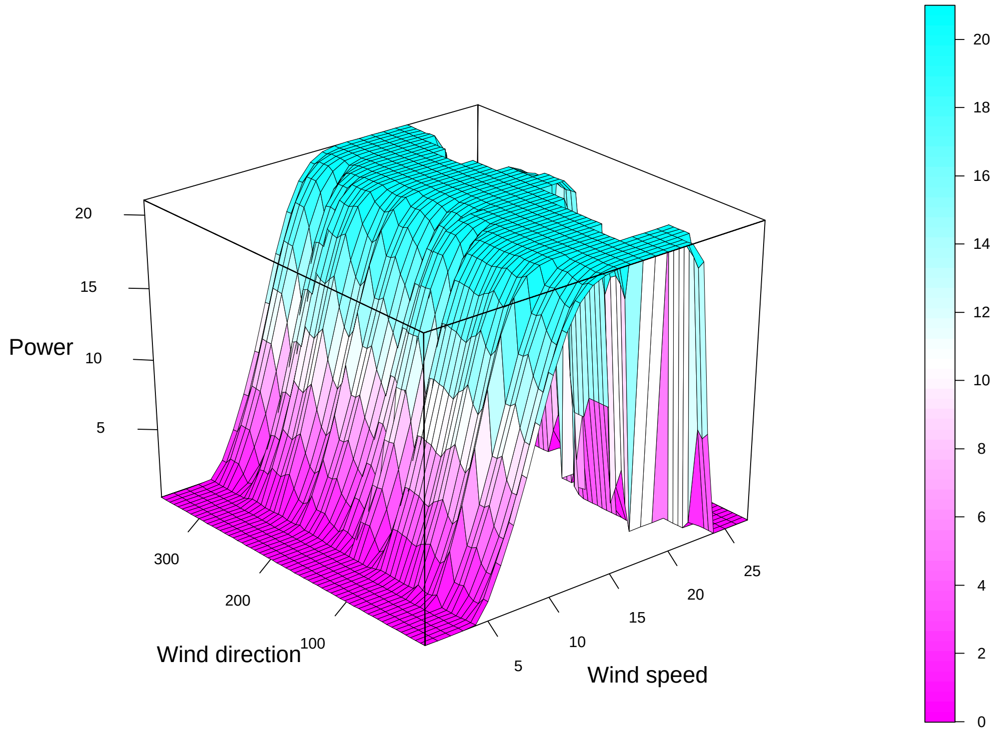

Figure 1: Deterministic power curve calculated by WAsP for the Klim wind farm.

Different lengths of initial period have been investigated ranging from 0 to 12 months. The benefits of the initialization have been investigated using the first 2 months, the first 6 months and the entire 20 months of the data set as evaluation period and the results are found in Table 1 using NRMSE (root mean square error normalised with rated power) as measure.

The first 4 lines of Table 1 state the results when the the power curve is initialized for all wind speeds and all data in the evaluation period is used. Using initialization results in a persistent improvement compared to no initialization, however the benefits are small. From the table it is also seen that the simulation period used in initialization should be short. This indicates that the estimated power curve must be allowed to adapt quickly to the observed data from the initial estimate based on simulated data.

The density of the observations decreases rapidly as the wind speeds increases, hence the period before sufficient data is available to give a firm estimate of the power curve is longer for high wind speeds compared to low wind speeds. It has therefore been investigated if it is beneficial only to initialize the upper part of the power curve corresponding to wind speeds above $10m/s$. As seen from lines 5 to 7 in Table 1, this turns out to be the case. Compared to no initialization a reduction in NRMSE of 14% is found for the first 2 months of the evaluation period.

As indicated above initialization is expected to be of more importance for the high wind speed part of the power curve compared to the low wind speed part. This is confirmed by the lower part of Table 1, which shows the effect of the initialization if only observations with wind speeds above $10m/s$ are considered in the evaluation. It is seen that the advantage of initialization is much more pronounced for high wind speeds and NRMSE is reduced by 33% and 31% for the first 2 and 6 months of the evaluation period, respectively.

Ini. length	Ini. ws	Eval. ws	Eval. period 2m.	6m.	20m.
0m.	–	All	18.3	14.7	14.0
2m.	All	All	16.3	14.0	13.8
4m.	All	All	16.7	14.2	13.9
8m.	All	All	17.1	14.4	14.0
2m.	$> 10m/s$	All	15.7	13.8	13.7
4m.	$> 10m/s$	All	15.7	13.9	13.8
8m.	$> 10m/s$	All	15.9	14.0	13.8
0m.	–	$> 10m/s$	26.6	24.6	21.9
2m.	All	$> 10m/s$	18.4	17.5	17.6
4m.	All	$> 10m/s$	18.9	17.9	17.8
8m.	All	$> 10m/s$	19.3	18.2	17.9
2m.	$> 10m/s$	$> 10m/s$	17.7	17.0	17.5
4m.	$> 10m/s$	$> 10m/s$	17.9	17.2	17.6
8m.	$> 10m/s$	$> 10m/s$	18.4	17.6	17.8

Table 1: NRMSE for different initialization lengths and evaluation periods for the Klim wind farm. Both the effect of only initializing as well as evaluting the upper part of the power curve above $10m/s$ is shown.

In conclusion, the results clearly show that benefits can be achieved by initializing the statistical power curve model based on a priori information, especially in the high wind speed regions. However care must be taken to give appropriately low emphasis to the a priori information in regions which are well represented by the observed data.

4 Improvement of the initial model

With the aim of improving the initial physical model, it is investigated if stability measures and mesoscale modelling can achieve this goal. Such an approach has the potential of identifying model structures by which the general performance of the prediction system can be improved, i.e. not limited to the initial phase. As an onset, the mesoscale model KAMM was used, and subsequently the mesoscale model WRF was used in order to verify the results.

4.1 Stability measures as input to statistical models

In this section ways to incorporate models with some physical basis into the prediction system are explored. A range of methods for using Global Forecasting System (GFS) forecast data to make wind speed predictions for a specific site are considered. The simplest way is to use the NWP wind speed directly, for example using U_{10} for the nearest NWP model grid point, and adjustments to different heights can be made using the logarithmic profile. The error for these kinds of predictions very often contains a significant systematic error, for example a bias, that can be reduced by *Model Output Statistics* (MOS). The MOS used in this first part of this study was

$$U_{pred} = A + BU_{NWP} \ , \tag{3}$$

where A and B are found by fitting a linear regression using a training period of the concurrent NWP and observed winds time series. A more novel way to produce a wind speed prediction is to introduce a stability parameter in the linear regression, as in the following,

$$U_{pred\_stab} = A + BU_{NWP} + CS_{NWP} \ , \tag{4}$$

where S_{NWP} represents a stability parameter based on data from the NWP system. A, B, and C can be a function of wind direction, using a training period of the concurrent NWP winds and stability parameter, and the observed winds time series. Although the stability parameter is very unlikely to have a purely linear influence, at this stage a linear relation is used in order to identify potentially useful stability measures. More sophisticated ways of making stability based adjustments could be employed in the near future.

The objective is to assess whether using a stability measure can reduce the mean absolute error (MAE), the mean of the absolute prediction errors. Various stability parameters, used to supply S_{NWP}, were employed. They are typically based on vertical gradients of horizontal wind speed or potential temperature. Among the parameters were the Brunt-Väisälä frequency, N^2, (evaluated as $(g\Delta\theta)/(\theta_0\Delta z)$), and the Froude number, Fr, (evaluated as $u\sqrt{\theta/(g\Delta\theta\Delta z)}$); and correspondingly, the value S_{NWP} is set to N^2 or Fr in (4). θ is the potential temperature, θ_0 is the mean potential temperature in the Δz layer, and g is the gravitational acceleration. The GFS output includes meteorological quantities at several heights, thus it is possible to use different vertical extents, Δz, in order to calculate stability parameters. Where the stability parameters have been calculated using quantities below 1000 m a.g.l., and Δz smaller than a few hundred metres, the stability measure is referred to as a 'shallow profile' measure. In contrast, for larger Δz, the term 'deep profile' measure is used.

Wind speed predictions have been made for the anemometer located at 76 m a.g.l. (corresponding approximately to a typical hub height for a 2 MW turbine) on the circa 120-m Risø mast. Half the data from the full period (01/09/2005-28/02/2006)was selected randomly for fitting the linear regression (training), the other half was used for evaluation. The error evaluations use the 0 hour leadtimes only and comparison to the 10-minute average measurement at the appropriate validation times.

When using the linear regression without the stability parameter employing the GFS product giving the mean wind speed for the layer between 0 mb and 30 mb above the surface (U_{30_0}) as U_{PWD} gave the most accurate wind speed predictions. Using a stability measure at best only gave a very marginal improvement in the accuracy of the prediction. These small improvements were achieved using 'shallow profile' stability parameters based on the absolute potential temperature gradient and the Brunt-Väisälä frequency (N^2).

Predictions based on using 10-m winds (U_{10}) from GFS as U_{PWD} gave a mean absolute error of 1.37 m/s, when no stability parameter was used. The accuracy was improved by the inclusion of the stability parameter in the linear regression, giving a reduction in mean absolute error of up to 5 %. However in absolute terms the best predictions based on U_{10} and a stability parameter were never better than the best prediction based on U_{30_0} and a stability parameter. For the Risø winds, the use of 'deep profile' stability measures did not bring about marked improvements in overall performance.

