Renewable Energy 159 (2020) 991–999

Contents lists available at ScienceDirect

Renewable Energy

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

Combination of meteorological reanalysis data and stochastic simulation for modelling wind generation variability

Matti Koivisto [a, *], Guðrún Margrét Jónsdóttir [b], Poul Sørensen [a], Konstantinos Plakas [a], Nicolaos Cutululis [a]

[a] Department of Wind Energy, Technical University of Denmark, Frederiksborgvej 399, 4000, Roskilde, Denmark
[b] School of Electrical and Electronic Engineering, University College Dublin, Belfield, Dublin 4, Ireland

ARTICLE INFO

Article history:
Received 20 October 2019
Received in revised form
13 May 2020
Accepted 5 June 2020
Available online 14 June 2020

Keywords:
Ramp
Reanalysis
Simulation
Stochastic
Variability
Wind

ABSTRACT

As installed wind generation capacities increase, there is a need to model variability in wind generation in detail to analyse its impacts on power systems. Utilization of meteorological reanalysis data and stochastic simulation are possible approaches for modelling this variability. In this paper, a combination of these two approaches is used to model wind generation variability. Parameters for the model are determined based on measured wind speed data. The model is used to simulate wind generation from the level of a single offshore wind power plant to the aggregate onshore wind generation of western Denmark. The simulations are compared to two years of generation measurements on 15 min resolution. The results indicate that the model, combining reanalysis data and stochastic simulation, can successfully model wind generation variability on different geographical aggregation levels on sub-hourly resolution. It is shown that the addition of stochastic simulation to reanalysis data is required when modelling offshore wind generation and when analysing onshore wind in small geographical regions.

1. Introduction

Wind power has grown rapidly over the last decade and is expected to grow even more in the future. Wind generation is dependent on weather patterns and thus variable, which can cause challenges to the planning and operation of power systems. Consequently, understanding the behaviour of wind generation is crucial in the analysis of modern power systems.

Wind generation variability needs to be modelled on different geographical and temporal resolutions. Modelling of large geographical areas is required, e.g., in interconnection expansion planning, where wind generation in multiple countries is analysed [1]. While hourly data may be sufficient in some applications, higher temporal resolution is required, e.g., in voltage stability studies, where short-term variability can affect the reactive power support that wind turbines can provide [2]. In addition to modelling the overall probability distribution of wind generation accurately, high temporal resolution applications require simulations where also the ramp behaviour of wind generation is well represented.

Two popular approaches for modelling the variability in wind are stochastic time series simulation and meteorological reanalysis techniques. Stochastic simulations that replicate the important spatiotemporal dependencies in wind speed or generation can be carried out by fitting a time series model to available measurements and simulating data from the model [3–6]. While stochastic time series simulations can represent the important statistical characteristics of wind generation, the simulations do not relate to actual historical weather patterns, and behaviour in specific cases cannot be verified directly. In addition, while modelling locations without measurements is possible [6], assumptions on how correlations depend on geographical distances are usually required, and such dependencies may not be generally applicable.

The reanalysis approach is based on past weather but can be applied on a specific set of current or future wind installations [7–10]. The approach is straightforward in application as long as reanalysis data are available. However, the approach depends on wind speed distributions and spatiotemporal dependencies being modelled accurately in the meteorological data set used as input. The reanalysis approach can capture important spatiotemporal dependencies in large-scale wind generation; however, wind variations are smoothened because of spatial and temporal averaging

* Corresponding author.
E-mail address: mkoi@dtu.dk (M. Koivisto).

https://doi.org/10.1016/j.renene.2020.06.033
0960-1481/

effects in mesoscale models [11,12]. In addition, the temporal resolution of the reanalysis data, e.g., hourly, may not be high enough for power system applications.

Models for building wind speed time series sampled with a high frequency that capture the short- and long-term wind speed variability have been proposed in the literature and some are compared in Ref. [13]. In Ref. [14], measured time series averaged every 10–15 min are coupled with realistic simulated turbulence modelled using a turbulence spectrum. Sub-hourly variability of wind speed is modelled using time series models and combined with measured data in Ref. [15]. Measured and simulated data are combined in Ref. [13]; however, the potential of using reanalysis data instead of measured data is highlighted as a future option.

This paper proposes a combination of the reanalysis and stochastic simulation approaches. In this novel approach for large-scale simulations, the reanalysis data models the majority of the spatiotemporal dependencies, while stochastic simulation models the short-term variability not modelled by the reanalysis approach. Stochastic wind simulation modelling [16–18], which is usually applied on turbine or wind power plant (WPP) level, is combined to the large-scale reanalysis data. Through the stochastic simulation, realistic sub-hourly ramp behaviour is captured in the simulations. In this way, the stochastic simulation model provides variability in the frequency range where the reanalysis data is lacking, even when downscaling is used.

The presented methodology allows large-scale simulation of many WPPs. Both the spectral information of individual locations and the dependencies between the locations using coherence functions are considered. The individually simulated WPPs can then be aggregated to the desired aggregation-level, e.g., regional or country-wise. However, the individual simulation of WPPs allows the simulation data to be used also, e.g., in power flow studies where each node with wind power installations needs data that are properly correlated with all the other nodes. This simulation of each WPP differentiates the methodology, e.g., from Ref. [15], where the spectral information is first taken to the aggregate level and time series simulation is then carried out directly for the aggregate.

The Correlations in Renewable Energy Sources (CorRES) tool is used for carrying out the simulations combining reanalysis data and stochastic simulation (called fluctuations) [19]. The tool is based on hourly mesoscale reanalysis data from the Weather Research and Forecasting (WRF) model [20]. Wind speed measurements from three locations are used to calibrate the spectral parameters for the stochastic simulation part of CorRES.

The importance of adding fluctuations to the mesoscale time series is tested for geographical areas with different sizes, going from a single offshore wind power plant (OWPP) to the aggregate onshore wind generation of western Denmark. Measured generation data covering two years with 15 min resolution are used for validating the CorRES simulations. Simulations with and without fluctuations are compared to the measured data to study the significance of stochastic simulation in the different cases.

The paper is structured as follows. Section 2 introduces the CorRES tool and Section 3 describes the selection of the spectral parameters for its stochastic simulation part. Section 4 presents simulation results for offshore and onshore wind generation in Denmark and compares the role of stochastic simulation in the different cases. Section 5 provides additional discussion and Section 6 concludes the paper.

2. CorRES simulation tool

This section presents the CorRES simulation tool with its two parts: the meteorological reanalysis data and stochastic simulation of fluctuations. The section then shows how the two parts are combined to provide the simulated wind speed output. CorRES can be used to simulate the wind speed and/or wind power output at specific locations and aggregated for a specific system level (e.g., regional or country-wise) [19]. Transformation from wind speeds to power generation is presented in Section 4.

2.1. Meteorological reanalysis data

CorRES is based on meteorological data obtained from the WRF model, a mesoscale modelling system [20], using the downscaling method presented in Ref. [21,22]. The coverage of the 10 km × 10 km WRF grid for the analysed area can be seen in Fig. 1. More information about the reanalysis data and WRF modelling in CorRES is given in Refs. [10,19]. The WRF data are hourly, and linear interpolation is used in-between the hourly WRF values.

