Seasonal Climate Prediction: A New Source of Information for the Management of Wind Energy Resources🖉

VERÓNICA TORRALBA

Barcelona Supercomputing Center, Barcelona, Spain

FRANCISCO J. DOBLAS-REYES

Barcelona Supercomputing Center, and Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

DAVE MACLEOD

Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, United Kingdom

ISADORA CHRISTEL

Barcelona Supercomputing Center, Barcelona, Spain

MELANIE DAVIS

Institut Català de Ciències del Clima, Barcelona, Spain

ABSTRACT

Climate predictions tailored to the wind energy sector represent an innovation in the use of climate information to better manage the future variability of wind energy resources. Wind energy users have traditionally employed a simple approach that is based on an estimate of retrospective climatological information. Instead, climate predictions can better support the balance between energy demand and supply, as well as decisions relative to the scheduling of maintenance work. One limitation for the use of the climate predictions is the bias, which has until now prevented their incorporation in wind energy models because they require variables with statistical properties that are similar to those observed. To overcome this problem, two techniques of probabilistic climate forecast bias adjustment are considered here: a simple bias correction and a calibration method. Both approaches assume that the seasonal distributions are Gaussian. These methods are linear and robust and neither requires parameter estimation—essential features for the small sample sizes of current climate forecast systems. This paper is the first to explore the impact of the necessary bias adjustment on the forecast quality of an operational seasonal forecast system, using the European Centre for Medium-Range Weather Forecasts seasonal predictions of near-surface wind speed to produce useful information for wind energy users. The results reveal to what extent the bias adjustment techniques, in particular the calibration method, are indispensable to produce statistically consistent and reliable predictions. The forecast-quality assessment shows that calibration is a fundamental requirement for high-quality climate service.

1. Introduction

The demand for renewable sources of energy as an alternative to fossil-fuel sources has increased for reasons such as the need to mitigate the climate change resulting from anthropogenic greenhouse gas emissions, the interest in the creation of new economic opportunities,

and the provision of energy access to people living in areas without access to other sources of energy (Renewable Energy Policy Network for the 21st Century 2015; IPCC 2007). Furthermore, the 21st Congress of the Parties for the United Nations Framework Convention on Climate Change (COP21) agreement has recently proposed several polices to promote energy efficiency and replace fossil fuels with renewable sources of energy (Lane 2016). Wind energy is the cheapest option for new sources of power-generating capacity and is the second-leading source of renewable energy worldwide, only exceeded by hydropower in terms of installed capacity (Pryor and Barthelmie 2010; Santos et al. 2015). In recent years, installed capacity for wind power has experienced rapid growth, with a total of 370 GW installed worldwide in 2014. As a consequence, wind energy has become a key element of the electricity supply in many parts of the world (World Wind Energy Association 2015).

Operational and economic issues related to wind energy, such as the need to match supply with demand at all times under the intermittent nature of wind, require the modeling and forecasting of generation processes for wind power at a range of temporal and spatial scales (Pinson 2013). Prediction of the variability of wind energy resources, which has been identified as a challenge to the grid integration of wind energy systems (Najafi et al. 2016; Füss et al. 2015), is a key piece of the decision-making processes because it allows end users to take informed, precautionary action with potential cost savings to their operations. Hence, more efficient energy management strongly depends on having accurate resource forecasts. Wind energy forecasting options have been traditionally limited to short time scales (from hours to a few days) because near-surface winds and thus wind energy production strongly depend on mesoscale and synoptic-scale variability (Graff et al. 2014; Pryor and Barthelmie 2010). At longer time scales, the assessment of the economic feasibility of future wind farms is a function of, among other things, the expected energy yield and the maintenance requirements over their life span of periods from a month to several decades. This information is not readily available to the relevant users, however, who therefore have to rely on past observation-based information, which is often available only as short time series. The need for climate information that is representative of the next few decades has raised the interest of the wind industry in climate projections, which are increasingly being used in long-term resource evaluation (Hueging et al. 2013; Reyers et al. 2015; Vautard et al. 2014).

With a focus on time scales from a month to a decade into the future, current energy practices use an approach that is based on the future climate being a repetition of an estimate of the current climatological behavior (Garcia-Morales and Dubus 2007). Advances in the science of climate prediction that cover the gap in climate information between weather forecasting and projections of climate change can be considered as an alternative to the state of the art by providing predictive information that helps users to make more informed decisions and move beyond using only climatological information.

It recently has been shown that climate predictions are capable of providing additional value for wind energy applications; information at these time scales can be especially beneficial for the management of power production plants (Clark et al. 2017; García-Bustamante et al. 2009; Lynch et al. 2014; Troccoli 2010). For instance, climate predictions could allow electricity-grid system operators to estimate the future production generated by wind farms and use it as input for load-balance models. Should this potential of climate prediction materialize, the matching of supply and demand could be optimized and significant cost savings be realized, with a better anticipation of market changes. This framework will favor greater penetration of renewable electricity into markets.

The scenario described is of great interest to the renewable energy community, but little progress had been made in practice. In recent years the skill of climate predictions has significantly improved (Doblas-Reyes et al. 2013), however. For instance, seasonal forecast systems (i.e., those that provide information for periods ranging from a month to slightly longer than a year into the future) are now providing skillful forecasts for extratropical regions where no substantial skill was found before (Clark et al. 2017; Dunstone et al. 2016; Scaife et al. 2014). This will promote their application in wind energy decision-making, as illustrated for different energy sources (De Felice et al. 2015; Garcia-Morales and Dubus 2007). At this time, however, there are very few instances of the application of seasonal predictions in the wind energy industry. Improved climate information that includes seasonal forecasts may change this situation, for example by allowing innovative wind energy insurance and helping to cover high-risk periods associated with persistent lower-than-expected wind resource.

Seasonal predictions will be beneficial if they are skillful enough, but also they must be tailored to the potential users in a decision-making context. In particular, seasonal predictions have systematic errors that make them unusable unless they are postprocessed to have statistical features that are similar to those of the observational reference employed. This problem has been recognized by the climate science community as

one of the main challenges for moving to a better use of climate predictions (Buontempo et al. 2014; Coelho and Costa 2010). Two recent European projects related to climate services that are sponsored by the European Commission Seventh Framework Programme (FP7) [European Provision of Regional Impacts Assessments on Seasonal and Decadal Timescales (EUPORIAS; see online at http://www.euporias.eu/) and Seasonal-to-Decadal Climate Prediction for the Improvement of European Climate Services (SPECS; see online at http://www.specs-fp7.eu/)] have tried to address these challenges and support the development of sectorial climate services in Europe through the involvement of stakeholders in defining effective ways to develop climate information.

This paper raises the limits associated with current seasonal forecast systems when used in wind energy applications. It focuses on the description of appropriate bias adjustment techniques to overcome some of these limits and to promote the use of climate forecast information on those occasions in which it can provide greater accuracy than current approaches. The method that is described recognizes that end users must be provided with information about the prediction uncertainty (Alessandrini et al. 2013), and therefore a probabilistic approach is adopted because it is more valuable in user-specific loss functions (Pinson and Tastu 2013).

An overview of the necessary steps to provide climate predictions to the wind energy sector is provided in Fig. 1, which summarizes the main challenges addressed in this paper. Section 2 of the paper introduces the datasets and describes one of the most widely used seasonal forecast systems and its limitations. Section 3 describes appropriate bias adjustment techniques. It also introduces measures to assess forecast quality and explains their relevance in a user context. Section 4 presents the impact of the bias adjustments on the wind speed seasonal forecasts, including an analysis of the changes in the statistical properties of the postprocessed predictions. Section 5 reports the concluding remarks and provides a wider context for future work in the dissemination of climate predictions in user-relevant formats.

2. Data

In this study, we use 10-m wind speed forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) System 4 (hereinafter System 4) operational seasonal forecast system (Molteni et al. 2011), which is based on a global climate model with coupled atmospheric and oceanic components. System 4

comprises an ECMWF atmospheric model—the Integrated Forecast System (IFS) CY36R4 with a T255 spectral truncation (horizontal resolution of approximately 80 km) and 91 vertical levels reaching up to 0.1 hPa—coupled to the Nucleus for European Modelling of the Ocean (NEMO), version 3.0, ocean model. The ocean model uses a grid with horizontal resolution of ~1° in the extratropics with equatorial refinement and 42 levels in the vertical direction. The atmosphere and ocean are coupled using a version of the Ocean Atmosphere Sea Ice Soil (OASIS3) coupler developed at the Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS).

