Using bias-corrected reanalysis to simulate current and future wind power output

Iain Staffell [a], [*], Stefan Pfenninger [b]

[a] Centre for Environmental Policy, Imperial College London, London SW7 2AZ, UK
[b] Department of Environmental Systems Science, ETH Zürich, 8092 Zürich, Switzerland

ARTICLE INFO

ABSTRACT

Reanalysis models are rapidly gaining popularity for simulating wind power output due to their convenience and global coverage. However, they should only be relied upon once thoroughly proven. This paper reports the first international validation of reanalysis for wind energy, testing NASA's MERRA and MERRA-2 in 23 European countries. Both reanalyses suffer significant spatial bias, overestimating wind output by 50% in northwest Europe and underestimating by 30% in the Mediterranean. We derive national correction factors, and show that after calibration national hourly output can be modelled with R^2 above 0.95. Our underlying data are made freely available to aid future research.

We then assess Europe's wind resources with twenty-year simulations of the current and potential future fleets. Europe's current average capacity factor is 24.2%, with countries ranging from 19.5% (Germany) to 32.4% (Britain). Capacity factors are rising due to improving technology and locations; for example, Britain's wind fleet is now 23% more productive than in 2005. Based on the current planning pipeline, we estimate Europe's average capacity factor could increase by nearly a third to 31.3%. Countries with large stakes in the North Sea will see significant gains, with Britain's average capacity factor rising to 39.4% and Germany's to 29.1%.

1. Introduction

Energy systems modellers require high resolution time series of power output from national fleets of wind farms, as their variable and unpredictable nature poses increasing challenges for the world's electricity systems. Power systems models such as WeSIM, Plexos and ANTARES, plus energy systems models such as TIMES, PRIMES and EnergyPLAN require external data to represent the contribution from wind as it cannot be controlled, and is not dispatched according to economic or market rationale. Such data are often difficult to acquire or simulate accurately, hindering research in this critical field.

Reanalysis — the output from global atmospheric simulations — is rapidly gaining popularity for simulating renewable energy resources due to its convenience and global coverage. However, as with any new technology, using reanalysis to synthesise wind outputs should only be relied upon once thoroughly proven. The prevailing approach within the wind research community is to use downscaled models and more detailed terrain data. Commercial tools such as WaSP, Virtual Met Masts, 3TIER and Vortex are widely used within the wind energy community as their results are significantly more accurate. This accuracy comes at the cost of complexity, and thus computational resource, intellectual capital and data requirements. Those in the wider energy research community may therefore see value in a simpler method for modelling wind power output aggregated to regional or national scales, rather than from individual farms. We demonstrate a technique using reanalysis with no downscaling and very limited wind farm characteristics, which is able to represent national fleet output very well across Europe with the use of simple correction factors.

In the last decade, wind power has achieved mainstream status and risen to dominate the international energy research agenda. Wind now commands a significant share of the world's electricity supply: global capacity stood at around 350 GW at the start of 2015, and is rising at 35% per year [1,2]. Around 135 GW of this is located in Europe, which hosts 87,000 wind turbines grouped into 17,000 farms, as mapped in Fig. 1. Wind now comprises 13% of installed capacity in Europe (a greater share than nuclear power), and can provide more than the entire national demand in five countries,

* Corresponding author.
    E-mail addresses: i.staffell@imperial.ac.uk (I. Staffell), stefan.pfenninger@usys.ethz.ch (S. Pfenninger).

including Germany and Spain.

Fig. 2 plots the recent growth in wind capacity against the range of hourly electricity demand in selected countries, showing that it is now physically possible for large parts of Europe to be completely powered by wind when peak output coincides with minimum demand, if it were not for issues with system control (stability, inertia, reserve, etc.). This figure makes no account for the timing of wind output, or its correlation with demand for electricity, but helps to put each country's installed capacity into context.

The rapid expansion in capacity has brought with it increased problems due to the variable and uncontrollable nature of wind output. Wind has profound effects on electricity markets: pushing down power prices and changing investment patterns [4], increasing the need for infrastructure upgrades [5], and forcing countries to increase cooperation between their electricity markets to maintain efficiency [6] and security [7].

The profiles of wind power output need to be understood at regional, national and continental scale, but this has traditionally been very difficult. Research into all aspects of renewable energy is reliant on a core foundation: high quality data. Wind power output is variable and weather-dependent, with complex correlations over space and time, and against human activity (and thus demand for electricity) [8]. Historic data is lacking due to commercial confidentiality, and significant time or finance is needed to produce a credible simulation, which poses a significant barrier to research across this field.

Conversely, meteorological data is often freely available, but significant time and knowledge is required to acquire, understand, process, correct and utilise it. A recent econometric study into wind farm degradation by the Renewable Energy Foundation [9] serves as an important warning: it lacked reliable data and any form of validation, but despite its results being widely discredited [10,11] it caused significant alarm among press and financial circles [12].

A recent development that could counter these problems is the use of freely-available reanalysis data produced by global weather models assimilating historic observations. Studies which use reanalysis to simulate wind outputs began to proliferate in 2014, but they either consider a broad region (Europe, US, China) with no validation against actual output [13–17], or are limited to validating in a single country [10,18–23].

This limits the confidence that can be placed in such studies, which dampens the impressive scope of their conclusions. Most energy systems models rely on exogenous inputs on the level and pattern of wind outputs, including the influential TIMES, PRIMES, EnegyPlan and IEA World Energy Model. Many modelling efforts the world over could therefore provide improved results and insights by using rigorously validated and calibrated reanalysis data as an input.

This paper provides the first international validation of reanalysis for simulating wind energy, drawing on historic data collected from 23 countries in Europe covering the period of 2005–2014. It shows that reanalysis can be a very powerful tool, simulating national fleet outputs with high accuracy and stability; however, this is only possible with a careful calibration, as we expose significant spatial bias in the underlying reanalysis data.

The next section introduces reanalysis, bias correction and previous studies in the field. Section 3 details our data sources and methods, and then Section 4 presents our Europe-wide validation. Section 5 summarises the long-term characteristics of the European wind resource and the rate at which capacity factors are improving because of better turbines and site selection. Section 6 concludes and considers further applications.

An extended validation and results section are provided as an online supplement to this article. Readers who wish to make use of this technique can download our core dataset and use a web-based interface for generating new results anywhere on the planet via www.renewables.ninja.

2. Background

2.1. Synthesising wind output data

The importance and complexity of wind output means it has been extensively studied over the last decade. Prominent studies include Sinden's assessment of the UK wind resource [8], the Pöyry and TradeWind reports on intermittency and expanding electricity markets [5,24], and studies into the feasibility of highly/fully renewable energy systems by Lu [13], Heide [25], Jacobsen [26] and Becker [16], for example.

Early studies such as Sinden [8] and Pöyry [5] were based on assumed long-run average capacity factors and wind speeds from ground-based metrological masts. These typically give observations from 10 m above ground (much lower than turbine hub heights), and the raw data requires substantial cleaning and cross-checking, as downtime is significant and jumps in the data are observed when measuring instruments are replaced. The location of these stations is also weighted towards population centres and airports for weather forecasting purposes, leaving the locations of wind-farms under-represented.

Within the wind research community, prognostic models such as MM5, RUC and WRF have been used to downscale reanalysis data into smaller and more accurate geographical grids [32]. These approaches take into account surface parameters such as topography and site-specific land use. Such methods are particularly relevant for accurate short-term forecasting of wind power at specific sites [32,33], as their results are significantly more accurate.

However, the downscaling process is costly in terms of time, intellectual capital, computing resources and data requirements. We show that this effort may be unnecessary for some applications, as using reanalysis directly with a simple correction factor yields
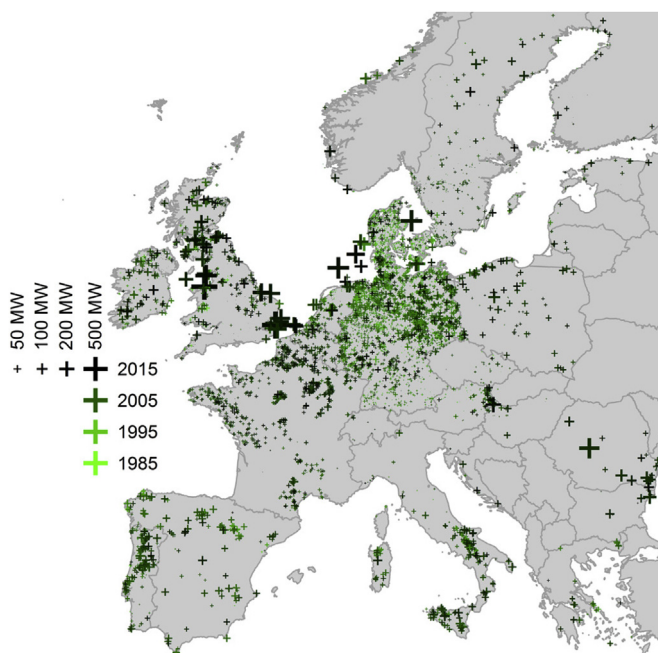

Fig. 1. Europe's wind farms as of 2015. Darker colours signify newer farms, and marker size is proportional to farm capacity. Data from Ref. [1]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
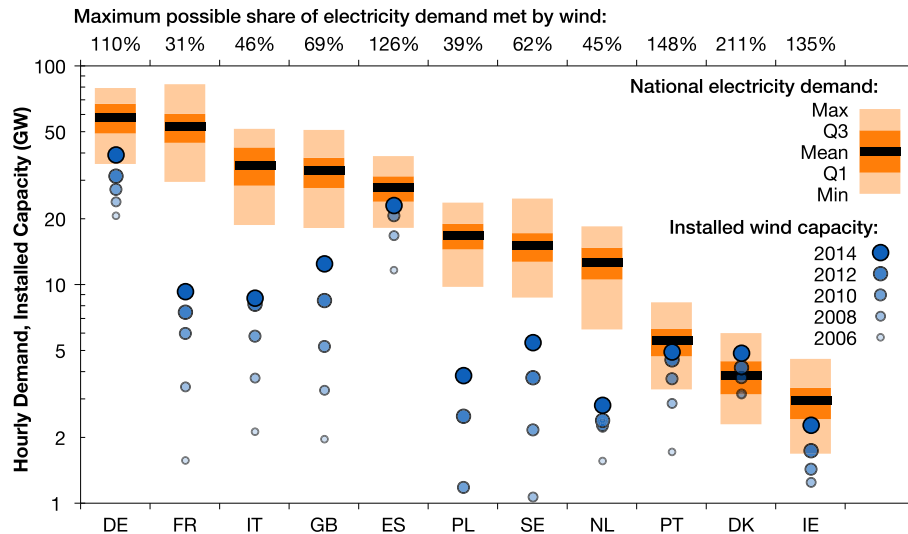
Fig. 2. Installed wind capacity in selected countries compared with the level of electricity demand. Countries are referred to by two-letter ISO codes, which are listed in the online supplement. Demand during 2006–2014 [3] is shown as stacked bars (Q1 and Q3 signify the 1st and 3rd quartiles). The evolution of wind capacity over this period is shown as circles. The top axis shows the hypothetical maximum percentage of demand wind could meet if peak output coincided with minimum demand.

national-scale results with high accuracy. While this is not sufficiently accurate for detailed modelling of site-specific conditions, it may be relevant to energy systems modellers working at spatially aggregated scales, for whom having 20+years of internally consistent renewable generation is valued, and where there are many other and more substantial sources of uncertainty.

