Power fluctuations from large wind farms - Final report

Risø-R-Report

Poul Sørensen, Pierre Pinson, Nicolaos Antonio Cutululis,
Henrik Madsen, Leo Enrico Jensen, Jesper Hjerrild,
Martin Heyman Donovan, Jesper Runge Kristoffersen,
Antonio Vigueras-Rodríguez
Risø-R-1711(EN)
August 2009

Risø DTU
National Laboratory for Sustainable Energy

Author: Poul Sørensen, Pierre Pinson, Nicolaos Antonio Cutululis, Henrik Madsen, Leo Enrico Jensen, Jesper Hjerrild, Martin Heyman Donovan, Jesper Runge Kristoffersen, Antonio Vigueras-Rodríguez
Title: Power fluctuations from large wind farms - Final report
Division: Wind Energy Division

Abstract (max. 2000 char.):

Experience from power system operation with the first large offshore wind farm in Denmark: Horns Rev shows that the power from the wind farm is fluctuating significantly at certain times, and that this fluctuation is seen directly on the power exchange between Denmark and Germany. This report describes different models for simulation and prediction of wind power fluctuations from large wind farms, and data acquired at the two large offshore wind farms in Denmark are applied to validate the models. Finally, the simulation model is further developed to enable simulations of power fluctuations from several wind farms simultaneously in a larger geographical area, corresponding to a power system control area.

Risø-R-1711(EN)
August 2009

ISSN 0106-2840
ISBN 978-87-550-3782-3
Group's own reg. no.:
1115044-01
Sponsorship:

Energinet.dk ForskEl

Contract no.:
6506

Pages: 49
Tables: 4
References: 44

Information Service Department
Risø National Laboratory for Sustainable Energy
Technical University of Denmark
P.O.Box 49
DK-4000 Roskilde
Denmark
Telephone +45 46774005
bibl@risoe.dtu.dk
Fax +45 46774013
www.risoe.dtu.dk

Contents

Preface

This is the final report of the PSO project “Power fluctuations from large wind farms”, funded by Energinet.dk as project number 6506. Risø National Laboratory – DTU (CVR DK42154113) has been responsible for the project.

1 Summary

The objective of the project has been to develop and validate models for characterisation, simulation and forecasting of the power variations in wind farms. The validation is based on wind and power measurements from individual wind turbines in the two large offshore wind farms in Denmark: Horns Rev and Nysted.

The data collection has been very continuous for the Nysted wind farm during the project period, while the data collection in Horns Rev has had some interruptions. Still, a unique database of measurement is now available for both wind farms, consisting of 1 second samples of wind speed, yaw angle (wind direction indicator) and power from each wind turbine in the wind farms. The collected 1 second values are instantaneous values, so since there is some variability within the seconds as well, especially for the nacelle wind speeds and for the power from fixed speed wind turbines (i.e. Nysted), the measurements should not be used to analyse the variability with time steps less than ½ - 1 minute.

The time scale of the fluctuations which has been in focus in the project is from a few minutes to 1-2 hours. This time scale has been selected based on observed variability of the power, especially from the Horns Rev wind farm. The time scale is between the turbulence time scale (typically less than a few minutes) and the electricity balancing markets operating on one hour averages. Instead, it addresses the time scale of interest for regulating power markets.

On the modelling site, two tracks have been followed. The first track has been to develop spatial correlation models. The advantage of this approach is that the models can be used to simulate non-existing wind farms, i.e. as a planning tool. The second track has been to develop regime-switching models. The advantage of this approach is that the models can be used as prediction tools, i.e. in the operation.

A detailed spatial correlation model has been developed and implemented. It was built on an original model, simulating wind speed fluctuations based on turbulence. Thus, the main focus has been on adding the variability in the time range between minutes and a few hours. This has also included the spatial correlation, which has been modelled in the frequency domain as coherence functions. The new model also takes into account the turbulence which is added inside the wind farm due to the upwind wind turbines.

The strong side of the spatial correlation model is that it can be applied to simulate non-existing wind farms. It only requires the spatial position of the wind turbines (x-y coordinates and hub height), the rotor diameters, and power curves. The basic assumption is that in the time scale of interest, the dynamics of the wind turbines influence only marginally the power fluctuations, and thus the main modelling challenge is in describing the wind speed variability.

The frequency domain input for the spatial correlation models has been estimated empirically based on measurements in Risøs test site for large wind turbines in Høvsøre. The data obtained from Høvsøre has been supplemented with the data from Horns Rev and Nysted.

Besides the detailed spatial correlation model, an aggregated model has also been proposed. The idea of the aggregated model is to use the average wind speed as input to an aggregated wind turbine, applying an aggregated power curve for that wind turbine.

The detailed spatial correlation model has been verified applying the Horns Rev as well as Nysted power measurements. For that purpose, methods to quantify ramp rates as well as reserve requirements based on power time series has been defined. Then these methods have been applied on the wind power measurements as well as simulations, for Horns Rev as well as Nysted.

The result of this verification is generally acceptable. The models simulates the measured ramp rates and reserve requirements relatively well when period times of 10 minutes and 30 minutes are applied, but for the small period time of 1 minute, the simulation result deviates relatively more from the measurements. This is as expected, firstly because the ramp rates and reserve requirements are quite small with 1 minute period time, secondly because the dynamics of the wind turbine within 1 minute is not modelled, and thirdly because the measurements are unfiltered instantaneous values and therefore significant noise is added due to aliasing.

In parallel, several approaches to regime-switching modelling has been introduced and evaluated. The motivation for applying such class of approaches comes from the observation of successive periods with power fluctuations of smaller and larger magnitude in the collected time-series of power production both at Horns Rev and Nysted. The regime-switching modelling approaches applied can be sorted into two different categories. On the one hand, they may rely on an observable regime sequence, which in the present case is defined by the last observed value of power generation at the wind farm. On the other hand, Markov-switching approaches rely on an unobservable regime sequence, which is modelled here with a first-order Markov chain. Such approach may permit to capture the influence of some complex meteorological phenomenon on the evolution of time-series of offshore wind power generation. Indeed, it has been shown in the frame of the project that the latter type of regime-switching approaches outperforms the former one, in terms of 1-step ahead forecasting performance. In addition, the proposed Markov-switching approach has the advantage that it allows for characterization of regime sequences. They can be compared in the future with the evolution of different meteorological variables, in order to see which of them could explain the different behaviours of power fluctuations, and then be used as input to advanced statistical models.

Focus has been given to enhancing introduced Markov-switching approach in order to make it more suitable for the wind power application. Mainly, a model with time-varying coefficients is described, so that employing the related adaptive estimation method allows for the tracking of long-term changes in the characteristics of the wind generation process. The proposed model and related estimation method can also be used for probabilistic forecasting, which has the advantage of informing on the probabilities of power fluctuations of smaller or larger magnitude.

2 Conditions for the project

The project has been funded by Energynet.dk as PSO project 6506. The participants have been the Wind Energy Department of Risø National Laboratory (now part of DTU), IMM DTU, and the wind farm owners. The wind farm ownership has changed during the project, as a result of the DONG Energy fusion. Thus, initially Tech-wise (part of Elsam, the owner of Horns Rev) and Energy E2 (owner of Nysted) participated. In the end of the

project, DONG Energy (owner of Nysted and 40 % owner of Horns Rev) and Vattenfall A/S (60 % owner of Horns Rev) have participated.

The project has formally been running from 1 October 2004 to 30 September 2008. The project kick-off meeting was held 3 February 2005. At that meeting, it was agreed to use the existing data infrastructures to collect the measurements, and it was agreed that a one second sampling time was preferred. Since then, project meetings have been held twice a year, and the remaining communication has been done via telephone and e-mail.

The budget was kkr.3410, with kkr.3071, originally accepted by Elkraft System (now Energinet.dk). This corresponded to 5150 man hours.

3 Objectives

According to the project proposal, "the objective of the project is to develop and validate models for simulation of the power variations in wind farms. The models should be applicable to determine the required spinning reserve to keep a safe and reliable operation of the power system. Focus will be on power fluctuations from large offshore wind farms, and measurements from the large offshore wind farms Horns Rev as well as Nysted will be applied to validate the models. The project is based on existing models for wind fluctuations in wind farms, which only includes the faster fluctuations due to turbulence. These models will be extended to include the slower variations, i.e. variations from minutes to hours, due to e.g. unstable weather conditions or frontline passages. During a couple of hours, these slow variations are often relatively large, and as they are strongly correlated over the wind farm, they will not be smoothed as much as the faster fluctuations as the power from all the wind turbines in the wind farm are summed up. Therefore, the slow variations are particularly important for large wind farms and there influence on the requirement for spinning reserve in the power system. Besides power fluctuations, the project will develop aggregated models for wind farms, i.e. reduction of detailed models including individual models for all wind turbines to models where the whole wind farm is represented by a single wind farm with appropriate characteristics. Such a reduction is necessary to simulate power systems with thousands of wind turbines."

4 Execution

The description of the execution in this chapter is organised to reflect the working tasks according to the project proposal.

4.1 Data collection

The data collection has been fully satisfactory compared to the expectations in the project description. Wind speeds, yaw errors and power have been measured at each individual wind turbine in the farms with 1 second sample time. Besides, power set points have been available for all the wind turbines, ensuring that only power variations in periods without power reduction have been included in the validation. According to the project description, the 1 second samples could have been done only in parts of the wind farm and only in shorter periods. However, these measurements have been made over the whole wind farms, and during very long periods.

4.2 Simulation of wind variability

The most important task here has been to include the slow wind variability in Risø's original PARKSIMU model, i.e. on the spatial correlation model. This extension has been done based on measurements in Høvsøre as well as the Horns Rev and Nysted wind farms. The results related to slow wind variability are described in more details in section 6.1.1.

The wind farm effects have also been included. The approach has been to apply a method specified in the wind turbine design requirement standard IEC 61400-1. This method has been verified based on the measurements on the wind turbine nacelles inside the wind farm, compared to those on the first upwind row of wind turbines.

The effect of different meteorological conditions in Horns Rev and Nysted has also been investigated. It is seen from the validation section 6.3 that the wind power fluctuates more in Horns Rev than in Nysted, even though Horns Rev uses variable speed wind turbines with generally more smooth power production.

Some of the results related to wind variability have been published in a peer reviewed paper in the journal Wind Energy [1].

4.3 Aggregated models

An aggregated model is described in section 6.2, reducing the detailed spatial correlation model. However, the main effort in the aggregated model task has been on the regime-switching simulation models that are described in chapter 7. The regime-switching models have been developed and described by IMM DTU. A number of different regime-switching models have been tested on Nysted data in the first place, and the Markov chain driven model has been identified as the best performing. This model has then been applied to Horns Rev as well. The model performed better on Nysted than on Horns Rev, which is probably related to the more unstable weather conditions in Horns Rev.

The regime-switching simulation models have been submitted for publication as a peer-reviewed paper in the journal Wind Engineering and Industrial Aerodynamics [2].

4.4 Validation and improvements

The spatial correlation model has been validated on Horns Rev data as well as Nysted as described in section 6.3. Some adjustment of the parameter quantifying the slow variations has been done to obtain good results in Horns Rev as well as Nysted.

The regime switching simulation model is trained on one data set and verified on another. This has been done independently for Nysted and Horns Rev.

The validation of the spatial correlation model on Horns Rev data has been published in a peer reviewed paper in IEEE Transactions on Power System [3].

4.5 Extension to power system region

The original project was extended with a task to simulate the wind power fluctuations in a power system region as supplement to the simulations of single wind farms, and to use this extension to simulate cut-out of wind farms at storm wind speeds. This is an extension of the spatial correlation model, and therefore the description of this work is included in chapter 6 as section 6.4.

The extension of the model has not yet been published, but there is substance enough in the work for a couple of peer-reviewed papers. It is the plan to use the model in other projects.

5 Observations

5.1 General observations

A fundamental issue in the operation and control of electric power systems is to maintain the balance between generated and demanded power. Scheduling of generation ensures that sufficient generation power is available to follow the forecasted load, hour by hour during the day [1]. Besides, sufficient reserves with response times from seconds to minutes must be available to balance the inevitable deviations from the schedules caused by failures and forecast errors.

In the Nordic power system the generation is scheduled on the NORDPOOL spot market [5] on the day ahead spot market, while the reserve capacities are agreed in the Nordel cooperation between the Nordic transmission system operators [6]. Those reserves include a “fast active disturbance reserve”, which is regulating power that must be available 15 minutes after allocation. The purpose of the regulating power is to restore the frequency controlled reserves, which are activated due to frequency deviations from the nominal value. The frequency controlled reserves are much faster than the regulating power, with required response times from a few minutes down to a few seconds, depending on the frequency error.

The wind power development influences the power balancing on all time scales. Wind power forecasts have been studied already when there was much less wind power in the systems than today, e.g. Troen and Landberg [7]. The wind power forecasts are used by the stakeholders on the day ahead market to reduce the forecast errors caused by wind variability and thereby optimize the bits on the market. On a shorter time scale, wind power variability also influences the power balancing, although this influence seems to become an issue only with a significantly higher amount of wind power in the system.

An example of this is the experience of the Danish transmission system operator, Energinet.dk, with the operation of the West Danish power system. According to Akhmatov et. al. [8], Energinet.dk has found that the active power supplied from the first large 160 MW offshore wind farm in this system, Horns Rev, is characterized by more intense fluctuations in the minute range than previously observed form the dispersed wind turbines on land, even though the installed power in Horns Rev is relatively small compared to the total 2400 MW wind power installation in the system.

The West Danish power system is DC connected to the Nordic synchronous power system, and it is a part of the Nordel cooperation. At the same time, the system is interconnected to the Central European UCTE system through the German AC connection. Therefore, the West Danish system also has obligations to the UCTE system, including the responsibility to keep the agreed power flow in this interconnection within acceptable limits.

With the present wind power capacity, the power flow in the interconnection can be kept within the agreed UCTE limits. However, a second neighbouring wind farm, the 200 MW Horns Rev B is already scheduled for 2008, and the Danish government has announced it’s energy policy, estimating the development with another 3000 MW

offshore wind power in Denmark 2025 [9]. Such a development will obviously influence the future demand for regulating power in the system.

In the United States, several detailed technical investigations of grid ancillary service impacts of wind power plants have been performed recently. Parson et. al. [10] summarizes the studies, which are also focused on the wind power variability. In the United States, the utilities deal with both load following and regulation on a control area basis, which is similar to the tasks of Energinet.dk in the West Danish control area. The studies in the United States quantify the impact of wind farms on load following as well as regulation power, based on measurements on several large wind farms.

On small island power systems, wind power fluctuations may have even more severe consequences for the operation and security of the system. An example of such a system is the island of Hawaii, which is not interconnected to other systems. The ability to integrate wind power in the system on the island of Hawaii is also limited by the generation resource mix, which includes a relatively large amount of non-regulating generating units [11]. The system operator HELCO has introduced wind power performance standards with limits for the wind power fluctuations in the seconds to minute range.

Wind power fluctuations are also a security issue in systems with very weak interconnections. The island of Ireland is a significant example of such a system, which has only a very weak connection to Scotland. Another example is the Iberian Peninsula with a weak connection to France.

5.2 Observations from measurements

Figure 1 shows the power generation in January 18 2005 of the wind turbines observed by the then owning power producer Elsam. It is divided into the power fluctuations from the Horns Rev wind farm and the fluctuations in the production of the “Onshore” wind turbines with a comparable installed capacity.

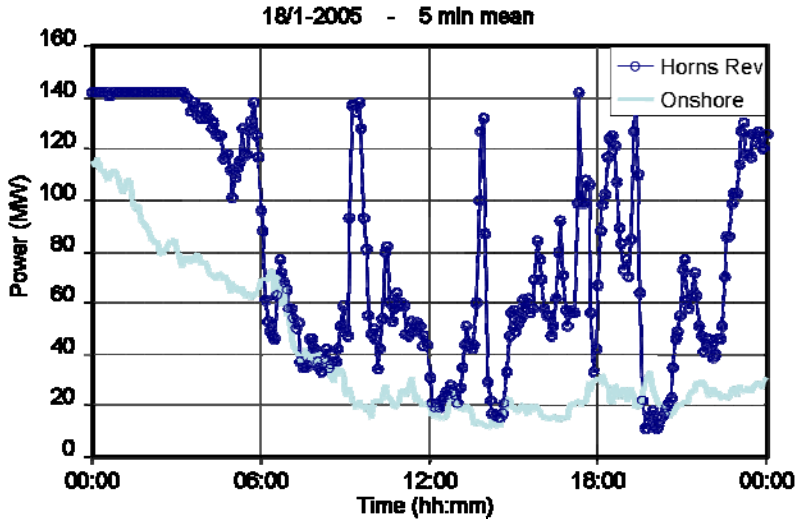

Figure 1. Power generation of Horns Rev offshore wind farm and onshore wind turbines January 18 2005.