System 4 is run in ensemble-prediction mode. Ensemble predictions are a way to account for uncertainties in the climate system, in particular those associated with the imperfections of the initial conditions and the model formulation (Slingo and Palmer 2011). For this reason, the operational System 4 forecasts are produced at the beginning of each month with 51-member ensembles. Each member of the ensemble uses slightly different initial conditions and different realizations of stochastic representations of subgrid physical processes in the atmosphere. This approach allows prediction of the forecast uncertainty (measured by the ensemble dispersion) along with the prediction itself. The simulations are performed for up to 7 months into the future.

Seasonal forecast systems traditionally do not produce operational forecasts of wind speeds at turbine height levels. Instead, wind speeds are made available at 10 m or at different pressure levels. It is difficult to interpolate directly to hub height because the physical height of pressure levels is not constant over time. For that reason, 10-m wind speeds have been selected for this analysis. It might be possible for the forecast systems to deliver wind speed at hub height, should the renewable-energy community show an interest in seasonal forecast systems.

The analysis in this paper focuses on boreal winter because the winter season has larger wind speed variability in the Northern Hemisphere (Archer and Jacobson 2013). In addition, the analysis of seasonal predictions of wind speed in winter can be relevant because of the higher variability in the supply of wind power in that particular season (Bett and Thornton 2016). This illustrates the potential of seasonal predictions for end users, because these predictions potentially have more impact where the interannual variability is the largest, although other seasons have also been analyzed (see Fig. S1 of the online supplemental material), and the conclusions apply equally. The predictions considered here are those issued on 1 November, for which 3-month statistics for the December–February
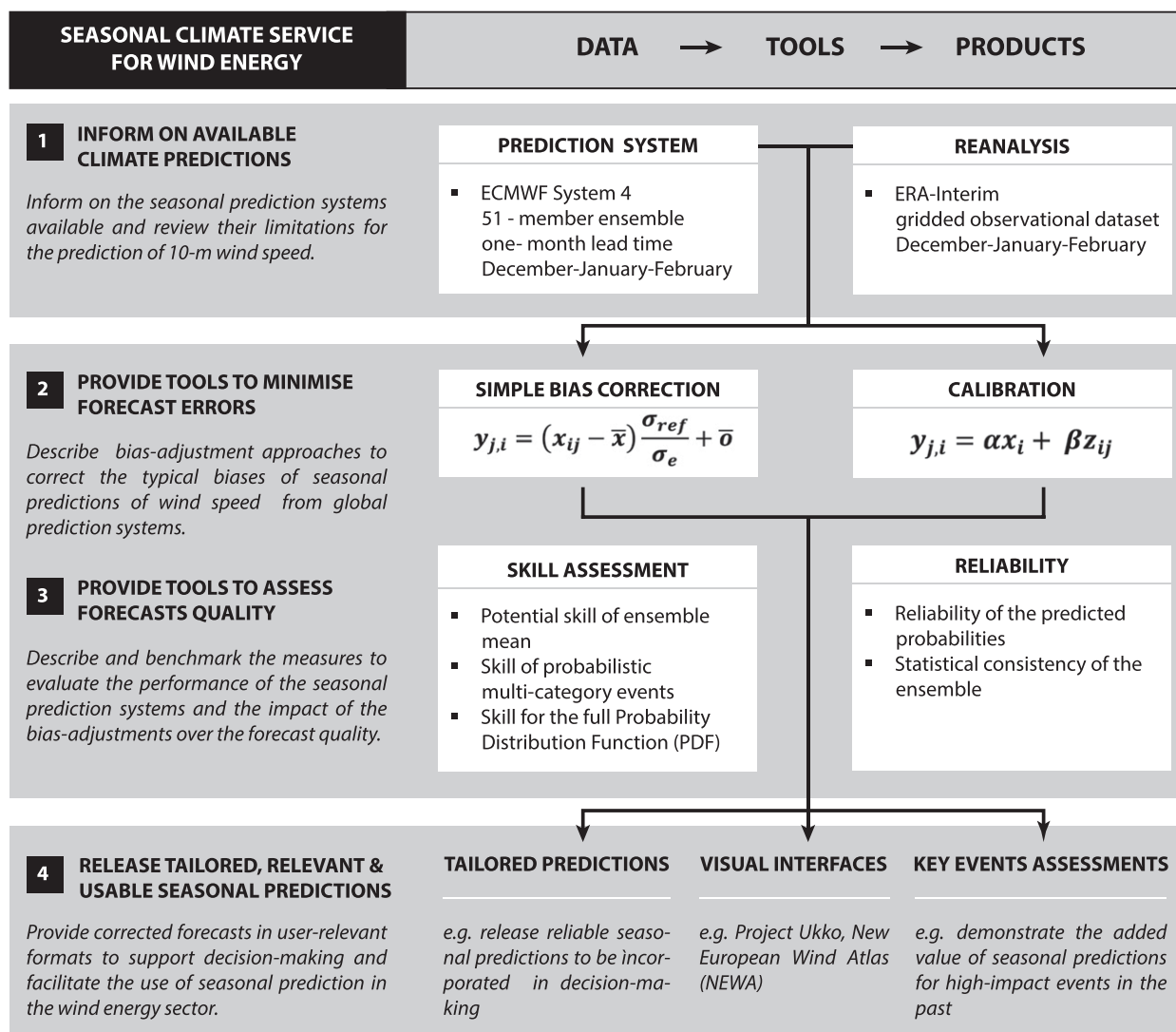
FIG. 1. Main steps for the development of a climate service for the wind energy sector that is based on seasonal climate forecasts. Steps 2 and 3 in the diagram outline the main challenges addressed in this paper.

period (DJF; also known as a 1-month-lead seasonal forecast) are made. Predictions over the period 1981–2013 have been used in the study. The prediction for DJF in 2013 has been used as an operational forecast, and the predictions over 1981–2012 have been used as the retrospective predictions (hindcasts) to be used in the validation process. This consideration aims to emulate true operational prediction conditions for which no observed information about the future is available.

To evaluate the System 4 prediction quality, we compare the predicted 10-m wind speed with the corresponding variable of the ERA-Interim reanalysis (Dee et al. 2011). This reanalysis uses the ECMWF IFS atmospheric model to assimilate observational data of

many types, including in situ observations and satellite retrievals, to produce a spatially and temporally complete "best guess" gridded observational dataset. ERA-Interim has the same resolution as System 4. This resolution is fairly coarse, but this product offers uniform global coverage in exchange. Given the sparsity of global wind observations, reanalyses have demonstrated their potential usefulness for large-scale wind energy applications (Cannon et al. 2015). The problems related to the lack of a long-enough historical data record have also promoted the use of reanalyses by the wind industry (Rose and Apt 2015).

For this reason, and being aware that reanalysis estimates can often be far from point observed values, the reanalysis has been used as the best available estimate of

wind speed. The choice of reanalysis dataset is arbitrary, and the conclusions are equally valid when using other reanalyses, either global or regional. Further work is needed to assess the seasonal predictions for the locations of specific wind farms.

3. Method

a. Data processing

Wind speed forecasts are affected by biases resulting from the inability to numerically reproduce all of the relevant processes that are responsible of climate variability (Doblas-Reyes et al. 2013). Apart from biases in the mean and other characteristics of the distribution of the simulated variables, for probabilistic forecasts there are additional difficulties, such as the lack of forecast reliability (Pinson 2012), which is a measure that quantifies the agreement between the predicted probabilities and observed relative frequencies of a particular event. This is important from a wind energy point of view since reliable probabilities are expected to be included in decision-making processes. Hence, climate predictions require a bias adjustment stage to statistically resemble the observational reference, minimize forecast errors, and formulate reliable probabilities. Bias adjustment of the wind speed has been identified as a requirement by the wind energy sector to fulfill acceptable reliability for use in its decision-making processes (Alessandrini et al. 2013).

