

Mind the (spectral) gap: how the temporal resolution of wind data affects multi-decadal wind power forecasts

Nina Effenberger, Nicole Ludwig, Rachel H. White

Angaben zur Veröffentlichung / Publication details:

Effenberger, Nina, Nicole Ludwig, and Rachel H. White. 2024. "Mind the (spectral) gap: how the temporal resolution of wind data affects multi-decadal wind power forecasts." *Environmental Research Letters* 19 (1): 014015. <https://doi.org/10.1088/1748-9326/ad0bd6>.

Nutzungsbedingungen / Terms of use:

CC BY 4.0



LETTER • OPEN ACCESS

Mind the (spectral) gap: how the temporal resolution of wind data affects multi-decadal wind power forecasts

To cite this article: Nina Effenberger *et al* 2024 *Environ. Res. Lett.* **19** 014015

View the [article online](#) for updates and enhancements.

You may also like

- [Hourly disaggregation of daily wind projections: an analogue-based, spatially coherent approach to support energy applications](#)
Zainab Benseddik, Hannah Bloomfield and Charles Rougé
- [Spatial and temporal patterns of global onshore wind speed distribution](#)
Yuyu Zhou and Steven J Smith
- [Assessing the stability of wind resource and operating conditions](#)
S C Pryor, T J Shepherd, M Bukovsky et al.

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

OPEN ACCESS

RECEIVED

14 September 2023

REVISED

2 November 2023

ACCEPTED FOR PUBLICATION

10 November 2023

PUBLISHED

1 December 2023

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Mind the (spectral) gap: how the temporal resolution of wind data affects multi-decadal wind power forecasts

Nina Effenberger^{1,*} , Nicole Ludwig¹ and Rachel H White²¹ Cluster of Excellence Machine Learning, University of Tübingen, 72076 Tübingen, Germany² Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

* Author to whom any correspondence should be addressed.

E-mail: nina.effenberger@uni-tuebingen.de**Keywords:** wind energy, climate change, downscaling, data resolutionSupplementary material for this article is available [online](#)

Abstract

To forecast wind power generation in the scale of years to decades, outputs from climate models are often used. However, one major limitation of the data projected by these models is their coarse temporal resolution—usually not finer than three hours and sometimes as coarse as one month. Due to the non-linear relationship between wind speed and wind power, and the long forecast horizon considered, small changes in wind speed can result in big changes in projected wind power generation. Our study indicates that the distribution of observed 10 min wind speed data is relatively well preserved using three- or six-hourly instantaneous values. In contrast, daily or monthly values, as well as any averages, including three-hourly averages, are almost never capable of preserving the distribution of the underlying higher resolution data. Assuming that climate models behave in a similar manner to observations, our results indicate that output at three-hourly or six-hourly temporal resolution is high enough for multi-decadal wind power generation forecasting. In contrast, wind speed projections of lower temporal resolution, or averages over any time range, should be handled with care.

1. Introduction

Wind is one of the main sources of renewable power and its utilization is on the rise in many countries, (e.g. Soares-Ramos *et al* 2020). It is therefore of utmost importance to have reliable wind power forecasts in the range of years to decades (i.e. a turbine's lifetime) for site assessment and reliable future power supply (Copernicus 2023). However, the high variability and stochasticity of weather and wind introduces uncertainty that makes multi-decadal planning difficult. Furthermore, computational complexity limits the resolution of forecasts; the temporal and spatial resolution³ of long-term and multi-decadal forecasts are therefore usually much coarser than that of short-term forecasts (Eyring *et al* 2016).

³ Notice that temporal and spatial resolution are often directly linked (e.g. Courant *et al* 1928).

But do the low temporal resolution outputs from e.g. climate models provide wind speeds representative of the true site-specific high resolution wind speed distribution? An indicator that this could be the case is the so called wind power spectral gap (Van der Hoven 1957). Wind speed variability can be assessed in the frequency domain in terms of power spectra by (Fourier-) decomposing high-resolution wind speed observations. This decomposition reveals an amplitude gap between high and low frequencies where the strong variability in high frequencies (on the order of seconds to minutes) is associated with turbulence, while strong low frequency variability (hours to days) is associated with synoptic weather systems (Van der Hoven 1957, Stull 1988). In the gap between the synoptic weather and turbulence peaks in the frequency spectrum are frequencies with little variability; this was found to be in the range of 10^{-3} Hz (~ 17 min) to 10^{-4} Hz (~ 3 h) by (Van der Hoven 1957). However, other research suggests the gap is

smaller (Kang and Won 2016) or may not exist at all (Larsén *et al* 2016). This spectral gap is described in many research papers (Horvath *et al* 2012, Kang and Won 2015, Larsén *et al* 2016, Lopez-Villalobos *et al* 2021) indicating that observations that fall into the corresponding frequencies do not add much knowledge. Given the inconsistency of the width of this spectral gap, the question remains of what temporal resolution we should aim for in multi-decadal wind power forecasting.

Due to the non-linear relationship between wind speed and wind power, small changes in the wind speed distribution can lead to significant changes in wind power generation. However, if the underlying wind speed characteristics are preserved, lower-resolution data is preferred for multi-decadal forecasting to reduce required storage space and potentially also computational costs.

Additionally, multi-decadal wind speed projections should also account for climate change and interannual climate variability (Pryor *et al* 2018). Several studies indicate that climate change will affect average wind speeds (McVicar *et al* 2012, Tobin *et al* 2015)) as well as wind speed variability (Pryor and Barthelmie 2010, Tobin *et al* 2016, Dunn *et al* 2019, Jeong and Sushama 2019, Ringkjøb *et al* 2020) which will impact wind power generation. In general, it must be assumed that current wind conditions may not be representative of future wind conditions (Jung and Schindler 2019); we thus often rely on output from climate models to help inform about future potential wind power. But which data (resolution) do we need for meaningful wind speed and power predictions?

There seems to be agreement among researchers that certain temporal resolutions are *too low*. It is therefore common to account for additional variability using so-called downscaling techniques (Pryor and Barthelmie 2010). In the past, different statistical downscaling approaches have been introduced, e.g. (Von Storch 1999, Tobin *et al* 2015, Shin *et al* 2018), as well as dynamical downscaling using regional climate models such as CORDEX (Giorgi and Gutowski Jr 2015), used by (e.g. Davy *et al* 2018) and (Yang *et al* 2022). The question remains, however: what temporal resolution of wind speed data is *high enough*?

