SCENARIOS AND TIME SERIES OF FUTURE WIND POWER PRODUCTION IN SWEDEN

REPORT 2015:141

WIND POWER

Energiforsk

Scenarios and time series of future wind power production in Sweden

JON OLAUSON, HANS BERGSTRÖM AND MIKAEL BERGKVIST, UPPSALA UNIVERSITY

ISBN 978-91-7673-141-3 | © 2015 ENERGIFORSK

Energiforsk AB | Phone: 08-677 25 30 | E-mail: kontakt@energiforsk.se | www.energiforsk.se

Authors' foreword

This project was conducted within the Vindforsk IV framework by researchers from the department of engineering sciences (division of electricity) and department of earth sciences (meteorology) at Uppsala University. The project comprises three main tasks:

1. Meteorologically based model of hourly wind power production
2. Scenarios of future wind power deployment (geographical and technical)
3. Statistical model of intra-hourly fluctuations

The first task built upon earlier work by two of the authors of this report. The second task relied heavily on interviews with knowledgeable persons in the Swedish wind community. The authors would like to express their gratitude for the valuable input from all interviewees. The project started in May 2014 and was concluded in May 2015.

The reference group for this project has been:

Jens Madsen, Vattenfall

Tobias Nylander, Vattenfall

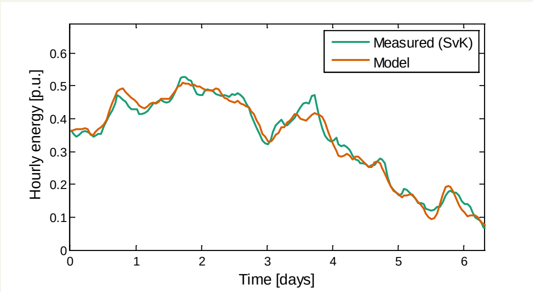
Johanna Lakso, Energimyndigheten

Erik Böhlmark, Svenska Kraftnät

Sammanfattning

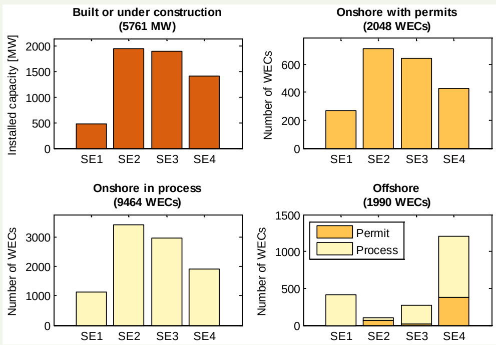
Omfattning och metodik i detta arbete sammanfattas nedan.

Modell för timvis aggregerad vindkraftsproduktion
Modellen baseras på information om individuella vindkraftverk samt data från den meteorologiska modellen MERRA. Eftersom modellen underskattade högfrekventa variationer i produktionen utvecklades en metod för kompensation av detta.



Scenarier av vindkraftsutbyggnad
Flera olika scenarier med en årsproduktion mellan 20 och 70 TWh togs fram. För att få realistiska antaganden om kapacitetsfaktorer, andel havsbaserad vindkraft etc., genomfördes intervjuer med sakkunniga personer.

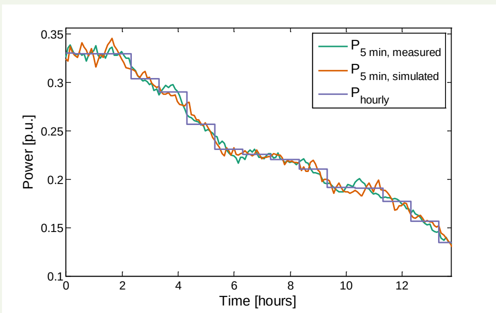
Scenarierna användes sedan som indata till den ovan beskrivna modellen och 36 år långa tidsserier skapades.



Vindkraftverk i den databas som användes som underlag för scenarierna. SE1-4 är elprisområdena i Sverige.

Modell av inomtimmen-fluktuationer
För vissa kraftsystemstudier kan produktionsdata med högre upplösning än en timme behövas. En statistisk modell för simulering av inomtimmen-variationer utvecklades därför.

Då mätningar av aggregerad vindkraftsproduktion i Sverige har upplösningen en timme användes data från kraftsystem i Tyskland och Danmark för att träna modellen.



Modellen för timvis aggregerad vindkraftsproduktion validerades med data från Svenska Kraftnät. Efter att den statistiska korrigeringen applicerats gav modellen både ett lågt fel i timvis produktion (medelfel strax under 3 % av installerad effekt) och en i det närmaste korrekt statistisk fördelning av såväl produktion som stegändring i produktion.

Med utgångspunkt i databasen av planerade projekt och antaganden om t.ex. andel havsbaserad vindkraft och förhöjd sannolikhet för projekt i riksintresseområde för vindkraft togs tre huvudscenarier fram med en årsproduktion av 20, 30 och 50 TWh (A1, B1 och C1). Inga explicita årtal sattes på scenarierna, men implicit antas 20 TWh uppnås ca år 2020 och 30 TWh tio år senare. Ett antal under-scenarier togs också fram med högre/lägre andel havsbaserade projekt, högre/lägre andel projekt i norr samt högre/lägre kapacitetsfaktor, vilket gav sammanlagt 23 olika scenarier. De två viktigaste faktorerna för den normaliserade variabiliteten (exempelvis uttryckt som standardavvikelse av en timmes stegändring i effekt eller spannet mellan första och sista percentilen i effekt) visade sig vara kapacitetsfaktor och andel havsbaserad vindkraft. Andra faktorer, exempelvis totalt installerad effekt, påverkade också men hade mindre betydelse. Ett variabilitetsindex, beräknat som ett medelvärde av ett flertal variabilitetsmått, var 0.98 för scenario A1, 0.88 för B1 och 0.80 för C1, vilket kan jämföras med 1 för modellerad produktion för vindkraftsflottan år 2015 och 1.15 för uppmätt produktion mellan 2007 och 2012. Med andra ord kan den relativa variabiliteten (normaliserat mot årlig produktion) förväntas minska i framtiden. Allt annat lika så skulle en geografiskt optimerad lokalisering av vindkraft kunna minska variabilitetsindexet med så mycket som 20-30 %.

Modellen för inomtimmen-variationer (fem respektive 15 minuters upplösning) tränades med data från Danmark och Tyskland eftersom dessa kraftsystem uppvisar en likartad karakteristik som det svenska i frekvensdomänen. Modellerna visade sig framgångsrikt kunna simulera såväl storlek på de snabbare variationerna som icke-stationäriteten hos dessa (vissa perioder är lugna medan andra är betydligt mer volatila).

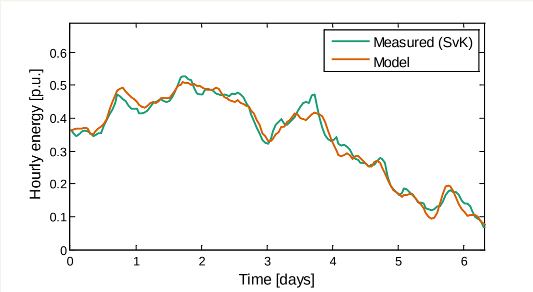
Även om en del analys av tidsserierna för de olika scenarierna presenteras (se framförallt bilaga A), så är huvudsyftet med detta arbete att producera och tillgängliggöra dessa serier för vidare studier. Tidsserier och metadata har därför lagts upp på en hemsida, se bilaga C för instruktioner för nedladdning.

Summary

The scope and methodology of this work is summarised in the graphical abstract below.

Model of hourly wind power production

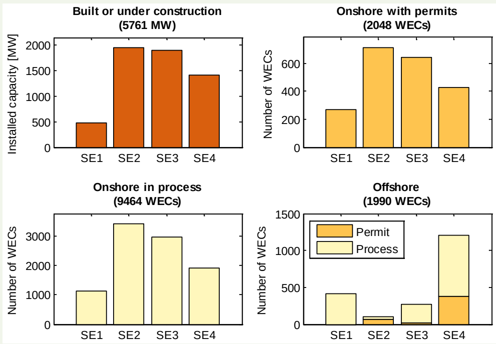
The model is based on data from the MERRA meteorological dataset and information on individual wind energy converters (WECs). A statistical improvement was also developed, compensating for missing high frequency variability in the original model.



Scenarios of WEC deployment

Several different scenarios of WEC deployment were developed, ranging from 20 to 70 TWh annual energy production. In order to have realistic assumptions on future capacity factors, share of offshore wind farms etc. interviews have been made with persons in the wind community.

The scenarios were used as input to the abovementioned model and 36 years long time series were generated.

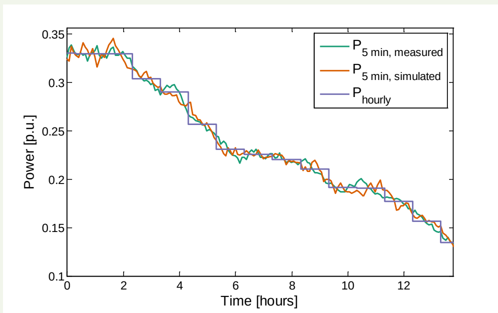


WECs in the database from which the scenarios were drawn. SE1-4 are the four electricity price areas in Sweden.

Model of intra-hourly fluctuations

For some studies it might be important to have time series of wind power production with higher resolution than one hour. A statistical model was therefore developed, simulating the intra-hourly fluctuations.

Since measurements of aggregated wind power production in Sweden has a temporal resolution of one hour, the model was trained on data from power systems in Denmark and Germany.



The model of hourly aggregated wind power production was validated with data from the Swedish Transmission System Operator (TSO). After applying the statistical correction, the model gave both low error in hourly production (mean absolute error was slightly below 3%) and a good representation of the statistical distribution of as well the hourly production as the step changes thereof.

With the database of planned projects as a starting point and assumptions on e.g. offshore wind power share and enhanced probability for farms in designated areas for wind power, three main scenarios were developed with an annual energy production of 20, 30 and 50 TWh (scenario A1, B1 and C1 respectively). No explicit dates were assigned to the scenarios, but implicitly is was assumed that 20 TWh will be reached around year 2020 and 30 TWh around ten years later. A number of sub-scenarios were also produced. Compared to the base scenarios, these are based on higher/lower share of offshore wind power, higher/lower share of capacity in the north and higher/lower capacity factors. In total 23 scenarios were developed. The two most important factors for the normalised variability (e.g. expressed as the standard deviation of one hour step changes or the inter-quantile range of production) proved to be the capacity factor and the share of offshore wind power. Other factors, such as the total annual energy production, had smaller impact. A variability index, calculated as an average of six normalised variability metrics, was 0.98 for scenario A1, 0.88 for B1 and 0.80 for C1, which should be compared to 1 for the modelled production of the current fleet (as of year 2015) of WECs and 1.15 for measured production years 2007 through 2012. In other words, the variability (normalised to annual energy production) is expected to be reduced in the future. It is also shown that by optimising the localisation of wind farms, the variability index could be reduced with as much as 20-30%.

The models for intra-hourly fluctuations (five and 15 minutes temporal resolution respectively) were trained on data from Denmark and Germany since these power systems exhibit similar characteristics to the Swedish system in the frequency domain. The models are shown to successfully simulate as well the magnitude as the non-stationarity of the faster fluctuations (some periods are calm while others are more volatile).

Although some analysis of the time series for the different scenarios are presented (mainly in Appendix A), the main purpose of this work is to produce and make available those series for further studies. The raw data has therefore been uploaded on a web page, see Appendix C for downloading instructions.

List of content

1 Introduction

This report has three main sections, dedicated to the model of hourly wind power production, development of scenarios and intra-hourly fluctuations respectively. The report is concluded with shorter sections comprising conclusions, discussion and future work. More detailed information on the scenarios are found in Appendix A. Appendix B is a reprint of a paper in the journal *Renewable Energy*, describing the model of hourly wind power production. Appendix C finally contains instructions for downloading the time series. This introduction section comprises a motivation of this work and some background on the Swedish power system and previous work in the area.

1.1 MOTIVATION

Wind energy converters (WECs) in Sweden are currently, as in many other countries, deployed at a fast pace. The variability and unpredictability of the power output from WECs, like that from other intermittent renewable energy sources, can be challenging. The magnitude of the integration problem and the costs related to it are conditioned on penetration level, meteorological conditions, WEC performance and the characteristics of the power system (e.g. hydro or thermal dominated). In order to study the future impact of wind power on the power system and to take proper actions in terms of e.g. grid reinforcements and procurement of reserves, production time series for projections/scenarios of future deployment and technology development are needed. It can also be relevant to study what factors are important for the resulting variability in the production. As will be shown, some parameters have small effect while others are very important. The results could be used as arguments to steer the development in a desired direction (i.e. a smaller impact on the grid). For all of these studies, long and meteorologically coupled time series with a relatively high temporal resolution are beneficial.

In wind integration studies, the time series of (aggregated) wind power production can be drawn from several sources, primarily upscaling of measured production, purely statistical models and finally physically based models, possibly in hybrid with some statistical fine-tuning. These methods have their pros and cons and are suitable in different situations. Upscaling is useful when a limited number of years are to be studied and the studied WEC fleet can be assumed to be relatively similar to the present in terms of installed capacity, geographical spread and technology used. The main advantage with this approach is the credibility of using actual measured data; no one will have any doubts that the data correctly represents the variability seen, at least for the past situation. For a Swedish example, see Ref. [1].

Purely statistical models (see e.g. Refs. [2] and [3]) have one primary advantage: their ability to generate multiple time series of an arbitrary length, and could find their use in e.g. reliability studies. The physical models excels when significant changes in the studied geographical area or WEC characteristics can be expected and can produce several decades long time series with a meteorological coupling. They also give a good possibility for parameter studies, i.e. how different factors influence the resulting variability. For credibility, meteorologically based models must prove their ability to adequately capture important characteristics of the system. Therefore, the output should be validated with actual aggregated wind power production. Some examples of recent research projects using physical methods to produce time series of wind power production are shown in Table 2 in Section 1.2.

1.2 BACKGROUND

As already mentioned, the installed wind power capacity has increased relatively fast in Sweden over the last years. At the end of 2014 the installed capacity was around 5400 MW, up from only 600 MW at the end of 2006. As will be shown in Section 3.3, the average hub height has increased significantly. This has enabled new geographical areas to be used; from the early deployment in open landscape, predominantly in the southern Sweden, most WECs are now installed in forested areas over the whole country. In 2014 wind power contributed to around 8% of the electricity consumption. Based on planned projects and available areas, there is however room for this figure to increase significantly. Sweden is part of the Nordic synchronous area. Relatively recently the country itself was divided into four bidding areas (price areas) in order to cope with transmission limitations. The produced time series in this report are given for these areas separately as well as for the whole of Sweden.

Several authors have produced scenarios and time series for future wind power production in Sweden. Carlsson [4] studied the future imbalance volumes and costs related to scenarios of 10 to 55 TWh annual energy production, based on planned wind power projects. The main focus was on the 10 TWh case, corresponding to 5 GW of installed capacity. This scenario illustrates the difficulties with predicting future deployment; although the projection was only a few years ahead (the report was published 2011, 5 GW installed capacity was reached in 2014), the mismatch is large, see Figure 1. The main explanations to the mismatch in this case are that an incomplete database with projects was used and that offshore deployment (in SE4) was overestimated.

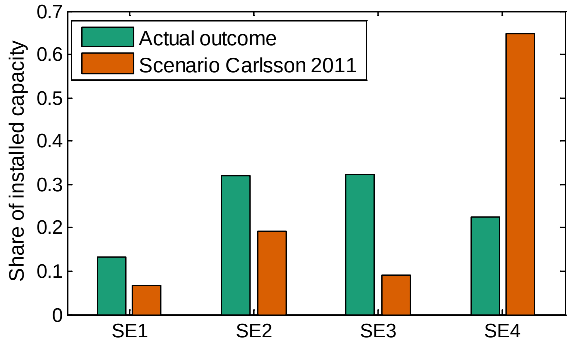


Figure 1. Comparison of installed capacity in scenario from Ref. [4] with actual outcome.

Magnusson et al [5] also focused on 10 TWh annual energy production (4 GW) in their scenarios from 2004. In the base scenario it was assumed that 75% of the installed capacity were offshore WECs, but in reality this figure was only around 3% at the time 4 GW was reached. The distribution among regions[1] in Sweden was however relatively accurate, especially for the scenario with more capacity in the north (see Figure 2). The ten year long time series produced in Ref. [5] has been used, sometimes in up-scaled form, in power system studies, see e.g. Refs. [6], [7]. When comparing to real production

[1] At this time there were no price areas in Sweden. The regions used in the report however correspond relatively well with today's price areas, especially for the two southern regions.

data, the maximum production was overestimated (96% of installed capacity compared to 83% in observations) and the energy in the high frequency variations were somewhat underestimated.

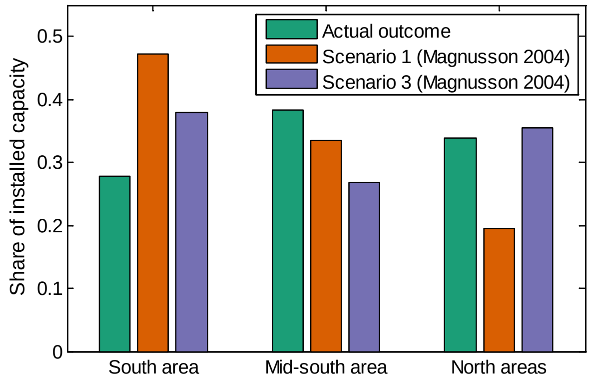


Figure 2. Comparison of installed capacity in scenarios from Ref. [5] with actual outcome. Scenario 1 is the base scenario and Scenario 3 is a scenario with more capacity in the north.

The Swedish TSO (Svenska Kraftnät, SvK) are using wind power scenarios for studies of the future power system. The scenarios are based on extrapolations of anticipated deployment from Svensk Vindenergi (the trade association for companies working with wind power in Sweden) and assumed capacity factors of 2800 full load hours for onshore WECs and 3500 full load hours for offshore WECs. The share of installed capacity in the different price areas are shown in Table 1.

Table 1. Share of installed capacity in different price areas in scenarios from SvK.

	Year 2020 (20.1 TWh/year)	Year 2030 (25 TWh/year)
SE1	11%	12%
SE2	34%	36%
SE3	30%	29%
SE4	24%	23%

A conclusion from the previous work with scenarios is that it is difficult to predict the future, also for a relatively short time period ahead. This is one of the reasons why several sub-scenarios are developed in the current work. We hope that the interviews with knowledgeable persons, the use of a database of planned projects and the studies of historical trends will contribute to realistic assumptions on geographical and technical development of Swedish wind power. At least, our intention is to be transparent with all assumptions and the parameter settings used.

This background section is concluded with a table showing some examples of methodology and datasets used for the production of time series of aggregated wind power output in recent research projects.

Table 2. Examples of research projects using physical models to produce time series of wind power production.

Country	Reference	Method	Temporal resolution	Validation
Sweden	Söder 2013 [6] Widén 2011 [7]	Meteorological model (MESAN), data from Ref. [5]	1h	Few farms and met. model with higher resolution
Sweden	Göransson & Johnsson 2013 [8]	Met. model (HIRLAM and NCEP/NCAR)	3/6h interpolated to 1h	No
Canada	Robitaille 2012 [9]	Met. measurements	1h interpolated to 1 minute	A few wind measurements
Germany	dena 2010 [10]	Met. model ("local model"/"local model Europe", DWD)	15 minutes	Large number of wind measurements and farms.
Switzerland (Europe)	Hagspiel et al 2012 [11]	Met. model (NCEP/NCAR) and statistical methods (copulas)	6h	Yearly production per country
USA, New England	NEWIS 2010 [12]	Met. model (NEWRAM)	10 minutes	Few wind measurements and farms.
Europe	TradeWind 2009 [13]	Met. model with added variability.	6h downscaled to 1h	Yearly energy for farms and hourly energy for western Denmark.
U.K.	Cannon et al 2015 [14]	Met. model (MERRA)	1h	Large number of wind measurements and TSO aggregated data

2 Model of hourly wind power production

Using meteorological models to simulate wind power production has several advantages over upscaling existing generation time series or using purely statistical methods. The most important of these is perhaps the possibility to adequately model future WEC fleets, including new types of WECs and deployment in new geographical areas. When comparing to upscaling the generated time series will also be significantly longer. For credibility, the physically based models needs to prove their ability by validation with historical aggregated generation data.

The model of hourly wind power production is based on information on WECs and reanalysis data from the meteorological dataset MERRA [15]. In Section 2.1 the underlying data, methodology and validation results for the model are described relatively briefly. A more detailed description can be found in Appendix B (reprint of journal article [16]). Based on the observed underestimation of fluctuations with periods ranging from two to ten hours, an improved model is suggested in Section 2.2.

For the rest of this report, average hourly power (sometimes expressed as hourly energy) is expressed in per unit (p.u.), where one p.u. corresponds to the installed capacity at the time.

2.1 "MERRA MODEL"

Data from 2007, 2009 and 2011 was used to train the model (i.e. tune the parameter values) and data from 2008, 2010 and 2012 was used for model evaluation. Information on WECs include coordinates, date of grid connection and possibly disconnection, installed capacity, rotor area and estimated annual energy production. Unfortunately there is no available, complete and reliable database with all necessary information on operating WECs, so data had to be combined from different sources, primarily Refs. [17] and [18], and manually corrected. In order to validate the model, data from SvK, the Swedish TSO, was used [19].

There are several global reanalysis datasets from meteorological models available. During the last years the performance of these have improved considerably. The MERRA dataset was chosen since it has a relatively high temporal and spatial resolution (one hour averages and 0.5×0.67° respectively) and has shown a good correlation with wind measurements at relevant heights; correlation coefficients are around 0.85 on an hourly basis and 0.94 on a monthly basis for measurements in terrain with low complexity [20]. The resolution of the MERRA model, as well as WECs built up to December 2012 and the four Swedish price areas, are shown in Figure 3.

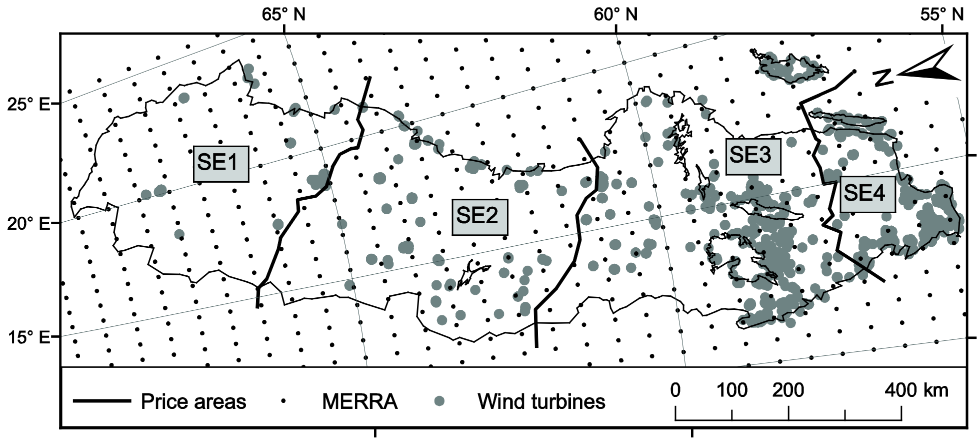


Figure 3. Visualisation of MERRA grid points, distribution of WECs (end of 2012) and the four price areas in Sweden. From Ref. [16].

The basic design of the model was:

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

A few factors contributing to the good results of the model were i) the mean wind speed for each WEC was determined by the anticipated annual energy production (as by the WEC owner) ii) Sweden had a relatively large number of WECs, spread over a large area, cancelling out errors in modelling of production for individual WECs iii) the model was calibrated using aggregated energy production time series (not wind measurements or production series from individual WECs as is often the case) iv) power curve smoothing was used and v) long-term trend as well as seasonal and diurnal bias were accounted for. It was however found that the use of time-varying air density and wind shear did not improve the model performance.

For the whole of Sweden the mean absolute error (MAE) was 2.9 % and the root mean square error (RMSE) was 3.8 % (related to the installed capacity). Correlation was 0.98. For the separate price areas the errors were higher and correlations were lower, especially for SE2. Eight weeks of actual and modelled production for the whole of Sweden are shown in Figure 4. Duration curves for one and four hour step changes are shown in Figure 5.

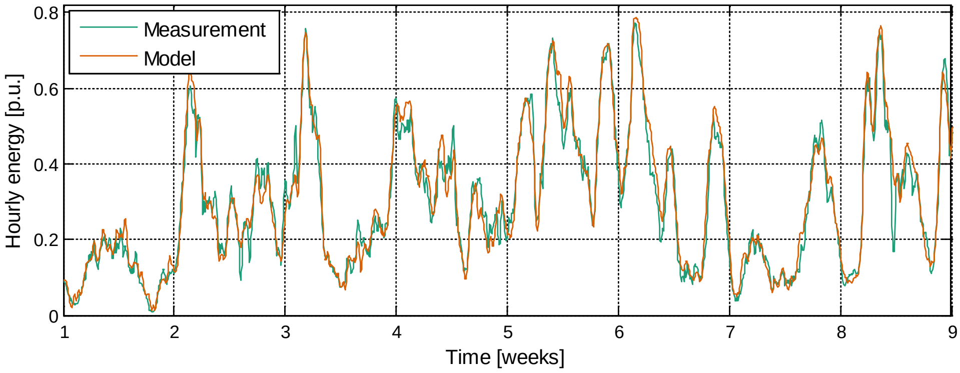


Figure 4. Model output and validation data for eight weeks. RMS error for the period is 3.9%, i.e. almost the same as the average for all three years of validation. From Ref. [16].

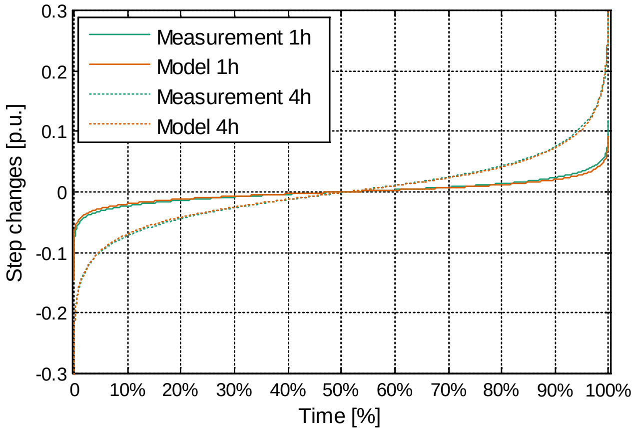


Figure 5. Duration curves for step changes (1 and 4 h) in hourly energy. From Ref. [16].

For the generation of time series of hourly production in this report, the following changes in methodology and parameter settings were performed as compared to Ref. [16]:

1. The whole period 2007 to 2012 was used to tune the parameters.
2. The term $f_{\Delta P}$ in the objective function (Eq. 5 in Ref. [16]) was omitted.
3. Time-varying air density and wind shear were not used.
4. WECs in smaller farms were aggregated to reduce computational time. Larger farms were split into four groups[2].
5. The optimized model for the whole of Sweden was used also for the separate price areas[3].
6. For future WECs producing more than 4000 full load hours (primarily in offshore farms), cut-out wind speed was increased from 25 to 30 m/s.

[2] The results using aggregated data were in principle identical as those using individual WECs. Individual coordinates will however still be available in the scenarios since this might be useful for future comparisons using more highly resolved meteorological data.

[3] When optimising and evaluating a model for a particular price area using the same years, the results were better than using the model for the whole of Sweden (SE). However, when using different optimisation and evaluation years the SE model was similar in performance. It was thus concluded that the SE model is more robust. An advantage with using the SE model is also that the contributions from the four price areas sum up to the production for the whole of Sweden.

It was also analysed weather the results could be improved by using wind data from the MESAN model [21] instead of MERRA data. The main advantage with MESAN is the higher spatial resolution; 11×11 km as compared to 56×(27-42) km for MERRA. It was however found that MESAN spans a much shorter time period and that the grid size was not homogeneous in time. Therefore MERRA data was used also in the present work.

2.2 IMPROVED MODEL

Although the errors in hourly production were low with the model described in Section 2.1, the produced time series lacked energy in the high-frequency variations. When comparing power spectral densities (PSD) of output from the model and from actual wind power production, a mismatch for frequencies above ca 0.1 hours^{-1} can be noticed, see Figure 6. One effect of the missing variability is that the step changes in hourly energy production will be somewhat underestimated by the model. The standard deviation of one hour step changes was for instance underestimated by 14%. This was our main incentive to develop a statistical model to restore the missing variability. For more details on the improved model, see Ref. [22].

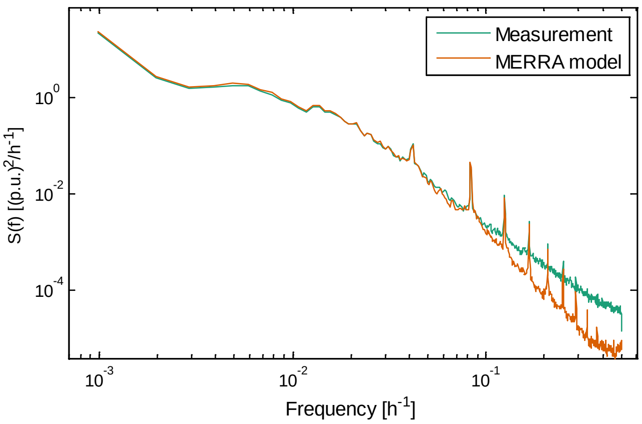


Figure 6. Comparison of the power spectral densities (PSD) of aggregated power for measurement and the MERRA model described in Section 2.1.