This paper illustrates the relative merits of different techniques for statistical bias adjustment of ensemble forecasts to address different aspects of the forecast error. Two approaches, a simple bias correction and a calibration method, have been selected.

1) SIMPLE BIAS CORRECTION

Simple bias correction is based on the assumption that both the reference and predicted distributions of seasonal wind speed are approximated well by a Gaussian (normal) distribution. The adjustment creates predictions that have the same mean and standard deviation as the reference dataset. This is a zero-order approach for correction of the systematic mean error that has been previously applied to correct temperature and precipitation (Leung et al. 1999). The Gaussian assumption is a limitation of the approach because the monthly and seasonal wind speed distribution can be, at times, slightly non-Gaussian.

The bias correction scheme can be summarized in this way:

$$y_{ij} = (x_{ij} - \overline{x}) \frac{\sigma_{\text{ref}}}{\sigma_e} + \overline{o}. \tag{1}$$

Seasonal mean anomalies are calculated by subtracting the ensemble mean $\overline{x}$ of the seasonal averages from the seasonal average x_{ij} of each forecast for each year i and member j. A new seasonal mean y_{ij} is calculated by multiplying the seasonal mean anomaly by the ratio of the standard deviation of the reference dataset σ_{ref} to the interannual standard deviation of the ensemble members σ_e and adding the climatological mean $\overline{o}$ of the reference dataset. This is done for each grid cell separately, resulting in a new wind speed forecast ensemble that has the same ensemble mean and standard deviation as the reference.

2) CALIBRATION METHOD

The calibration can be considered as a way of obtaining predictions with an interannual variance that is equivalent to that of a reference dataset in a similar way to the simple bias correction method but at the same time ensuring an increased reliability of the probability predictions. Here we apply the technique of variance inflation (von Storch and Zwiers 2001). This calibration strategy has been selected because an inflation of the ensemble spread is required to obtain reliable probabilities, and it is applied as in (Doblas-Reyes et al. 2005).

If x_i is the ensemble-mean prediction for any grid point at year i and z_{ij} is the difference of ensemble member j from the ensemble mean, then the calibrated estimate of the ensemble member j can be expressed as

$$y_{ij} = \alpha x_i + \beta z_{ij}. \tag{2}$$

The coefficients α and β are defined as follows:

$$\alpha = \text{abs}(\rho) \frac{\sigma_{\text{ref}}}{\sigma_{\text{em}}} \quad \text{and} \tag{3}$$

$$\beta = (1 - \rho^2)^{1/2} \frac{\sigma_{\text{ref}}}{\sigma_e}. \tag{4}$$

The σ_{em} is the standard deviation of the ensemble mean (the time series of x_i), σ_e is the standard deviation of the ensemble, σ_{ref} is the standard deviation of the reference, and ρ is the correlation between the ensemble mean of the retrospective forecasts and the reference dataset. The α and β coefficients are found under two constraints: The first is that the standard deviation of the inflated prediction is the same as that for the reference, and the second is that the predictable signal after the inflation is made equal to the correlation of the ensemble mean with the reference dataset.

b. Assessment of forecast quality

Seasonal forecast systems, as in any other forecasting process, have to be systematically compared with a

reference, preferably observations, to assess their overall quality in a multifaceted process known as forecast-quality assessment (Mason and Baddour 2008). This is a fundamental step in the prediction problem because a prediction has no value without an estimate of its quality that is based on past performance (Doblas-Reyes et al. 2013). Moreover, quantification of the uncertainty is one of the most crucial aspects for successful development of the wind industry and minimization of the financial risk.

Three sources of uncertainty in common scoring metrics of probabilistic forecasts should be considered: improper estimates of probabilities from small-sized ensembles, an insufficient number of forecast cases, and imperfect reference values as a result of observation errors. A way to alleviate these problems is to use several scoring measures to offer a comprehensive picture of the forecast quality of the system (Jolliffe and Stephenson 2012) and to apply statistical inference as often as required. Note that these sources of uncertainty are independent of the uncertainty of the individual forecasts: the user should consider and be provided with both types of uncertainty when making decisions when this information is included.

Several scoring measures are used in this paper, including skill and reliability measures such as the reliability diagram and the rank histogram. Forecast quality has been used to evaluate the performance of the seasonal forecast system as well as the impact of the two bias adjustment techniques on forecast quality. The goal is to offer the most general and, a priori, relevant information for a user in the wind energy sector instead of the traditional view offered by climate scientists in which the information provided to the users is mainly based on correlation, which is very useful but gives only a small part of the information user requires.

1) SKILL SCORES

Skill estimates that are based on the performance of the system in the past may guide users about the expected performance of future forecasts (Weisheimer and Palmer 2014), always with the caveat that the predictability of the climate system may change over time. Skill scores are a tool for end users that can be used to develop alternative strategies to their baseline information to minimize the risk and to perform optimal management (Pinson et al. 2009). Skill scores for both deterministic (ensemble mean) and probabilistic predictions are considered.

The Pearson correlation coefficient between the ensemble mean and the reference dataset has been used as a measure of the linear correspondence between the forecasts and the reference. This deterministic skill measure is invariant to changes in scale; hence the bias

correction and calibration of the forecasts do not change the correlation of the ensemble mean with the observations. The bias adjustment techniques that are illustrated in this paper (defined in section 3a) have been applied in leave-one-out cross validation to mimic as closely as possible an operational context in which new coefficients might be estimated to predict each year. In cross-validation mode, the prediction to be adjusted is removed from the sample used to estimate the coefficients. As a result, the correlation of the postprocessed forecast changes relative to the correlation computed directly with the uncorrected forecasts.

A comprehensive measure of the predictive skill for probabilistic seasonal predictions of categorical events is the ranked-probability skill score (RPSS) (Epstein 1969; Wilks 2011). This is a squared distance between the cumulative probabilities of the categorical forecast and reference vectors relative to a naive forecast strategy, which in our case has been taken as the climatological statistics (hereinafter referred to as "climatology"; derived from all of the possible events recorded in the past) because this is the preferred current choice of the users targeted by this analysis. The RPSS is based on the rank probability score, which is a measure of the squared distance between the forecast and the reference cumulative probabilities. In the case that is presented here, the RPSS has been computed on the basis of categorical forecasts for terciles. Three equiprobable events are associated with the two terciles of the climatological distribution of the reference: wind speed exceeding the upper tercile (above-normal category), wind speed not exceeding the lower tercile (below-normal category), and values occurring between the two terciles (normal category). The probabilities have been computed as the fraction of ensemble members in the corresponding category. This is only one example; other categories could be defined if they better represent the decisions involved in precautionary climate action. The individual values of the reference dataset in the verification time series can fall in any of the three categories, with probability determined by the probability density function for the target season.

The continuous ranked probability skill score (CRPSS) is a probabilistic skill score (Jolliffe and Stephenson 2012) that has commonly been used to evaluate the predictive skill of the full probability distribution. It is based on the continuous ranked probability score (CRPS), which is a score that reduces to the mean absolute error if a deterministic forecast is used. The CRPS measures the difference between the predicted and observed cumulative distributions, and it can be converted into a skill score that measures the performance of a forecast relative to the climatology.

The RPSS and CRPSS range from 1 to $-\infty$. Skill scores below 0 are defined as unskillful, those equal to 0 are equal to the climatology forecast, and anything above 0 is an improvement upon climatology, through to 1, which indicates a "perfect" forecast.

Fair scores to ensemble forecasts have been recently introduced (Fricker et al. 2013; Ferro 2014). A skill score is fair when it favors predictions with ensemble members that perform as if they have been sampled from the same distribution as the reference dataset. The fair versions of the RPSS and CRPSS have been used to give an estimate of what the skill is when an infinite ensemble size is used (a measure of potential skill). The differences between the results of the fair scores and the basic scores are small, as has been shown for the RPSS in Fig. S2 of the supplementary material.