Outputs from climate models are often available in various temporal aggregations, where some datasets represent temporal averages and other datasets consist of instantaneous values. In the CMIP6 datasets (Eyring *et al* 2016), one of the most widely used set of global circulation models (GCMs), wind projections are available as temporal means and instantaneous values (CMIP6 Data Request 2016). Currently, however, little focus has been placed on the choice of data and many studies are performed without explicitly stating whether averages or instantaneous values are being used.

With this study we show that the type of data (instantaneous vs averages) influences the wind speed distributions and thus the estimated wind power generation. We also conduct analysis to determine whether there is a temporal resolution that is *high enough*, i.e. for which added temporal resolution provides little additional information and accuracy. We give recommendations regarding the choice of data and the temporal resolution that downscaling techniques, and climate model outputs, should aim for. To do this, we conduct empirical data analysis using data from eight different mid-latitude sites across Europe and North America. In section 2.1 we describe the data and the methodology used. Our results are presented in section 3 and we discuss their implications in section 4. Finally, we conclude in section 5.

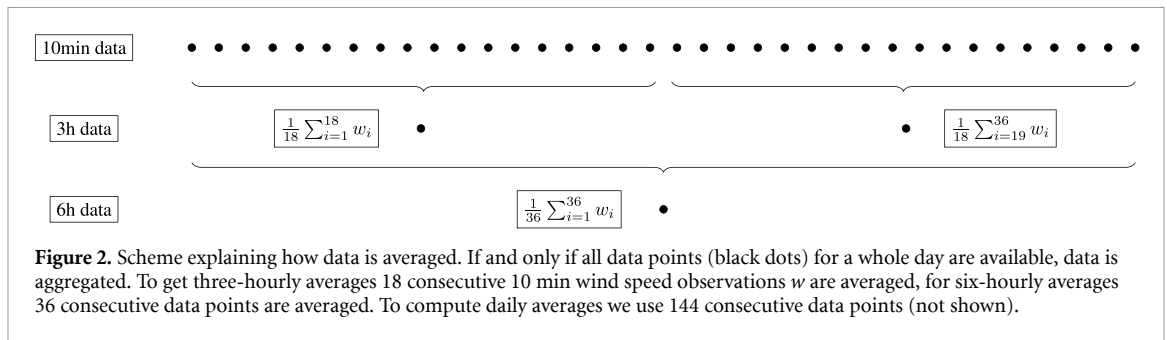
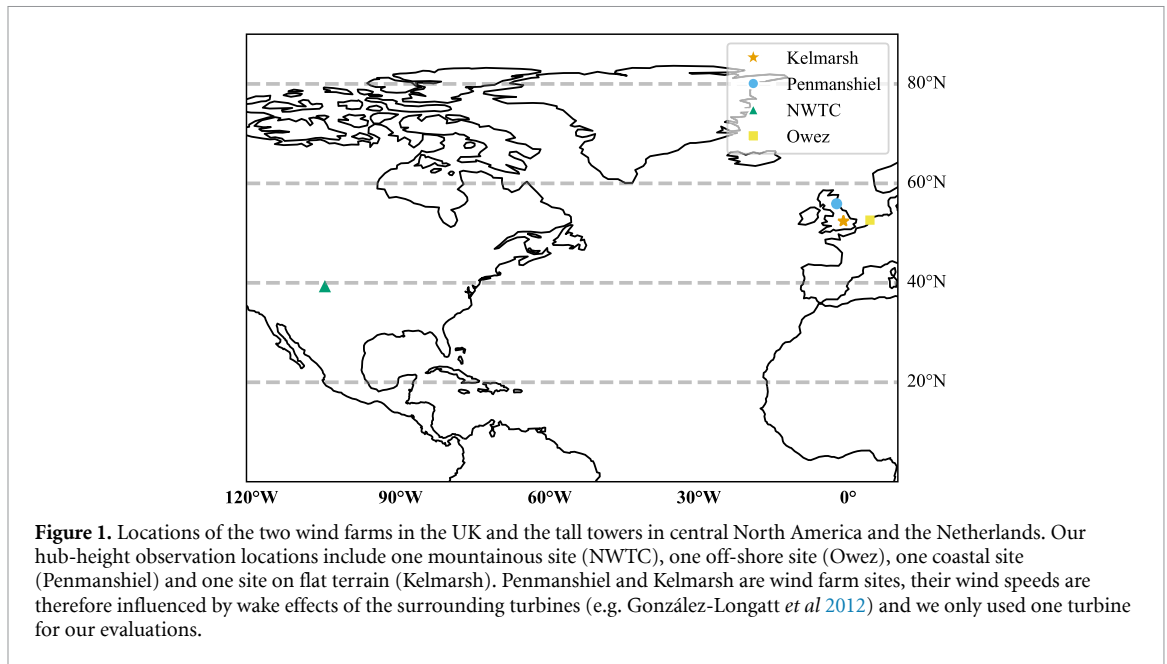
2. Methods

To investigate how aggregating wind speed data affects the wind speed distribution we use both parametric and non-parametric approaches. In the main manuscript we focus on observations of turbine-hub-height winds from four sites, with locations shown in figure 1. We use these as our primary data as they are hub-height data; however, all four sites have a relatively limited observational period of only a few years (see table 1). We thus also use 10 m wind speeds from an additional four observational met masts from locations across Germany (Wetterdienst 2023) (see figure A1 for locations) that have between 18 and 31 years of data available. These data show very similar results to the hub-height datasets (analysis presented in the supplementary material in section A). We also use these longer datasets to analyze multi-decadal tendencies of wind power generation in section 3.3. In the following we first give a description of the data used and then describe our methodology.

2.1. Data

We investigate wind speed observations using open-source high met mast wind data of four mid-latitude locations. All of the wind speeds are either measured at wind turbine hub-height directly, i.e. by nacelle anemometers (sites Penmanshiel and Kelmarsh) or by high met masts (sites NWTC and Owez). The observation heights are between 59 m and 116 m and all of the measurements are provided as 10 min averages, a very common aggregation-level of wind resource data (Harper *et al* 2010). In table 1 we present static information of the observation sites. Static information of the longer datasets (at 10 m height) is presented in table A.1.

