

# Modelling the Swedish wind power production using MERRA reanalysis data



Jon Olauson<sup>\*</sup>, Mikael Bergkvist

Division of Electricity, Department of Engineering Sciences, Uppsala University, Sweden

## ARTICLE INFO

### Article history:

Received 17 April 2014

Accepted 25 November 2014

Available online

### Keywords:

Wind power

Physical model

Wind variability

MERRA reanalysis data

## ABSTRACT

The variability of wind power will be an increasing challenge for the power system as wind penetration grows and thus needs to be studied. In this paper a model for generation of hourly aggregated wind power time series is described and evaluated. The model is based on MERRA reanalysis data and information on wind energy converters in Sweden. Installed capacity during the studied period (2007–2012) increased from around 600 to over 3500 MW. When comparing with data from the Swedish TSO, the mean absolute error in hourly energy was 2.9% and RMS error was 3.8%. The model was able to adequately capture step changes and also yielded a nicely corresponding distribution of hourly energy. Two key factors explaining the good results were the use of a globally optimised power curve smoothing parameter and the correction of seasonal and diurnal bias.

Because of bottlenecks in the Swedish transmission system it is relevant to model certain areas separately. For the two southern areas the MAE were 3.7 and 4.2%. The northern area was harder to model and had a MAE of 6.5%. This might be explained by a low installed capacity, more complex terrain and icing losses not captured in the model.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

## 1. Introduction

The variability of renewable energy sources is important to study since it affects load following costs, required investments in transmission capacity, emissions of CO<sub>2</sub> and other pollutants and electricity prices. In Sweden, installed wind power capacity has been growing fast during the last years. There are plans for a continued expansion of wind energy but also an ongoing debate on whether this is desirable and economically and technically feasible. 2012 was a record year when it comes to net export of electricity. One important question is whether Sweden should aim at becoming a large exporter of renewable energy, whether we should shut down nuclear power plants (which today contribute with around 45% of the electricity production) or whether it is better to keep the share of wind power below e.g. 10%. In order to have a fruitful debate and make rational decisions trustworthy models of wind power production are necessary.

In power system studies wind power production could be modelled either with statistical methods, e.g. auto-regressive [1–3]

or Monte Carlo [4,5] models, or with physical models. One advantage with statistical models is their ability to create arbitrary long time series. Physical models can be based either on meteorological measurements [6–8] or on meteorological models [9–15]. The use of meteorological models could in its turn be either direct or use statistical or dynamic downscaling to increase the resolution [16]. Some authors had also used simple upscaling of measured energy production to model future expansion [17,18].

A benefit with physical models compared to statistical is that hidden correlations with load will be observable. An example could be that during the coldest hours in winter wind energy converters (WECs) might produce more (or less) than average during that season. This would not be captured by most statistical models although seasonal and diurnal trends are taken into account. Still exactly those hours could be the most important to correctly capture since the load will then be at its maximum.

Sweden has the majority of its hydro power located in the northern part but around 85% of the inhabitants are living in the southernmost third. This means north-south bulk transmission and sometimes congestion due to bottle-necks. Because of this Sweden is divided into four electricity price areas (SE1–4) which are relevant to model separately (see Fig. 1). The basic design of the model presented here is:

<sup>\*</sup> Corresponding author. The Ångström Laboratory, Box 534, 751 21 Uppsala, Sweden. Tel.: +46 734 125930.

E-mail address: [jon.olauson@angstrom.uu.se](mailto:jon.olauson@angstrom.uu.se) (J. Olauson).

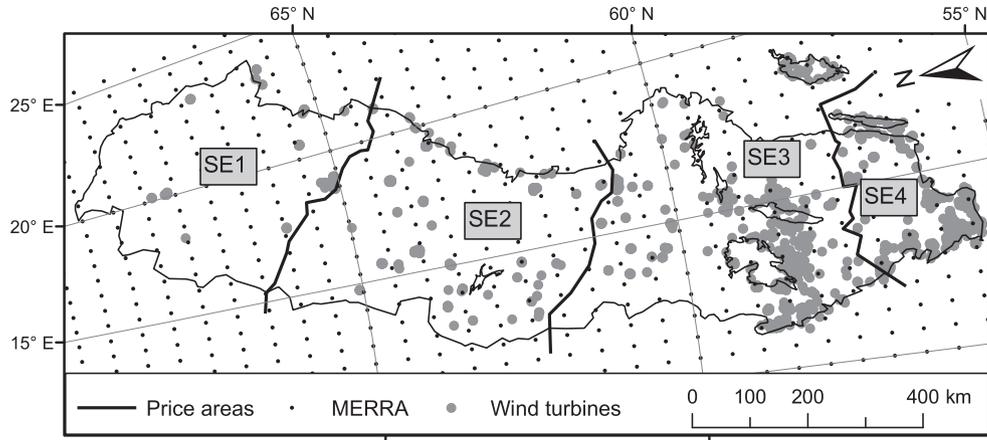


Fig. 1. Figure visualising MERRA grid points, distribution of WECs (end of 2012) and the four price areas in Sweden.

1. Start with MERRA meteorological time series and WEC information.
2. Calculate hourly wind at turbine hub height.
3. Calculate hourly energy for each WEC.
4. Aggregate hourly energy for Sweden or price areas.
5. Use bias correction to improve results.

The main purpose with this paper is to present a novel model where wake losses, multi-turbine smoothing etc. are taken into account, although they cannot be modelled on an individual turbine level. We want the model to perform well on the aggregated level and also to correct for systematic errors, both long-term trends and seasonal/diurnal bias. Data from years 2007, 2009 and 2011 was used for calibrating the parameter sets and data from 2008, 2010 and 2012 was used for evaluation. Section 2 describes the data used and Section 3 the model structure. In Section 4 an evaluation of model performance and parameter influence is presented. The paper is concluded with a discussion, conclusions and some suggestions for future work.

## 2. Data

Reanalysis data from the MERRA (Modern Era Retrospective-Analysis for Research and Applications) project [19] and information about WECs in Sweden were used in the model. Hourly averaged production data for the four price areas has been obtained from the Swedish Transmission System Operator (TSO) [20]. In Fig. 1 MERRA grid points and operating WECs in Sweden are shown.

### 2.1. MERRA

There are several global reanalysis datasets from meteorological models available. During the last years the performance of these has improved considerably. The MERRA dataset was chosen since it has a relatively high temporal and spatial resolution (one hour averages and  $0.5 \times 0.67^\circ$  respectively) and has shown a good correlation with wind measurements at relevant heights; Pearson's correlation coefficients are around 0.85 on an hourly basis and 0.94 on a monthly basis for measurements in terrain with low complexity [21]. Data is available from 1979 and onwards. Besides wind speed at different heights/pressure levels, wind direction, displacement height, temperature, moisture content, air pressure etc. can be downloaded free of charge [22]. Although the data used is a reanalysis and consequently has lower error compared to forecasts there are of course still uncertainties. These are partly due to the limited resolution, especially spatial.