Wind speed predictions have been made for the 55 m a.g.l. anemometer at the Alaiz wind farm, near Pamplona in Northern Spain. Again half the data from the full period (01/07/2004-31/05/2005) was selected randomly for fitting the linear regression (training), and the other half was used for evaluation. The error evaluations use the leadtimes from 0 to 48 hours, with a 6-hour interval, and comparison to the 10-minute average measurement at the appropriate validation times.

For predictions made without using the stability parameter, U_{30_0} and U_{10} gave a mean absolute error of 2.98 m/s and 3.19 m/s respectively. When including a stability parameter a large reduction in mean absolute error was achieved (between 15% and 20%) when using 'deep profile' stability parameters based on the absolute potential temperature gradient and the Brunt-Väisälä frequency (N^2). For the Aliaz winds, the use of 'shallow profile' stability measures brought about improvements to the mean absolute error in the range 1 - 8 %. The 'deep profile' stability measures had much more impact.

Wind speed predictions have been made for two 46 m a.g.l. anemometers at the Klim wind farm in Northern Jutland, Denmark, (period: 01/12/2005-31/03/2006). Including a stability parameter yielded no significant reduction in mean absolute error irrespective of the leadtime was used. This may be related to the NWP stability measures being inaccurate or alternatively because the GFS model predicts wind speed very well (mean absolute error of 1.2 m/s).

Risø		Alaiz	
no Stab	Stab	no Stab	Stab
1.26	1.25 0.73 %	2.98	2.48 16.6 %

Table 2: Summary of prediction performance for Risø and Alaiz given by mean absolute error [m/s] for predictions made without and with stability parameter. The percentage improvement is also given.

Table 2 summarizes the main results. The 'deep profile' stability measure that improved the prediction for Alaiz was based on a vertical extent, $\Delta z \simeq 3000$ m. This suggests that stability and flow interaction with terrain, with a comparable vertical extent, is important for the Alaiz case. Mesoscale modelling will be described in the next section that explores the role of terrain and stability in the region of the Alaiz wind farm.

4.2 Mesoscale modelling

The importance of stability in determining the wind conditions at the Alaiz wind farm site is supported by mesoscale simulations using KAMM Adrian and Fiedler (1991) using a wide range of different atmospheric conditions, i.e. wind speed, wind direction and potential temperature profiles.

Idealized KAMM studies have been performed using sets of wind profiles representing different wind directions and different atmospheric stabilities. Four different geostrophic wind speed profiles (5, 10, 15, 20 m/s) and three different stabilities, neutral, typical (near neutral) and stable, were investigated.

The geostrophic wind and temperature profiles were defined at 0, 1500, 3000, 5500 m above sea level. The geostrophic wind speeds forcing the model were constant with height. For the stable and typical stability cases the temperature profiles were evaluated using NCEP/NCAR reanalysis data.

Figure 2 summarizes the results from five sets of the idealized KAMM simulations. The thick rectangles show the forcing directions, e.g. the sector centred on 30° is red. For a given large-scale forcing the simulated winds at 50 m for five locations neighbouring the wind farm are show by lines of the same colour. The direction on the diagram indicates the direction where the wind comes from and the length indicates the wind speed speed-up. The dotted-line circle represents a speed-up of 1, meaning that the wind speed at

50m is the same as the geostrophic forcing. A speed-up above 1 indicates a wind faster than the forcing wind speed, and when it is below 1 this indicates a wind slower than the wind forcing.

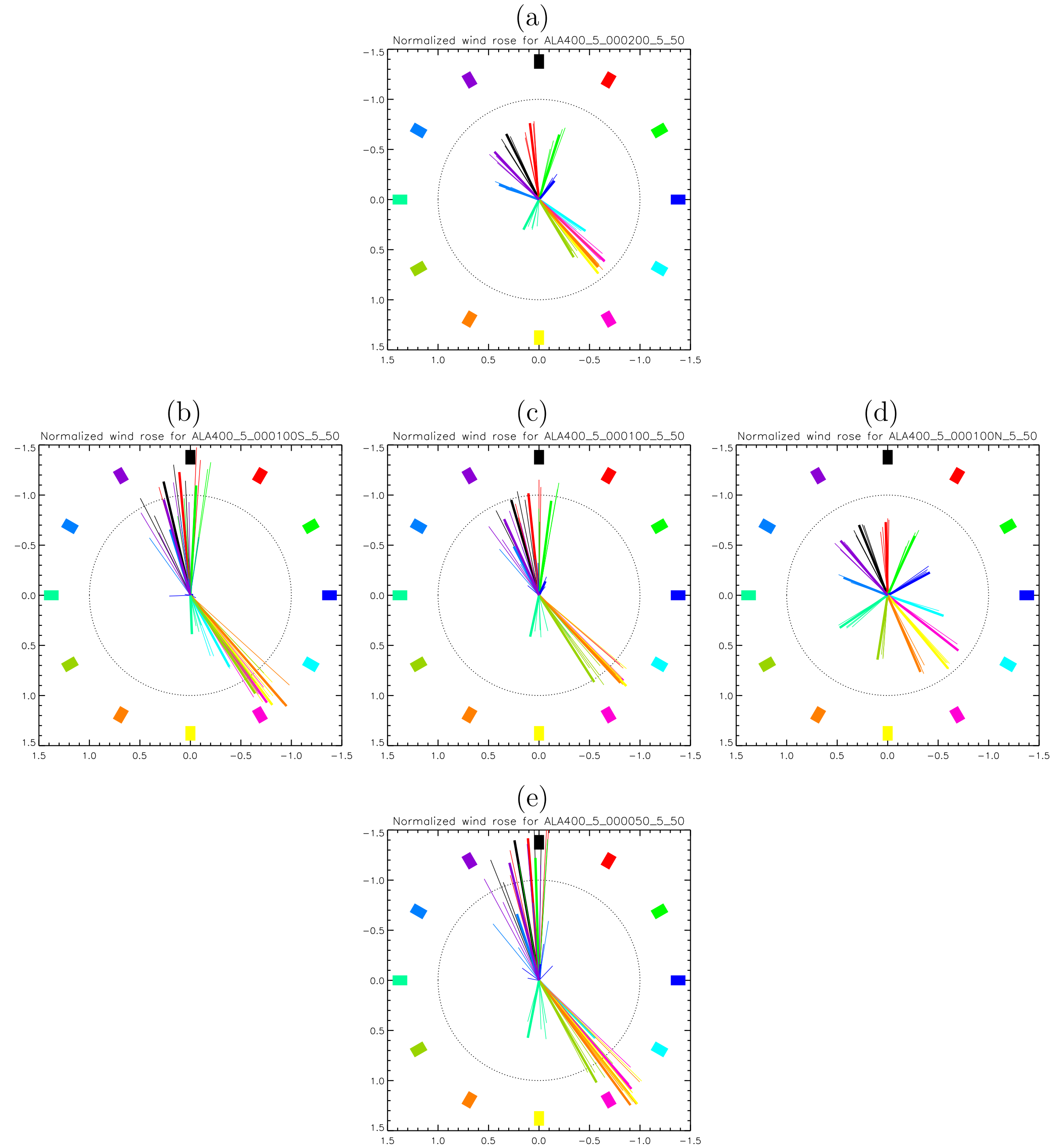

Figure 2: Diagrams showing the mesoscale effects on the geostrophic wind forcings for five sets of KAMM simulation using different wind speed and thermal stratifications, (a) 20 ms^{-1} and typical stability, (b) 10 ms^{-1} and stable conditions, (c) 10 ms^{-1} and typical stability, (d) 10 ms^{-1} and neutral conditions, (e) 5 ms^{-1} and typical stability. Each set is made up of twelve simulations using different wind directions indicated by the colours. Please refer to the main text for an explanation of the figures.

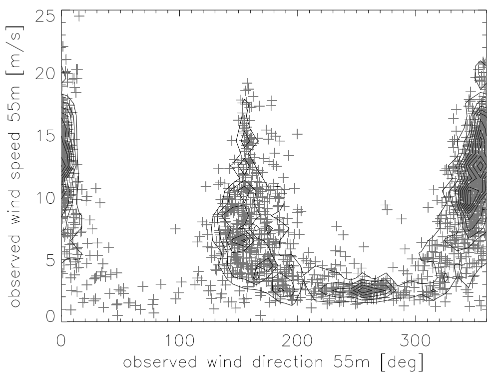

Figure 3: The observed wind speed and direction at 55 m a.g.l. from data for 2001.

Looking at Fig. 2(c) first, the set using 10 ms^{-1} geostrophic winds and typical stability, it is possible to see that the wind directions at the wind farm site tend to be concentrated into the northern and south-eastern sectors, regardless of the forcing wind direction. Examining the next Fig. 2(e), the set using 5 ms^{-1} geostrophic winds and typical stability, it is possible to see that the direction concentration effect is slightly enhanced and that there is a greater degree of speed-up of the winds. Similar behaviour is seen when examining Fig. 2(b), the set using 10 ms^{-1} winds and stable stratification. Figure 2(a) shows the set of simulations using 20 ms^{-1} winds and typical stratification. Here it is possible to see a reduction in the direction concentrating effect of the winds at the wind farm, and a reduced speed-up effect. Figure 2(d) shows the set using 10 ms^{-1} winds and neutral stratification. More so than in the case of Fig. 2(a), it shows a further reduction in the direction concentrating effect and speed-up.

Figure 3 shows the observed wind speed and direction at the 55 m a.g.l. mast at the Alaiz wind farm. It can been seen from the observations that there is a strong concentration of winds, and especially the more powerful winds, from the directions centred on 350°and 150°. This corresponds strongly with the findings of the idealized mesoscale simulations.