More recently, reanalysis datasets have been explored as a means of simulating wind power production. These are the product of an atmospheric model set to match historic weather observations, and contain estimated weather parameters on a regular grid, often with global coverage, spanning several decades.

The first studies to use reanalysis modelled the power outputs at individual locations in Hungary [27] and Northern Ireland [28]. In the last two years, studies have expanded their scope to validate models against regional or national aggregate output in the UK [10,18–20,23], Denmark [21] and Sweden [22].

These studies focus on relatively small and geographically similar areas — coastal countries in northern Europe. Furthermore, many of these studies are validated against wind speeds rather than power outputs — neglecting the complexity of transforming from meteorological input to electrical output. Sharp provides a comprehensive overview of recent studies and their validation methods [20].

Another growing body of literature uses reanalysis to quantify and explore the wind resources of regions and countries, but with no effort made to validate the results. For example:

- Lu estimated the global potential of wind energy, but without validation [13].
- Gunturu and co-authors characterised the wind resources in the US [29], Europe [14] and Australia [30], all apparently without validation.
- Huber simulated wind and solar outputs across Europe, validating the ramping rates between hours (i.e. the fine-grain pattern) as opposed to the overall levels [15].
- Becker et al. simulated 100%-renewable wind and solar mixes in the US [16], but only hinted at approximate visual comparison to resource maps as a means of validation.
- McKenna et al. (2015) estimated the levelised cost of wind energy across Europe, suggesting that "future work should focus

on [...] validating estimated potentials with outputs from actual wind turbines" [17].

The results presented by Lu, Cosseron, Huber and McKenna all appear to feature the same spatial bias across Europe that we reveal in Section 4.1; most easily identified by capacity factors that are unreasonably high in Britain and Northern Germany (30–50%) and unreasonably low in Romania (5–15%), Portugal and Spain (10–25%). We contest that the results of such uncorrected studies cannot be relied upon, as the capacity factors derived from uncorrected reanalysis models may be out by up to ±50%.

2.2. Reanalysis — a convenient source of weather data

Reanalysis combines a system for assimilating historic weather observations with an atmospheric circulation model to infer the state of the global weather system. The model is set to replicate historic observations from satellites, ground observatories, ships, aircraft, etc., producing a hindcast as opposed to a forecast. In essence, reanalyses take difficult to use observations, apply automated quality control, and transform them into a standardised dataset with uniform and complete spatial and temporal coverage [31]. For this reason, reanalyses are widely used for commercial applications in the energy sector for understanding the availability of renewable resources (wind, solar and hydro).

Several reanalysis products are available, as listed in Table 1. Wind speeds are most commonly available at a fixed height of 10 m above ground, only MERRA and ERA-20C provide other heights closer to those used by wind turbines. Wind speed variables are also available at other model heights, usually based on fixed pressure or isothermal levels. The height of these levels above ground is not constant, and often well outside the region of interest (>250 m or <0 m).

A key benefit of reanalyses is that they can infer variables for which there are no observations; for example wind speeds at 50 m (met masts are usually only 10 m tall), in locations that are either remote or out to sea (where met masts are not present). This raises a fundamental issue with reanalysis data: whilst it is very convenient, it is just the output of a coarse model. Put simply: can it be trusted?

Table 1
Overview of publicly available reanalysis datasets and the parameters most relevant to wind power synthesis.

Institution/model	Released	Coverage	Spatial resolution (lat x lon, degrees)[a]	Time resolution (hours)	Wind speed heights	Other model heights
ECMWF/ERA40	2004	1957−2002	1.125 × 1.125	6	10 m	60
ECMWF/ERA-Interim	2006	1979 − present	0.75 × 0.75[b]	6	10 m	60
ECMWF/ERA-20C	2012	1900−2010	1.125 × 1.125[b]	3	10, 100 m	91
ECMWF/ERA5	2017	1979 − present	0.281 × 0.281	1	10, 100 m	137
JMA/JRA-25	2004	1979−2004	1.125 × 1.125	6	10 m	40
JMA/JRA-55	2013	1958 − present	0.5625 × 0.5625	6	10 m	60
NASA/MERRA	2009	1979 − present	0.5 × 0.667	1	2, 10, 50 m	72
NASA/MERRA-2	2015	1980 − present	0.5 × 0.625	1	2, 10, 50 m	72
NCEP/R2	2001	1979−2012	2.5 × 2.5	6	10 m	28
NCEP/CFSR	2009	1979−2010	0.3125 × 0.3125	1	10 m	6
NCEP/CFSv2	2011	2011 − present	0.205 × 0.205	1	10 m	6
NOAA/20CRv2	2010	1871−2011	2 × 2	6	10 m	28

[a] Native model resolutions are presented, data may be available at lower resolutions using statistical downscaling.
[b] ECMWF uses reduced Gaussian Grids with lower horizontal resolution (longitude) closer to the poles.

2.3. The need for bias correction

A key factor with the studies listed in §2.1 is the need for calibration, or bias correction, to bring simulated capacity factors in line with reality. While reanalyses may be able to replicate the pattern of output over time, they are not able to accurately assess the overall level of output, or its variation over space. The reasons for this are three-fold [31,34]:

1. Reanalyses are less than perfect computer models, and are known to contain systematic errors (biasses) due to errors in the underlying weather forecast model;
2. Their spatial coarseness means they are unable to resolve the detailed topography of a particular region, missing out on speed-up and blockage effects; and
3. The wind speed observations they attempt to replicate are not representative of wind farm sites, being primarily inferred from satellite data, and ground observations from short met masts.

Stickler notes that "while reproducing quite well the interannual variability, reanalysis products have been found to contain major biases" [35]. Decker adds that "at monthly time scales, the bias terms in reanalysis products are the dominant source of mean-square errors, as opposed to the correlation term which becomes the dominant source at hourly time scales" [36].

The bias in wind speeds appears to have received little attention, as the primary focus has been on temperature and precipitation [35,36]. However, errors in wind speed are typically as significant as those for temperature and precipitation, and worse than those for surface irradiance [36]. The wind speeds from several reanalyses exhibited biases in the region of 2−4 m/s relative to 33 North American meteorological masts [36], which can translate to the difference between a 20% and 80% capacity factor for a typical wind turbine as output is so sensitive to wind speed.

Statistical methods for bias correction are widely used in climate modelling to bring the frequency distribution of modelled outputs into line with historic observations [37,38]. Several methods of bias correction are employed, ranging in complexity from additive and linear scale factors to quantile mapping [39,40].

Previous studies which use reanalysis for wind energy use simple statistical calibrations which have been derived for small geographic regions. Model calibration takes either the form of a formula-driven reduction in wind speeds [22,23,41−43], an empirical regression to adjust the shape of the power curve [18,21,24] or post-processing to adjust energy output [44].

In the UK and Denmark, observed capacity factors are 26−32% lower than those estimated from uncorrected reanalyses [10,21,42], and for offshore farms in the North Sea they are 12−15% lower

[23,41]. To correct for this, previous studies have reduced wind speeds by 1.2−1.3 m/s [21,41]. Such calibration factors have not been produced for other countries, and the underlying trends in them are not known.

The hourly wind speeds from reanalyses cannot be relied on to predict the long-run average capacity factor for a given location. It therefore follows that these models cannot be used over a wider region (e.g. Europe, China, India) or globally without prior calibration. We provide the first such calibration for Europe.

3. Methods

3.1. The Virtual Wind Farm model

The *Virtual Wind Farm* (VWF) model was used for this study, which is described in Ref. [10] and validated for Great Britain in Refs. [23] and [45]. It takes hourly wind speed data from NASA's MERRA [46] and MERRA-2 [47] reanalyses, chosen for their ease of access, good spatial and temporal resolution and stability over long time-scales [48].

As illustrated in Fig. 3, the VWF model:

(a) acquires wind speeds at 2, 10 and 50 m above ground at each MERRA grid point;
(b) interpolates speeds to the specific geographic coordinates of each wind farm using LOESS regression;
(c) extrapolates speeds to the hub height of the turbines at each site using the logarithm profile law; and then
(d) converts speeds to power outputs using manufacturers' power curves, which are smoothed to represent a farm of several geographically dispersed turbines.

A fuller mathematical description is given in the online supplement §4.

The model applies a smoothing transform to the turbine power curve in stage (d) to account for there being a distribution of wind speeds within any given hour, and between the individual turbines of a geographically dispersed farm. Rather than using a static pre-defined curve such as those given by National Grid [18] or Trade-Wind [24], we use a Gaussian filter that can be applied to the power curve for any model of turbine. The width of this filter (σ) is a function of wind speed (w):

$$\sigma = 0.6 + 0.2\ w\ \text{m/s} \tag{1}$$

to capture the increase in spatial and temporal variation at high wind speeds. The parameters we use in Equation (1) were determined empirically to give the best representation of historic output.
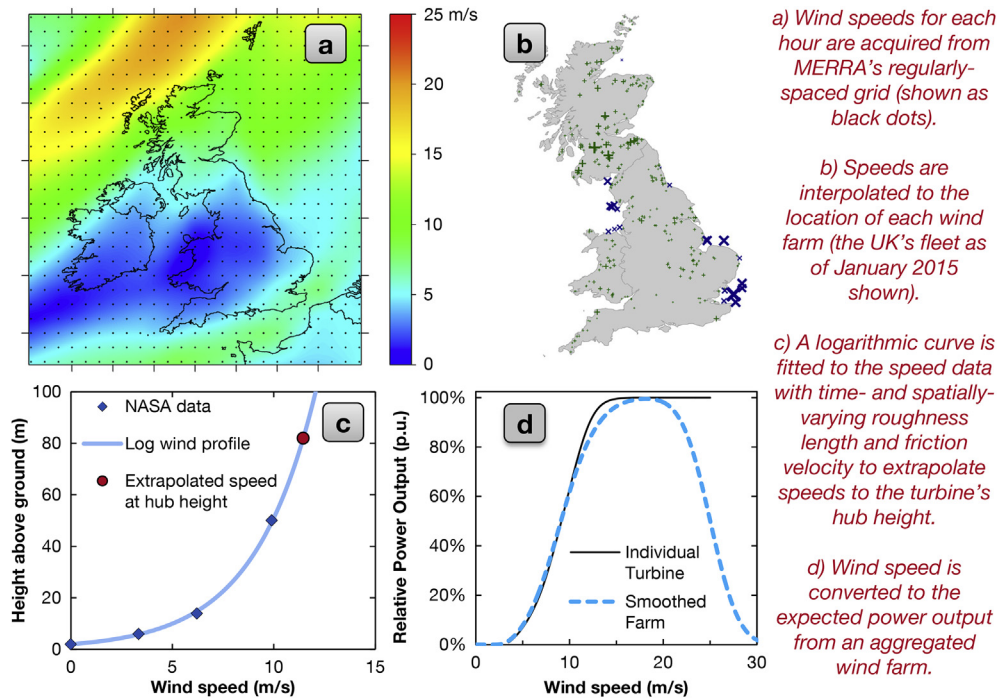
a) Wind speeds for each hour are acquired from MERRA's regularly-spaced grid (shown as black dots).

b) Speeds are interpolated to the location of each wind farm (the UK's fleet as of January 2015 shown).

c) A logarithmic curve is fitted to the speed data with time- and spatially-varying roughness length and friction velocity to extrapolate speeds to the turbine's hub height.

d) Wind speed is converted to the expected power output from an aggregated wind farm.