Throughout the rest of this section, when referring to the "MERRA model" we mean the model described in Section 2.1, i.e. a model that uses MERRA reanalysis data and information on Swedish WECs to calculate the aggregated, hourly wind power production. With "Measurement" or "SvK" we mean actual, hourly wind power production measured by the Swedish TSO. With "Improved model" we are referring to the MERRA model combined with the statistical correction described in this section.

One straightforward solution to the problem would be to add simulated noise in order to fill the gap in the PSD curve. We however notice that the error in volatility of the hourly energy is non-stationary; sometimes the MERRA model data gives an accurate description of the fluctuations, but during many periods there is much more volatility in the measured data. We therefore look for a statistical correction to increase the high-frequency fluctuations, taking into account this non-stationary behaviour.

2.2.1 Method

To put it shortly, the idea is to mimic the high-frequency variations in measured data, conditioned on hourly energy values and its derivative as well as seasonal and diurnal patterns, and to add the simulated noise to low-pass filtered time series from the MERRA model. The method involves the following steps:

1. Separate hourly data from SvK and the MERRA model into their low frequency and high frequency components, e.g. SvK_{HF} and $MERRA_{LF}$.
2. Transform the SvK_{HF} data to an approximately stationary time series y_{SvK}.
3. Find the magnitudes of the frequency domain representation of y_{SvK} using FFT (Fast Fourier Transform).
4. Generate y_{sim} of the same length as the MERRA time series, using inverse FFT with interpolated magnitudes from 3) and random phases.
5. Transform y_{sim} to $MERRA_{HF,sim}$ using the inverse of the transformation used in 2).
6. The time series for the improved model, $MERRA_{sim}$, is achieved as $MERRA_{LF}$ + $MERRA_{HF,sim}$.

For the separation of frequencies a windowed sinc filter with a 5000 samples long kernel was used. The cutoff frequency was set to 0.1 hours^{-1}. An example of the separation of the signal into its high and low frequency components is shown in Figure 7.

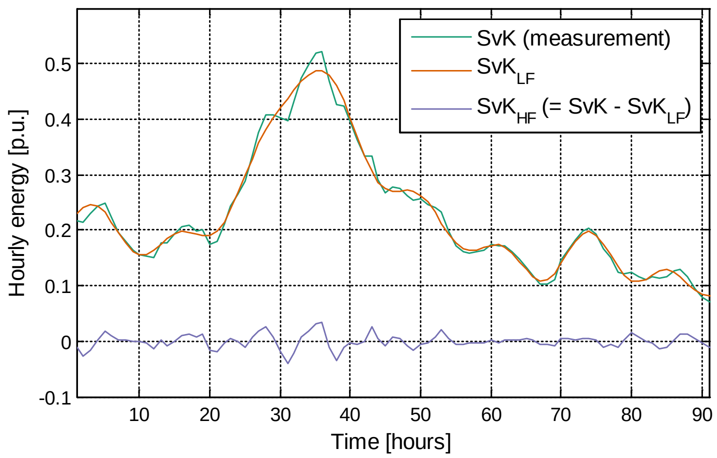


Figure 7. Example of separation of a signal into its low and high frequency components (subscript LF and HF respectively). It is the SvK_{HF} signal that should be modelled and subsequently added to $MERRA_{LF}$.

The volatility of the high frequency signal vary in time and thus needs to be transformed to an approximately stationary time series. We observed that a large part of the variations in volatility can be explained by the (instantaneous) value of the low passed filtered energy production and the derivative thereof. For each coordinate in the hourly energy – energy derivative space, the volatility of the high frequency component should be estimated. This was done by collecting the closest samples and fitting a t location-scale distribution. This distribution resembles the normal distribution, but have an additional parameter specifying the heaviness of the tails. The scale parameters, corresponding to the standard deviation in the normal distribution, of the fits for different coordinates are shown in Figure 8. Generally a low energy production and low derivative give a mildly fluctuating high-frequency signal, while the highest fluctuations were found when energy production is in the intermediate range and the derivative is high.

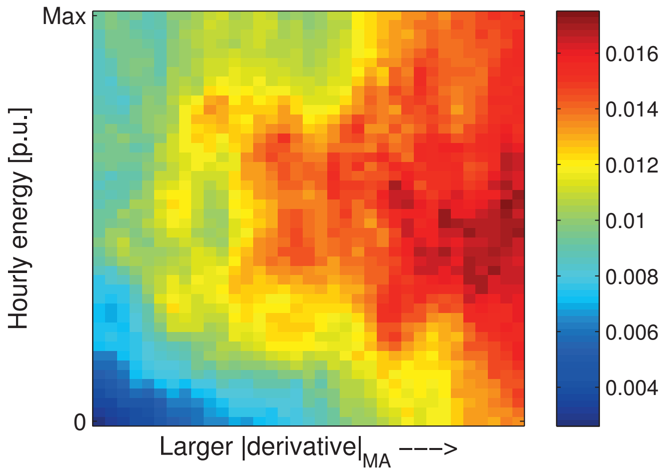


Figure 8. The t location-scale distribution scale parameter (σ) as a function of SvK_{LF} and its derivative. For the derivative, a 15 hours moving average (MA) of the absolute value was used.

All high frequency data were transformed (mapped) to a zero-mean normal distribution with variance one: $SvK_{HF} \rightarrow y_{SvK,0}$. After removing the fluctuations' dependence on hourly energy and its derivative, a seasonal and diurnal pattern still existed in the resulting time series. The transformed high frequency signal had non-zero mean values related to particular months of the year and time of the day. The mean values followed a quasi-sinusoidal pattern with three of four periods per day. After removing the seasonal/diurnal bias, an approximately stationary time series (y_{SvK}) with $N(0,1)$ distribution resulted.

The next step was to find the frequency domain representation of y_{SvK} using the FFT (Fast Fourier Transform) algorithm. By using interpolated magnitudes from y_{SvK} and random phases, simulated time series y_{sim} of desired length could be generated with the inverse FFT algorithm. In order to get real valued time series the magnitudes and phases in the spectra were made symmetrical. The DC-component was forced to be zero.

Finally y_{sim} was converted to non-stationary high frequency noise with the inverse seasonal/diurnal and t locations-scale transformations and the noise was added to the low-pass filtered time series from the MERRA model. A few samples of the resulting series were below zero or above the maximum of the original time series. These were simply changed to zero and the maximum value respectively.

2.2.2 Results

In this section, output from the improved model (MERRA model with added noise) are compared to measurements from the Swedish TSO. The results are given for the evaluation years 2008, 2010 and 2012.

An important question is whether the magnitudes of the simulated high-frequency fluctuations have a correct temporal pattern. Remember that the desire to correctly capture this pattern was the reason to choose a more complex method and not simply add (stationary) noise corresponding to the missing energy in the PSD. Figure 9 shows that the model to a relative large extent can capture the varying magnitude of the fluctuations, e.g. the calm periods during week 5 and 6 are present also in the model.

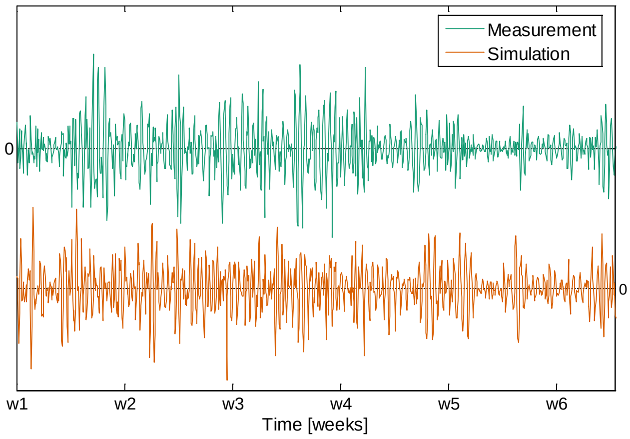


Figure 9. Measured and simulated high frequency component over a couple of weeks (the vertical axes has the same scale but are shifted for better readability). The model is able to capture some of the trend of the magnitude of the fluctuations, but the match is not perfect. The period shown was chosen since it is representative with respect to its correlation in 24-hour moving average of the absolute fluctuation magnitudes.

We now move to the resulting time series of actual hourly energy production from the improved model. Figure 10 gives an example on how the short-term fluctuations have increased with the filtering and addition of noise; the improved model and measurement now have similar smoothness. This effect is however much more obvious in the PSD plot shown in Figure 11. With the raw MERRA model the PSD starts to deviate from measurement PSD at around 0.1 hours^{-1} and was significantly lower above 0.2 hours^{-1}, but with the improved model there is an almost perfect match.

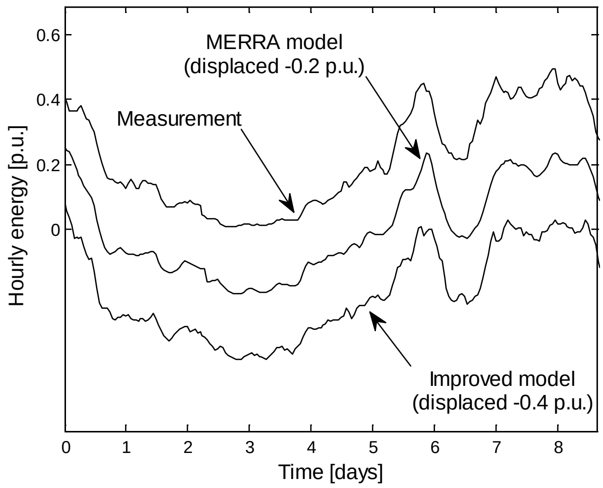


Figure 10. Hourly energy time series of measurement and models. The time series from the MERRA model appears to be too smooth while the output from the improved model has similar smoothness as the measurement. This is confirmed by the spectra shown in Figure 11.

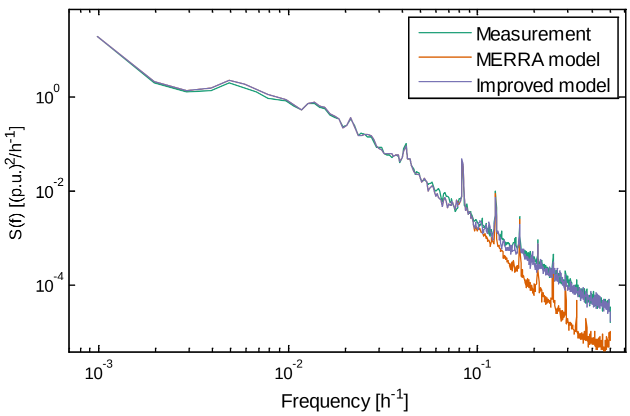


Figure 11. PSDs for measurement, the MERRA model and the improved model for the validation years (2008, 2010 and 2012).

The main motive to develop the statistical model was to correct for the observed underestimation in one hour step changes in the MERRA model. Thus Figure 12 contain the most important results of this section; with the improved model the duration curve of one hour step changes closely follow that for measurements. Naturally, low-pass filtering and adding noise to a signal will increase the error. The error increment was however acceptable, only 0.14 %-units (RMSE up from 3.7% to 3.9% of installed capacity).

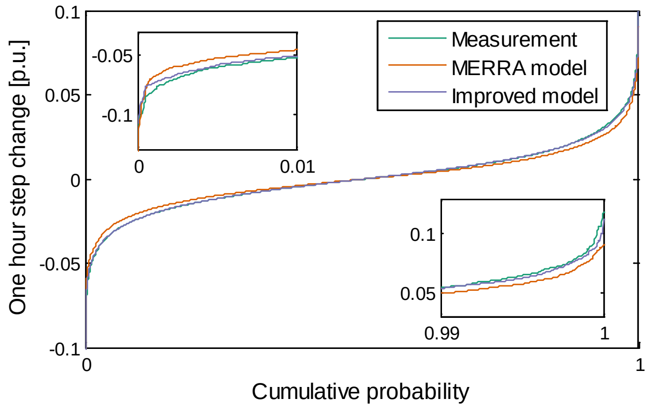


Figure 12. The figure shows the improvement in duration curve for one hour step change (with the tails as insets). Only the curves for one hour step changes are shown, since the four hour curve of the MERRA model was already in good agreement with measurements.

2.2.3 Application on scenarios

The time series for the scenarios have different PSDs, depending on e.g. geographical area, installed capacity and capacity factor. The generated high-frequency noise was therefore multiplied with a scaling constant in order to get a smooth transition between frequencies below and above cut-off frequency (0.1 hours^{-1}). The scaling constants were chosen so that the average ratio of the PSDs for $MERRA_{sim}$ and SvK for frequencies above 0.1 hours^{-1} was the same as the average ratio of PSDs for $MERRA$ and SvK between 0.04 to 0.1 hours^{-1}.

In some cases the simulated high frequency time series had a few samples with extreme magnitude. Therefore a limit on $MERRA_{HF,sim}$ was set. The limit was based on the observed ratios of maximum value over 99.9th percentile and minimum value over 0.1th percentile. For a longer time series (e.g. 36 simulated years compared to six in the training dataset), this ratio is expected to be slightly higher. This was accounted for, see Figure 32 in Section 4.3 for a demonstration of the methodology used.

Improved time series for the four different price areas were generated independently and added together to produce the time series for the whole of Sweden.

3 Scenarios

In this section the underlying data and methodology for developing scenarios are described. More details for all scenarios can be found in Appendix A. Guidelines for accessing the time series are presented in Appendix C.

A scenario of future WECs require in principle two components: coordinates and characteristics of the turbines. The general methodology used in this work was to build a large database with current and planned wind farms and to interview different actors in the wind power sector in Sweden. Wind farms in the database were given different probability weights based on e.g. wind conditions, geographical location and whether the farm has all necessary permits in place or not. Realisations of scenarios were finally randomly drawn from the database given the desired total annual energy production from onshore and offshore turbines. No explicit date were put on the scenarios, but implicitly it was assumed that 20 TWh annual production is reached around year 2020; all WECs built year 2000 or later were kept in their present form in the 20 TWh scenarios. For the 30 TWh scenarios, WECs built 2010 or later were kept. The 50-70 TWh scenarios are for a more faraway future and no current WECs were kept. They were however added to the "repower" category in the database and could therefore reappear with upgraded performance in the scenarios.

This section starts with a description of the database of wind farms. Subsequently the methodology for the "scenario generator" is described in some more detail. Section 3.3 and 3.4 comprises historical trends of WECs in Sweden and results from parameter studies respectively. The section is concluded with tables of the chosen parameter values for the different scenarios and some example results for one of the scenarios.

3.1 DATABASE OF WIND FARMS

For information on already built WECs, the database developed in Ref. [16] was re-used and extended with WECs built in 2013 and early 2014. Information on projects under construction (firm and unconditional order) were received from Svensk Vindenergi[4] in March 2014. In total, the capacity for WECs already built or under construction is 5761 MW, which correspond well with the 5830 MW given in Ref. [23].

Information on wind farms with all permits ready (but no official decision for construction start) and in the planning process was received primarily from Vindbrukskollen[5] and Havsvind.se [24]. These two categories will be referred to as "permit" and "process" in this report. The database was reviewed manually to correct for errors, duplicates and unrealistically dense park layouts[6], which resulted in the removal of around 1800 WECs. For the bulk of the wind farms, coordinates for the individual WECs were available. When this was not the case, the WECs were placed in an array around the centre coordinate of the farm. The capacity and distribution of WECs are shown in Figure 13 and Figure 14.

[4] "Swedish Wind Energy", trade association for companies working with wind power in Sweden.
[5] Email from Suvi Hänninen, Vindbrukskollen (Länsstyrelsen, Västra Götalands län), 2013-10-28
[6] The latter was only reviewed for large farms.

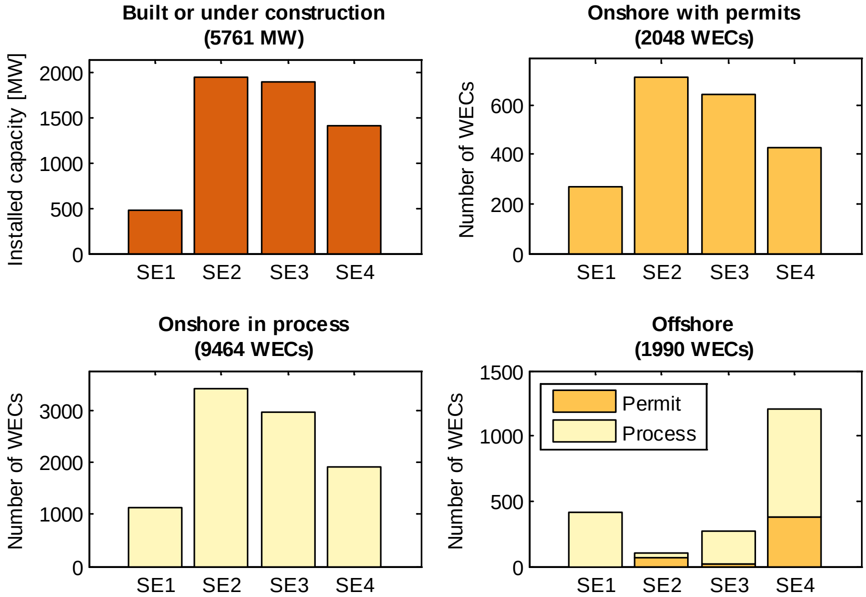


Figure 13. WECs in the database, sorted by type and price area. The geographical extent of the price areas are shown in Figure 14 (SE1 in north to SE4 in south).

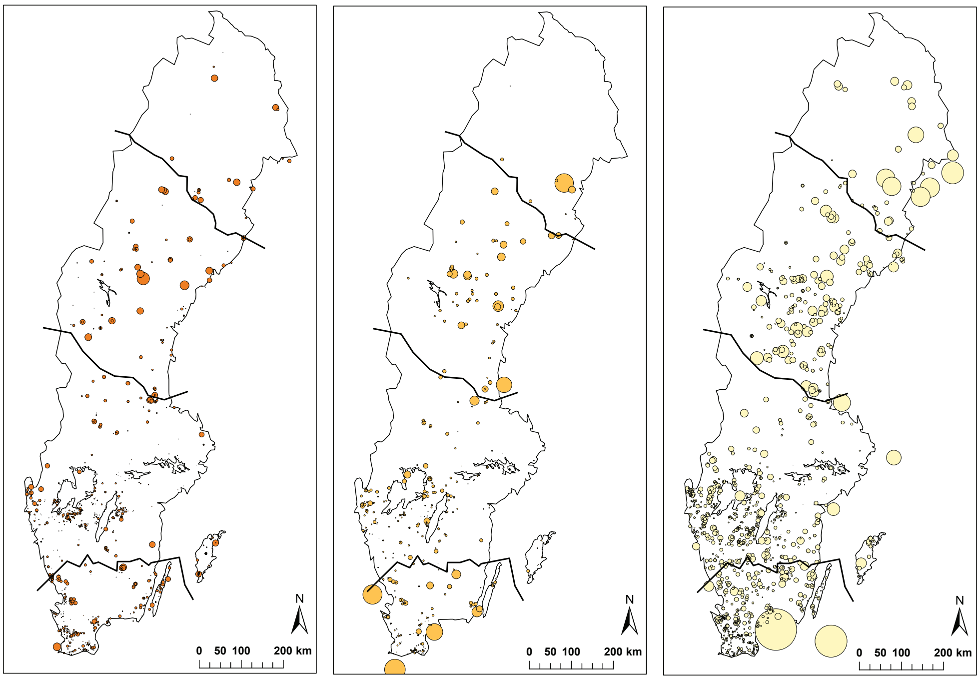


Figure 14. Wind farms in the database (built or under construction, with permits and in planning process respectively). The circle areas are proportional to capacity. The largest circles in each figure corresponds to 270, 768 and 3000 MW respectively.

For WECs already built or under construction, the installed capacity and estimated annual energy yield as given by the farm owner was used. For WECs in the "permit" and "process" category, the available information was however deemed too uncertain, especially for projects that will be realised decades from now. These variables were instead determined by chosen parameter values in the "scenario generator", see Section 3.2.

3.2 METHODOLOGY

A scenario generator was programmed in Matlab which randomly select wind farms from the database described in Section 3.1. Farms with a total capacity below 1 MW was however not considered. A number of parameters have to be set by the user, governing the total annual energy production for the scenario, characteristics of the WECs and probability weights for the different farms. Before the final values of the parameters were set, a study was conducted to examine the impact from and relative importance of the different parameters, see Section 3.4. The study show that some of these are of virtually no importance for the resulting time series, and consequently less effort had to be put into finding realistic values for these parameters. Short explanations for all parameters are found in Table 3 and more detailed explanations are given in the subsequent paragraphs.

Table 3. Parameters in the model.

Parameter	Short explanation
$TWh_{onshore}$	Annual energy production from onshore WECs [TWh].
$TWh_{offshore}$	Annual energy production from offshore WECs [TWh].
Keep old	Keep existing WECs built this year or later. Older WECs might be repowered (see below).
$Flh_{onshore}$	Average full load hours for new onshore WECs [hours].
$Flh_{offshore}$	Average full load hours for new offshore WECs [hours].
$P/A_{onshore}$	Rated power per swept rotor area [W/m²].
$P/A_{offshore}$	Rated power per swept rotor area [W/m²].
SD flh	Size of random component in flh for each farm [hours].
Wind impact 1	Impact from mean wind speed on farm full load hours (see explanation below).
Wind impact 2	Impact from mean wind speed on probability weight (1 = linear, 2 = quadratic etc.).
Separation	Separation distance for calculation of onshore WEC capacity [rotor diameters].
Max P	Maximum capacity for onshore WECs [kW].
$P_{unknown}$	Capacity for onshore WECs when no coordinates are known [kW].
Park share 1	Share of WECs in farms with permit actually built.
Park share 2	Share of WECs in farms in process actually built.
Weight north	Probability weight for farms in north (SE1 and SE2) relative farms in south.
Weight process	Probability weight for farms in process relative farms with permit.
RI wind 1	Probability weight for farms with 1-50% of WECs in designated area for wind power.
RI wind 2	Probability weight for farms with >50% of WECs in designated area for wind power.
Repower 1	Probability weight for repowering farms relative farms with permit.
Repower 2	New capacity of repowered farms (relative old capacity).
Aggregate	Yes for aggregating WECs in a farm to reduce computational time.

The impact from the mean wind speed on the probability weight was calculated according to $w_{wind} = \bar{u}^{a}$, where w_{wind} is the probability weight due to mean wind speed,

$\bar{u}$ is mean wind speed according to the MIUU model[7] and a is the *Wind impact 2* parameter. The number of full load hours for each onshore farm was calculated as

$$flh = flh_{mean} + (\bar{u} - \bar{u}_{mean}) \cdot b + e, \quad (1)$$

where e is normally distributed noise with standard deviation according to *SD flh*, subscript *mean* stands for the mean of all new onshore WECs in the scenario and b is the *Wind impact 1* parameter[8]. The mean wind speed for each WEC was subsequently adjusted in order to give the desired capacity factor, given the power curve and wind speed distribution at site.

In the database of planned projects, the given (if any) rated capacity and annual energy yields was deemed too uncertain to be useful. The capacity factor was instead determined by equation (1). The installed capacity for a farm was calculated based on the median of the distance to the closest neighbouring WEC, assumed rated power per swept rotor area and assumed separation distance expressed in rotor diameters. The installed capacity was however never allowed to be larger than *Max P* for onshore WECs.

> **Example**: We have set the parameters *Separation, Max P and P/A$_{onshore}$* to 4 rotor diameters, 3500 kW and 280 W/m^2 respectively. A farm has median distance to the closest neighbour of 420 metres which gives a rotor diameter of 420/4 = 105 metres. The installed capacity for each WEC is thus
>
> $$\min\left[3500\,kW, 280\frac{W}{m^2}\cdot\left(\pi\frac{(105m)^2}{4}\right)\right] = 2425\,kW.$$

The reason to why the (for future WECs) seemingly low value of 3500 kW was chosen as maximum capacity was that the coordinates are given for present WECs; if the farms will be built in a more distance future with larger WECs, the farm layouts are expected to be less dense and the total farm capacity will thus be similar to that achieved with the abovementioned calculation method.

In Sweden there are designated areas for wind power ("riksintresse vindbruk"[9]), see Figure 15 for distribution over the different price areas. It is more likely that farms in these areas will be realised. This increased likelihood was governed by the two *RI wind* parameters; one for farms with below 50%

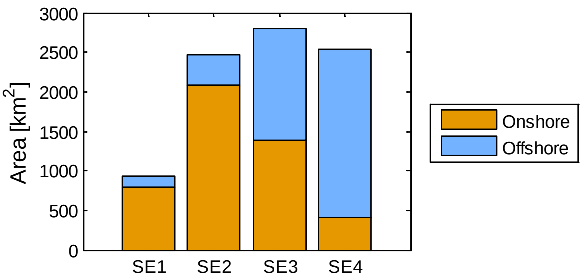


Figure 15. Designated areas for wind power ("riksintresse vindbruk"). The bar plot shows the distribution over the four price areas.

---

[7] MIUU is a mesoscale model developed at Uppsala University. Datasets of mean wind speed at different heights with spatial resolution of 500m and 1km are freely available for Sweden. See e.g. [25], [26].

[8] A calculation using the same WEC and hub height will result in larger difference between good and poor sites than using this equation with b = 500. For poorer sites it is however likely that the hub height will be higher and the rotor area larger for a given installed capacity and hence the difference is reduced. As will be shown in Section 3.4, the distribution of capacity factors among the farms is of minor importance; what really matters is the assumed mean capacity factor.

[9] https://www.energimyndigheten.se/Om-oss/Var-verksamhet/Framjande-av-vindkraft/Riksintresse-vindbruk-/

and one for farms with above 50% of the WECs in the designated areas.

Many of the current wind farms might very well be repowered after e.g. 20 years of operation. In the scenario generator this was handled by the parameters *Keep old*, *Repower 1* and *Repower 2*. WECs built before the year *Keep old* were moved to the database of possible new projects after an increase of the capacity with a factor of *Repower 1*. The repower projects were given a probability weight of *Repower 2* relative to farms with permit.

> **Example**: Farm A was constructed year 1995 and the parameter *Keep old* is set to year 2000, thus the farm is removed from the category of existing WECs and moved to the repower category. Mean wind speed is 6.8 m/s and the farms is located in southern Sweden. Farm B has all permits in place and is located in northern Sweden. Mean wind speed is 7.7 m/s and the entire farm is planned in a dedicated area for wind power. Using probability weights 0.9 for northern projects, 0.8 for repower projects, 1.5 for farms in dedicated areas for wind power and a linear wind impact, the relative probability weights for the two farms are calculated as
>
> $$w_A = 0.8 \cdot 6.8^1 = 5.4$$
>
> $$w_B = 7.7^1 \cdot 0.9 \cdot 1.5 = 10.4$$
>
> In other words it is roughly twice as likely that farm B will be added to the scenario compared to farm A.

Offshore farms were treated differently than onshore in the development of scenarios, partly because there are only 19 new projects in the database. The installed capacity of the farms were taken from Ref. [24] and from web pages for the developers. No probability weights were given to the farms, instead they were put in a manually ordered list. Factors influencing the order were whether the farms have all permits ready, distance to shore, water depth and a judgement of the "seriousness" of the developer. The majority of the more realistic offshore projects are located in the southern part of Sweden (predominantly in SE4).

The choice of parameter values were based on historical trends, interviews with persons in the Swedish wind power sector and educated guesses, see Section 3.5 for the chosen values. The main objective with the interviews was to get information and feedback on the assumptions needed for the development of scenarios of future WEC deployment and technology. The interviewees and persons in the reference group are listed in Table 4. The most important questions were:

- How many full load hours will future onshore and offshore WECs produce?
- Will we see a continued trend towards larger swept rotor area as compared to the rated capacity?
- Will offshore wind power become significant (and if so, when)?
- Is it less (or more) likely that projects in the north will be realised?[10]
- Will really large projects (e.g. >100 WECs) be realised?
- Will current wind farms be repowered?

[10] Sweden was relatively recently divided into four price areas to cope with congestion due to transmission bottlenecks. This might lead to lower electricity prices in the north. More severe icing problems might also lead to lower deployment in the north. At the other hand it is easier to get permission to build large farms in the north since this part of the country is (even) less densely populated. Generally it is also easier to get permits for higher WEC heights in the north.

- General comments on the scenarios, e.g. is it valuable with intra-hourly data and which levels on annual energy production are most relevant for the scenarios.

Table 4. Interviewees and members of the reference group (the latter are marked with an asterisk).

Organization	Persons
Vattenfall[11] (Power company)	Daniel Gustafsson, Roy Lilleberg, Tobias Nylander*, Jens Madsen* and members from a project team (Energy system 2050).
SvK (Swedish TSO)	Petter Glantz and Erik Böhlmark*
Svensk Vindenergi (Swedish Wind Energy)	Anton Steen
KTH (Royal Institute of Technology)	Lennart Söder
KVA (Royal Swedish Academy of Sciences)	Harry Frank
Sweco (Engineering consultants)	Olle Viotti and employees from the wind analysis and energy market groups.
Energimyndigheten (The Swedish Energy Agency)	Johanna Lakso*

3.3 HISTORICAL TRENDS

In this section the historical trends of rated power per swept rotor area, hub height, estimated and capacity factor are shown for Swedish WECs. The database is not complete for some of these variables; information on rotor diameter was available for WECs corresponding to 83% of the installed capacity and information on hub height was available for 69%. The information was mostly lacking for WECs built in the 90s and in the years 2014 and 2015. The section is concluded by a comparison of estimated capacity factor and mean wind speed from the MIUU model and some thoughts on offshore farms.

Figure 16 show the trends for rated power per swept rotor area and hub heights respectively. Obviously the rotor areas have become relatively larger and the hub heights have increased significantly over the years.