Climatically representative profiles of the atmospheric flow based on the GFS data over the period 01/07/2004-31/05/2005 have been used to drive the mesoscale model. These profiles were created using a similar method to that given in Frank and Landberg (1997) and the method draws on the method of statistical-dynamical downscaling as used in Frey-Buness et al. (1995). Figure 4(a) shows the GFS forecasted geostrophic wind speed and direction for the aforementioned period. Each point in the figure represents a forecast based on the nearest GFS grid points to the wind farm. The figure shows the distribution of the forcing winds in the region. Figure 4(b) shows the 103 representative wind classes that have been defined based on data shown in Fig. 4(a). Each wind class in fact describes the way the geostrophic wind speed and direction, and potential temperature changes with height.

The KAMM modelling results are shown in Fig. 4(c) for each wind class. Comparing the wind speed and direction distribution in Fig. 4(c) and Fig. 3, it is possible to see that the

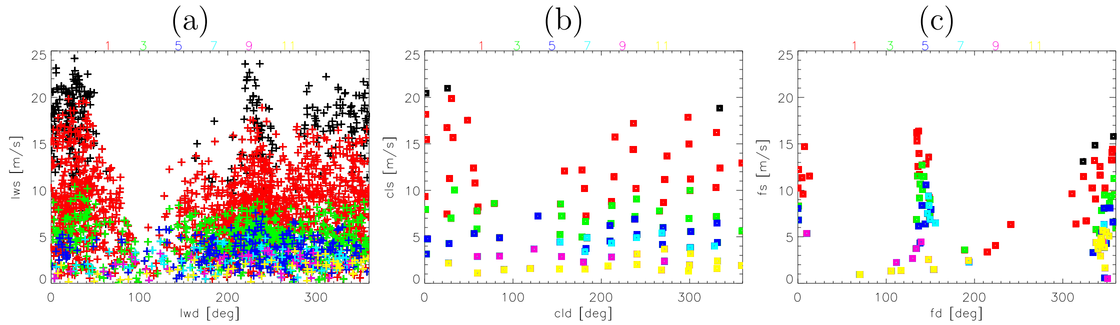

Figure 4: (a) The geostrophic wind and direction derived from GFS forecasts for the period 01/07/2004 - 31/05/2005 derived from grid points close to the Alaiz wind farm. (b) The geostrophic wind and direction for the 103 wind classes that have been calculated to represent the data in (a). (c) The KAMM modelled wind speed and direction at 50 m a.g.l. at the wind farm site resulting from the simulations using the wind classes shown in (b). The colours of the plotted points indicate the values of $1/Fr^2$, the colour key is provided by the colour of the values above the plots.

modelling results reproduce to a fair degree the wind speed and direction distributions observed at the wind farm site.

The modelling results can be applied in a prediction system as a downscaling look-up table. By finding the entry in the downscaling look-up table, via reference to the most appropriate wind class for the driving flow situation provided by the GFS forecast, the wind conditions at the wind farm site are given.

Without application of MOS, the mean absolute error using the KAMM mesoscale look-up table was 3.46 m/s. With MOS the mean absolute error is brought down to 2.60 m/s. This compares favourably with the predictions made using GFS and MOS alone, giving a mean absolute error of 2.98 ms^{-1} for Alaiz.

4.3 Validation of the mesoscale results using WRF

The mesoscale simulations described and used so far here have been performed using the model KAMM. Validation has been carried out by comparing wind speed prediction with wind speed measurements at Risø, Aliaz and Klim. As a second step in the validation process model results from KAMM are compared with a second mesoscale model. The second mesoscale model is the WRF model, which stands for the Weather Research and Forecasting model. There are in fact two versions of the model: the ARW (Advanced Research WRF) and the NMM (Non-hydrostatic Mesoscale Model). The ARW version is used here because it features an option to use an initialization method rather like that used for KAMM. Skamarock et al. (2005) gives a detailed description of the WRF/ARW model.

To the extent possible for two different models, the same idealized simulations that were presented in Fig. 2 were carried out using WRF in place of KAMM. Qualitatively, the results from the idealized WRF simulations are very similar to the results from the idealized KAMM simulations. In Fig. 5 the results from both KAMM and WRF for the idealized simulations using the 10 ms^{-1}profile with typical stability is shown, a full comparison is given in Badger et al. (2007). A very similar wind direction concentration and wind speed enhancement effects are exhibited. Similarly the wind direction concentration and wind speed enhancement effects are strongest for the low wind speed and stability cases. The dominant angles for the wind direction concentrations are 330 - 0° and 130 - 150°. This is also in very good agreement with the KAMM results.

(a) (b)

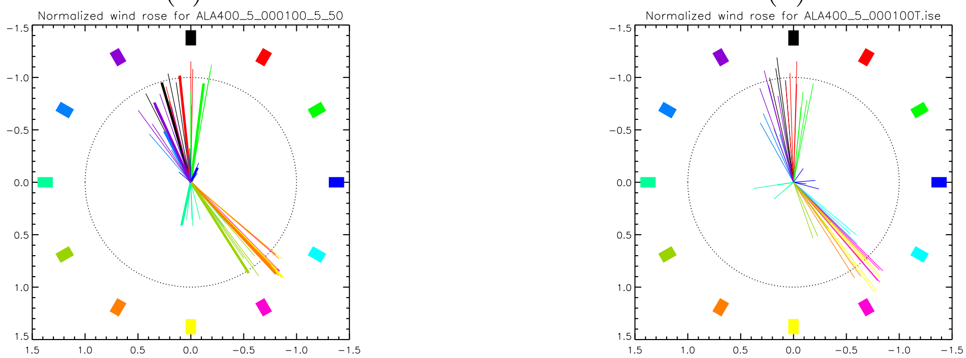

Figure 5: Diagrams showing the mesoscale effects on the geostrophic wind forcings from 12 directions using the 10 ms^{-1} and typical stability profile given by (a) KAMM and (b) WRF. Each set is made up of 12 simulations using different wind directions indicated by the colours. Please refer to the main text for an explanation of the figures.

The WRF mesoscale model was also used to create a downscaling look-up table using the wind classes shown in Fig. 4(b). As in the KAMM case, the prediction performance was evaluated using data from the period 01/07/2004-31/05/2005. Table 3 summarizes the performance of the KAMM and WRF look-up table prediction systems. First, looking at the mean error and mean absolute error for the two systems without using MOS, it can be seen that the WRF look-up table system does not perform as well as the KAMM look-up table system. However it does still perform better than when using the wind class geostropic winds alone which gives a mean absolute error of 4.19 ms^{-1}. After MOS the performance of the two models' look-up tables is rather similar. It should be repeated that this performance is better than using GFS surface wind data (U_{30-0}) for the wind farm location and MOS.

4.4 Application of stability correction in WPPT

The stability correction and the information contained within the mesoscale look-up tables can be incorporated in the WPPT system by adding a preprocessor step. Figure 6 shows how the preprocessor would fit into the system. The NWP provided data, including

Model	no MOS ME	no MOS MAE	with MOS ME	with MOS MAE
KAMM	-0.66	3.46	-0.11	2.60
WRF	-1.18	3.78	-0.15	2.65

Table 3: Mean error (ME) and mean absolute error (MAE) in m/s for predictions made using KAMM and WRF look-up tables without MOS and with MOS.

the appropriate stability measure, is fed into the preprocessor and the linear regression is applied. The measured wind speed and measured farm power are also fed into the preprocessor so that the linear regression coefficients can be updated recursively. The recursive features of WTTP bring many advantages to the system, such as accommodation of seasonal variations, dirtying of turbine blades, changes to the operation of the NWP system, etc. Thus a recursive adjustment of a stability correction is considered advantageous.

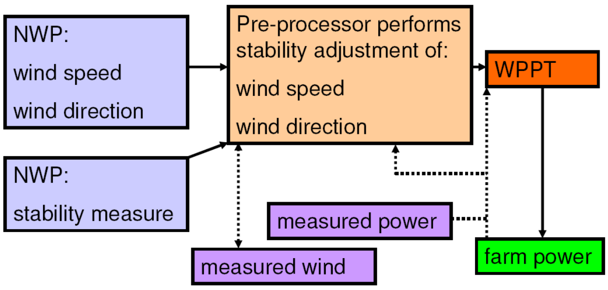

Figure 6: Schematic diagram showing the conceptual development of the WPPT preprocessor.

5 Selection of tuning parameters

The adaptiveness of the estimation methods are controlled by a constant $\lambda \in (0, 1)$ called the forgetting factor. Furthermore, the method used for estimation in conditional parametric models requires the selection of a so called bandwidth $h > 0$. As a consequence, in order to set up WPPT two forgetting factors and one bandwidth must be selected. In fact for the conditional parametric models it is even possible to use different bandwidths and forgetting factors for each fitting point, i.e. for each point at which the functions involved are estimated (Nielsen et al., 2000a).

Here we consider how the forgetting factors and bandwidths can be tuned over time. In this way an additional layer of adaptiveness is added which ultimately should allow for easy setup of the software because default values can be supplied and the software will then automatically adjust these to the specific site or region.

5.1 Forgetting factor

In this section we consider linear regression models or auto regressive models with exogenous variables (2).