Fig. 3. Overview of the VWF methodology.

3.2. Wind speed bias correction

In a departure from pure meteorological studies such as [36], we measure bias in terms of the derived power output from wind farms rather than the wind speeds which are directly taken from the reanalysis. This reflects the assumption that power output is the important metric for energy systems modellers. We then apply our corrections to the underlying wind speeds, assuming that the fundamental error lies in the wind speeds rather than the method of converting them to output. Correcting the wind speeds also avoids implausible results, such as scaling up the simulated CFs to over 100%.

The bias, or systematic error (ε_{CF}), for a country is defined as the ratio of observed to simulated capacity factors:

$$\varepsilon_{CF} = \frac{CF_{obs}}{CF_{sim}} \tag{2}$$

We calculate ε_{CF} at the national level due to data availability, as individual farm or turbine outputs are only published for Britain [10], Finland [49], Germany and Denmark [50]. We therefore assume that all farms within a country experience the same bias, except where countries have both onshore and offshore capacity.

Previous work has shown that the value of ε_{CF} is between 9% and 23% higher for offshore than for onshore farms in Britain [23,41]. The spatial coarseness of current reanalyses cannot account for local terrain, trees and buildings which may obstruct the flow of air on land, reducing CF_{obs} and thus ε_{CF}. This is less of an issue out to sea, and so for offshore farms we assume the relevant country's bias should be multiplied by 1.16, the average of the above two findings. We derive separate bias values for onshore and offshore farms (ε_{CF}^{on} and ε_{CF}^{off}) such that their average when weighted by the amount of onshore and offshore capacity (Q^{on} and Q^{off}) yields the overall bias calculated for that country (ε_{CF}):

$$Q^{on}\varepsilon_{CF}^{on} + Q^{off}\varepsilon_{CF}^{off} = \left(Q^{on} + Q^{off}\right)\varepsilon_{CF} \quad \text{where } \varepsilon_{CF}^{off} = 1.16\,\varepsilon_{CF}^{on} \tag{3}$$

For example, MERRA predicts an average CF of 39.2% in Britain compared to the 29.0% that has been observed over the period 2005–14. Therefore $\varepsilon_{CF} = 0.74$ which is notably higher than previous findings for Britain's onshore fleet: 0.69 [42], 0.69 [43] and 0.68 [10]. When averaged over our validation period (2005–2015) 73% of Britain's capacity has been onshore and 27% offshore. When this is applied to the above formula, we find $\varepsilon_{CF}^{on} = 0.70$ and $\varepsilon_{CF}^{off} = 0.82$; which is more consistent with previous findings. We therefore assume that the uncorrected CF should be reduced by 30% for each onshore farm in Britain, and by 18% for each offshore farm.

To correct this bias, we derive a time series of modified wind speeds (w') such that the energy yield (and thus the CF) resulting from these speeds equals the expected value:

$$P(w') = \varepsilon_{CF}P(w') \quad \text{where } \overline{P(w)} = CF_{sim} \text{ and } \overline{P(w')} = CF_{obs} \tag{4}$$

Here, the function P represents the conversion from wind speed to capacity factor (stage d in Fig. 3), and the mean value of P over our validation period yields the values of CF we report. P depends on the power curve for the specific turbine being modelled, and is defined numerically rather than analytically.

Wind speeds are corrected using both a multiplicative factor (α) and a linear offset (β m/s):

$$w' = \alpha w + \beta \tag{5}$$

This correction scheme was previously found to give the best replication of historic outputs in the UK [23,45], as using two parameters allows both the level and the variability of CF to be controlled. One parameter used alone could exaggerate or dampen the variability of wind speeds, inadvertently impacting the diurnal and seasonal trends.

One approach would be to find specific values of α and β for each country that minimise the error in seasonal or hourly CF, or the distribution of CF. For the sake of simplicity, and to reduce the likelihood of over-fitting our model, we instead define a value for α for each country based on the observed bias:

$$\alpha = 0.6\varepsilon_{CF} + 0.2 \tag{6}$$

This simplicity allows our bias correction process to be applied to countries where only the long-run average observed CF is known, and still yields a good compromise for representing the seasonal and hourly variability across the countries of Europe.

The wind speeds at each farm are multiplied by α, and then the model seeks the value of β for each farm that yields the desired CF. These values must be site-specific as they depend on the distribution of wind speeds and the model of turbine installed (which determines the function P). As there is a monotonic relationship between wind speed and power output under normal conditions (i.e. w < 25 m/s), β is found through a simple iterative search.

In every country, this results in a scale factor of α<100% and an offset of $\beta > 0$, fitting with Decker's observation that reanalyses "have a strong tendency to overestimate the variability in the wind" [36]. For example, the average calibration for British wind farms is $w' = 0.66\ w + 2.64$ m/s, the values for other countries are given in the online supplement §2.2. The impossible situation of negative wind speeds is avoided as β is always positive.

3.3. Historic wind output data

Validation requires historic output data from wind farms, which is generally difficult to acquire due to commercial sensitivities. Two kinds of accuracy are important: the ability to predict the level of output at different locations (spatial accuracy) and the timing of that output to hourly or better resolution (temporal accuracy). The ideal data set for testing both simultaneously would consist of high-frequency observations from individual turbines across the whole continent of Europe for several years. Such data does not exist outside the industry, and so we performed two phases of validation: first using monthly and annual data from 23 countries (wide geographic coverage); and then using hourly data from selected countries (high temporal resolution).

3.3.1. Data with wide geographic coverage

Data on national annual energy output and installed capacity were collected for the period 2005 to 2014 from EuroStat [51], ENTSO-E [3], EurObserv'er [52] and BP [53]. EuroStat data for 2014 were not available at the time of writing; and for thirteen countries the ENTSO-E monthly output data were only available from 2010 onwards. All ENTSO-E outputs were reduced by 5% to account for transmission and distribution losses, bringing their averages into line with the other three sources [54]. Further details on these data, their processing, and a comparison between sources are given in the online supplement §1.1.

Energy output was reported as the sum over a year whereas capacity was a snapshot at points in time. When combining these to estimate CF we assume that capacity grows at a constant rate during each year. The average capacity factor (CF) for a year (y) was estimated from total energy output over the course of the year (E_y) and the geometric mean of the installed capacity at the start of the year (P_y) and the start of the following year (P_{y+1}) as in equation (7).

$$CF_{2014} = \frac{E_{2014}}{\sqrt{P_{2014} \times P_{2015}}} = \frac{E_{2014}}{e^{(\ln P_{2014} + \ln P_{2015})/2}} \tag{7}$$

For example, wind power in Finland produced 1124 GW h during 2014, during which capacity grew from 428 to 611 MW [53]; if the year-start or the year-end capacity were used in isolation we would estimate CF = 30.0% or 21.0%, using the year-average capacity estimate of 511 MW we yield CF = 25.1%, which is in keeping with data from previous years.

The long-run average capacity factors for 23 countries are shown in Fig. 4, averaged over the four sources. Their estimates for the Europe-wide average CF, weighted by each country's installed capacity, lie in the range of 22.3–22.6%. Agreement between the sources is generally good for countries with established wind sectors (DE, DK, ES, FR, GB), but begins to break down for the more recent entrants (GR, HU, RO), as shown later in Fig. 6.

3.3.2. Data with high temporal resolution

Nationally aggregated wind output data with hourly or better resolution were acquired from system operators in the 8 countries listed in Table 2, covering 72% of Europe's installed wind capacity. Data files were downloaded from operator's websites where available, and in some cases Flash-based graphs of output were reverse-engineered to recover the underlying data values. These data series were converted to Greenwich Mean Time and aggregated to hourly resolution centred on half past each hour, for compatibility with MERRA and MERRA-2.

With the exception of France, these sources only gave power output with no corresponding data on the amount of capacity being monitored. In some countries, this capacity differed significantly from the values obtained in the previous section, as not all farms are monitored by the system operators. We therefore estimate the evolution of capacity over time using the start and retirement dates for individual farms from Ref. [1], and infer the percentage of this capacity that was monitored by aligning the resulting CFs with those from the four sources used in the previous section. The final column of Table 2 gives our estimate for the monitored capacity as of January 2015, in GW and as a percentage of the total fleet.

3.4. Simulations performed

In this study, we simulate the hourly capacity factors from both the current and planned future wind fleets in Europe, aggregated to national level. We note that capacity factors are influenced by the models of turbine installed, and so they may differ between countries because of location and available wind resource, and also
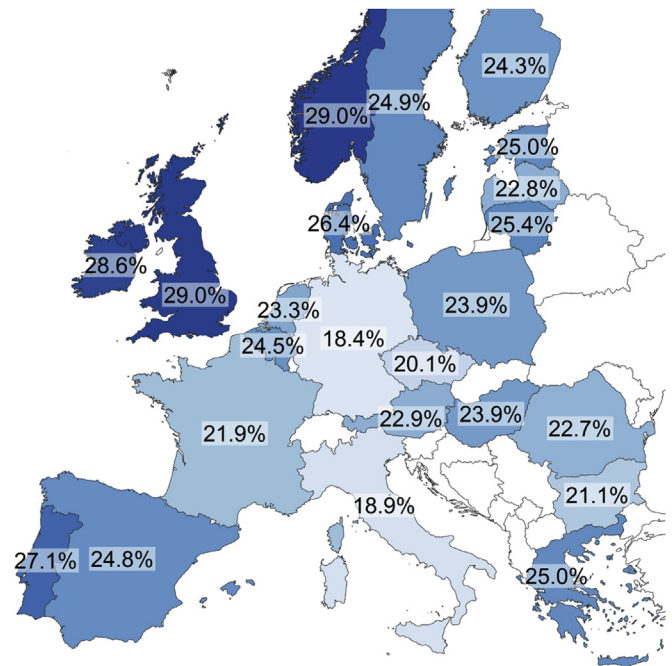

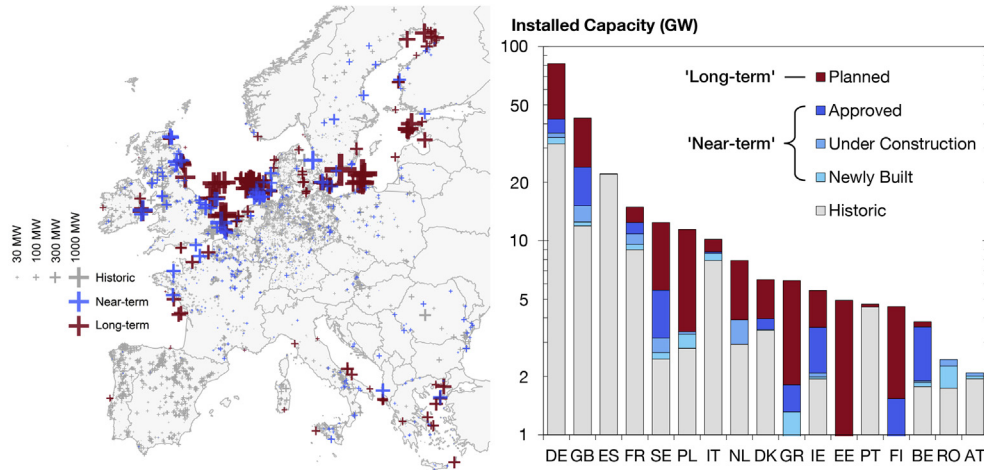**Fig. 4.** Historic national average capacity factors for 23 countries in Europe over the period of 2005–14.