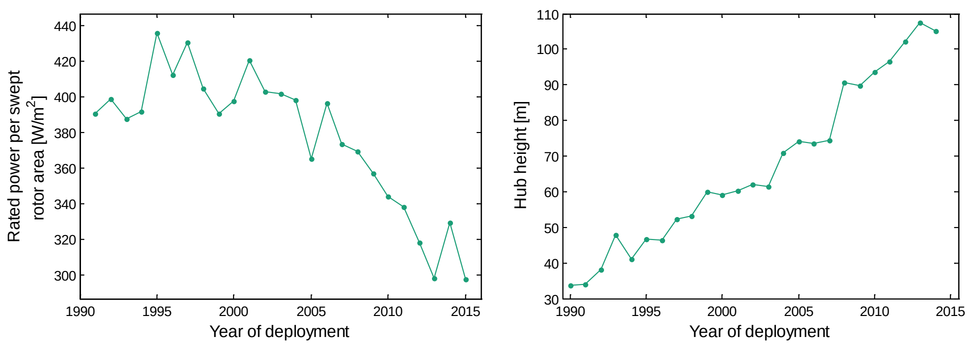


Figure 16. Historical, capacity weighted, trend of rated power per swept rotor area and hub height. The year 1990 was removed as an outlier for the left figure since there was only three installed WECs whereof two with unusually large rotor.

[11] Vattenfall recently made a survey with manufacturers on near-future trends for WECs. This information (shared with us in generalized form) was very useful.

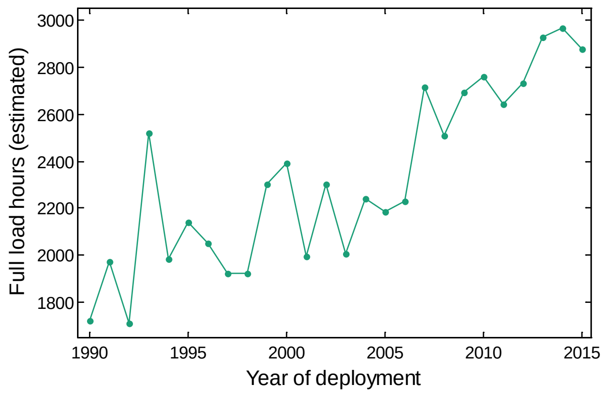


Figure 17. Historical, capacity weighted, estimated capacity factor (expressed in full load hours).

As a results of higher towers, larger rotors for a given rated capacity and, to a smaller extent, improvements in WEC technology, the capacity factors have increased substantially. Estimated capacity factors (full load hours) for different deployment years are shown in Figure 17. Since the capacity factors was systematically overestimated by turbine owners in the 90s, the improvement is in reality larger. For the aggregate fleet as of the end of 2012, the overestimation of production corresponds to around 7%[12].

The linear correlation between estimated capacity factor and mean wind speed according to the MIUU model is very low, only 0.07 for WECs built after year 2000 (also 0.07 for WECs built after 2005). Two factors that can explain the low correlation are:

- The MIUU model cannot resolve the small scale terrain, and therefore some projects in complex terrain will have better wind resource than anticipated by the MIUU model.
- In sites with lower mean wind speed, higher towers and larger rotors will be used, increasing the capacity factors.

Because of the weak observed correlation, the MIUU wind speed was given a relatively small impact on the farm probability and capacity factor in the scenarios (see Section 3.5).

Currently the installed capacity of offshore WECs is small in Sweden and none of the interviewees believed that offshore wind power will become significant in the near future. There are however some factors pointing towards increased levels of offshore wind in the future: mean wind speed is high (around 8-9 m/s at 100 metres), the Baltic Sea does not have as harsh climate as e.g. the Atlantic (extreme winds and waves are moderate and salinity is low) and the visual and acoustic impact is probably lower than for onshore wind. An additional support mechanism for offshore wind power is also currently under investigation [27], which may change the investment game plan significantly.

3.4 PARAMETER STUDIES

In this section results from some controlled experiments are presented. These were conducted in order to determine the impact on the resulting variability from different parameters. The results can be interesting in their own, but were also valuable for understanding which parameters to focus on and which were of less importance in the

[12] This fact was kept in mind when assigning capacity factors in the scenarios of future WECs; the full load hours given in the interviews were slightly adjusted downwards.

development of the scenarios. All results in this section are compared to a base case with the following parameter values:

- 20 TWh new onshore (no old WECs or offshore).
- In average 3000 full load hours. Standard deviation was 200 hours.
- Installed capacity per swept rotor area was 280 W/m^2.
- Other parameter values as described in Section 3.5.

In order to reduce computational time, ten years were considered and only the base model was used (i.e. not the improved model for hourly production or the model for intra-hourly fluctuations). It was however validated that the normalised results changed very little by using the improved model.

All variability metrics were normalised to the energy production and to the base case. This means that a simple upscaling of a production time series will give the same resulting variability. A discussion on whether the variability should best be normalised to the installed capacity (in MW) or the energy production (in TWh) can be found in Section 6.

A multitude of metrics are available for quantifying the variability of wind power, see e.g. Ref. [28] for a review. Most of these are highly correlated and for most of the parameter studies only the standard deviation of one hour step changes was considered[13]. For variability related to different capacity factors some other metrics are also shown. Note that the vertical axes in the figures in this section have different scales, this must be taken into account when studying the impact from the parameters on the variability.

We begin by evaluating the distribution of standard deviation in one hour step change for 100 realisations of the base case. Figure 18 show that there is a considerable spread among the different realisations. In order to reduce the random influence on the generation of scenarios, it was decided to produce 20 realisations for each scenario and to choose one with medium variability.

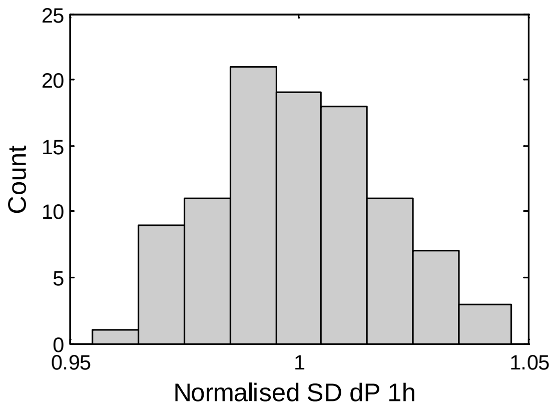


Figure 18. Histogram of standard deviation of one hour step change for 100 realisations of the base case.

Figure 19a shows how the variability reduces with more installed capacity. More than 30 TWh does not seem to give a further decrease. For comparison Figure 19b shows how variability can be reduced with an optimised spread of the farms[14]. Already with around five farms the variability is reduced to a level comparable with the base case. After adding more farms the variability is saturated at

[13] Although the metrics are highly correlated, the range differs. The metrics primarily used in this report, standard deviation in one hour step change, is one of the metrics with smallest range (see Figure 26).

[14] The optimisation was performed in a stepwise manner.

around 70-80% of that of the base case (depending on the grid points allowed for farm placement).

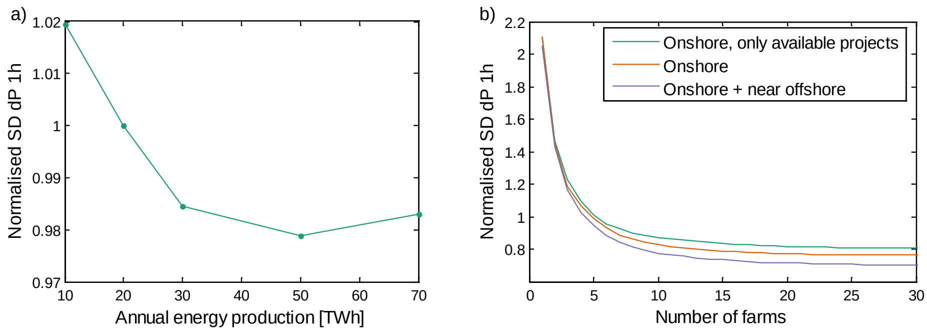


Figure 19. (a) Influence from annual energy production on the normalized standard deviation of one hour step change. (b) Variability for a fleet optimised to reduce variability. In order to isolate the effect of geographical spread, the same capacity factor was used for offshore and onshore farms. Note the different ranges of the vertical axes.

The capacity factor has a strong influence on the variability. With 2000 full load hours the maximum over mean value was 3.7. This ratio gradually decreases with higher capacity factor and is only 2.0 with 4000 flh. For comparison the ratio in actual production data was 3.4 year 2007 and 3.1 year 2012. Figure 20 shows how five other variability metrics change when the number of full load hours is varied between 2000 and 4000 hours. The results raise the question whether a higher capacity factor should be encouraged by some policy support instrument (see Section 6 for a further discussion).

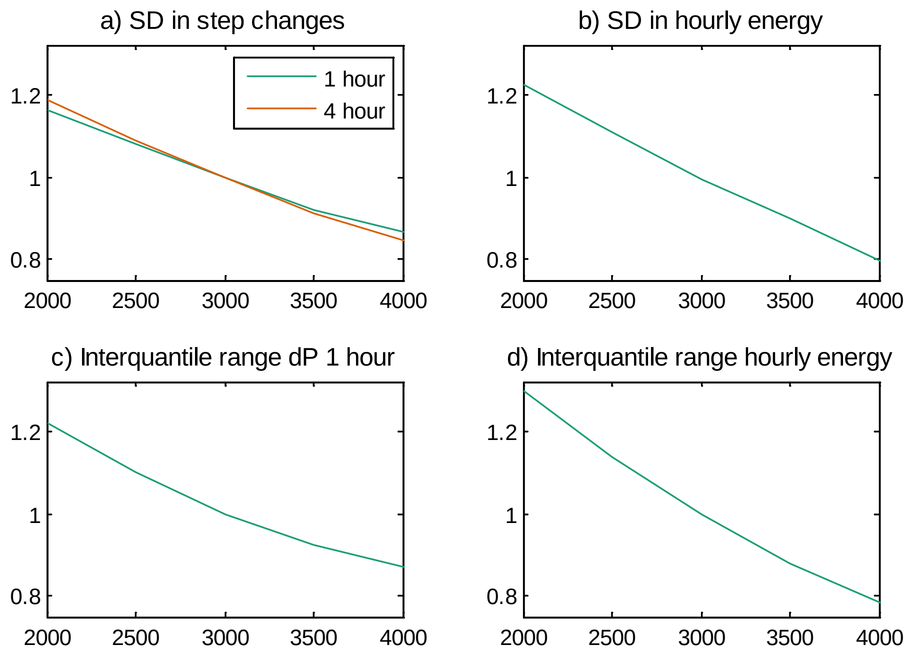


Figure 20. Influence from capacity factor on several variability metrics. Horizontal axes show number of full load hours and the vertical axes show normalised variability metrics. SD is short for standard deviation and dP for step change. The interquantile ranges are given between the 1st and 99th percentiles.

Figure 21a-d shows the influence from several parameters on the normalised standard deviation in one hour step change. In a) the standard deviation in number of full load hours is shown to have a relatively small impact. More variations around the mean value seem to be slightly positive in terms of reduced variability. Figure 21b proves that the rated power per swept rotor area has a negligible impact. Keep in mind however that this is for a fixed capacity factor (3000 full load hours); increasing the rotor area at fixed wind conditions and rated power would increase the capacity factor and hence reduce variability as shown in Figure 20. In c) the influence from the share of capacity in northern Sweden (SE1 and SE2) is shown. Not surprisingly variability is lowest with an approximate equal distribution between north and south[15]. Figure 21d finally shows that a larger offshore share would reduce the aggregated variability (although the variability in the regional grid could probably be an issue). This is primarily due to a higher assumed capacity factor for offshore WECS (4000 full load hours), but could also be attributed to a slightly larger total geographical area used. When the share of offshore increases above around 50%, the variability increases again due to the massive concentration to a few large projects.

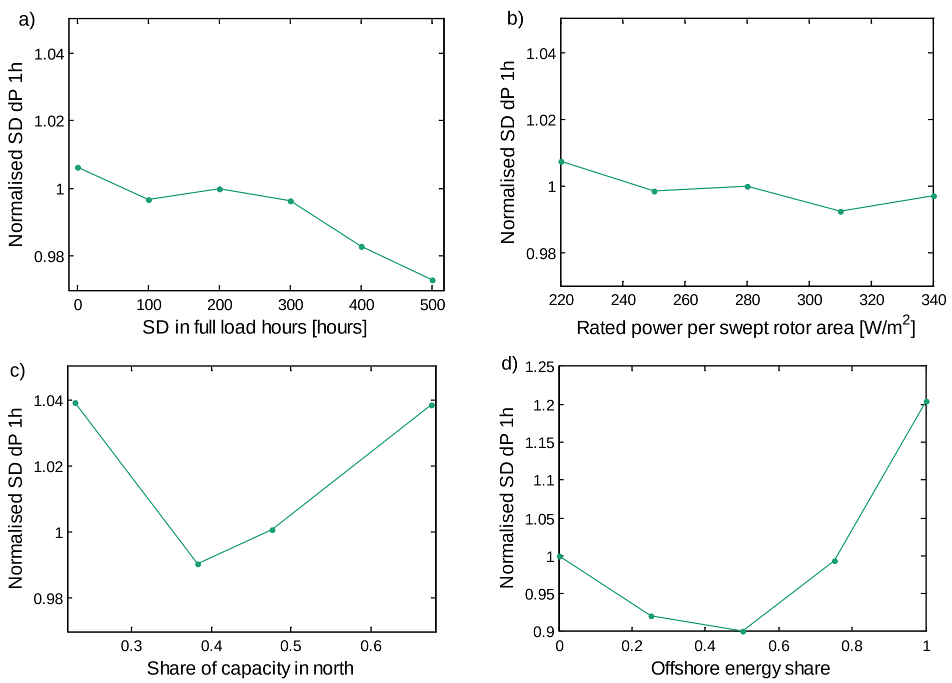


Figure 21. Influence from several parameters on the normalised standard deviation in one hour step change. Note the different ranges of the vertical axes.

3.5 RESULTING SCENARIOS

In this section the chosen parameter values are given for all scenarios and some examples of the resulting time series are shown. The scenarios are divided into four groups, A-D, based on their annual energy production (20, 30, 50 and 70 TWh respectively). A1, B1 and C1 are the base scenarios, based on the most likely parameter sets. Separate scenarios are available with different assumptions on capacity factor (±500 full load hours as

[15] We did not looked deeper into the influence from distribution over all four price areas. Results from the 100 realisations of the base case (Figure 18) however suggests that it would be beneficial with more capacity in SE1 than we have today.

compared to the base scenarios), probability weight for farms in north and offshore share. Scenario A8 consist of the current fleet (built and under construction) while scenario D1 has a very high wind penetration.

From the parameter studies in the previous section, it is clear that the most important parameter for the variability is the capacity factor. In Table 5 parameter values common for all scenarios are given. Table 6 shows the most important parameters differing between the scenarios. Apart from the parameter values given in Table 5 and 6, the following values were specified:

- *Keep old* was set to year 2000 for scenario A1-A6, 1900 for A8, 2010 for the B scenarios and 2020 for the remaining scenarios.
- $P/A_{onshore}$ was set to 300 W/m^2 for the A scenarios and 280, 270 and 260 W/m^2 for B, C and D scenarios respectively.
- $P/A_{offshore}$ was set to 350, 330, 310 and 290 W/m^2 for the A, B, C and D scenarios respectively.
- *Weight north* was normally set to 0.8. For "more in north" and "less in north" scenarios values of 1.8 and 0.4 were used.

Table 5. Parameter values (common for all scenarios). See Section 3.2 for explanations.

Parameter	Value	Parameter	Value
SD flh	300	*Park share 2*	0.7
Wind impact 1	500	*Weight process*	0.4
Wind impact 2	1	*RI wind 1*	1.3
Separation	4	*RI wind 2*	1.5
Max P	3500	*Repower 1*	0.8
$P_{unknown}$	3000	*Repower 2*	1.3
Park share 1	0.8	*Aggregate*	yes

Table 6. The most important parameters and features of the different scenarios. The capacity factors (CF) are expressed in full load hours. See Section 3.2 for explanations.

	Annual energy [TWh]	Whereof new offshore [TWh]	CF new onshore [hours]	CF new offshore [hours]	Comments
A1	20	0	3000	-	Base scenario
A2	20	0	3000	-	More WECs in north
A3	20	0	3000	-	Fewer WECs in north
A4	20	5	3000	4000	More offshore
A5	20	0	3500	-	Higher capacity factor
A6	20	0	2500	-	Lower capacity factor
A7	20	0	3000	-	Only new WECs
A8	14.3	0	-	-	Only WECs already built or under construction
B1	30	5	3200	4100	Base scenario
B2	30	5	3200	4100	More WECs in north
B3	30	5	3200	4100	Fewer WECs in north
B4	30	12	3200	4100	More offshore
B5	30	0	3200	-	No offshore
B6	30	5	3700	4600	Higher capacity factor
B7	30	5	2700	3600	Lower capacity factor
C1	50	13	3300	4300	Base scenario
C2	50	13	3300	4300	More WECs in north
C3	50	13	3300	4300	Fewer WECs in north
C4	50	25	3300	4300	More offshore
C5	50	5	3300	4300	Less offshore
C6	50	13	3800	4800	Higher capacity factor
C7	50	13	2800	3800	Lower capacity factor
D1	70	20	3400	4400	High penetration scenario

An example of a generated scenario (B1) is shown in Figure 22. Similar maps for all scenarios can be found in Appendix A together with bar plots of installed capacity per area, duration curves of production and some key variability metrics. Common for all scenarios is that SE1 has the smallest installed capacity. For scenarios with a large offshore fleet (e.g. all C scenarios except C5), SE4 is generally the area with most installed capacity. Scenarios with less offshore capacity often have a relatively equal distribution between the areas SE2, SE3 and SE4.

Information on the WECs from the scenarios were subsequently fed to the model described in Section 2, and time series spanning 1979-2014 were generated. An example of a small part of the time series for scenario B1 is shown in Figure 23. Duration curves of the production (for all data as well as for the winter months December - February) is shown in Figure 24. The minimum production is zero or very close to zero for all scenarios. The first percentiles of the production is however significantly higher for the later scenarios as compared to the 20 TWh scenarios or, even more, when comparing to historical data. Expressed in p.u. (i.e. normalised to installed capacity) the 1st percentile is 2.3% for the measured production between 2007 and 2012 and 3.0%, 3.9% and 5.0% for scenario A1, B1 and C1 respectively. The 1st, 5th and 10th percentiles for all scenarios and for historical production are given in Figure 25. In this figure the values are normalised to the mean production, so the differences between the scenarios are smaller.

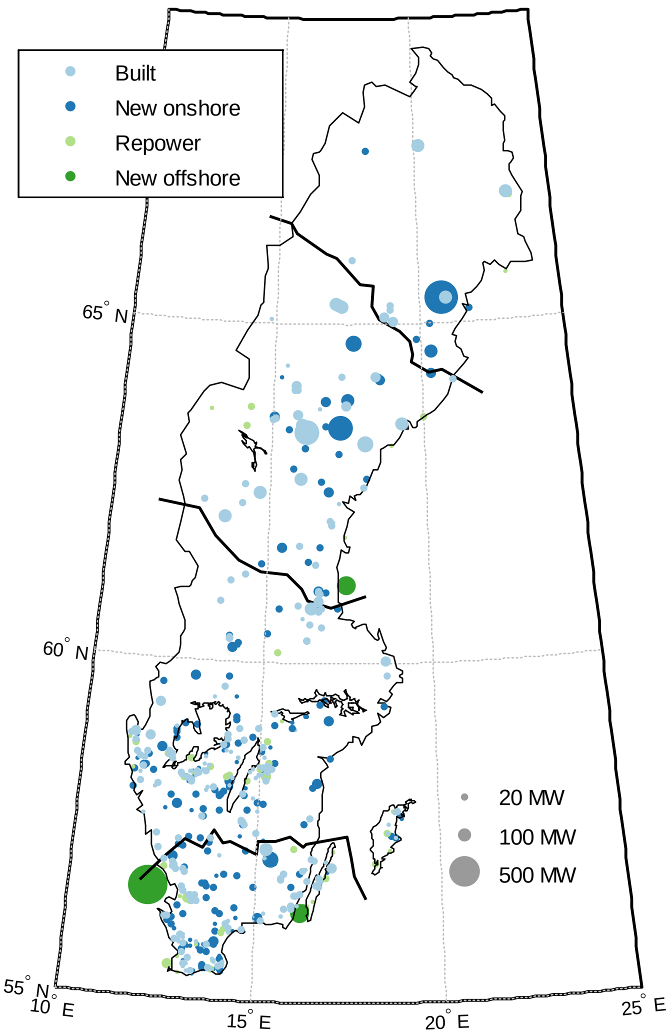


Figure 22. An example of a scenario (B1). The circles areas are proportional to farm capacity.

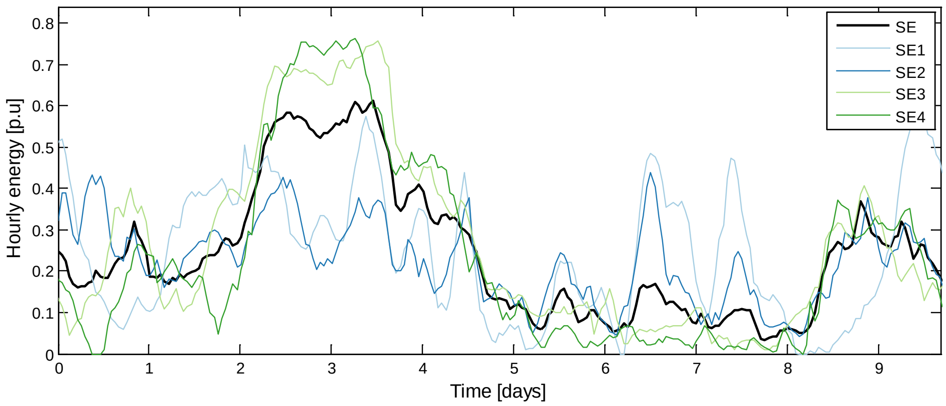


Figure 23. Example of time series of hourly energy production (scenario B1). Note that the aggregated output for the whole of Sweden (SE) is smoother than the output for the individual price areas; the range is smaller and the ramping rate is slower.

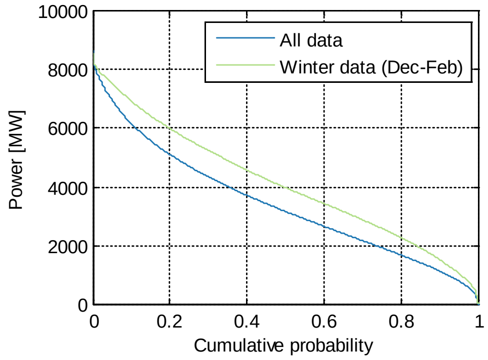


Figure 24. Duration curves for hourly energy production (scenario B1) for the whole of Sweden. Similar figures for all scenarios are found in appendix A.

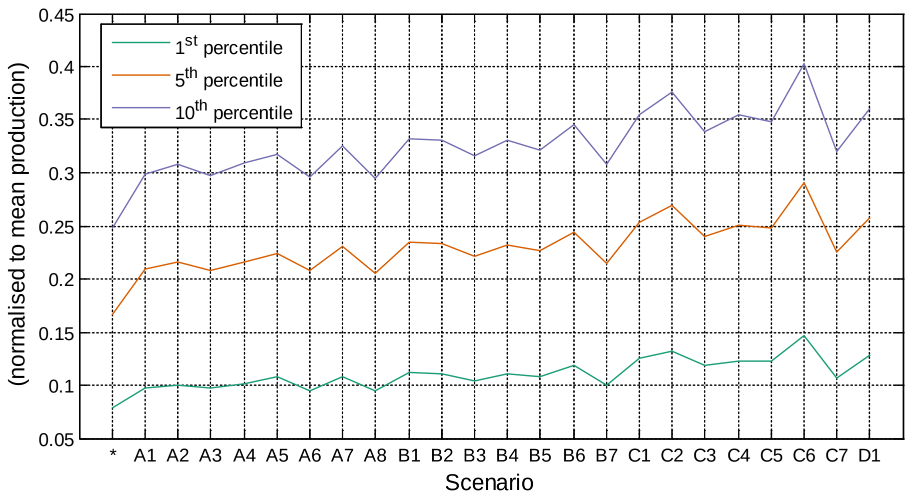


Figure 25. The 1st, 5th and 10th percentiles of hourly production for historical production 2007 through 2012 (marked with an asterisk) and the different scenarios. The percentiles are normalised to the mean production.

An example of the correlation between the production in the different areas can be found in Table 7. As a result or the relatively low correlation between production in SE1 and the other areas, variability is generally lower in scenarios with more capacity in this area. Note that the correlation in production averaged over longer time horizons[16] is significantly higher than the correlation in hourly production.

Table 7. Correlation between the hourly production in the different areas (scenario B1).

	SE	SE1	SE2	SE3	SE4
SE	-	0.57	0.78	0.87	0.79
SE1	0.57	-	0.72	0.25	0.13
SE2	0.78	0.72	-	0.50	0.28
SE3	0.87	0.25	0.50	-	0.75
SE4	0.79	0.13	0.28	0.75	-

[16] For instance, the correlation in 2-5 days averaged production might be of interest since this is the time scale of many large scale weather patterns.

Some key variability metrics for all scenarios can be found in Figure 26. All values are normalised to the annual energy production and then to the value for scenario A8 (the WEC fleet as of 2015). For some of the metrics expressed in MW, see Appendix A. Note that although the range of the normalised metrics vary, the correlation between them is above 0.97 for all pairs. A variability index, calculated as the average value of all six metrics, is 0.98 for scenario A1, 0.88 for B1 and 0.80 for C1. For actual production year 2007-2012, the same index was 1.15.

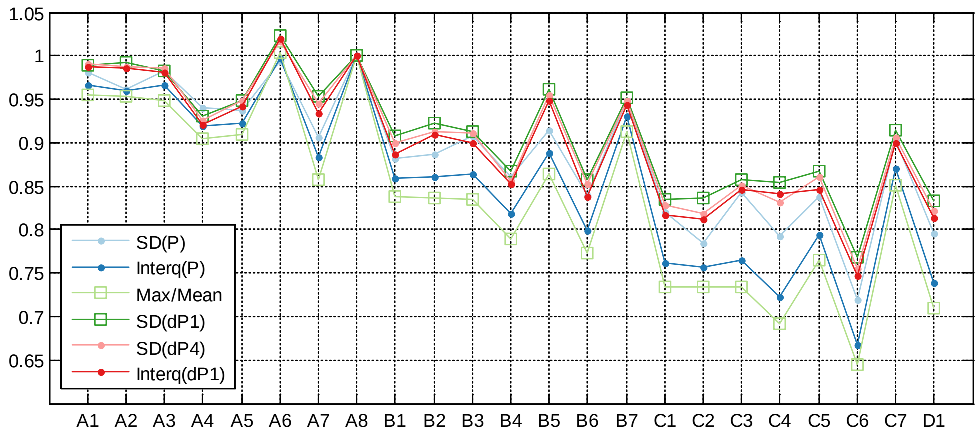


Figure 26. Normalised variability metrics for the different scenarios (normalised to energy production and scenario A8). SD is short for standard deviation, P for hourly energy production, dP1 for one hour step change in hourly energy production and interq for range between the 1st and the 99th percentile.

We conclude this section by showing how the wind power production for scenario B1 correlates with the hourly Swedish load (which had an average value of 16.0 GW during the studied period 2001 to 2013). Wind power production is generally higher during wintertime when the load is higher. This leads to a weakly positive correlation between wind production and load (R=0.20). Judging from Figure 27 it however seems that the average wind power have a maximum for loads around 19-20 GW, but declines for higher loads. A load below 9.5 GW and above 24 GW are relatively rare (below 1 % of the time respectively) so the values corresponding to these levels must be interpreted with caution.

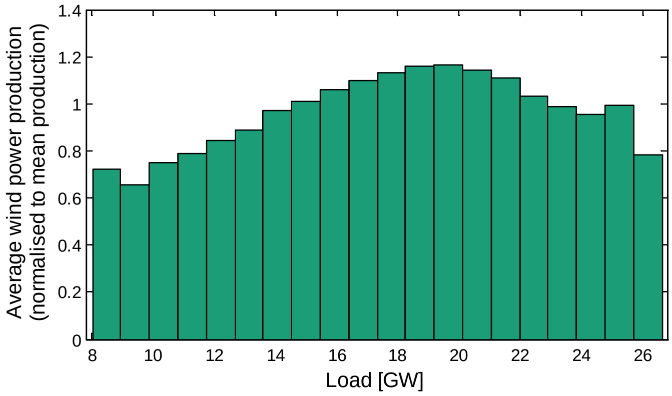


Figure 27. Average wind power production for different levels of electric load in Sweden (year 2001 to 2013). Equally spaced load bins with a width of around 0.93 GW was used.

4 Intra-hourly fluctuations

For some power system studies it is relevant with wind power time series with higher resolution than one hour. There is however no publically available intra-hourly measurements on aggregated wind power production in Sweden, neither is this possible to produce from the meteorological dataset used in earlier sections (the MERRA dataset). In order to develop and optimise a statistical model of the intra-hourly fluctuations, aggregated production data from Denmark, Germany, Ireland, Canada, USA and smaller parts of Sweden was studied. The data was retrieved from the system operators webpages (Refs. [29]–[34]) or by personal communication.

Key metadata for the studied areas is shown in Table 8. Dimensions of the areas are approximate and wind farms are not always distributed over the entire areas (in particular AESO have most farms concentrated in a smaller area).

Table 8. Metadata for wind power in the studied power systems. The last two rows show data for Sweden (measured and simulated with the methods described in Section 2).