To bring the data to a format where we can compare different temporal resolutions, we pre-process the data by excluding all days where at least one observation is missing. We then average the 10 min wind



speed observations to three-hourly, six-hourly and daily data: the n 'th wind speed value w_n in the time-series averaged over t consecutive time steps is computed as

$$w_n^t = \frac{1}{t} \sum_{i=nt}^{nt+t} w_i^1, \quad (1)$$

where $t = 18$ (three-hourly), 36 (six-hourly) and 144 (daily) respectively; for $t = 1$ we get the original 10 min resolution time series w^1 , see figure 2 for a visualization of this averaging process.

In addition to calculating averages, we also consider wind speed time series of lower resolution, which we call instantaneous values. To do so, we use wind speed measurements every t 'th time step only and exclude all other wind speed measurements w_i where $i \neq nt$. The results are eight observations per day (every three hours), four observations per day (every six hours) and one observation per day respectively.

2.2. Comparing wind speed distributions

To determine whether wind speed distributions from data of different temporal resolution are statistically different, we compute pairwise

Kolmogorov–Smirnov test statistics of cumulative density distributions

$$F_W(w) = P(W \leq w). \quad (2)$$

The Kolmogorov–Smirnov statistic D is given by:

$$D = \sup_w |T(w) - S(w)| \quad (3)$$

where w are the wind speed values, T and S are the wind speed distributions to be compared, and the supremum, $\sup$, is the largest value of the set of values $|T(w) - S(w)|$ across all w . The Kolmogorov–Smirnov test only takes the largest absolute difference between the two distributions across all w values into account and we identify a statistically significant difference if the p -value of an individual test is $p \leq 0.05$.

2.3. Modeling wind speeds using Weibull distributions

The Kolmogorov–Smirnov test can tell us whether wind distributions of different temporal resolution differ; in order to quantify the differences found, we model the wind speeds, w , as Weibull distributions. This is done by fitting the parameters of a three-parameter Weibull distribution,

Table 1. Static data of the four different sites with hub-height measurements. Our chosen datasets cover a large range of observation heights as well as different mean wind speeds and variances.

Name	Country (Lat, Lon)	Temporal resolution in min	Observation period (Number of days)	Observation height in m	Mean wind speed in $\frac{m}{s}$ (variance in $\frac{m^2}{s^2}$)	Data source
Kelmarsh	United Kingdom (52.40, −0.95)	10	2016–2021 (1687)	78.5	6.25 (7.70)	Plumley (2022a)
Penmanshiel	United Kingdom (55.90, −2.31)	10	2016–2021 (1522)	59	7.27 (15.30)	Plumley (2022b)
Owez	Netherlands (52.61, 4.39)	10	2012–2017 (1148)	116	4.82 (13.02)	Ramon <i>et al</i> (2020, 2022a)
NWTC	United States (39.21, −105.23)	10	2005–2010 (1083)	87	8.37 (16.55)	Ramon <i>et al</i> (2020, 2022b)

$$f(w; \beta, \lambda, \theta) = \frac{\beta}{\lambda} \left(\frac{w - \theta}{\lambda} \right)^{\beta-1} e^{-\left(\frac{w - \theta}{\lambda} \right)^\beta} \quad (4)$$

to the different temporal resolution datasets. The three-parameter Weibull distribution described in equation (4) is defined by $w \geq \theta$ and $f(w; \beta, \lambda, \theta) = 0$ for $w < \theta$ where $\beta > 0$ is the shape parameter, $\lambda > 0$ is the scale parameter and θ is the location parameter of the distribution which equals the lowest possible value of the distribution. For $\beta \approx 3$, the Weibull distribution approximates a Gaussian distribution, while $3 > \beta \geq 1$ corresponds to a right-skewed distribution, and $\beta > 3$ corresponds to a left-skewed distribution, for $\beta < 1$ the density values are steadily decreasing with increasing w . λ represents the variability, i.e. smaller values of λ are associated with less variability (Rinne 2008). We fit the parameters using maximum likelihood estimation (MLE). This approach requires maximizing the likelihood function

$$L = \prod_{i=1}^n f(w_i; \beta, \lambda, \theta). \quad (5)$$

We can then evaluate the change of the parameters when wind speeds are averaged or discarded to produce datasets with different temporal resolution. While Weibull distributions are commonly used to model wind speed distributions (e.g. Mert and Karakuş 2015), we additionally use a generalized Gamma distribution to test the sensitivity of our results to the choice of underlying distribution. The conclusions are unchanged (figures shown in the appendix) and we therefore consider our results to be insensitive to the distribution choice.

2.4. Validating the Weibull parametrization

Using a kernel density estimation we confirm that Weibull distributions are a reasonable representation of our data, with the exception of monthly averages. The kernel density estimator $\hat{f}$ of an unknown density f at a point x is defined by

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad (6)$$

where we choose K to be the Gaussian kernel

$$K(x, h) = \exp\left(-\frac{x^2}{2h^2}\right). \quad (7)$$

The band-width h is selected using Scott's rule (Scott 2015).

Figures A.1–A.4 show kernel density estimations for averaged and instantaneous wind speeds respectively; it is clear that monthly values cannot be described by a Weibull distribution and thus we do not include this temporal resolution in subsequent analysis. To validate the fit of the Weibull distributions to the original wind speed distributions we generate quantile-quantile plots of the observations of length $\frac{1}{t}$ against $\frac{1}{t}$ randomly drawn samples from the corresponding Weibull distribution; these plots are shown in figures A.5 and A.6. The Weibull distributions generally provide a good fit to the data, although in some locations the fit is less good at higher wind speeds.

2.5. Power generation and transferability of results

As a last step, we relate the insights from the wind speed distributions to multi-decadal wind power forecasts. Wind power generation is often forecasted using hub-height wind speed forecasts with a turbine-specific wind power curve (Wang *et al* 2019) that describes the relationship between wind (speed) and potential wind power generation, and is highly non-linear. We apply the Enercon E92/2350 wind power curve of a 79 m high turbine, visualized in figure 3. Given the relatively short duration of the main observational datasets we use in this study (around 3–5 years of non-missing data) and our interest in multi-decadal forecasts, we study potential power generation using four 18 to 31 year long wind speed datasets.

Additionally, to determine whether the results we find for observational datasets are also applicable to climate model data, we repeat our analysis using data generated by the historical run of the MPI-ESM1-2-LR general circulation model (Eyring *et al* 2016) which was the only model with three-hourly data available on the ESGF node (WCRP CMIP6 2023) at the time of research.