The selected day is characterized by very unstable weather conditions, and therefore the power fluctuations from the Horns Rev wind farm are exceptionally high. At the same time, the fluctuations in the production of the “Onshore” wind turbines are much less.

Generally, the onshore wind power fluctuates much less than the Horns Rev wind farm, first of all because the offshore turbines are concentrated in a very small area where the wind speed fluctuations are much more correlated than the wind speeds at the turbines dispersed over a much bigger area. Another reason which may increase the fluctuations from Horns Rev is that the meteorological conditions are different offshore than onshore.

To some extent, these fluctuations can be handled by the Wind Farm Main Controller, using the power rate limiter in that controller [12]. The power rate limiter can only reduce the power, i.e. produce less than what is available according to e.g. Figure 1. Therefore, the power rate limiter can efficiently limit the positive ramp rate, while the negative ramp rate is more difficult to limit. Still, also the negative power ramping would be significantly reduced with a limited positive ramp rate, because it would cut the power peaks in Figure 1.

Another example of unstable power production is shown in Figure 2. From 13:00 to 14:30, the power ramps fast up and down a couple of times.

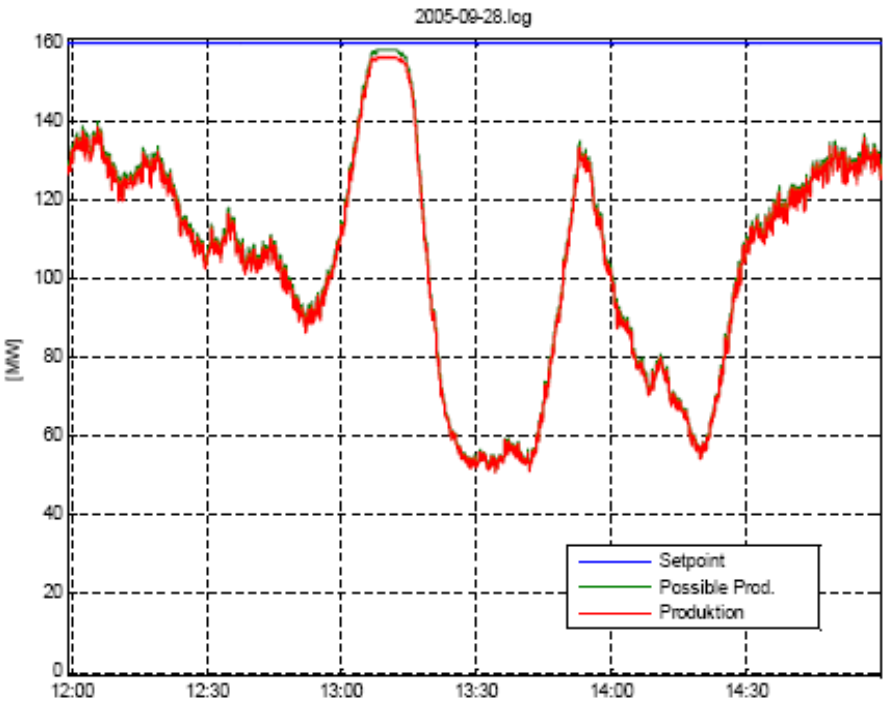

Figure 2. Plot of production at Horns Rev on 28/9-2005.

The webcam image in Figure 3, taken at the end of the period at 14:30, shows that at this time a rain shower (the area marked with red) passes the southern mostly turbines.

Figure 3. Webcam image from 28/9-2005 14:30:05.

The rain shower activity in the period has been confirmed by downloads from Danish Meteorological Institute. It is the impression of the wind farm owner that there is a significant correlation between rain showers and power fluctuations, but this issue has not been analysed systematically in the project.

6 Spatial correlation models

This chapter describes simulation models that use the location and a description of the wind conditions on the sites where the wind turbines are located. The location is used to quantify the correlation between the power fluctuations at different wind turbines, which can be very important for the total power fluctuations as observed in chapter 5.

Two types of wind farm power fluctuation models are described: the detailed model representing each wind turbine, and an aggregated wind model which lumps the wind turbines. The detailed model is validated, based on measurements in Horns Rev and Nysted wind farms.

Finally, an extension of the model to simulate wind power from several wind farms spread over a larger region is described. The extended model is also able to simulate over a longer period, allowing the mean wind speed to vary in time as well as space. The extended model is used to make a case study comparing power fluctuations from 4 closely sited wind farms (Horns Rev) to power fluctuations if two of the wind farms are located in the Djursland – Anholt area.

6.1 Detailed wind farm model

Figure 4 shows a block diagram of the detailed model with individual representation of a number N of wind turbines. The PSD $S_{u[i]}(f)$ of the wind speed in a fixed point in the hub height of each wind turbine i is given as input. For each wind turbine, a rotor wind block calculates the PSD $S_{u_{\mathrm{eq}}[i]}(f)$ of an equivalent wind speed $u_{\mathrm{eq}[i]}(t)$, which includes the smoothing of the wind speed due to the weighted averaging over the rotor. The kernel of the model is the time series simulator, which outputs time series $u_{\mathrm{eq}[i]}(t)$ of the equivalent wind speeds, taking into account the correlation between the wind speeds expressed in the $N \times N$ coherence functions $\boldsymbol{\gamma}_{[r,c]}(f)$. The wind turbine blocks then calculate the produced electrical power $P_{\mathrm{wt}[i]}(t)$. Finally, the power of all wind turbines is summed to provide the wind farm output power $P_{\mathrm{wf}}(t)$.

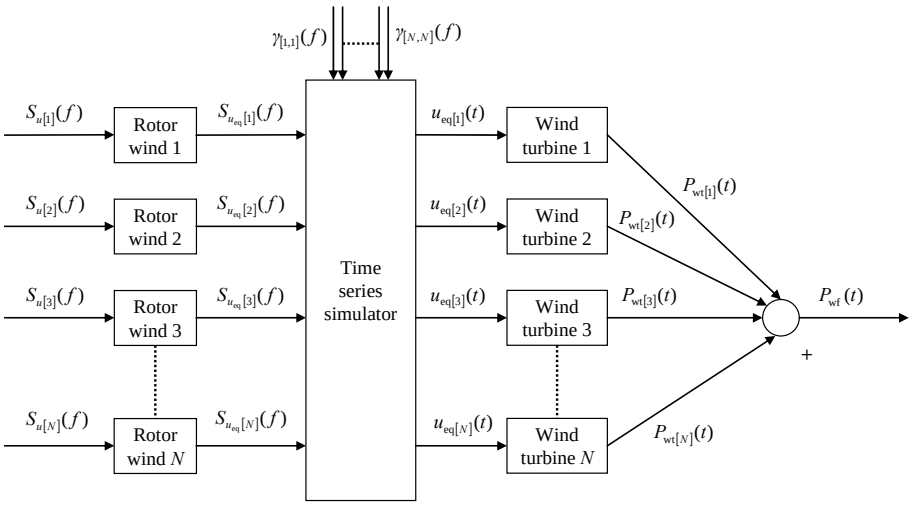

Figure 4. Overall structure of detailed model for simulation of power fluctuations.

The present detailed model is a further development of the wind models for simulation of power fluctuations from wind farms, described by Sørensen et al. in [12]. The model in [12] is focused on simulation of the power quality characteristics defined in IEC 61400-21 [14]. The simulated power quality characteristics are maximum 0.2 sec power, reactive power and flicker emission from the wind farm. The model was verified for that purpose in [15]. The present model has focus on longer term fluctuations as described in the introduction.

The main development of the present model compared to [12] is on the input side, where the PSD and coherence functions applied in the present paper are changed significantly by including empirical results from Risøs test site for large wind turbines in Høvsøre. So far, only temporary results of the Høvsøre analyses have been published [16], but in the present paper, updated analysis results for power spectral densities and coherences will be presented and applied.

Another difference is that the rotor wind blocks in Figure 4 are moved to the frequency domain (i.e. before the time simulator box) in the present model, whereas the model described in [12] applied the time series simulator on $S_{u[i]}(f)$ and therefore had the hub height wind speed $u_{[i]}(t)$ as outputs from the time series simulator.

Finally, the present model simplifies the dynamics because the focus is on fluctuations with one minute period time or longer. Thus, the present model neglects the periodic (3p) components of power fluctuation, which are crucial for simulation of flicker emission, but much faster than the time scale in focus here. Another simplification is that the present model uses a steady state representation of the wind turbine, whereas the wind turbine model in [15] included dynamics of the mechanical transmission system as well as the electrical system.

6.1.1 Power spectral density of wind speed

Analytical expressions for PSD functions of wind speed in a fixed point can be found in the literature and in standards. van Karman [17] proposed a PSD for wind speeds in 1948, and Kaimal [18] proposed another type of PSD in 1972. The international standard IEC 61400-1 [19] for design requirements to wind turbines specifies a Kaimal type PSD function $S_{\mathrm{IEC}}(f)$, which can be used in wind turbine design. The same PSD is chosen for the present work. $S_{\mathrm{IEC}}(f)$ can be expressed as a double sided spectrum according to

$$S_{\mathrm{IEC}}(f) = \sigma^2 \cdot \frac{2 \cdot \frac{L_1}{V_0}}{\left(1 + 6 \cdot \frac{L_1}{V_0} \cdot f\right)^{5/3}},\ L_1 = \begin{cases} 5.67 \cdot z & ,\mathrm{z} \le 60\mathrm{m} \\ 340.2\ \mathrm{m} & ,\mathrm{z} > 60\mathrm{m} \end{cases} \tag{1}$$

V_0 is the average wind speed and σ is the standard deviation of the wind speed in a 10 minutes interval, L_1 is the length scale and z is the height above ground, i.e. the hub height in our case. The literature also

Inside a wind farm, IEC 61400-1 recommends to apply the model proposed by Frandsen [20] to include the added turbulence generated by upwind turbines in the structural design of wind turbines. Frandsen proposes to calculated the turbulence $\sigma = \sigma_{\mathrm{wf}}$ according to

$$\sigma_{\mathrm{wf}} = \frac{1}{2}\left(\sqrt{\left(\frac{0.36 \cdot V_0}{1 + 0.2 \cdot \sqrt{s_1 \cdot s_2 / C_{\mathrm{T}}}}\right)^2 + \sigma_{\mathrm{am}}{}^2} + \sigma_{\mathrm{am}}\right) \tag{2}$$

where σ_{am} is the turbulence in the ambient flow. s_1 and s_2 are the separations between rows and columns respectively in the farm, normalized by the wind turbine rotor diameter. C_{T} is the wind turbine thrust coefficient.

For simulation of power fluctuations, it is chosen to use $\sigma = \sigma_{\mathrm{am}}$ for the wind turbines in the front with expected free inflow, and $\sigma = \sigma_{\mathrm{wf}}$ for the remaining wind turbines inside the wind farm.

According to e.g. Courtney et al. [21] there is a significant variability of the wind speed at lower frequencies, which is not included in the Kaimal spectrum. The Kaimal type PSD functions are valid only for shorter time scales, corresponding to what is normally considered in mechanical design of wind turbines, i.e. from 0.02 sec to 600 sec. For simulations of wind farm power fluctuations, the PSD functions are required on a longer time scale (up to several hours). The Høvsøre measurements have been applied to fit the PSD $S_{\mathrm{LF}}(f)$ for low frequencies, and the fit is expressed as

$$S_{\mathrm{LF}}(f) = (\alpha_{\mathrm{LF}} \cdot V_0 + \beta_{\mathrm{LF}})^2 \frac{\frac{z}{V_0}}{\left(\frac{z \cdot f}{V_0}\right)^{5/3} \cdot \left(1 + 100 \cdot \frac{z \cdot f}{V_0}\right)} \tag{3}$$

Assuming $\beta_{\mathrm{LF}} = 0$, $\alpha_{\mathrm{LF}} = 0.0046$ has been estimated based on Høvsøre measurements.

The input spectra $S_{u[i]}(f)$ in Figure 4 include high frequencies as well as low frequencies according to

$$S_{u[i]}(f) = S_{IEC}(f) + S_{\mathrm{LF}}(f) \tag{4}$$

6.1.2 Coherence function

In [12], the empirical coherence functions suggested by Schlez and Infield [22] were used. However, these expressions are mainly based on measurements with 18 m high masts with distances up to 102 m. Based on the measurements in 80 m height with up to 1.2 km distances in Høvsøre, parameters for the decay factor $A_{[r,c]}$ are fitted to a coherence function $\gamma_{[r,c]}(f)$ between wind turbine number r and wind turbine number c according to

$$\gamma_{[r,c]}(f) = \exp\left(-\left(A_{[r,c]}\frac{d_{[r,c]}}{V_0} + j2\pi\tau_{[r,c]}\right)\cdot f\right) \tag{5}$$

Here, $d_{[r,c]}$ is the distance between wind turbine r to wind turbine c, and $\tau_{[r,c]}$ is the time it takes a wave to travel from wind turbine r to wind turbine c, which depends on the wind speed and the inflow angle $\alpha_{[r,c]}$ of the wind speed to the line between wind turbine r and wind turbine c as defined in [12]. For flow with inflow angles $\alpha_{[r,c]}$ between these points, Schlez and Infield suggest to use the expression

$$A_{[r,c]} = \sqrt{\left(A_{\mathrm{long}}\cos\alpha_{[r,c]}\right)^2 + \left(A_{\mathrm{lat}}\sin\alpha_{[r,c]}\right)^2} \tag{6}$$

A_{long} is the decay factor when the flow is longitudinal, i.e. when the wind direction points directly from wind turbine r to wind turbine c. A_{lat} is the decay factor when the flow is lateral, i.e. when the wind direction perpendicular to the line between wind turbine r and wind turbine c.

Based on the Høvsøre measurements representing horizontal distances between 300 m and 1200 m, the coherence function has been fitted to the decay factors

$$A_{\mathrm{long}} = 4 \tag{7}$$

$$A_{\mathrm{lat}} = \frac{V_0}{2\mathrm{m/s}} \tag{8}$$

6.1.3 Rotor wind model

The idea in the rotor wind model is to generate an equivalent wind speed which can be applied to a simplified aerodynamic model to simulate the aerodynamic torque or power on the wind turbine shaft. The rotor wind block includes the smoothing of the wind speed due to the weighted averaging over the rotor.

Neglecting the periodic components, the rotor block smoothing of wind turbine can be expressed as a wind turbine admittance function $F_{\mathrm{wt}}(f)$ defined as

$$F_{\mathrm{wt}}(f) = \frac{S_{u_{\mathrm{eq}}}(f)}{S_u(f)} \tag{9}$$

Assuming different coherence functions and wind speed weightings in the rotor plane, the admittance function has been solved numerically by Sørensen in [23]. IEC 61400-1 specifies the coherence function $\gamma_{\mathrm{IEC}}(f)$ in the rotor plane according to

$$\gamma_{\mathrm{IEC}}(f) = \exp\left(-12 \cdot \sqrt{\left(\frac{d \cdot f}{V_0} \right)^2 + \left(0.12 \cdot \frac{d}{L_1} \right)^2} \right) \tag{10}$$

Calculating the admittance function numerically with the coherence function in (10), this admittance function is fitted to the analytical expression

$$F_{\mathrm{wt}}(f) = \frac{1}{\left(1 + \left(\sqrt{f^2 + f_1^2} \,/\, f_0 \right)^{4/3} \right)^{3/2}} \;,\; f_0 = \frac{\sqrt{2}}{A} \cdot \frac{V_0}{R} \;,\; f_1 = 0.12 \cdot \frac{V_0}{L_1} \tag{11}$$

Here, R is the radius of the wind turbine rotor disk and A is the coherence decay factor, i.e. $A \cong 12$ using the coherence function $\gamma_{\mathrm{IEC}}(f)$.

6.1.4 Time series simulator

The kernel of the model is the time series simulator, which simulates simultaneous time series $u_{\mathrm{eq}[i]}(t)$ of the N wind speeds indexed i, based on the N PSD's $S_{u_{\mathrm{eq}}[i]}(f)$ of the wind speeds and the $N \times N$ coherence functions $\gamma_{[r,c]}(f)$ between wind speeds indexed r and c.

The applied method is in principle the same as used by Veers [24] to simulate the wind speed field in a wind turbine rotor. However, the coherence function (5) used in the present model is complex as a consequence of the horizontal travel time for wind speed from one wind turbine to another.