Fig. 5. Summary of the near-term and long-term future European wind fleets considered in this study.
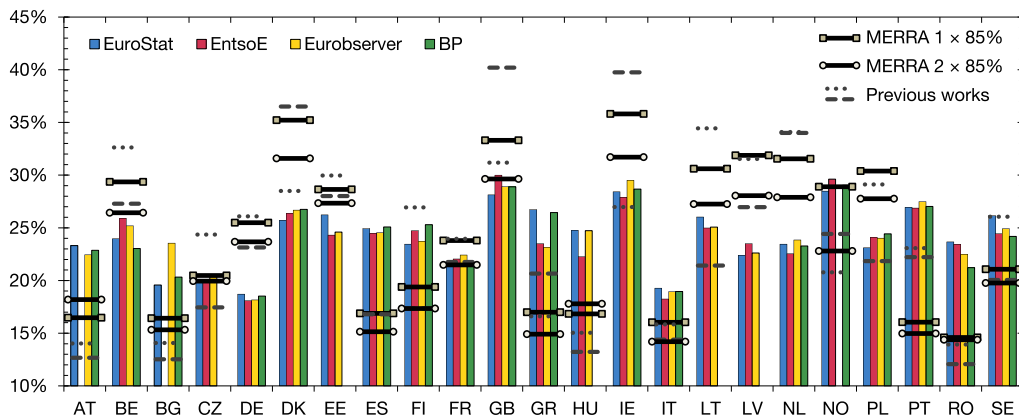

Fig. 6. Long-run average capacity factors (2005–14) given by historic statistics (as bars), compared to reanalysis models without bias-correction (lines).

Table 2
Overview of data sources used for hourly validation.

Country	Source[a]	Time period	Temporal resolution	Estimated capacity covered by data	
Germany	EEX	2010–14	15 min	34.1 GW	(87%)
Spain	Red	2007–14	10 min	21.9 GW	(95%)
UK	NG + Elexon	2009–14	30 min	11.5 GW	(92%)
France	RTE	2011–14	60 min	8.8 GW	(94%)
Italy	Terna	2010–14	60 min	8.4 GW	(100%)
Sweden	SVK	2007–14	60 min	5.4 GW	(100%)
Denmark	Energienet	2000–14	60 min	4.9 GW	(100%)
Ireland	EirGrid	2002–14	15 min	2.2 GW	(96%)

[a] Links to all data sources given in the online supplement §1.2.

because of their relative approaches to policy, markets, subsidy, planning constraints, etc. The latter will not influence the bias correction factors we derive, as we simulating the specific mix of turbine models installed in each country.

3.4.1. The current European wind fleet

For validation and assessing the long-term wind resource we simulated all of the wind farms in Europe over 1 MW with known latitude and longitude data, which amounted to 8736 wind farms and 110 GW of capacity (82% of Europe's total). Simulations were performed using both MERRA and MERRA-2 wind speed data, but after finding no material differences between the two (as reported in the next section) we focussed on the MERRA simulation, as this is

the more widely recognised and understood reanalysis.

Of the 360 models of wind turbine employed in Europe, we collected manufacturers' power curves for the 100 most popular, representing 81% of installed capacity. The remaining turbines were assigned to the most similar model based on the age of the turbine model and the power density (peak output divided by swept area — W/m^2), as in Ref. [10].

The tower height was not known for 62% of farms, and so was estimated using a regression of known heights against the logarithm of turbine capacity and the date of installation. The start date was not known for 16% of farms, and so was inferred from other farms in the same country with turbines of the same capacity. Having to estimate input parameters for such large portions of the

installed fleet is a potential source of error, so we note that the model validation may be improved with better knowledge of the turbine population.

Each farm was simulated over the 20-year period from 1995 to 2014, to give a long-term view on the average statistics the current fleet would give. MERRA data is available going back to 1979; however, we choose not to use the earlier years as fewer meteorological observations were available to assimilate into the reanalysis, potentially compromising its accuracy.

When validating for the period 2005–14, the output from each farm was zeroed at times when it did not exist (before its start and after its retirement date), so that the time-evolution of each country's wind fleet was correctly represented. Later on, Section 4.2 highlights the importance of simulating the evolving fleet of farms (i.e. only those that were available in 2005 when estimating 2005 capacity factors) as opposed to a static snapshot. The simulation was performed on a standard workstation and required around 2000 gigabytes of storage for input data and 1200 CPU-hours (at 3.4 GHz) to complete.

The simulation results were aggregated to country level, and are available to download from www.renewables.ninja.

3.4.2. The future wind fleet

In addition to modelling Europe's current fleet, we consider two possible snapshots of Europe's future wind fleet. For the 'near-term' future, we group together farms that have been built during 2015 (312 farms, 6.5 GW), those that are under construction as of December 2015 (227 farms, 8.8 GW) and those which have obtained legal approval (75 farms, 29.2 GW). For the 'long-term' future, we add farms that are at various pre-approval stages in planning pipeline (237 farms, 93.4 GW). In total, the near-term fleet contains 150 GW and the long-term fleet 248 GW of capacity, which are broken down by country in Table S1 of the online supplement.

We do not attempt to estimate when these farms will come online, and so do not attribute a date to these snapshots. We also assume that all of the current fleet remains online, with no attempt made to estimate which existing farms may retire or be repowered in the coming years. Note also that we ignore any effects of ageing (as explored in Ref. [10]), and assume that the average age of wind turbines in each country remains constant over time.

The list of speculative farms were taken from Ref. [1], and while it may not be completely exhaustive, it ought to give a representative view of the evolution in Europe's capacity, based on developers' current intentions.

The 'near-term' fleet represents a 36% increase in capacity from the 'current' fleet and the 'long-term' fleet represents a further 65% increase. Fig. 5 shows the geographic distribution of these new farms against the backdrop of Europe's current capacity, highlighting the increasing prominence of the North and Baltic Seas. The move towards offshore farms is evident: 7.5% of current capacity (pre-2015) is offshore, rising to 26% of recently built capacity (during 2015), and 38% of the capacity currently under construction. All of the capacity that is approved or planned is offshore, although this is limited by data availability.

Understandably, a significant portion of these future farms were missing meta-data. Using the same processes as in the previous section, tower heights were inferred for 86% of farms. Turbine models also had to be inferred for 54% of farms, based on the turbine capacity and manufacturer (if known). For each turbine capacity, missing models were drawn randomly from the population of known models of that capacity. Power curves were not available for many of the largest prototype turbines, such as the Siemens SWT-7.0-154, so these were matched to known curves based on power density (as above).

These fleets were simulated using bias-corrected wind speeds,

which were derived from the current wind fleet. It is reasonable to assume that MERRA's performance and bias when simulating future onshore farms should be comparable to the current onshore fleet, as their locations are similar. The same cannot be said for the offshore farms in new areas, deep into the North and Baltic Seas, as validation against metered output data is not possible. However [41], and [10] show that MERRA replicates offshore wind speeds recorded on oil rigs and buoys with greater accuracy than onshore speeds. Sharp's review of 16 reanalysis studies reports similar or slightly better correlations, root mean square (RMS) errors and biases for offshore than for onshore studies [20].

4. Validation

This section first looks at the bias present in the long-run estimates of capacity factors from MERRA and MERRA-2. We derive a set of bias correction factors for each EU country, and then compare the estimated power outputs from the corrected MERRA dataset to historic data. MERRA-2 is not considered further in this paper, as we find only minimal differences between the MERRA and MERRA-2 wind speeds, other than those in MERRA-2 being systematically lower. For completeness, the validation for MERRA-2 is presented in the online supplement §2.3 onwards.

4.1. Bias in national capacity factors

Fig. 6 compares the national average capacity factors estimated from the uncalibrated MERRA and MERRA-2 data with two previous reanalysis studies (Huber [15], and McKenna [17]) and historic statistics. The MERRA and MERRA-2 capacity factors were reduced to 85% to bring them in line with the two previous works. This gave EU-average capacity factors (weighted by national installed capacity) of 22.4% from MERRA and 21.0% from MERRA-2, compared with 22.5% [15] and 23.1% [17], and 22.4% from historical sources.

We find that reanalysis shows little spatial correlation with historic data across Europe, with R^2 values across countries of just 0.19 for MERRA and 0.15 for MERRA-2, compared with 0.32 in Ref. [15] and 0.08 in Ref. [17]. The RMS errors range from 6.4 to 7.2% points. The fact that this is evident in previous works employing MERRA and ERA-Interim suggests that the problem lies in the underlying reanalysis wind speed data rather than the specific methodology used in the VWF model. This sadly casts doubt on the many previous studies that have used reanalysis to simulate renewable energy resources without prior calibration.

Historical sources suggest that the national CFs across Europe range from 18.4% up to 29.0%. The three reanalysis models exaggerate this spatial variation, ranging from minima of 12.1–13.9% up to maxima of 36.7–40.2%.

We derive bias correction factors for the EU countries from the unweighted average of the four historic sources, taking the broadest view of the available data with no judgement on which is best. These corrections are presented in Fig. 7, showing the multiplier that must be applied to the raw MERRA (left) and MERRA-2 (right) outputs to yield the historic average capacity factors.

While the national correction factors are relatively stable over time, across space they have a standard deviation of ±32%, signifying large heterogeneous bias over a relatively small portion of the world's surface (~2%). The correction factors for MERRA fit reasonably with previous findings in the literature: 68–74% for the UK and Denmark [10,21,42,43]. Critically we find that these previously-known scale factors cannot be applied elsewhere across continental Europe. The correction factors for MERRA-2 are the first we know of, and have a similar structure to those for MERRA, albeit around 10% higher.

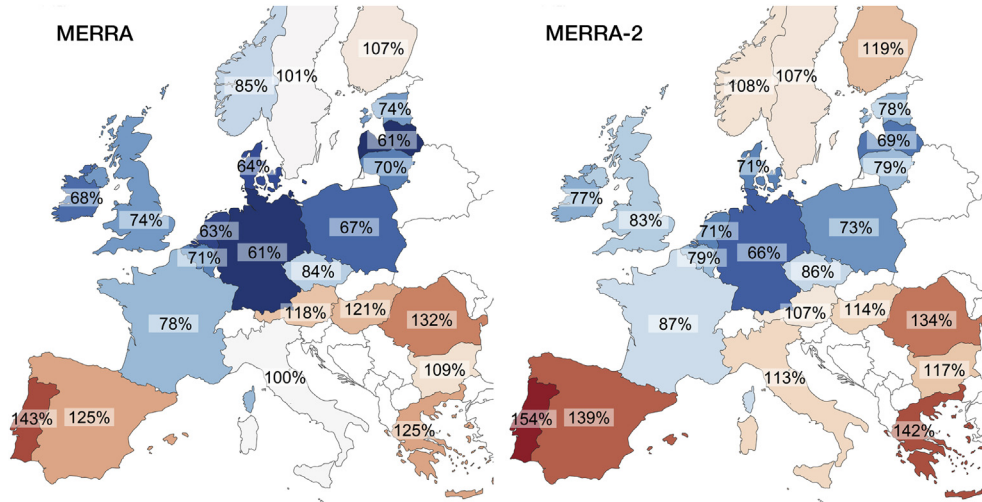Two approximate trends are evident: a north/south divide

Fig. 7. Bias correction factors (ε_{CF}) for MERRA and MERRA-2 across 23 European countries.

separated by the Pyrenees and Alps; and an inland/coastal divide between those countries which border the North, Baltic and Celtic seas, and those which do not. We make tentative suggestions for the causes for this phenomenon in our conclusions, and signpost this as an interesting avenue for further research.