### 2.2. Wind energy converters

Unfortunately there is no complete, available database with information on the around 2400 WECs built in Sweden until the end of 2012. Therefore data from different sources had to be put together. A lot of effort was put into filling data gaps and fixing erroneous data. Important information includes coordinates, rated power, hub height, rotor diameter, date of connection to (and possibly disconnection from) the grid, estimated annual energy production and price area. Although the coordinates are inexact for some WECs, the most dominant sources of error are expected to be the date of connection and, to less extent, the estimated annual energy production. The inexactness of the coordinates however influenced the model parameterisation, see Section 3.1.

As also noted in Ref. [23] there seem to be some error in the data for the price area SE1 from 2010 onwards. Hourly production is often higher than installed capacity. Contacts have been established with the Swedish TSO and Energy Agency, but the problem has not yet been resolved. Because of this no model for SE1 was calibrated and the data from SE1 2010–12 was excluded from the model for the whole of Sweden.

Two sets of WEC data were constructed, one primarily based on information from the Swedish electricity certificate system [24] and one based on a database earlier funded by the Swedish Energy Agency [25]. Complementary data sources were "Vindbrukskollen" [26] and communication with wind power owners. In the first dataset information on diameter and hub height is lacking for

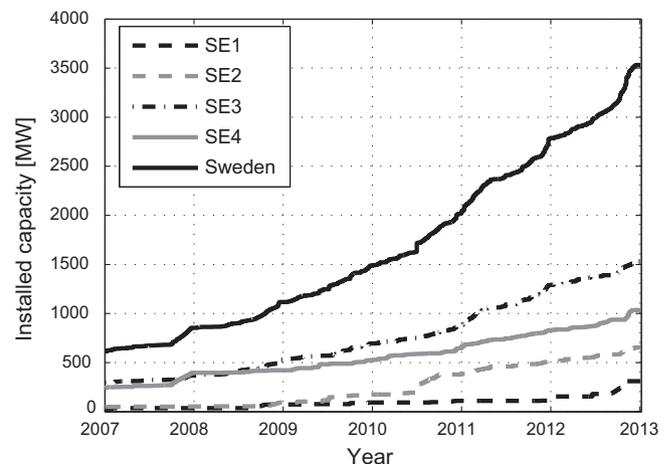


Fig. 2. Installed capacity of WEC in Sweden.

around 20% of the WECs. When no data on hub height was available this was calculated from the other WECs based on a second order polynomial fit of the hub height as a function of rated power. It was noted already at an early stage that the first dataset outperformed the second, and therefore only the first dataset was used in the later stages of modelling. In Fig. 2 the wind power capacity in the different price areas are shown. As can be seen SE2 had only a few WECs installed in the beginning of the period (around 50 MW) which make this area more challenging to model correctly.

### 3. Model

In this section the key functionalities of the wind power model are described. Firstly the choice of model structure, i.e. calculation procedures and parameterisation, is explained. After that the calculation of hourly wind speed and energy (including bias correction) are described and finally the objective function and optimisation technique are presented.

#### 3.1. Model structure

Using reanalysis data as input for the generation of wind power time series is not a new idea, see e.g. Refs. [9–15]. These works differ in level of complexity in the transformation from wind speed to aggregated wind power, in the possible use of parameters (that are optimised) in this process and in the method, if any, for evaluation of the model accuracy. We saw the following main reasons to use a different approach than the ones used in the earlier studies:

1. In order to model wake losses and power curve smoothing [27] for individual turbines, detailed information on the WECs are needed. In particular exact coordinates are necessary, but in the present dataset coordinates are often available only for the wind farm centre and sometimes only the county of the farm is known.
2. In earlier work, the models were often optimised and evaluated using data from individual met mast and wind farms or annual energy production for the power system. It cannot be guaranteed that this methodology gives the best performance on the aggregated power system scale.
3. The performance can be enhanced by introducing more parameters and an improved bias correction.

Based on physical considerations we identified several parameter candidates, some of them binary structure parameters and

**Table 1**  
Parameters used in the model.

Parameter	Range	Description
Dataset	{1, 2}	Data on WEC, described in Section 2.2.
Loss <sub>w</sub>	[0.06, 0.25]	Loss used when calculating scaling constant for wind, see Section 3.2.
Loss <sub>int</sub>	[0.00, 0.40] × 12	“Internal” losses for each wind sector, see Section 3.4.
Loss <sub>ext</sub>		2nd order polynomial describing “external” losses, see Section 3.4.
σ <sub>u</sub>	[0, 3]	Standard deviation [m/s] used for smoothing the power curve, see Section 3.4.
ρ <sub>var</sub>	{0, 1}	Structural parameter determining weather time-varying air density should be used.
α <sub>var</sub>	{0, 1}	Structural parameter determining weather time-varying wind shear exponent should be used.
P <sub>A,unknown</sub>	[229, 472]	Assumed rated power per rotor area [W/m <sup>2</sup> ] for WECs which lacks this information.
Bias correction		12 × 24 matrix correcting for seasonal/diurnal bias in aggregated energy.

some with a continuous range. A summary of the parameters and their purpose can be found in Table 1 and more detailed explanations are found in the subsequent sections. As will be shown in Section 4, some of these can hardly be justified since they did not significantly improve (or even deteriorated) the performance.

General ideas governing the choice of parameter candidates were that: i) wake losses and power curve smoothing are important phenomena that needs to be parameterised even though this cannot be done on an individual turbine level (i.e. global parameters are necessary), ii) the losses are likely to differ between prevailing and other directions, iii) there might be systematic errors in the underlying meteorological model giving rise to bias depending on season and time of day, iv) it can be important to distinguish losses that reduces the maximum WEC power and losses that reduces the power in the wind, but where full WEC output is still possible if the wind speed is high enough (called “external” and “internal” losses), v) that the (long-term corrected) energy production of the WECs as compared to what was anticipated by their owner is not constant in time and vi) that it would be interesting to see whether taking into account time-varying wind shear and air density improves the model performance.