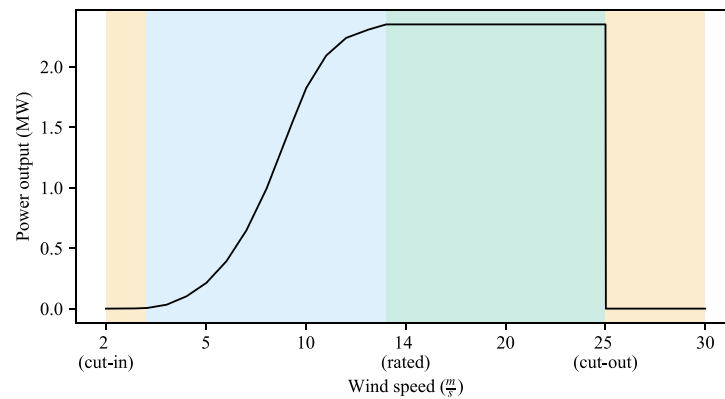


Figure 3. Wind power curve of an Enercon E92/2350 turbine of 79 m height. The relationship between wind speed and wind power can be roughly divided into four different regions: No power is generated if the wind speed is below the cut-in wind speed or if the wind speed is above the cut-out wind speed where the turbine is shut down to protect it from damage. In between, wind power generation first increases rapidly with increasing wind speed. Once the maximum wind speed that the wind turbine can convert to power is reached, the power output is usually constant until the wind speed exceeds the cut-out wind speed.

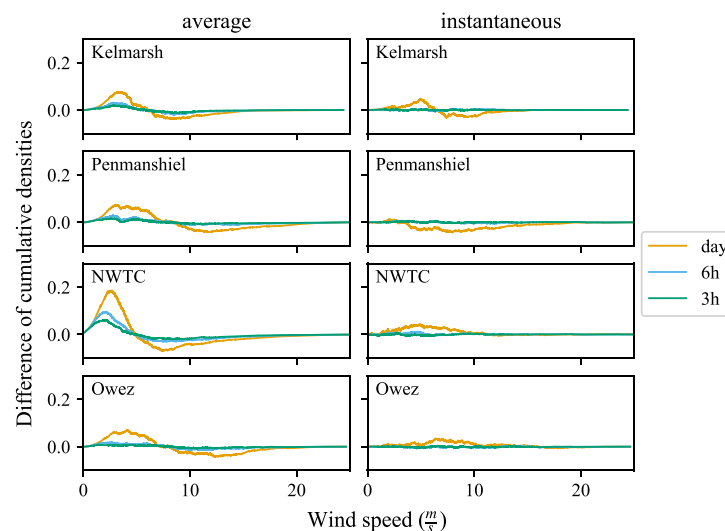


Figure 4. Difference of cumulative densities from the 10 min data to the other temporal resolution datasets for average wind speeds (left) and instantaneous wind speeds (right). It can be seen that the averaged wind speeds are visually distinguishable, which is less the case for instantaneous wind speeds, particularly for data with a temporal resolution of six-hourly or higher. The four datasets plotted are Kelmarsh, Penmanshiel, NWTC and Owez.

3. Results

We analyze the distributions of wind speed averages and instantaneous wind speed time series, and find consistent results across all sites investigated: averaging introduces shifts to the wind speed distributions, while three-hourly and six-hourly instantaneous data are usually close to the original 10 min data. This is clearly demonstrated in figure 4, with differences between 10 min wind speeds and averaged data (left hand column) larger than the differences between 10 min wind speeds and instantaneous wind

speeds of lower resolution (right hand column). Similar patterns are observed in the four longer datasets, see figure A.7.

The impact of averaging can also be seen in the variance of the data: while averaging does not affect the mean, it reduces the variance of the averaged time series (figure 5). In contrast, three-hourly and six-hourly instantaneous values (dashed lines in figure 5) preserve the variance, with daily instantaneous values preserving substantially more of the variance than daily averages, and at some sites, more than six-hourly averages.

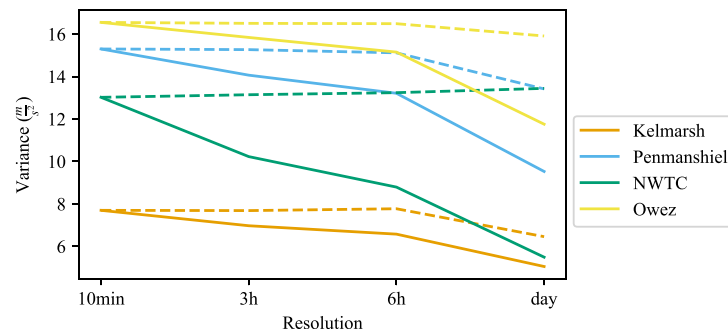


Figure 5. Variances of averaged (solid lines) and instantaneous (dashed lines) values. The variances steadily decrease when wind speeds are averaged and stay close to the 10 min variances for the instantaneous three-hourly and six-hourly distributions.

Table 2. Test statistics of Kolmogorov–Smirnov test for Penmanshiel averages. We reject the hypothesis that the wind speed distributions are equal if the p -value $p \leq 5 \times 10^{-2}$. Therefore, only the three-hourly averages are not significantly different from six-hourly averages. Values in **bold** indicate no statistically significant difference in the distributions.

	Daily	Six-hourly	Three-hourly	10 min
Daily	1	1.12×10^{-3}	2.94×10^{-5}	1.40×10^{-7}
Six-hourly		1	4.41×10^{-1}	5.83×10^{-3}
Three-hourly			1	5.83×10^{-3}

Table 3. Test statistics of Kolmogorov–Smirnov test for Penmanshiel instantaneous data. We reject the hypothesis that the wind speed distributions are equal if the p -value $p \leq 5 \times 10^{-2}$ which reveals that only daily instantaneous values are significantly different from all other distributions.

	Daily	Six-hourly	Three-hourly	10 min
Daily	1	3.03×10^{-2}	2.34×10^{-2}	6.62×10^{-3}
Six-hourly		1	1.00	9.58×10^{-1}
Three-hourly			1	8.07×10^{-1}

3.1. Kolmogorov–Smirnov tests

To quantify whether the difference between the 10 min wind speed distribution and the averaged and instantaneous distributions of lower resolution are statistically significant we compute Kolmogorov–Smirnov test statistics of their cumulative density distributions pairwise. The results for the Penmanshiel site are presented in table 2 (for averaged data) and table 3 (instantaneous data), where values in **bold** indicate no statistically significant difference in the distributions, i.e. that using the dataset of lower temporal resolution may retain all information about the wind speed distribution. We observe that all averaged distributions, as well as daily instantaneous distributions, differ significantly from the 10 min wind speeds (right-most column), whilst three- and six-hourly instantaneous values are not distinguishable from the 10 min data. The results for all other locations are very similar (see tables A.2–A.15): in general, averages do not preserve the wind speed distribution, and three- and six-hourly instantaneous values are almost always not statistically distinguishable from the 10 min data.