System operator	Time period	Resolution [minutes]	Capacity factor	Installed capacity [GW]	Dimensions [km]
BPA (US)	2008-2013	5	0.28	1.4 – 4.5	1100×800
ERCOT (US)	2012-2014	5	0.25	9.4 – 12	1000×700
Energinet (DK)	2010-2013	5	0.27	3.5 – 4.8	350×300
TenneT (DE)	2009-2013	15	0.19	10 – 12	800×400
50 Hertz (DE)	2009-2013	15	0.17	9.6 – 14	500×350
Eirgrid (IE)	2008-2013	15	0.28	0.81 – 1.8	450×300
AESO (CA)	2008-2012	10	0.32	0.50 – 1.1	1200×600
Gotland (SE)	2011	5	0.25	0.11	120×50
Vattenfall (SE)	2011-2014	10	0.39	0.24	
Sweden (measured)	2007-2012	60	0.24	0.61-3.5	1600×400
Sweden (simulated)	1979-2014	60	differs	differs	1600×400

4.1 DATA

The time series received from the system operators were expressed in MW. For easier comparison the data were transformed to per units (p.u.), i.e. the MW values were normalised to the installed capacity. The quality of the time series of installed capacity differed between the areas; for some the installed capacity for each day were available, but in some cases interpolation had to be used since only monthly or yearly values were available. For the current application the error related to the interpolation is of virtually no importance. As an example, overestimating the installed capacity will result in an underestimation of the p.u. power. The effect is a relatively uniform reduction of the power spectral density (PSD). Since a constant was used to compensate for the difference in PSDs up to 0.5 hours^{-1}, the simulated high-frequency fluctuations will be more or less unaffected by the error. The time series were reviewed manually and suspicious data were removed or altered.

4.2 POWER SPECTRAL DENSITY

Figure 28 shows power spectral density estimates for the various power systems. There is a relatively large difference in PSDs between the areas, especially for higher frequencies. In comparison, the variations in PSDs between different years are small, even though the capacity for some areas increased with more than 100% during the studied period. In general, there is more volatility in production in USA/Canada than in Europe for similar installed capacity and area dimensions.

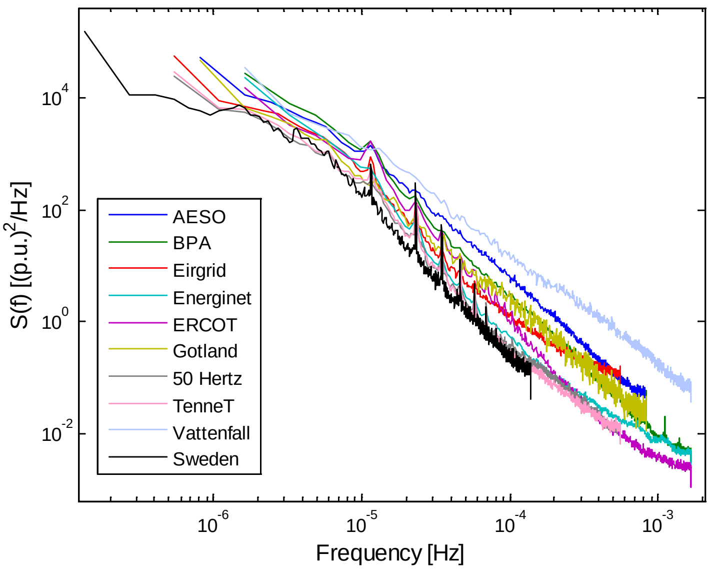


Figure 28. Power spectral densities for all areas (Welch estimates with segments of 2048 samples).

For the continued modelling, the most similar areas to Sweden with respect to PSDs were chosen (TenneT and Energinet). TenneT, with 15 minutes temporal resolution, is the most similar area at higher frequencies. Energinet has 5 minutes temporal resolution and a relatively constant PSD ratio to Sweden for all frequencies. PSDs for the chosen areas are shown in Figure 29.

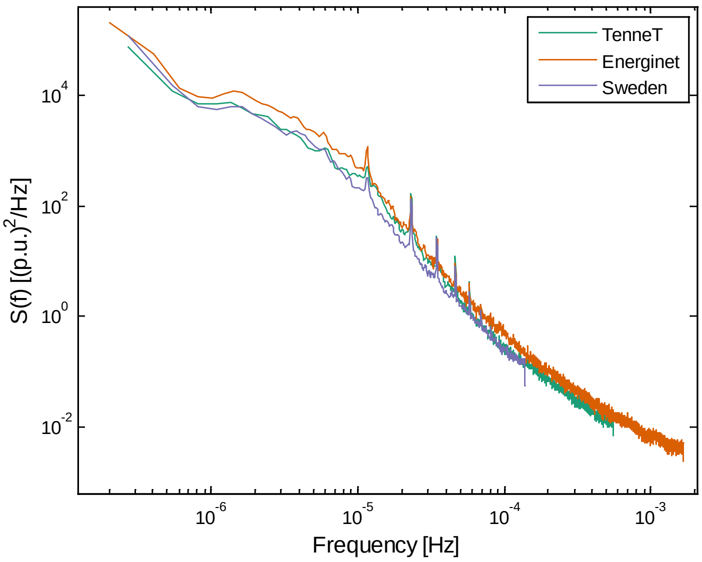


Figure 29. Power spectral densities for selected areas (Welch estimates with one month long segments).

4.3 METHODOLOGY

The methodology used for training a model and simulate high-frequency variations has a lot in common with that presented in Section 2.2. The general idea was to isolate the higher frequency variation for Energinet and TenneT, transform these time series to approximately stationarity and find the corresponding Fourier series. Subsequently similar noise was generated, back-transformed and added to the smoothened curve from the hourly model for Sweden. The following steps were performed:

1. Fit cubic splines to hourly data from Energinet, TenneT and the hourly model for Sweden described in Section 2.
2. Isolate the higher frequency variations: $x(t) = P_{5/15\ min} - P_{cubic\ splines}$.
3. Transform $x(t)$ to stationarity: $x(t) \rightarrow y(t)$.
4. Find the magnitudes of the frequency domain representation of $y(t)$ using FFT.
5. Generate $y_{sim}(t)$ using inverse FFT with interpolated magnitudes from 4) and random phases.
6. Transform $y_{sim}(t)$ to $x_{sim}(t)$ using the inverse of the transformation used in 3).
7. The time series for the intra-hourly model was achieved as $z(t) + c{\cdot}x_{sim}(t)$, where z(t) is the cubic splines fit for the objective system and c is a scaling constant.

The reason for modelling $x(t) = P_{5/15\ min} - P_{cubic\ splines}$ instead of $x(t) = P_{5/15\ min} - P_{hourly\ mean}$ was that the latter would give larger fluctuations and a strange autocorrelation pattern. This can easily be understood by studying Figure 30. Filtered hourly data (like in Ref. [35]) was also evaluated, but this gave slightly poorer performance than using cubic splines[17].

[17] The cubic splines were calculated based on one sample per hour, while the filter was applied on hourly mean data at sampling frequency (i.e. 5 or 15 minutes).

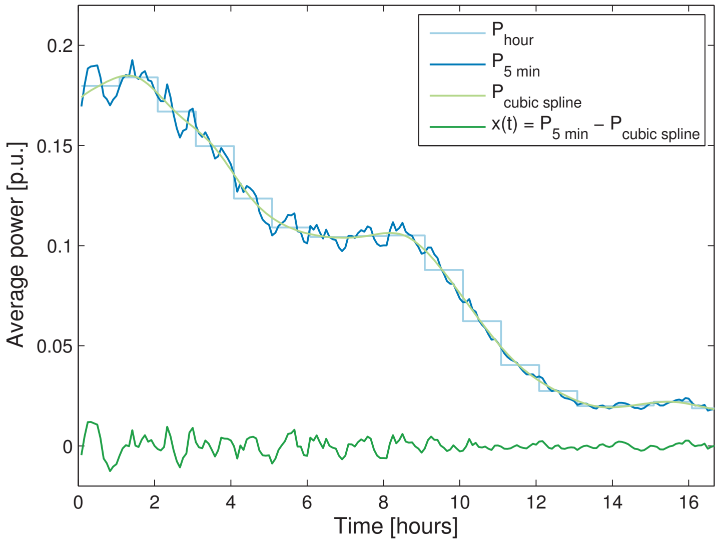


Figure 30. An example from the Energinet time series showing average hourly and 5 minute power, cubic spline fit for hourly power and difference between 5 minute power and cubic spline fit. The last variable, *x(t)*, is the one to be modelled. Note that *x(t)* is non-stationary and thus first needs to be transformed.

The transformation to approximately stationarity was performed exactly like in Section 2.2.1. For each bin in the power – power derivative space, a t location-scale distribution was fitted to the fluctuation magnitudes *x(t)* and the data was transformed to a normal distribution with zero mean and standard deviation one. The fluctuations dependence on power and its derivative is shown in Figure 31. Subsequently the seasonal and diurnal bias was removed and an approximately stationary time series *y(t)* resulted.

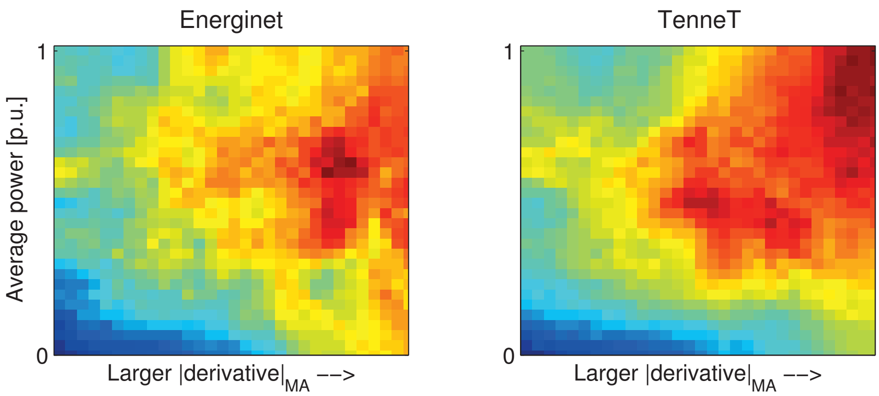


Figure 31. Sigma parameter in the t location-scale distribution for *x(t)* as a function of power and its derivative. The fluctuations are approximately an order of magnitude smaller at low power and low derivative than at medium power and high derivative.

Noise with the same frequency domain magnitudes and random phases was finally generated and back-transformed using the same seasonal/diurnal bias and t location-scale distributions. Since the power spectral densities differs between Energinet, TenneT and the different scenarios for Sweden, a scaling constant was applied to the simulated noise. The constant was chosen so that the average ratio between the PSDs above 0.5 hours^{-1} was the same as the average ratio for PSDs of hourly production for frequencies between 1/12 and 0.5 hours^{-1}.

Example: Energinet data was used to train a model for intra-hourly fluctuations for scenario B3. When comparing the PSDs of hourly power, the average ratio for frequencies between 1/12 to 0.5 hours^{-1} was 0.92, i.e. there was slightly less variability in the B3 time series than in that for Energinet. A scaling constant was hence chosen so that the simulated time series for B3 (temporal resolution 5 minutes) has a PSD with, in average, 92% of that of Energinet for frequencies above 0.5 hours^{-1}. This approach guarantees that there is no abrupt change in the PSD of the simulated time series (see Figure 36).

A few samples in the simulated high frequency fluctuations time series sometimes got unrealistically high magnitude. Therefore limits were set based on the observed ratio of maximum over 99.9th percentile and minimum over 0.1th percentile. It was taken into account that for a longer simulated time series (e.g. 36 years for the scenarios), the ratios should be increased, see Figure 32 for extrapolation for Energinet. A limit was also set for the resulting power time series; zero downwards and an upward limit based on the observed ratio of maximum 5/15 minutes value over maximum hourly value.

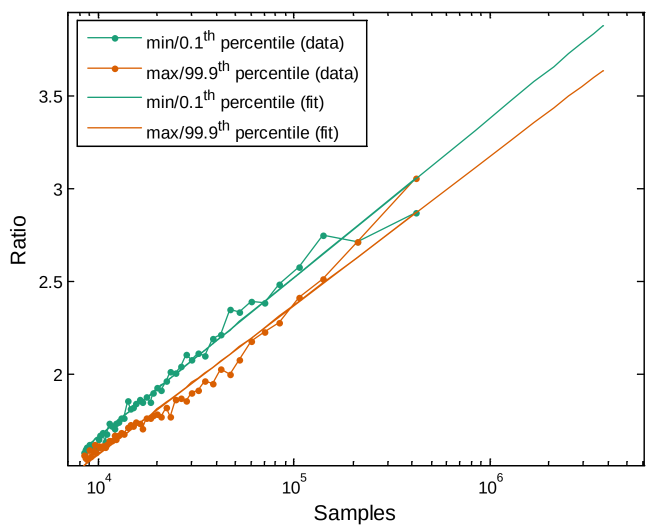


Figure 32. Ratios as a function of number of samples (Energinet). The linear fits were extrapolated to the number of samples for the simulated time series (36 years of data).

4.4 RESULTS

In this section evaluation results are given for Energinet and TenneT. In both cases, the first half of the time period was used for training the model and the second half was used for evaluation. For Sweden there is no validation data with five or 15 minutes temporal resolution. Some exemplifying results are however given in Section 4.5.

Figure 33 show distributions of measured and simulated high frequency fluctuations, i.e. $x(t)$ and $x_{sim}(t)$. As can be seen from the figure the match is good also for more extreme events. Figure 34 shows an example of the timing of the magnitude of the fluctuations. When comparing to the improved model for hourly fluctuations in Section 2.2, the intra-hourly model is better at capturing the non-stationarity. Correlations in one day moving averages of the absolute values of the fluctuation magnitudes were 0.76 for Energinet and 0.79 for TenneT. Figure 35 shows the power spectral densities of the measured and simulated power time series. The distribution of step changes in 5/15 minute average power was also very well reproduced, the standard deviation was for example 0.54%

and 0.55% for measurements and simulation for Energinet and 0.82% and 0.80% for TenneT.

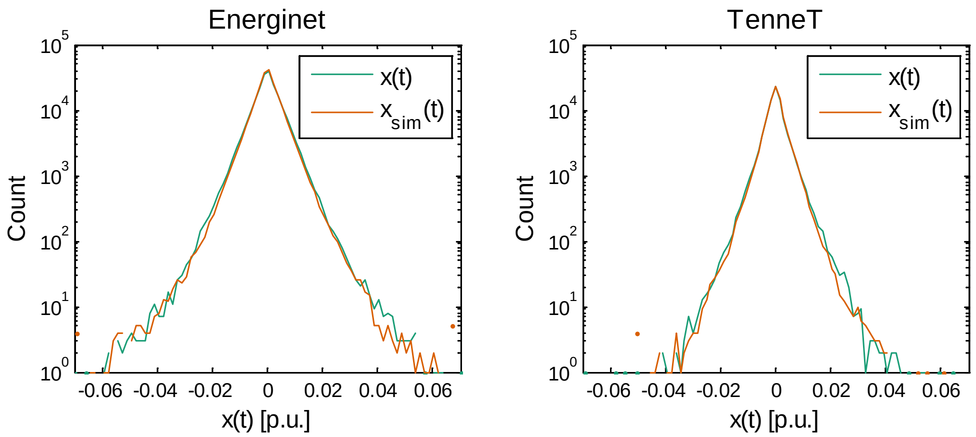


Figure 33. Distribution of *x(t)* and *$x_{sim}(t)$*, i.e. measured and simulated difference between power (5 and 15 min average) and cubic spline fit for hourly power.

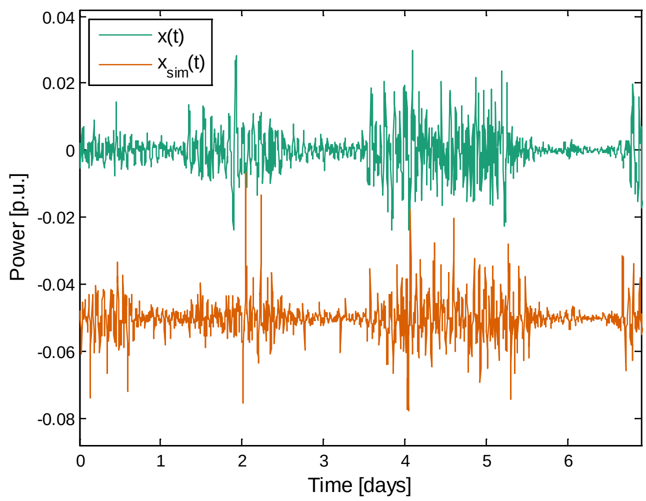


Figure 34. Example time series of *x(t)* and *$x_{sim}(t)$* for Energinet. *x_{sim}* is displaced -0.05 p.u for better readability. As can be seen, the non-stationarity of the fluctuation magnitude is relatively well reproduced.

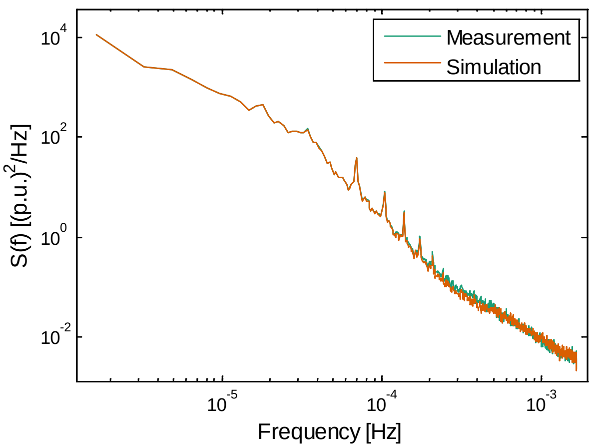


Figure 35. PSDs for the evaluation period (Energinet).

4.5 APPLICATION ON SCENARIOS

The models of intra-hourly fluctuations applied on the hourly time series for the different scenarios were trained on the whole datasets from Energinet and TenneT. In order to assure that the combined production for the separate price areas matches the production for the whole of Sweden, 5 and 15 minute time series were first produced for the separate areas and subsequently summed to give the production for Sweden as a whole. Figure 36 and Figure 37 show examples of resulting power spectral densities and time series of 5 minute aggregated production.

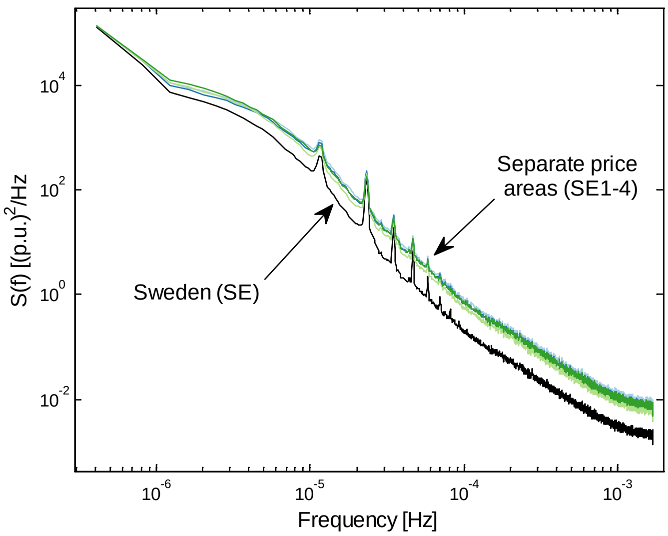


Figure 36. Power spectral densities for the 5 minute time series for scenario B1.

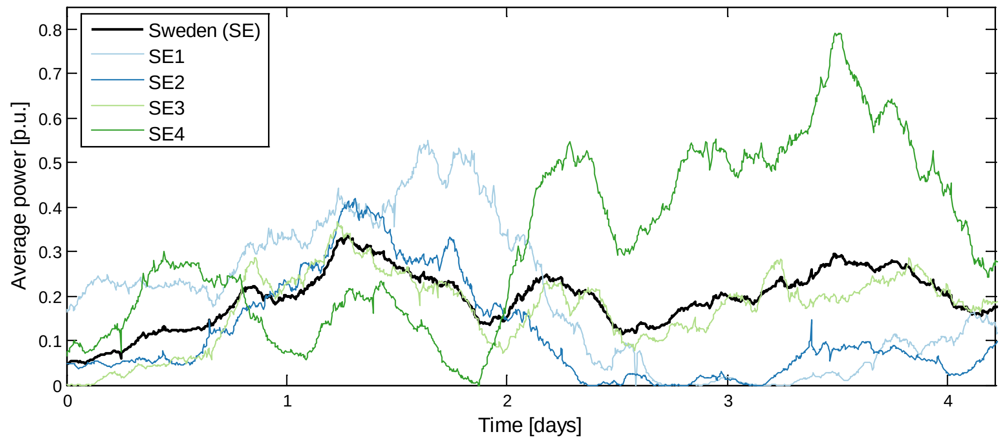


Figure 37. A few days example of the resulting 5 minute time series for scenario B1.

5 Conclusions

This report contain three main sections: one on the modelling of hourly wind power production, one on the development on scenarios for wind power in Sweden and one on the statistical modelling of intra-hourly fluctuations. The model of hourly aggregated production is based on the MERRA reanalysis dataset and shows good agreement with actual measurements from the Swedish TSO (mean absolute error was below 3% of installed capacity). The original model, first presented in Ref. [16], however underestimates the energy in fluctuations with periods shorter than 10 hours, which leads to an underestimation of the one hour step changes. Therefore a statistical correction was developed, compensating for the missing high-frequency fluctuations. With the improved model, the simulated time series had a very similar power spectral density as measurements over all frequencies and the duration curve for one hour step changes was in good agreement with measurements.

Three main scenarios were developed, with annual energy production of 20, 30 and 50 TWh respectively (A1, B1 and C1). Sub-scenarios with i) more/less deployed in the northern part ii) higher/lower capacity factor and iii) more/less offshore compared to the base scenarios were also produced. One scenario of very high penetration (70 TWh) as well as a reference scenario with the current fleet were finally generated. The scenarios were based on a database of planned projects (over 16,000 WECs) and rules governing the probabilities for the different farms to be built. As an example, farms with all permits ready are thought to be more likely to be built and so are farms in designated areas for wind power. 36 years long time series with one hour temporal resolution were produced for all 23 scenarios. These are available online, see Appendix C.

Before the final parameter values were set for the scenarios, studies of historical trends and impact from the different parameters were conducted. It was shown that the two most important parameters with respect to the resulting variability were capacity factor and share of offshore capacity. By optimizing the location of the farms, the variability could be reduced substantially. However, just installing more capacity in a more random manner gave a very small decrease in variability compared to the situation today. The average rated power per swept rotor area and share of production in northern Sweden also have a relatively small impact on the variability.

A variability index, calculated as the average value of six different normalised variability metrics, was 0.98 for scenario A1, 0.88 for B1 and 0.80 for C1 (compared to 1 for modelled production for the WEC fleet as of year 2015 and 1.15 for measured production year 2007 through 2012). Thus the relative variability is expected to be reduced in the future, mostly as a consequence of higher capacity factors. The minimum production will still be zero or very near zero for all scenarios. The 1st, 5th and 10th percentiles, however, are considerably higher than for historical data. The maximum over mean production is gradually reduced when comparing the historical production 2007-2012 and modelled production for scenario A8 (WEC fleet year 2015), A1, B1 and C1. An example highlighting the importance of using realistic scenarios instead of simple upscaling of historical production is that the maximum production for scenario C1 (50 TWh/year) is 12.6 GW while the same figure for upscaled production from year 2007 to 2012 (to 50 TWh/year) is 19.4 GW. It is thus likely that transmission costs as well as problems related to high wind / low load will be severely overestimated with the upscale approach.

Since Sweden does not have measurements of aggregated wind power production with a sub-hourly resolution, statistical models of the intra-hourly fluctuations were trained on data from Germany (TenneT) and Denmark. These systems were chosen since their power spectral densities of aggregated wind power production are similar to that for Sweden. When evaluating the models of 5 and 15 minute production, they gave similar characteristics in terms of deviation from the hourly mean and 5/15 minute step changes. Not only was the magnitude of the shorter fluctuations reproduced, but they were also well captured in time (some periods show very small fluctuations while others are more volatile). Time series with 5 and 15 minute resolution were subsequently generated for the three base scenarios.

6 Discussion

In this report, variability is always normalised to the annual energy production. Most readers would probably find some kind of normalisation justified, but some might argue that it would be better to use the installed capacity instead of the produced energy. This would change the results and conclusions so that the normalised variability would be higher rather than lower for a fleet with higher capacity factor. We however believe that these results would be misleading; the benefit that wind power brings to the power system is not the installed capacity but the amount of produced energy. Most political targets we are aware of are set as goals in e.g. TWh or share of annual electricity consumption. Consequently the (negative) impact from wind power on the grid in terms of added variability should be compared to the benefits in form of energy production.

An interesting question not addressed in this work is how much variability costs and what it would be worth to reduce it. It would likely only imply a small increase in cost per produced kWh for the developer to construct wind farms with significantly higher capacity factors. Maybe this extra cost could be more than compensated by reduced investments in grid infrastructure and balancing costs? Offshore wind power is currently not competitive in Sweden but a special support system is currently under investigation. Based on the results from this study, more offshore would lead to a reduction of variability, which should be taken into account. Preferably, studies should first be made on the impact on the power system, not only looking at the isolated variability for different wind power scenarios.

The three factors contributing most to a lower variability were higher average capacity factor, more offshore wind and an optimised location of future farms. In order to get an idea of the benefits of combining these factors we compared two cases:

- 20 TWh new onshore wind power with 3000 full load hours, distributed in a quasi-random manner as described in Section 3.2
- 10 TWh new onshore (distributed in a random manner) plus ten 1 TWh farms optimised for variability reduction. The optimised farms were drawn from available onshore and offshore projects. Furthermore, the capacity factors were increased with 20% to 3600 full load hours for onshore farms and 4800 for offshore.

The variability index (see Section 3.5) was reduced with 29% for the latter case. The optimised farms were located relatively uniformly around the perimeter of Sweden, primarily at offshore sites.

An important issue when dealing with scenarios of future WECs is whether the model of aggregated production is still valid. This is of course impossible to answer today. We however believe that the approach with a validated physically based model (rather than using upscaling of historical output or purely statistical methods) is the best method for dealing with the production from future wind power fleets. It should also be noted that the training and evaluation of the model were performed for a period when the installed capacity increased almost with a factor of six. It is theoretically possible that the model will overestimate the variability of future fleets since the same power curve smoothing parameter (see Ref. [16] for details) was used as for the training period. However, when optimizing the model on different years (from 2007 to 2012), no trend towards increased smoothing parameter could be found; a standard deviation of around 1.1 m/s was optimal for all individual years.

As noted by members of the reference group, the most important factor not (directly) taken into account for the generation of the scenarios was the cost for grid connection. This cost could differ between projects and is important for the overall cost of the projects and consequently the probability that the farm will actually be constructed. Unfortunately there is no simple way to estimate the cost; the distances to the closest point of grid connection is not readily available, and even if they were, a site-specific analysis would have to be made in order to determine whether the farm could be easily connected or if additional grid reinforcements would be required. One factor reducing the impact of not taking into account grid connection costs is that a substantial amount of money have been invested in projects with all permits. Most of these projects, which are the most likely to be added to the scenarios, therefore probably have acceptable grid connection costs. The average grid connection cost is likely higher in the northern part of Sweden. The cost was therefore to some extent implicitly accounted for by the probability weight for northern projects.

A final comment is on curtailment of wind power. So far, this option has not been used in Sweden, but it might be more relevant with a higher wind power penetration. For scenario B1, as an example, reducing the maximum output for the two northern areas with 10% would cost 0.3% of the produced energy. Another example is that it could be necessary to limit the maximum instantaneous ratio of wind power production over load. According to load data from 2001 to 2013 and wind power production according to scenario B1, the highest ratio would be 72% for the whole of Sweden. Setting a limit at 50% would lead to a spill of only 0.3% of the wind energy production. This is of course simplified analyses, but the produced time series provide good opportunities for more in-depth studies of curtailment and other strategies to reduce negative effects from wind power production.

7 Future work

Some possible extensions of the current work include:

- Expand the wind power models and scenarios to the Nordic synchronous area (Norway, Sweden, Finland and eastern Denmark). This area is tightly interconnected and power system studies should preferably not look at Sweden as an isolated system. Therefore realistic time series of future wind power production in the whole area would be an important contribution.
- Develop an accompanying model for forecast errors. These errors are conditioned on improvements in forecasting methods as well as the varying meteorological conditions (i.e. also these time series are non-stationary). The forecast errors are important for power system operation, e.g. for reserves.
- Improved intra-hourly model, e.g. by using ARMAX (autoregressive moving average models with exogenous inputs) or artificial neural networks.
- Combined study of solar, wind, wave and tidal power for the Nordic synchronous area. How do these sources correlate? Are there benefits from combining them in some particular way?
- More in-depth study of the (net) variability as a function of WEC characteristics and geographical spread. What is reduced variability from wind power worth? Could it be justified to steer the development towards e.g. higher capacity factors and more offshore? Should we procure wind farms at proper locations in order to reduce the variability?

Acknowledgements

This work was conducted within the StandUP for Energy strategic research framework. The study was funded by Vindforsk through contract 38515-1. The authors would like to express their gratitude to the interviewees and the reference group (see Table 4), to Anders Björck, Åsa Elmqvist and Erik Thunberg at Energiforsk/Vindforsk, to Magnus Hedlund, Morgan Rossander and the “variability group” (solar, wind, wave and tidal power) at Uppsala University, to Hans Ohlsson (wpd), Martin Johansson (Energimyndigheten), Math Bollen and Rujiroj Leelaruji (STRI), to Peiyuan Chen (Chalmers), to Helena Nielsen (Vattenfall) and to Martin Johansson, (Energimyndigheten).

References

[1] J. Olsson, L. Skoglund, F. Carlsson, and L. Bertling, 'Future wind power production variations in the Swedish power system', in *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, 2010, pp. 1–7.