#### 3.2. Wind speed and wind shear

The hourly wind speed has to be horizontally interpolated to the WEC position and vertically extrapolated to hub height. Horizontal interpolation was performed with the bilinear method. Vertical extrapolation was performed with the power law

$$u(z) = u(z_{ref}) \left( \frac{z - d}{z_{ref} - d} \right)^\alpha, \quad (1)$$

where  $u$  is wind speed,  $z$  is height above ground,  $d$  is displacement height and  $\alpha$  is the shear exponent. The displacement height describes the elevation of the zero level of the wind in and near cities, forest and other vegetation and is often around 0.6–1.0 of the canopy height.  $\alpha$  is a function of surface roughness, orography, atmospheric stability etc. [28]. Two different model structures were used, controlled by the parameter  $\alpha_{var}$ . In the first structure the shear exponent was a constant, adapted so that the average calculated annual energy production (AEP) using MERRA from 1979 to 2012 equals that specified by the WEC owner. In the second structure the shear exponent was allowed to vary according to MERRA wind speeds 10 and 50- $d$  metres above displacement height. In order to achieve the “correct” AEP, Equation (1) was multiplied with a constant. In both cases losses in calculation of AEP were controlled by the parameter Loss<sub>w</sub>.

#### 3.3. Wind direction

The hourly wind direction at WEC position was calculated as a weighted vector sum of the four surrounding MERRA grid points. The energy production is likely to be dependent on the wind direction e.g. through direction dependent park losses and topographic effects not captured by MERRA. With production series for individual WECs, a statistical model accounting for this effect could be built, e.g. using empirical power curves for the different wind directions. Since only aggregated energy was available, this effect was instead parameterised by allowing different losses in different wind sectors (bins of 30°). Sector 1 is the prevailing sector for each WEC and sector 2,3,...,12 follows clockwise. The prevailing sector was defined as the one with highest energy production calculated with a reference power curve (Vestas V90 2 MW) and MERRA data from 1979 to 2012.

### 3.4. Power curves and losses

Power curves from wind turbine manufacturers are given in form of power as a function of wind speed. The most important factor for the shape of the power curve is the ratio of rated power to rotor area ( $P_A$ ). By using four reference WECs, ranging from 229 to 472 W/m<sup>2</sup>, arbitrary power curves in that range could be interpolated. The chosen way to model wake effects, air density dependence etc. with only a few parameters was to transform the power curves into functions of power in the incoming wind ( $P_u$ , measured in Watts per square metre swept rotor area).

There are several reasons to believe that more smoothed power curves (i.e. higher power around cut-in wind speed, lower power around rated wind speed and a more smooth transition from rated to zero power at cut-out wind speed) are suitable:

- Power curves are certified using 10-minute averages of wind speed. In the MERRA model hourly averages are given.
- Sweden has generally higher turbulence levels compared to power curve test conditions.
- The limited spatial resolution of MERRA could lead to an underestimation of the smoothing effect of wind variability on aggregated power.

The smoothing effect was parameterised with a standard deviation on the incoming wind speed ( $\sigma_u$ ); the power curves were re-calculated using a normal distribution of wind speeds. To summarise, the WEC output power is expressed as in Equation (2). Some examples of power curves are shown in Fig. 3.

$$P_{\text{WEC}} = f(P_u, P_A, \sigma_u). \quad (2)$$

The use of a power curve dependent on incoming power in the wind gave the possibility to have both external and internal losses. An example of the former could be transmission losses and of the latter wake losses. The hourly energy fed in to the grid was hence modelled as

$$P_{\text{grid}} = (1 - \text{loss}_{\text{ext}}) \cdot P_{\text{WEC}} \\ = (1 - \text{loss}_{\text{ext}}) \cdot f\left((1 - \text{loss}_{\text{int}}) \cdot \frac{1}{2} \rho u^3, P_A, \sigma_u\right), \quad (3)$$

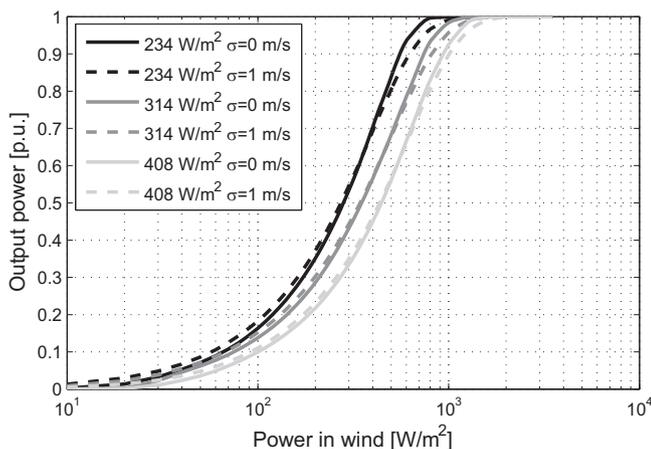


Fig. 3. Power curves for three selected WECs with different ratio of rated power to rotor area. The multi-turbine/smoothing effect accomplished by using a normal distribution with standard deviation of 1 m/s is also visualised. Cut-out region is not shown.

where  $\rho$  is air density. The model structure parameter  $\rho_{\text{var}}$  determines if air density should be constant (1.225 kg/m<sup>3</sup>) or calculated hour by hour from MERRA temperature and sea level air pressure. For the latter case the ideal gas law and the barometric equation (neglecting impact of moisture content) were used, combining to

$$\rho = \frac{\rho_0 T_0}{p_0} \cdot \frac{p_{\text{hub}}}{T} = \frac{\rho_0 T_0}{p_0 T} \cdot p_{\text{sea}} \cdot e^{-\frac{Mg}{R} h} \approx 0.003484 \cdot \frac{p_{\text{sea}}}{T} \cdot e^{-\frac{0.003416 \cdot h}{T}}, \quad (4)$$

where  $T$  is temperature,  $p$  is air pressure,  $h$  is hub height above sea level and  $M$ ,  $g$  and  $R$  are molar mass of air, gravitational acceleration and the universal gas constant. Subscript “0” indicates standard conditions and “sea” indicates sea level.

The internal losses were represented by 12 parameters, one for each wind sector, and were applied for individual turbines. The external losses were applied on the aggregated energy production in the form of a second order time-dependent polynomial function using Matlab robust fit with least absolute residuals. Two reasons for introducing time-dependent losses were identified. Firstly that the calculated energy production for WECs built in the 90's and early 00's were systematically overestimated. This was seen by analysing production for year 2–4 after commission as compared to the anticipated production, using data from Ref. [25]. Long-term correction was performed with the “wind index” method [29], and the results showed a ratio around 0.9 before year 2003 and slightly below 1 for recent years. The second reason is that WEC performance might deteriorate with age, see Ref. [30] for a UK study. Since the capacity weighted average turbine age has decreased from around 5.8 to 3.8 years during the studied period, it can be expected that the ratio between observed and modelled production should increase (we are of course here referring to production before application of the time-dependent loss term).

The combined effect of these two factors can be seen as a gradual increase in the production ratio, see Fig. 4. To make the trend easily visible, the figure shows monthly energy ratio and fit for the full six years, with seasonal bias removed. The actual fit in the model was however made on hourly energy for the optimisation years only.