4.2. Inter-annual variability

In keeping with previous studies of other weather variables, we find that although MERRA contains significant spatial biases, the inter-annual variability within each country is well replicated. The historical sources show a marked variation from one another, making a quantitative assessment of MERRA's accuracy difficult. Fig. 8 shows Denmark as an example, the thin coloured lines represent the four data sources we use, which at times follow each other precisely (implying the shared use of a common source) but at other times differ substantially (ENTSO-E in particular).

When matching historic outputs, it is necessary to simulate only the farms which existed at any given time, as opposed to simulating
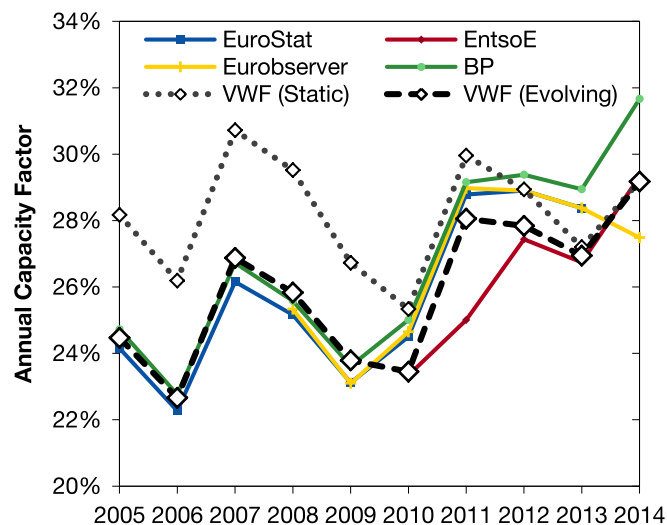

Fig. 8. Annual capacity factors in Denmark over ten years, comparing historic sources (thin lines) with the VWF Model when simulating the wind farms that existed during each year (dashed line) and the static 2015 fleet (dotted line).

a static snapshot of installed capacity over the whole period. To illustrate this, the dotted line in Fig. 8 (VWF Static) shows the simulated output of the current 2015 fleet of farms, which would have produced a 28.2% capacity factor during 2005 if they had existed back then. In comparison, the dashed line shows represents the "fleet of the day": the farms that were actually operating in 2005 would have produced a 24.5% capacity factor during 2005, which is much closer to reality. This effect is strongest in Denmark and the UK due to the strong shift towards offshore wind farms, but it is visible to a smaller extent in other countries, as shown later in Section 5.3.

Fig. 9 shows the year-to-year variation in capacity factors across Europe. The shaded areas summarise the distribution of percentiles across all countries, and the thin lines trace the three countries with the greatest installed wind capacity. The error bars on the historic data series show the standard deviation across the four historic data sources. The capacity-weighted mean is consistently around 2% lower than the simple average across countries, pulled down by the large amount of unproductive capacity in Germany, and the relatively limited capacity in the most productive countries (Ireland, Norway, Portugal).

While the simulation is not perfect, it appears to capture the bulk trends in Europe's wind resources; for example the spread between countries. The anti-correlation between countries is particularly notable during 2010 – a particularly low wind year for most of Northern Europe, but one of the better years for Spain.

Fig. 10 shows that the simulation error (its deviation from historic annual CFs) is comparable to the uncertainty in those historic CFs. Error is taken to be the RMS deviation between the simulated CFs and the mean of the historic sources; uncertainty is measured by the standard deviation across those sources. The unweighted average error across all countries was 1.37%, only slightly higher than the average uncertainty (1.35%). The interpretation is that for a typical country in Europe, the simulated annual capacity factors are on average 1.37% away from our best estimate of the true value; but our uncertainty on that true value is ±1.35%.

4.3. Monthly validation

We compared MERRA to ENTSO-E data for the 13 countries for which installed capacity data could be reliably discerned. The correlation between simulated and actually monthly capacity factors, averaged across all countries was $R^2 = 0.91$. In four countries the
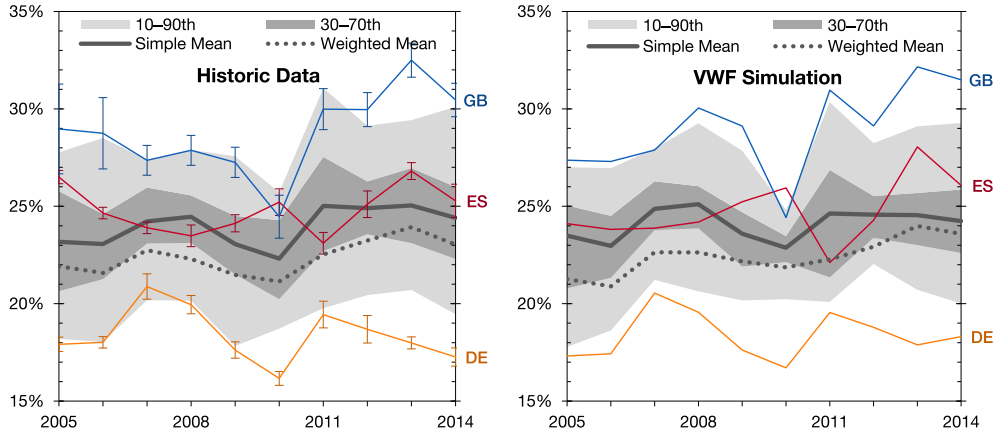
Fig. 9. Annual capacity factors across Europe, comparing historic data averaged over four sources (left) with our simulation (right).
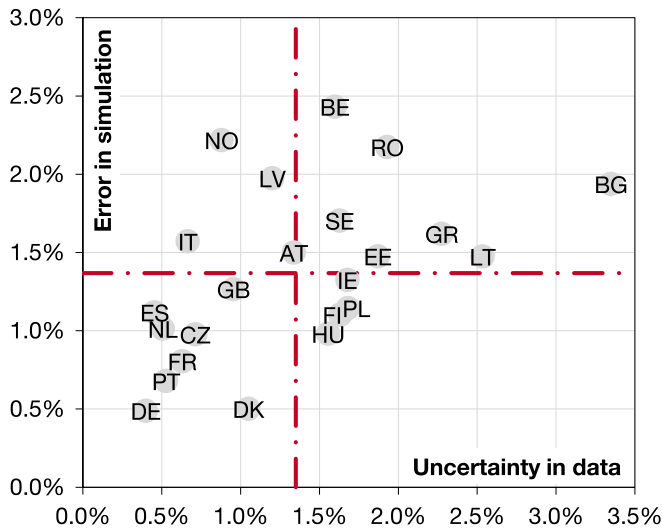

Fig. 10. RMS error when simulating annual capacity factors from 2005 to 14. Dashed lines show the mean across all countries.
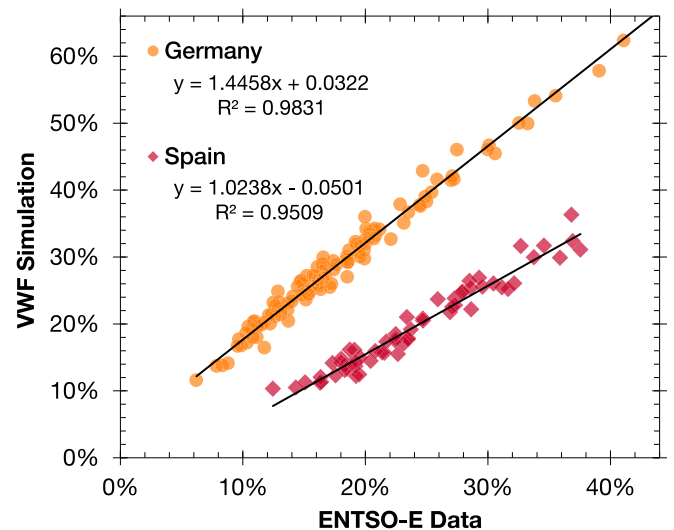

Fig. 11. Monthly validation for two countries (before correction) showing very different bias, but similar accuracy.

correlation was over 0.95, in a further five it was over 0.90; the worst correlations were 0.75 in Greece and 0.78 in Portugal. Further details, including monthly plots for each country are given in the online supplement §2.3.

Fig. 11 shows the correlation for two countries at the extremes of our calibration: Germany and Spain. Three features are of note: the good representation that the calibrated VWF model offers; the significant difference between the scale factors required for calibration; and the different nature of the two calibrations. In Germany, the regression yields a negative offset and a steep gradient, implying that MERRA over-estimates the variability in power output (and thus presumably in wind speeds). In Spain, the opposite is true, and so MERRA under-estimates this variability.

The seasonal profile of Europe's wind output is shown in Fig. 12. Based on the ENTSO-E data, the EU-wide average capacity factor ranges from 30.3± 5.2% in winter down to 17.0± 2.4% in summer (simple averages), or from 28.6% in winter down to 16.0% in summer (capacity-weighted averages). At the extremes, Britain ranges from 39.2% to 19.1%, and Germany from 26.0% to 12.3% across the seasons. The calibrated VWF model (Fig. 12 right) replicates these statistics to within ±0.7%.

Warmer Mediterranean countries have lower seasonal variation than colder northern European countries. This is seen in Fig. 12

comparing Spain and Britain. The ratio of winter to summer CF averages 1.52:1 across Spain, Italy, Portugal and Greece; compared with 2.04:1 across Britain, Germany, Denmark and Norway. This could benefit the integration of wind into electricity systems, as it provides a better match to the seasonality of demand for electricity in cold and temperate climates.

4.4. Hourly validation

Fig. 13 presents a selection of comparisons between the VWF model's estimation of the aggregate wind output in Germany to historic data from EEX over the period 2010−14.