3.2. Changes in Weibull parameters

The cumulative density plots and Kolmogorov–Smirnov tests indicate that wind speed distributions are likely to *change* when wind speeds are averaged and likely to stay similar when measurements are discarded, at least until around six-hourly resolution. The aim of the next steps is to *quantify these changes* by parameterizing the distributions as Weibull distributions.

Figure 6 shows the values of the three Weibull parameters for the different aggregation levels for averaging (top row) and instantaneous (bottom row). The shape parameter β stays approximately constant across all aggregation levels and types, for both averaged and instantaneous data. For averaging, the location parameter θ increases with higher aggregation levels. This is consistent with the lowest values of the dataset increasing as they are averaged. Conversely, with lower resolution instantaneous data, the lowest values remain similar, leading to similar θ for all resolutions studied. The scale parameter λ decreases with increased averaging length, and remains relatively constant for instantaneous data, consistent with

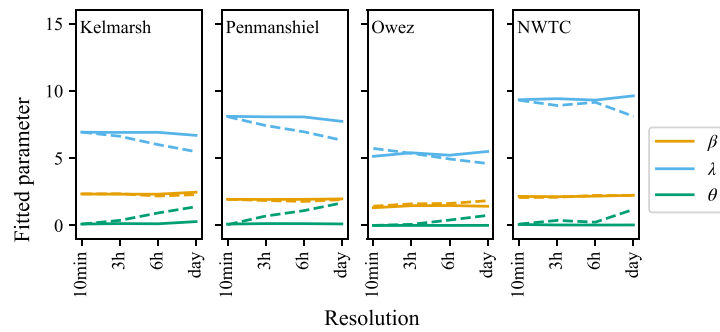


Figure 6. Parameters of Weibull distribution fitted to observational data averages (dashed lines) and instantaneous values (solid line). The parameters for the four datasets Kelmarsh, Penmanshiel, NWTC and Owez stay similar when data are discarded and change when data are averaged. The corresponding non-parameterized wind speed distributions are visualized in figures A.2 and A.4.

the changes in variance shown in figure 5. We observe very similar results for ~ 30 years of data measured at 10 m height (see figures A.8 and A.9). In addition, two of our observational sites, NWTC and Owez, have wind speed data at multiple heights, ranging from 10 m to 130 m; the changes in Weibull parameters seen in figure 6 are found at all different heights studied (see figure A.10).

To test for robustness of our results we also use an MLE to fit a generalized Gamma distribution (see appendix for details); both Weibull and Gamma distributions are regarded as suitable statistical models for wind speed data (e.g. Mert and Karakuş 2015). We find very similar results (see figures A.12 and A.13), with large changes in parameters when averaging data, and relatively small changes for three- and six-hourly instantaneous values. This suggests our conclusions are not sensitive to our choice of parameterization.

3.3. Implications for multi-decadal wind power forecasting

As introduced in section 1, wind speeds and wind speed variability are subject to interannual variability and climate change. Hence, for multi-decadal wind power forecasts climate models can provide useful information (e.g. Tobin *et al* 2015). To understand how our results apply on multi-decadal timescales, we repeat our analysis using data from four sites where multiple decades of 10 m wind speed observations are available. We also repeat our analysis on climate model output data to determine whether our conclusions are applicable to model data. We extract 10 m horizontal zonal and meridional wind speeds, u and v from the historical run of the CMIP6 model MPI-ESM1-2-LR for grid points closest to these four longer-term observational sites (locations shown in figure A.1) to compute wind speeds $w(l)$ at location l using the formula

$$w(l) = \sqrt{u(l)^2 + v(l)^2}. \quad (8)$$

We only use direct output from the model, and thus at daily temporal resolution we do not have instantaneous values, only averages.

For these multi-decadal datasets the Kolmogorov–Smirnov tests and Weibull parameterization analysis produce results comparable to those for the hub-height observations shown in the previous sections. The parameters of fitted Weibull distributions behave similarly in response to changing temporal resolution, with a decrease of λ as temporal resolution decreases indicating a decrease in variability (see figure 7). The climate model data do not show an increase in θ with decreasing temporal resolution, although for three-hourly and six-hourly averages θ fitted to the climate model data is very close to the parameters fitted to the observational data.

So far we have only looked at wind speeds and their distributions. However, these are just a proxy for wind power—our variable of interest—and wind power depends non-linearly on wind speed. To estimate the power a wind turbine could generate over its lifetime, we apply the Enercon E92/2350 power curve (see figure 3) to the four multi-decadal observational wind speed datasets as well as to the wind speeds between 1990 and 2010 from the corresponding closest grid-points in the CMIP6 dataset. We then compare the expected cumulative power generation of the highest available resolution (10 min in observations and three-hourly instantaneous values in CMIP6 model) to lower resolutions (three-hourly, six-hourly, and daily averages and instantaneous values). Although the highest available resolution in the CMIP6 model is only three-hourly, our previous analysis shows that three-hourly instantaneous values are closely aligned with 10 min data.

Figure 8 shows the expected cumulative power generation using wind speed observations of different resolutions. We show this as a fraction of the total power generation achieved when applying the wind power curve to the 10 min observational wind speed data and integrating over the whole time period. The dotted gray line shows 100% and thus

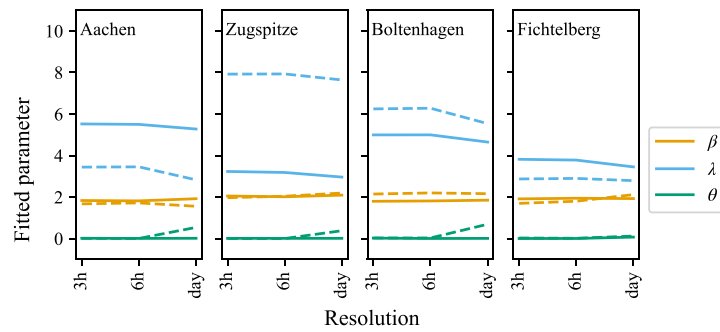


Figure 7. Parameters of Weibull distribution fitted to observational data averages (dashed lines) and to the closest grid points in the CMIP6 MPI-ESM1-2-LR dataset (solid lines). The parameter λ decreases in all cases when comparing three-hourly to daily averages, indicating a decrease in variability. The location parameter β does not show any consistent trends. θ increases with daily averages in observational data, but not in the CMIP6 data. The datasets used are from Aachen, Zugspitze, Boltenhagen and Fichtelberg.