Since the underlying system of which (2) is an approximation may change over time it is often beneficial to track the coefficients. Given a constant called the forgetting factor $\lambda \in (0, 1)$ exponential weighted adaptive recursive least squares (Ljung, 1987) accomplish this by the following recursions:

$$
\begin{aligned}
\boldsymbol{k}_t &= \frac{\boldsymbol{P}_{t-1}\boldsymbol{x}_t}{\lambda + \boldsymbol{x}_t^T \boldsymbol{P}_{t-1}\boldsymbol{x}_t} && (5)\\
\xi_t &= y_t - \boldsymbol{x}_t^T \boldsymbol{\theta}_{t-1} && (6)\\
\hat{\boldsymbol{\theta}}_t &= \hat{\boldsymbol{\theta}}_{t-1} + \boldsymbol{k}_t \xi_t && (7)\\
\boldsymbol{P}_t &= \lambda^{-1}(\boldsymbol{I} - \boldsymbol{k}_t \boldsymbol{x}_t^T)\boldsymbol{P}_{t-1} && (8)
\end{aligned}
$$

where $\mathbf{P}_t$ is the inverse correlation matrix, $\boldsymbol{k}_t$ is the gain, ξ_t is the prediction error, and $\hat{\boldsymbol{\theta}}_t$ is the vector of estimates at time t. The memory time constant $1/(1 - \lambda)$ (Ljung, 1987) is an indication of how many observations the estimates are based on.

Several methods allowing the forgetting factor to vary over time exist in the literature (Ljung and Gunnarsson, 1990). Here we focus on methods which directly aim at minimizing the expected value of the prediction error. In (Haykin, 1996, Sec. 16.10) a steepest decent algorithm for updating λ is presented.

Since the admissible values of λ are bounded, a clipper function is added to the algorithm on an ad hoc basis. Initial investigations revealed that the algorithm does not perform

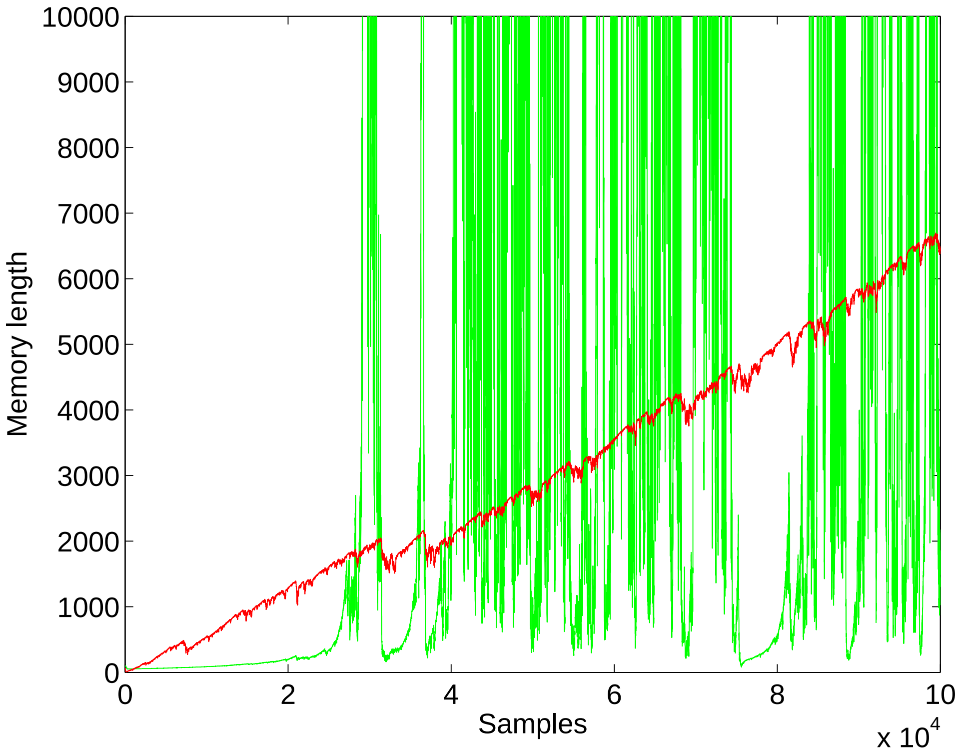

Figure 7: Comparing unbound (red) and bounded (green) versions of the steepest descent algorithm on a simulation with a constant model.

well for noisy systems which only change slowly over time, i.e. where the appropriate λ is close to 1. For the application to the dynamical model of WPPT (Madsen et al., 2005a) this would pose problems. For this reason an unbounded version of the algorithm was derived, i.e. λ is expressed as a function of g and the function is constructed so that $g \in \mathbb{R}$ results in λ being restricted to $(1 - 1/N_{min}, 1)$. This can be obtained using:

$$\lambda(g) = 1 - \frac{1}{N_{min} + \exp(g)} \quad , \; g \in \mathbb{R} \text{ and } N_{min} > 1 \tag{9}$$

Naturally, the resulting algorithm is very similar to the algorithm in (Haykin, 1996, p. 735), the difference being that the derivative of λ w.r.t. g is now included in the algorithm.

What may not be that obvious is how much the stability differs between the two versions. As an example consider the constant model $y_t = 1.5x_t + e_t$, where x_t is autoregressive: $x_t = 0.975x_{t-1} + 0.025s_t$ with s_t iid. uniformly distributed on $[1;2]$. Finally, e_t is Gaussian iid. noise with zero mean and standard deviation 0.7. Figure 7 depicts the trace of the memory length over time. It is seen that the clipper function causes very unstable behavior, whereas our revised method results in a steady increase of the memory length. In both cases the initial forgetting factor was set to 0.99.

For both versions of the algorithm, a constant controlling the step size must be set, i.e.

instead of λ a new tuning parameter, α, is introduced. However, we have demonstrated that for a wide range of α the method is not sensitive to the actual value. For further details see (Christiansen et al., 2007).

5.2 Bandwidth

The power curve model and the upscaling model of WPPT is described in (Madsen et al., 2005a). The models are conditional parametric models (1) for which the coefficient functions are tracked (estimated) over time as described in (Nielsen et al., 2000a). Basically, the estimation procedure works by estimating the coefficient functions for a number of fixed values of the arguments, i.e. the vector $\mathbf{u}$; the so called fitting points. For each of these fitting points the bandwidth, the scaling of the individual elements of $\mathbf{u}$, the forgetting factor, and the degree of the local approximating polynomial, c.f. (Nielsen et al., 2000a), must be selected. Here we consider the bandwidth, the remaining quantities will normally be constant across fitting points.

In the project it has been investigated how the bandwidth at each fitting point can be updated over time. An approach similar to the approach in Section 5.1 is applied. More precisely the expected value of $w(\mathbf{u}_t)\xi_t^2$ is minimized, where $w(\mathbf{u}_t)$ is the weight on observation t given the fitting point, and ξ_t is the error when predicting the output based on the fitting point considered. Given this criteria function, it is important to define the weight function so that its integral is independent of the bandwidth.

The bandwidth h_t at time t is expressed as

$$h_t = h_0 + \exp(g_t) \tag{10}$$

which for $g_t \in \mathbb{R}$ restricts the bandwidth to be larger than h_0. The minimum bandwidth should be related to the grid size of the fitting points. In (Christiansen et al., 2007) an update formula for g is derived. The formula can be used for any weight function as long as the partial derivative of the weight function w.r.t. g is available.

To demonstrate the ability of the method to adjust the bandwidth, consider the following model

$$y_t = x_t\theta(u_t) + e_t \ , \tag{11}$$

i.e. a simple model resembling the up-scaling model of WPPT. Here we define the coefficient function as

$$\theta(u) = \begin{cases} 1 & , \ 0 \le u \le 1 \\ u & , \ 1 < u \le 2 \end{cases} \tag{12}$$

and estimate the function using a local linear approximation of the function. Furthermore, x_t *iid.* $U[1;2]$, u_t *iid.* $U[0;2]$, and e_t *iid.* $N(0, 0.25^2)$. The result is shown in Figure 8. It is seen that even though the initial bandwidth is not appropriate for the central points (purple, light blue, light green), the algorithm changes the bandwidths within 1000 samples, whereas the remaining bandwidths are only modified to a minor extent.

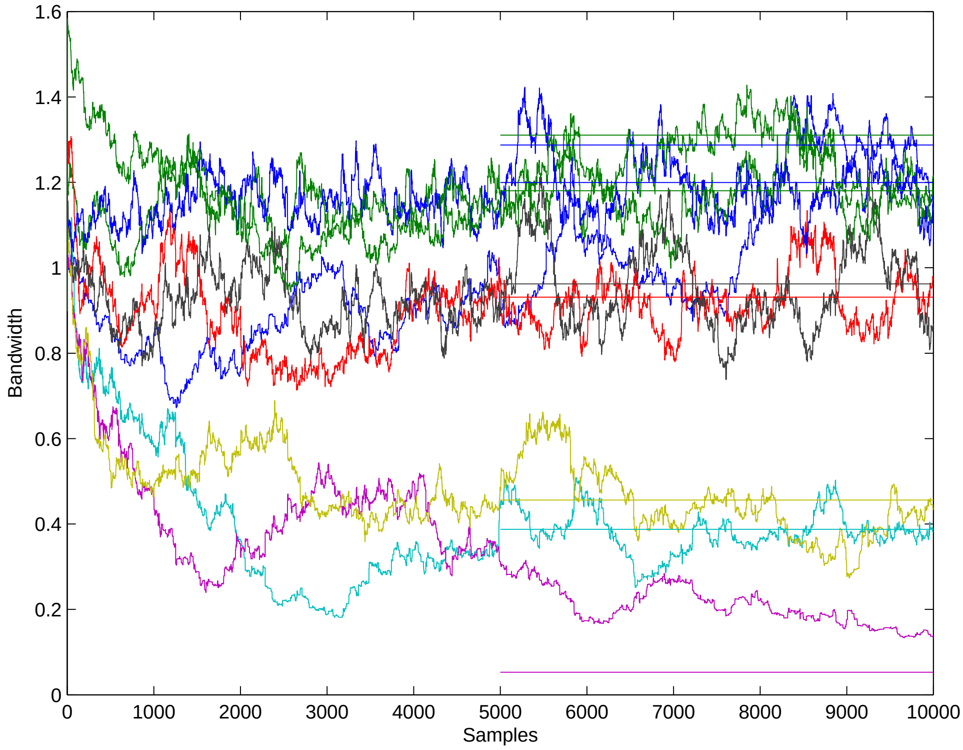

Figure 8: Using a tri-cube weight function to optimize the bandwidth at nine fitting points spread equidistantly from 0 to 2. Both steepest descent traces and fix bandwidth optimized on the last half of the data are shown. The 4 boundary points are blue and green, the central point is purple, the neighbour points of the central point are light blue and light green, and the remaining two points are black and red.