A bias correction was performed to account for seasonal and diurnal bias in the aggregated production. This is motivated by an observed systematic error depending on month of year and time of the day, see Fig. 5. These errors could have several reasons

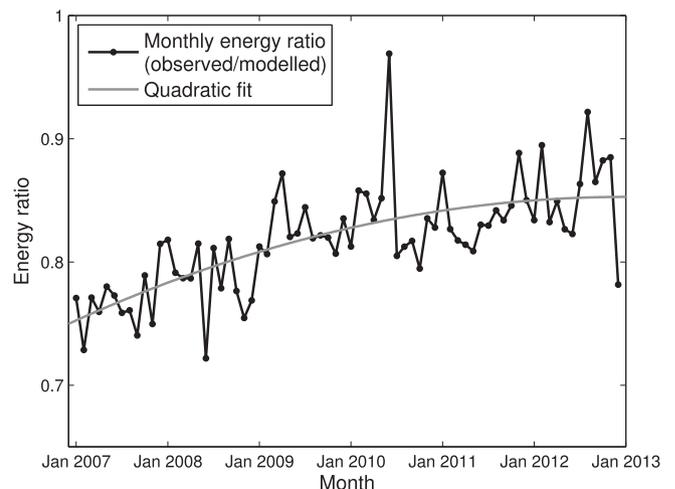


Fig. 4. Example of ratio of observed over modelled (before application of external losses) monthly energy. Seasonal bias has been removed.

including icing losses in wintertime and inability of MERRA to correctly capture the seasonal and diurnal dependence of the wind speed and wind shear.

3.5. Model optimisation

There are several desirable properties of a model of wind power production. Three of them were identified as the most important:

- Low error in hourly energy (P).
- Low error in energy step change ( $\Delta P$ ).
- Good match in statistical distributions.

A well calibrated model should fulfil all these criteria. One way to achieve this target is to create an objective function (OF) that combines goodness measures from the different categories mentioned above:

$$OF = f_P \cdot f_{\Delta P} \cdot f_{\text{statistical}} = E_P^2 \cdot (E_{\Delta P,1h}^2 + E_{\Delta P,4h}^2) \cdot (E_{S1} + E_{S2} + E_{S3} + E_{S4}) \quad (5)$$

where the different  $E$ :s are RMS errors. The statistical measures  $S1$ – $S4$  are duration curves for 1 h and 4 h step changes, histogram of hourly energy and monthly capacity factor, see Figs. 8–10.

The surface of the objective function in the parameter space has a lot of local minima but is locally relatively smooth. The parameters are dependent in a non-trivial way. Based on this the “random restart hill-climb optimisation”-technique was chosen to tune the parameters. Hill-climb optimisation starts with random parameter values and randomly changes one parameter. Structural parameters are not changed during the hill-climb. If a better parameter set was achieved, i.e. a lower value on the objective function, the change is accepted and a new parameter is randomly changed. The optimisation continues with consecutive smaller steps until no further improvement is possible. Subsequently a new starting point in parameter space is randomly chosen and the process starts over again. In a first stage internal losses were represented by only one parameter value, but subsequently losses were allowed to vary for the different sectors. 350 hill-climb searches were performed for the model of the whole of Sweden, evaluating in total 81,000 parameter sets. For the separate areas (SE2–4) similar amounts of runs were executed.

It is obvious that several parameter sets give similar performance, see Section 4.2. It could also be noted that some of the goodness measures have a negative correlation, i.e. trying to optimise the model for one measure deteriorate others. This was most obvious for the errors in step changes which, at the end of the hill-climb, counteracts the other measures.

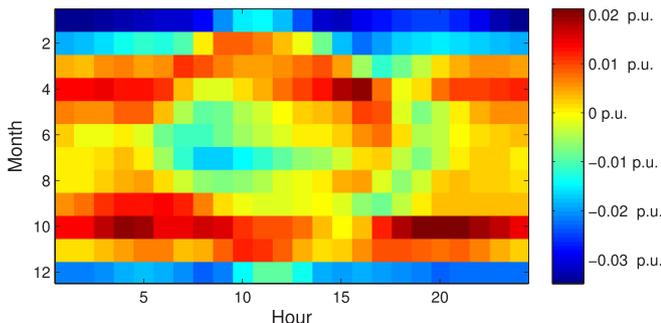


Fig. 5. Example of mean error as function of month and hour of the day (before bias correction).

Table 2

Model performance; results for the evaluation years 2008, 2010 and 2012. All errors are given as percentage of the installed power.

	Sweden	SE2	SE3	SE4
Mean absolute error	2.9%	6.5%	3.7%	4.2%
RMS error	3.8%	9.1%	5.0%	5.9%
Mean error	−0.1%	−0.7%	−0.5%	0.4%
RMS error $\Delta P, 1\text{ h}$	1.5%	3.0%	1.9%	2.8%
RMS error $\Delta P, 4\text{ h}$	3.6%	6.6%	4.7%	6.1%
Correlation	0.98	0.89	0.97	0.97

4. Results

In this section results for the whole of Sweden (SE) and the separately modelled price areas SE2, SE3 and SE4 are presented. In Table 2 some goodness measures for each area are shown. Recall that results are given for three evaluation years while the model is calibrated using three different years. All results are given in per unit (p.u.), where one p.u. represents the installed capacity as given in Fig. 2. The mean absolute error of hourly energy is slightly below 3% (i.e. 0.03 p.u.) for SE, around 4% for SE3 and SE4 and 6.5% for SE2. The same trend is visible in the RMS errors and mean error, i.e. lowest for the SE and highest for SE2. It is clear from these results that it is easier to model a larger area or an area with more installed power; the errors are largest for SE2 which had only around 50 MW installed during the beginning of the studied period. Other reasons for SE2 being harder to model could include more complex terrain and more extensive icing losses during winter.

When comparing with earlier work the errors are small. Aigner & Gjengedal [9] modelled the Danish and German (TenneT) system using COSMO data. These systems have large installed wind power capacity (3100 and 10,400 MW respectively), but the RMS errors were still relatively high; 7.1 and 6.5% respectively. Kubik et al. [15] modelled the small North Ireland system (290 MW) using MERRA data and got a RMSE of 11.9%. Two reasons for the smaller errors in the present work are more detailed information on the WECs and a more detailed model with more parameters.

Fig. 6 shows model output and measured data for eight weeks. As for all subsequent figures in this paper, results are given for SE. The RMSE for this period is 3.9%, i.e. almost identical to the RMSE for the full three years of evaluation. It is evident from the figure that the model can capture the observed energy levels and fluctuations well. Note in particular that the model performs well during periods of fast ramping. A histogram of the errors for the entire evaluation period is shown in Fig. 7.