Wind speed variations in the WRF data are smoothed due to spatial and temporal averaging effects [11,12]. Additionally, as the WRF data are hourly, inter-hour variations are not modelled. Therefore, the short-term wind speed fluctuations are captured through the stochastic model added to the WRF data as shown in Sections 2.2 and 2.3. If fluctuations are not included, sub-hourly values are obtained directly from the liner interpolation.

2.2. Stochastic model

The fluctuations are added to the WRF data to obtain realistic sub-hourly simulations when using hourly reanalysis data. They also model the short-term variability that may be lacking in WRF data on hourly and even lower frequencies. The temporal behaviour of fluctuations at each WPP is modelled by specifying a fluctuation power spectral density (PSD) for each location. As has been shown in Ref. [12], the wind speed PSDs at frequencies that lack variability in the WRF data can be expected to follow

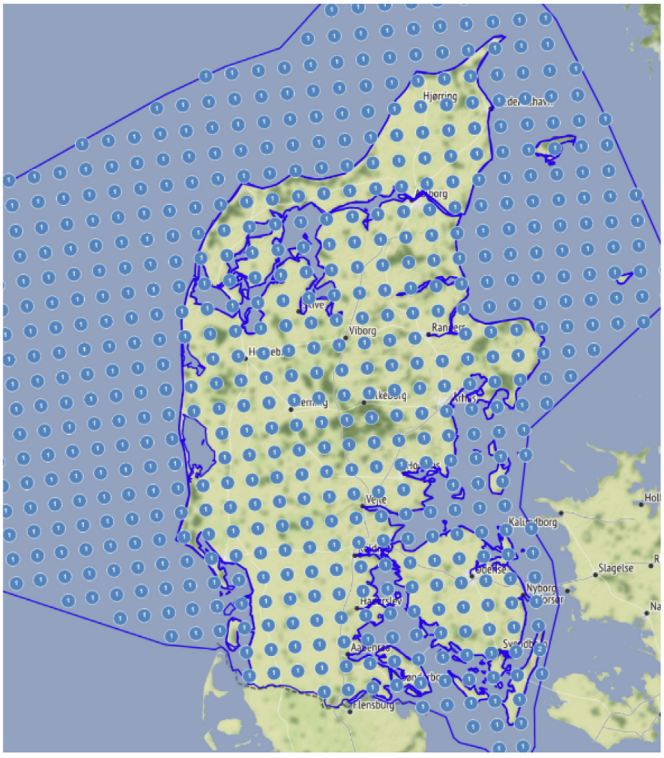

Fig. 1. The meteorological data grid points for the analysed area.

$$S_{\text{meso}}(f) = \frac{a_1}{f^{5/3}}, \tag{1}$$

where f is frequency and a_1 is the coefficient of the spectra. However, as the fluctuations are added to mesoscale reanalysis data, they should not add variability on the frequencies where the WRF data successfully capture the wind variability. The parameter f_0 is introduced to specify the frequency range where the stochastic simulation adds variability. Thus, the PSD for the stochastic simulation, providing the additional variability to the mesoscale reanalysis data, is defined as

$$S_{\Delta\text{meso}}(f) = \frac{a_1}{{f_0}^{5/3} + f^{5/3}}. \tag{2}$$

In Fig. 2, the PSD of the WRF data with and without fluctuations is presented. It shows that for $f < f_0$, WRF provides practically all variability, and for $f > f_0$ the fluctuations, modelled as in (2), increase the PSD. In addition to (2), the fluctuation model in CorRES considers turbulence, as shown in Ref. [18]; however, the addition of turbulence PSD has a very small influence on the 10 and 15 min resolution data studied in this paper and is thus not covered further.

As multiple WPPs are analysed, the dependencies between locations are included in the fluctuation modelling. This is achieved by specifying coherence functions between the WPPs. Coherence function between two locations i and j is defined as shown in Ref. [16]:

$$\gamma_{i,j}(f) = \mathrm{e}^{-\left(A_{i,j}\frac{d_{i,j}}{V_0}\right)f}, \tag{3}$$

where f is frequency, $d_{i,j}$ is the distance between the locations, V_0 is average wind speed and $A_{i,j}$ is the decay factor. V_0 is calculated from the WRF data for the different time steps. V_0 is defined as 10 min mean wind speed [18]; in this paper, as hourly WRF data are used, V_0 is calculated using linear interpolation from the hourly WRF data to get a representative approximate 10 min mean wind speed for each time step (this is discussed more in Section 5). The decay factor is specified as

$$A_{i,j} = \sqrt{\left(A_{\text{long}}\cos(\alpha_{i,j})\right)^2 + \left(A_{\text{lat}}\sin(\alpha_{i,j})\right)^2}, \tag{4}$$

where $\alpha_{i,j}$ is the inflow angle (as specified in Ref. [16]), A_{long} is the decay factor when the flow is longitudinal (i.e., wind direction is from location i to j) and A_{lat} is the decay factor when the flow is lateral [18]. Decay factor parameters $A_{\text{long}} = 4$ and $A_{\text{lat}} = \frac{V_0}{2\text{m/s}}$ are used, as suggested in Ref. [18].

Time series simulation from the specified PSDs and coherence functions is carried out as shown in Ref. [18]. When calculating the cross power spectral density functions using the PSDs and (3), the delay time for wind field to travel from location i to j is also considered, as shown in Ref. [16]. Fluctuations are simulated only for frequencies higher than f_0.

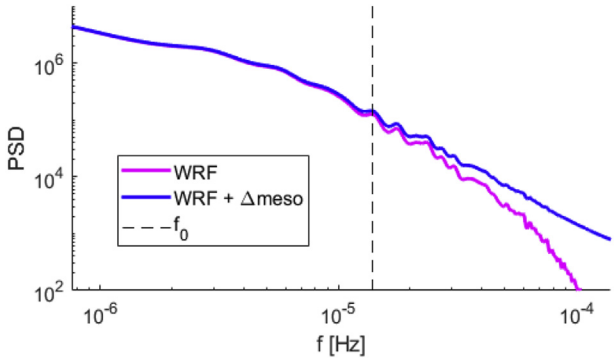

Fig. 2. PSD of WRF data with and without the fluctuation PSD, as defined in (2); $a_1 = 3{\cdot}10^{-4}$ and $f_0 = \frac{1}{20}$ h were chosen for this illustrative example. The WRF PSD was estimated using Matlab function pcov [23].

2.3. Combining the reanalysis data and the fluctuations

The simulated wind speeds for n locations $\boldsymbol{v}_t = [v_{1,t}, \ldots, v_{n,t}]$ are

$$\boldsymbol{v}_t = \boldsymbol{v}_t^{\text{WRF}} + \boldsymbol{v}_t^{\text{fluctS}}, \tag{5}$$

where $\boldsymbol{v}_t^{\text{WRF}} = [v_{1,t}^{\text{WRF}}, \ldots, v_{n,t}^{\text{WRF}}]$ are the WRF mesoscale wind speeds and $\boldsymbol{v}_t^{\text{fluctS}} = [v_{1,t}^{\text{fluctS}}, \ldots, v_{n,t}^{\text{fluctS}}]$ are the simulated fluctuations. The components of $\boldsymbol{v}_t^{\text{WRF}}$ are correlated in space and time, as are the components of $\boldsymbol{v}_t^{\text{fluctS}}$. This allows (5) to adjust the spatiotemporal correlations of $\boldsymbol{v}_t$ compared to $\boldsymbol{v}_t^{\text{WRF}}$ by specifying the correlation structure of $\boldsymbol{v}_t^{\text{fluctS}}$.