6 Estimation criteria

6.1 General aspects of estimation criteria

The commonly used estimation procedures in wind power forecasting minimize the squared sum of the prediction errors of the wind power forecast. In non-parametric adaptive estimation Nielsen et al. (2000a), these squared errors are furthermore weighted. Nielsen and Ravn (2003) use a criterion based directly on regulation prices. It turns out that this criterion is closely related to quantile regression (Koenker, 2005). To realize this, reference is made to (Bremnes, 2004; Morthorst, 2003). Let e_p be the actual electricity produced at the future time point under consideration, let e_b be the bid at the spot marked, which must be placed in advance, and let p_s be the spot price. Furthermore, let c_- be the unit cost of production less than the bid and let c_+ be the unit cost of excess production. With this setup, the income corresponding to the future time point will be

$$\begin{cases} e_p p_s - (e_b - e_p)c_- & \text{if } e_p \leq e_b \\ e_p p_s - (e_p - e_b)c_+ & \text{if } e_p > e_b \end{cases} \tag{13}$$

Following the arguments by Bremnes (2004) the income structure above leads to the conclusion that the bid should be the $c_+/(c_+ + c_-)$ quantile in the conditional distribution of the future power production. Using quantile regression this corresponds to minimizing the criterion

$$\frac{1}{c_+ + c_-}\begin{cases} \ c_+ e & , \quad e \geq 0 \\ -\ c_- e & , \quad e < 0 \end{cases} \tag{14}$$

Which is equivalent to minimizing the criterion in (Nielsen and Ravn, 2003). These criteria are related to the criterion used in robust estimation which was considered as part of this project, cf. Section 6.2 below.

However, it also leads to the idea that instead of focusing on a single point prediction focus should be on forecasting the conditional distribution of the future power production, so-called probabilistic forecasts. Providing an overview of this subject is beyond the scope of this report. It has been considered in the PSO-project "Vindkraftforudsigelse med ensemble forecasting"(Wind power prediction with ensemble forecasting), see (Giebel et al., 2005; Nielsen et al., 2006a) and the references therein. For a more complete overview see (Pinson et al., 2007b).

One conclusion from the project just mentioned is that the estimation procedures normally used lead to bias of the estimated power curve. This bias is problematic when using meteorological ensemble forecasts to produce probabilistic forecasts of the wind power production. As part of this project it has been investigated whether non-parametric inverse regression techniques can help solve the bias issue (Nolsøe, 2006). However, it proves to be problematic to preserve the monotonicity of the power curve in this case, making non-parametric inverse regression problematic. Instead it seems to be more beneficial to use orthogonal regression as described in (Pinson et al., 2007a), which describes work which has been carried out as part of the PSO-project "High Resolution Ensembles for Horns Rev".

As part of this project is has also been investigated whether a bidding strategy derived from the discussion above will result in more critical situations. The result is described in (Holttinen and Ikäheimo, 2007). See also Linnet (2005), which describes the results of a master thesis project closely linked to this project.

6.2 Robust non-parametric regression

On-line systems are potentially vulnerable to erroneous data. WPPT applies a wealth of initial checks and validation steps of the data it receives and models are only updated if newly received data are considered as valid. Despite this, these valid data may still include a non-negligible noise component. This will result in contaminating the model estimates and consequently in an increase of the level of prediction error in the short run. It has then been envisaged to robustify the estimation method used in WPPT. Robust estimation in ARX-models can be performed using the method described in (Sejling et al., 1994). However, the power curve modeling part of WPPT, which is based on a local polynomial regression (Nielsen et al., 2000a), has not been robustified so far. This is discussed in the present paper.

Let ε denote a model residual. The classical criterion function ρ used for estimation in local polynomial regression is a quadratic criterion, i.e. $\rho(\varepsilon) = \varepsilon^2$. It thus gives a high weight to large model residuals when recursively adapting model estimates. Write $\hat{\theta}$ the estimator based on the ρ criterion function. Following (Huber, 1981), this function is modified in order to downweight these large residuals. Write ρ^* the criterion function for robust estimation. ρ^* is made linear for residuals larger than a certain threshold value c. It is defined as

$$\rho^*(\varepsilon, c) \;=\; \begin{cases} \varepsilon^2/2 & , \ |\varepsilon| \leq c \\ c|\varepsilon| - c^2/2, & |\varepsilon| > c \end{cases} \tag{15}$$

and the related robust estimator is referred to as $\hat{\theta}^*$.

The power curve for a wind farm may change over time owing to e.g. changes in the surroundings or aging of the turbines. Therefore, it is considered that the criterion function may also be non-stationary. And, because distributions of model residuals may be asymmetric and skewed, the symmetry constraint on the definition of ρ^* is relaxed. Instead of specifying c, one defines a proportion α of model residuals to be considered as suspicious (thus downweighted). The related threshold points are estimated by using the empirical distribution of recent model residuals, so that the same proportion of positive and negative model residuals are downweighted. A benefit of this approach is that the threshold parameters are made a function of the time of the year. The resulting criterion

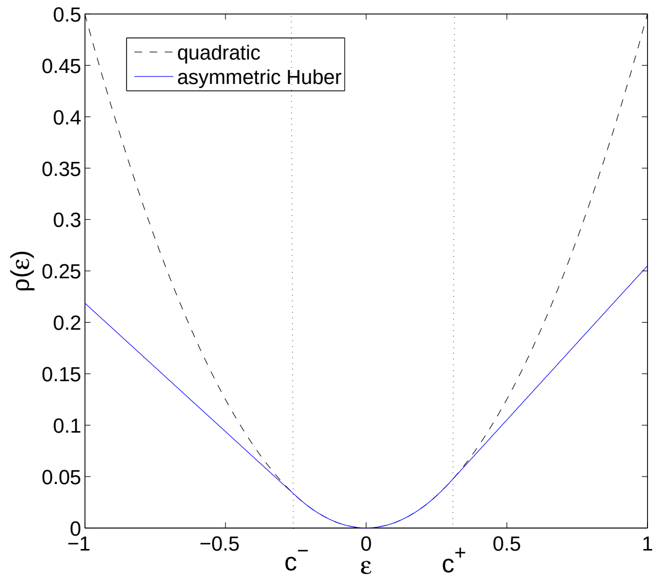

Figure 9: The 'usual' quadratic (ρ) and asymmetric Huber ($\rho^\dagger$) criterion functions. The thresholds points c^- and c^+ locate the negative and positive transitions from quadratic to linear criteria. Here these points are such that c^- = -0.25 and c^+ = 0.3. Negative residuals larger than c^- (in absolute value) and positive residual larger than c^+ are then downweighted when updating the model estimates.

function $\rho^\dagger$ writes

$$\rho^\dagger(\varepsilon, \alpha) \; = \; \begin{cases} c_\alpha^- \varepsilon - {c_\alpha^-}^2/2, & \varepsilon < c_\alpha^- \\ \varepsilon^2/2 & , \; \varepsilon \in [c_\alpha^-, c_\alpha^+] \\ c_\alpha^+ \varepsilon - {c_\alpha^+}^2/2, & \varepsilon > c_\alpha^+ \end{cases} \tag{16}$$

where $\mathbf{c}$ is the vector of negative and positive threshold values, denoted by c^- and c^+, respectively, uniquely defined by α. The related robust estimator is $\hat{\theta}^\dagger$. The two criterion function ρ and $\rho^\dagger$ are depicted in Figure 9. All the mathematical developments and details about the proposed method for the robustification of local polynomial regression are described in (Pinson et al., 2007c).

Simulation results based on semi-artificial datasets allow highlighting of the properties and performance of the robust estimators. Focus is given to the modelling of the power curve of the Klim wind farm (21MW) located in Northern Jutland. This regression function is nonlinear, bounded and non-stationary. By semi-artificial is meant that the wind speed measurements are the real measurements from the meteorological mast at the wind farm, but that the related power values are obtained by transformation through a modelled power curve. Both time-series cover a period of N = 10000 time steps with an hourly time resolution. They are normalized so that they take values in the unit interval. At any time step, the relation between wind speed and the 'true' power output is given by a power curve modelled as a double exponential function whose parameters linearly vary over time. The resulting non-stationary power curve is depicted in Figure 10. Note that by considering that the power curve is a function of wind speed only, we assume that other variables e.g. wind direction do not influence this power curve. This may not be true for real-world test cases. Though, the interest of this semi-artificial dataset is that the true power curve, which is the target regression, is available and can be used for evaluating the various estimators. The true wind speed and power data are then corrupted by additive and impulsive Gaussian noises in order to obtain simulated but realistic wind speed and power measurements. The characteristics of the noise have been derived based on the expertise gained from a large number of wind power modelling and forecasting applications. The way the data have been generated is further detailed in (Pinson et al., 2007c). The corrupted data are also shown in Figure 10.