In power system analysis the step changes and ramping rates are of great importance. The distribution of one and four hour step changes for model and measurement is visualised in Fig. 8 while Table 3 shows a comparison of extreme step change values. In general the match is good, but there is a weak tendency of the model to underestimate the one hour changes. It should be noted that the maximum and minimum step change values are very sensitive to potential errors in the production data from the Swedish TSO. The histogram of hourly energy (Fig. 9) shows an excellent fit between model and observation. The monthly capacity factor, which could be of interest in studies of e.g. long-term hydro power reserves, is also well reproduced (see Fig. 10).

4.1. Comparison with a more basic model

As was shown in the section above, the results from the proposed model show good agreement with observations on an aggregated level. However, if similar results can be achieved with a simpler model this would be preferable. We therefore set up the following model for comparison (only for SE):

1. MERRA hourly wind speed for the 50 m level was bilinearly interpolated to the turbine positions.
2. The mean wind speed for each site was adjusted in order to give the annual energy production calculated by the wind turbine owner. A generic power curve (based on Vestas V90 2 MW) was used for these calculations.
3. Hourly production for each turbine was calculated using the same power curve.
4. The aggregated production was adjusted with a constant based on the observed bias during the optimisation years 2007, 2009 and 2011.

Without using the last step (bias correction), the RMSE in hourly data was 6.4%, i.e. 67% higher than for the proposed model. With bias correction the simple model had 8% larger error in hourly energy and 3–4% larger error in one and four hours step changes. The RMS error in the histogram was 28% higher. For duration curves, both for hourly energy and for one and four hours step changes, the RMS errors were substantially higher; 75, 34 and 147% respectively. The simple model performs particularly poorly in the upper end of the duration curve; maximum hourly energy was only 0.72 p.u. compared to 0.82 p.u. in observations. The RMS error in monthly capacity factor, finally, was 31% higher than for the more advanced model.

#### 4.2. Parameter influence

In this section the influence from parameter values and bias corrections on the model performance is presented. As mentioned above, a multitude of parameter sets with similar overall performance resulted from the optimisation procedure. For some parameters, good results could be found in the entire range while other parameters had a more direct (first-order) impact on the objective function (OF). Parameters and bias corrections are presented in descending order of importance (for SE).

The introduction of seasonal and diurnal bias correction significantly improved the results. For SE the OF value decreased with 30% and the RMS value of the error in hourly energy decreased with 6%. For the separate areas, the importance of seasonal and diurnal bias correction varied greatly; for SE4 there was actually no reduction at all in the OF while for SE2 there was an improvement of 21%. The use of time-varying external losses reduced the OF with 21% relative using an optimised but constant value. In SE3 and SE4 the reductions were 15% and 7% respectively. For SE2 the OF increased with 24%, see Section 5 for a discussion.

Although the smoothing effect could not be calculated on an individual wind farm level, the introduction of a global standard deviation parameter, applied on the power curves, considerably

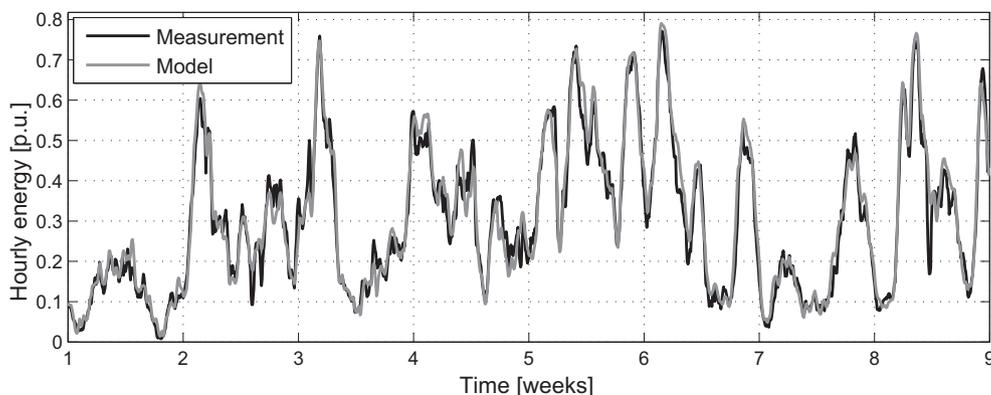


Fig. 6. Model output and validation data for eight weeks. RMS error for the period is 3.9%, i.e. almost the same as the average for all three years of validation.

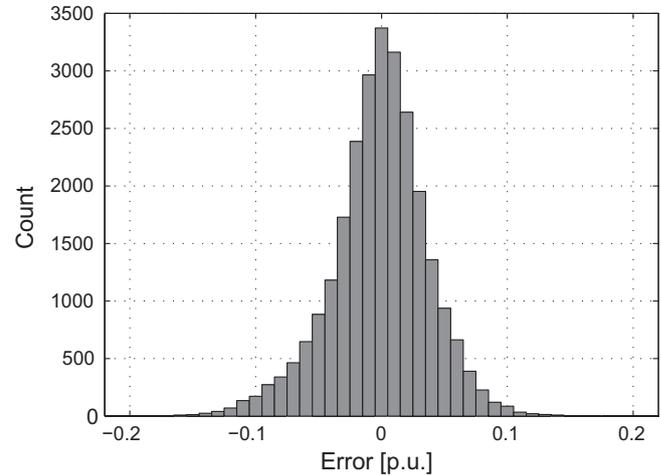


Fig. 7. Distribution of errors, i.e. the difference between model and observation.

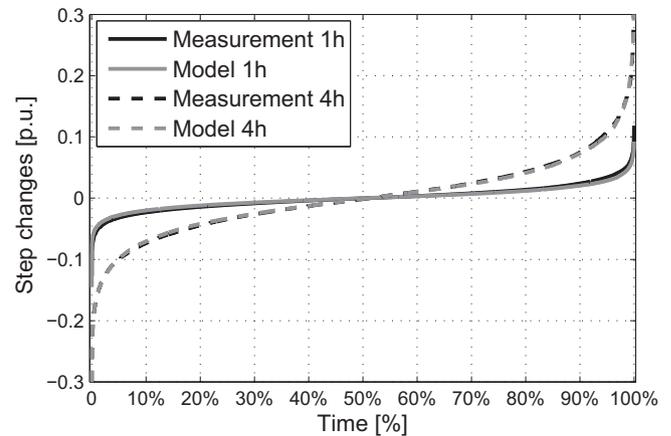


Fig. 8. Energy step change (1 and 4 h).

improved the results. For the SE and SE2 area the OF reduction was around 17% for optimum value of  $\sigma_u$  as compared to not using smoothing at all. For SE3 and SE4 the reductions were around 30%. As can be seen in Fig. 11, the optimum value of  $\sigma_u$  was around 1.0 m/s. For SE2–4 the optima were in the range 0.8–1.3 m/s.