2) RELIABILITY

Reliability analysis of prediction systems remains as a prime concern for the wind energy sector, as it is for any user of probability predictions, because of the risks and uncertainties involved in the forecasting of wind resources (Chaudhry and Hughes 2012). Rank histograms are a simple tool to evaluate the reliability of ensemble forecasting systems (Elmore 2005). They are generated by dividing the observations among a limited number of bins, thereby defining a set of exhaustive and mutually exclusive events. Then the observed frequencies for these bins are compared with the corresponding forecast probabilities. Rank histograms help to determine whether the forecast is assumed to be reliable, and in that case it is expected to be flat. Some deviations from uniformity can appear for reliable forecasts because of randomness, however. The rank histograms have been displayed on probability paper (Bröcker 2008). On the y axis, rank histograms display cumulative probabilities (instead of the traditional observed frequency) that indicate how probable that observed frequency would be if the prediction were reliable. This information is useful to identify whether the deviations from reliable behavior are systematic or merely random. In addition, the readability of the rank histogram is further improved by scaling the ordinate by a logit transformation that has the effect of displaying both small and large probabilities equidistantly. On the right, the 90%, 95%, and 99% simultaneous confidence intervals have been represented.

Rank histograms illustrate whether the ensemble members and the verifying observation come from the same probability distribution, in which case the forecasts are statistically consistent and no calibration of the ensemble is needed. This happens when the rank histogram is flat (as if coming from a uniform distribution). Because of sampling variations, the histograms are almost never flat, however. To assess whether the deviations from flatness are attributed to chance or to deficiencies in the forecasts, goodness-of-fit test statistics are computed: the Pearson χ^2, the Jolliffe–Primo test statistic for slope (JP slope), and the Jolliffe–Primo test statistic for convexity (JP convex) (Jolliffe and Primo 2008). The Jolliffe–Primo statistics are obtained from the decomposition of the Pearson χ^2 in components that allow the identification of bias (slope) or under-/overdispersion (convexity) in the forecast ensemble. The detailed mathematical definition of this goodness-of-fit test can be found in the appendix of Jolliffe and Primo (2008).

Reliability diagrams are a common diagnostic of probabilistic predictions that assess both reliability and skill. They consist of a plot of the observed relative frequency against the predicted probability of a dichotomous event, providing a quick visual assessment of the impact of tuning probabilistic forecast systems. A perfectly reliable system should result in a line that is as close as possible to the diagonal, within a certain measure of uncertainty.

The information provided by the reliability diagram should be interpreted with care because even a perfectly reliable forecast system is not expected to have an exactly diagonal reliability diagram because of the limited samples that are typical of seasonal forecast systems (Jolliffe and Stephenson 2012). To address this problem, we have included consistency bars (Bröcker and Smith 2007) in these diagrams. They indicate how likely the observed relative frequencies are, under the assumption that predicted probabilities are accurate.

To draw a reliability diagram, discretization and grouping into probability bins (10 in this paper) of the probability forecasts have to be done. A reliability diagram also includes the frequency of the forecast probabilities included in each bin, which is known as a sharpness diagram. Sharpness gives an indication of the variation in forecast probabilities issued by the prediction system, independent of the observations.

The rank histogram and the reliability diagram are complementary tools to assess the reliability of the system. The former assesses the full forecast ensemble and does not require the formulation of forecast probabilities—an aspect that is necessary in the case of the reliability diagram, for which one assesses the features of both the forecast system and the statistical model that transforms the ensemble into probabilities.

4. Results

Total installed wind power indicates the power capacity available on each wind farm. It has been
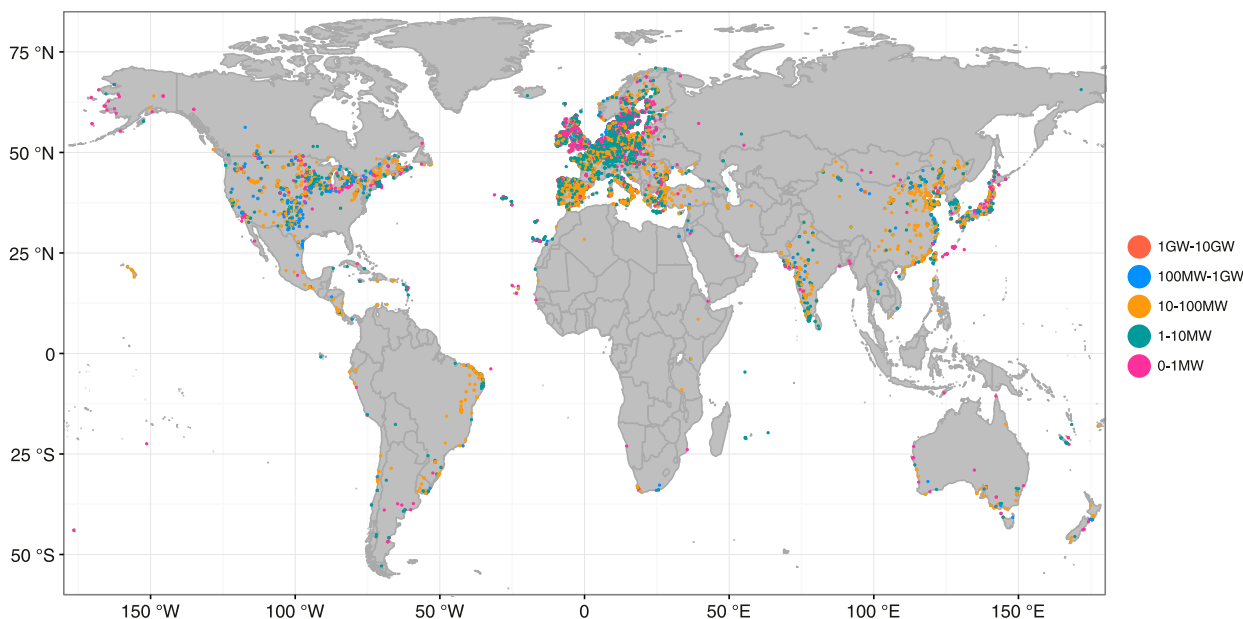
FIG. 2. Total installed wind-power capacity for each individual wind farm (those that are operational and under construction have been included) in 2015 (source: http://www.thewindpower.net).

represented in Fig. 2 to identify which are the most important locations from the point of view of a wind energy user. To illustrate the performance of seasonal predictions, we selected two regions that are key for the wind energy sector because of the concentration of wind farms at those locations. For the selection of the regions we have also taken into account the potential skill available in such regions (supplemental Fig. S1). The first region is in Canada (longitude from 112.5° to 113.2°W and latitude from 50.3° to 51.0°N). Canada is an important player in terms of energy resources (Vaillancourt et al. 2014) and is a global leader in the sustainable development of wind energy. This region had an exceptional year in 2014 for wind energy development, ranking seventh globally in terms of newly installed capacity (Canadian Wind Energy Association: http://canwea.ca/wind-energy/installed-capacity/) for that year. The region in the North Sea (longitude from 9.8° to 10.6°E and latitude from 58.0° to 58.7°N) is the second region that was considered. It is the most important region for offshore energy activities in Europe because of the large and consistent wind resource, the relatively shallow water that minimizes the cost of the wind farms, and the proximity to developed electricity markets (Schillings et al. 2012).

Figure 3 displays the predictions for the uncorrected, bias-corrected, and calibrated sets for these two regions. The effect of the bias adjustment on the predictions is that, when the corrections are applied the hindcasts (gray dots) and the reference dataset (black dots) show similar mean and variance. After the bias adjustment,

the probabilities in each category differ as a result of the changes in the ensemble distribution. The skill changes accordingly with the bias adjustment, showing a decrease in the correlation and an increase in the probabilistic skill scores. The decrease of the correlation is due to the cross validation, which leads to an implicit leakage of information and a degeneracy in this measure of potential skill (Barnston and van den Dool 1993; Barnston et al. 2012). The improvement of the fair RPSS and CRPSS are associated with reduction of the systematic errors. Contrary to the correlation, the RPSS and the CRPSS are both sensitive to the systematic differences in the statistical properties (mean and variance) of the predicted variables with respect to those in the observations and to the inadequacy of the ensemble dispersion to act as a prediction of the forecast error (the lack of reliability). This is a useful example of the importance of using more than one measure of forecast quality, in particular when dealing with user-relevant variables.