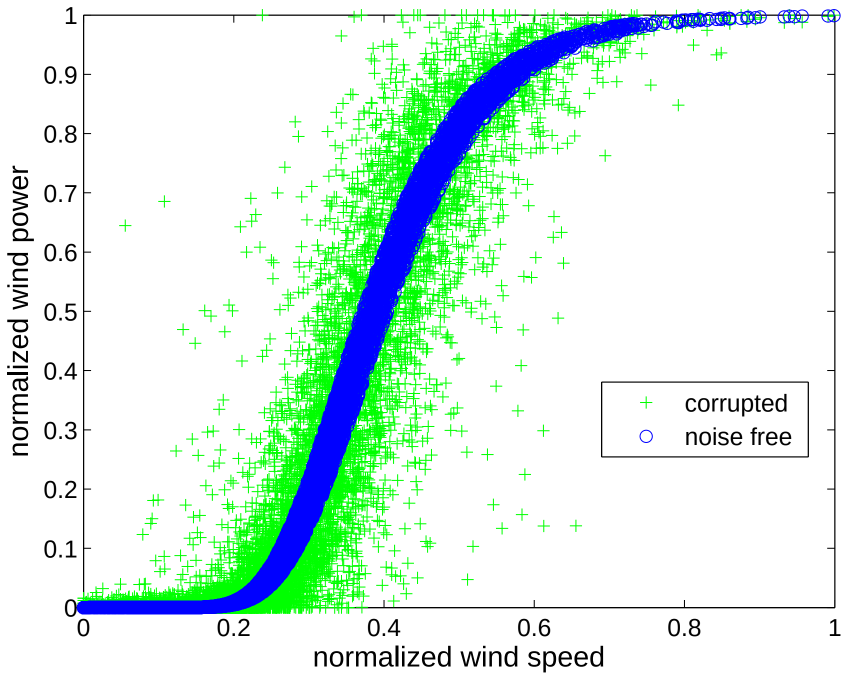

Figure 10: Noise-free and corrupted power curves.

The dataset is split into a learning set (2000 points), a cross validation set (2000 points), and an evaluation set (6000 points). The first two parts are used for determining the optimal parameters for $\hat{\theta}$, while the latter serves for independently evaluating the performance of the various estimators. Following (Madsen et al., 2005b), the criteria used for performance evaluation are the Normalized Mean Absolute Error (NMAE) and Normalized Root Mean Square Error (NRMSE). They are calculated against the true power data (NMAE_t and NRMSE_t) and also against the corrupted power data (NMAE_r and NMAE_r), which would correspond to operational conditions.

The robust estimators $\hat{\theta}^*$ and $\hat{\theta}^\dagger$ are applied with various c and α values in order to minimize the NRMSE_t criterion on the evaluation set. These minima are reached for $c = 0.15$ and $\alpha = 0.28$. This means that for $\hat{\theta}^*$, model residuals larger than 0.15 (in absolute value) are downweighted, while a proportion of 28% of the largest model residuals to be downweighted is optimal for the case of $\hat{\theta}^\dagger$. The minimum NRMSE_t for all competing estimators and related values of the other criteria are gathered in Table 4.

	$\hat{\theta}$	$\hat{\theta}^*$	$\hat{\theta}^\dagger$
NMAE_r	7.1408	6.9742	6.9168
NMAE_t	2.3928	2.0047	1.7426
NRMSE_r	11.4830	11.4844	11.4960
NRMSE_t	2.9793	2.5067	2.1958

Table 4: Minimum values of the NRMSE_t and related values of the other evaluation criteria for the competing estimators.

Both robust estimators allow for better approximation of the true power curve model. Indeed, error criteria calculated against the 'true' power data (NMAE_t and NRMSE_t) exhibit significant decreases when going from the classical towards the robust estimators. For instance, the decrease in NRMSE_t is 15.9% when going from $\hat{\theta}$ to $\hat{\theta}^*$ and 26.3% when going from $\hat{\theta}$ to $\hat{\theta}^\dagger$. However, when considering the error criteria calculated against the corrupted power data (which would correspond to what happens when dealing with real-world test cases), one sees that NRMSE_r stays at a similar level. This shows that if concentrating on a NRMSE criterion only when evaluating a prediction model in operational conditions, the benefits of robust estimation would not visible. Instead, the NMAE_r seems more appropriate since it is also lowered significantly. Further benefits of the robust estimators in operational conditions are discussed in (Pinson et al., 2007c), where they are used for forecasting on real-word data at the Middelgrunden wind farm in Denmark. It is shown that the classical adaptive estimator $\hat{\theta}$ is already fairly robust against outliers that may be used for model adaptation. In addition, lower values of error criteria were recorded when using both robust estimators. An advantage of $\hat{\theta}^\dagger$ over $\hat{\theta}^*$ is that its performance is less sensitive to the choice of its robustification parameter α.

6.3 System consequences

The above discussion presented ways of generating estimates for future wind power production and applied various criteria for measuring their quality. One common aspect of these criteria is that there is no formalised linkage between the consequences of application of a particular estimation methodology and system aspects. Here, the latter refers to aspects of the electricity system as a whole, in particular the total imbalance between the forecasted and realised amount of wind power and the overall consequences (in terms of economy, stability and otherwise).

The question here is what the relations between the way that wind power producers make their forecasts for the electricity market and the overall operation of the electricity system for which the transmission system operator (TSO) is responsible.

Wind power producers need a prediction method to bid their energy to a day-ahead market. They pay balancing costs for the energy that has been forecasted incorrectly. These prices, in turn, reflect the cost of balancing electricity production and/or demand elsewhere in the system. If the prices for up- and down-regulation are asymmetric, it is worthwhile for wind power producers to bid a different amount than the forecasted amount. In this way the balancing costs are minimised. Therefore minimising the costs may result in different bids than minimising the errors in energy.

This has been analysed as part of this project. The result, based on data from Western and Eastern Denmark, is described in (Holttinen and Ikäheimo, 2007). It shows that the balancing cost as covered by wind power producers is in general reduced slightly (between 2 and 7 %) by using such a bidding strategy. Hence, there is a motivation to use such strategy.

It was also investigated, using the same data, whether a bidding strategy derived from the discussion above will have consequences at system level. It was shown that indeed there will be consequences, as seen from Figure 11. In particular, it was found that for large errors downwards (forecast production larger than realized production), the information (i.e., the wind power bids) received by the transmission system operator (TSO), would result in larger errors and therefore more up-regulation would be needed at system level.

Whether this is critical, in e.g. stability or economic terms, depends on the available up-regulation capacity, its costs and other aspects, however, it has been beyond the scope of this project to go into this.

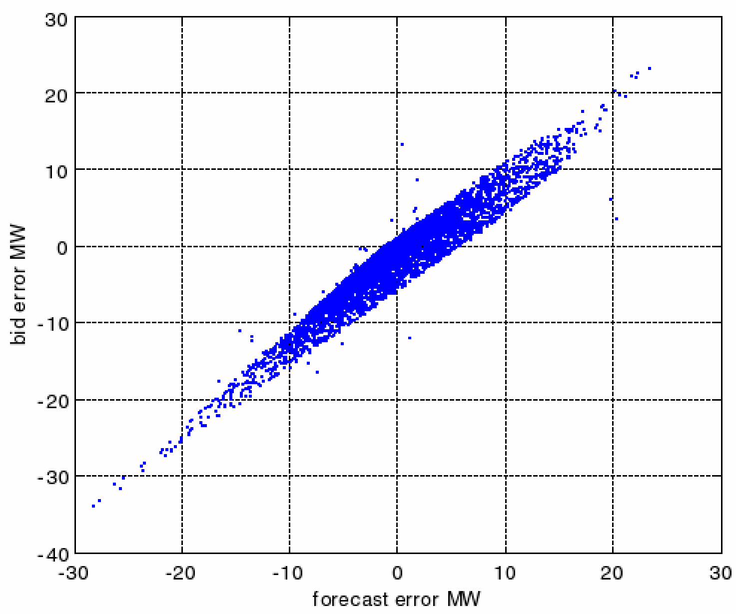

Figure 11: Summed errors for both parks calculated from optimal bids (on ordinate) plotted against summed errors for both parks calculated from power forecasts (abscissa).

7 Using several meteorological forecasts

Operational robustness can be improved by having access to meteorological forecasts from several suppliers. However, most of the time the wind power forecasting system will have access to all meteorological forecasts and we consider how, and under which circumstances, this extra information can be used to improve the forecasts.

The use of multiple meteorological forecasts will, especially when these come from different suppliers, result in operational robustness. This is obtained by running multiple instances of the forecasting software, e.g. WPPT, and thereby producing multiple wind power forecasts. Most of the time all these forecasts will be available and the question then arises if additional benefits can be obtained by combining the individual wind power forecasts.