Fig. 3 shows time series of WRF data and simulations with fluctuations using two different random seeds for an example location. It can be seen that both simulations with fluctuations follow the WRF data, but with additional short-term (or high frequency) variability. The difference in random seed results in a different fluctuation pattern around the WRF data; however, the two fluctuation time series are simulated using the same spectral parameters, so when long time series are simulated, they add the same level of variability to the WRF data.

3. Selecting the spectral parameters

This section shows the selection of the spectral parameters for (2) based on wind speed measurements from multiple locations. When the spectral parameters are selected, the wind speed modelling part of CorRES is defined; the transformation to power generation is presented in Section 4.

3.1. Measured wind speed data

Measured wind speed data from three locations are analysed.

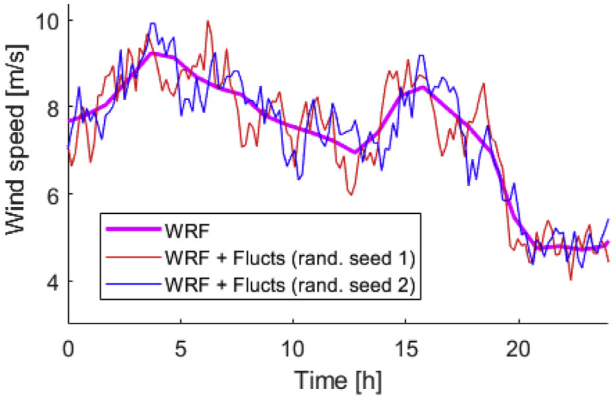

Fig. 3. An example time series of WRF data and simulations with fluctuations using two different random seeds (but with the same spectral parameters). The time series have a 10 min resolution, with hourly WRF data linearly interpolated to reach the intra-hour values.

The locations and the available data are shown in Table 1. The Hovsore and Risø data sets include measurements collected for the DTU Online Meteorological Data base [24]. The Cabauw data set includes measurements gathered at the Cesar Observatory located in the western part of the Netherlands; the data are publicly available in Ref. [25].

3.2. Impact of the spectral parameters

Wind speed fluctuations, modelled using (2), are dependent on the two parameters, namely f_0 and a_1. The parameters define the intensity (a_1) and the frequency range (f_0) of the generated fluctuations. Example time series of WRF data and simulations with fluctuations are shown in Fig. 4. The two spectral parameter sets show different fluctuation behaviour around the WRF data, with parameter set B fluctuating more. This highlights how the modelled wind speed fluctuations vary depending on the selected parameters.

3.3. Finding the optimal parameters

The selection of the optimal fluctuation parameters for each analysed wind speed location is carried out by comparing the autocorrelation function (ACF) of the measured and simulated data while varying both f_0 and a_1. It has been shown that WRF data may lack variability in frequency scales up to multiple hours [12]; in this paper, f_0 values from $\frac{1}{4h}$ to $\frac{1}{14h}$ were tested.

In [26], $a_1 = 3\times10^{-4}$ has been suggested for an offshore location in Denmark for modelling (1) for the frequency range of interest for fluctuations. However, when fluctuations resulting from (2) were applied in (5), the resulting $\boldsymbol{v}_t$ showed too high variability for the test locations when using 3×10^{-4}. Thus, a_1 values 1.5×10^{-4}, 2×10^{-4}, and 2.5×10^{-4} were also tested.

The comparison of the different f_0 and a_1 values was carried out by varying f_0 from $\frac{1}{4h}$ to $\frac{1}{14h}$ with a step of $\frac{1}{1h}$ and considering all the a_1 values described in the previous paragraph. A simulation was carried out for each location for each combination of these f_0 and a_1 values and the ACF of the resulting simulated time series was compared to the ACF of the measured data. The ACFs were compared by calculating

$$RMSE_{\mathrm{ACF}} = \sqrt{\frac{\sum_{k=0}^{K}\left(\rho_{\mathrm{meas}}(k) - \rho_{\mathrm{sim}}(k)\right)^2}{K}}, \tag{6}$$

where $\rho_{\mathrm{meas}}(k)$ is the ACF of the measured data and $\rho_{\mathrm{sim}}(k)$ is the ACF of the simulated data for the studied location at lag k. ACFs are compared up to lag K. $K = 10$ h was considered in the comparison, as it is the time frame in which fluctuations are expected to have a significant impact.

ACF was chosen as the comparison metric as there is a strong link between the ACF of a time series and its ramp behaviour. Ramp behaviour can be analysed at different lags k as

Table 1
The wind speed data.

Name	Location	Height [m]	Time interval
Hovsore (Denmark)	56.44° N, 8.15° E	80	4 years (2005–2008)
Risø (Denmark)	55.69° N, 12.09° E	77	5 years (2000–2005)
Cabauw (Netherlands)	51.97° N, 4.93° E	80	6 years (2001–2006)

All wind speed data have 10 min resolution.

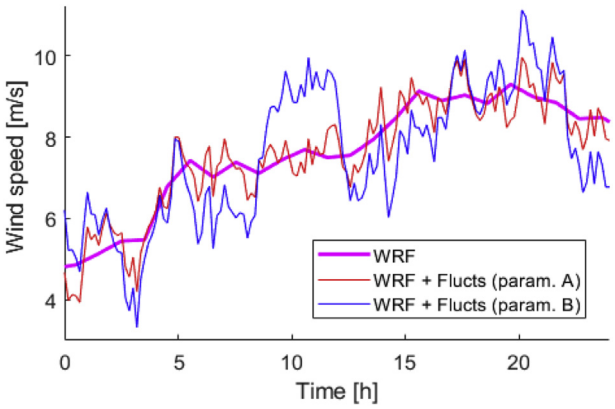

Fig. 4. An example time series of the WRF data and simulations with fluctuations. The different fluctuation parameter sets are: A, where $a_1 = 2\times10^{-4}$ and $f_0 = \frac{1}{5h}$; and B, where $a_1 = 3\times10^{-4}$ and $f_0 = \frac{1}{13h}$. The time series have a 10 min resolution, with hourly WRF data linearly interpolated to reach the intra-hour values.

$$v'_{i,t}(k) = v_{i,t} - v_{i,t-k}, \tag{7}$$

where $v_{i,t}$ is wind speed of location i at time t and k is the lag of the difference. The standard deviation (SD) of the difference at lag k, when assuming that $\mathrm{SD}(v_{i,t}) = \mathrm{SD}(v_{i,t-k}) = \sigma_v$, is

$$\mathrm{SD}\left(v'_{i,t}(k)\right) = \sqrt{\sigma_v^2 + \sigma_v^2 - 2\rho_k\sigma_v\sigma_v} = \sqrt{2}\sigma_v\sqrt{1-\rho_k}, \tag{8}$$

where $\rho_k = cor(v_{i,t}, v_{i,t-k})$ is ACF(k) of $v_{i,t}$. As fluctuations have an impact only on high frequencies, the overall wind speed SD of the location, σ_v, is defined mostly by the mesoscale reanalysis data. However, the fluctuations change significantly the ACF, as can be seen in Fig. 5. As ACF is strongly linked to ramp behaviour (8), and as it was considered important to model the ramp behaviour accurately, minimizing (6) was chosen as the metric when choosing the optimal fluctuation parameter set. Fig. 5 also shows that the addition of fluctuations to the mesoscale WRF data is required, as