The information provided by global forecast systems is relatively coarse. In a global context, the sizes of the two selected regions are small. Besides, for a small region the skill is expected to be noisier and less robust than for a larger one. To explore how the size of the region affects the forecast quality we have estimated the forecast quality for larger regions (supplemental Fig. S2). The comparison shows that the skill differences are small when a larger region is considered. Future work will focus on the formulation of predictions for specific sites. This is a nontrivial task because the bias
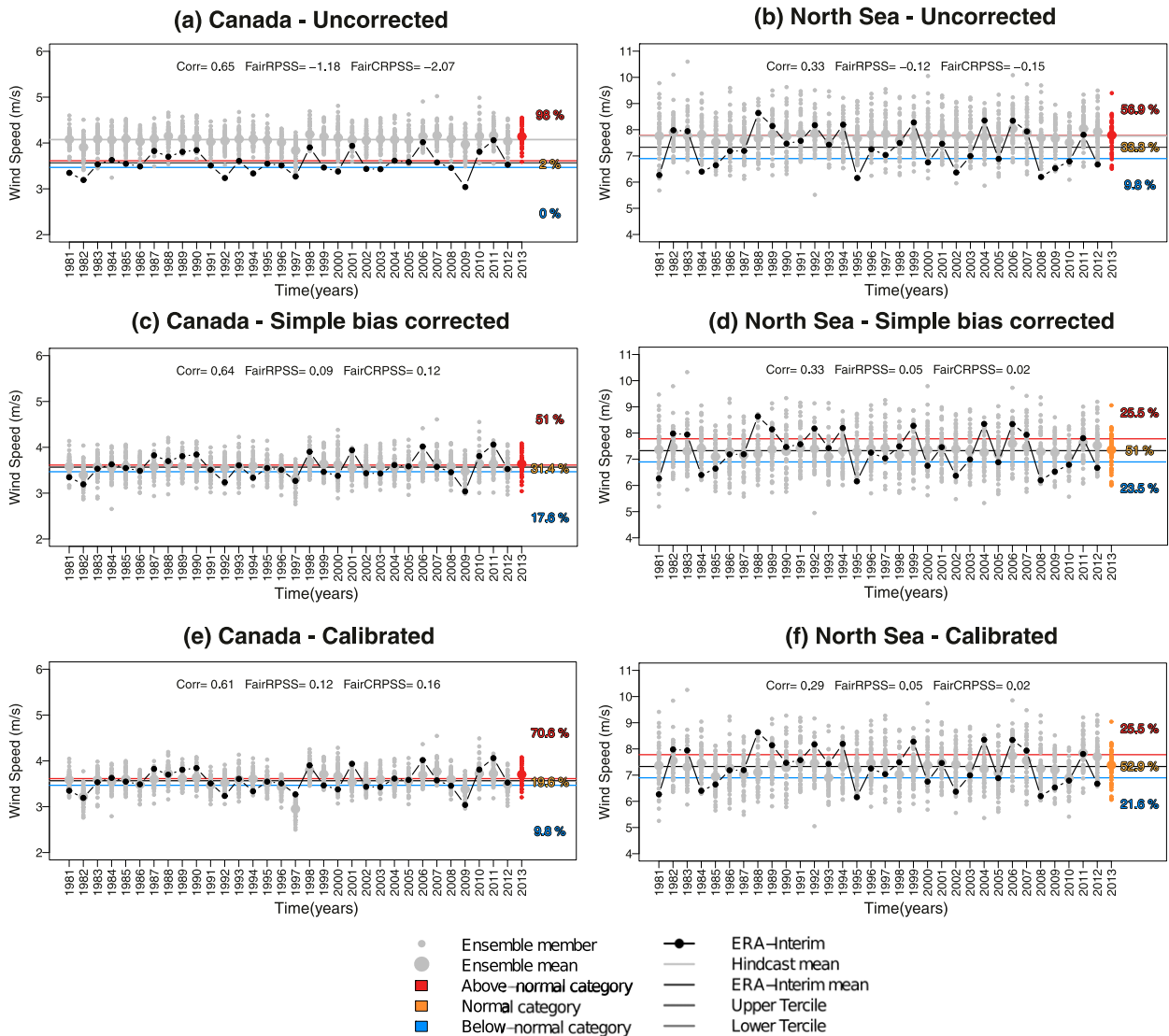
FIG. 3. Time series of 10-m wind speed from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on 1 Nov for the period of 1981–2013. The ensemble members of the hindcasts are represented as small gray dots, and the ensemble mean is represented with a large gray dot for each start date. The gray horizontal line shows the mean of the hindcast for the whole period (1981–2012), and the blue and red horizontal lines show its lower and upper terciles, respectively. The ensemble members of the forecast year (2013) are represented as red dots. The percentages indicate the fraction of members in each category, which is limited by the terciles. The black dots represent the 10-m wind speed values of ERA-Interim. The black horizontal line shows the mean of the ERA-Interim for the 1981–2012 period. Correlation, RPSS, and CRPSS are shown in the upper part of each panel.

adjustment techniques necessary in seasonal forecasting require sufficiently long observational references that are not readily available.

The fair RPSS maps for the uncorrected, bias-corrected, and calibrated wind speed are shown in Fig. 4. The uncorrected predictions (Fig. 4a) display very low scores all around the world. The highest values are found in tropical regions, in particular in some regions of northeastern South America and northwestern Africa.

This maximum can be explained because the largest predictability at seasonal time scales is attributed to anomalies in the tropical sea surface temperatures resulting from coupled ocean–atmosphere phenomena, in particular those related to El Niño–Southern Oscillation events (Kirtman and Pirani 2009) that affect mainly the regions mentioned above.

Figures 4b and 4c show that the fair RPSS increases globally when bias adjustment is applied. This kind of

(a) Uncorrected

(b) Simple bias corrected   (c) Calibrated

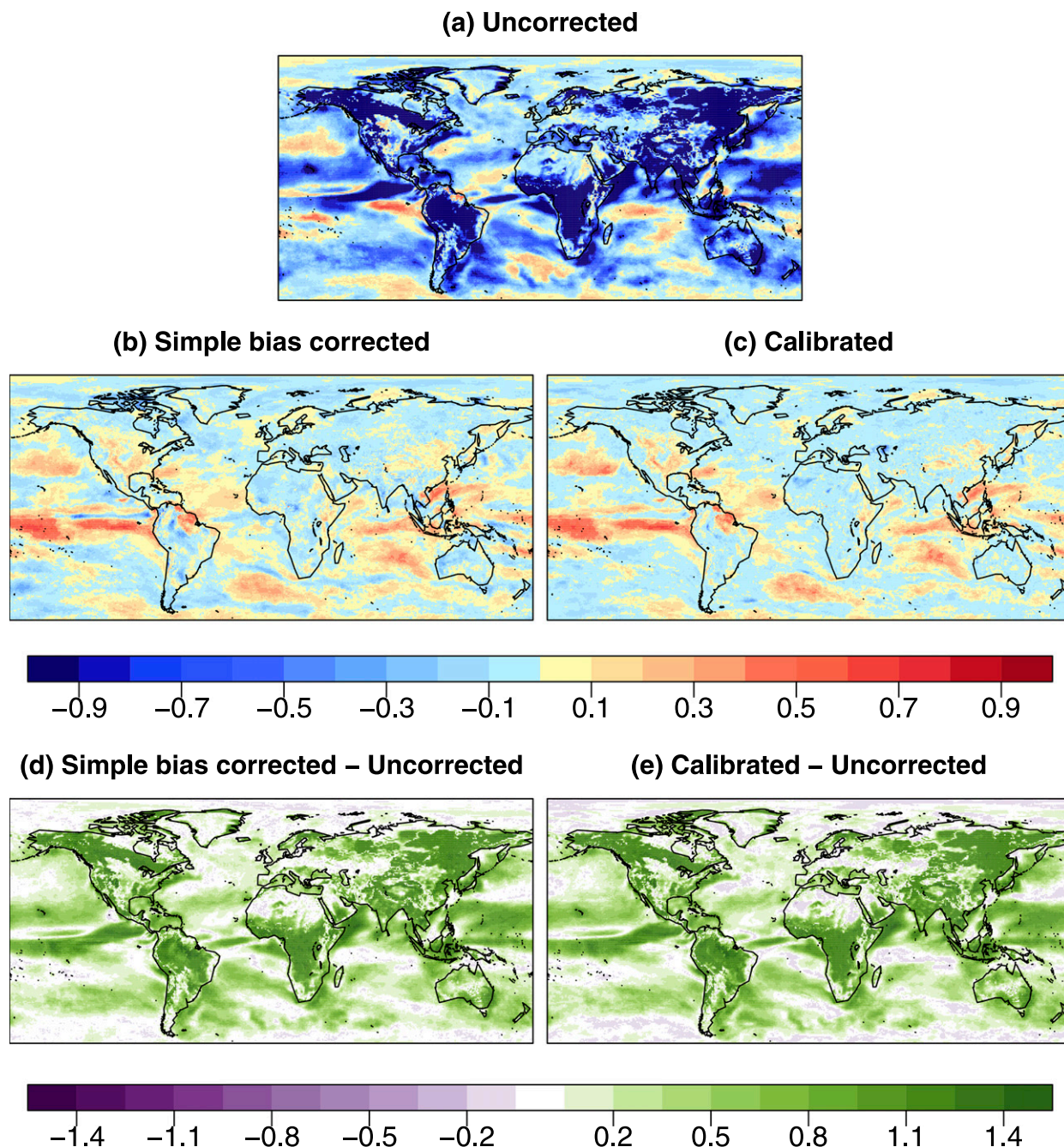## (d) Simple bias corrected – Uncorrected   (e) Calibrated – Uncorrected