It has already been mentioned that the first dataset (see Section 2.2) yielded markedly better results than the second. Since the second dataset was abandoned in a relatively early stage, the effect

**Table 3**  
Extreme step change values in p.u.

	1 h	4 h
	Measurement/model	Measurement/model
Max	0.118/0.092	0.390/0.341
Min	-0.127/-0.148	-0.320/-0.314
Max percentile	0.055/0.049	0.177/0.176
Min percentile	-0.052/-0.046	-0.159/-0.158

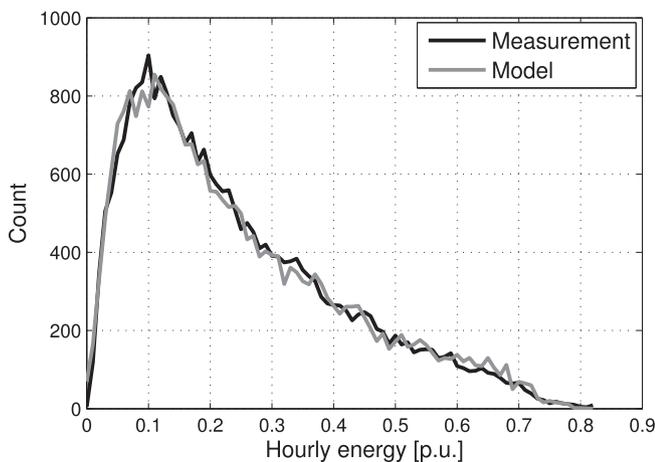
on the OF cannot be quantified. It is however strongly believed that the quality of the WEC data is one of the more important factors for successful modelling. Direction dependent internal losses contributed with a relatively small improvement in the OF: 6% for SE and 1–9% for the separate areas as compared to using the same loss in all directions.

The value of  $Loss_w$ , i.e. the loss applied when calculating mean wind speed, seem to have very little importance for SE and SE4. For SE2 and SE3 there was however a tendency that losses in the lower part of the allowed range gave poorer performance (around 10% increase in OF). The average values of internal losses of good parameter sets were found in the whole range, but they were strongly correlated to  $Loss_w$ . This gave the effect that most good parameter sets had similar optimised external losses (around 23% in the beginning of the period and 14% in the end for SE).

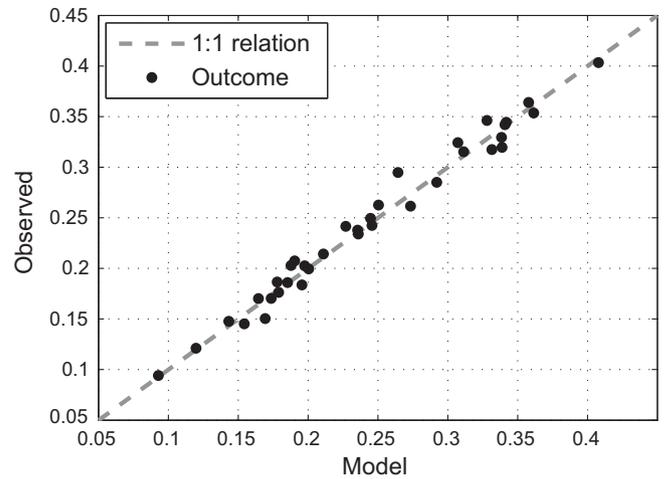
Using time-varying values of the wind shear and air density did not prove very successful. Although there are good physical reasons to believe that a more detailed modelling of these variables should enhance the performance, in practice the effect was negligible. Comparing the best parameter sets with and without time-varying wind shear and air density gave differences in OF ranging from -1 to 1% and -3 to 3% respectively. Potential explanations for the failure can be found in the next section. A parameter with little impact on the SE model performance is also  $P_{A,unknown}$ . For the separate areas, especially for SE2, a higher assumed ratio between installed power and rotor gave slightly better performance.

**5. Discussion**

In a wind farm there are several factors influencing the shape of the power curves. Because of the lack of sufficient detailed data, these factors were represented by only a few, sometimes global, parameters/variables using the idea of transforming the power curves to functions of incoming energy. Losses of different kind (e.g. due to wakes, availability, icing, blade degradation, high wind



**Fig. 9.** Histogram calculated in bins of 0.01 p.u.

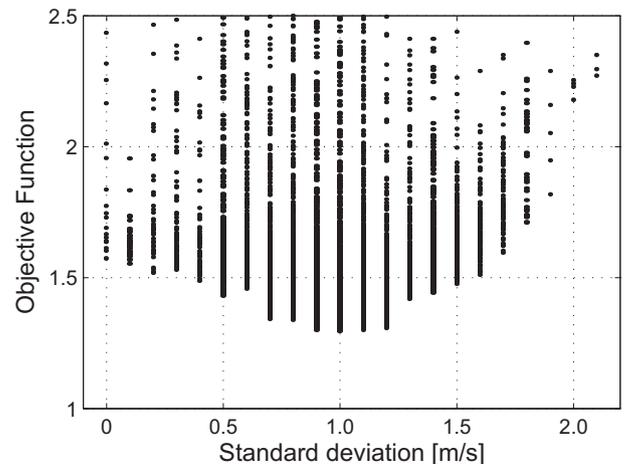


**Fig. 10.** Monthly capacity factor (36 months in evaluation period).

hysteresis and losses in transformer and internal electric grid) will have their own unique impact, and the question is if the simplified model could represent those to some extent. The most significant losses are due to wake effects from upstream turbines. There are several methods used to represent those [31], including the Katic/Jensen model

$$1 - \frac{V}{U} = \frac{(1 - \sqrt{1 - C_t})}{\left(1 + 2k \frac{X}{D}\right)^2}, \tag{6}$$

where  $V$  and  $U$  are disturbed and undisturbed wind speeds,  $C_t$  is the thrust coefficient,  $k$  is the wake decay constant,  $X$  is the distance between the two turbines and  $D$  is the turbine diameter. Since the thrust coefficient is largest for low wind speeds the reduction of wind speed using Equation (6) is also largest at low winds. Because of the shape of the power curve the losses will however be zero for undisturbed winds below cut-in, increasing to a maximum in the steepest part of the power curve and then again decrease towards zero for undisturbed wind speeds a little above rated. A very similar shape of wind speed dependent losses will be accomplished by reducing the incoming wind energy with a fixed percentage



**Fig. 11.** Objective function dependence on the parameter  $\sigma_u$ . The discrete nature of  $\sigma_u$  was chosen to increase computational speed in Matlab.

(controlled by  $Loss_{int}$  in equation (3)). The curve will however be shifted around 1 m/s so that losses are underestimated for low winds and overestimated for high wind speeds compared to the Katic/Jensen model. In Fig. 12 optimised internal losses for different wind sectors are presented. The results are given as averages for the endpoints of the 350 hill-climb searches. Losses are in average lowest for sector 1 and 2, while maximum losses are found in sector 3, 4 and 10 which are orthogonal to the prevailing sector. This result could mainly be explained by the fact that wind farms often have smaller separation distance between the WECs in sectors containing little energy. Another plausible explanation is that sectors with little energy receive a large share of the total energy from wind speeds in the range 5–10 m/s, i.e. exactly where the wake losses are highest.