[2] R. Billinton, H. Chen, and R. Ghajar, 'Time-series models for reliability evaluation of power systems including wind energy', *Microelectronics Reliability*, vol. 36, no. 9, pp. 1253–1261, Sep. 1996.

[3] G. Papaefthymiou and B. Klöckl, 'MCMC for Wind Power Simulation', *IEEE Transactions on Energy Conversion*, vol. 23, no. 1, pp. 234–240, 2008.

[4] F. Carlsson, 'Wind power forecast errors - Future volumes and costs', Elforsk rapport 11:01, 2011.

[5] M. Magnusson, R. Krieg, M. Nord, and H. Bergström, 'Effektvariationer av vindkraft, En studie av vindenergiproduktionens tidsvariation', Elforsk rapport 04:34, 2004.

[6] L. Söder, 'På väg mot en elförsörjning baserad på enbart förnybar el i Sverige, En studie om behovet av reglerkraft, Version 2.0', KTH, 2013.

[7] J. Widen, 'Correlations Between Large-Scale Solar and Wind Power in a Future Scenario for Sweden', *IEEE Transactions on Sustainable Energy*, vol. 2, no. 2, pp. 177–184, 2011.

[8] L. Göransson and F. Johnsson, 'Cost-optimized allocation of wind power investments: a Nordic–German perspective', *Wind Energy*, vol. 16, no. 4, pp. 587–604, 2013.

[9] A. Robitaille, I. Kamwa, A. Oussedik, M. de Montigny, N. Menemenlis, M. Huneault, A. Forcione, R. Mailhot, J. Bourret, and L. Bernier, 'Preliminary Impacts of Wind Power Integration in the Hydro-Quebec System', *Wind Engineering*, vol. 36, no. 1, pp. 35–52, Feb. 2012.

[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', Deutsche Energie - Agentur GmbH (dena), Berlin, Germany, 2010.

[11] S. Hagspiel, A. Papaemannouil, M. Schmid, and G. Andersson, 'Copula-based modeling of stochastic wind power in Europe and implications for the Swiss power grid', *APPLIED ENERGY KIDLINGTON*, vol. 96, p. 33, Aug. 2012.

[12] NEWIS, 'New England Wind Integration Study', 2010.

[13] F. van Hulle et al, 'TradeWind final report: Integrating Wind, Developing Europe's power market for the large-scale integration of wind power', 2009.

[14] D. J. Cannon, D. J. Brayshaw, J. Methven, P. J. Coker, and D. Lenaghan, 'Using reanalysis data to quantify extreme wind power generation statistics: A 33 year case study in Great Britain', *Renewable Energy*, vol. 75, pp. 767–778, Mar. 2015.

[15] R. Lucchesi, 'File Specification for MERRA Products. GMAO Office Note No. 1 (Version 2.3)', 2012.

[16] J. Olauson and M. Bergkvist, 'Modelling the Swedish wind power production using MERRA reanalysis data', *Renewable Energy*, vol. 76, pp. 717–725, Apr. 2015.

[17] 'Vindstat.com - statistics on Swedish wind power production'. [Online]. Available: http://vindstat.com/. [Accessed: 22-Oct-2014].

[18] 'Energimyndigheten - Marknadsstatistik'. [Online]. Available: http://www.energimyndigheten.se/Foretag/Elcertifikat/Marknadsstatistik/. [Accessed: 22-Oct-2014].

[19] 'Svenska Kraftnät - Elstatistik per elområde'. [Online]. Available: http://www.svk.se/Drift-och-marknad/Statistik/Elstatistik-per-elomrade/. [Accessed: 22-Oct-2014].

[20] S. Liléo, E. Berge, O. Undheim, R. Klinkert, and R. E. Bredesen, 'Long-term correction of wind measurements - State-of-the-art, guidelines and future work', Elforsk, 13:18, Jan. 2013.

[21] 'Analysmodell (MESAN) | SMHI'. [Online]. Available: http://www.smhi.se/klimatdata/oppna-data/meteorologiska-data/analysmodell-mesan-1.30445. [Accessed: 01-Jan-2015].

[22] J. Olauson, H. Bergström, and M. Bergkvist, 'Restoring the missing high-frequency fluctuations in a wind power model based on reanalysis data', *Submitted to Renewable Energy*, Feb. 2015.

[23] 'Svensk Vindenergi - Statistik'. [Online]. Available: http://www.vindkraftsbranschen.se/statistik/2014-2/. [Accessed: 09-Oct-2014].

[24] 'Havsbaserade vindkraftsprojekt'. [Online]. Available: http://havsvind.se/wp-content/uploads/2012/02/Havsbaserad_vindkraft_sammanst%C3%A4llning_2012.pdf. [Accessed: 02-Jan-2015].

[25] H. Bergström and S. Söderberg, 'Wind mapping of Sweden – Summary of results and methods used', 2009.

[26] 'WeatherTech Scandinavia AB'. [Online]. Available: http://www.weathertech.se/. [Accessed: 03-Jan-2015].

[27] 'Satsning på havsbaserad vindkraft'. [Online]. Available: http://www.regeringen.se/sb/d/19695/a/251961. [Accessed: 15-Jan-2015].

[28] J. Olauson, D. Lingfors, M. Bergkvist, and J. Widén, 'Quantifying Variability: A Review of Metrics and a Case Study of Net Load Variability', in *Proceedings for 13th Wind Integration Workshop*, Berlin, Germany, 2014.

[29] 'BPA - Load & Total Wind Generation'. [Online]. Available: http://transmission.bpa.gov/Business/Operations/Wind/default.aspx. [Accessed: 23-Oct-2014].

[30] 'Energinet - Five minutes' metering'. [Online]. Available: http://www.energinet.dk/EN/El/Engrosmarked/Udtraek-af-markedsdata/Sider/Femminutters-maalinger.aspx. [Accessed: 23-Oct-2014].

[31] 'TenneT - Actual and forecast wind energy feed-in'. [Online]. Available: http://www.tennettso.de/site/en/Transparency/publications/network-figures/actual-and-forecast-wind-energy-feed-in. [Accessed: 23-Oct-2014].

[32] '50Hertz - Archive Wind power'. [Online]. Available: http://www.50hertz.com/en/Grid-Data/Wind-power/Archive-Wind-power. [Accessed: 23-Oct-2014].

[33] 'Eirgrid - Wind Generation'. [Online]. Available: http://www.eirgrid.com/operations/systemperformancedata/windgeneration/. [Accessed: 23-Oct-2014].

[34] 'AESO - Wind Power and Internal Load Data'. [Online]. Available: http://www.aeso.ca/gridoperations/20544.html. [Accessed: 23-Oct-2014].

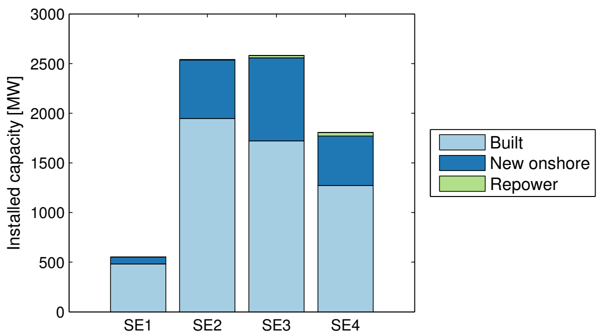
[35] S. Rose and J. Apt, 'Generating wind time series as a hybrid of measured and simulated data', *Wind Energ.*, vol. 15, no. 5, pp. 699–715, 2012.

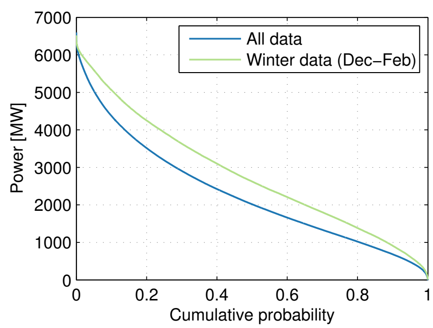
Appendix A - Scenarios

Scenario A1 (base scenario)

Installed capacity: 7478 MW
Yearly energy: 20.1 TWh
Whereof new offshore: 0 TWh
Full load hours new onshore: 3000 h
Full load hours new offshore: –

Keep old WECs built year 2000 or later
Probability weight north relative south: 0.8



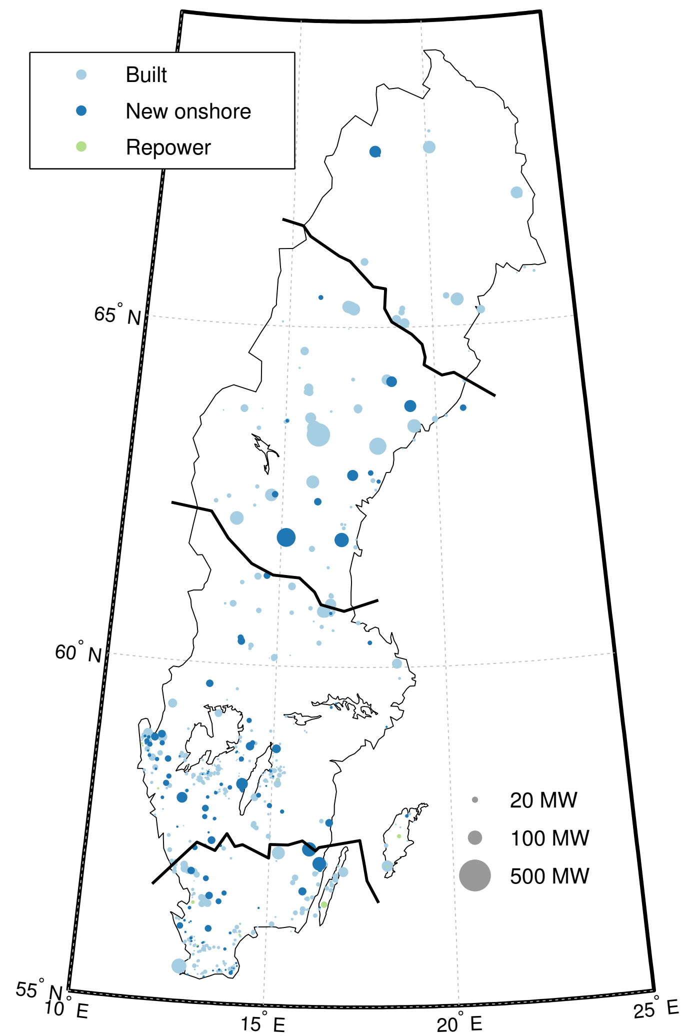


Results

Full load hours: 2689 h
Capacity factor: 0.31
SD of hourly energy: 1399 MW
SD of dP: 157 MW

1st percentile: 224 MW
1st percentile (winter): 284 MW
1st percentile dP: –387 MW
99th percentile dP: 411 MW

(SD is short for standard deviation, dP for one hour step change)



Scenario A2 (more in north)

Installed capacity: 7464 MW
Yearly energy: 20.1 TWh
Whereof new offshore: 0 TWh
Full load hours new onshore: 3000 h
Full load hours new offshore: –

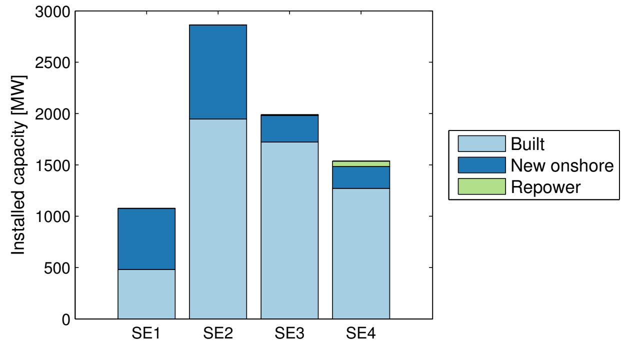
Keep old WECs built year 2000 or later
Probability weight north relative south: 1.8

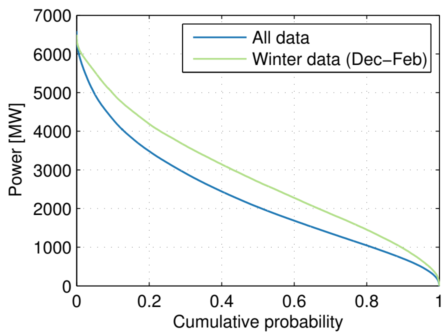
Results

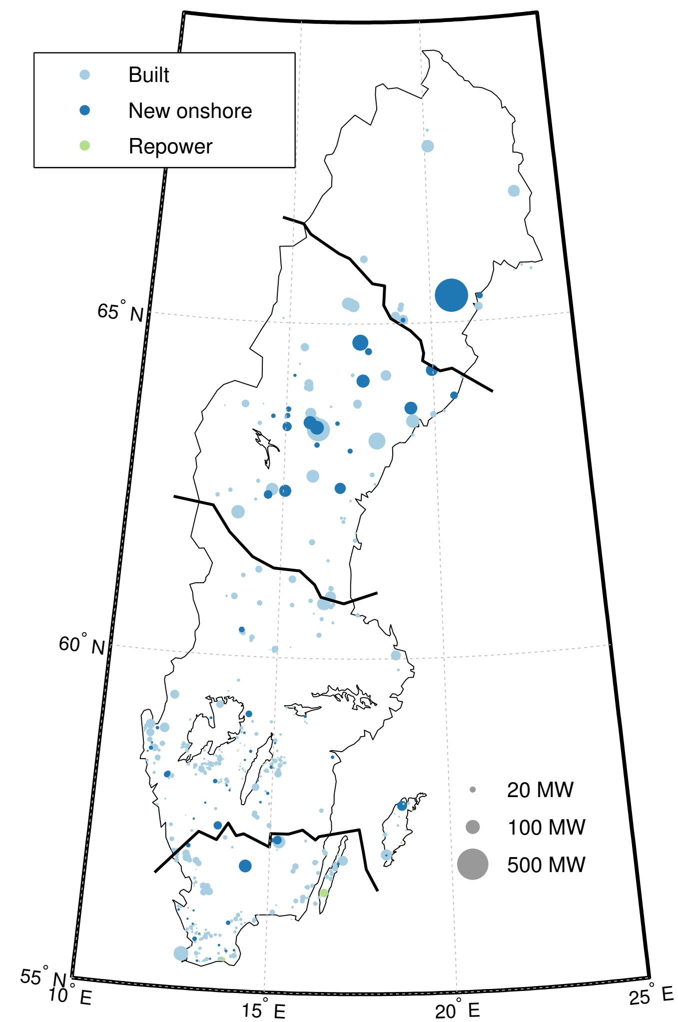
Full load hours: 2695 h
Capacity factor: 0.31
SD of hourly energy: 1372 MW
SD of dP: 157 MW

1st percentile: 230 MW
1st percentile (winter): 319 MW
1st percentile dP: –387 MW
99th percentile dP: 409 MW

(SD is short for standard deviation, dP for one hour step change)



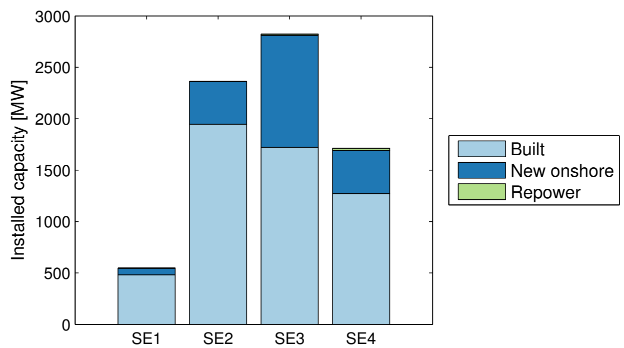


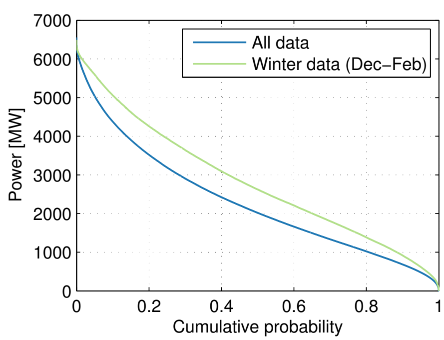


Scenario A3 (less in north)

Installed capacity: 7445 MW
Yearly energy: 20.1 TWh
Whereof new offshore: 0 TWh
Full load hours new onshore: 3000 h
Full load hours new offshore: –

Keep old WECs built year 2000 or later
Probability weight north relative south: 0.4



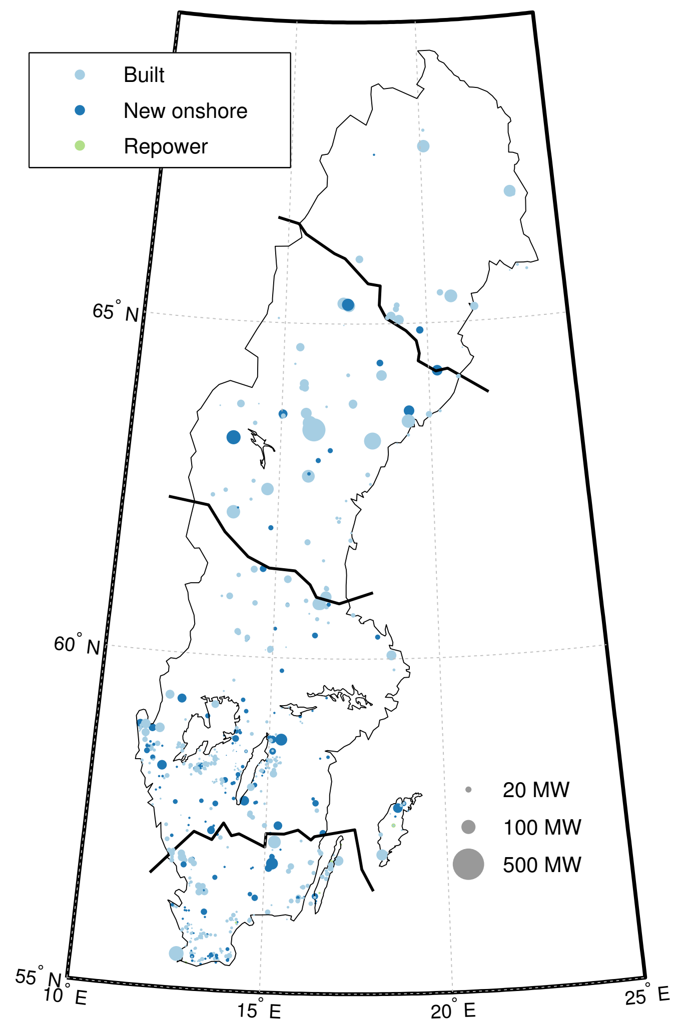


Results

Full load hours: 2701 h
Capacity factor: 0.31
SD of hourly energy: 1400 MW
SD of dP: 156 MW

1st percentile: 223 MW
1st percentile (winter): 289 MW
1st percentile dP: –384 MW
99th percentile dP: 408 MW

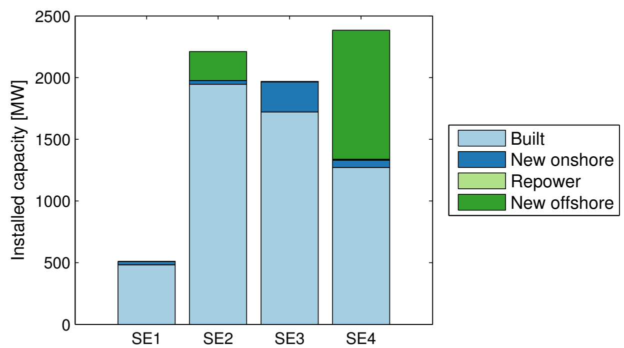
(SD is short for standard deviation, dP for one hour step change)

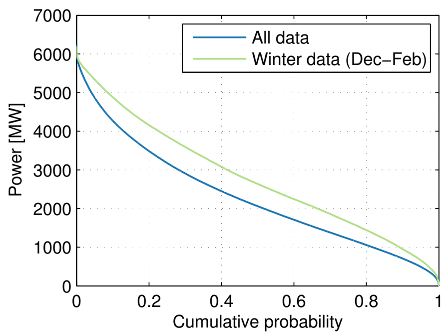


Scenario A4 (more offshore)

Installed capacity: 7073 MW
Yearly energy: 20.1 TWh
Whereof new offshore: 5 TWh
Full load hours new onshore: 3000 h
Full load hours new offshore: 4000 h

Keep old WECs built year 2000 or later
Probability weight north relative south: 0.8



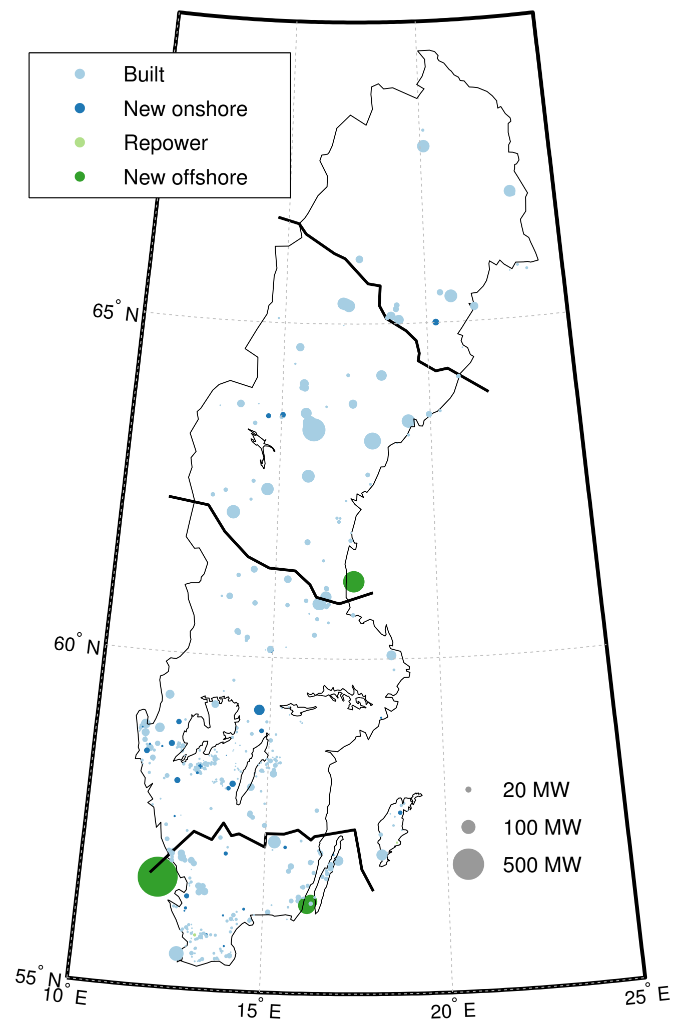


Results

Full load hours: 2839 h
Capacity factor: 0.32
SD of hourly energy: 1340 MW
SD of dP: 147 MW

1st percentile: 232 MW
1st percentile (winter): 318 MW
1st percentile dP: –361 MW
99th percentile dP: 383 MW

(SD is short for standard deviation, dP for one hour step change)



Scenario A5 (higher c.f.)

Installed capacity: 7138 MW
Yearly energy: 20.1 TWh
Whereof new offshore: 0 TWh
Full load hours new onshore: 3500 h
Full load hours new offshore: –

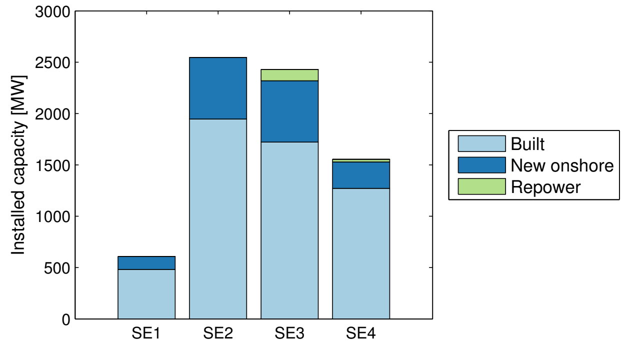
Keep old WECs built year 2000 or later
Probability weight north relative south: 0.8

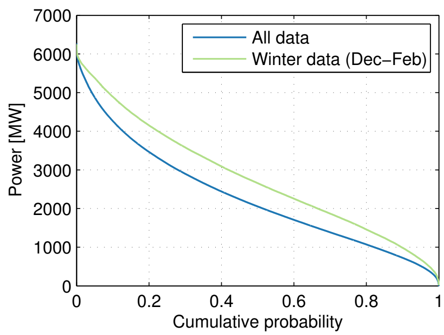
Results

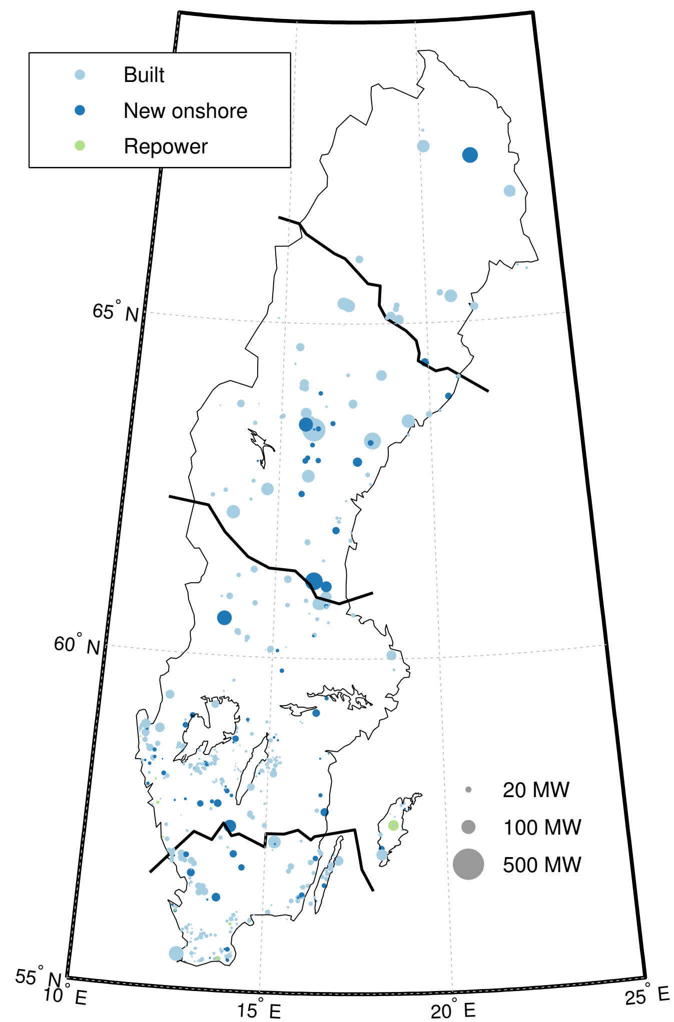
Full load hours: 2813 h
Capacity factor: 0.32
SD of hourly energy: 1335 MW
SD of dP: 150 MW

1st percentile: 247 MW
1st percentile (winter): 332 MW
1st percentile dP: –368 MW
99th percentile dP: 393 MW

(SD is short for standard deviation, dP for one hour step change)



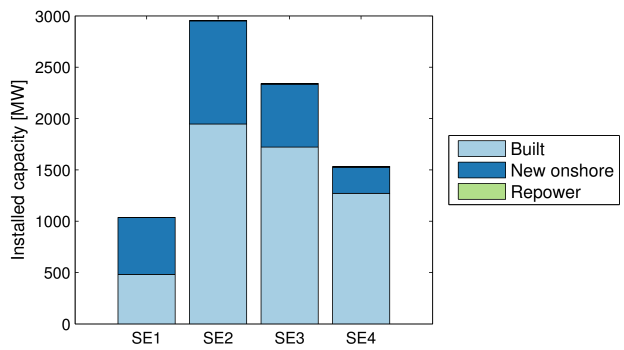


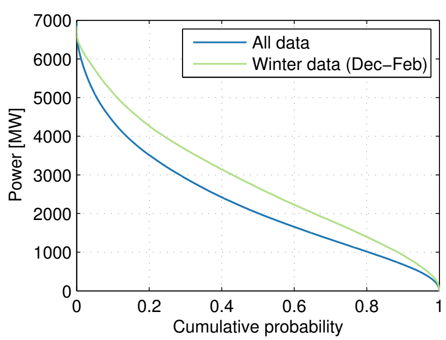


Scenario A6 (lower c.f.)