FIG. 4. Fair RPSS for tercile events of 10-m wind speed forecasts from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on 1 Nov for the period of 1981–2012.

assessment is widely available for variables such as temperature and precipitation but is not available for wind speed. The skill improvement has been quantified in Figs. 4d and 4e, which indicate that the skill scores for the bias-adjusted predictions increase by more than 1 relative to the uncorrected ones. The fair RPSS maps (Figs. 4b,c) for the postprocessed predictions have their maximum values in the tropics. Although the skill is relatively low at extratropical latitudes, some positive skill is found in those regions. For instance, some regions in Europe such as the North Sea or Scandinavia display positive values. Wind speed predictions show the highest skill in northern Europe; in southern Europe negative RPSS values are found. This result is in agreement with

previous work (e.g., Weisheimer et al. 2011) that indicated that seasonal dynamical predictions have limited forecast quality over Europe.

The skill improvement is also present in southeastern Asia, the central United States, and northeastern South America, where positive values appear when bias-correction and calibration techniques are applied. The bias adjustment allows the skill in those regions associated with ENSO teleconnections (Hamlington et al. 2015; Quan et al. 2006), as well as that associated with other sources of seasonal to interannual predictability such as the persistence of the North Pacific decadal oscillation (Gershunov and Cayan 2003), to emerge. Wind speed with positive skill in North American regions has important implications for the wind energy sector in this economically active region.

The differences between the correlation and CRPSS before and after the bias adjustment of the wind speed forecasts have been included in online supplemental Figs. S3 and S4. The correlation of the uncorrected forecasts is always higher because of the cross-validation leakage mentioned above. It is noticeable that the correlation spatial distribution in the calibrated hindcasts is noisier than in the two other types of forecasts considered. This noise is due to the coefficients estimated in the calibration having a smaller spatial decorrelation length and being less robust than the mean and variance used in the simple bias correction.

For the uncorrected predictions (Figs. 5a,b), the overpopulated lower ranks and the negative slope in the rank histogram illustrate that a positive unconditional bias is present in the data. These biases appear for the predictions of both regions, although the effect of this deficiency seems more important in Canada (Fig. 5a) where all the observations are exceeded by the majority of the ensemble members, leaving the highest rank categories almost empty. The simple bias-corrected and calibrated forecasts show more homogeneously populated ranks, indicating that the reliability of the ensemble improves when the bias adjustment is applied. The deviation of the flatness of these rank histograms could be the result of some forecast deficiencies still remaining after the bias adjustment. For instance, for the calibrated forecasts in Canada (Fig. 5e), the rank 50 shows a very large value that might indicate that the ensemble overestimates the true uncertainty range.

To assess whether the deviations from flatness of the rank histograms are attributed either to chance or to deficiencies in the forecasts, goodness-of-fit test statistics, with the null hypothesis being that the rank histogram is uniform, are computed (Table 1). The three statistical tests—Pearson χ^2, JP slope, and JP convex—allow us to identify whether the forecasts are biased or whether the ensemble has over- or underdispersion.

Table 1 shows that departures from flatness exist for the uncorrected forecasts, especially in Canada, where the tests take very high values, showing that the ensembles are underdispersive, as evidenced by the high JP convex values. The high values of the JP slope show that the forecasts are also affected by biases. The uncorrected forecasts in the North Sea also have biases and are underdispersive, although the statistical tests have smaller values than those in Canada. The results are statistically significant, with the significance p values being virtually 0.

The tests applied to the simple bias-corrected forecasts and the calibrated forecasts indicate that the deviation from flatness is minimized when the bias adjustment is applied. The Pearson χ^2 for the calibrated data in the Canada region has higher values than the simple bias-corrected ones (p value of 0.01), and the JP tests provide no evidence of departures from flatness, with p values higher than 0.01. This result consequently shows that the biases and the underdispersion in the raw ensemble are corrected and that the deviations from uniformity are independent of these specific problems. Making sure that the ensemble is well calibrated is a critical aspect of the forecast for the user because it suggests that the ensemble predictions represent the forecast error, within statistical sampling, and can be trusted in specific applications that have been developed using meteorological observational references.

To further analyze the impact of the bias adjustment on reliability, reliability diagrams (Fig. 6) allow comparison between the observed frequencies and forecast probabilities (obtained from the ensemble forecasts) for binary events. The events are defined by the thresholds of the lower and upper terciles, as for the RPSS but in a dichotomous way. If the prediction system is reliable, then good agreement should exist between forecast probabilities and observed relative frequencies and the graph should be close to the diagonal.