The time series simulator has two main parts. The first part calculates Fourier Coefficients of the wind speeds at the individual wind turbines, using the PSD's to ensure the right amplitudes of the Fourier Coefficients and the Coherence function to ensure the right correlation between wind speeds at different wind turbines. The second part uses a (conventional or fast) Inverse Fourier Transform to convert the Fourier Coefficients to time series of wind speeds. A more detailed description of the method was published in [1].

The simulated time series have the specified PSD's and coherence functions. An infinity of solutions fulfil that criteria. Different time series can be obtained by selecting different seed for the random generator that is applied in the simulation.

6.1.5 Wind turbine model

Finally, the wind turbine blocks in Figure 1 are needed to model the conversion of wind fluctuations to power fluctuations. For the present purpose, a steady state model of the wind turbine is applied, based on the power curve of the wind turbine. A more advanced approach including the dynamics of rotor speed and generator power control could be applied if higher frequencies of power fluctuations were in focus. This issue will be discussed further in the analysis part of the paper.

6.2 Aggregated wind farm model

The simulation times of the detailed model in Figure 4 increases almost with the square of the number of wind turbines. The main reason for this is the required calculation time of the time series simulator in Figure 4. Besides, if the power fluctuation model would be applied in a model to simulate the power and frequency control in a power system, the

wind turbines would not be represented individually, but rather by a single or a few aggregated wind turbines. Therefore, an aggregation of the detailed model is suggested.

A simple aggregation is illustrated in Figure 5. For simplicity, the present aggregation is based on the assumption that the rated power of all wind turbines is the same. If this is not the case, the averaging can be replaced by a weighted average. Moreover, the aggregation assumes that the wind turbines are linear, which is a rough approximation when the wind speed is close to rated. Still, this aggregation is a reasonable compromise between accuracy and simulations time.

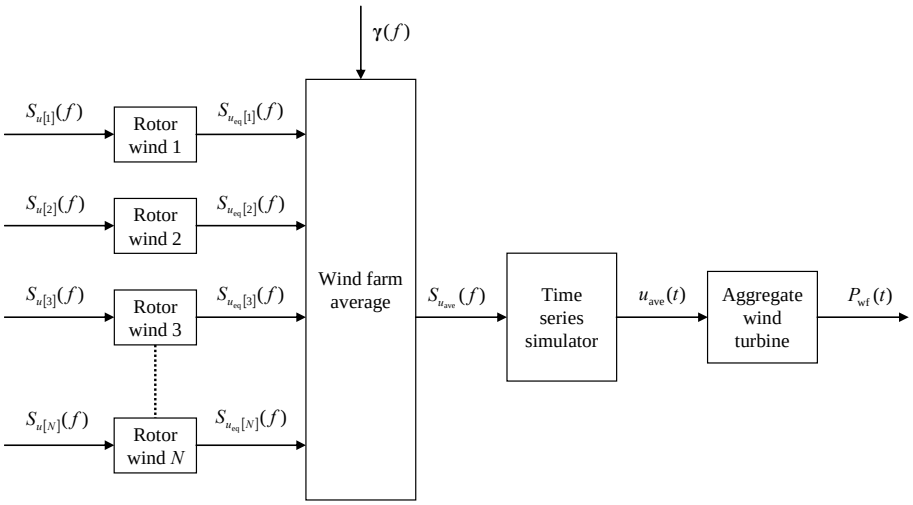

Figure 5. Overall structure of aggregated model for simulation of power fluctuations.

6.2.1 Wind farm average wind speed

The Wind farm average block in Figure 5 calculates the PSD function of the average of N wind speeds, based on the PSD functions of the N wind speeds and the coherence functions between them. In the time domain, the relation is simply

$$u_{ave}(t) = \frac{1}{N}\sum_{i=1}^{N} u_{eq[i]}(t) \tag{12}$$

More details about how the "Wind farm average" can be implemented in the frequency domain is available in [1].

6.2.2 Aggregated wind turbine

To mitigate the error due to the non-linear power curve, Nørgaard and Holttinen [25] proposed an aggregated power curve to represent multiple wind turbines, where the sharp curves of the single wind turbine power curve (i.e. the strong non-linearities) are partly smoothed.

The aggregated power curve proposed by Nørgaard and Holttinen was intended for electricity market modelling, applying large time scales (one hour average values) and geographic areas (200 – 500 km regions). In the wind farm cases studied in the present paper, the distances are much smaller, and therefore using one hour averages, the difference between aggregated and single wind turbine power curves should be less than in [25]. But in the present case, the averaging time used is only one minute, and therefore it could still be relevant to replace the single turbine power curve with an aggregated power curve.

6.3 Validation of detailed wind farm model

The validation is done in terms of ramp rates and reserves requirements, defined for the wind farm power. Duration curves are used to compare measured and simulated power fluctuations.

The definitions of ramp rates and reserve requirements are involving a statistical period time T_{per}, which reflects the time scale of interest. The time scales of interest will depend on the power system size, load behaviour and specific requirements to response times of reserves in the system. In order to study the wind variability and model performance in different time scales, the analysis of data is performed with one minute, ten minute and thirty minute period times.

For the present analysis, the measured wind farm power is calculated in p.u. as the average power of the available turbines in each 2 hour segment. Thus, the reduction of the wind farm power due to non availability of wind turbines is removed. This choice is justified because missing data from a turbine is not necessarily indicating that the turbine is not producing power, but can also be because of failures in the SCADA system.

For each measured 2 hour segment, the average wind speed and wind direction is calculated, and a 2 hour wind farm power time series is simulated. The simulations are performed with the detailed model including low frequency fluctuations and wind farm generated turbulence.

6.3.1 Ramp rates

The definition of ramp rates applied in this paper is quite similar to the definition of load following applied by Parson et. al. [10]. The intention is to quantify the changes in mean values from one period T_{per} to another, which specifies the ramp rate requirement that the wind farm power fluctuation causes to other power plants.

The definition of ramp rates is illustrated for period time T_{per} = 600 s in Figure 6.

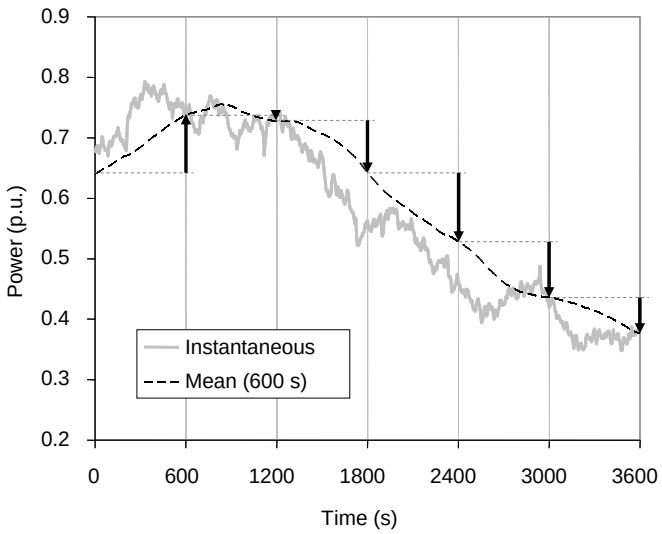

Figure 6. Definition of ramp rated for period time Tper = 10 min. The ramp rates are indicated with arrows.

The instantaneous time series of power can be either measured or simulated, in our case both are with on second time steps. Then the mean value of the power is calculated at the

end of each period, although it is illustrated for all time steps in Figure 6. The ramp rate is simply the change in mean value from one period to the next, i.e.

$$P_{\text{ramp}}(n) = P_{\text{mean}}(n+1) - P_{\text{mean}}(n) \tag{13}$$

Note that this definition specifies the ramping of the wind farm power. Thus, negative ramp rate means decreasing wind power, which requires positive ramping of other power plants.

When the ramp rates have been calculated for each set of neighbour periods n and n+1 for all segments, the ramp rates are binned according the corresponding initial power $P_{\text{mean}}(n)$. This is because the statistics of the ramping will depend strongly on the initial power. For instance, the power is not likely to increase very much when it is already close to rated. A power bins size 0.1 p.u. has been selected.

The ramping is sorted in each power bin, and a duration curve is obtained. This is done for the measurements and for the simulations. As an example, the duration curves for ten minute ramp rates in the initial power range between 0.8 p.u. and 0.9 p.u. is shown in Figure 7. There is a good agreement between the simulated and the measured duration curves.

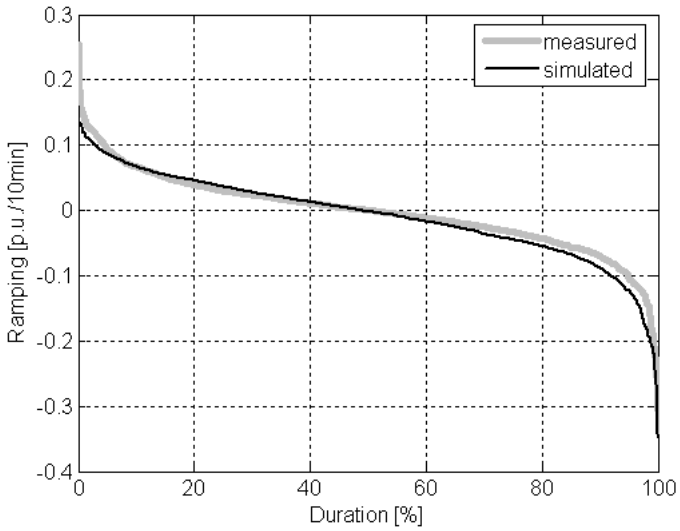

Figure 7.Duration curves of 10 minutes ramp rates in the initial power range from 0.8 p.u. to 0.9 p.u..

The most interesting point of the duration curves is the highest wind farm negative ramp rate, i.e. around 100% on the duration curve, because this quantifies the highest requirement to the ramp rates of other power plants. The wind farm positive ramp rates are not so interesting here because they can be limited directly by the wind farm main controller.

In Figure 8, the 99% percentile of the 10 minutes ramp rates duration curve for all power ranges is shown. In order to assess the model performances, both Horns Rev and Nysted simulated and measured ramp rates are plotted.

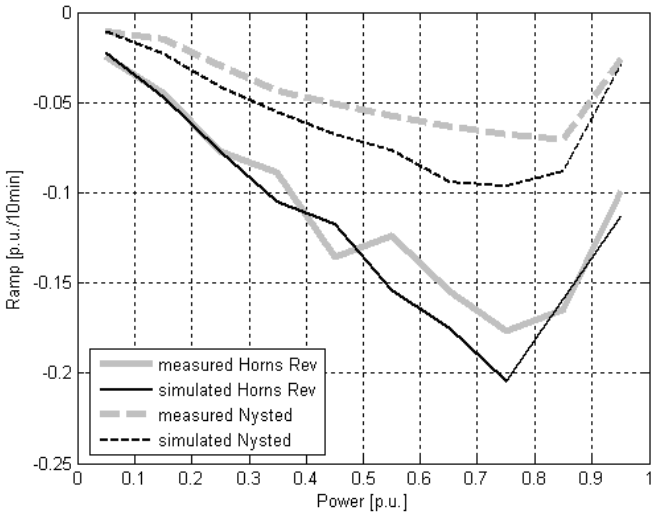

Figure 8. The 99% percentiles of 10 minutes ramp rates in all power ranges for Horns Rev and Nysted wind farm.

There is an acceptably good match, with differences generally less than 0.03 p.u./10 min for both Horns Rev and Nysted, although Nysted simulated ramp rates are systematically bigger the measured ones.

The 99% percentile for 1 minute and 30 minutes period is shown in Figure 9 and Figure 10, respectively. The match between simulated and measured power fluctuations is similar to the 10 minutes periods, with the simulated power fluctuations still being systematically bigger than the measured ones. For the 1 minute period, Although the difference between simulated and measured looks significant, it is less than 0.01 p.u./1min.

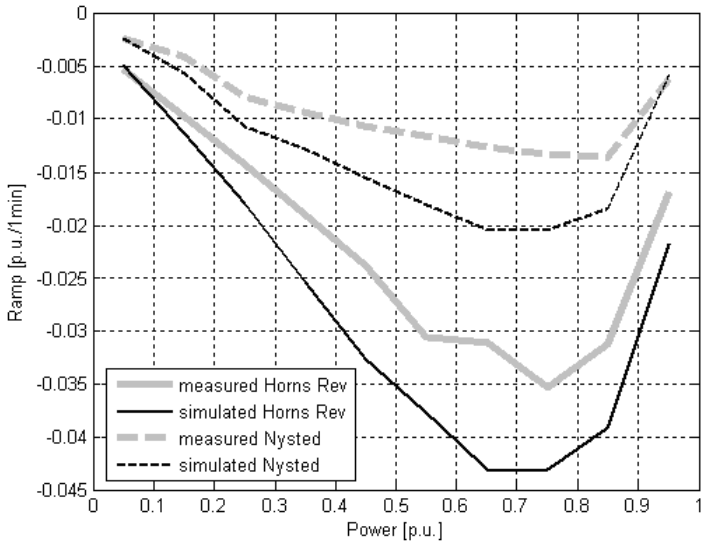

Figure 9. The 99% percentiles of 1 minute ramp rates in all power ranges for Horns Rev and Nysted wind farms.

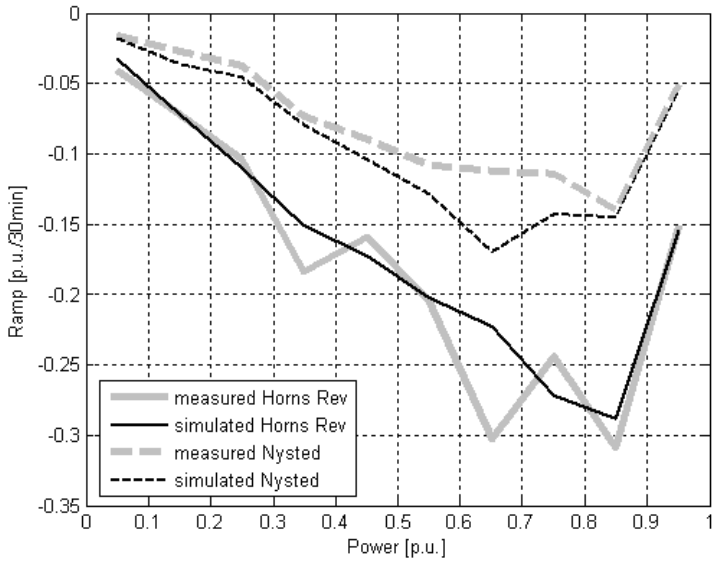

Figure 10. The 99% percentiles of 30 minutes ramp rates in all power ranges for Horns Rev and Nysted wind farms.

Summarizing, the ramping analyses, for Nysted wind farm, show a good match between simulated and measured time series.

6.3.2 Reserve requirements

The intention is to quantify the difference between the instantaneous power and the mean value which are dealt with as ramping. Since the reserves must be allocated in advance, the positive reserve requirement is defined as the difference between the initial mean value and the minimum value in the next period.

This definition of reserve requirements is illustrated for period time T_{per} = 600 s in Figure 11. Formally, the reserve requirements are defined as

$$P_{\text{res}}(n) = P_{\text{mean}}(n) - P_{\text{min}}(n+1) \tag{14}$$

Note that with this definition, positive reserves means decreasing wind power that requires positive reserves form other power plants.

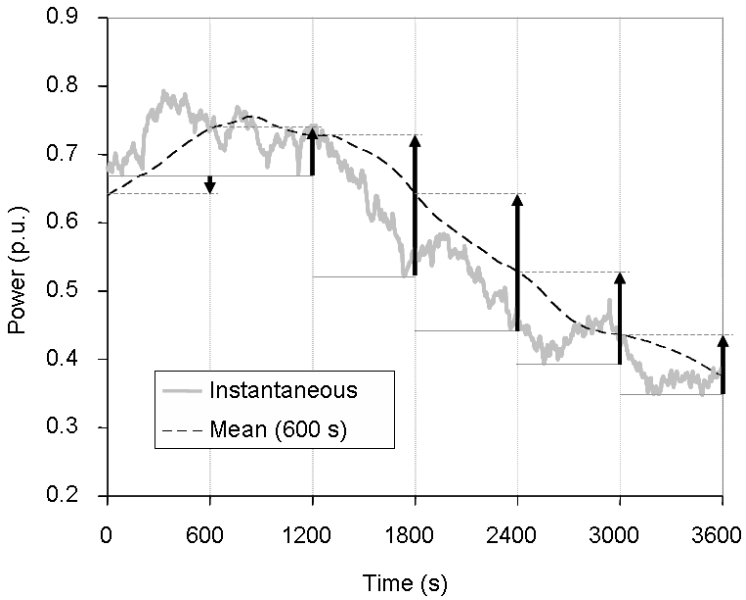

Figure 11. Definition of reserve requirements for period time Tper = 10 min. The reserves are indicated with arrows

The duration curves are then calculated for the reserves in all power bins, and the 1 % percentiles are found in each power bin.