Installed capacity: 7862 MW
Yearly energy: 20.1 TWh
Whereof new offshore: 0 TWh
Full load hours new onshore: 2500 h
Full load hours new offshore: –

Keep old WECs built year 2000 or later
Probability weight north relative south: 0.8



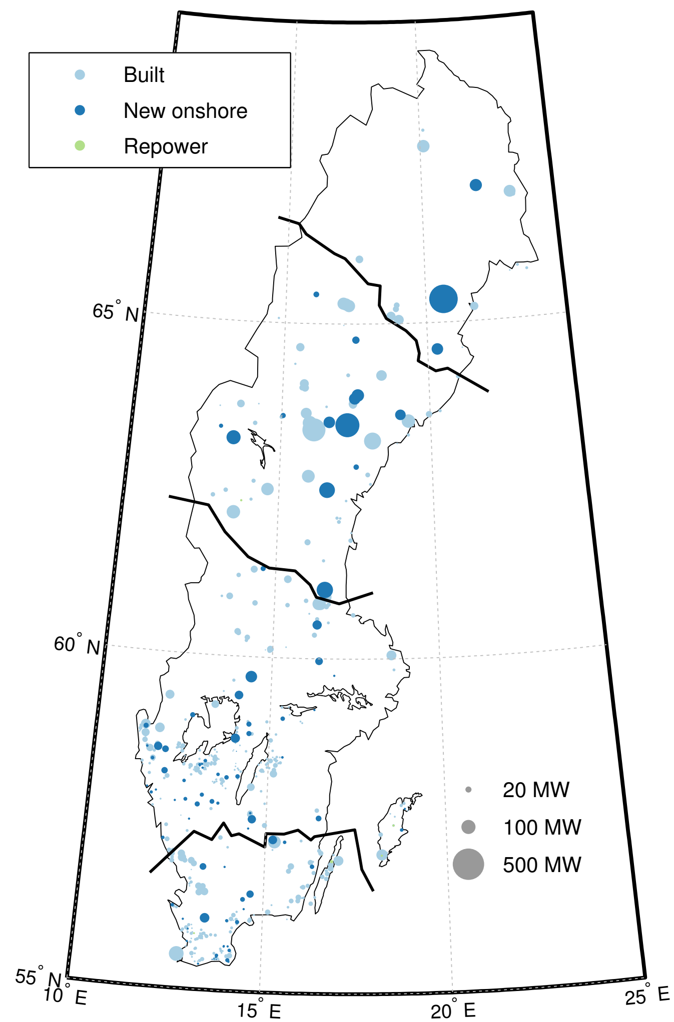


Results

Full load hours: 2563 h
Capacity factor: 0.29
SD of hourly energy: 1420 MW
SD of dP: 162 MW

1st percentile: 216 MW
1st percentile (winter): 287 MW
1st percentile dP: –401 MW
99th percentile dP: 426 MW

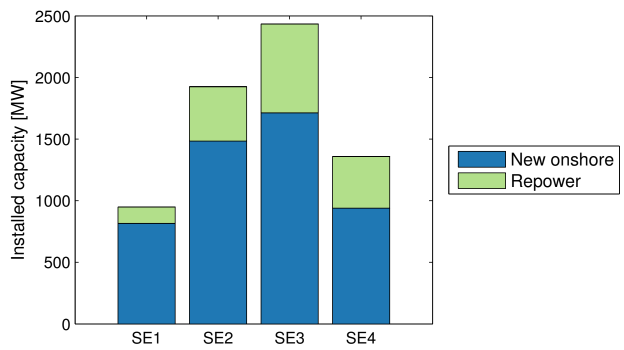
(SD is short for standard deviation, dP for one hour step change)

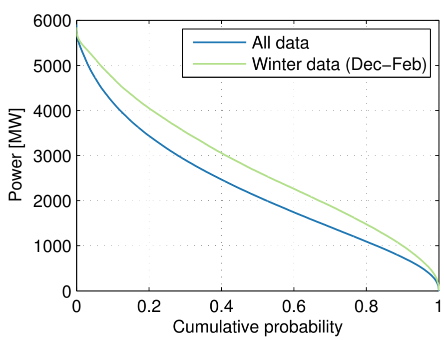


Scenario A7 (only new WECs)

Installed capacity: 6666 MW
Yearly energy: 20 TWh
Whereof new offshore: 0 TWh
Full load hours new onshore: 3000 h
Full load hours new offshore: –

Keep old WECs built year 2020 or later
Probability weight north relative south: 0.8



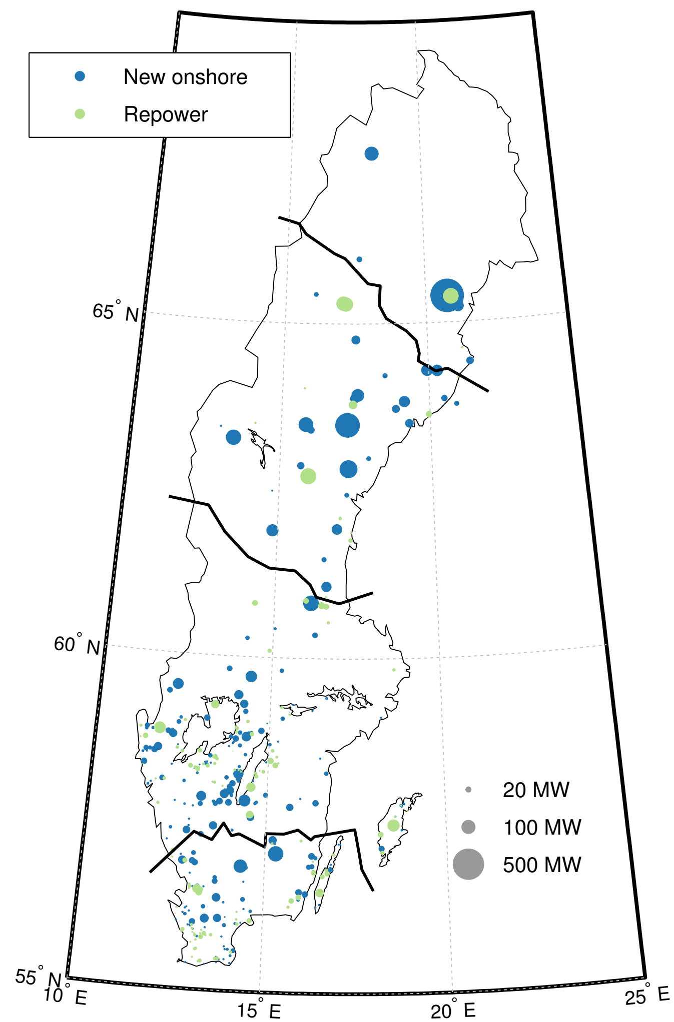


Results

Full load hours: 3007 h
Capacity factor: 0.34
SD of hourly energy: 1287 MW
SD of dP: 150 MW

1st percentile: 248 MW
1st percentile (winter): 331 MW
1st percentile dP: –365 MW
99th percentile dP: 388 MW

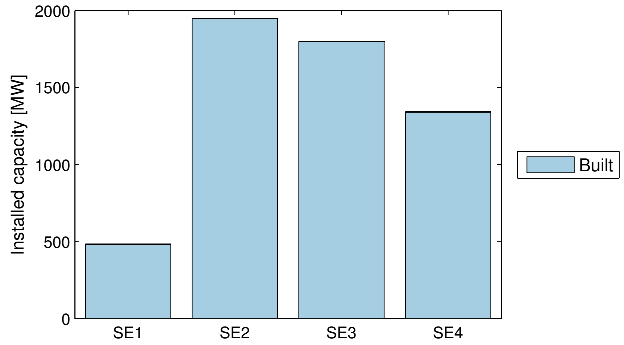
(SD is short for standard deviation, dP for one hour step change)

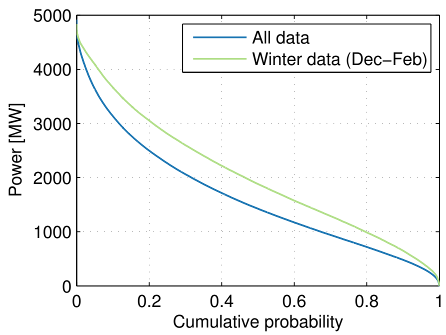


Scenario A8 (existing WECs)

Installed capacity: 5572 MW
Yearly energy: 14.3 TWh
Whereof new offshore: 0 TWh
Full load hours new onshore: –
Full load hours new offshore: –

Keep old WECs built year 1900 or later
Probability weight north relative south: 0.8



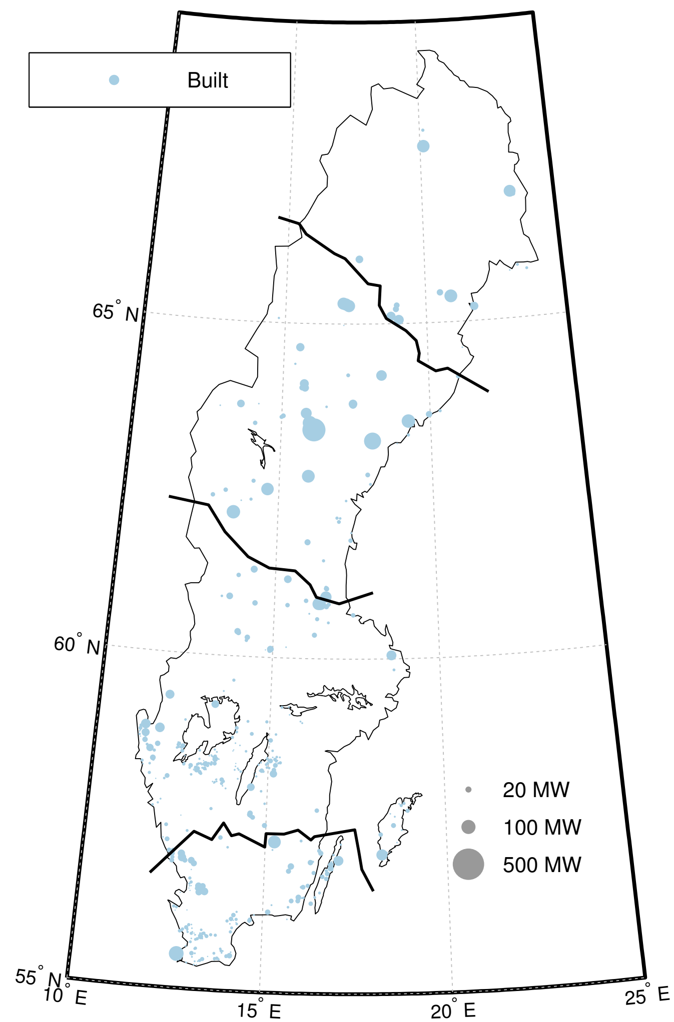


Results

Full load hours: 2569 h
Capacity factor: 0.29
SD of hourly energy: 1015 MW
SD of dP: 113 MW

1st percentile: 155 MW
1st percentile (winter): 205 MW
1st percentile dP: −278 MW
99th percentile dP: 298 MW

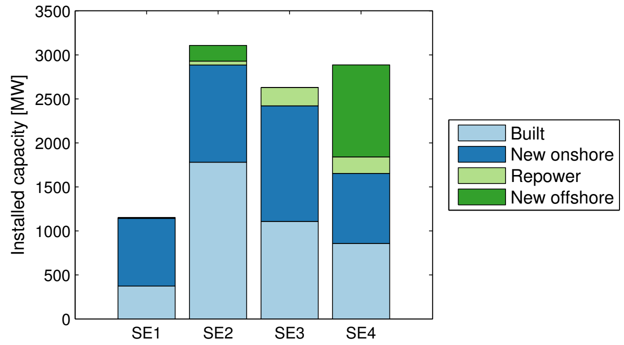
(SD is short for standard deviation, dP for one hour step change)

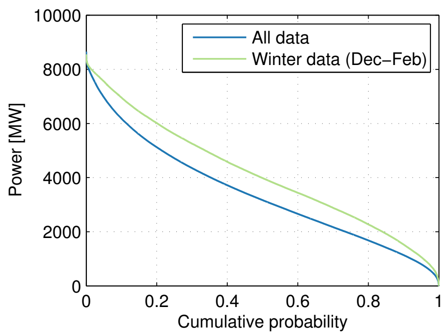


Scenario B1 (base scenario)

Installed capacity: 9773 MW
Yearly energy: 30 TWh
Whereof new offshore: 5 TWh
Full load hours new onshore: 3200 h
Full load hours new offshore: 4100 h

Keep old WECs built year 2010 or later
Probability weight north relative south: 0.8



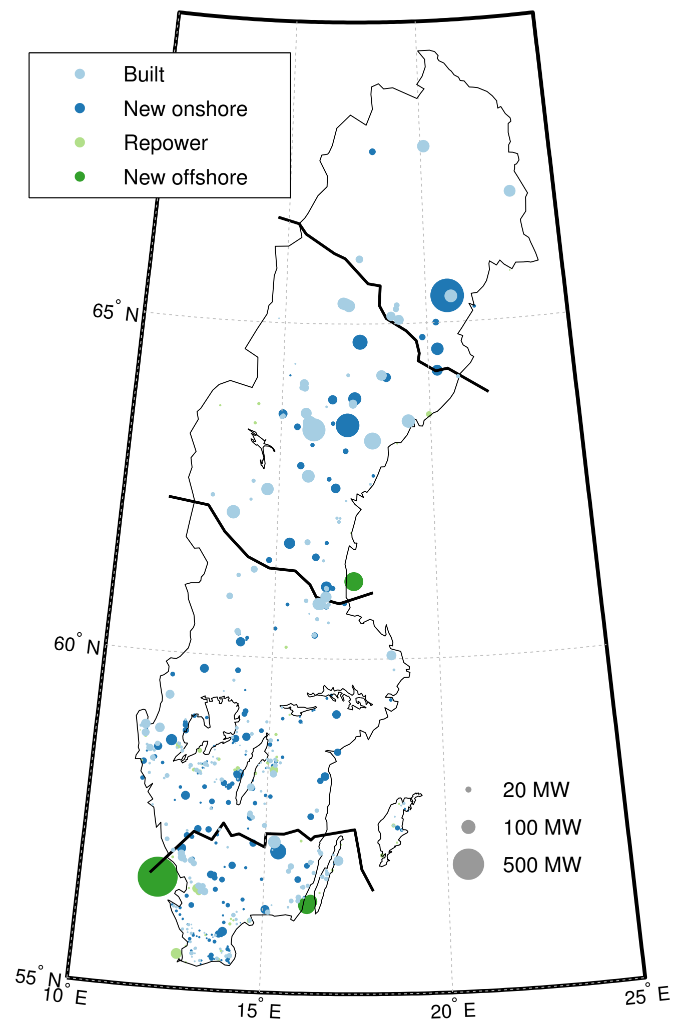


Results

Full load hours: 3076 h
Capacity factor: 0.35
SD of hourly energy: 1881 MW
SD of dP: 215 MW

1st percentile: 383 MW
1st percentile (winter): 528 MW
1st percentile dP: −522 MW
99th percentile dP: 550 MW

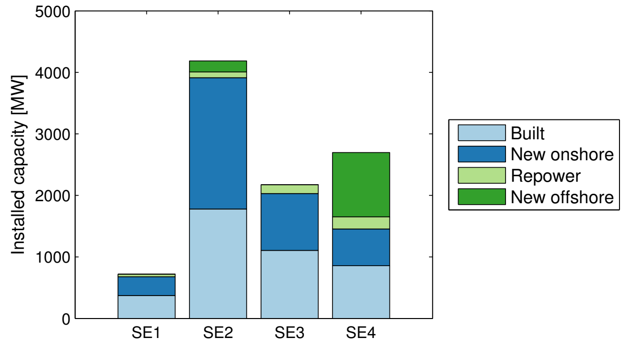
(SD is short for standard deviation, dP for one hour step change)

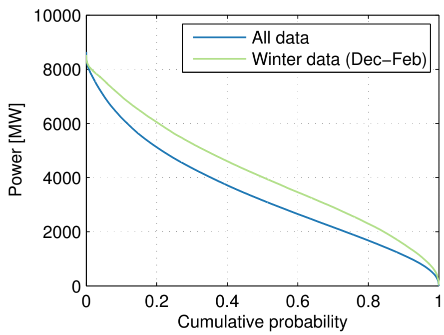


Scenario B2 (more in north)

Installed capacity: 9774 MW
Yearly energy: 30.1 TWh
Whereof new offshore: 5 TWh
Full load hours new onshore: 3200 h
Full load hours new offshore: 4100 h

Keep old WECs built year 2010 or later
Probability weight north relative south: 1.8



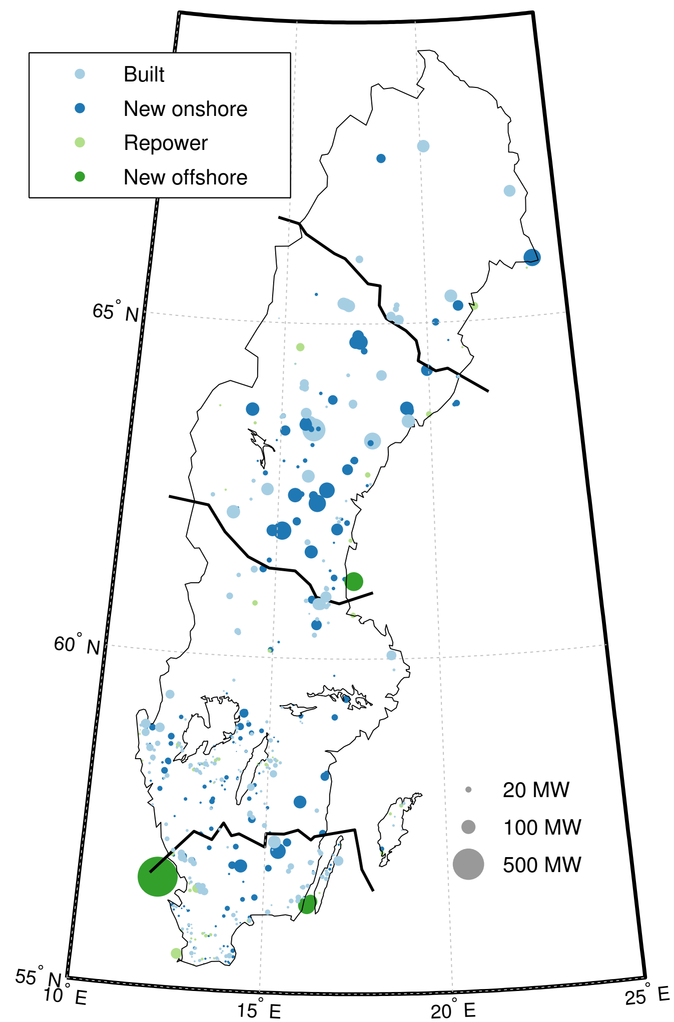


Results

Full load hours: 3078 h
Capacity factor: 0.35
SD of hourly energy: 1890 MW
SD of dP: 219 MW

1st percentile: 379 MW
1st percentile (winter): 528 MW
1st percentile dP: –535 MW
99th percentile dP: 565 MW

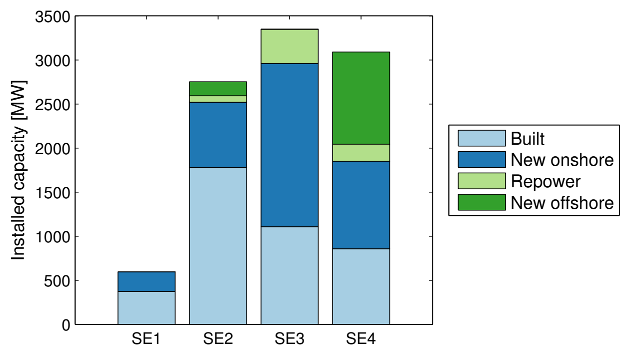
(SD is short for standard deviation, dP for one hour step change)

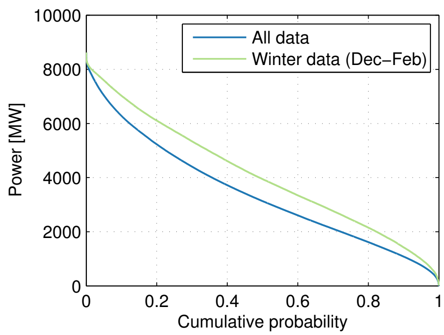


Scenario B3 (less in north)

Installed capacity: 9787 MW
Yearly energy: 30 TWh
Whereof new offshore: 5 TWh
Full load hours new onshore: 3200 h
Full load hours new offshore: 4100 h

Keep old WECs built year 2010 or later
Probability weight north relative south: 0.4



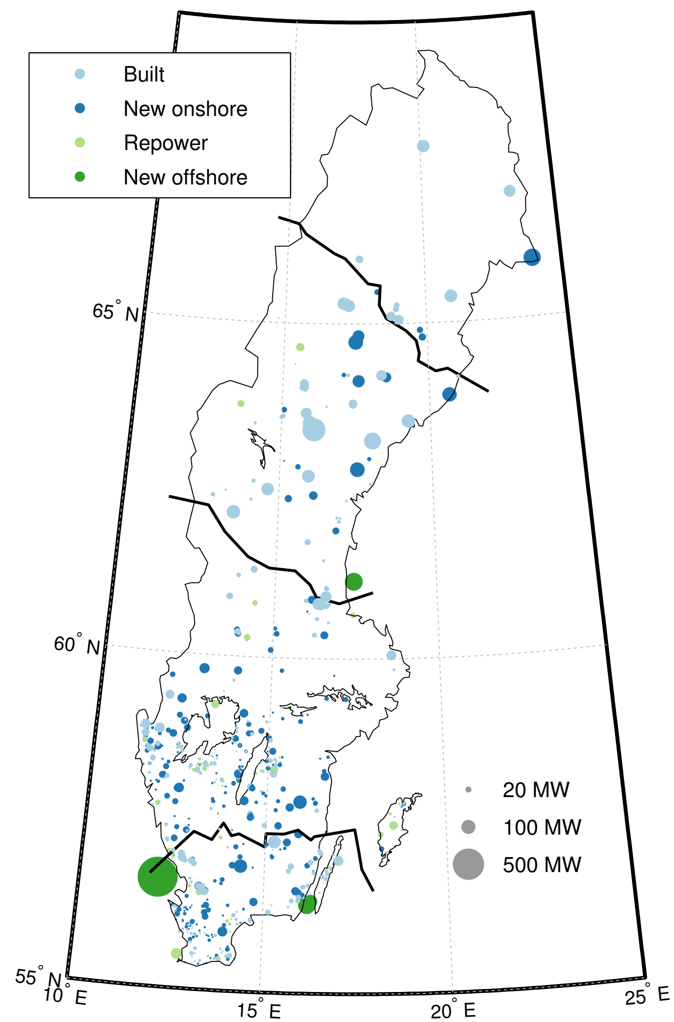


Results

Full load hours: 3072 h
Capacity factor: 0.35
SD of hourly energy: 1935 MW
SD of dP: 216 MW

1st percentile: 357 MW
1st percentile (winter): 463 MW
1st percentile dP: –531 MW
99th percentile dP: 556 MW

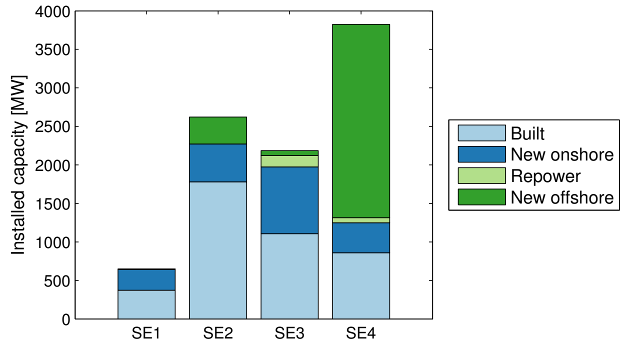
(SD is short for standard deviation, dP for one hour step change)

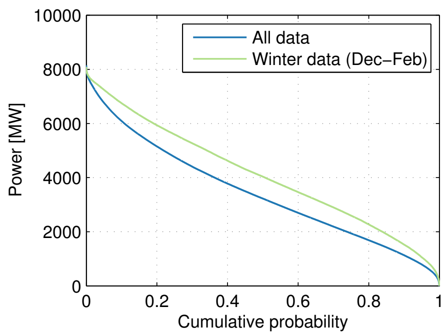


Scenario B4 (more offshore)

Installed capacity: 9282 MW
Yearly energy: 30.1 TWh
Whereof new offshore: 12 TWh
Full load hours new onshore: 3200 h
Full load hours new offshore: 4100 h

Keep old WECs built year 2010 or later
Probability weight north relative south: 0.8



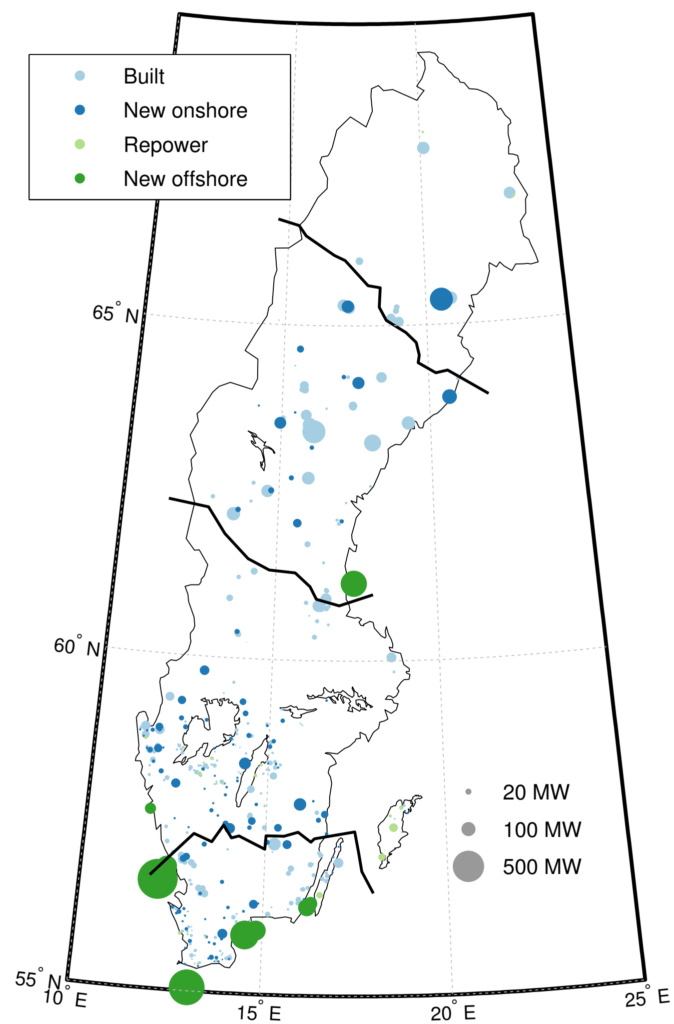


Results

Full load hours: 3240 h
Capacity factor: 0.37
SD of hourly energy: 1836 MW
SD of dP: 205 MW

1st percentile: 379 MW
1st percentile (winter): 515 MW
1st percentile dP: –505 MW
99th percentile dP: 527 MW

(SD is short for standard deviation, dP for one hour step change)



Scenario B5 (no offshore)

Installed capacity: 10097 MW
Yearly energy: 30.1 TWh
Whereof new offshore: 0 TWh
Full load hours new onshore: 3200 h
Full load hours new offshore: –

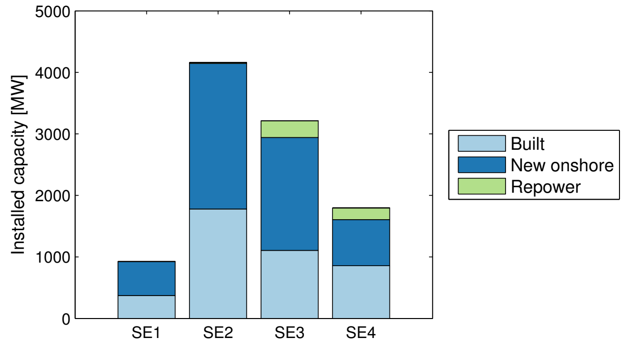
Keep old WECs built year 2010 or later
Probability weight north relative south: 0.8

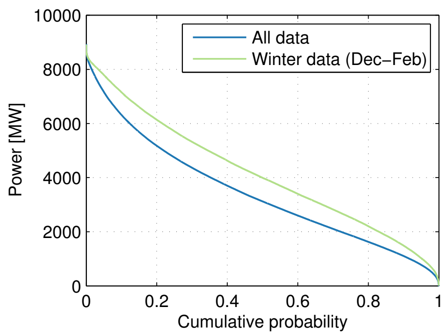
Results

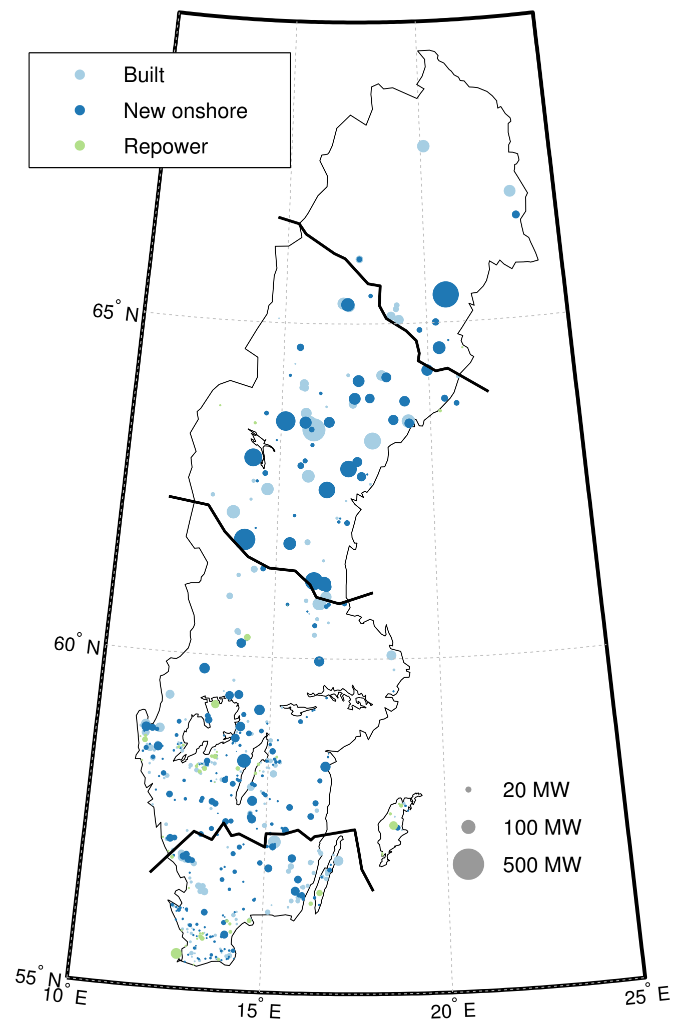
Full load hours: 2981 h
Capacity factor: 0.34
SD of hourly energy: 1952 MW
SD of dP: 228 MW

1st percentile: 369 MW
1st percentile (winter): 492 MW
1st percentile dP: –559 MW
99th percentile dP: 588 MW

(SD is short for standard deviation, dP for one hour step change)



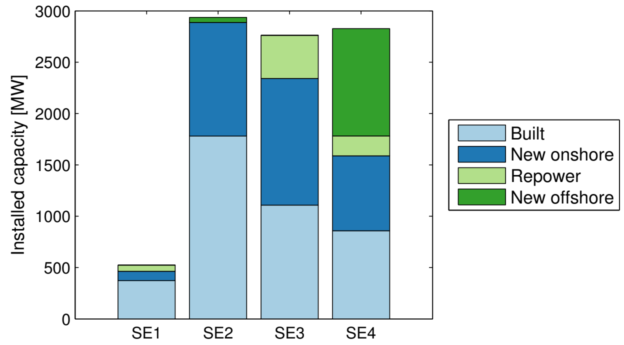


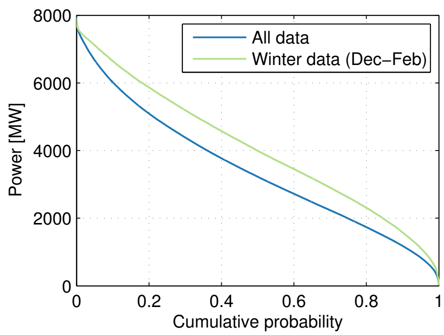


Scenario B6 (higher c.f.)