The root mean square error (RMSE) between simulated and actual capacity factors is 3.11%; implying that an estimated CF of 50% for a given hour can be treated as accurate to within 46.89−53.11%. Germany's wind farms produce less than 6.1% CF for a quarter of the time and more than 24.9% CF for a quarter of the time, the model is able to replicate this distribution of output to within ±0.4% except at the extreme upper tail (the 95th percentile is under-estimated by 0.9%). The change in CF from hour to hour is normally distributed with a standard deviation of 1.8%. The model under-estimates this slightly, as it does for the 4-hourly power swings.
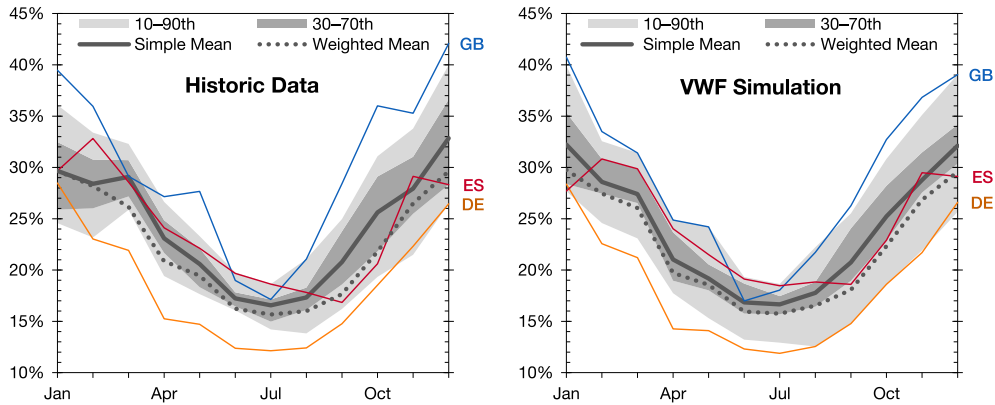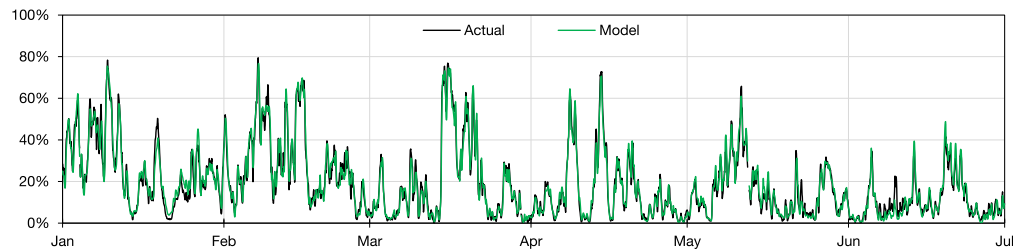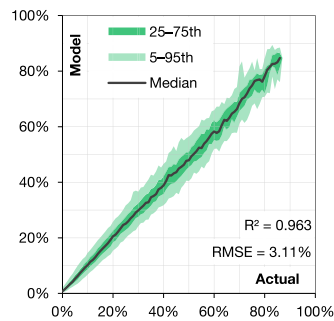
Fig. 12. Monthly average capacity factors across Europe, comparing ENTSO-E data (left) with simulation (right).
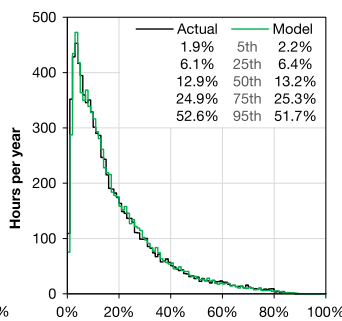
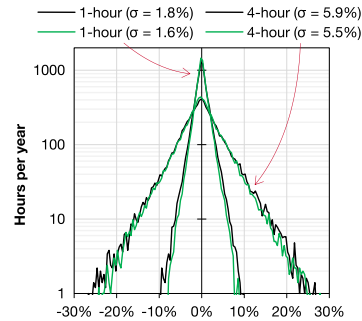**Fig. 13.** Comparison of the simulated and actual hourly capacity factors in Germany, showing: (a) A short segment of the 5-year time series; (b) their correlation, where shaded areas represent the distribution of percentiles around the median; (c) the histogram of capacity factors, showing the number of hours per year that CF lies within 1% width bins, with the inset table giving the 5th to 95th percentile of output; and (d) the histogram of the rate of change of output (or power-swings), showing the number of hours per year that ΔCF lies in 0.2% width bins over 1-h and 4-h windows, with the inset values giving the standard deviation of each distribution.

Corresponding figures for the other seven countries are given in the online supplement §2.4. Table 3 summarises their main statistics:

- the RMS error and correlation between actual and simulated hourly CFs;
- the shape of the CF distribution, represented by the RMS error on the percentiles, $RMS(P_n^{act} - P_n^{sim})$ for $n = 0..100$, where P_n denotes the nth percentile of the distribution;
- the error on the standard deviation of the 1-h power swing (ΔCF).

The performance of the VWF model is excellent across north-west Europe, with correlations to hourly CFs of above 0.90. Performance is best in Germany, possibly because of the large number of farms and their wide geographic dispersion, meaning that no one anomaly will have significant impact.

The model's performance is markedly worse in the Mediterranean countries, which appears to be due in part to a poor representation of the distribution of CF. The simple bias correction process we use is unable to capture both the seasonal and hourly variability simultaneously. Adjusting the α parameter in equations (5) and (6) to represent the seasonal trend means that hourly variability is under-estimated so CF never falls below 4%. Adjusting them to match the hourly variability means the seasonal variability becomes exaggerated.

This distribution of percentiles (their histograms) are generally well represented, with the error on the value of any given percentile being less than ±1.25% points in five of the eight countries. In Spain, Italy and Denmark the errors were substantially higher, as the VWF model under-estimated the low tails and over-estimated the high tails of the distributions (see Supplementary Figures S7, S10 and S12).

Finally, we see that the MERRA systematically under-estimates the rate of change of wind speeds, as the width of ΔCF is 0.16%–1.13% smaller than in reality. The error is generally greater in small

Table 3
Summary of performance from hourly validation. Error values are given in absolute CF.

	RMS error	R^2	RMS error on percentiles	Error on σÄCF (1-h)
Germany	3.11%	0.963	0.42%	−0.24%
Spain	6.07%	0.841	2.16%	−0.16%
Britain	4.68%	0.935	1.24%	−0.61%
France	4.39%	0.913	0.63%	−0.24%
Italy	7.44%	0.761	3.19%	−0.34%
Sweden	5.66%	0.906	0.93%	−0.46%
Denmark	6.75%	0.912	3.19%	−0.81%
Ireland	6.65%	0.905	1.22%	−1.13%

countries, which have more dramatically varying wind conditions due to their size relative to weather fronts, and thus the largest values of ΔCF.

Fig. 14 shows the combined seasonal and diurnal trends in Germany. Each line shows the mean CF for each hour of the day across all days in a given season. Winter sees high CFs in Germany, which are reasonably constant throughout the day. Summer CFs are less than half those of winter, and show a strong day/night cycle due to heating from the sun.

The VWF model (shown by the dotted lines) is able to replicate these patterns in Germany, as well as Spain, Sweden and Ireland, as shown in the online supplement §2.5. The impact of solar heating in the summer and shoulder seasons appears to be slightly over-estimated in France, slightly under-estimated in Britain, and dramatically under-estimated in Denmark and Italy. In Denmark, average summer CFs range from 15% at night to 24% in mid-afternoon, whereas the model predicts a range from 16% to just 19%. The situation in Italy is similar. Correcting for these problems would improve both the RMS error and the error on percentiles, which are reasonably poor in Denmark, in contrast to other northern-EU countries.

5. Results from an EU-Wide analysis

After validating the bias-corrected VWF model, we used it to simulate the current European fleet of wind farms (pictured in Fig. 1) operating with weather data from the last 20 years (1995–2014). This is the first international assessment to be made using a rigorously validated model, and the underlying raw data
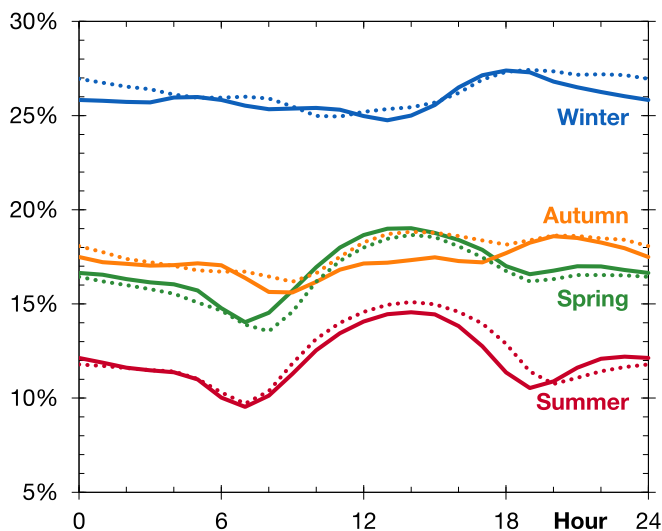
(hourly national capacity factors) are available to download from www.renewables.ninja.

5.1. Long-run average capacity factors

The current European wind fleet has an estimated long-run average CF of 24.2%. This varies from year to year by ±1.0%, which is less pronounced than for individual countries due to the diversity benefit of geographical smoothing. For comparison, the current British wind fleet averages 32.4± 2.1%, and Germany's fleet averages 19.6± 1.5%, as shown in Fig. 15.

The poor wind year experienced across northern Europe is clearly visible in Fig. 15; however these low wind speeds were not felt across the whole continent, and 2010 was in fact one of the better years for Spain. This variability is purely due to weather effects, as the simulation of a static fleet of farms strips out the influence of technological improvement increasing CF over time (as highlighted earlier in Fig. 8).

Fig. 15 also shows the average CFs estimated for the current offshore fleet in each country. For the countries with over 1 GW of installed capacity (Britain, Germany and Denmark) offshore CFs average 35.8%, 35.7% and 33.0% respectively. In other countries they vary more widely as there are so few farms installed to date. The British result fits well with the current average CF of Round 1 and 2 offshore farms (~36%) [55], lending confidence to our method of bias correcting for onshore and offshore farms separately.

5.2. Distributions and correlations

This smoothing effect is evident in the frequency distribution of CFs shown in Fig. 16. Individual countries have a wide distribution of CF, whereas the aggregate European distribution is much narrower, spending more time in the range of 15−35%.

The plot also quantifies the tails of each distribution: on a typical weather year, hourly CFs would fall below and rise above these values for 24 h of the year. The European average CF can be expected to remain between 7 and 63% for 363 days of the year, making spatially-aggregated wind power a more reliable source of generation than it can be in any single country.

Germany sits at the geographic centre of Europe and hosts the largest wind capacity, therefore the correlation between German and European output is strongest at $R^2 = 0.66$. In contrast, countries on the edges of Europe have almost no correlation to the continent-wide output: for Greece, Romania and Finland R^2 lies between 0.01 and 0.06. The results for all countries are presented in the online supplement §3.1.

Adding more wind capacity in these peripheral countries would help to further diversify the continent's profile of output, although this would only translate into a practical benefit for system operation if there existed sufficient transmission capacity to move this power between countries.

5.3. Measuring technical improvement

The productivity of national wind fleets has increased gradually over the last ten years, as seen earlier in Fig. 8 (dashed line) and Fig. 9 (dotted line). This is partly due to technical improvement: the move to taller towers and bigger blades. In the last 10 years, Europe's average wind turbine has grown 15 m−66 m in height [1], granting access to 7% higher wind speeds on average. Some countries, particularly Denmark and Britain, have seen significant gains from moving to better locations, with more capacity moving offshore and to sites further from shore.