The resulting 1 % percentiles with 10 min, 30 min and 1 min period times are shown in Figure 12, Figure 13 and Figure 14, respectively.

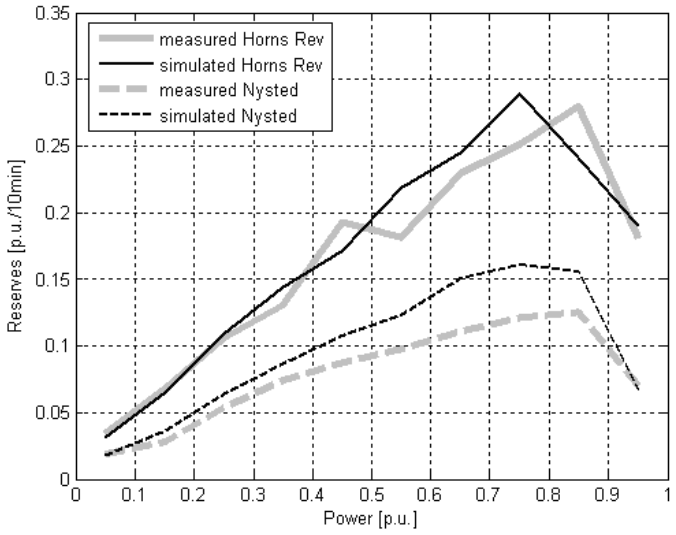

Figure 12. 1 % percentiles of 10 minutes reserve requirements in Horns Rev and Nysted wind farm

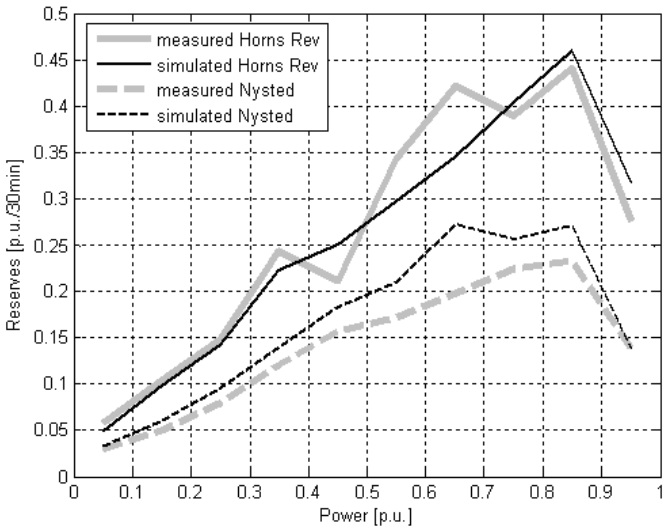

Figure 13. 1 % percentiles of 30 minutes reserve requirements in Horns Rev and Nysted wind farm

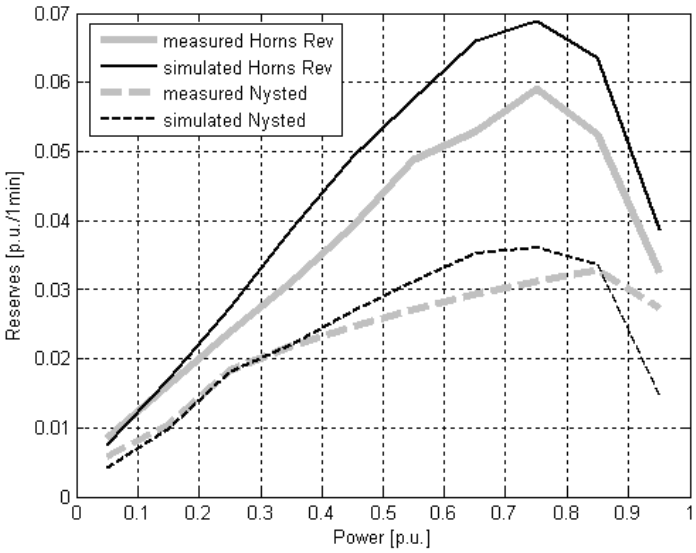

Figure 14. 1 % percentiles of 1 minutes reserve requirements in Horns Rev and Nysted wind farm

It is concluded that the simulated time series agree quite well with the measured in the prediction of the reserves. As expected, the reserve requirement decreases in the highest power bin because the wind speed is above rated.

As the ramp rates, the reserve requirements in Nysted wind farm are smaller than ones in Horns Rev.

6.4 Extension to power system region

The above described and validated models were developed to simulate the power fluctuations from a single wind farm, and the length of the simulated time series were limited to two hours. This section describes an extension of the simulation model, which

now enables the model to simulate wind power fluctuations in a large power system region, and with unlimited time series length.

For a transmission system operator, it is off course important to know how much the power from a large wind farm is fluctuating. As mentioned in the introduction, Energinet.dk became interested in this issue based on the experience with Horns Rev, the first large offshore wind farm. Now, a new large wind farm (Horns Rev 2) is under construction, and two more are on the way in West Denmark. In this situation, the concern is not only the power fluctuation from the single wind farms, but rather the sum of power fluctuations from all the wind farms.

Energinet.dk's System Plan 2007 [26] pointed at geographical spreading of the future offshore wind farms as a mean to meet the challenges of future wind power integration. A commission under the Danish Energy Authority has issued "Future Offshore Wind Farm Sites – 2025" [27], appointing a number of potential sites for large offshore wind farms, taking into account water depth, wind resources, grid connection and other important issues. The extension of the simulation model provides a new tool to simulate the wind power fluctuations assuming different sites.

6.4.1 Modification of time series simulator

The mathematical reason for the limitations in the original models is that the mean wind speed and wind direction was assumed to be constant during the simulation period. With this assumption, the Fourier Coefficients of the wind speeds at the individual wind turbines would be constant during the (up to 2 hour) simulation period, and the succeeding Inverse Fourier Transform (see paragraph 6.1.4) can be done by a standard (conventional or fast) Inverse Fourier Transform.

With variable mean wind speeds and directions, the Fourier Coefficients will also become variable in time. To handle this, the modified time series simulator operates with two time steps. The primary time step is the original one (typically 1-600 s) defining the time steps of the simulated time series. The secondary time step (typically 1 hour) is the resolution of the mean value time series.

The modified time series simulator calculates the Fourier Coefficients for each secondary time step. This is a computer time consuming task, especially with many wind turbines. Still, significant modelling and implementation effort has been made to reduce this calculation time. With this effort, it has become possible to simulate quite many individual wind turbine wind speeds simultaneously.

As an example, 331 wind turbines were simulated 4 days with 1 minute time steps, and the total computer simulation time was 28 minutes. With the original model, it took 30 s to simulate 80 wind turbines in 1 hour, so it is estimated that it would have taken 13-14 hours with the original model. If the number of wind turbines is increased further, then the benefit would be even higher, because the calculation time increases with approximately number of wind turbines for the modified method, while it increased with approximately the square of the number of wind turbines for the modified method.

With time varying Fourier Coefficients, the Inverse Fourier Transform also needs to be modified. First, the Fourier Coefficients for the individual primary time step are found by interpolation between the calculated values at the secondary time steps, and subsequently, the Inverse Fourier Transform is done for the individual primary time step. This is also much more computer time consuming than the original Inverse Fourier Transform.

It is planned to describe the modified time series simulator in more details in a future peer-reviewed paper. That description will include the applied approach to reduce the computer calculation time.

6.4.2 Mean values from climate model

According to the above description, the mean wind speed time series must be given as input to the simulation with the secondary time step (1 hour) resolution. This is obtained by available simulations with macro scale weather models.

A number of weather model data series have been considered. Figure 15 shows the available weather model results that have been considered. The NCEP/NCAR and ERA data is world wide, but has a very high resolution. Max Plank Institute (MPI) in Germany has provided Risø DTU with data for Germany (including southern part of Denmark) with high resolution (10 km), and data for all Europe with 50 km resolution.

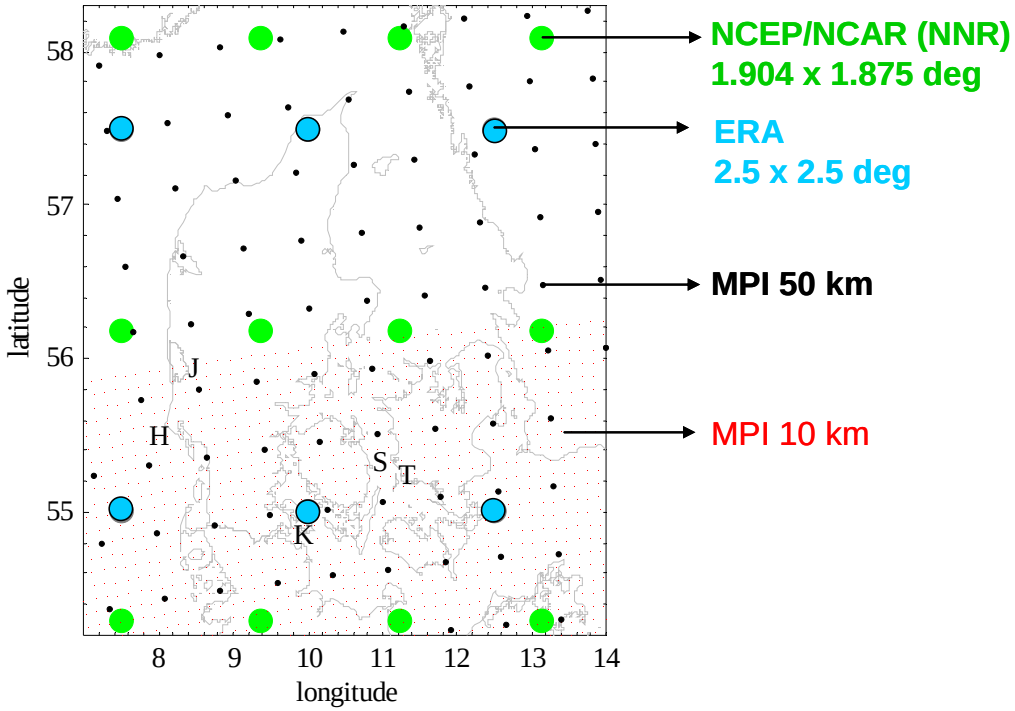

Figure 15. Available weather model results

In order to include all Denmark with the highest possible resolution, the MPI 50 km data has been selected. This data includes 25 years (1979 – 2003) of historical data with 1 hour resolution.

6.4.3 Simulation case

A simulation case has been specified to illustrate the relevance of geographical spreading of the future offshore wind farms as a mean to mitigate the impact of wind farm cut-outs due to a storm front passage.

Figure 16 shows the 6 wind farms that are simulated. The first wind farm is the existing wind farm in Horns Rev. The second wind farm is the Horns Rev 2, which is expected to be commissioned by the end of 2009. The last 4 wind farms are selected possible positions from the Danish Energy Authority report [27].

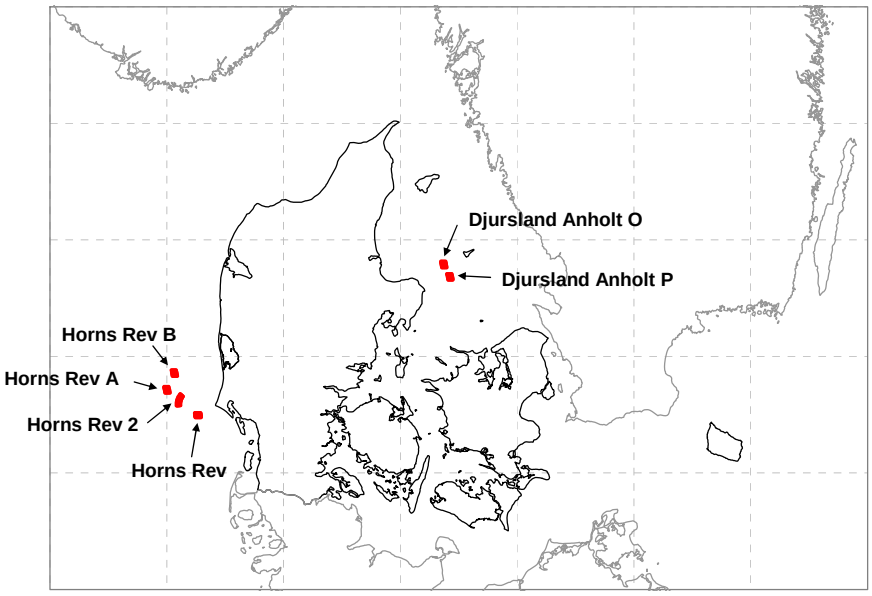

Figure 16. Considered offshore wind farms in Danmark

Data for Horns Rev and for Horns Rev 2 is given in details, including the positions sizes and positions of the individual wind turbines. The wind turbine size and relative positions have some influence on the simulation result, but the main most important parameter is the total power and the geographical position of the wind farm (Figure 16). Another assumption is the annual mean wind speed, which is used to calibrate the MPI weather model date. The calibration is necessary, because the weather model data represents an average area of 50 km × 50 km in 10 m height. The main wind farm data is summarised in Table 1

Table 1. Data for the 6 simulated wind farms.)the assumed annual mean wind speeds are not justified*

Name	Symbol	Wind turbine power	Total power	Annual mean wind speed
Horns Rev	HR1	80 × 2.0 MW	160 MW	9.6 m/s*)
Horns Rev 2	HR2	91 × 2.3 MW	209 MW	10.4 m/s*)
Horns Rev A	HRA	40 × 5.0 MW	200 MW	10.6 m/s*)
Horns Rev B	HRB	40 × 5.0 MW	200 MW	10.5 m/s*)
Djursland Anholt O	DAO	40 × 5.0 MW	200 MW	9.0 m/s*)
Djursland Anholt P	DAP	40 × 5.0 MW	200 MW	9.0 m/s*)

All 6 wind farms are now simulated for 96 hours from 28/1-2000 to 31/1-2000. This period is interesting, because it includes the passage of a storm front. All the wind farms are simulated in one run to ensure the right correlation between the wind speeds. The primary time step for the simulations is selected to 1 minute.

The idea is now to compare two scenarios:

- The concentrated scenario: Horns Rev and Horns Rev 2 are supplemented with 2 new wind farms Horns Rev A and Horns Rev B.
- The spread scenario: Horns Rev and Horns Rev 2 are supplemented with 2 new wind farms Djursland Anholt O and Djursland Anholt P.

The concentrated scenario is clearly beneficial from the point of view of annual energy production, because the annual mean wind speed is significantly higher in the Horns Rev area than in the Djursland-Anholt area. However, from the point of view of wind power fluctuations, the concentrated scenario will provide faster and larger variations. The simulation tool is used to quantify this difference.

6.4.4 Simulation results

Figure 17 shows the individual wind farm averages of the 331 simulated wind speeds. It is clear from the graph that the wind speeds at the wind farms are much more correlated if they are close than if they are separate in distance. It is also observed that there is some delay of the variations at Djursland Anholt compared to Horns Rev. This is because the wind direction is western, which is typical, especially for storms. Finally, the graph also indicates that the wind speeds are higher at Horns Rev than at Djursland – Als. In this context, it should be noticed that the wind speed exceeds the 25 m/s cut-out wind speed for a long period in Horns Rev, but only for a short period in Djursland – Anholt.

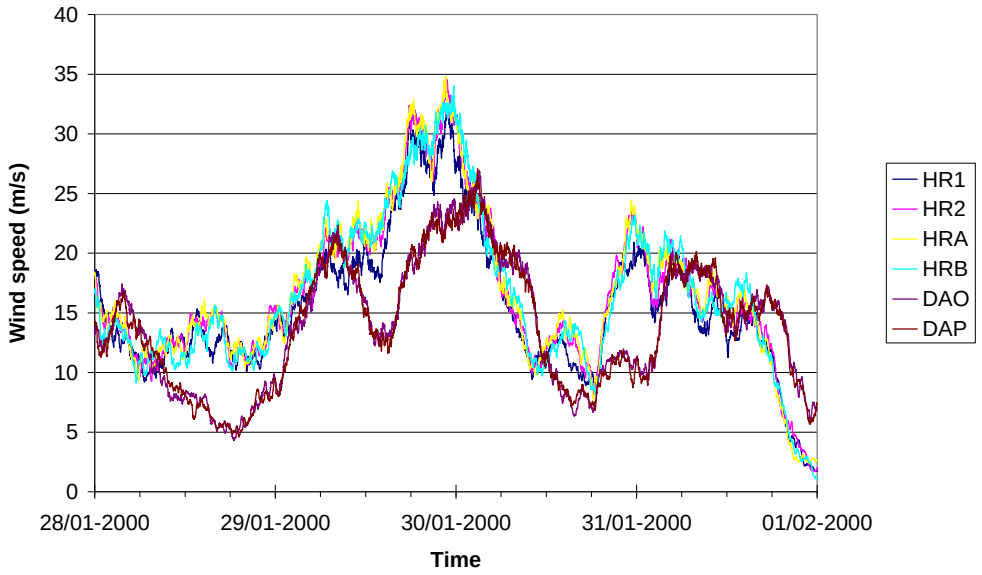

Figure 17. Wind farm averages of simulated wind speeds

Looking at the cut-out of wind turbines, it is important to notice that the wind speeds at the individual wind turbines fluctuate more than the average wind farm wind speed. This is shown in Figure 17, where the wind speed of turbine A1 is compared to the average wind speed in Horns Rev (HR1) wind farm. The figure also shows the weather model mean wind speed, which was used as input for the simulation. It is clearly seen how the simulation model adds faster fluctuations to the weather model date, but keeps the long term mean value trend.