An issue with the proposed model is the assumption that air density dependency could be described by using a single power vs. incoming power relation. This is equivalent to the calculation method used in IEC 61400-12 [32] where wind speed is corrected with a factor  $(\rho/1.225)^{1/3}$ . Comparison with density dependent power curves as well as previous work [33] however shows that for large deviations from standard air density this assumption does not hold very well in the approximate range 75–110% of rated wind speed. This might be one reason why the use of time-varying air density does not improve model performance. The most important factor explaining this lack of success is however thought to be the correlation of low air temperature (and thus high density) and icing losses; until icing losses are accounted for in the model it is unlikely to improve the results by using time-varying air density. After such a feature is implemented the performance could be further improved by using air density dependent power curves from manufacturers.

Almost all the best parameter sets for SE had time-varying shear exponents, although the improvement compared to using a fixed value was small. As mentioned earlier MERRA has a good correlation with measurements on an hourly scale. It has however been noted that MERRA and measurements have an extremely weak correlation in wind shear exponents and that the distribution of exponent value is much narrower for MERRA than for measurements. This mismatch is likely to have a relatively large effect on model performance, in particular for recent years when the bulk of the WECs have hub heights around 100 m and the interpolation error from 50 m can be expected to be large. Potential ways to

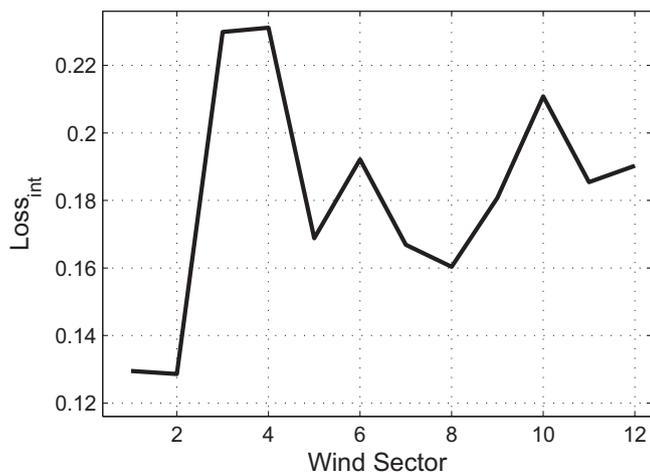


Fig. 12. Mean values of direction dependent  $Loss_{int}$  for the endpoints of the 350 hill-climb searches. Sector 1 is the prevailing sector.

alleviate this problem is to parameterise the wind shear using e.g. atmospheric stability or to evaluate other reanalysis datasets.

When comparing the results for Sweden to the ones for separate areas, it is obvious that parameters influence on model performance varies. For some parameters, clearer optima can be found for SE2-4 than for SE. Examples are the smoothing effect and the ratio between generator power and rotor area. For the systematic errors, both the time-dependent external losses and the seasonal/diurnal bias, the largest improvements were however seen for SE. This is likely because the bias is more robust for a larger area with more capacity. For SE2 e.g., the polynomial fit of the trend in losses is very sensitive to errors in the data and longer downtime periods for individual wind farms. For SE2 and SE4 it would therefore actually be better to use a linear function instead of a quadratic.

Based on the results from Sweden, which parameters should be included in future models for other countries or areas? It depends both on available WEC data quality and the size/capacity of the area. If more exact coordinates are available, wake losses and multi-turbine smoothing can be calculated hour by hour for each WEC. If coordinates are inexact however, a lot can still be gained by using a global smoothing parameter. When calculating the mean wind speed from estimated AEP it seems wise to use around 20% losses (i.e. a higher mean wind speeds will result as compared to using zero losses). It is however questionable if the parameterisation of this value can be justified. Seasonal and diurnal bias can obviously be significant, and we suggest that this should always be included in future models. A long-term trend in modelled versus observed aggregated production can also be important to take into account. Care should be taken, though, not to use a too high order function if the training period is short or the installed capacity is low. Introducing direction dependent losses can enhance the model performance slightly, and could therefore be considered. The use of a time-varying wind shear and air density is however hard to justify. Finally it is important to design the objective function with care. If e.g. good match of statistical distributions are of interest, then this must be reflected in the OF; optimising only for a low error in hourly energy does not automatically give the best fit in the distributions.

## 6. Conclusions

In this paper a model of hourly wind power production in Sweden is presented and evaluated. The model is based on MERRA reanalysis data and relatively detailed information on individual WECs. Overall, the model for the whole of Sweden had small errors in hourly energy and step changes and showed good ability to capture monthly capacity factors, histogram of hourly energy and distribution of step changes. The mean absolute error in hourly energy was 2.9% and the RMS error was 3.8%. In comparison with a more basic model, the proposed one performed better for all evaluated metrics; not very much when it comes to hourly errors but a great deal for the statistical distributions. The simulation of sub-areas was, as expected, more difficult and yielded larger errors.

Reduction in production was modelled by internal and external losses, where the former reduces the incoming energy in the wind and the latter reduces the aggregated produced energy. The power output was also multiplied with a correction factor for each hour and month due to observed systematic errors. The time-varying external losses and the correction term substantially reduced the model errors. The wind shear exponent and air density were allowed to vary from hour to hour, but this did not improve the results. A final conclusion is that smoothing the power curves by adding a standard deviation of the incoming wind of around 1 m/s was optimal.

## 7. Future work

The next step will be the development of scenarios of future wind power expansion in Sweden. These scenarios will be fed into the model to give information on future variability and the effect on the net load. It would be desirable to include the other Nordic countries in the model since these are tightly interconnected with the Swedish grid. One idea is also to try to simulate sub-hourly changes in power output using a statistical model.

Some possible improvements of the model have been identified. Firstly, modelling of hourly icing losses could be implemented. After this has been done, measured air density dependent power curves could be used. Finally, the model is expected to benefit from improved modelling of the time-dependent wind shear exponent.

## Acknowledgements

This work was conducted within the StandUP for Energy strategic research framework. The study was funded by the Swedish Civil Contingencies Agency (MSB) through contract 2010-2787. The authors would like to thank Olle Viotti and Per Edström (Sweco) and Matthias Mohr (Uppsala University) for valuable comments on the text.