The Europe-wide capacity factor is also raised by the shift of capacity away from Germany (the least productive nation) to
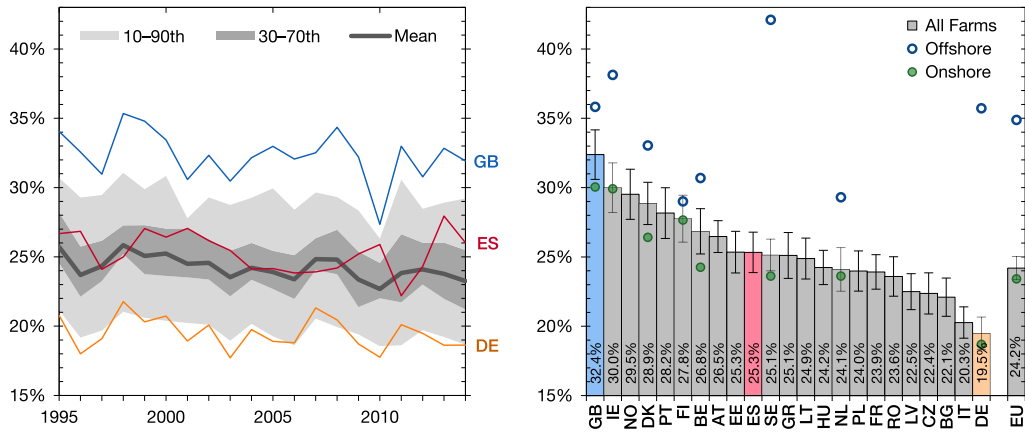

Fig. 14. Seasonal and diurnal trends in the German wind output, comparing actual (solid lines) and simulated (dotted).

Fig. 15. Estimated capacity factors for the present-day EU fleet of wind farms with the last twenty years of weather data, showing (left) the variability between years, and (right) the long-run averages for each country. Error bars signify the standard deviation in annual CF over the twenty years, and for the relevant countries circles denote the averages for onshore and offshore farms.
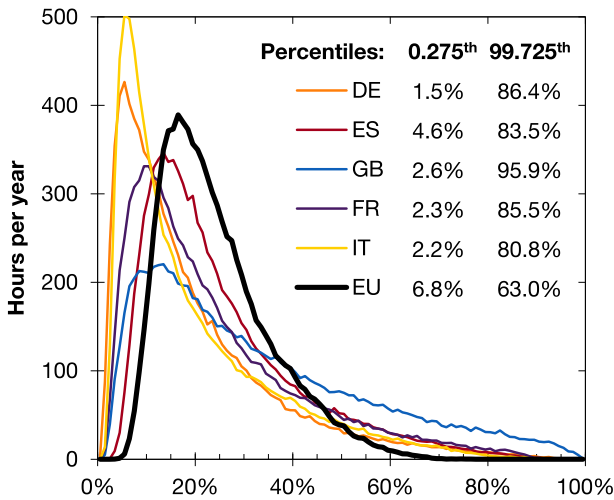

Fig. 16. The frequency distribution of capacity factors in five countries and Europe as a whole.

windier countries such as Britain and Ireland. Germany was a first-mover, holding a 43% share of Europe's installed capacity in 2005 but only 29% by 2014. In this time, Britain's share has grown from

3% to 11%. Given that British wind farms are 65% more productive (from Fig. 15), this shift alone has resulted in the European average capacity factor rising by 1.3% points.

Simulating Europe's wind fleet as it evolved over the last twenty years yields a long-run average CF of 22.4%, which matches with the historic statistics given in §3.3. When simulating instead Europe's current wind fleet as of January 2015, this increases to 24.2% (as in Fig. 15), implying that when weather conditions are held constant, the modern fleet can operate at 2% points higher CF than seen over the last two decades.

In Fig. 17, we estimate how the long-run average capacity factor has changed in six countries as their wind fleets have evolved, and may continue to evolve in the future. Spain and Sweden could not be included due to insufficient data on farm construction dates. In the left half of the left-hand plot, we isolated the farms which were operating at the start of each year from 2005 to 2015, and then simulated their long-run average CF over the period 1995–2015. In the right half, we simulate the 'near-term' future fleet, which adds to the current fleet those farms built during 2015, under construction and with legal approval; and 'long-term' future fleet, which adds all farms in the planning pipeline. The right-hand plot summarises the future capacity factors for all of the countries we simulated.

Looking at the historic assessment, Britain and Denmark have seen the largest rises in productivity due to their growing offshore
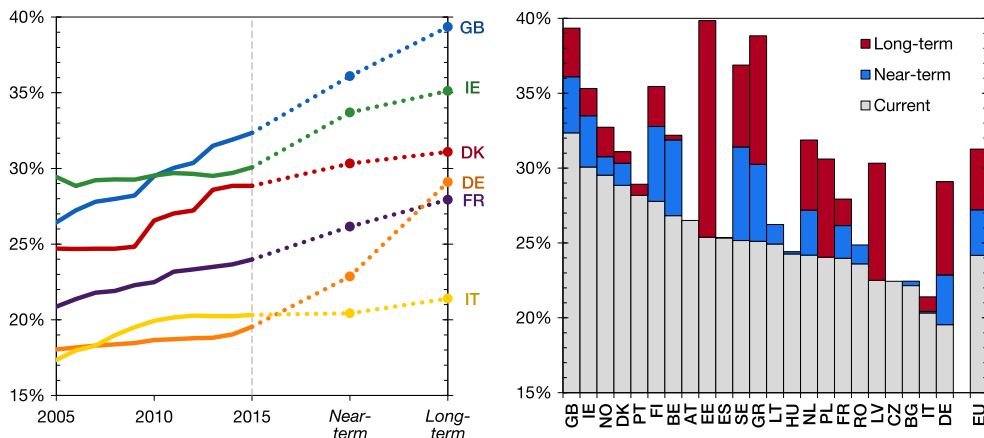

Fig. 17. The increase in long-run average capacity factors due to improving technology and locations.

capacity. Britain's current fleet operates with a 32.4% CF, but if the country had stopped building new farms at the start of 2005, this would be just 26.4%. Over the last ten years, Britain and Denmark's CF have increased by 19% and 17% (in relative terms), France and Italy's have increased 12 and 13%, while Germany and Ireland's have increased by only 7 and 4%.

Looking forwards, the 'near-term' and 'long-term' future fleets have dramatically higher CFs than seen by today's fleet. In most countries future CFs continue following their historic trend, implying that the planning pipeline in most countries is a natural evolution from the new capacity that was developed over the last decade. For example, the capacity factor for Britain's near-term farms averages 39.7%, and the long-term farms average 43.4%. This means Britain's CF continues rising from 32.4% today (with 11.9 GW installed) to 36.1% in the near-term (23.4 GW) and 39.4% in the long-term (42.3 GW). In contrast, Denmark's CF remains relatively flat, increasing from 28.9% to only 31.1% in the long-term while capacity rises from 3.4 GW to 6.3 GW).

This agrees with the findings of Drew et al. who estimated a one-fifth increase in British capacity factors circa 2025 (to 39.7%) by simulating the "Round 3" offshore sites (~26 GW of capacity) [19]. They disagree with the findings of Andresen et al. however, who estimate a one-third increase in Danish capacity factors circa 2025 (to 37.3%) by simulating ~6.5 GW of capacity, with both new and all existing farms replaced by new, tall turbines [21].

One country dramatically defies the historic trend: the strong shift towards new offshore capacity in Germany is expected to raise its CF from 19.5% today (with 31.5 GW) to 22.9% in the near-term (42.3 GW) and 29.1% in the long-term (81.6 GW). It of course remains to be seen whether all of this offshore capacity will be developed, but the prize of doing so will be a German wind fleet that is 50% more productive than today's.

6. Conclusions

The pattern of power output from national fleets of wind farms has risen to vital importance for both the research and operation of power systems. The complexity of the weather makes wind output challenging to synthesise accurately, and commercial confidentiality means historic data is often limited. We present a model for simulating the hourly power output from wind farms located anywhere in the world and validate it across 23 countries in Europe.

Both the MERRA and MERRA-2 reanalyses contain a systematic bias in wind speeds with a strong spatial gradient across Europe. A naïve simulation using reanalysis without correction would therefore yield significant errors in average wind capacity factors; ranging from a 30% under-estimate in Portugal and Romania to a 60% over-estimate in Germany and Denmark. This significant bias appears to be present in previous works which have used reanalysis to assess the potential for renewables at continental and global scale, and sadly casts doubt on the results they reach.

We develop a simple linear equation for correcting this bias, which depends on only one input: the ratio of historic to simulated capacity factors in that location. Once this correction is applied, the simulated hourly capacity factors are able to replicate historic data with exceptional accuracy. For example, the simulated hourly outputs for Germany and other northwest European countries have a correlation of above 0.95 to historic data, with accurate replication of seasonal and diurnal trends, the overall distribution, and the power swings.

There is little to distinguish MERRA from MERRA-2 in terms of wind speeds. MERRA-2 yields capacity factors that are around 10% lower across Europe, but the spatial heterogeneity in bias is slightly greater than for MERRA. Compared to MERRA, the hourly capacity factors from MERRA-2 are slightly better correlated in some

countries and worse in others; but the difference in results is immaterial at less than ±1%, so we currently see no strong reason to favour MERRA-2.

We use the Virtual Wind Farm (VWF) model to estimate Europe's national aggregated wind output over the last twenty years using MERRA. The EU-wide output from the current fleet of turbines is found to be more stable than individual countries' output due to geographic smoothing, and higher than historic levels due to the improvement in turbine siting and technology. The Europe-wide capacity factor has increased by a tenth over the last ten years, while in individual countries it has risen by up to a fifth due to the move offshore.

We also simulate two snapshots of Europe's future wind fleet, based on those farms under construction, with legal approval and in early-stage planning. These future fleets continue the trend of historic technical improvement and exhibit much higher capacity factors than seen today. The strongest increases are seen in Britain and Germany due to large developments in the North Sea. Capacity factors could reach as high as 40.9% and 32.8% in these two countries, if all of the currently planned farms end up being developed.

Such an increase in the productivity of wind farms would have far-reaching consequences: increasing the economic viability of wind power, reducing or eliminating the need for subsidies, decreasing the energy sales from conventional generators and thus their revenues, and increasing the contribution that wind makes towards carbon mitigation.

6.1. Why is bias correction required?

Reanalysis is no substitute for detailed micro-scale wind resource modelling. It is spatially coarse and cannot represent the detailed orography in mountainous regions. So, it is perhaps not surprising that some form of correction is required.

The impact of orography on wind speeds in the mountainous regions of Spain, Italy, Greece and Scandinavia can exceed ±50%, whereas in the flatter plains of France, northern Germany, Poland and England it is ±5% [56]. Reanalyses such as MERRA cannot represent these speed-up effects as each grid cell is treated as a flat plane. If we assume that farm developers are rational and install turbines in the windiest locations, we should expect reanalyses to under-estimate wind speeds in southern Europe and Scandinavia relative to northern Europe − which is the broad trend observed in Fig. 7.

The spatial coarseness of reanalyses relative to micro-scale models may also be a factor. When the atmosphere is discretised into 50 × 50 km cells it is not possible to accurately represent the underlying atmospheric processes which generate wind speeds. For example, if a reanalysis incorrectly simulates the general trend in storm tracks or over-estimates the strength of the westerly winds moving in from the Atlantic over the British Isles and into north-west Europe, it would assume higher than observed wind speeds in these areas. It may be found that other reanalyses listed in Table 1 exhibit lower − or different − bias, and so the use of ensemble datasets could prove beneficial. However, as we note in Section 2.1, previous studies which use ERA-Interim appear to display a geographical trend in bias.