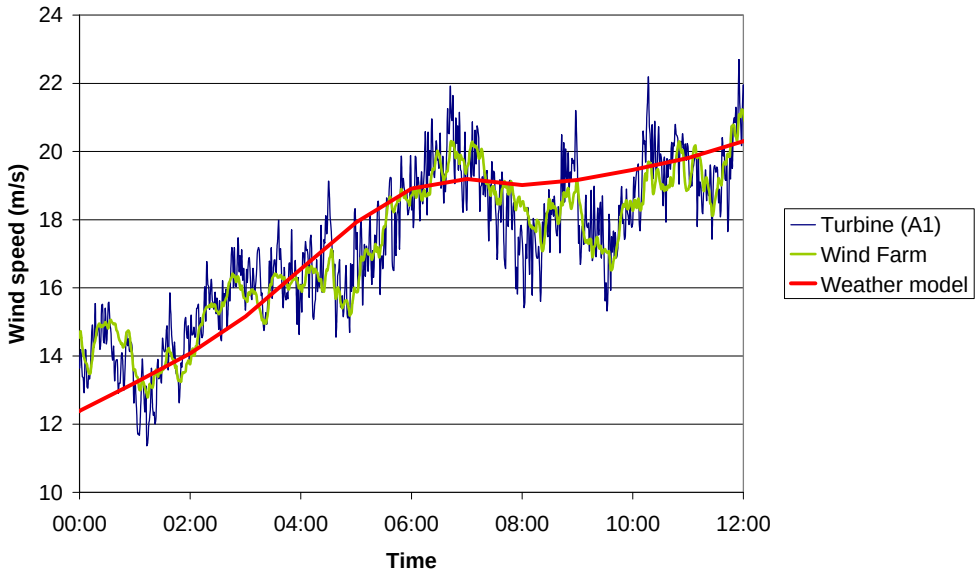

Figure 18. Comparison of variability of a single wind turbine wind speed, the average wind speed in the wind farm, and the weather model input mean value.

Another important observation is the impact that the fluctuations have on the power curves. Figure 19 shows the wind farm power vs. wind farm average wind speed, compared to the input power curve of the individual wind turbines. The curves are quite close up to cut-out wind speed, but then it is seen that if the average wind speed is close to 25 m/s, then some wind turbines are connected and other wind turbines are disconnected, because some individual wind speeds are below and others are above 25 m/s. The figure also could indicate that the cut-out can be simulated with reasonable accuracy using an aggregated wind turbine model with the modified power curve. However, it should be noticed that the slope of the wind farm power curve will depend on wind speed ramp rate, faster wind fluctuations, and of the size of the wind farm.

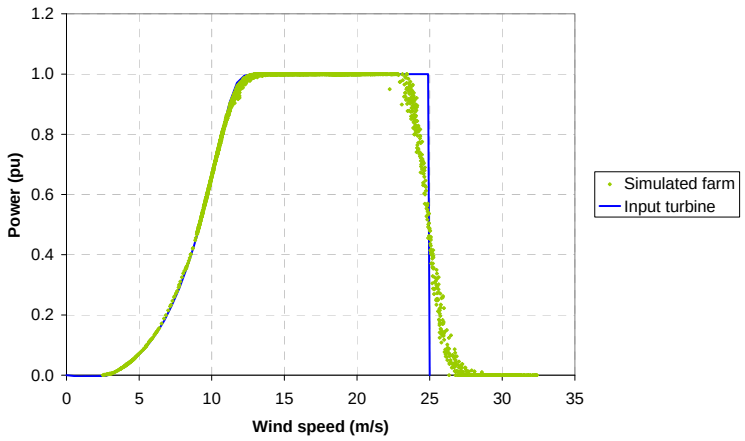

Figure 19.Simulated wind farm power curve compared to input power curve for wind turbine.

Finally, the simulated wind power is illustrated for the two scenarios in Figure 20 and Figure 21 respectively. It is clearly illustrated how the storm front hits all 4 wind farms

almost simultaneously and with full power in the afternoon 29/1 with the concentrated scenario, while the front is delayed and has lost intensity when it hits Djursland Anholt wind farms with the spread scenario.

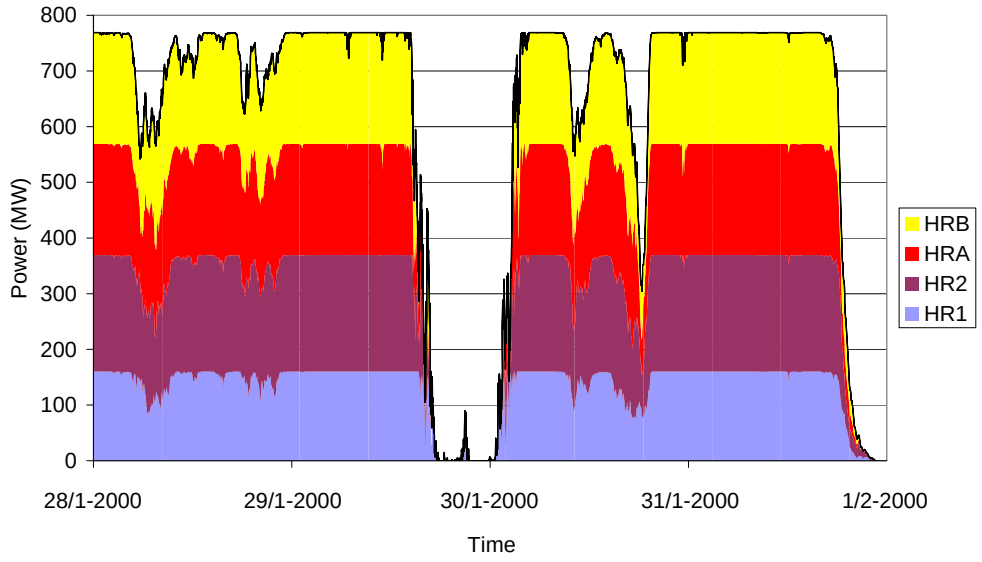

Figure 20. Simulated wind power fluctuations in the concentrated scenario.

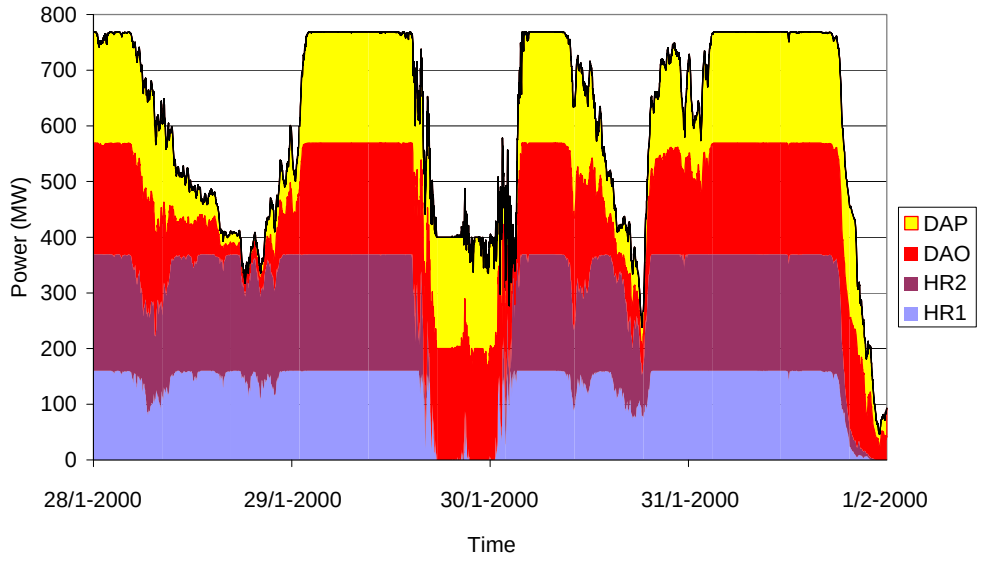

Figure 21. Simulated wind power fluctuations in the spread scenario

The above simulation results are only one example, but they illustrate quite well the possibility to use the simulation tool

6.5 Conclusion

The simulation model predicts the statistical behaviour of the power time series quite well. Duration curves for ramp rates and reserve requirements are generally simulated to

slightly higher values than measured. The extreme values though obtained from simulations agree well with the measured ones.

The extended model provides the opportunity to compare the power fluctuations with different scenarios for the siting of future wind farms. A single case study illustrates that the siting of large offshore wind farms is very important for the power fluctuations, especially in the case of a storm passage.

The spatial correlation model developed in this project is based on the PARKSIMU model which was originally developed with focus on local and short time scale (up to 10 minutes) power quality issues such as flicker emission from smaller wind farms. It was first developed to simulate large wind farms in longer time scales (up to 2 hours), and later to simulate wind power fluctuations in a larger power system region over several days. Therefore, the software has changed name to WPTS (Wind Power Time Series simulator).

In principle, WPTS could be applied to even larger areas like the European power system. However, that would require substantial aggregation of the wind turbines. An aggregated model for a single wind turbine has already been developed, but to include a wind farm aggregate in a model that takes into account the correlation between aggregates is another issue. However, the model developer has a clear idea as to how this should be done.

Finally, the here presented simulations used calibration of the weather model data with a constant scaling factor that gave the right annual mean wind speed. In order to take into account the effect of different terrains with different direction dependencies, it would be more correct to do the calibration in individual wind direction sectors. Such a calibration could be base on WAsP wind resource calculations.

7 Regime-switching models

Large offshore wind farms concentrate a high wind power capacity at a single location. Onshore, the same level of installed capacity is usually spread over an area of significant size, which yields a smoothing of power fluctuations [28]. This spatial smoothing effect is hardly present offshore, and thus the magnitude of power fluctuations may reach very significant levels.

Modelling the power fluctuations for the specific case of offshore wind farms is a current challenge [29], for better forecasting offshore wind generation, developing control strategies, or alternatively for simulating the combination of wind generation with storage. This chapter investigates the applicability and performance of some statistical models.

Operators of offshore wind farms often observe abrupt changes in power production. The fast variations can be related to the turbulent nature of the wind. They are smoothed out when considering the cumulative production for the wind farm, since turbines are spread over a pretty large area. In addition, when inspecting power production data averaged at a few minutes rate, one observes variations that are due to slower local atmospheric changes e.g. frontline passages and rain showers. These meteorological phenomena add complexity to the modelling of wind power production, which is already non-linear and bounded owing to the characteristics of the wind-to-power conversion function. Such succession of periods with power fluctuations of lower and larger magnitudes calls for the use of regime-switching models.

In the frame of the project, focus has been given to the application of Self-Exciting Threshold AutoRegressive (SETAR) model, the Smooth Transition AutoRegressive (STAR) model, as well as the Markov-Switching AutoRegressive (MSAR) model for that purpose. Their performance are evaluated on a one-step ahead forecasting exercise, and compared to those of linear models, i.e. AutoRegressive Moving Average (ARMA) models. The available data consist in time-series of power production averaged at a 1, 5, and 10 minutes rate, for the Horns Rev and Nysted wind farms. These models and the results from their application to the wind power production data at these two offshore wind farms are presented in Section 7.1 and 7.2, respectively.

Then, since it is found that the Markov-switching approach outperforms the others, it has been decided to focus on this class of models, and to enhance them so they account for some specific characteristics of the wind generation process e.g. nonstationarity. The developments towards MSAR models with time-varying coefficients are described in Section 7.3. The results corresponding to the application of such models for forecasting short-term wind power fluctuations at Horns Rev and Nysted are shown in Section 7.4. Finally, Section 7.5 gathers the conclusions from these research works and gives perspectives on future developments.

7.1 From linear to regime-switching models

Generated wind power is considered hereafter as a stochastic process for which statistical models are set up in order to describe its temporal evolution. The notation y_t is used for denoting both the state of the stochastic process at time t and the measured value at that time. All the measured power values over the considered period are gathered in the time-series $\{y_t\}$, $t=1,\ldots,T$, where T is the total number of successive observations. The set Ω_t = $(y_1,y_2,\ldots,y\_t)$ that contains all the observations up to time t, is referred to as the information set. Our framework is that of univariate time-series modelling, i.e. no explanatory variable is used.

The well-known linear ARMA model is used as a benchmark. It also serves as a basis for constructing the regime-switching models. The term `regime' originates from the assumption such that the considered stochastic process switches between a finite number of distinct (and most often linear) models. Denote by r the number of these regimes. The SETAR and STAR models are consequently presented. For these two families of models, the switches from one regime to the other are governed by an observable process, i.e. by some function of lagged values of $\{y_t\}$.

7.1.1 The baseline ARMA model

The linear ARMA(p,q) model encompasses an autoregressive (AR) part of order p and a moving average (MA) part of order q,

$$y_t = \theta_0 + \sum_{i=0}^{p} \theta_i y_{t-i} + \sum_{j=0}^{q} \varphi_j \varepsilon_{t-j} + \varepsilon_t \qquad (15)$$

where ε_t is a white noise process, i.e. a purely random process with zero mean and finite variance. For the theory related to ARMA models we refer to eg. [30]. The model parameters can be determined for instance with Maximum Likelihood (ML) estimation and with a Gaussian assumption on the distribution of residuals.

7.1.2 The SETAR model

A Self-Exciting Threshold AutoRegressive (SETAR) model is a piecewise linear model with an AR part for each of the *r* regimes, and for which the current regime is determined by a function of lagged values of the time-series [31]. This may yield abrupt switches from one regime to the other. Threshold values r_k, $k=1,2,\ldots,r-1$ define the intervals on which the various AR parts are active.

The SETAR$(r;p_1,p_2,\ldots,p_r)$ model is given by

$$y_t = \theta_0^{(s_t)} + \sum_{i=0}^{p_{m_t}} \theta_i^{(s_t)} y_{t-i} + \sigma_{s_t} \varepsilon_t \tag{16}$$

where for a given regime *k*, p_k and σ_k denote the order of the AR model and the related variance of the noise sequence. ε_t is a white noise process with unit variance, such that ε_t is independent of Ω_{t-1}. $\{s_t\}$ is the sequence of regimes, taking values in {1,2,…,r}, for which each s_t is defined by

$$s_t = \begin{cases} 1, & if\ y_{t-1} \in]-\infty, r_1] \\ 2, & if\ y_{t-1} \in]r_1, r_2] \\ \vdots & \\ r, & if\ y_{t-1} \in]r_{r-1}, \infty[\end{cases} \tag{17}$$

The parameters of the SETAR model are estimated with the Minimum Mean Square Error (MMSE) estimation method, as described in [2].

7.1.3 The STAR model

Often, abrupt changes between regimes are not satisfactory, even though separate regimes have been clearly identified. Smooth Transition AutoRegressive (STAR) models have been introduced in the literature in order to feature smooth (and controllable) transitions between regimes [32]. The Smooth Transition Bilinear (STBL) model, which belongs to the family of STAR models, has already been successfully applied for describing wind speed variations [33]. Focus is given here to the multiple-regime STAR (that we will, for convenience, refer to as STAR only), for which the value of the considered stochastic process $\{y_t\}$ at time *t* is given as a weighted average of several AR parts. The weights assigned to the AR parts are a function of the previous process value y_{t-1}. These past process values are sorted in bins as it is done for SETAR model. However, a transition function controls here the transition from one regime to the other. The choice of the transition function depends on which type of behaviour is to be modelled. The two most popular transition functions are the exponential and logistic ones. The latter is chosen here, since it permits to more clearly separate the different regimes. For a detailed description of the STAR models employed and estimation of their parameters, we refer to [2].