## References

- [1] Billinton R, Chen H, Ghajar R. Time-series models for reliability evaluation of power systems including wind energy. *Microelectron Reliab* 1996;36(9): 1253–61. [http://dx.doi.org/10.1016/0026-2714\(95\)00154-9](http://dx.doi.org/10.1016/0026-2714(95)00154-9).
- [2] Li Y, Xie K, Hu B. Copula-ARMA model for multivariate wind speed and its applications in reliability assessment of generating systems. *J Electr Eng Technol* 2013;8(3):421–7. <http://dx.doi.org/10.5370/JEET.2013.8.3.421>.
- [3] Sturt A, Strbac G. Time series modelling of power output for large-scale wind fleets. *Wind Energy* 2011;14(8):953–66. <http://dx.doi.org/10.1002/we.459>.
- [4] Goić R, Krstulović J, Jakus D. Simulation of aggregate wind farm short-term production variations. *Renew Energy* 2010;35(11):2602–9. <http://dx.doi.org/10.1016/j.renene.2010.04.005>.
- [5] Papaefthymiou G, Klöckl B. MCMC for wind power simulation. *IEEE Trans Energy Convers* 2008;23:234–40. <http://dx.doi.org/10.1109/TEC.2007.914174>.
- [6] Cox J. Impact of intermittency: how wind variability could change the shape of the British and Irish electricity markets. Summary report, Tech. rep., Pöyry. Jul. 2009.
- [7] Papaefthymiou G, Kurowicka D. Using copulas for modeling stochastic dependence in power system uncertainty analysis. *IEEE Trans Power Syst* 2009;24(1):40–9. <http://dx.doi.org/10.1109/TPWRS.2008.2004728>.
- [8] Kubik ML, Coker PJ, Barlow JF, Hunt C. A study into the accuracy of using meteorological wind data to estimate turbine generation output. *Renew Energy*;51. <http://dx.doi.org/10.1016/j.renene.2012.08.084>.
- [9] Aigner T, Gjengedal T. Modelling wind power production based on numerical prediction models and wind speed measurements. In: 17th Power Systems Computation Conference, Stockholm, Sweden; 2011.
- [10] Dena. Dena grid study II – integration of renewable energy sources in the German power supply system from 2015–2020 with an outlook to 2025. Tech. rep. Berlin, Germany: Deutsche Energie - Agentur GmbH (dena); 2010
- [11] Hagspiel S, Papaemannouil A, Schmid M, Andersson G. Copula-based modeling of stochastic wind power in Europe and implications for the Swiss power grid. *Appl Energy* 2012;96:33. <http://dx.doi.org/10.1016/j.apenergy.2011.10.039>.
- [12] Magnusson M, Krieg R, Nord M, Bergström H. Effektvariationer av vindkraft, en studie av vindenergiproduktionens tidsvariation. Tech. Rep. Elforsk rapport 04:34. 2004.
- [13] Göransson L, Johnsson F. Cost-optimized allocation of wind power investments: a Nordic-German perspective. *Wind Energy* 2013;16(4):587–604. <http://dx.doi.org/10.1002/we.1517>.
- [14] NEWIS. New England wind integration study. Tech. rep., 2010
- [15] Kubik ML, Brayshaw DJ, Coker PJ, Barlow JF. Exploring the role of reanalysis data in simulating regional wind generation variability over northern Ireland. *Renew Energy* 2013;57:558–61. <http://dx.doi.org/10.1016/j.renene.2013.02.012>.
- [16] Foley AM, Leahy PG, Marvuglia A, McKeogh EJ. Current methods and advances in forecasting of wind power generation. *Renew Energy* 2012;37(1):1–8. <http://dx.doi.org/10.1016/j.renene.2011.05.033>.
- [17] Holttinen H. Hourly wind power variations in the Nordic countries. *Wind Energy* 2005;8(2):173–95. <http://dx.doi.org/10.1002/we.144>.
- [18] Olsson J, Skoglund L, Carlsson F, Bertling L. Future wind power production variations in the Swedish power system. In: Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES; 2010. p. 1–7. <http://dx.doi.org/10.1109/ISGTEUROPE.2010.5638945>.
- [19] Lucchesi R. File specification for MERRA products. GMAO office note no. 1 (version 2.3), Tech. rep. 2012.
- [20] [Online], Electricity statistics for Sweden. URL: <http://www.svk.se/Drift-och-marknad/Statistik/Elstatistik-for-hela-Sverige/>, [accessed 04.11.14].
- [21] Liléo S, Berge E, Undheim O, Klinkert R, Bredesen RE. Long-term correction of wind measurements – state-of-the-art, guidelines and future work. Tech. rep., Elforsk report 13:18. Jan. 2013.
- [22] [Online], MERRA Download. URL: [http://disc.sci.gsfc.nasa.gov/daac-bin/FTPSubset.pl?LOOKUPID\\_List=MAT1NXSLV](http://disc.sci.gsfc.nasa.gov/daac-bin/FTPSubset.pl?LOOKUPID_List=MAT1NXSLV), [accessed 10.02.13].
- [23] Scharff R. On distributed balancing of wind power forecast deviations in competitive power systems. Licentiate thesis. Stockholm: KTH; Oct. 2012.
- [24] [Online], Electricity certificate statistics. URL: <http://www.energimyndigheten.se/Foretag/Elcertifikat/Marknadsstatistik/>, [accessed 10.03.13].
- [25] [Online], Statistics on Swedish wind turbines. URL: <http://vindstat.com/>, [accessed 04.11.14].
- [26] [Online], Online planning tool for wind power permissions in Sweden. URL: <http://www.vindlov.se/sv/>, [accessed 10.04.13].
- [27] Nørgaard P, Holttinen H. A multi-turbine power curve approach. In: Proceedings of Nordic Wind Power Conference, Gothenburg, Sweden; 2004.
- [28] Emeis S. Wind energy meteorology. Springer; 2013.
- [29] Thøgersen ML, Motta M, Sørensen T, Nielsen P. Measure-correlate-predict methods: case studies and software implementation. In: EWEA Conference Proceedings. EMD International A/S; 2007.
- [30] Staffell I, Green R. How does wind farm performance decline with age? *Renew Energy* 2014;66:775–86. <http://dx.doi.org/10.1016/j.renene.2013.10.041>.
- [31] Duckworth A, Barthelmie R. Investigation and validation of wind turbine wake models. *Wind Eng* 2008;32(5):459–75. <http://dx.doi.org/10.1260/030952408786411912>.
- [32] Power performance measurements of electricity producing wind turbines. Tech. Rep. IEC 61400-12-1:2005(E). Geneva, Switzerland: International Electrotechnical Commission; 2005.
- [33] Svenningsen L. Proposal of an improved power curve correction, Poster EWEC 2010. Warsaw. Apr. 2010. URL: [http://proceedings.ewea.org/ewec2010/posters/PO.310\\_EWEC2010presentation.pdf](http://proceedings.ewea.org/ewec2010/posters/PO.310_EWEC2010presentation.pdf).