The slope of the reliability diagrams is positive. This shows that as the forecast probability of the event occurring increases so does the verified chance of observing the event and that therefore the forecasts have some reliability. The reliability curves for the three events have a steeper slope than the diagonal in both regions, suggesting that the probability forecasts are over-confident. For the uncorrected forecasts in Canada (Fig. 6a), the curve for the below-normal category (blue line) flattens when the forecast probability is above 0.45. This means that, when the forecast probability is higher than 0.45, there is no relationship between the forecast probabilities and the frequency of
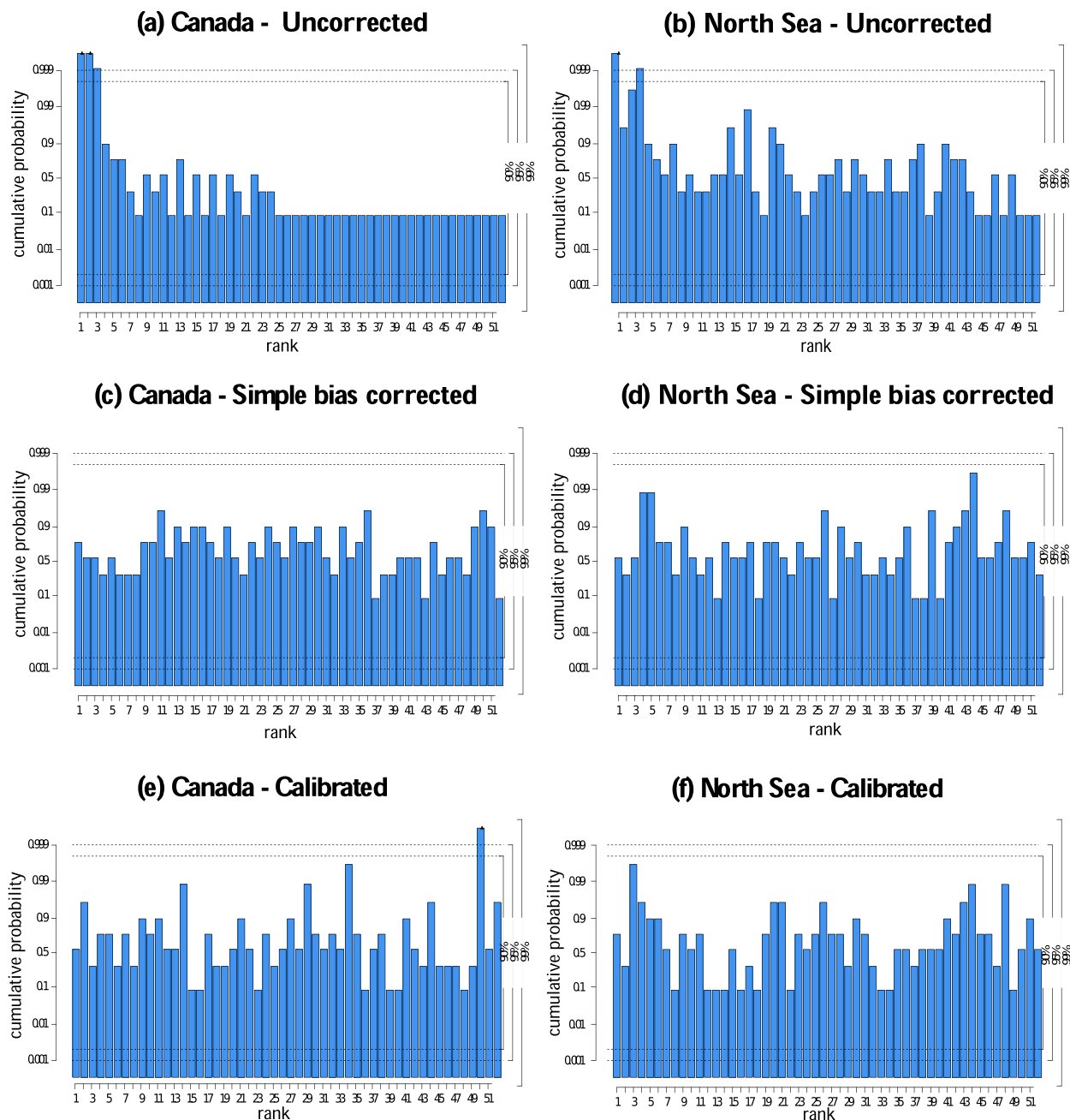
FIG. 5. Rank histograms of 10-m wind speed forecasts from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on 1 Nov for the period of 1981–2012. These rank histograms have been represented on probability paper to show whether the deviations from reliable behavior are systematic or random. The x axis represents the ranks. The probabilities of the cumulative observed frequency on a log-it scale are shown on the y axis. On the right side of each panel, 90%, 95%, and 99% simultaneous confidence intervals are indicated. If all ranks were equally likely on average, approximately 90% of all rank histogram would be contained in the 90% confidence interval and approximately 10% of all rank histograms would have at least one bar that falls outside this interval.

the observed below-normal wind speeds. The reliability diagram for the uncorrected predictions in the North Sea (Fig. 6b) shows only a narrow set of probabilities issued, with values ranging from 0.1 to 0.5 for the above-normal (red line) and below-normal (blue line)

categories and from 0.4 to 0.7 for the normal category (orange line). In addition, the above-normal category is so steep that it falls outside the consistency bars. This result illustrates the poor reliability for that event in the North Sea when the predictions are uncorrected.

TABLE 1. Goodness-of-fit tests: Pearson χ^2, JP-slope, and JP-convex statistics formulated by Jolliffe and Primo (2008). They have been computed from the rank histograms (Fig. 4) of 10-m wind speed forecasts from ECMWF System 4 in winter (DJF) for the period of 1981–2012. The p values are in parentheses for each combination of statistic and region.

	Uncorrected		Simple bias corrected		Calibrated	
	Canada	North Sea	Canada	North Sea	Canada	North Sea
Pearson χ^2	2080.37 (0.00)	148.25 (2.01×10^{-11})	36.12 (0.94)	57.25 (0.25)	77.56 (0.01)	68.62 (0.05)
JP slope	291.69 (2.12×10^{-65})	38.46 (0.56×10^{-11})	0.17 (0.68)	0.1 (0.75)	0.03 (0.87)	0.12 (0.72)
JP convex	319.69 (1.69×10^{-71})	11.32 (0.76×10^{-4})	1.15 (0.28)	1.77 (0.18)	0.73 (0.39)	3.07 (0.06)

The reliability curves of the bias-corrected predictions (Figs. 6c,d) show features that are similar to those of the uncalibrated ones. One should bear in mind that, apart from correcting the mean and standard deviation of the forecast distribution, the simple bias correction does not have any additional impact on the predictions and, hence, no substantial changes beyond the effect of the cross validation should be expected in the reliability diagram.

The calibrated predictions for the above-normal and below-normal events (Figs. 6e,f) have reliability diagrams that have their points lying closer to the diagonal than is found for the uncorrected and bias-corrected predictions. This result corresponds to a better agreement between the forecast probabilities and the probability of the observed event than in the other two cases, suggesting that the overconfidence has been corrected. For the North Sea (Fig. 6f), the slope of the curve for the normal category (orange line) (Fig. 6f) becomes horizontal, suggesting that the system cannot discriminate between predictable and unpredictable normal wind speeds in this region, which is not surprising because normal events might not have strong signals, which are the kinds of signals that are associated with the predictability of the system.

In addition, for the predictions of below-normal and above-normal wind speeds after calibration the sharpness diagrams (Figs. 6e,f) show more homogeneously populated bins for both regions. This means that the forecast system is able to predict those events with a larger range of forecast probability values. Conversely, the uncorrected and simple bias-corrected predictions display their frequency peaks near the climatological frequency so that they often predict the event with a climatological probability. These results show the improvement in the reliability of the predictions obtained when calibration is applied—improvements that are particularly relevant to the users.

5. Conclusions

Seasonal predictions have not yet been widely taken into account by the wind energy sector. Some applications in the energy sector of this type of forecasts have been recently identified, however. They illustrate that predictions at seasonal time scales can be used as input by the industry in decision-making processes to replace the current naive climatological information. In this paper we illustrate a strategy for the use of wind speed seasonal predictions by the wind energy sector.

After describing one of the most popular operational seasonal forecast systems, ECMWF's System 4, and its forecast-quality characteristics, two different bias adjustment techniques to correct the typical deficiencies of the predictions from global forecast systems are described. It is shown that bias adjustment is indispensable for the predictions to be usable. The System 4 predictions have skill in predicting wind speed at seasonal time scales, especially in the tropics, but also in extratropical regions of relevance to the wind energy sector. This is an encouraging result that has not been documented elsewhere. Dynamical seasonal predictions do suffer from a number of important systematic errors that also affect wind speed predictions. Bias adjustment methods are required for the predictions to have the same statistical properties as the observational reference and hence to be applicable by users. With regard to the bias adjustment, the simple bias correction and the calibration method produce predictions with statistical properties that allow their actual application. The most important gain in forecast quality for seasonal predictions comes through the increase in their skill and reliability, the latter being a critical aspect of the forecasts from the user perspective. These gains in forecast quality cannot be evidenced using correlation, which suggests that more than one measure of forecast quality is needed, even in a user context.