7.1.4 A regime-switching model governed by a hidden Markov chain

The models described above rely on an observable process for determining the actual regime, which is determined as a function of past values of the process. Markov

Switching AutoRegressive (MSAR) models propose an alternative to this observable regime-switching modelling, by allowing the switches to be governed by an unobservable process. It is assumed to be a Markov chain. A nice feature of such approach is that it permits to reflect the impact of some external factors on the behaviour of certain time-series [34]. Indeed, it has been found particularly suitable for modelling the temporal evolution of weather variables, such as daily rainfall occurrences [35] or wind fields [36] especially because it manages to capture the influence of some complex meteorological features eg. related to the motion of large meteorological structures. For the specific case of the fluctuations of offshore wind generation, our aim is to use this hidden Markov chain for describing meteorological features governing the regimes that cannot be determined from past values of measured power production only. MSAR models and the estimation of their parameters are briefly presented here. An extended description is available in [2].

MSAR models resemble SETAR models in their formulation. If considering r regimes and AR models of orders $p_1,p_2,\ldots,p_r$ for each of these regimes, the corresponding MSAR$(r;p_1,p_2,\ldots,p_r)$ model is indeed given by

$$y_t = \theta_0^{(s_t)} + \sum_{i=0}^{p_{m_t}} \theta_i^{(s_t)} y_{t-i} + \sigma_{s_t} \varepsilon_t \tag{18}$$

where ε_t is a white noise process with unit variance, σ_k the standard deviation of the noise sequence in the k^{th} regime, and $\{s_t\}$ the regime sequence. Even though the regime sequence for MSAR models is unobservable, it is assumed that $\{s_t\}$ follows a first order Markov chain on the finite space $\{1,\ldots,r\}$: the regime at time t is determined from the regime at time $t-1$ only, in a probabilistic way

$$P(s_t=j \mid s_{t-1} = i, s_{t-2}, \ldots, s_0) = P(s_t = j \mid s_{t-1} = i) \tag{19}$$

All the probabilities governing the switches from one regime to the other are gathered in the so-called transition matrix P, for which the element p_{ij} represents the probability of being in regime j given that the previous regime was i. P is such that: *(i)* all the elements on a given row sum to 1 since the r regimes represent all regimes that can be reached at any time; *(ii)* all p_{ij} are positive in order to ensure ergodicity, which means that any regime can be reached eventually.

Estimating the parameters of MSAR models is more complicated than for the case of the regime-switching models with observable regime sequences. The method employed is based on the Expectation Maximization (EM) and is extensively described in [2].

7.2 Results from offshore case studies

The models presented above are used for describing the fluctuations of offshore wind generation on two real-world case studies. The exercise consists in one-step ahead forecasting of time-series of wind power production. The data for these two offshore wind farms are firstly described. Then, the configuration of the various models and the setup of estimation methods are given. Finally, a collection of results is shown and commented.

7.2.1 Case studies

The raw power data consist in one-second measurements for each wind turbine. It has been chosen to model each wind farm as a single wind turbine, the production of which consists in the average of the power generated by all the available wind turbines. These turbines are of nominal capacity 2000 kW and 2300 kW for Horns Rev and Nysted, respectively.

A sampling procedure has been developed for obtaining time-series of 1, 5, and 10 minutes power averages. These sampling rates are selected so that the very fast fluctuations related to the turbulent nature of the wind disappear and reveal slower fluctuations at the minute scale. Because there may be some erroneous or suspicious data in the raw measurements, it has been considered that a minimum of 75% of the raw measurement within a time interval needed to be available so that the related average is seen as valid.

At Horns Rev, the available raw data are from 16th February 2005 to 25th January 2006. And, for Nysted, these data have been gathered for the period ranging from 1st January to 30th September 2005. The time-series have been split into learning and testing sets. The former serves for the fitting of statistical models while the latter allow us to appraise what the performance of these models may be in operational conditions. Sufficiently long periods without any invalid data are identified in order to define the necessary datasets. For Horns Rev, the training set relates to September 2005. The testing set is composed by 19 periods whose lengths are between 2 and 16 days, identified in the remaining of the whole dataset.

Regarding Nysted, the training set corresponds to the period from the 15th February 2005 to the 9th March 2005, while the test set gathers 14 periods of length 6-27 days from the rest of the available data.

7.2.2 Models, estimation setup and evaluation criteria

The various time-series are modelled with the linear ARMA and regime-switching SETAR, STAR and MSAR models. The order of the AR and MA parts are chosen to vary between 1 and 5. This yields 25 competing ARMA models.

The optimal threshold values for SETAR and STAR models are determined from the nonlinear optimization procedures described in [2]. We impose the number of regimes to be $r=3$. Our choice for 3 regimes is motivated by the influence of the wind-to-power conversion function on the variance of wind generation: this variance is lower in the low and high power range, while it is much larger in the steep slope part of the power curve [37]. Thresholds are initialized by considering various combinations of lower and higher threshold values, which yields 9 combinations of initial threshold values for each wind farm. For the particular case of STAR models, we fix the shape of the logistic functions by setting the slope parameter to 1. There are then 1125 competing models in each of the SETAR and STAR model families.

The initialization of the EM algorithm for the case of MSAR models consists in picking an initial transition matrix P, as well as initial AR parts, by specifying their parameters and their variances, such that the resulting MSAR model is stationary. The way MSAR models are initialized is further described in [2]. The final number of competing models in this MSAR class reaches 1250.

Either with the ML estimation method and a Gaussian assumption on the residual distributions, or with the Minimum Mean Square Error (MMSE) estimation method, the parameters of the models are determined with the aim of minimizing a quadratic error criterion. Therefore, in order to be consistent with the way parameters are estimated, models are also evaluated with a quadratic criterion on the testing set. More precisely, from the large panel of error measures available for evaluating wind power predictions [38], the Root Mean Square Error (RMSE) criterion is chosen.

7.2.3 Results and discussion

Table 2 lists the best models of each class - best in terms of a minimum RMSE on the testing set - for the time-series related to the Nysted wind farm. For instance for the 1 minute averaged data, the best of the 25 competing ARMA models has been found to be the ARMA(5,4). In addition, models are ranked from minimum to maximum RMSE. Table 2 also gives the characteristics of the optimal SETAR and STAR models, i.e. the thresholds that were determined from the optimization procedure. Note that for the 5 and 10 minutes averaged data, the thresholds related to the lower regime for the SETAR models are very low (equal to 2.2 and 6 kW, respectively), thus isolating the no-production cases as a regime itself.

Table 2. Performance evaluation for the various models for Nysted. Results are for the 3 time-series averaged at different rates. The left column gives the optimal model of each class. The optimal threshold values for SETAR models (r) and STAR models (c) are also given. The models are ranked as a function of their RMSE on the testing set.

(a) 1-minute averaged data

Model	RMSE [kW]	r [kW]	c [kW]
MSAR(3; 4, 4, 3)	13.3	–	–
STAR(3; 5, 5, 5)	16.1	–	(920.9, 2096.6)
SETAR(3; 4, 4, 4)	16.5	(203.6, 2006.3)	–
ARMA(5, 4)	16.5	–	–

(b) 5-minute averaged data

Model	RMSE [kW]	r [kW]	c [kW]
MSAR(3; 4, 5, 5)	35.5	–	–
STAR(3; 5, 5, 5)	48.9	–	(827.3, 1638.7)
ARMA(4, 5)	50.8	–	–
SETAR(3; 1, 3, 3)	50.9	(2.2, 2149.8)	–

(c) 10-minute averaged data

Model	RMSE [kW]	r [kW]	c [kW]
MSAR(3; 2, 4, 4)	60.6	–	–
STAR(3; 5, 5, 5)	86.2	–	(579.4, 1545.4)
SETAR(3; 3, 5, 5)	88.6	(6.0, 1595.4)	–
ARMA(5, 1)	88.9	–	–

Whatever the sampling rate, the ARMA, SETAR and STAR models have a similar level of performance, while the RMSE for the MSAR models is much lower. The improvement obtained with the Markov-switching models with respect to the three other types of models ranges from 19% to 32% depending on the sampling rate. Then, STAR models have an advantage versus the two others since the best STAR model is ranked second in all cases. The performance of the SETAR and ARMA models are a lot alike. It appears that considering separate regimes does not give any improvement against the

classical linear models unless the switches between regimes are smoothed and controlled by some transition function. And, the hypothesis of some succession of regimes that could be captured with a first order Markov chain is validated by these results.

The testing set for Nysted is composed by 14 periods of different lengths and with different characteristics eg. various mean production levels. The detail of the performance of the various models is shown in Figure 22, which gives the RMSE of the models listed in Table 2 for each period. There are only few periods for which the level of performance of the MSAR is worse than that of the other models. In general, the performance of all models is similar from one period to the other, and it does not seem that certain type of conditions would advantage such or such type of model. Also, by noticing that the curves for ARMA, SETAR and STAR models lie on top of each other whatever the period, one understands that modelling the regime-switching with a lagged value of measured wind power output does not yield a more dynamic modelling of the power fluctuations.

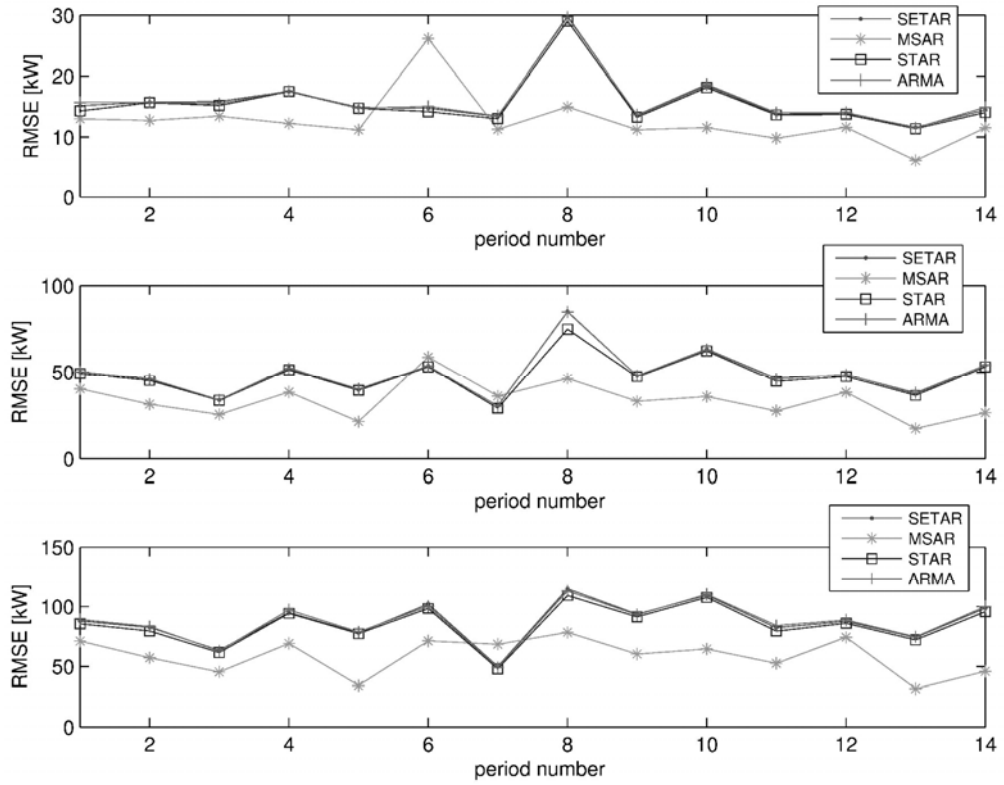

Figure 22. The RMSE on all test data sets from Nysted for each model. Top: 1 minute. Middle: 5 minute. Bottom: 10 minute.

Similar results are obtained for the Horns Rev wind farm. They are not extensively described in [2]. For both wind farms, the application of the various regime-switching modelling approaches leads to the conclusion that Markov-switching models permit to capture some complex behaviour induced by meteorological processes, which cannot be captured with regime-switching models relying on an observable regime sequence. For that reason, it is decided in the following to focus on this modelling approach only, and to enhance it so it better accounts for the specific characteristics of the wind generation process.

7.3 A MSAR model with time-varying coefficients

A drawback of the method described above (or more extensively in [2]) is that model coefficients are not time-varying, while it is known that wind generation is a process with long-term variations due to eg. changes in the wind farm environment, seasonality or climate change, see eg. [39]. The main objective of the present paper is to describe a method for adaptive estimation in Markov-switching autoregression that indeed allows for time-varying coefficients. This method utilizes a parameterization inspired by those proposed in [40][41] and in [42]. Adaptivity in time is achieved with exponential forgetting of past observations. In addition, the formulation of the objective function to be minimized at each time-step includes a regularization term that permits to dampen the variability of the model coefficient estimates. A recursive estimation procedure permits to lower computational costs by updating estimates based on newly available observations only. In parallel, advantage is taken of the possibility to express predictive densities from Markov-switching autoregressive models for associating one-step ahead forecasts with prediction intervals. Predictive densities are given as a mixture of conditional densities in each regime, the quantiles of which can be obtained by numerical integration methods. A thorough description of the proposed models and related adaptive estimation methods is given in [43].

7.4 Performance and advantages of MSAR models with time-varying parameters for modeling and forecasting offshore wind power fluctuations

Case-studies are the same than those considered in Section 3. Focus is given here to time-series of power production data averaged at a 10 minutes rate. Model setup is described in a first stage, followed by point forecasting and probabilistic forecasting results.

7.4.1 Model configuration and estimation setup

From the averaged data, it is necessary to define periods that are used for training the statistical models and periods that are used for evaluating what the performance of these models may be in operational conditions. These two types of datasets are referred to as learning and testing sets. We do not want these datasets to have any data considered as not valid. Sufficiently long periods without any invalid data are then identified and permit to define the necessary datasets. For both wind farms, the first 6000 data points are used as a training set, and the remainder for out-of-sample evaluation of the 1-step ahead forecast performance of the Markov-switching autoregressive models. These evaluation sets contain N_n=20650 and N_h=21350 data points for Nysted and Horns Rev, respectively. Over the learning period, a part of the data is used for one-fold cross validation (the last 2000 points) in order to select optimal values of the forgetting factor and regularization parameter. The autoregressive order of the Markov-switching models is arbitrarily set to $p=3$, and the number of regimes to $r=3$. For more information on cross validation, we refer to [44]. The error measure that is to be minimized over the cross validation set is the Normalized Root Mean Square Error (NRMSE), since it is aimed at having 1-step ahead forecast that would minimize such criterion over the evaluation set.

For all simulations, the autoregressive coefficients and standard deviations of conditional densities in each regime are initialized as

$$\begin{aligned}\theta_0^{(1)} &= [0.2, 0, 0, 0], \quad \sigma_0^{(1)} = 0.15 \\ \theta_0^{(2)} &= [0.5, 0, 0, 0], \quad \sigma_0^{(2)} = 0.15 \\ \theta_0^{(3)} &= [0.8, 0, 0, 0], \quad \sigma_0^{(3)} = 0.15\end{aligned} \tag{20}$$

while the initial matrix of transition probabilities is set to

$$P_0 = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.1 & 0.8 & 0.1 \\ 0 & 0.2 & 0.8 \end{bmatrix} \tag{21}$$

It is considered that the forgetting factor for adaptive estimation of model parameters cannot be less than $\lambda = 0.98$, since lower values would correspond to an effective number of observations smaller than 50 data points. Such low value of the forgetting factor would then not allow for adaptation with respect to slow variations in the process characteristics, but would serve more for compensating for very bad model specification. No restriction is imposed on the potential range of values for the regularization parameter.

7.4.2 Point forecasting results

The results from the cross-validation procedure, i.e. the values of both the forgetting factor and the regularization parameter that minimize the 1-step ahead NRMSE over the validation set, can be found in [43]. In both cases, the forgetting factor takes value very close to 1, indicating that changes in process characteristics are indeed slow. Fast and abrupt changes are dealt with thanks to the Markov-switching mechanism. In addition, regularization parameter values are not equal to zero, showing the benefits of the proposal. Note that one could actually increase this value even more if interested in dampening variations in model estimates, though this would affect forecasting performance.

For evaluation of out-of-sample forecast accuracy, we follow the approach presented in [38] for the evaluation of short-term wind power forecasts. Focus is given to the use of error measures such as NRMSE and Normalized Mean Absolute Error (NMAE). In addition, forecasts from the proposed Markov-switching autoregressive models are benchmarked against those obtained from persistence. Persistence is the simplest way of producing a forecast and is based on a random walk model. A 1-step ahead persistence forecast is equal the last power measure. Despite its apparent simplicity, this benchmark method is difficult to beat for short-term look-ahead time such as that considered here.