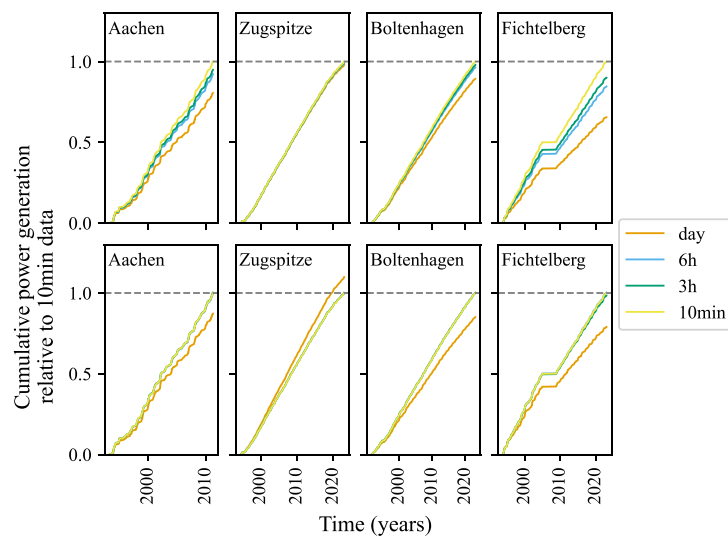


Figure 8. Cumulative power generation of daily, six-hourly and three-hourly values relative to the total 10 min wind power generation computed by feeding the wind speeds into a power curve. The datasets used are from Aachen, Zugspitze, Boltenhagen and Fichtelberg. Top row: average values underestimate wind power generation—the longer the estimation the higher the error. Bottom row: instantaneous values every three or six hours only introduce minor errors. Daily data is in no case a good proxy. The power generation gap at the Fichtelberg site between ~ 2005 and 2010 stems from various missing observations. For longer time periods more errors accumulate and we therefore report the absolute errors in table A.16.

if the expected cumulative power generation reaches this threshold without overshooting we consider the change introduced by the temporal resolution to be small. The top row of figure 8 shows that averaging values, particularly to daily, but even to three- or six-hourly at some locations, can lead to relatively large errors in estimated power generation, with errors of up to -34.48% (daily average), -15.45% (six-hourly average), and -10.06% (three-hourly average). Conversely, three-hourly and six-hourly instantaneous values, shown in the bottom row, reveal very similar results to the 10 min data, with errors less than 2% .

Figure 9 demonstrates that very similar results are found using the output from the MPI-ESM1-2-LR global climate model, with results given relative to the total amount of power generated using three-hourly instantaneous values. In all cases six-hourly instantaneous values are closest to three-hourly instan-

eous values, followed by three-hourly averages and six-hourly averages; daily averages differ substantially.

Reducing temporal resolution leads to an underestimation of expected power generation at all sites studied except for daily instantaneous values at Zugspitze (figure 8 bottom row, site (f)), a site with relatively high mean wind speeds and high wind speed variance, situated at a high elevation above sea level (2956 m). This is very likely a function of the particular wind turbine power curve we have chosen. It is important to note that in all cases, both with observational data and climate projections, the difference in power generation compared to higher-resolution data increases with an increasing forecast horizon and does not average out—the shift in the wind distribution leads to a systematic error in wind power estimation. In general, averaging tends to result in an underestimation of expected power generation, while discarding values has only minor impacts.

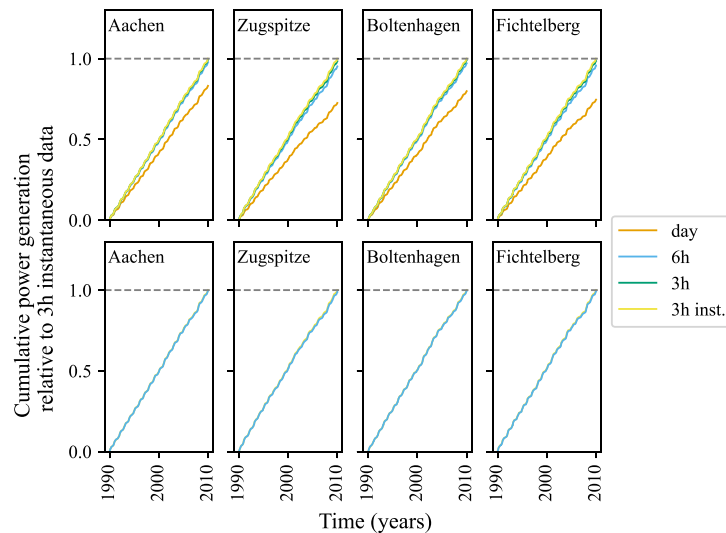


Figure 9. Cumulative power generation of daily, six-hourly and three-hourly values relative to three-hourly instantaneous (3 h inst.) wind power generation at the four CMIP6 locations closest to Aachen, Zugspitze, Boltenhagen and Fichtelberg computed by feeding the wind speeds into a power curve. Top row: average values underestimate wind power generation. Bottom row: instantaneous values (note that daily instantaneous values were not available as direct output, and so are not included). Using six-hourly instantaneous values introduces only minor errors relative to three-hourly instantaneous values.

4. Discussion

Global climate models simulate climate dynamics physically using partial differential equations. Their long forecast horizons make them computationally expensive with large storage space needs and their output is usually provided with a temporal resolution ranging from three hours to one month, either as instantaneous or averaged values. Using hub-height wind speed observations, this study investigates how the temporal resolution of wind speed data affects the wind speed distributions. Using multi-decadal observational datasets at 10 m height we study the corresponding potential power generation and which time resolution is actually necessary.