Let $\hat{y}_1$ and $\hat{y}_2$ be the two individual wind power forecasts and assume for simplicity that these are unbiased. Bates and Granger (1969) consider the case where the combined forecast $\hat{y}_c$ is obtained as a weighted average:

$$\hat{y}_c = w_1\hat{y}_1 + (1 - w_1)\hat{y}_2 \ , \tag{17}$$

where w_1 is the weight on forecast number 1. The variance of the forecast error of the combined forecast depend on the weight, the variances of the individual forecast errors (σ_1^2 and σ_2^2) and the correlation ρ between the individual forecast errors. Bates and Granger (1969) provide a formula for the value of w_1 minimizing the variance σ_c^2 of the combined forecast error together with an expression of the minimal variance.

If, arbitrarily, the individual forecast number 1 is assumed to be the best $\sigma_1 \leq \sigma_2$ and the relative improvement over forecast number 2 is expressed as $r_1 = (\sigma_2 - \sigma_1)/\sigma_2$, then the relative improvement of the optimal combination over the best individual forecast $r_c = (\sigma_1 - \sigma_c)/\sigma_1$, can be expressed as:

$$r_c = 1 - \sqrt{\frac{1 - \rho^2}{(1 - r_1)^2 - 2\rho(1 - r_1) + 1}} \ , \tag{18}$$

which is defined for $|\rho| < 1$. Figure 12 depicts r_c as a function of ρ and r_1 . It is seen that given two forecasts, with identical performance ($r_1 = 0$), for which the individual forecast errors are correlated with a correlation coefficient of $\rho = 0.7$, the combined forecast will perform 8% better than the best individual forecast. If the forecasts do not perform equally well ($r_1 > 1$) the correlation must be lower in order to achieve the same improvement.

The method can, in principle, be extended to an arbitrary number of individual, possibly biased, forecasts (de Menezes and Taylor, 2000; Granger and Ramanathan, 1984; Nielsen et al., 2007a). In practice the means, variances, and covariances/correlations of the individual forecast errors must be estimated.

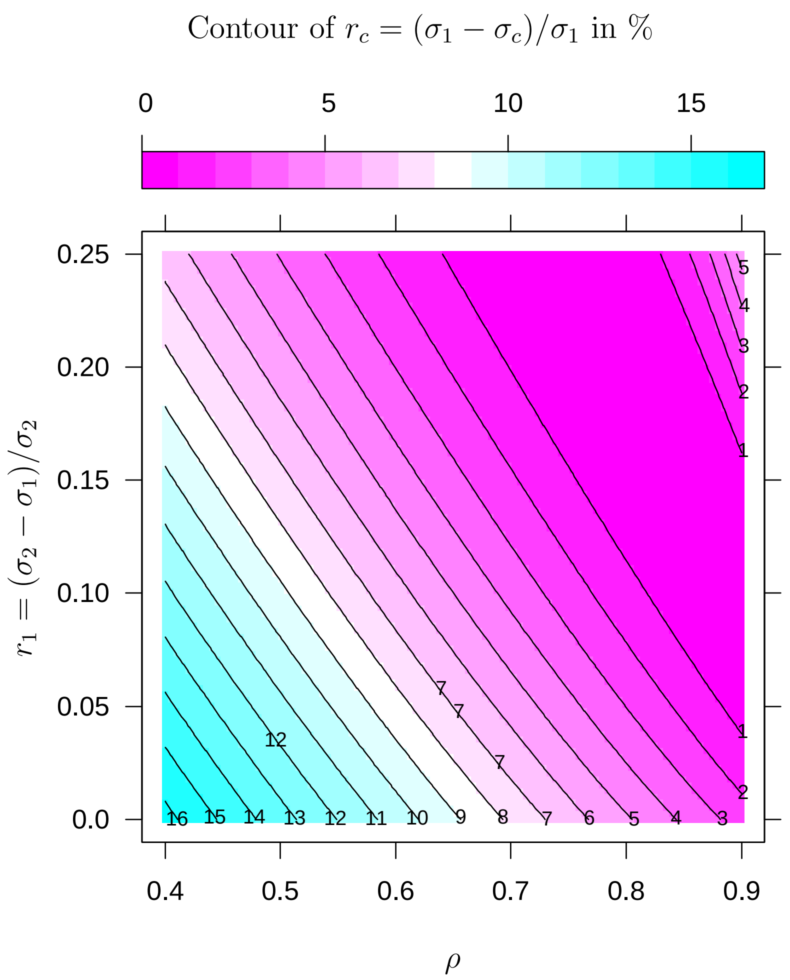

Figure 12: Combination of two unbiased individual forecasts with performance $\sigma_1 < \sigma_2$. The plot shows the relative improvement r_c as a function of the correlation ρ and the relative performance r_1 of method 1 as compared to method 2. In the top-right corner the weight on method 2 is negative.

For application to wind power forecasting the following questions arise:

(i) What level of correlation and relative performance can be expected for well calibrated wind power forecast errors based on different meteorological forecasts?

(ii) What will be the effect of estimation error on the quality of the combined forecast?

As part of the project these aspects were investigated for two wind farms, considering forecast horizons up to 24 hours (Nielsen et al., 2007a). An adaptive estimation procedure was applied.

With respect to (i) it has been found that when the meteorological forecasts originate from models nested in different global models, then the correlation is between 0.45 and 0.75 for the wind farm Klim in Denmark and between 0.65 and 0.80 for the wind farm Alaiz in Spain. In many cases the performance of the individual forecasts are similar. Based on Figure 12 it is therefore suggested that improvements of 5 to 15% over the best individual forecast can be obtained. In fact, this is the range of improvements which is obtained. For Klim an overall level (across horizons) of 9% is obtained and for Alaiz an overall level of 4% is obtained (Nielsen et al., 2007a). These results are obtained when combining three individual forecasts based on meteorological models nested in different global models.

Given the true means, variances, and covariances of the individual forecast errors, the combined forecast can not perform worse than any of the individual forecasts. With respect to (ii), when estimating the quantities the combined forecast can in fact perform worse than the best of the individual forecasts. Hence, based on the investigations in (Nielsen et al., 2007a) it is recommended that two or three good meteorological forecasts are used and the forecast errors of these should have low correlation (less that approximately 0.8). This seems to be the case for meteorological forecasts originating from different global models. Since this also markedly improves the operational robustness, we find this combination procedure to be sufficient. Sánchez (2006) considers methods which seem to allow for many forecasts, possibly with high correlation, to be combined.

8 Conclusion and discussion

We have demonstrated that it is indeed possible to achieve further automation and robustness by modification of the methods underlying tools such as WPPT. Operational robustness is easily achieved by using meteorological forecasts from two or three different suppliers. However, most of the time all meteorological forecasts will be available. When combining these in an optimal manner, large improvements in performance are observed. This is especially true when the forecasts originate from different global meteorological models; this further adds to the operational robustness.

With respect to initialization of the statistical models, it has been demonstrated that WPPT quickly learns the relation between meteorological forecasts and power production. However, for some combinations of wind speed and direction which seldom occur, it is beneficial to use a good physically based power curve as initialization. In locations around the globe, cut-out due to high wind speeds occur so seldom that, in practice, the cut-out characteristics of the wind farms can not estimated from on-site observations. The proposed method of initialization, in principle, solves this problem by using the physical power curve in regions where data is limited. In practice, further refinement of the statistical models are needed in order to accommodate the characteristics of the physical power curve near the cut-out wind speed.

For complex sites such as Alaiz in Spain it has been demonstrated that stability measures derived from the meteorological forecasts can markedly improve the forecasts of wind speed. This has been shown in two ways. The first way is through a simple linear regression based wind prediction system, where one of a range of candidate stability measures is used in the regression. It was found that a stability measure defined over a vertical extent similar to the vertical extent of the terrain in the region was most effective. The stability measure improved the mean absolute error on the wind speed predictions by nearly 17 % for the Alaiz wind speed predictions. Therefore this part of the study indicates a benefit of using more of the information within operational weather forecasting models. A preprocessor system for WPPT is suggested to carry out a stability adjustment to the input wind speed. The second way is through mesoscale simulations using both KAMM and WRF mesoscale models. The mesoscale modelling, through idealized simulations, demonstrated the ability of the models to capture the strong features of the observed wind speed and direction distributions. A mesoscale downscaling look-up table prediction system was tested. The look-up table entries were based on 103 wind class simulations giving the wind speed and direction at the wind farm location for different wind forcing and stability conditions. Although the performance was not as good as the predictions made by the linear regression using stability measures, the performance achieved was better than simply using GFS predictions and MOS. This demonstrates that the mesoscale modelling, either by KAMM or WRF, is able to add information about the local flow around Aliaz and improve wind speed predictions.

Adaptive and non-parametric methods such as those used in WPPT apply a number of tuning parameters. Normally, these have to be setup for each installation. We have presented methods which add an additional layer of adaption, by adapting these tuning parameters. The criteria used for adapting these directly address the prediction performance.

The issue of robust and adaptive estimation of a wind farm power curve has been addressed. It has been shown that existing non-parametric estimators can be modified in order to limit the influence of highly suspicious data when adapting the parameters of a model. The ability of the resulting robust estimators to better approximate the true power curve of a wind farm has been demonstrated from simulations, and the related benefits in operational conditions have been discussed. Such robust methods will allow for enhancement of the quality of the estimation of the direction dependent power curve in on-line situations where the data quality cannot always be insured.

It has been shown that the choice of methods for generating wind power prognoses for use in the electricity market may have economic consequences for wind power producers. In particular it was shown that it may be profitable to make prognoses that are biased in order to minimize balancing costs. It was also shown that such prognoses may have consequences at the level of the electricity system as whole; however, it has been beyond the scope of the project to go further into this.

References

Adrian G. and Fiedler F. Simulations of unstationary wind and temperature fields over complex terrrain and comparion with observations. *Beitr. Phy. Atmosph.*, **64**:27–48, 1991.