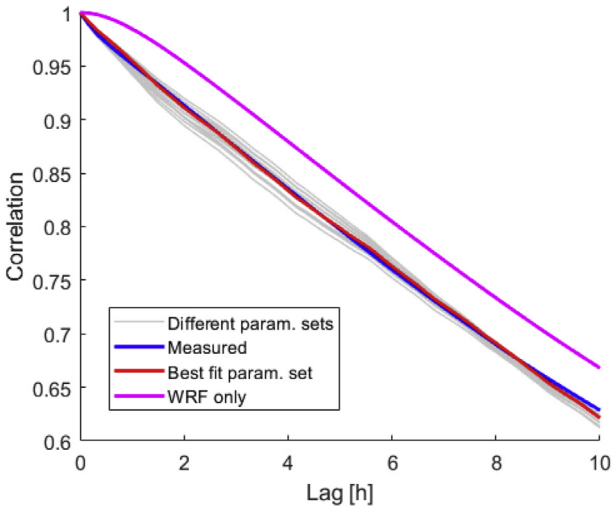

Fig. 5. ACFs of simulated wind speeds using different fluctuation parameter combinations for the Hovsore dataset are shown in grey. ACF of the measured data is plotted in blue and ACF of the best fit parameter set in red. ACF of the WRF data, with linear interpolation to reach intra-hour values, is plotted in magenta. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

only using WRF data shows too high ACF values up to lags of multiple hours.

3.4. The resulting spectral parameters

Table 2 shows the spectral parameters which provide the closest fit to the measured wind speed data based on (6) for the three analysed wind speed locations. It can be seen that for all locations, the best fitting a_1 parameter is 2×10^{-4}, with f_0 on average $\frac{1}{10h}$. As quite similar fluctuation parameters show the best fit for all locations, the parameter set $a_1 = 2\times10^{-4}$ and $f_0 = \frac{1}{10h}$ is used for all consequent runs presented in the paper. Further comparison of the wind speed fluctuation parameters for individual locations is given in Ref. [27]. The selection of the parameters is discussed in Section 5.

4. Wind generation simulation results

In this section, wind generation for different cases in Denmark are simulated utilizing the CorRES tool presented in Section 2 and the results are compared to measured data. The different cases include the generation output of individual OWPPs and the aggregation of onshore wind generation for different geographical areas. With comparison to the measured data, the importance of fluctuations is assessed in the different studied cases.

4.1. Data and CorRES simulation setup

The measured data considered in this section consist of wind generation measurements gathered in Denmark. The data were collected with a 15 min resolution (15-min averages) over two years, 2010 and 2011. Offshore wind generation is measured on the OWPP-level, while onshore wind generation is measured on regional level; the regions are shown on a map in Fig. 6. Due to data restrictions, details of the two OWPPs cannot be provided. For the remainder of this paper they will be referred to as OWPP 1 and OWPP 2.

Onshore wind power plant (WPP) installation data were taken from Ref. [28]; installations until 2011 were used to model the years 2010 and 2011. Wind turbine types of the WPPs were linked to power curve data from Ref. [29]. For OWPPs, wake effects were modelled using the PyWake software [30]. All generation data are standardized to values between 0 and 1, where 1 means generation at installed capacity, by dividing the generation time series data by installed capacity time series.

The remainder of this section is divided into three parts. In each part, wind generation is simulated using the CorRES simulation tool presented in Section 2, utilizing the stochastic model parameters found in Section 3. The best fitting fluctuation parameter set found in Section 3, i.e., $a_1 = 2\times10^{-4}$ and $f_0 = \frac{1}{10h}$, is used in all runs. The simulations are compared to measured generation data; for individual OWPPs in Section 4.2, for different onshore wind regions in Denmark in Section 4.3 and for the aggregate onshore wind generation of western Denmark in Section 4.4. In all studied cases, each WPP is simulated and the resulting time series are then aggregated to the appropriate level of aggregation. As each WPP, rather than each turbine, is simulated, the wind farm average block described in Ref. [18] is used to capture the smoothening effect of the wind farm. To assess the importance of fluctuations, simulations with and without fluctuations are compared for each studied case. If fluctuations are not added, linear interpolation of the WRF data is used directly to reach the sub-hourly resolution wind speeds (for each WPP individually).

Table 2
The best fitting fluctuation parameter sets for each analysed wind speed location.

Name	a_1	f_0[1/h]
Hovsore	2×10^{-4}	1/10
Risø	2×10^{-4}	1/11
Cabauw	2×10^{-4}	1/9

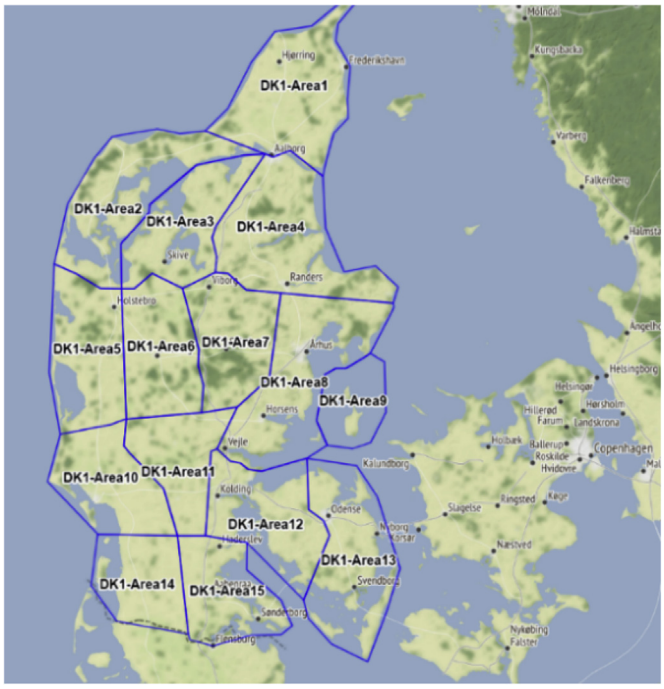

Fig. 6. The analysed onshore wind regions in western Denmark.

4.2. Offshore wind generation

PSDs for the two analysed OWPPs are shown in Fig. 7, for

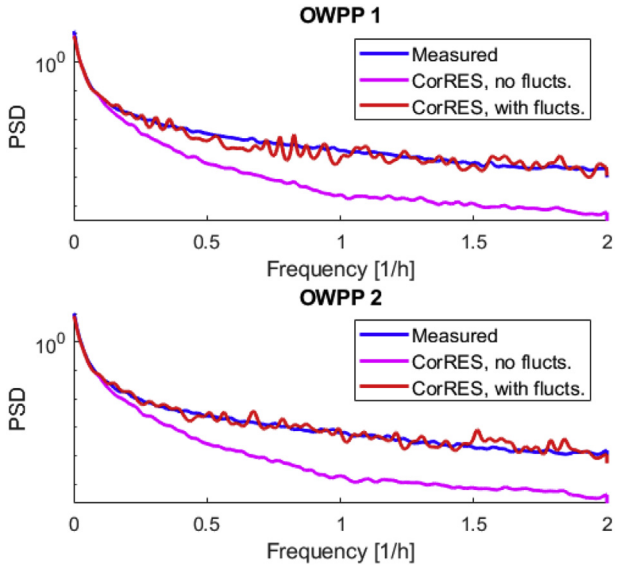

Fig. 7. Generation PSDs of OWPP 1 and OWPP 2 for the measured data and CorRES simulations with and without fluctuations. The PSDs were estimated using Matlab function pcov [23].