Using hub-height data we find that in all cases investigated, three-hourly and six-hourly instantaneous observations well-preserve the underlying 10 min wind speed distribution, whilst three- and six-hourly averaging leads to distributional shifts. However, whether a significant wind speed distribution shift results in a significant change in wind power generation depends on the turbine and its corresponding power curve. Our results, using an example turbine power curve, indicate that the differences in wind distribution highlighted in this study can lead to accumulating errors when power generation is forecasted for years to decades (see figures 8 and 9).

Our results are consistent across different observational sites and a GCM. We can therefore give two suggestions when working with wind speed projections for wind power modelling. First, instantaneous wind speed projections should be preferred over wind speed averages, as sub-sampling wind speed data introduces relatively minor errors in contrast

to averaging wind speeds, where we observe a characteristic distributional shift. This shift associated with averaging data indicates that we might consistently over- or underestimate wind power generation. Second, instantaneous wind speed projections of six or three hours suffice, whilst daily data may be too low resolution, even with instantaneous values. For the sites studied, temporal downscaling of either three- or six- hourly data is unlikely to provide substantial improvement in accuracy. For example, instantaneous observational wind speeds with a three-hourly temporal resolution lead to errors of less than 0.29%, with six-hourly data leading to errors of less than 1.57%. The experiments conducted using climate model data support these results. This knowledge can be useful to reduce storage space almost loss-free and decrease computational complexity in further applications using the data.

The primary shortcomings of our investigations include a potential lack of generalizability across turbines, sensitivity to our underlying highest temporal data resolution, and uncertain transferability to other climate models. More specifically, as wind speed data measured at hub-height is often confidential, we are limited to relatively few datasets. Furthermore, for site assessment, wind speeds are usually transformed to hub-height which is non-trivial (e.g. Bañuelos-Ruedas *et al* 2010). However, the close agreement of results from stations across a range of different locations, including different local geography (off-shore, on-shore, different altitudes and local topography), suggests that our results likely hold for the majority of locations. For 7 out of the 8 locations studied, using daily instantaneous or average values underestimates the power generation (see figure 8), while

six-hourly instantaneous values are good approximations. For one station, however, the daily instantaneous data overestimates power; understanding the conditions under which daily data over- or underestimates the power generation would be useful for studies in which only daily data is available. In addition, while the Enercon power curve we use in this study has the characteristic shape of any modern horizontal wind turbine power curve, it is not necessarily representative of the turbines at a particular site.

Regarding the sensitivity to the underlying data resolution, we use data in 10 min resolution, hence variability on shorter time scales—often associated with turbulence (e.g. Stull 1988)—is not preserved. However, higher temporal resolution data is rarely available or used in wind power forecasting (Effenberger and Ludwig 2022, Tawn and Browell 2022), and thus we assume that this is a minor issue. This claim is also supported by (Lopez-Villalobos *et al* 2021) who investigate wind power spectra (Van der Hoven 1957) and find only small differences between power production given wind speeds of different resolution between 1 min and 6 hours.

Lastly, climate projections are characterized by different pathways that describe anthropogenic climate change (Eyring *et al* 2016). This makes handling the data cautiously even more important, especially in wind power forecasting, where non-linearly dependent wind speed projections are often used as a proxies for power generation. Our results using one historical CMIP6 run indicate that changes in observational wind speed distributions for different temporal resolution data can be seen in data from climate models as well. However, future research has to investigate the sensitivity of these results to different climate models.

5. Conclusion

Wind power generation depends non-linearly on wind speeds. For multi-decadal wind power forecasts in the order of years to decades, small changes in modelling the wind speed distributions can result in large systematic errors in power estimation, with absolute errors increasing with longer forecast horizon. Using wind speeds from hub-height observations and 10 m climate model output, this study investigates how the temporal resolution of wind speed data affects wind speed distributions and corresponding potential power generation.

We show that instantaneous wind speeds of lower resolution more accurately represent the underlying distribution of higher resolution data when compared to averaged wind speeds. Three- and six-hourly instantaneous values preserve the wind speed distribution of 10 min wind speed averages well. Small changes in the wind speed distribution, through averaging or using daily data, has significant impacts on the estimated wind power generation of a turbine over its lifetime. These results hold true across several

observational sites and a global climate model. Based on our results, we argue that modelling wind speed *distributions* correctly is what we should aim for in multi-decadal wind power forecasting.

Data availability statement

The data that support the findings of this study will be openly available following an embargo at the following URL/DOI: <https://github.com/NinaEffenberger/wind-variability> (Effenberger 2023). Data will be available from 31 December 2023.

Acknowledgments

This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC Number 2064/1—Project Number 390727645 and the Athene Grant of the University of Tübingen. The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Nina Effenberger, and the Natural Sciences and Engineering Research Council of Canada (NSERC) [RGPIN-2020-05783] for supporting Rachel H White. We acknowledge support by Open Access Publishing Fund of University of Tübingen. Open Access funding enabled and organized by Projekt DEAL. WOA Institution: Eberhard Karls Universität Tübingen Consortia Name : Projekt DEAL. Nina wants to thank Roland Stull and the Weather Forecast Research Team at UBC for (intellectual) resources!

ORCID iD

Nina Effenberger  <https://orcid.org/0000-0002-0713-1164>

References

- Bañuelos-Ruedas F, Angeles-Camacho C and Rios-Marcuello S 2010 Analysis and validation of the methodology used in the extrapolation of wind speed data at different heights *Renew. Sustain. Energy Rev.* **14** 2383–91
- CMIP6 Data Request 2016 CMIP6 documentation (available at: <https://clipc-services.ceda.ac.uk/dreq/u/4b97609a32d53dff5b8e73729e4f258b.html>) (Accessed 7 July 2021)
- Copernicus 2023 Decadal energy predictions (available at: <https://climate.copernicus.eu/decadal-predictions-energy>)
- Courant R, Friedrichs K and Lewy H 1928 Über die partiellen Differenzengleichungen der mathematischen Physik *Math. Ann.* **100** 32–74
- Davy R, Gnaniuk N, Pettersson L and Bobylev L 2018 Climate change impacts on wind energy potential in the European domain with a focus on the Black Sea *Renew. Sustain. Energy Rev.* **81** 1652–9
- Dunn R J H, Willett K M and Parker D E 2019 Changes in statistical distributions of sub-daily surface temperatures and wind speed *Earth Syst. Dyn.* **10** 765–88
- Effenberger N 2023 Wind speed data subset *Wind-variability* (available at: <https://github.com/NinaEffenberger/wind-variability>)