The forecast performance assessment over the evaluation set is summarized in Table 3. NMAE and NRMSE criteria have lower values when employing Markov-switching models. This is satisfactory as it was expected that predictions would be hardly better than those from persistence. The reduction in NRMSE and NMAE is higher for the Nysted wind farm than for the Horns Rev wind farm. In addition, the level of error is in general higher for the latter wind farm. The Horns Rev wind farm is located in the North Sea (while Nysted is in the Baltic sea, south of Zealand in Denmark). It may be more exposed to stronger fronts causing fluctuations with larger magnitude, and that are less predictable.

Table 3. One-step ahead forecast performance over the evaluation set for Nysted and Horns Rev. Results are both for persistence and Markov-switching models. Performance is summarized with MAE and NRMSE criteria, given in percentage of the nominal capacity of the representative single turbine.

	persistence			Markov-switching model		
	Nbias	NMAE	NRMSE	Nbias	NMAE	NRMSE
Nysted	-0.0001	2.37	4.11	-0.0005	2.20	3.79
Horns Rev	0.0	2.71	5.06	-0.0004	2.70	4.96

An expected interest of the Markov-switching approach is that one can better appraise the characteristics of short-term fluctuations of wind generation offshore by studying the estimated model coefficients, standard deviations of conditional densities, as well as transition probabilities. Autoregressive coefficients may inform on how the persistent nature of power generation may evolve depending on the regime, while standard deviations of conditional densities may tell on the amplitude of wind power fluctuations depending on the regime. Finally, the transition probabilities may tell if such or such regime is more dominant, or if some fast transitions may be expected from certain regimes to the others.

The set of model coefficients at the end of the evaluation set for Nysted can be summarized by the model autoregressive coefficients and related standard deviations of related conditional densities

$$\theta_{N_n}^{(1)} = [0, 1.361, -0.351, -0.019], \quad \sigma_{N_n}^{(1)} = 0.007$$
$$\theta_{N_n}^{(2)} = [0.013, 1.508, -0.778, 0.244], \quad \sigma_{N_n}^{(2)} = 0.041$$
$$\theta_{N_n}^{(3)} = [-0.001, 1.435, -0.491, 0.056], \quad \sigma_{N_n}^{(3)} = 0.011$$

while the final matrix of transition probabilities is

$$P_{N_n} = \begin{bmatrix} 0.888 & 0.036 & 0.076 \\ 0.027 & 0.842 & 0.131 \\ 0.051 & 0.075 & 0.874 \end{bmatrix}$$

In parallel for Horns Rev, the autoregressive coefficients and related standard deviations are

$$\theta_{N_h}^{(1)} = [0.002, 1.253, -0.248, -0.008], \quad \sigma_{N_h}^{(1)} = 0.023$$
$$\theta_{N_h}^{(2)} = [0.022, 1.178, -0.3358, 0.123], \quad \sigma_{N_h}^{(2)} = 0.066$$
$$\theta_{N_h}^{(3)} = [0.069, 0.91, -0.042, -0.022], \quad \sigma_{N_h}^{(3)} = 0.005$$

while the final matrix of transition probabilities is

$$P_{N_h} = \begin{bmatrix} 0.887 & 0.069 & 0.044 \\ 0.222 & 0.710 & 0.068 \\ 0.173 & 0.138 & 0.689 \end{bmatrix}$$

For both wind farms, the first regime is dominant in the sense that it has the highest probability of keeping on with the same regime when it is reached. However, one could argue that the first regime is more dominant for Horns Rev, as the probabilities of staying in second and third regimes are lower, and as the probabilities of going back to first regime are higher. The dominant regimes have different characteristics for the two wind farms. At Nysted, it is the regime with the lower standard deviation of the conditional density, and thus the regime where fluctuations of smaller magnitude are to be expected. It is not the case at Horns Rev, as the dominant regime is that with the medium value of standard deviations of conditional densities. Such finding confirms the fact that power fluctuations seem to be of larger magnitude at Horns Rev than at Nysted.

Let us study an arbitrarily chosen episode of power generation at the Horns Rev wind farm. For confidentiality reason, the dates defining beginning and end of this period cannot be given. The episode consists of 250 successive time-steps with power measurements and corresponding one-step ahead forecasts as obtained by the fitted Markov-switching autoregressive model. These 250 time steps represent a period of around 42 hours. The time-series of power production over this period is shown in Figure 23, along with corresponding one-step ahead forecasts. In parallel, Figure 24 depicts the evolution of the filtered probabilities, i.e. the probabilities given by the model of being in such or such regime at each time step. Finally, the evolution of the standard deviation of conditional densities in each regime is shown in Figure 25.

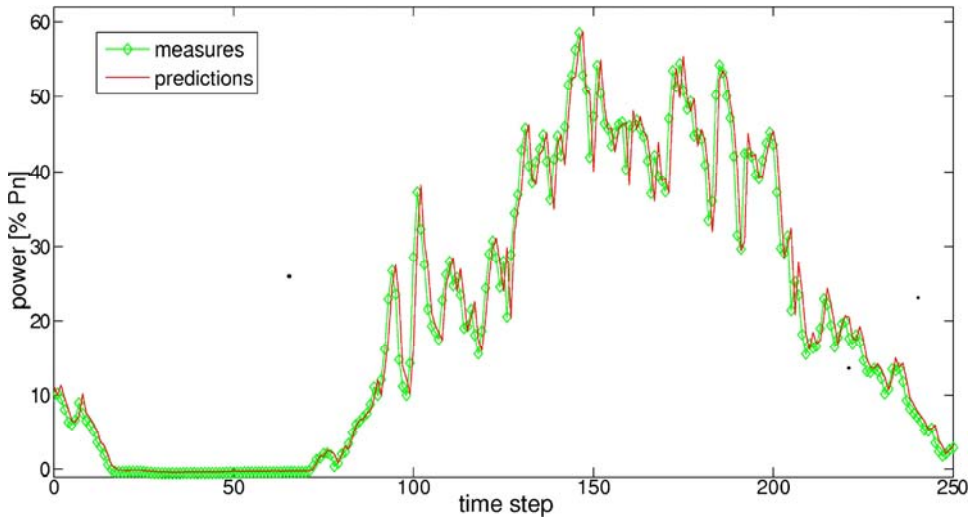

Figure 23. Time-series of nomalised power generation at Horns Rev (both measures and one-step ahead predictions) over an arbitrarily chosen episode

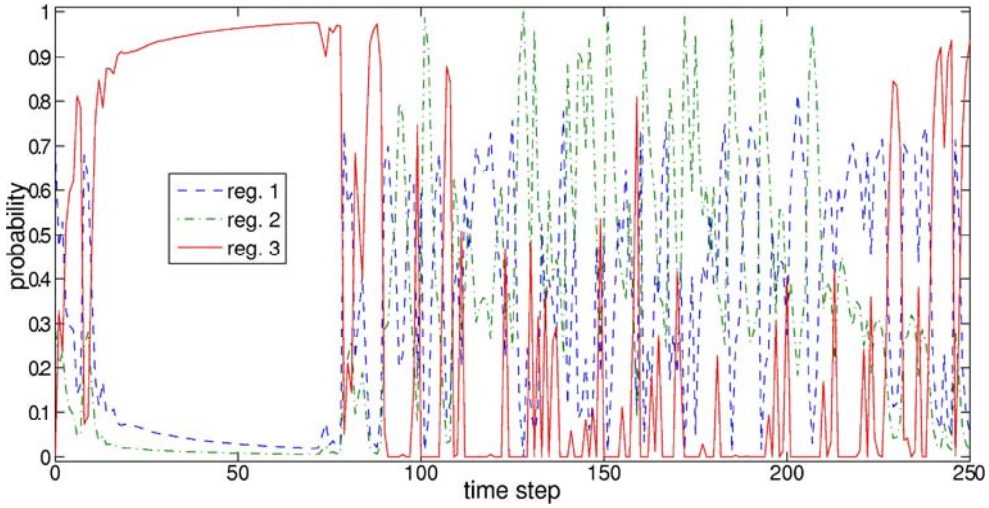

Figure 24. Time-series of filtered probabilities given by the Markov-switching model over the same period as in Figure 23.

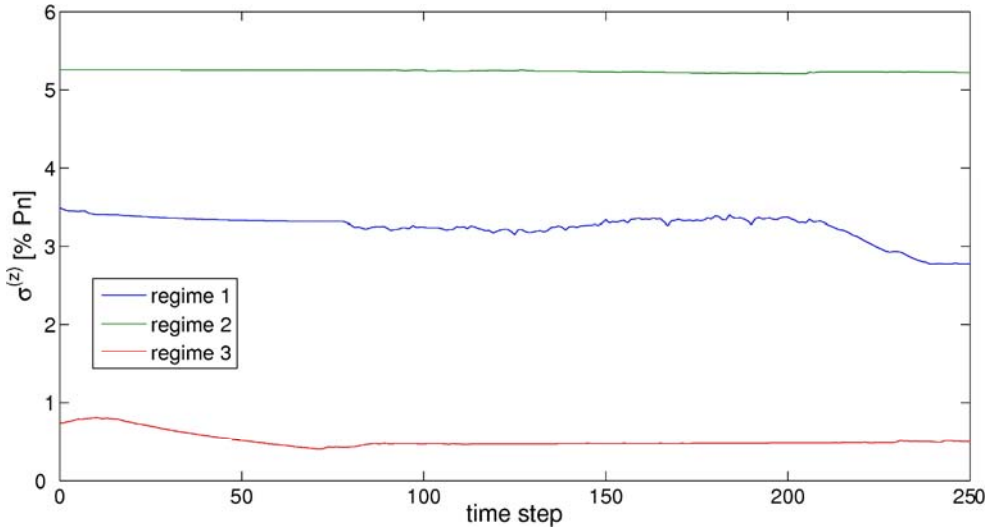

Figure 25. Time-series of the standard deviation of conditional densities in the various regimes for the same episode as in Figure 23.

First of all, it is important to notice that there is a clear difference between the three regimes in terms of magnitude of potential power fluctuations. There is a ratio 10 between the standard deviations of conditional densities between regime 2 and 3. In addition, these regimes are clearly separated, as there is a smooth evolution of the standard deviation parameters over the episode. If focusing on the power time-series of Figure 23, one observes successive periods with fluctuations of lower and larger magnitude. Then, by comparison with the evolution of filtered probabilities in Figure 24, one sees that periods with highly persistent behavior of power generation are all associated with very high probability of being in the first regime. This is valid for time steps between 20 and 80 for instance. This also shows that regimes are not obviously

related to a certain level of power generation, as it would be the case of using eg. SETAR models. If looking again at the autoregressive coefficients in each regime given above for Horns Rev at the end of the evaluation period, one clearly sees that intercept coefficients are almost zero. While regime 1 appears to be the regime with low magnitude fluctuations, both regime 2 and 3 contribute to periods with larger ones. Studying obtained series of filtered probabilities along with the evolution of some meteorological variables is expected to give useful information for better understanding meteorological phenomena that govern such behavior. This would then permit to develop prediction methods taking advantage of additional explanatory variables.

7.4.3 Interval forecasting results

In a second stage, focus is given to the uncertainty information provided by the Markov-switching autoregressive models. Indeed, even if point predictions in the form of conditional expectations are expected to be relevant for power management purposes, the whole information on fluctuations will actually be given by prediction intervals giving the potential range of power production in the next time-step, with a given probability i.e. their nominal coverage rate. Therefore, the possibility of associating point predictions with central prediction intervals is considered here. Central prediction intervals are intervals that are centred in probability around the median. For instance, a central prediction interval with a nominal coverage rate of 80% has its bounds consisting of the quantile forecasts with nominal proportions 0.1 and 0.9. Therefore, for evaluating the reliability of generated interval forecasts, i.e. their probabilistic correctness, one has to verify the observed proportions of quantiles composing the bounds of intervals. For more information on the evaluation of probabilistic forecasts, and more particularly for the wind power application, we refer to [42][43].

Prediction intervals are generated over the evaluation set for both Horns Rev and Nysted. The nominal coverage of these intervals range from 10% to 90%, with a 10% increment. This translates to numerically calculating 18 quantiles of the predictive densities [43]. The observed coverage for these various prediction intervals are gathered in Table 4.

Table 4. Empirical coverage of the interval forecasts produced from the Markov-switching autoregressive models for Horns Rev and Nysted.

nominal coverage [%]	empirical cov. Horns Rev [%]	empirical cov. Nysted [%]
10	10.09	10.38
20	21.23	19.55
30	31.48	28.69
40	41.67	38.16
50	51.36	48.59
60	61.39	59.18
70	70.45	69.59
80	79.84	79.92
90	89.59	90.92

The agreement between nominal coverage rates and observed one is good, with deviations from perfect reliability globally less than 2%. However as explained above, this valuation has to be carried further by looking at the observed proportions of related quantile forecasts, in order to verify that intervals are indeed properly centred. Such

evaluation is performed in Figure 26 by the use of reliability diagrams, which gives the observed proportions of the quantiles against the nominal ones. The closer to the diagonal the better. For both wind farms, the reliability curve lies below the diagonal, indicating that all quantiles are underestimated (in probability). This underestimation is more significant for the central part of predictive densities. Note that for operational applications one would be mainly interested in using prediction intervals with high nominal coverage rates, say larger than 80%, thus corresponding to quantile forecasts that are more reliable in the present evaluation. It seems that the Gaussian assumption for conditional densities allows to have predictive densities (in the form of Normal mixtures) that appropriately capture the shape of the tails of predictive distributions, but not their central parts. Using nonparametric density estimation in each regime may allow to correct for that.

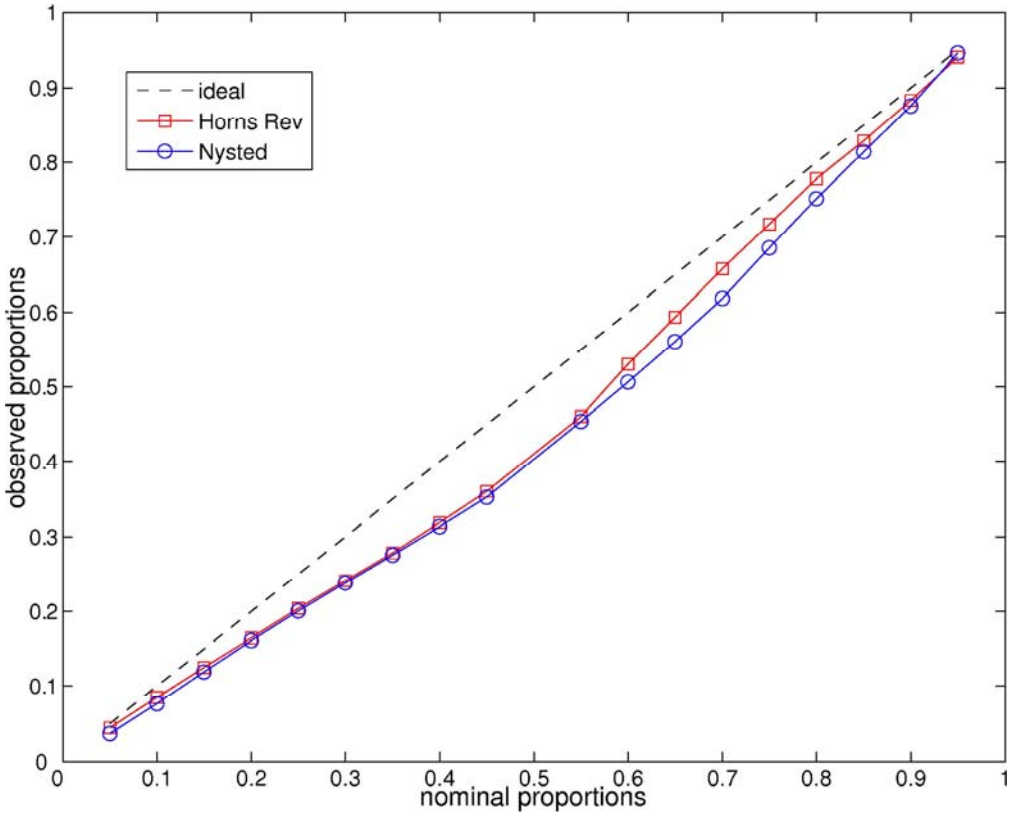

Figure 26. Reliability evaluation of quantile forecasts obtained from the Markov-switching autoregressive models for both Horns Rev and Nysted. Such reliability diagram compares nominal and observed quantile proportions.