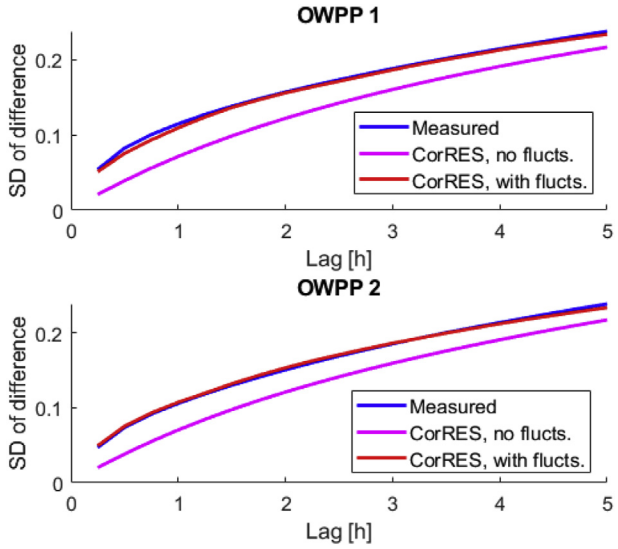

Fig. 8. Generation ramp SDs on different lags for the two OWPPs for the measured data and CorRES simulations with and without fluctuations.

measured data and CorRES simulations with and without fluctuations. The simulation with fluctuations shows PSDs similar to the measured data for both locations, whereas without fluctuations the PSDs are low for high frequencies compared to the measurements. It is notable that variability is lacking for up to a few hours. Fig. 8 shows generation ramp behaviour for the same OWPPs for measurements and simulations. The ramp SDs are visualized by plotting

$$\mathrm{SD}\left(p'_{i,t}(k)\right) = \mathrm{SD}\left(p_{i,t} - p_{i,t-k}\right), \tag{9}$$

where $p_{i,t}$ is generation from location i at time t, on different lags k. When fluctuations are not added, the ramp SDs are significantly lower compared to the measured data, for lags up to a few hours. The ramp behaviour of the simulation with fluctuations is similar to the measurements for both OWPPs.

As can be seen in Table 3 and Table 4, without fluctuations the 15 min ramp SDs are significantly lower in the simulation compared to the measurements. When fluctuations are added, the ramp SDs are close to the measured data for both OWPPs. All analysed 15 min ramp percentiles are too close to zero in "CorRES, no flucts" compared to the measurements. When fluctuations are included in "CorRES with flucts", the ramp percentiles are on average closer to the measured data; however, especially for OWPP 2, the 5th and 95th percentiles are somewhat too far from zero compared to the measurements. This is discussed further in Section 5.

4.3. Regional onshore wind generation

Fig. 9 shows PSDs for the measured data and the simulations for two example onshore wind regions (the regions can be seen on map in Fig. 6). The simulation without fluctuations has less high frequency components than the measurements. Addition of the fluctuations provides simulated time series with PSDs similar to the measured data. Fig. 10 shows the generation ramp behaviour of the same regions. It can be seen that when fluctuations are not considered, the ramp SDs are lower than in the measured data for lags up to a few hours. However, the differences are smaller than for the individual OWPPs analysed in the previous section. The ramp SDs of the simulation with fluctuations are closer to the measured

Table 3
Descriptive statistics of OWPP 1 generation 15 min ramp behaviour.

15 min ramp statistic	Measured	CorRES, no flucts	CorRES with flucts
SD	σ_{ramp}	$0.39\sigma_{\mathrm{ramp}}$	$0.95\sigma_{\mathrm{ramp}}$
5th, 95th percentile	p_5, p_{95}	$0.38p_5$, $0.39p_{95}$	$1.19p_5$, $1.17p_{95}$
1st, 99th percentile	p_1, p_{99}	$0.36p_1$, $0.37p_{99}$	$0.87p_1$, $0.86p_{99}$

The values are shown in relation to measured data; exact values for individual OWPPs cannot be shared due to data restrictions.

Table 4
Descriptive statistics of OWPP 2 generation 15 min ramp behaviour.

15 min ramp statistic	Measured	CorRES, no flucts	CorRES with flucts
SD	σ_{ramp}	$0.43\sigma_{\mathrm{ramp}}$	$1.05\sigma_{\mathrm{ramp}}$
5th, 95th percentile	p_5, p_{95}	$0.43p_5$, $0.43p_{95}$	$1.31p_5$, $1.28p_{95}$
1st, 99th percentile	p_1, p_{99}	$0.38p_1$, $0.42p_{99}$	$1.00p_1$, $0.99p_{99}$

The values are shown in relation to measured data; exact values for individual OWPPs cannot be shared due to data restrictions.

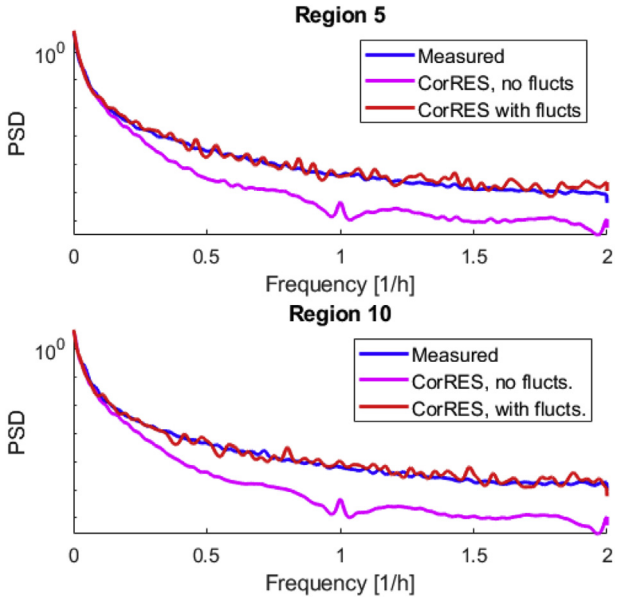

Fig. 9. Generation PSDs for two example onshore wind regions for the measured and simulated data. The PSDs were estimated using Matlab function pcov [23].

data. For region 10 shown in Fig. 10, the ramp SDs of the simulation with fluctuations deviates somewhat from measured data for lags higher than 2 h; this is discussed in Section 5.

Table 5 and Table 6 show that both the 15 min ramp SDs and the analysed percentiles of "CorRES with flucts" are closer to the measured data compared to "CorRES, no flucts". The 5th and 95th percentiles are slightly too far from zero in the simulation with fluctuations for both regions; however, the 1st and 99th percentiles are very similar to the measurements.