Badger J., Giebel G., Larsen X.G., Nielsen T.S., Nielsen H.A., Madsen H., and Tøfting J. Report on the use of stability parameters and mesoscale modelling in short-term prediction. Technical Report Risø-R-1614, Risø National Laboratory, Technical University of Denmark, Frederiksborgvej 399, DK-4000 Roskilde, 2007. `http://www.risoe.dk/rispubl/VEA/ris-r-1614.pdf`.

Bates J.M. and Granger C.W.J. The combination of forecasts. *Operational Research Quarterly*, **20**(4):451–468, 1969.

Bremnes J.B. Probabilistic wind power forecasts using local quantile regression. *Wind Energy*, **7**(1):47–54, 2004.

Christiansen L.E., Nielsen H.A., Nielsen T.S., and Madsen H. Automatic selection of tuning parameters in wind power prediction. Technical Report IMM-2007-12, Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2007. `http://www.imm.dtu.dk/pubdb/views/publication_details.php?id=5316`.

Cleveland W. *Multivariate Analysis and Its Applications*, chapter Coplots, Nonparametric Regression, and conditionally Parametric Fits, pp. 21–36. Institute of Mathematical Statistics, Hayward, 1994.

de Menezes L.M. and Taylor J.W. Review of guidelines for the use of combined forecasts. *European Journal of Operational Research*, **120**(1):190–204, 2000.

Frank H.P. and Landberg L. Modelling the wind climate of ireland. *Boundary-Layer Meteorology*, **85**:359–378, 1997.

Frey-Buness F., Heimann D., and Sausen R. A statistical-dynamical downscaling procedure for global climate simulations. *Theor. Appl. Climatol.*, **50**:117–131, 1995.

Giebel G., Badger J., Landberg L., Nielsen H.A., Nielsen T.S., Madsen H., Sattler K., Feddersen H., Vedel H., Tøfting J., Kruse L., and Voulund L. Wind power prediction using ensembles. Technical Report Risø-R-1527(EN), Risø National Laboratory, 2005. ISBN 87-550-3464-0.

Giebel G., Badger J., Louka P., Kallos G., Perez I.M., Lac C., Palomares A.M., and Descombes G. Short-term forecasting using advanced physical modelling: The results of the anemos project. In *Proceedings of the European Wind Energy Conference & Exhibition*. Athens, Greece, 2006. `http://www.ewec2006proceedings.info/allfiles2/967_Ewec2006fullpaper.pdf`.

Granger C.W.J. and Ramanathan R. Improved methods of combining forecasts. *Journal of Forecasting*, **3**:197–204, 1984.

Haykin S. *Adaptive Filter Theory*. Prentice Hall, 3rd edition, 1996.

Holttinen H. and Ikäheimo J. Wind prediction and bids in Denmark. VTT Energia, 2007. `http://www.imm.dtu.dk/~han/pub/FU4101_DKvind_bids_and_balance_VTT_2006.pdf`.

Huber P. *Robust Statistics*. Wiley & Sons, New York, 1981.

Koenker R. *Quantile Regression*. Cambridge University Press, 2005.

Linnet U. *Tools supporting wind energy trade in deregulated markets*. Master's thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2005. URL `http://www.imm.dtu.dk/pubdb/views/publication_details.php?id=3969`.

Ljung L. *System Identification, Theory for the User*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

Ljung L. and Gunnarsson S. Adaption and tracking in system identification – a survey. *Automatica*, **26**:7–22, 1990.

Madsen H., Nielsen H.A., and Nielsen T.S. A tool for predicting the wind power production of off-shore wind plants. In *Proceedings of the Copenhagen Offshore Wind Conference & Exhibition*. Danish Wind Industry Association, Copenhagen, 2005a. `http://www.windpower.org/en/core.htm`.

Madsen H., Pinson P., Kariniotakis G., Nielsen H.A., and Nielsen T.S. Standardizing the performance evaluation of short-term wind prediction models. *Wind Engineering*, **29**(6):475–489, 2005b. URL `http://www2.imm.dtu.dk/pubdb/p.php?4262`.

Morthorst P. Wind power and the conditions at a liberalized power market. *Wind Energy*, **6**(3):297–308, 2003.

Nielsen C.S. and Ravn H.F. Criteria in short-term wind power prognosis. In *Proceedings of the European Wind Energy Conference & Exhibition*. Madrid, Spain, 2003. `http://www.ewea.org`.

Nielsen H.A., Nielsen T.S., Joensen A.K., Madsen H., and Holst J. Tracking time-varying coefficient-functions. *Int. J. of Adaptive Control and Signal Processing*, **14**(8):813–828, 2000a.

Nielsen H.A., Nielsen T.S., and Madsen H. On on-line systems for short-term forecasting for energy systems. In *Proceedings of the OR 2002 conference*, pp. 265–271. Springer, Klagenfurt, Austria, 2002a.

Nielsen H.A., Nielsen T.S., and Madsen H. Using meteorological forecasts for short term wind power forecasting. In *Proceedings of the IEA R&D Wind Annex XI Joint Action Symposium on Wind Forecasting Techniques*, pp. 49–58. Norrköping, Sweden, 2002b.

Nielsen H.A., Nielsen T.S., Madsen H., and Giebel G. Candidate prediction models and methods. Informatics and Mathematical Modelling, Technical University of Denmark, 2005. `http://www.imm.dtu.dk/~han/pub/FU4101_Candidate_pred_models.pdf`.

Nielsen H.A., Nielsen T.S., Madsen H., Giebel G., Badger J., Landberg L., Sattler K., Voulund L., and Tøfting J. From wind ensembles to probabilistic information about future wind power production – results from an actual application. In *Prooceedings of the 9th International Conference on Probabilistic Methods Applied to Power Systems*. 2006a. Invited paper.

Nielsen H.A., Nielsen T.S., Madsen H., San Isidro M.J., and Marti I. Optimal combination of wind power forecasts. *Wind Energy*, 2007a. Accepted, `http://www.imm.dtu.dk/~han/pub/FU4101_WEpaperCFCaccepted.pdf`.

Nielsen H.A., Pinson P., Christiansen L.E., Nielsen T.S., Madsen H., Badger J., Giebel G., and Ravn H.F. Improvement and automation of tools for short term wind power forecasting. In *Scientific Proceedings of the European Wind Energy Conference & Exhibition*. Milan, Italy, 2007b. `http://www.ewec2007proceedings.info/allfiles2/475_Ewec2007fullpaper.pdf`.

Nielsen T., Madsen H., and Christensen H. WPPT – a tool for wind power prediction. In *Proceedings of the Wind Power for the 21st Century Conference*. Kassel, Germany, 2000b.

Nielsen T., Madsen H., and Christensen H.S. WPPT - a tool for wind power prediction. In *Proceedings of the Wind Power for the 21st Century Conference*. Kassel, 2000c.

Nielsen T.S., Nielsen H.A., and Giebel G. Initialisation of power curve models using à priori information. Technical Report IMM-2007-13, Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2007c. `http://www.imm.dtu.dk/pubdb/views/publication_details.php?id=5329`.

Nielsen T.S., Nielsen H.A., Madsen H., San Isidro M.J., and Marti I. Optimal wind power forecasting by model combination. Poster presented at the European Wind Energy Conference & Exhibition 2006, Athens, Greece, 2006b. `http://www.imm.dtu.dk/~han/pub/FU4101_EWEC06poster.pdf`.

Nielsen T.S., Nielsen H.A., Pinson P., Madsen H., Giebel G., and Badger J. Improvement and automation of tools for short term wind power forecasting. Poster presented at Global Windpower 2006, Adelaide, South Australia, 2006c. `http://www.imm.dtu.dk/~han/pub/FU4101_GW06poster.pdf`.

Nolsøe K. Estimating the power curve. Informatics and Mathematical Modelling, 2006. `http://www.imm.dtu.dk/~han/pub/FU4101_EstimatingThePowerCurve.pdf`.

Pinson P., Nielsen H., Madsen H., and Nielsen T. Local linear regression with adaptive orthogonal fitting for the wind power application. *Statistics and Computing*, 2007a. Under revision.

Pinson P., Nielsen H., Møller J., Madsen H., and Kariniotakis G. Nonparametric probabilistic forecasts of wind power: required properties and evaluation. *Wind Energy*, 2007b. In press.

Pinson P., Nielsen H.A., and Madsen H. Robust estimation of time-varying coefficient functions – application to the modeling of wind power production. Technical Report IMM–2007-09, Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2007c. `http://www.imm.dtu.dk/pubdb/views/publication_details.php?id=5275`.

Ravn H.F. Short term wind power prognosis with different success criteria. In *Proceedings of the 9th International Conference on Probabilistic Methods Applied to Power Systems*. 2006. Invited paper. `http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?isnumber=4202205&arnumber=4202303&count=196&index=97`.

Sánchez I. Adaptive combination of forecast with application to wind energy forecast. In *9th International Conference on Probabilistic Methods Applied to Power Systems*. Power Engineering Society, IEEE, KTH, Stockholm, Sweden, 2006.

Sejling K., Madsen H., Holst J., Holst U., and Englund J. Methods for robust estimation of ar-parameters. *Computational Statistics and Data Analysis*, **17**:509–536, 1994.

Skamarock W.C., Klemp J.B., Dudhia J., Gill D.O., Barker D.M., Wang W., and Powers J.G. A description of the advanced research wrf version 2. *NCAR Technical Note NCAR TN 226468+STR*, 2005.