- Effenberger N and Ludwig N 2022 A collection and categorization of open-source wind and wind power datasets *Wind Energy* **25** 1659–83
- Eyring V, Bony S, Meehl G A, Senior C A, Stevens B, Stouffer R J and Taylor K E 2016 Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization *Geosci. Model Dev.* **9** 1937–58
- Giorgi F and Gutowski Jr W J 2015 Regional dynamical downscaling and the CORDEX initiative *Ann. Rev. Environ. Res.* **40** 467–90
- González-Longatt F, Wall P and Terzija V 2012 Wake effect in wind farm performance: steady-state and dynamic behavior *Renew. Energy* **39** 329–38
- Harper B A, Kepert J D and Ginger J D 2010 *Guidelines for Converting Between Various Wind Averaging Periods in Tropical Cyclone Conditions* (World Meteorological Organization)
- Horvath K, Koracin D, Vellore R, Jiang J and Belu R 2012 Sub-kilometer dynamical downscaling of near-surface winds in complex terrain using WRF and MM5 mesoscale models *J. Geophys. Res. Atmos.* **117** D11
- Il Jeong D and Sushama L 2019 Projected changes to mean and extreme surface wind speeds for North America based on regional climate model simulations *Atmosphere* **10** 497
- Jung C and Schindler D 2019 Changing wind speed distributions under future global climate *Energy Convers. Manage.* **198** 111841
- Kang S-L and Won H 2015 Intra-farm wind speed variability observed by nacelle anemometers in a large inland wind farm *J. Wind Eng. Ind. Aerodyn.* **147** 164–75
- Kang S-L and Won H 2016 Spectral structure of 5 year time series of horizontal wind speed at the Boulder Atmospheric Observatory *J. Geophys. Res. Atmos.* **121** 11946
- Larsén X G, Larsen S E and Petersen E L 2016 Full-scale spectrum of boundary-layer winds *Bound.-Layer Meteorol.* **159** 349–71
- Lopez-Villalobos C A, Rodriguez-Hernandez O, Martinez-Alvarado O and Hernandez-Yepes J G 2021 Effects of wind power spectrum analysis over resource assessment *Renew. Energy* **167** 761–73
- McVicar T R et al 2012 Global review and synthesis of trends in observed terrestrial near-surface wind speeds: implications for evaporation *J. Hydrol.* **416–417** 182–205
- Mert I and Karakuş C 2015 A statistical analysis of wind speed data using Burr, generalized gamma and Weibull distributions in Antakya, Turkey *Turk. J. Electr. Eng. Comput. Sci.* **23** 1571–86
- Plumley C 2022a Kelmarsh wind farm data (available at: <https://doi.org/10.5281/zenodo.7212475>) (Accessed 7 November 2023)
- Plumley C 2022b Penmanshiel wind farm data (available at: <https://doi.org/10.5281/zenodo.5946808>) (Accessed 7 November 2023)
- Pryor S C and Barthelmie R J 2010 Climate change impacts on wind energy: a review *Renew. Sustain. Energy Rev.* **14** 430–7
- Pryor S C, Shepherd T J and Barthelmie R J 2018 Interannual variability of wind climates and wind turbine annual energy production *Wind Energy Sci.* **3** 651–65
- Ramon J, Lledó L, Pérez-Zanón N, Soret A and Doblas-Reyes F J 2020 The Tall Tower Dataset: a unique initiative to boost wind energy research *Earth Syst. Sci. Data* **12** 429–39
- Ramon J, Lledó L, Pérez-Zanón N, Soret A, and Doblas-Reyes F J 2022a Owez wind farm data (available at: <https://talltowers.bsc.es/node/4856>) (Accessed 7 November 2023)
- Ramon J, Lledó L, Pérez-Zanón N, Soret A, and Doblas-Reyes F J 2022b NWTC wind farm data (available at: <https://talltowers.bsc.es/node/4846>) (Accessed 7 November 2023)
- Ringkjøb H-K, Haugan P M, Seljom P, Lind A, Wagner F and Mesfun S 2020 Short-term solar and wind variability in long-term energy system models—a European case study *Energy* **209** 118377
- Rinne H 2008 *The Weibull Distribution: A Handbook* (CRC press)
- Scott D W 2015 *Multivariate Density Estimation: Theory, Practice and Visualization* (Wiley)
- Shin J-Y, Jeong C and Heo J-H 2018 A novel statistical method to temporally downscale wind speed Weibull distribution using scaling property *Energies* **11** 633
- Soares-Ramos E P P, de Oliveira-Assis L, Sarrias-Mena R and Fernández-Ramírez L M 2020 Current status and future trends of offshore wind power in Europe *Energy* **202** 117787
- Stull R B 1988 *An Introduction to Boundary Layer Meteorology* vol 13 (Springer Science & Business Media)
- Tawn R and Browell J 2022 A review of very short-term wind and solar power forecasting *Renew. Sustain. Energy Rev.* **153** 111758
- Tobin I et al 2016 Climate change impacts on the power generation potential of a European mid-century wind farms scenario *Environ. Res. Lett.* **11** 034013
- Tobin I, Vautard R, Balog I, Bréon F-M, Jerez S, Michele Ruti P, Thais F, Vrac M and Yiou P 2015 Assessing climate change impacts on European wind energy from ENSEMBLES high-resolution climate projections *Clim. Change* **128** 99–112
- Van der Hoven I 1957 Power spectrum of horizontal wind speed in the frequency range from 0.0007 to 900 cycles per hour *J. Atmos. Sci.* **14** 160–4
- Von Storch H 1999 On the use of “inflation” in statistical downscaling *J. Clim.* **12** 3505–6
- Wang Y, Qinghua H, Linhao Li, Foley A M and Srinivasan D 2019 Approaches to wind power curve modeling: a review and discussion *Renew. Sustain. Energy Rev.* **116** 109422
- WCRP CMIP6 2023 ESGF data node (available at: <https://esgf-data.dkrz.de/search/cmip6-dkrz/>) (Accessed 27 November 2023)
- Wetterdienst D 2023 Open data german meteorological service (available at: https://opendata.dwd.de/climate_environment/CDC/) (Accessed 11 July 2023)
- Yang Y, Javanroodi K and Nik V M 2022 Climate change and renewable energy generation in Europe—long-term impact assessment on solar and wind energy using high-resolution future climate data and considering climate uncertainties *Energies* **15** 302