Table 7 shows that addition of fluctuations provides 15 min ramp SDs closer to the measured data for all the studied onshore wind regions. Although for some regions, such as 13 and 15, the 15 min ramp SD is somewhat higher in "CorRES with flucts" than in the measured data, the ramp SDs are still closer to the measurements than when fluctuations are not added.

4.4. Aggregate onshore wind generation

This section considers the aggregation of all the 15 onshore

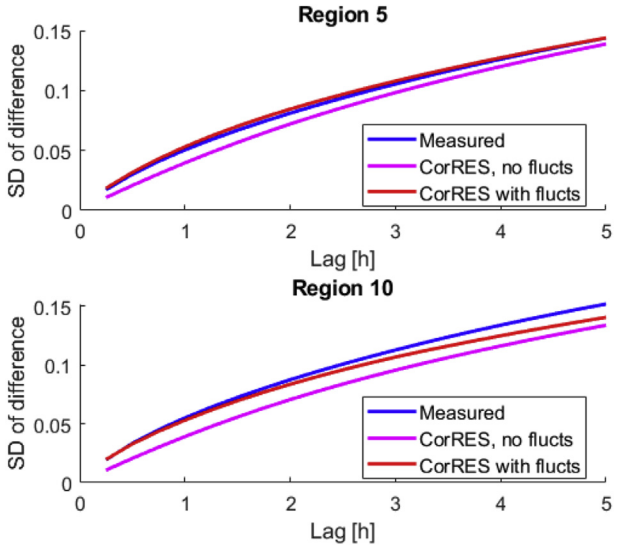

Fig. 10. Ramp SDs for different lags for the two example onshore wind regions for the measured and simulated data.

Table 5
Descriptive statistics of onshore wind region 5 generation 15 min ramp behaviour.

15 min ramp statistic	Measured	CorRES, no flucts	CorRES with flucts
SD	0.017	0.011	0.018
5th, 95th percentile	−0.026, 0.026	−0.015, 0.017	−0.030, 0.031
1st, 99th percentile	−0.048, 0.049	−0.030, 0.032	−0.050, 0.050

Table 6
Descriptive statistics of onshore wind region 10 generation 15 min ramp behaviour.

15 min ramp statistic	Measured	CorRES, no flucts	CorRES with flucts
SD	0.019	0.011	0.020
5th, 95th percentile	−0.029, 0.029	−0.015, 0.016	−0.033, 0.032
1st, 99th percentile	−0.056, 0.057	−0.030, 0.032	−0.058, 0.055

Table 7
Generation 15 min ramp SDs for the onshore wind regions.

Region	Size (km^2)	Measured	CorRES, no flucts	CorRES with flucts
1	3385	0.017	0.010	0.018
2	3364	0.016	0.010	0.016
3	2544	0.018	0.011	0.020
4	3850	0.019	0.010	0.018
5	2782	0.017	0.011	0.018
6	2517	0.016	0.010	0.017
7	2460	0.018	0.010	0.018
8	4014	0.014	0.009	0.016
9	1329	0.045	0.014	0.045
10	2556	0.019	0.011	0.020
11	2114	0.016	0.010	0.016
12	2903	0.016	0.011	0.017
13	3605	0.015	0.010	0.019
14	2135	0.020	0.011	0.019
15	2367	0.018	0.012	0.021

The area sizes are calculated using the borders shown in Fig. 6. The sum of the areas is 41924 km^2 (a more detailed area specification for western Denmark, considering only land area, gives a total area of 33275 km^2).

wind regions shown in Fig. 6, which means aggregate onshore wind generation of western Denmark. Fig. 11 shows that when modelling such aggregated level, the PSDs show some difference between the measured data and the simulation without fluctuations. However,

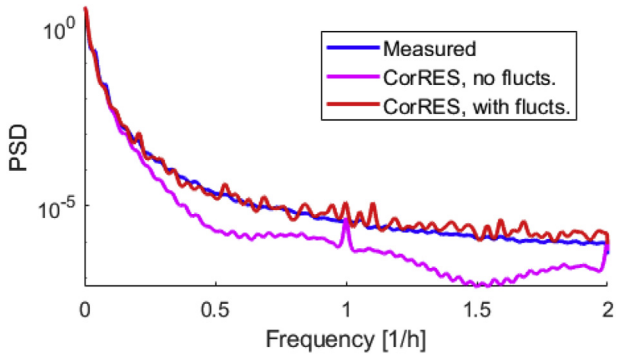

Fig. 11. Generation PSD for the aggregate onshore wind generation of western Denmark for the measured and simulated data. The PSDs were estimated using Matlab function pcov [23].

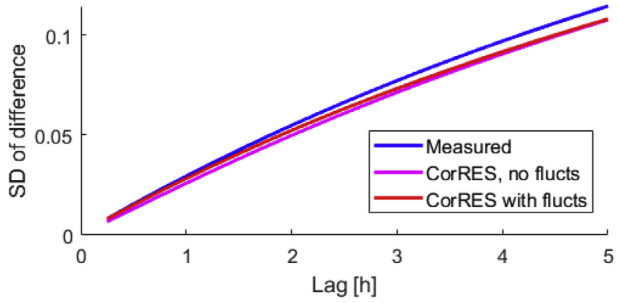

Fig. 12. Ramp rate SDs for different lags for the aggregate onshore wind generation of western Denmark for the measured data and the simulations.

the PSD at 1 1/h, the original resolution of the WRF data, is similar in "CorRES, no flucts" and the measured data. The ramp behaviours of "CorRES, no flucts" and "CorRES with flucts" in Fig. 12 are quite similar; thus, the fluctuations do not have a significant role in this simulation case.

A comparison of measured data and the simulations with and without fluctuations is shown in Table 8. The measured data and the two simulations have similar 15 min ramp SDs. For the studied percentiles, the simulation with fluctuations is closer to the measured data; however, even without fluctuations, the simulated data shows 15 min ramp behaviour similar to the measurements.

It can be concluded that on the aggregate level of western Denmark, the fluctuations add only little to the variability in the WRF data. The WRF data can model most of the ramping behaviour on this level of geographical aggregation when analysing 15 min resolution data by applying simply linear interpolation to reach the sub-hourly resolution. However, in both Fig. 12 and Table 8, the simulation with fluctuations still shows results closer to the measured data.

5. Discussion

The differences between the measured data and CorRES with fluctuations for region 10 in Fig. 10 could be seen as a reason to consider changing the fluctuation parameter set (a_1 and f_0). However, the high frequency variability up to around 2 h is well modelled by the selected parameter set. As fluctuations should impact mainly high frequency variability, focus was given to differences in ramp SDs up to around 2 or 3 h. Even though fluctuations impact also higher lags (Fig. 5), variability for lags higher than 3 h depend also on the technical WPP parameters (such as hub heights) and on the WRF model specification (e.g., roughness

Table 8
Descriptive statistics of the aggregate onshore wind generation 15 min ramp behaviour.

15 min ramp statistic	Measured	CorRES, no flucts	CorRES with flucts
SD	0.008	0.007	0.008
5th, 95th percentile	−0.013, 0.013	−0.010, 0.011	−0.013, 0.013
1st, 99th percentile	−0.024, 0.023	−0.018, 0.020	−0.022, 0.023

The land area of the analysed aggregate onshore area (western Denmark) is approximately 33275 km^2.

parametrization). Thus, rather than reconsidering the fluctuation parameter set due to one region, it was considered that the WRF model specification and the technical WPP parameters should be checked first; this will be done in future research.