Installed capacity: 9047 MW
Yearly energy: 30 TWh
Whereof new offshore: 5 TWh
Full load hours new onshore: 3700 h
Full load hours new offshore: 4600 h

Keep old WECs built year 2010 or later
Probability weight north relative south: 0.8



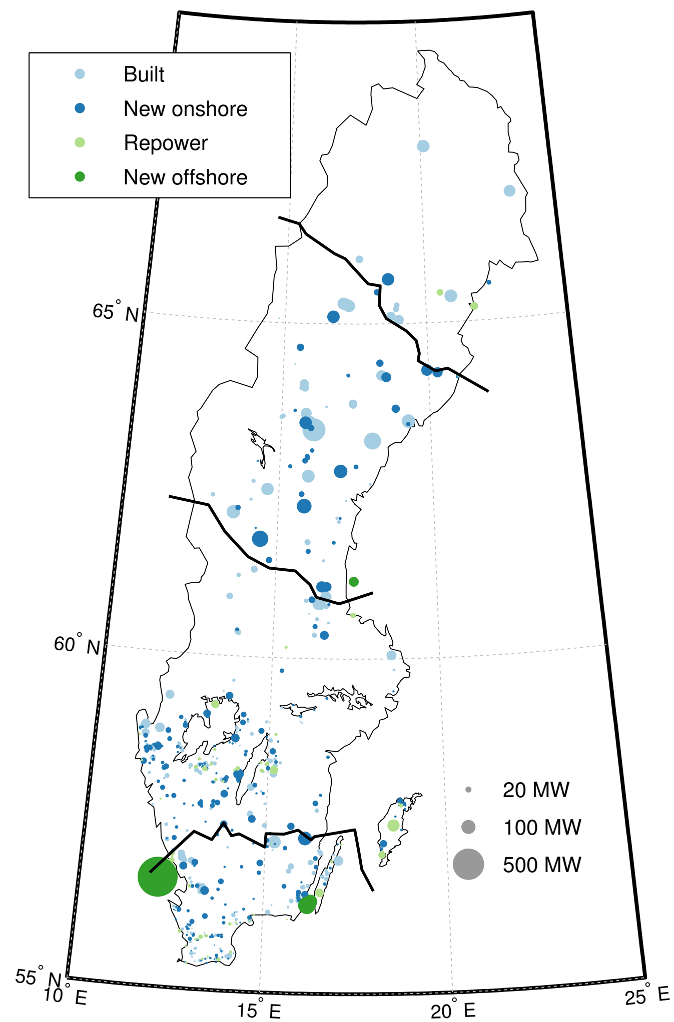


Results

Full load hours: 3320 h
Capacity factor: 0.38
SD of hourly energy: 1787 MW
SD of dP: 203 MW

1st percentile: 405 MW
1st percentile (winter): 549 MW
1st percentile dP: –494 MW
99th percentile dP: 518 MW

(SD is short for standard deviation, dP for one hour step change)



Scenario B7 (lower c.f.)

Installed capacity: 10757 MW
Yearly energy: 30.1 TWh
Whereof new offshore: 5 TWh
Full load hours new onshore: 2700 h
Full load hours new offshore: 3600 h

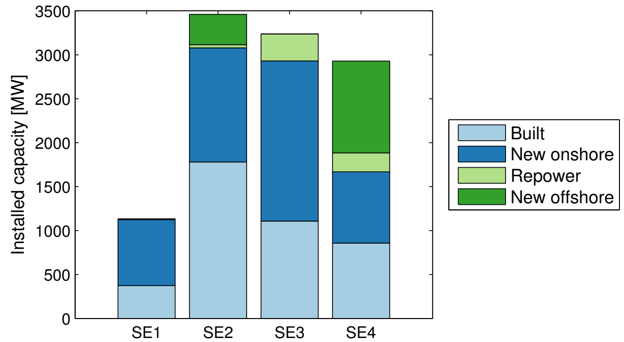
Keep old WECs built year 2010 or later
Probability weight north relative south: 0.8

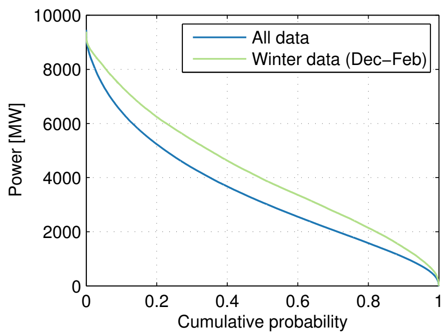
Results

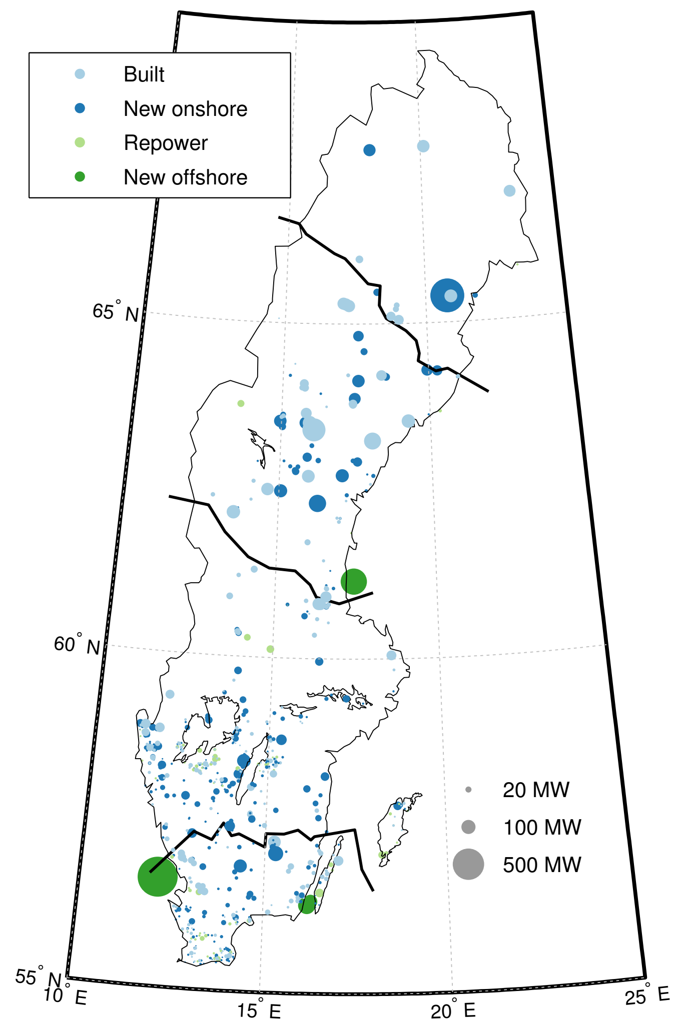
Full load hours: 2805 h
Capacity factor: 0.32
SD of hourly energy: 2028 MW
SD of dP: 226 MW

1st percentile: 346 MW
1st percentile (winter): 466 MW
1st percentile dP: −557 MW
99th percentile dP: 588 MW

(SD is short for standard deviation, dP for one hour step change)



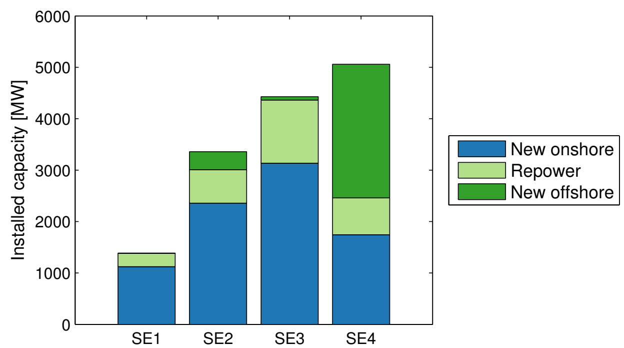


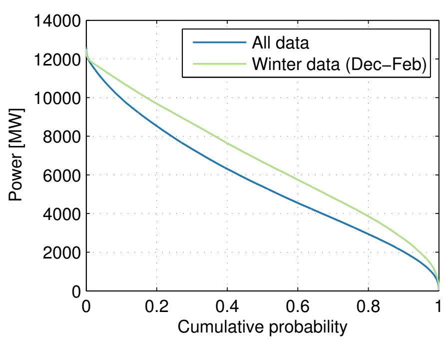


Scenario C1 (base scenario)

Installed capacity: 14231 MW
Yearly energy: 50 TWh
Whereof new offshore: 13 TWh
Full load hours new onshore: 3300 h
Full load hours new offshore: 4300 h

Keep old WECs built year 2020 or later
Probability weight north relative south: 0.8



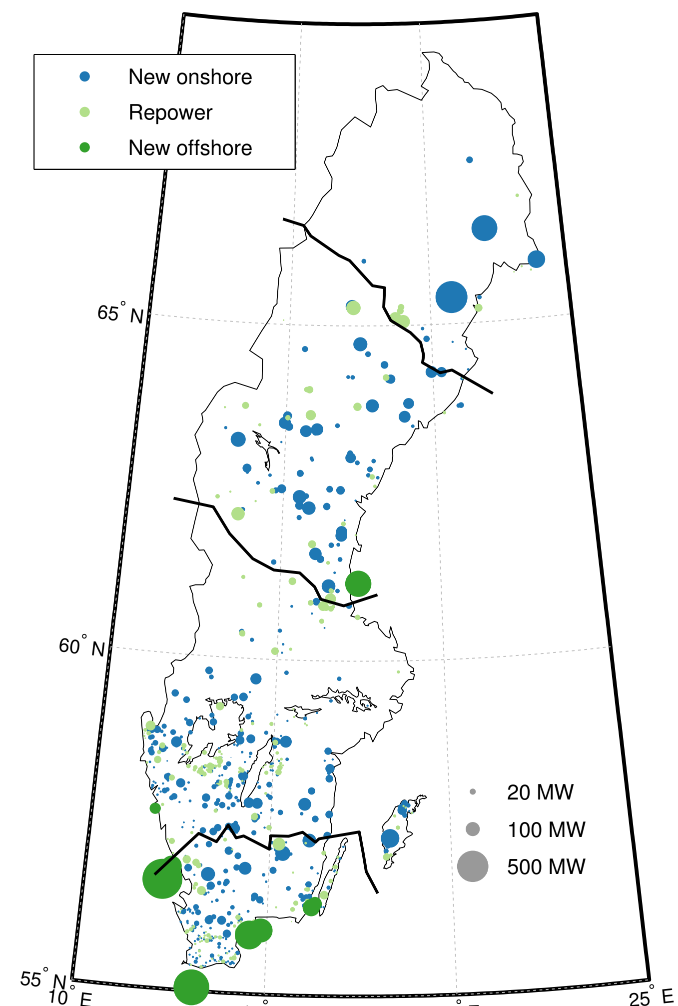


Results

Full load hours: 3516 h
Capacity factor: 0.4
SD of hourly energy: 2909 MW
SD of dP: 329 MW

1st percentile: 717 MW
1st percentile (winter): 971 MW
1st percentile dP: –801 MW
99th percentile dP: 841 MW

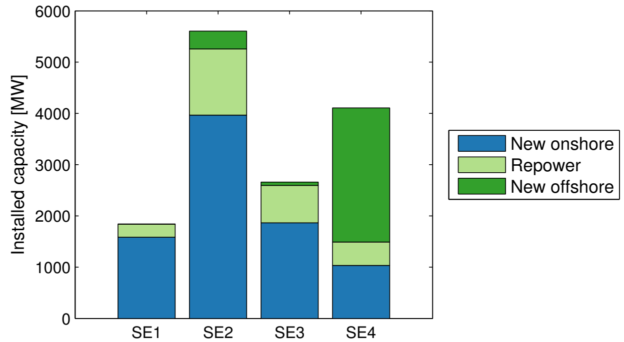
(SD is short for standard deviation, dP for one hour step change)

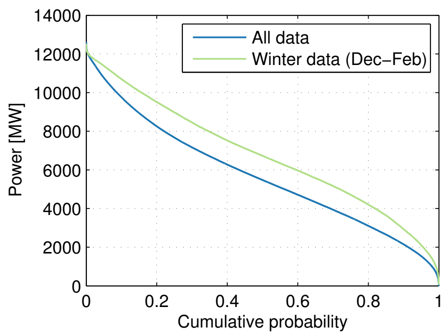


Scenario C2 (more in north)

Installed capacity: 14214 MW
Yearly energy: 50.1 TWh
Whereof new offshore: 13 TWh
Full load hours new onshore: 3300 h
Full load hours new offshore: 4300 h

Keep old WECs built year 2020 or later
Probability weight north relative south: 1.8



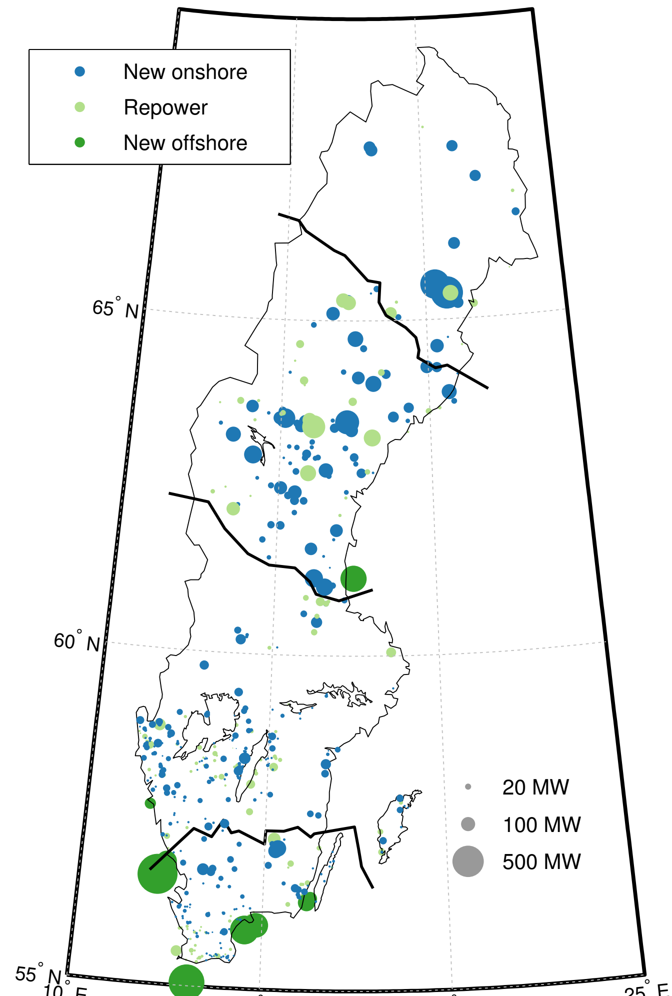


Results

Full load hours: 3524 h
Capacity factor: 0.4
SD of hourly energy: 2786 MW
SD of dP: 330 MW

1st percentile: 756 MW
1st percentile (winter): 1083 MW
1st percentile dP: −801 MW
99th percentile dP: 834 MW

(SD is short for standard deviation, dP for one hour step change)



Scenario C3 (less in north)

Installed capacity: 14208 MW
Yearly energy: 50 TWh
Whereof new offshore: 13 TWh
Full load hours new onshore: 3300 h
Full load hours new offshore: 4300 h

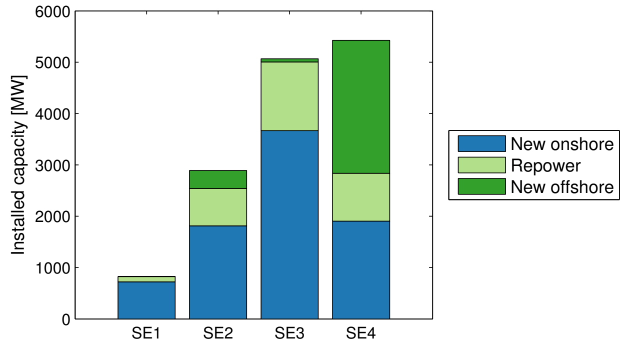
Keep old WECs built year 2020 or later
Probability weight north relative south: 0.4

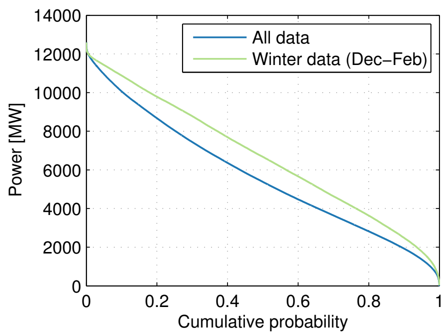
Results

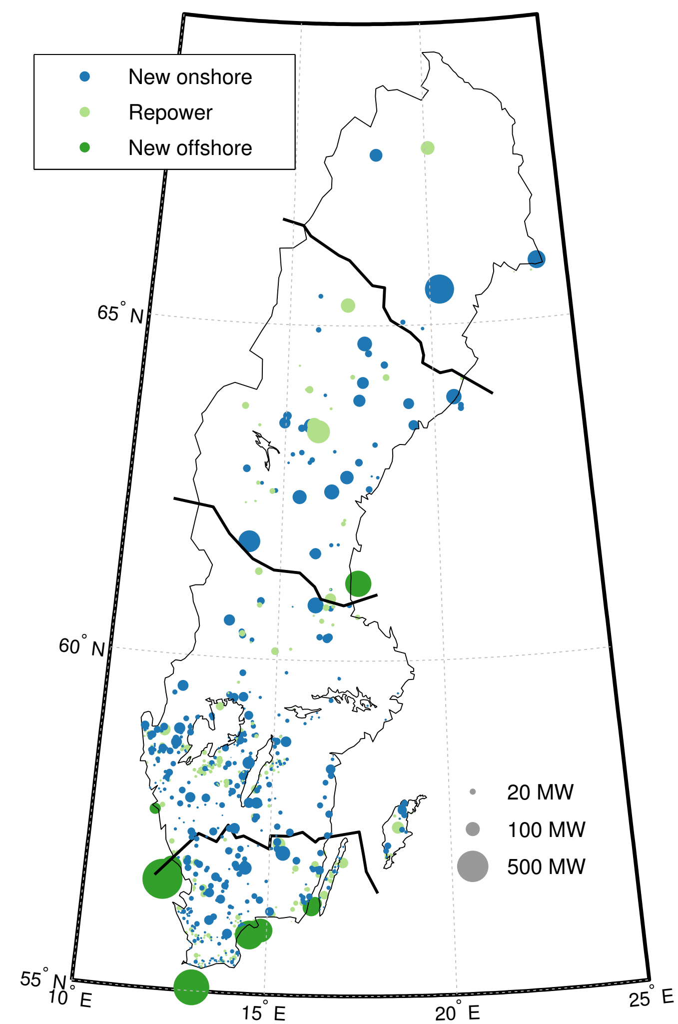
Full load hours: 3521 h
Capacity factor: 0.4
SD of hourly energy: 2989 MW
SD of dP: 338 MW

1st percentile: 681 MW
1st percentile (winter): 851 MW
1st percentile dP: −831 MW
99th percentile dP: 870 MW

(SD is short for standard deviation, dP for one hour step change)



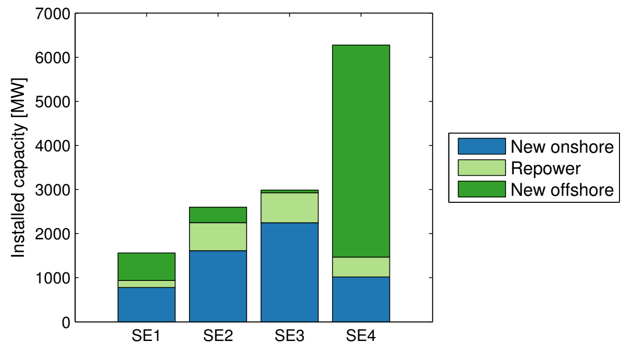


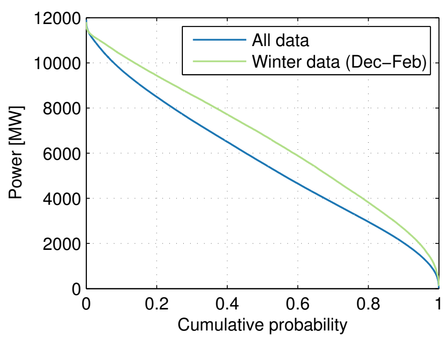


Scenario C4 (more offshore)

Installed capacity: 13426 MW
Yearly energy: 50.1 TWh
Whereof new offshore: 25 TWh
Full load hours new onshore: 3300 h
Full load hours new offshore: 4300 h

Keep old WECs built year 2020 or later
Probability weight north relative south: 0.8



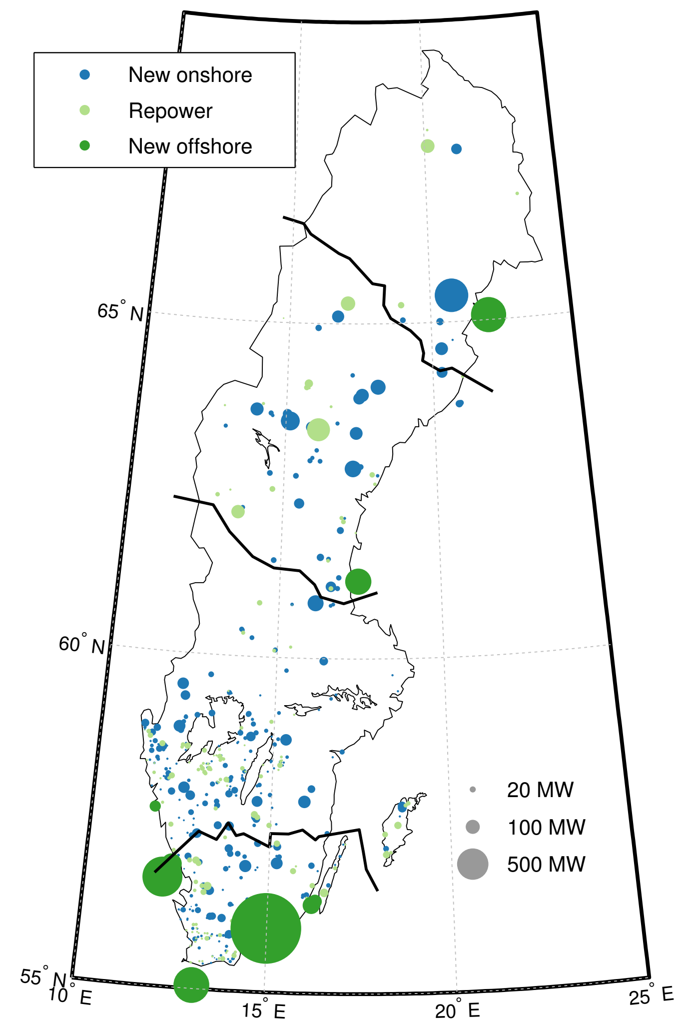


Results

Full load hours: 3731 h
Capacity factor: 0.43
SD of hourly energy: 2817 MW
SD of dP: 337 MW

1st percentile: 697 MW
1st percentile (winter): 935 MW
1st percentile dP: –832 MW
99th percentile dP: 863 MW

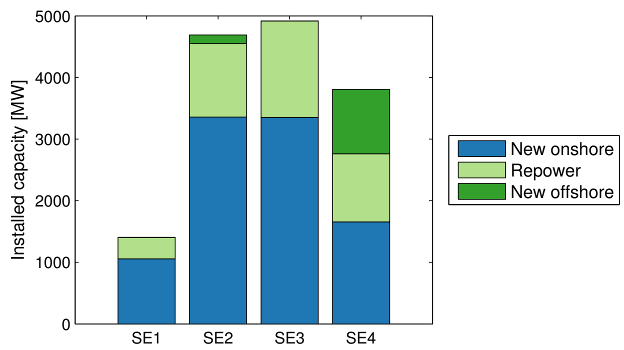
(SD is short for standard deviation, dP for one hour step change)

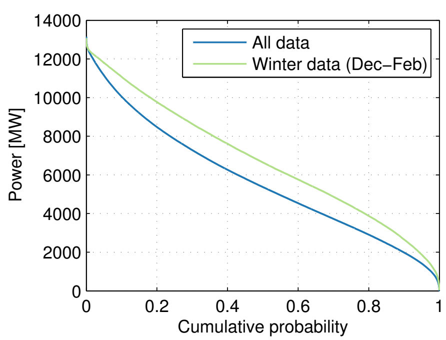


Scenario C5 (less offshore)

Installed capacity: 14818 MW
Yearly energy: 50 TWh
Whereof new offshore: 5 TWh
Full load hours new onshore: 3300 h
Full load hours new offshore: 4300 h

Keep old WECs built year 2020 or later
Probability weight north relative south: 0.8



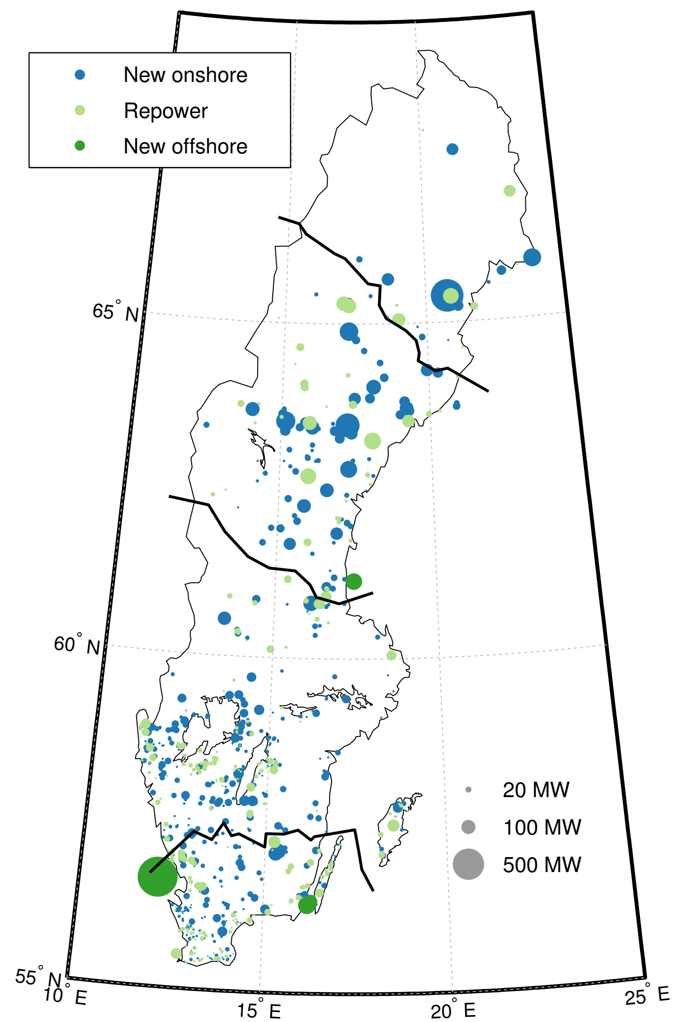


Results

Full load hours: 3377 h
Capacity factor: 0.39
SD of hourly energy: 2971 MW
SD of dP: 342 MW

1st percentile: 700 MW
1st percentile (winter): 960 MW
1st percentile dP: –826 MW
99th percentile dP: 877 MW

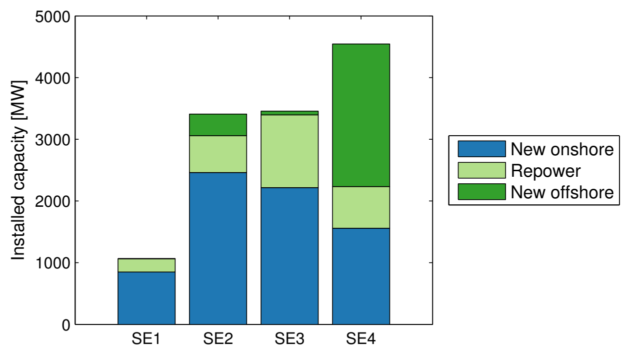
(SD is short for standard deviation, dP for one hour step change)

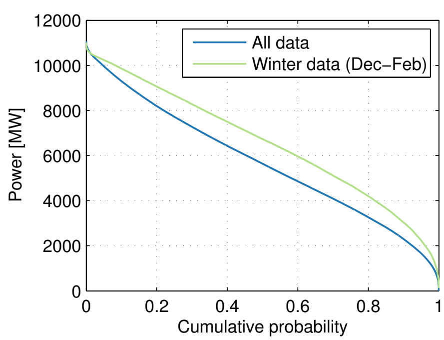


Scenario C6 (higher c.f.)