Finally, Figure 27 depicts the same episode with power measures and corresponding one-step ahead point prediction that than shown in Figure 23 for the Horns Rev wind farm, except that here point predictions are associated with prediction intervals with a nominal coverage rate of 90%. Prediction intervals with such nominal coverage rate are the most relevant for operation applications, and they have been found to be the most reliable in practice. The size of the prediction intervals obviously varies during this 250 time-step period, with their size directly influenced by forecasts of filtered probabilities and standard deviations of conditional densities in each regime. In addition, prediction intervals are not symmetric, as even if conditional densities are assumed to be Gaussian in each regime, the resulting one-step ahead predictive densities are clearly not. In this

episode, prediction intervals are wider during periods with power fluctuations of larger magnitude. Even though point predictions may be less accurate (in a mean square sense) during these periods of larger fluctuations, Markov-switching autoregressive models can provide this valuable information about their potential magnitude.

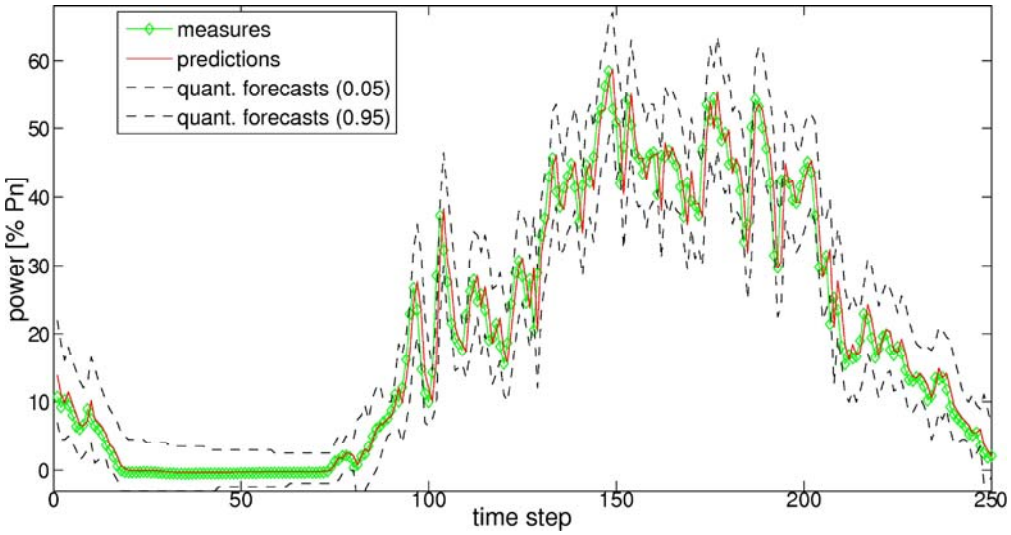

Figure 27. Time-series of normalized power generation at Horns Rev (both measures and one-step ahead predictions) over an arbitrarily chosen episode, accompanied with prediction intervals with a nominal coverage of 90%.

7.5 Conclusions

Particular attention has to be given to the modelling of the fluctuations of offshore wind generation, since dedicated models are needed for enhancing the existing control and energy management strategies at offshore wind parks. This issue has been addressed by applying chosen statistical models. The choice for regime-switching approaches has been motivated by the succession of periods with fluctuations of lower and larger magnitudes that can be easily noticed when inspecting time-series of offshore wind power production averaged at a minute rate.

Two different types of regime-switching approaches have been applied. On the one hand, SETAR and STAR models rely on explicit rules for determining what the current regime is. It is in practice given by some function of past values of measured power. On the other hand, MSAR models are based on the idea that the regime-switching is governed by a hidden Markov process. This, from theoretical point of view, may allow one to capture some complex influence of meteorological conditions on the wind power fluctuations. For verifying this (a priori) nice feature of MSAR models, they have been compared to SETAR, STAR and ARMA models on a one-step ahead forecasting exercise, with the aim of minimizing a quadratic error criterion. In all cases, it has been found that MSAR models significantly outperform the other ones: the error reduction ranges between 19 and 32% depending on the test case and the sampling rate. The gain of applying SETAR or STAR models instead of simple linear ARMA models does exist, but is relatively small. MSAR models indeed manage to capture the influence of some complex meteorological features on the power fluctuations. It will be of particular interest to study the relation between the temporal evolution of some meteorological

variables and the regime sequences of MSAR models in order to determine which of these variables have a direct impact on the magnitude of power fluctuations. Integrating this knowledge in existing forecasting methods will permit to significantly increase their skill for the specific case of very short-term prediction (from some minutes to few hours) at offshore sites.

The MSAR class of models has been generalized in a second stage so that they are allowed to have time-varying coefficients, though slowly varying, in order to follow the long-term variations in the wind generation process characteristics. An appropriate estimation method using recursive maximum likelihood and Tikhonov regularization has been introduced. The proposal for including a regularization term comes from the aim of application to noisy wind power data, for which the use of non-regularized estimation may result in ill-conditioned numerical problems. The convergence and tracking abilities of the method have been shown from simulations. Then, Markov-switching autoregressive models have been employed for characterizing the 10 minutes power fluctuations at Horns Rev and Nysted, the two largest offshore wind farms worldwide. The models and related estimation method have been evaluated on a one-step ahead forecasting exercise, with persistence as a benchmark. For both wind farms, the forecast accuracy of the proposed approach is higher than that of persistence, with the additional benefit of informing on the characteristics of such fluctuations. Indeed, it has been possible to identify regimes with different autoregressive behaviors, and more importantly with different variances in conditional densities. This shows the ability of the proposed approach to characterize periods with lower or larger magnitudes of power fluctuations.

In addition to generating point predictions of wind generation, it has been shown that the interest of the approach proposed also lied in the possibility of associating prediction intervals or full predictive densities to point predictions. Indeed when focusing on power fluctuations, even if point predictions give useful information, one is mainly interested in the magnitude of potential deviations from these point predictions. It has been shown that for large nominal coverage rates (which are the most appropriate for operational applications) the reliability of prediction intervals was more acceptable than for low nominal coverage rates. It is known that for the wind generation process, noise distributions are not Gaussian, and that the shape of these distributions is influenced by the level of some explanatory variables [34]. Therefore, in order to better shape predictive densities, the Gaussian assumption should be relaxed in the future. Nonparametric density estimation may be achieved with kernel density estimators, though this may introduce some problems in a recursive maximum likelihood estimation framework eg. multi-modality of conditional densities.

One-step ahead predictive densities of power generation have been explicitly formulated. Such densities consist of finite mixtures of conditional densities in each regime. However, it has been explained that the issue of parameter uncertainty was not considered, and that this may also affect the quality of derived conditional densities, especially in an adaptive estimation framework where the quality of parameter estimation may also vary with time. Novel approaches accounting for such parameter uncertainty should hence be proposed. The derivation of analytical formula might be difficult. In contrast, one may think of employing nonparametric block bootstrap procedures. This will be the focus of further research.

Broader perspectives regarding follow-up studies include the development of stochastic models for simulating the interaction of offshore wind generation with conventional

generation or storage, used as a backup for smoothing the fast power fluctuations at offshore wind farms. Better control strategies will result from the application of these models, which will significantly reduce the potential large costs induced by unwanted large power fluctuations.

Finally, it should be stressed that the presented model actually only simulate the available wind power. Control of the wind farms can mitigate the power fluctuations on the cost of reduced energy production. The challenge is to develop control strategies with an appropriate balance between lost energy production and benefit of less severe power fluctuations. Especially "storm control" is relevant, because storms are relatively rare, and therefore the lost energy production due to storm control will be limited. The control issue will be dealt with in the PSO Mesoscale project, and other future projects.

8 Conclusion

Conclusions are presented for the correlation models and regime switching models in sections 6.5 and 7.5 respectively. Below, a few more general conclusions are made.

The presented models are statistical models, based only on the observed wind speed and power measurements. Future models could take advantage of including other observations such as rain showers and other indicators of unstable weather.

Wind power fluctuations are expected to be of an increasing importance to the balancing of future power systems, taking into account that the plans for wind power development in many European countries are suggesting large offshore solutions with wind power concentrated in a number of relatively small areas with a combination of appropriate water depths, wind resources and grid connection possibilities. In this perspective, it is particularly relevant to use simulation and prediction tool that can take into account realistic correlation between the wind power variability inside and between these areas.

References

[1] Sørensen, P.; Cutululis, N.A.; Vigueras-Rodriguez, A.; Madsen, H.; Pinson, P.; Jensen, L.E.; Hjerrild, J.; Donovan, M.H., Modelling of Power Fluctuations from Large Offshore Wind Farms. Wind Energy (2008) 11 , 29-43.

[2] Pinson, P., Christensen, L.E.A., Madsen, H., Sørensen, P.E., Donovan, M.H., Jensen, L.E. 2006. Regime-switching modelling of the fluctuations of offshore wind generation. Journal of Wind Engineering and Industrial Aerodynamics (2008) 96 (no.12) , 2327-2347.

[3] Sørensen, P.; Cutululis, N.A.; Vigueras-Rodriguez, A.; Jensen, L.E.; Hjerrild, J.; Donovan, M.H.; Madsen, H., Power fluctuations from large wind farms. IEEE Trans. Power Systems (2007) 22 , 958-965

[4] B. M. Weedy, B.J. Cory, *Electric power systems*, Chichester, United Kingdom: John Wiley & Son, Fourth Edition 1998, ch. 4.

[5] http://www.nordpool.com

[6] Nordic grid code: Nordel, June 2004

[7] I. Troen, L. Landberg, "Short term prediction of local wind conditions," in *Proc. European Community Wind Energy Conference Madrid 1990,* pp. 76-78.

[8] V. Akhmatov, H. Abildgaard, J. Pedersen, P. B. Eriksen, "Integration of Offshore Wind Power into the Western Danish Power System," in *Proc. Copenhagen Offshore Wind 2005*, CD.

[9] En visionær dansk energipolitik 2025. Faktaark – Vedvarende energi. http://www.ens.dk/graphics/Energipolitik/dansk_energipolitik/Energistrategi2025/Faktaark_VE_190107Endelig.pdf. 19. januar 2007.

[10] B. Parson, M. Milligan, B. Zavadil, D. Brooks, B. Kirby, K. Dragoon, J. Caldwell, "Grid Impacts of Wind Power: A summary of recent studies in the United States," *Wind Energy* , vol. 7, Apr./Jun. 2004, pp 87-108.

[11] D. Brooks, J. Smith, E. Lo, S. Ishikawa, L. Dangelmaier, M. Bradley, "Development of a methodology for the assessment of system operation impacts of integrating wind generation on a small island power system," in *Proc. European Wind Energy Conference London 2004,* CD.

[12] J. R. Kristoffersen, "The Horns Rev wind farm and the operational experience with the wind farm main controller," in *Proc. Copenhagen Offshore Wind 2005*, CD.

[13] P. Sørensen, A. D. Hansen, P. A. C. Rosas, "Wind models for simulation of power fluctuations from wind farms", *J. Wind Eng. Ind. Aerodyn.* (2002) (no.90) , 1381-1402

[14] IEC 61400-21. Wind turbine generator systems. Part 21. Measurement and assessment of power quality characteristics of grid connected wind turbines. Geneva November 2001.
[15] A.D. Hansen, P. Sørensen, L. Janosi, J. Bech, Wind Farm Modelling for Power Quality, Proc. of IECON '2001
[16] P. Sørensen, J. Mann , U. S. Paulsen and A. Vesth, "Wind farm power fluctuations", *Euromech 2005*, Kassel, October 2005.
[17] von Karman, T.: 1948, Progress in the statistical theory of turbulence, Proc. National Akad. Sci. 34, 530–539.
[18] J. C. Kaimal, J.C. Wyngaard, Y. Izumi, O.R. Cote, Spetral characteristics of surface-layer turbulence, Q.J.R. Meteorol. Soc. vol 98, 1972, pp. 563-598.
[19] IEC 61400-1. Wind turbines – Part 1. Design requirements. 3rd edition. Geneva August 2005.
[20] S. Frandsen. Turbulence and turbulence-generated structural loading in wind turbine clusters. Risø-R-1188(EN). Roskilde, November 2005.
[21] M. Courtney, I. Troen. Wind speed spectrum from one year of continuous 8 Hz measurements.
[22] W. Schlez and D. Infield. Horizontal, two point coherence for separations greater than the measurement height. *Boundary-Layer Meteorology 87.* Kluwer Academic Publishers, Netherlands 1998. pp.459-480.
[23] P. Sørensen. Frequency domain modelling of wind turbine structures. Risø-R-749. Roskilde 1994.
[24] P. S. Veers. Three-dimensional wind simulation. SAND88-0152. Sandia National Laboratories. Albuquerque 1988.
[25] P. Nørgaard, H. Holttinen, "A Multi-Turbine Power Curve Approach", Nordic Wind Power Conference, NWPC 2004. Gothenburg. March 2004
[26] System Plan 2007. Energinet.dk. Available in Danish and in Enelish
[27] Fremtidens havmølleplaceringer - 2025. Udvalget for fremtidens havmølleplaceringer. Energistyrelsen (Danish Energy Authority) April 2007.
[28] Focken, U., Lange, M., Monnich, M., Waldl, H.-P., Beyer, H.-G., Luig, A., 2002. Short-term prediction of the aggregated power output of wind farms - A statistical analysis of the reduction of the prediction error by spatial smoothing effects. Journal of Wind Engineering and Industrial Aerodynamics 90: 231-246.
[29] Hendersen, A.R., Morgan, C., Smith, B., Sørensen, H.C., Barthelmie, R.J., Boesmans, B., 2003. Offshore wind energy in Europe - A review of the state-of-the-art. Wind Energy 6: 35-52.
[30] Madsen, H., 2004. Time Series Analysis (2nd edition). Technical University of Denmark: Kgs. Lyngby, Denmark. (ISBN 87-643-00098-6)
[31] Tong, H., 1990. Non-Linear Time Series - A Dynamical Approach (1st edition). Oxford University Press: Oxford, UK.
[32] Chan, K.S., Tong, H., 1986. On estimating thresholds in autoregressive models. Journal of Times Series Analysis 7: 178-190.
[33] Madsen, H., 1996. Models and Methods for Wind Power Forecasting. Eltra, Skærbæk, Denmark. ISBN 87-87090-29-5.
[34] Hamilton, H., 1989. A new approach to the economic analysis of nonstationay time-series and business cycles. Econometrica 57: 357-38.
[35] Robertson, A.W., Kirshner, S., Smyth, P., 2003. Hidden Markov models for modeling daily rainfall occurence over Brazil. Report UCI-ICS-03-27, Information and Computer Sciences, University of California, Irvine (California).
[36] Ailliot, P., Monbet, V., 2006. Markov switching autoregressive models for wind time series. Journal of Statistical Planning and Inference (submitted).
[37] Pinson, P., 2006. Estimation of the Uncertainty in Wind Power Forecasting. Ph.D. Thesis, Ecole des Mines Paris, France.
[38] Madsen, H., Pinson, P., Nielsen, T.S., Nielsen, H.Aa., Kariniotakis, G., 2005. Standardizing the performance evaluation of short-term wind power prediction models. Wind Engineering 25: 475-489.
[39] Giebel, G., Kariniotakis, G., Brownsword, R., 2003. The state of the art in short-term prediction of wind power - A literature overview. Position paper for the ANEMOS project, deliverable report D1.1, available online: http://www.anemos-project.eu, 2003.
[40] Collings, I.B., Krishnamurthy, V., Moore, J.B., 1994. On-line identification of hidden Markov models via prediction error techniques. IEEE Transactions on Signal Processing 42: 3535-3539.
[41] Collings, I.B., Ryden, T., 1998. A new maximum likelihood gradient technique algorithm for on-line hidden Markov model identification. Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, Seattle.
[42] Holst, U., Lindgren, G., Holst, J., Thuvesholmen, M., 1994. Recursive estimation in switching autoregressions with a Markov regime. Journal of Time Series Analysis 15: 489-506.
[43] Pinson, P., Madsen, H., 2007. Markov-switching autoregression with time-varying coefficients for offshore wind power fluctuations. Technical Report, Technical University of Denmark, Informatics and Mathematical Modelling.
[44] Stone, M., 1974. Cross-validation and assessment of statistical predictions (with discussion). Journal of the Royal Statistical Society B 36: 111-147.

Risø DTU is the National Laboratory for Sustainable Energy. Our research focuses on development of energy technologies and systems with minimal effect on climate, and contributes to innovation, education and policy. Risø has large experimental facilities and interdisciplinary research environments, and includes the national centre for nuclear technologies.

Risø DTU
National Laboratory for Sustainable Energy
Technical University of Denmark

Frederiksborgvej 399
PO Box 49
DK-4000 Roskilde
Denmark
Phone +45 4677 4677
Fax +45 4677 5688

www.risoe.dtu.dk