While the 15 min ramp SDs are modelled quite well in the simulation with fluctuations in the analysed cases, the studied percentiles are somewhat off compared to measurements, especially for the OWPPs as shown in Tables 3 and 4. Having the SDs well modelled but the percentiles incorrect suggests that the ramp distribution shape could be improved. This is further visualized in Fig. 13, where "CorRES with flucts" shows PDF and cumulative distribution function (CDF) closer to measured data compared to "CorRES, no flucts"; however, the shape of the distribution is somewhat different. Both analysed OWPPs show similar behaviour. The generation ramp rate distribution in Fig. 13 is non-Gaussian even though the simulated wind speed fluctuations have a Gaussian shape; this is mainly because wind speed to generation transformation (power curve) is highly non-linear. However, as the resulting simulated ramp distribution shows a different shape compared to the measured data in Fig. 13, simulation of non-Gaussian wind speed fluctuations is planned for future research.

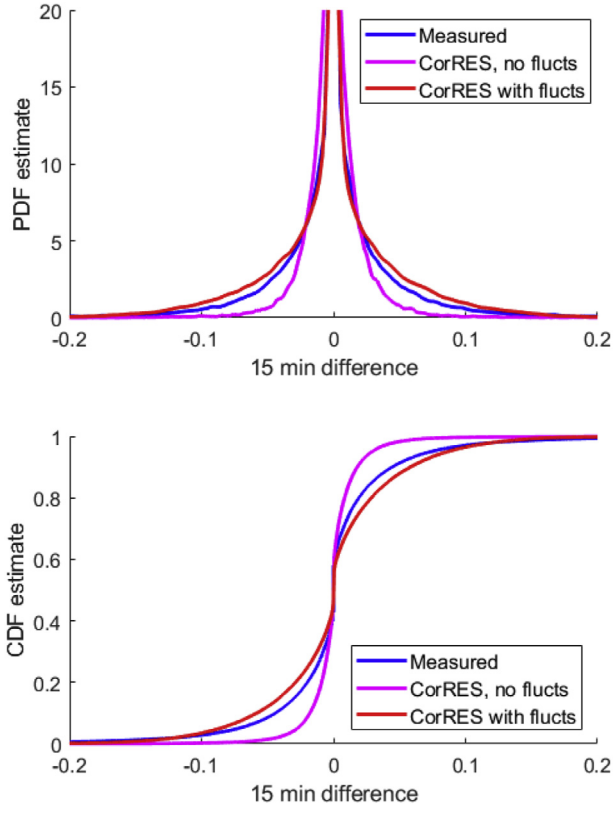

Fig. 13. 15 min ramp rate PDFs and cumulative distribution functions (CDFs) for one of the analysed OWPPs for measured data and the simulations.

When selecting the spectral parameters based on (6), multiple combinations of a_1 and f_0 give relatively similar $RMSE_{ACF}$ values. E.g., $f_0 = \frac{1}{10h}$ and $f_0 = \frac{1}{11h}$ show very similar $RMSE_{ACF}$, and even changing a_1 slightly does not have a drastic impact on $RMSE_{ACF}$. The selection of the two parameters is also correlated: if a_1 is increased, a shorter frequency range (f_0) can be selected to reach a relatively similar $RMSE_{ACF}$ [27]. Finding parameters that are in a strict sense optimal is thus challenging. Added difficulty is that different heights and meteorological years show somewhat different optimal parameter sets [27]. However, the spectral parameters used in Section 4 in the studied cases show a good fit to measurements. Although technically possible, the fluctuation parameters f_0 and a_1 were not set for each location separately when modelling wind generation in Section 4. Rather, generally applicable parameters were used. However, other significantly different geographical regions may require different fluctuation spectra parameters, and selection of location-specific parameters can be considered in future work. In addition, possible time-dependency of the spectral parameters can be considered; currently, the parameters f_0 and a_1 are not assumed to depend, e.g., on hour of the day or month. Parameter-dependency on height should also be considered.

It is important to note that the selection of the fluctuation spectra parameters depend on the mesoscale data set used. If the mesoscale data include more high frequency components, e.g., due to using a newer reanalysis data set, less additional variability may be required from the fluctuations. On the other hand, if reanalysis data would be used directly without the downscaling described in Section 2.1, more variability may be required from the fluctuations. In general, the fluctuation parameter set should be recalibrated whenever the meteorological data set is changed or if a very different geographical area is analysed.

The presented methodology adds the fluctuations using stochastic simulation based on PSDs and coherence functions; simulation is carried out in frequency domain and the data are then transformed to time series [18]. Similar simulations could be carried out using ARMA-type models. The reason for choosing to simulate based on PSDs and coherence functions is that information is readily available on frequency domain. E.g. Ref. [12], describes variability in wind using PSDs, and turbulence is usually defined using a spectrum. For applying autoregressive–moving-average (ARMA) type models, such frequency domain information would first need to be transformed to parameters of the ARMA type model used. Reference [15] shows how this is achieved for univariate data (either one plant or aggregate of many plants). However, in the presented methodology it was considered more convenient that PSDs and coherence functions can be applied directly.

The presented methodology is based on WRF data, which are downscaled from the raw reanalysis data [21,22]. However, a more detailed downscaling (and/or more accurate raw reanalysis data) could provide the required additional high frequency information missing in the WRF data used in this paper. This is something that will be considered in future research. However, the presented methodology allows simulation of even higher frequency data than analysed in this paper; e.g., 5 min resolution data can be simulated efficiently using the presented approach, whereas it could be challenging to obtain 5 min resolution simulations over a large geographical area using downscaling.

The coherence functions between locations were estimated using (3), which depends on the mean wind speed V_0 (calculated from the hourly WRF data using linear interpolation). As V_0 is defined as 10 min average [18], after the addition of the fluctuations, (3) can change, as fluctuations impact the 10 min mean wind speeds. However, this was not taken into account. The linearly

interpolated WRF data were considered enough to estimate the changes in the coherence functions on different mean wind speeds.

6. Conclusion

This paper has shown that a combination of meteorological reanalysis data and stochastic simulation can successfully model variability in wind generation on different geographical levels of aggregation. A spectral parameter set was found for the fluctuation model based on measured wind speed data from three locations. With the calibrated fluctuation model, CorRES was used to simulate wind generation time series for different cases in Denmark, going from single OWPP to the aggregate onshore wind generation of western Denmark. These simulations were compared to measured data covering two years with 15 min resolution.

The results show that the fluctuations have a crucial role in modelling ramp behaviour when analysing offshore wind generation with geographically concentrated OWPP installations. When analysing individual onshore wind regions in Denmark (on average 2795 km^2 in size), the addition of fluctuations is required to accurately model the short-term variability. On the level of aggregate onshore wind generation of the whole of western Denmark (approximate land area of 33275 km^2), the WRF data can model most variability seen in the measurements even without fluctuations on the 15 min resolution using simply linear interpolation to reach intra-hour values. However, also on this level of aggregation, the simulation with fluctuations showed ramp behaviour closer to the measured data.