Installed capacity: 12475 MW
Yearly energy: 50 TWh
Whereof new offshore: 13 TWh
Full load hours new onshore: 3800 h
Full load hours new offshore: 4800 h

Keep old WECs built year 2020 or later
Probability weight north relative south: 0.8



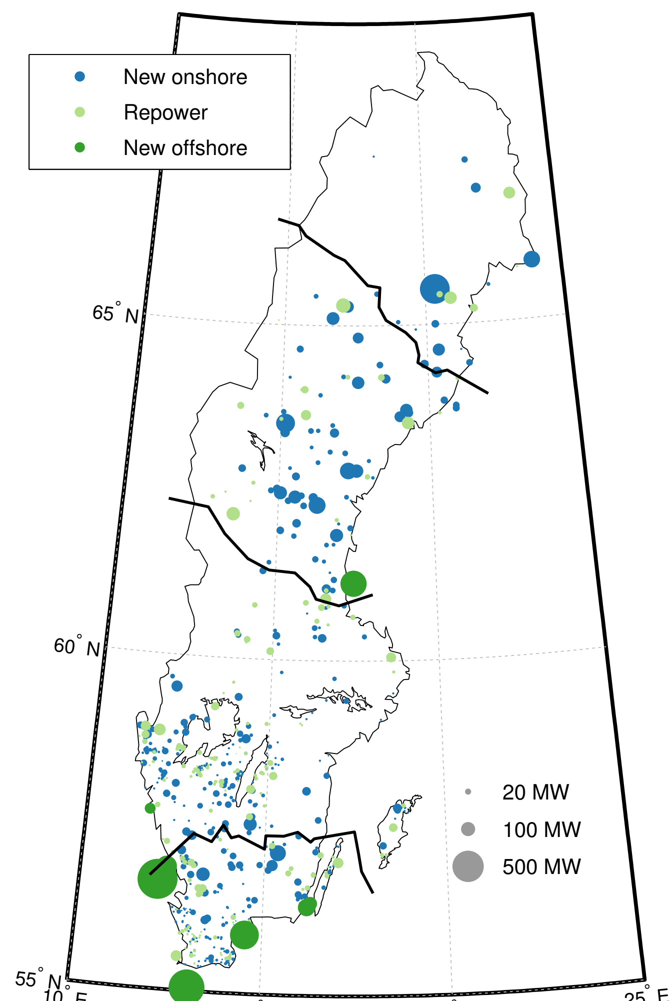


Results

Full load hours: 4014 h
Capacity factor: 0.46
SD of hourly energy: 2554 MW
SD of dP: 303 MW

1st percentile: 837 MW
1st percentile (winter): 1162 MW
1st percentile dP: −735 MW
99th percentile dP: 769 MW

(SD is short for standard deviation, dP for one hour step change)



Scenario C7 (lower c.f.)

Installed capacity: 16625 MW
Yearly energy: 50.1 TWh
Whereof new offshore: 13 TWh
Full load hours new onshore: 2800 h
Full load hours new offshore: 3800 h

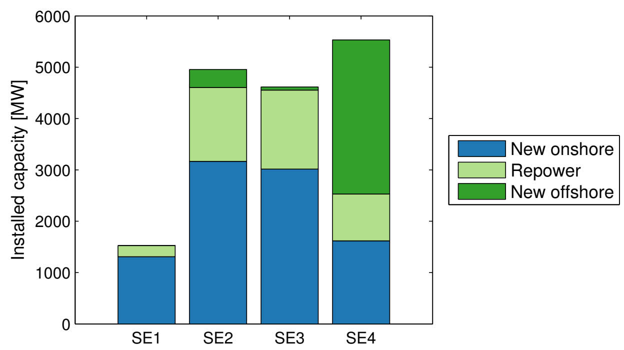
Keep old WECs built year 2020 or later
Probability weight north relative south: 0.8

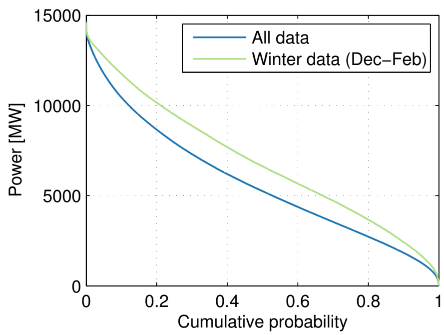
Results

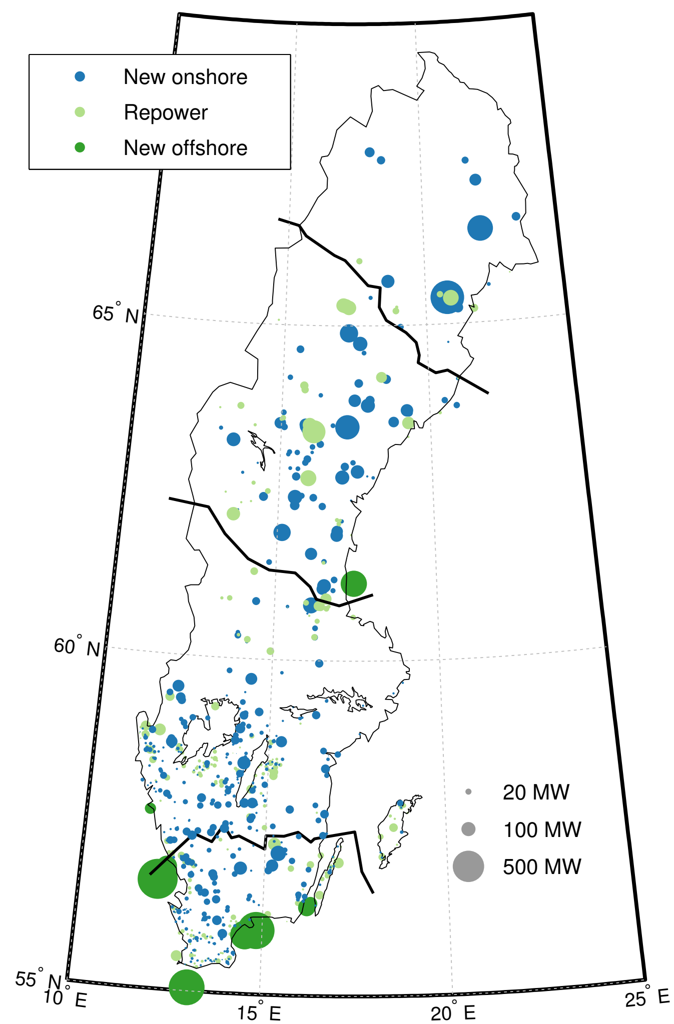
Full load hours: 3018 h
Capacity factor: 0.34
SD of hourly energy: 3211 MW
SD of dP: 361 MW

1st percentile: 614 MW
1st percentile (winter): 796 MW
1st percentile dP: −888 MW
99th percentile dP: 927 MW

(SD is short for standard deviation, dP for one hour step change)



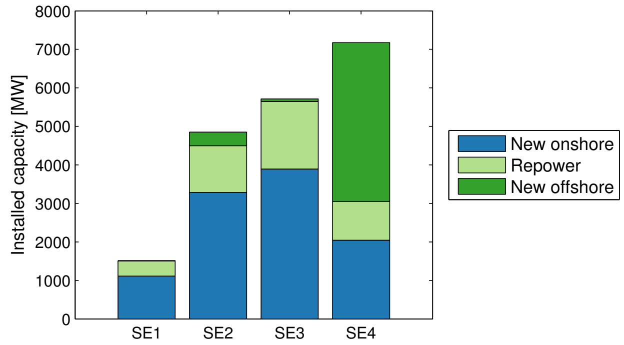


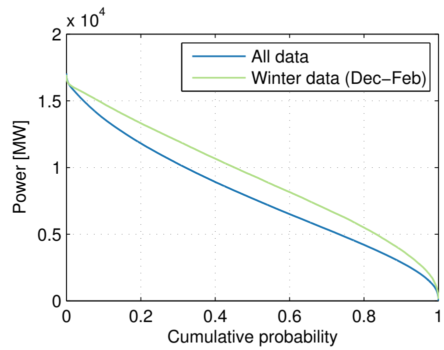


Scenario D1 (70 TWh)

Installed capacity: 19247 MW
Yearly energy: 70 TWh
Whereof new offshore: 20 TWh
Full load hours new onshore: 3400 h
Full load hours new offshore: 4400 h

Keep old WECs built year 2020 or later
Probability weight north relative south: 0.8



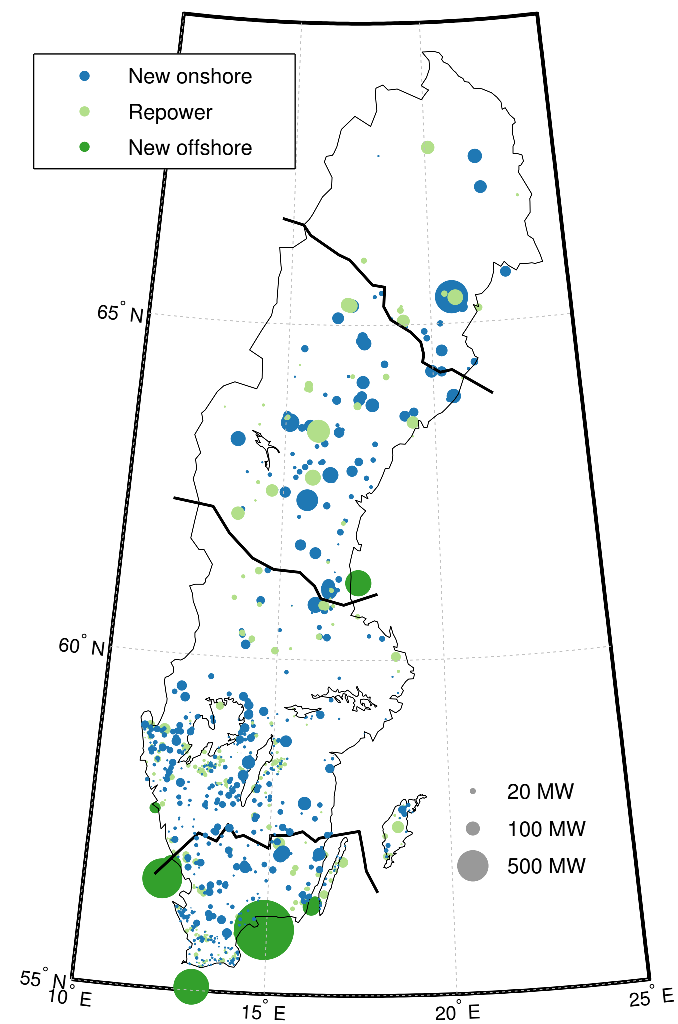


Results

Full load hours: 3641 h
Capacity factor: 0.42
SD of hourly energy: 3957 MW
SD of dP: 459 MW

1st percentile: 1019 MW
1st percentile (winter): 1341 MW
1st percentile dP: −1122 MW
99th percentile dP: 1170 MW

(SD is short for standard deviation, dP for one hour step change)



Appendix B

Renewable Energy 76 (2015) 717–725

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Modelling the Swedish wind power production using MERRA reanalysis data

Jon Olauson*, 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.

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_2 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:

* Corresponding author. The Ångström Laboratory, Box 534, 751 21 Uppsala, Sweden. Tel.: +46 734 125930.
E-mail address: jon.olauson@angstrom.uu.se (J. Olauson).

http://dx.doi.org/10.1016/j.renene.2014.11.085

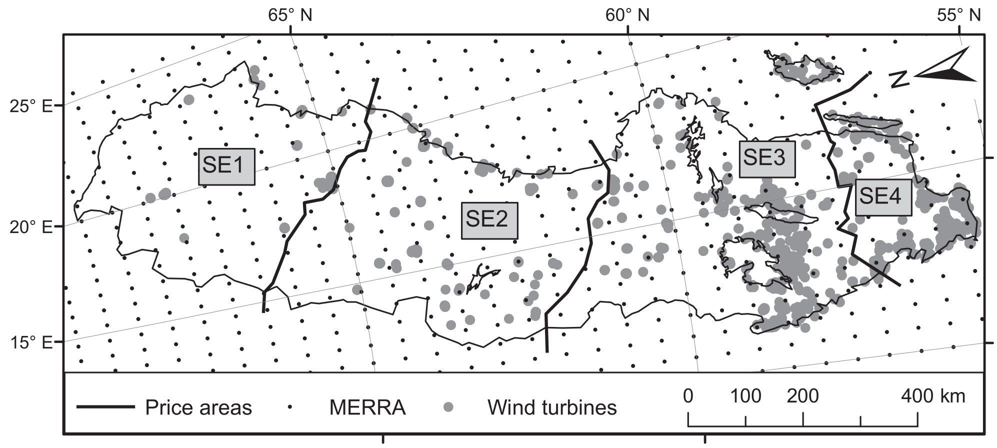


Fig. 1. Figure visualising MERRA grid points, distribution of WECs (end of 2012) and the four price areas in Sweden.

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 $0.5 \times 0.67°$ 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

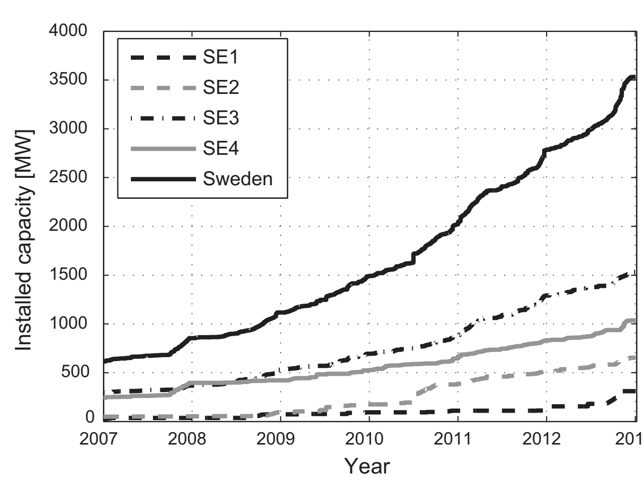


Fig. 2. Installed capacity of WEC in Sweden.

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 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.

Table 1
Parameters used in the model.

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

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

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

where u is wind speed, z is height above ground, d is displacement height and α 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. α is a function of surface roughness, orography, atmospheric stability etc. [28]. Two different model structures were used, controlled by the parameter α_{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 $Loss_w$.

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^2, 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 smoothened 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 smoothening effect of wind variability on aggregated power.

The smoothening effect was parameterised with a standard deviation on the incoming wind speed (σ_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.

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

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

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

where ρ is air density. The model structure parameter ρ_{var} determines if air density should be constant (1.225 kg/m^3) 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

$$\rho = \frac{\rho_0 T_0}{p_0} \cdot \frac{p_{hub}}{T} = \frac{\rho_0 T_0}{p_0 T} \cdot p_{sea} \cdot e^{-\frac{Mg}{RT}h} \approx 0.003484 \cdot \frac{p_{sea}}{T} \cdot e^{-\frac{0.003416 \cdot h}{T}}, \tag{4}$$

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

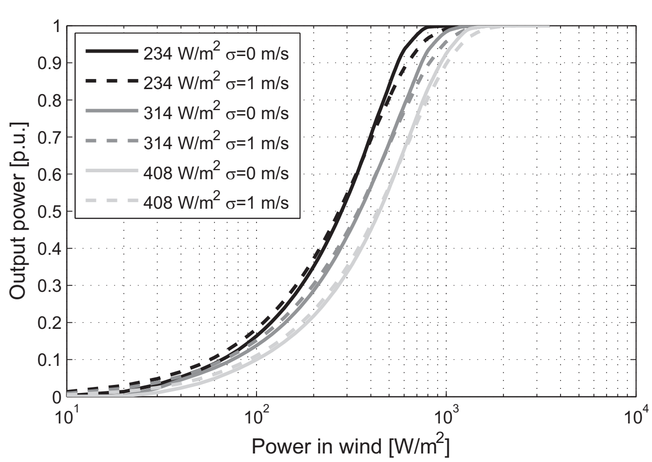


Fig. 3. Power curves for three selected WECs with different ratio of rated power to rotor area. The multi-turbine/smoothening effect accomplished by using a normal distribution with standard deviation of 1 m/s is also visualised. Cut-out region is not shown.

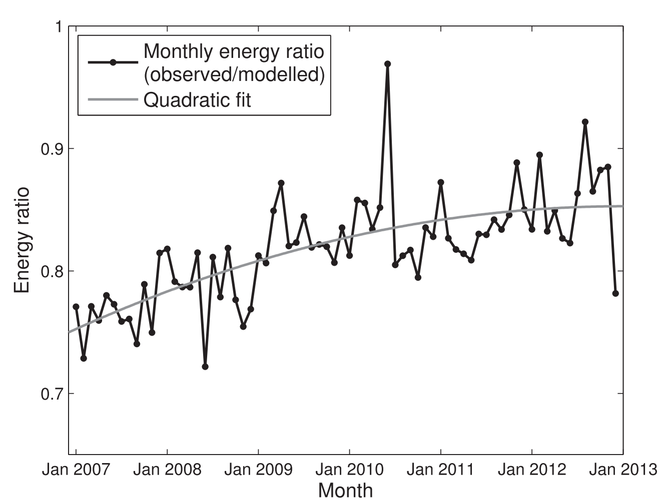


Fig. 4. Example of ratio of observed over modelled (before application of external losses) monthly energy. Seasonal bias has been removed.

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 (ΔP).
- 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:

$$\begin{aligned} \mathrm{OF} &= f_P \cdot f_{\Delta P} \cdot f_{\mathrm{statistical}} \\ &= E_P^2 \cdot \left(E_{\Delta P,1h}^2 + E_{\Delta P,4h}^2\right) \cdot \left(E_{S1} + E_{S2} + E_{S3} + E_{S4}\right) \end{aligned} \tag{5}$$

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.

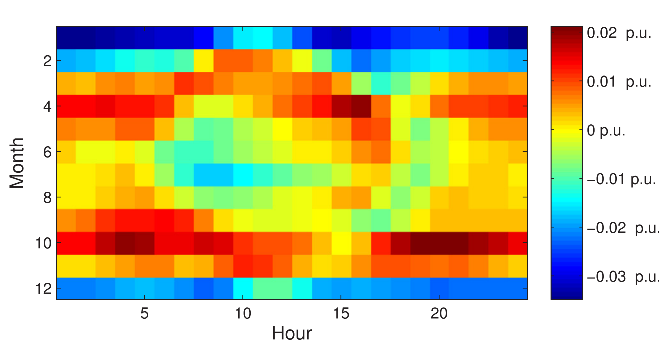


Fig. 5. Example of mean error as function of month and hour of the day (before bias correction).

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 ΔP,1 h	1.5%	3.0%	1.9%	2.8%
RMS error ΔP,4 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 improved the results. For the SE and SE2 area the OF reduction was around 17% for optimum value of σ_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 σ_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

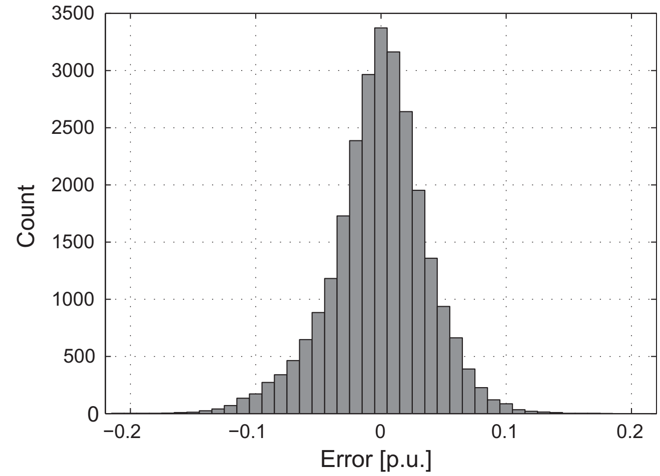


Fig. 7. Distribution of errors, i.e. the difference between model and observation.

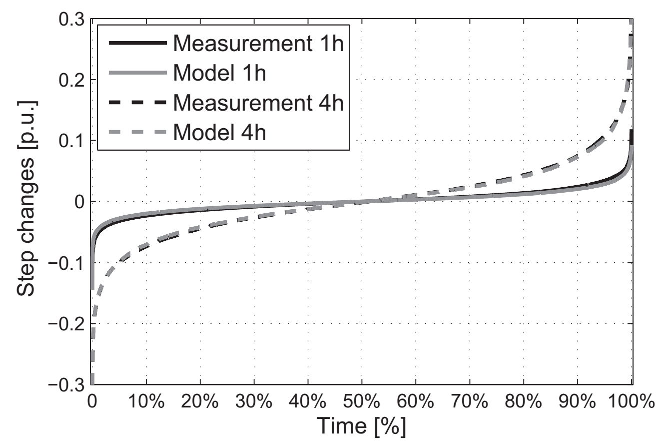


Fig. 8. Energy step change (1 and 4 h).

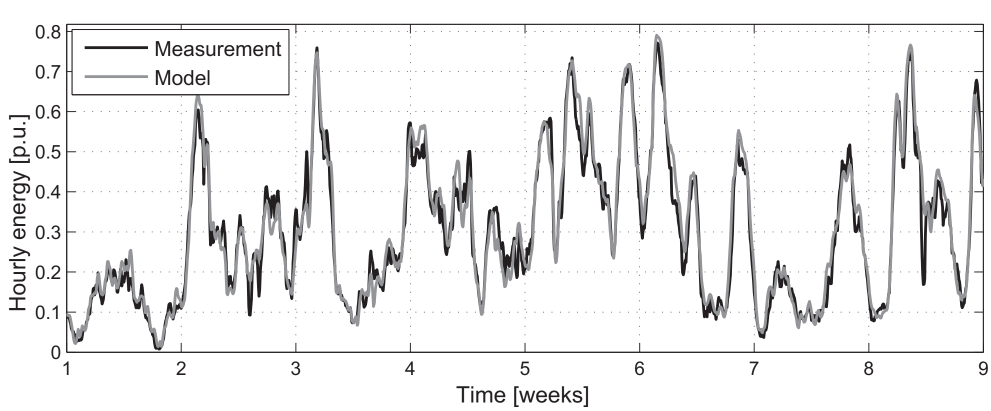


Fig. 6. Model output and validation data for eight weeks. RMS error for the period is 3.9%, i.e. almost the same as the average for all three years of validation.

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 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

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

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

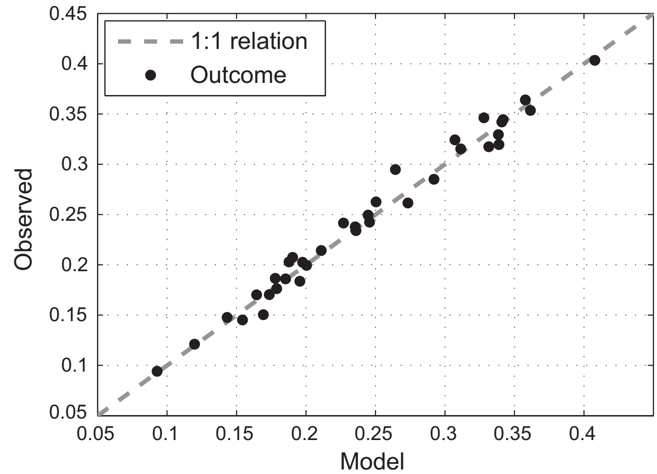


Fig. 10. Monthly capacity factor (36 months in evaluation period).

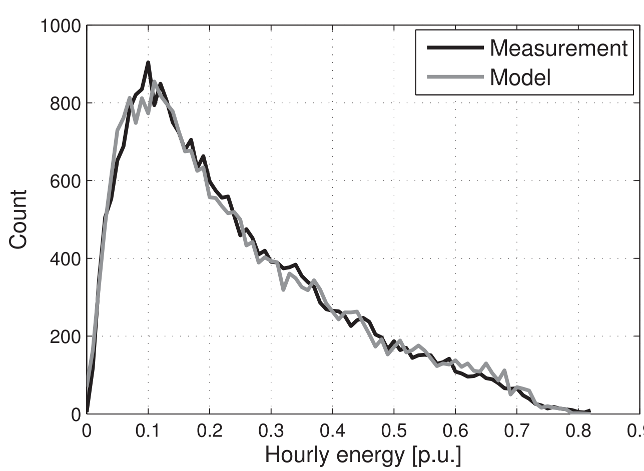


Fig. 9. Histogram calculated in bins of 0.01 p.u.

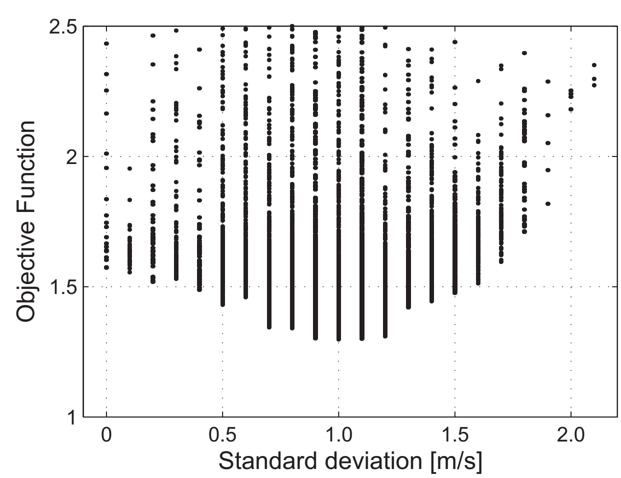


Fig. 11. Objective function dependence on the parameter σ_u. The discrete nature of σ_u was chosen to increase computational speed in Matlab.

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

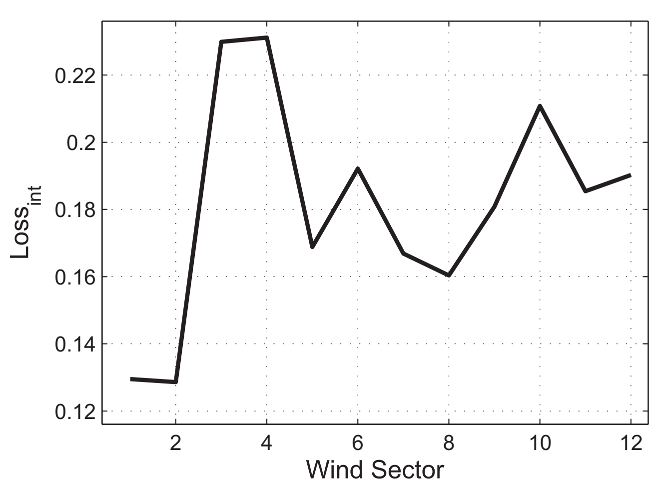


Fig. 12. Mean values of direction dependent $Loss_{int}$ for the endpoints of the 350 hill-climb searches. Sector 1 is the prevailing sector.

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.
[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, [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.

Appendix C – Instructions for downloading time series

The files can be found on:

https://www.dropbox.com/sh/gfbk52a40l5red4/AACTtj4G_x6aJWgnqk_vMqDEa?dl=0

If the link does not work, please contact Vindforsk or Uppsala University: jon.olauson (at) angstrom.uu.se or mikael.bergkvist (at) angstrom.uu.se.

Hourly averaged power time series are available for all scenarios (year 1979-2014). For scenario A1, B1 and C1, time series with 5 and 15 minutes temporal resolution are also available. Time stamps are given in separate files (Swedish time zone; UTC +1h). The time series are given for the period starting with the given time stamps, e.g. a time stamp of 04:15 in the 15 minute data implies production between 04:15 and 04:30. The file 'WECs.csv' contains coordinates and some other information on all WECs in the different scenarios.

Production is given in MW (average value for the time period). Installed capacity (MW) in the different areas are:

	SE1	**SE2**	**SE3**	**SE4**	**SE**
A1	552	2539	2581	1807	7478
A2	1077	2863	1988	1536	7464
A3	547	2362	2822	1715	7445
A4	510	2211	1968	2384	7073
A5	609	2546	2429	1555	7138
A6	1037	2953	2340	1532	7862
A7	949	1925	2433	1359	6666
A8	484	1947	1799	1342	5572
B1	1151	3108	2629	2886	9773
B2	719	4186	2173	2696	9774
B3	596	2753	3349	3090	9787
B4	651	2622	2185	3824	9282
B5	927	4160	3213	1797	10097
B6	524	2936	2761	2826	9047
B7	1132	3460	3236	2928	10757
C1	1386	3358	4428	5060	14231
C2	1841	5605	2660	4107	14214
C3	824	2891	5067	5426	14208
C4	1561	2600	2988	6276	13426
C5	1405	4689	4916	3807	14818
C6	1065	3409	3457	4543	12475
C7	1525	4954	4615	5530	16625
D1	1510	4850	5713	7174	19247

Scenarios and time series of future wind power production in Sweden

The introduction of large amounts of wind power is a challenge for the operation of the power system. In order to facilitate future studies of the effects of significantly higher penetration levels, scenarios and accompanying time series of hourly wind power production in Sweden were developed. The annual production in the scenarios range from 20 to 70 TWh and the time series spans the period 1979 to 2014. The model for hourly wind power production was validated with historical measurements and is shown to give a small error (MAE below 3%) and a good match of the statistical distributions of hourly power and step changes thereof. A statistical model for simulation of intra-hourly fluctuations (5 and 15 minutes temporal resolution) was also developed. An important conclusion is that the normalised variability will be reduced in the future, mainly because of higher capacity factors and an increased share of offshore wind power.

Another step forward in Swedish energy research

Energiforsk – Swedish Energy Research Centre is a research and knowledge based organization that brings together large parts of Swedish research and development on energy. The goal is to increase the efficiency and implementation of scientific results to meet future challenges in the energy sector. We work in a number of research areas such as hydropower, energy gases and liquid automotive fuels, fuel based combined heat and power generation, and energy management in the forest industry. Our mission also includes the generation of knowledge about resource-efficient sourcing of energy in an overall perspective, via its transformation and transmission to its end-use. Read more: www.energiforsk.se

Energiforsk