A further consideration is that reanalysis models are calibrated to satellite observations of air pressure and ground based measurements of wind speed. ERA-Interim features variational bias corrections for irradiance, but not for pressure (from which wind speeds are derived) [57]. The ground based measurements come from airports, military bases and other sites of weather stations, which are not representative of wind farms. If wind farm operators could agree to share their SCADA data with NASA, ECMWF and other agencies then these observations, which are not dissimilar to

tall-tower wind speeds, could be added to the assimilation model, giving a better representation of the metrics that the energy community are interested in.

6.2. Future developments

A fundamental limitation of reanalysis is the need to calibrate to ensure that capacity factors match the level and distribution seen in historic data. This correction scheme we propose could be extended by using bespoke parameters for each individual farm, or by varying parameters by season, hour of day, and so forth. This would improve the model's ability to replicate historic statistics, but would run the risk of over-fitting, and reduce the ease with which the model can be transferred to other regions of the world.

Ultimately, this method is limited by the availability of historic data. Future effort should be directed at developing physics-based methods for bias correction. If a suitable correction scheme could be found, this could ultimately be incorporated into the host reanalysis, improving the accuracy of wind speed estimation at source.

6.3. Open access

In an effort to support future research in this area, we make the core dataset from this paper available online. This comprises twenty-year time series of the estimated hourly output from Europe's wind farms aggregated at national level — some $175,320 \times 23$ observations. This is available for the current, near-term and long-term future fleets that we simulate; allowing the evolution of Europe's wind fleet to be represented in other energy systems models.

A version of the Virtual Wind Farm model has also been made open-access and is available to use via a web interface at www.renewables.ninja, which is described further in Ref. [58]. We hope that making this data and model available to the community will help to overcome barriers to research in this area, avoid duplicated efforts and enable new questions to be answered.

Acknowledgements

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.energy.2016.08.068.

References

[1] Pierrot M. The wind power database. 2015. URL: http://www.thewindpower.net.
[2] REN21. Renewables global status report. 2014. URL: http://www.ren21.net/gsr.
[3] ENTSO-E. Data portal > consumption data (hourly load values) and production data (detailed monthly production). 2015. URL: https://www.entsoe.eu/data/data-portal/consumption/.
[4] Sensfuß F, Ragwitz M, Genoese M. The merit-order effect: a detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. Energy Policy 2008;36(8):3086–94.
[5] Pöyry Energy Consulting. Implications of intermittency: Oxford. 2009. URL: http://www.uwig.org/ImpactofIntermittency.pdf.
[6] Green R, Vasilakos N. Storing wind for a rainy day: what kind of electricity does Denmark export? Energy J 2012;33(3):1–22.
[7] UCTE. Final report: system disturbance on 4 November 2006. 2006. URL: http://tinyurl.com/6qtxq8v.
[8] Sinden G. Characteristics of the UK wind resource: long-term patterns and relationship to electricity demand. Energy Policy 2007;35(1):112–27.
[9] Hughes G. The performance of wind farms in the United Kingdom and Denmark. London: Renewable Energy Foundation; 2012. URL: http://tinyurl.com/cn5qnqg.
[10] Staffell I, Green R. How does wind farm performance decline with age? Renew Energy 2014;66:775–86.
[11] MacKay D. On the performance of wind farms in the United Kingdom. 2013. URL: http://tinyurl.com/o4cj6ut.
[12] Staffell I, Green R. High quality wind data for energy systems models. In: Presented at energy systems conference. London; 2014. URL: http://tinyurl.com/hq-wind-data.
[13] Lu X, McElroy MB, Kiviluoma J. Global potential for wind-generated electricity. Proc Natl Acad Sci 2009;106(27):10933–8.
[14] Cosseron A, Gunturu UB, Schlosser CA. Characterization of the wind power resource in Europe and its intermittency. Energy Procedia 2013;40(0):58–66.
[15] Huber M, Dimkova D, Hamacher T. Integration of wind and solar power in Europe: assessment of flexibility requirements. Energy 2014;69(0):236–46.
[16] Becker S, Frew BA, Andresen GB, Zeyer T, Schramm S, Greiner M, et al. Features of a fully renewable US electricity system: optimized mixes of wind and solar PV and transmission grid extensions. Energy 2014;72(0):443–58.
[17] McKenna R, Hollnaicher S, v. d. Leye P Ostman, Fichtner W. Cost-potentials for large onshore wind turbines in Europe. Energy 2015;83(0):217–29.
[18] Cannon DJ, Brayshaw DJ, Methven J, Coker PJ, Lenaghan D. Using reanalysis data to quantify extreme wind power generation statistics: a 33 year case study in Great Britain. Renew Energy 2015;75(0):767–78.
[19] Drew DR, Cannon DJ, Brayshaw DJ, Barlow JF, Coker PJ. The impact of future offshore wind farms on wind power generation in Great Britain. Resources 2015;4(1):155–71.
[20] Sharp E, Dodds P, Barrett M, Spataru C. Evaluating the accuracy of CFSR reanalysis hourly wind speed forecasts for the UK, using in situ measurements and geographical information. Renew Energy 2015;77(0):527–38.
[21] Andresen GB, Søndergaard AA, Greiner M. Validation of Danish wind time series from a new global renewable energy atlas for energy system analysis. Energy 2015;93:1074–88.
[22] Olauson J, Bergkvist M. Modelling the Swedish wind power production using MERRA reanalysis data. Renew Energy 2015;76(0):717–25.
[23] Staffell I. Developing an open-access wind profile for Great Britain. London: Department of Energy & Climate Change; 2014.
[24] TradeWind. Integrating Wind: developing Europe's power market for the large-scale integration of wind power. 2009.
[25] Heide D, von Bremen L, Greiner M, Hoffmann C, Speckmann M, Bofinger S. Seasonal optimal mix of wind and solar power in a future, highly renewable Europe. Renew Energy 2010;35(11):2483–9.
[26] Jacobson MZ, Delucchi MA. Providing all global energy with wind, water, and solar power, Part I: technologies, energy resources, quantities and areas of infrastructure, and materials. Energy Policy 2011;39(3):1154–69.
[27] Kiss P, Varga L, Janosi IM. Comparison of wind power estimates from the ECMWF reanalyses with direct turbine measurements. J Renew Sustain Energy 2009;1(3):033105–11.
[28] Kubik M, Brayshaw DJ, Coker PJ, Barlow JF. Exploring the role of reanalysis data in simulating regional wind generation variability over Northern Ireland. Renew Energy 2013;57:558–61.
[29] Gunturu UB, Schlosser CA. Characterization of wind power resource in the United States. Atmos Chem Phys 2012;12(20):9687–702.
[30] Hallgren W, Gunturu UB, Schlosser A. The potential wind power resource in Australia: a new perspective. PLoS One 2014;9(7):e99608.
[31] Ebisuzaki W, Zhang L. Reanalyes, best practices. In: Presented at 12th climate prediction applications science workshop (CPASW). Fairfax, VA; 2014.
[32] Slade R, Bauen A, Gross R. Global bioenergy resources. Nat Clim Change 2014;4(2):99–105.
[33] Carbon Trust. Conversion factors. 2011. URL: https://www.carbontrust.com/media/18223/ctl153_conversion_factors.pdf.
[34] Berg AA, Famiglietti JS, Walker JP, Houser PR. Impact of bias correction to reanalysis products on simulations of North American soil moisture and hydrological fluxes. J Geophys Res Atmos 2003;108(D16):1–15.

[35] Stickler A, Brönnimann S. Significant bias of the NCEP/NCAR and twentieth-century reanalyses relative to pilot balloon observations over the West African Monsoon region (1940–1957). Q J R Meteorological Soc 2011;137: 1400–16.

[36] Decker M, Brunke MA, Wang Z, Sakaguchi K, Zeng X, Bosilovich MG. Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. J Clim 2012;25:1916–44.

[37] Christensen JH, Boberg F, Christensen OB, Lucas-Picher P. On the need for bias correction of regional climate change projections of temperature and precipitation. Geophys Res Lett 2008;35(20):L20709.

[38] Piani C, Haerter JO, Coppola E. Statistical bias correction for daily precipitation in regional climate models over Europe. Theor Appl Climatol 2010;99(1–2): 187–92.

[39] Hagemann S, Chen C, Haerter JO, Heinke J, Gerten D, Piani C. Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. J Hydrometeorol 2011;12(4): 556–78.

[40] Hagemann S, Haerter JO, Piani C. Bias correction of climate model data – the golden solution for impact models or cursed black magic?. In: Presented at GCI conference. Bonn; 2010. URL: http://tinyurl.com/l5jtu2g.

[41] Ofgem. Electricity capacity assessment. Ref 126/12: London. 2012. URL: http://tinyurl.com/97cuuxt.

[42] Hawkins S, Eager D, Harrison GP. Characterising the reliability of production from future British offshore wind fleets. In: Presented at IET conference on renewable power generation (RPG 2011). Edinburgh; 2011.

[43] Kubik M, Brayshaw D, Coker P. Reanalysis: an improved data set for simulating wind generation?. In: Presented at WREF 2012. Denver, CO; 2012. URL: http://tinyurl.com/c4ge72x.

[44] Rose S, Apt J. What can reanalysis data tell us about wind power? Renew Energy 2015;83:963–9.

[45] Staffell I, Green R. Is there still merit in the merit order Stack? The impact of dynamic constraints on optimal plant mix. IEEE Trans Power Syst 2016;31(1): 43–53.

[46] Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, et al. MERRA: NASA's modern-era retrospective analysis for research and applications. J Clim 2011;24(14):3624–48.

[47] Molod A, Takacs L, Suarez M, Bacmeister J. Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2. Geosci Model Dev 2015;8(5):1339–56.

[48] Liléo S, Petrik O. Investigation on the use of NCEP/NCAR, MERRA and NCEP/CFSR reanalysis data in wind resource analysis. In: Presented at EWEA. Brussels, Belgium; 2011. URL: http://proceedings.ewea.org/annual2011/papers/81.pdf.

[49] VTT. Wind energy statistics in Finland. 2015. URL: http://tinyurl.com/wind-energy-finland.

[50] Staffell I. Using WindStats archives to assess the impact of cold climate on turbine output. Wind Rep 2015;28(3):1–3.

[51] European Commission. Eurostat database: electricity infrastructure (nrg_113a) and supply of renewable energies (nrg_107a). 2015. URL: http://ec.europa.eu/eurostat/data/database.

[52] EurObserv'ER. Wind energy barometer. 2015. URL: http://eurobserv-er.info/.

[53] BP. 64th statistical review of world energy. 2015. URL: http://bp.com/statisticalreview.

[54] ENTSO-E. Guidelines for monthly statistics data collection. 2015. URL: http://tinyurl.com/entsoe-ms-guidelines.

[55] Crabtree CJ, Zappala D, Hogg SI. Wind energy: UK experiences and offshore operational challenges. Proc IMechE, Part A J Power Energy 2015;229(7): 727–46.

[56] DTU Wind Energy. Global wind atlas. 2016. URL: http://globalwindatlas.com/methods.html.

[57] Dee D. Reanalysis at ECMWF. In: Presented at NOAA climate reanalysis task force technical workshop. College park, MD; 2015. URL: http://tinyurl.com/jbh24l9.

[58] Pfenninger S, Staffell I. Validated PV output simulations across Europe using reanalysis and satellite data. Energy 2016. http://dx.doi.org/10.1016/j.energy.2016.08.060.