CRediT authorship contribution statement

Matti Koivisto: Conceptualization, Methodology, Formal analysis, Software, Validation, Data curation, Writing - original draft, Visualization. **Guðrún Margrét Jónsdóttir:** Methodology, Formal analysis, Software, Validation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Poul Sørensen:** Conceptualization, Methodology, Software, Writing - review & editing, Funding acquisition. **Konstantinos Plakas:** Formal analysis, Software, Validation, Data curation, Writing - review & editing. **Nicolaos Cutululis:** Conceptualization, Methodology, Writing - review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The work has been supported by the NSON-DK (Danish Energy Agency, EUDP, grant 64018-0032; previously ForskEL), Flex4RES (Nordic Energy Research, grant 76084) and PSfuture (La Cour Fellowship, DTU Wind Energy) projects. Guðrún Margrét Jónsdóttir is supported by the Science Foundation Ireland, under Investigator Programme, grant no. SFI/15/IA/3074. The authors acknowledge Energinet for providing the measurements, and Andrea Hahmann at DTU Wind Energy for running WRF and providing the data.

References

[1] T. Brown, D. Schlachtberger, A. Kies, S. Schramm, M. Greiner, Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system, Energy 160 (2018) 720–739.
[2] T. Souxes, I.-M. Granitsas, C. Vournas, Effect of stochasticity on voltage stability support provided by wind farms: application to the Hellenic interconnected system, Elec. Power Syst. Res. 170 (2019) 48–56.
[3] B. Klöckl, G. Papaefthymiou, Multivariate time series models for studies on stochastic generators in power systems, Elec. Power Syst. Res. 80 (3) (March 2010) 265–276.
[4] D. Villanueva, A. Feijóo, J.L. Pazos, Simulation of correlated wind speed data for economic dispatch evaluation, IEEE Trans. Sustain. Energy 3 (1) (January 2012) 142–149.
[5] M. Koivisto, J. Ekström, J. Seppänen, I. Mellin, J. Millar, L. Haarla, A statistical model for comparing future wind power scenarios with varying geographical distribution of installed generation capacity, Wind Energy 19 (4) (April 2016) 665–679.
[6] J. Ekström, M. Koivisto, I. Mellin, R.J. Millar, M. Lehtonen, A statistical model for hourly large-scale wind and photovoltaic generation in new locations, IEEE Trans. Sustain. Energy 8 (4) (October 2017) 1383–1393.
[7] I. González-Aparicio, F. Monforti, P. Volker, A. Zucker, F. Careri, T. Huld, J. Badger, Simulating European wind power generation applying statistical downscaling to reanalysis data, Appl. Energy 199 (August 2017) 155–168.
[8] I. Staffell, S. Pfenninger, Using bias-corrected reanalysis to simulate current and future wind power output, Energy 114 (November 2016) 1224–1239.
[9] G.B. Andresen, A.A. Søndergaard, M. Greiner, Validation of Danish wind time series from a new global renewable energy atlas for energy system analysis", Energy 93 (1) (December 2015) 1074–1088.
[10] E. Nuño, P. Maule, A. Hahmann, N. Cutululis, P. Sørensen, I. Karagali, Simulation of transcontinental wind and solar PV generation time series, Renew. Energy 118 (April 2018) 425–436.
[11] R. Frehlich, R. Sharman, The use of structure functions and spectra from numerical model output to determine effective model resolution, Mon. Weather Rev. 136 (April 2008) 1537–1553.
[12] X.G. Larsén, S. Ott, J. Badger, A.N. Hahmann, J. Mann, Recipes for correcting the impact of effective mesoscale resolution on the estimation of extreme winds, J. Appl. Meteorol. Climatol. (March 2012) 521–533.
[13] J. Olauson, M. Bergkvist, J. Rydén, Simulating intra-hourly wind power fluctuations on a power system level, Wind Energy 20 (2017) 973–985.
[14] S. Rose, J. Apt, Generating wind time series as a hybrid of measured and simulated data, Wind Energy 15 (2017) 699–715.
[15] E. Fertig, Simulating subhourly variability of wind power output, Wind Energy 22 (2019) 1275–1287.
[16] P. Sørensen, A. Hansen, P. Rosas, Wind models for simulation of power fluctuations from wind farms, J. *Wind Eng. Ind. Aerod.* 90 (12–15) (2002) 1381–1402.
[17] P. Sørensen, N.A. Cutululis, A. Vigueras-Rodriguez, L.E. Jensen, J. Hjerrild, M.H. Donovan, H. Madsen, Power fluctuations from large wind farms, IEEE Trans. Power Syst. 22 (3) (July 2007) 958–965.
[18] P. Sørensen, N.A. Cutululis, A. Vigueras-Rodriguez, H. Madsen, P. Pinson, L.E. Jensen, J. Hjerrild, M. Donovan, Modelling of power fluctuations from large offshore wind farms, Wind Energy 11 (1) (February 2008) 29–43.
[19] M. Koivisto, K. Das, F. Guo, P. Sørensen, E. Nuño, N. Cutululis, P. Maule, Using time series simulation tool for assessing the effects of variable renewable energy generation on power and energy systems, WIREs Energy Environ. 8 (3) (May/June 2019) e329.
[20] W. Skamarock, J. Klemp, J. Dudhia, D. Gill, D. Barker, M. Duda, X. Huang, W. Wang, J. Powers, Description of the Advanced Research WRF Version 3, 2008. Boulder, Colorado, USA.
[21] A.N. Hahmann, D. Rostkier-Edelstein, T.T. Warner, F. Vandenberghe, Y. Liu, R. Babarsky, S.P. Swerdlin, A reanalysis system for the generation of mesoscale climatographies, J. Appl. Meteorol. Climatol. (May 2010) 954–972.
[22] A.N. Hahmann, C.L. Vincent, A. Peña, J. Lange, C.B. Hasager, Wind climate estimation using WRF model output: method and model sensitivities over the sea, Int. J. Climatol. 35 (12) (October 2015) 3422–3439.
[23] MathWorks, Autoregressive power spectral density estimate - covariance method". https://se.mathworks.com/help/signal/ref/pcov.html. (Accessed 1 October 2019).
[24] DTU online meteorological data. http://rodeo.dtu.dk/. (Accessed 30 April 2019).
[25] Cesar, Cabauw experimental site for atmospheric research. http://www.cesar-database.nl/. (Accessed 30 April 2019).
[26] X.G. Larsén, C. Vincent, S. Larsen, Spectral structure of mesoscale winds over the water, Q. J. R. Meteorol. Soc. 139 (672) (April 2013) 685–700.
[27] G.M. Jónsdóttir, M. Koivisto, P. Sørensen, F. Milano, Modelling Wind Speeds Using CorRES: Combination of Mesoscale Reanalysis Data and Stochastic Simulations, Wind Integration Workshop, Dublin, Ireland, October 2019. http://faraday1.ucd.ie/archive/papers/windmeso.pdf. (Accessed 27 April 2020).
[28] The Wind Power, Onshore wind farm database, purchased online from. http://www.thewindpower.net/. (Accessed 15 November 2015).
[29] The Wind Power, Turbine and power curve database. http://www.thewindpower.net/. (Accessed 16 January 2018).
[30] PyWake software, DTU Wind Energy, available at: http://doi.org/10.5281/zenodo.2562662 (accessed on 1 October 2019).