The predictions and the impact of the bias adjustment are illustrated on two skillful regions that are crucial for the wind energy sector, the North Sea and central Canada. A further analysis of the predictions reveals that both the bias correction and the calibration method produce an improvement in the consistency of the ensemble. In addition, the reliability diagrams demonstrate that the calibration method, which also corrects the deficiencies in the ensemble spread, provides more
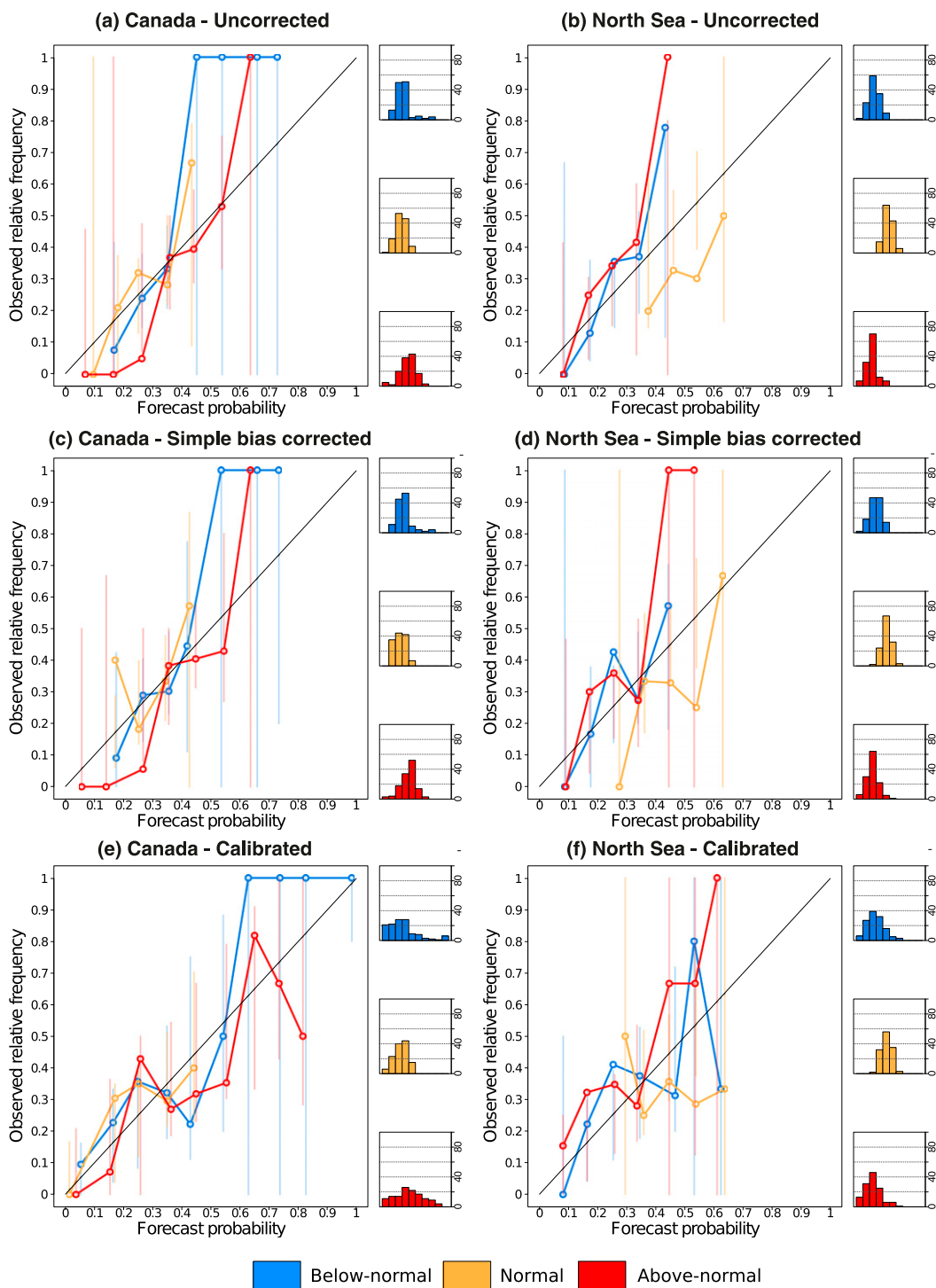
FIG. 6. Reliability diagrams of 10-m wind speed forecasts from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on 1 Nov for the period of 1981–2012. Three events are represented: above-normal wind speeds (red line), normal wind speeds (orange), and below-normal wind speeds (blue). To the right of each panel are shown the sharped diagrams with the distribution of samples for each bin and each event. The consistency bars have been represented as vertical lines to illustrate how likely the observed relative frequencies are under the assumption that predicted probabilities are reliable.

reliable predictions than does the simple bias correction technique. Improvements in reliability are fundamental from a user perspective because reliability guarantees the trustworthiness of the predictions.

Our work demonstrates that calibration is necessary because it produces an improvement in both skill and reliability, making this technique essential for seasonal predictions to be usable. The development of these strategies is part of a recent initiative undertaken by the climate community in which climate services are developed to provide more relevant, reliable, and action-oriented climate information (Buontempo et al. 2014). This paper illustrates the fact that seasonal predictions of near-surface wind speed have skill in several regions where there is substantial installed power and that after bias adjustment the predictions are reliable for use.

Future improvements include the combination of seasonal predictions from different sources, based on both dynamical and empirical–statistical forecast systems. The global and illustrative character of this paper requires the use of a reanalysis as reference data. Verification against other reanalyses and regional observed wind speed data might offer slightly different results because of the observational uncertainty, which is an additional factor that will be taken into account in future analyses, but the need for a bias adjustment process will be unavoidable. There are simple ways to convert the wind speed into energy density that will also be explored from the point of view of seasonal prediction, and the use of empirical downscaling could offer additional benefits when considering seasonal predictions for specific power plants.

The work described here opens the field to the next step in the development of a climate service: the creation of tailored products that facilitate the widespread use of climate predictions by the wind energy sector (step 4 in Fig. 1). The release of climate services can range from knowledge transfer (informing, documenting, and providing training in the best bias adjustment techniques) to the creation of operational online interactive interfaces that allow users in the wind industry to easily explore probabilistic predictions. An example of a prototype of an interactive platform that incorporates bias-adjusted predictions can be found at the Project Ukko (http://www.project-ukko.net) online interface designed in the framework of the EUPORIAS project. In addition, the New European Wind Atlas (NEWA; see online at http://euwindatlas.eu/), which is currently in development, will provide access to skill evaluations of climate predictions. Further interactions between the climate science community and the renewable energy community are also indispensable to quantify the actual economic value of climate predictions

and to evaluate how the predictions have performed in the past. This is a necessary step to demonstrate to energy stakeholders the saliency of climate forecast outcomes.

REFERENCES

Alessandrini, S., S. Sperati, and P. Pinson, 2013: A comparison between the ECMWF and COSMO ensemble prediction systems applied to short-term wind power forecasting on real data. *Appl. Energy*, **107**, 271–280, doi:10.1016/j.apenergy.2013.02.041.

Archer, C. L., and M. Z. Jacobson, 2013: Geographical and seasonal variability of the global ''practical'' wind resources. *Appl. Geogr.*, **45**, 119–130, doi:10.1016/j.apgeog.2013.07.006.

Barnston, A. G., and H. M. van den Dool, 1993: A degeneracy in cross-validated skill in regression-based forecasts. *J. Climate*, **6**, 963–977, doi:10.1175/1520-0442(1993)006<0963:ADICVS>2.0.CO;2.

——, M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.*, **93**, 631–651, doi:10.1175/BAMS-D-11-00111.1.

Bett, P. E., and H. E. Thornton, 2016: The climatological relationships between wind and solar energy supply in Britain. *Renewable Energy*, **87**, 96–110, doi:10.1016/j.renene.2015.10.006.

Bröcker, J., 2008: On reliability analysis of multi-categorical forecasts. *Nonlinear Processes Geophys.*, **15**, 661–673, doi:10.5194/npg-15-661-2008.

——, and L. A. Smith, 2007: Scoring probabilistic forecasts: The importance of being proper. *Wea. Forecasting*, **22**, 382–388, doi:10.1175/WAF966.1.

Buontempo, C., C. D. Hewitt, F. J. Doblas-Reyes, and S. Dessai, 2014: Climate service development, delivery and use in Europe at monthly to inter-annual timescales. *Climate Risk Manage.*, **6**, 1–5, doi:10.1016/j.crm.2014.10.002.

Cannon, D., D. Brayshaw, J. Methven, P. Coker, and D. Lenaghan, 2015: Using reanalysis data to quantify extreme wind power generation statistics: A 33 year case study in Great Britain. *Renewable Energy*, **75**, 767–778, doi:10.1016/j.renene.2014.10.024.

Chaudhry, N., and L. Hughes, 2012: Forecasting the reliability of wind-energy systems: A new approach using the *RL* technique. *Appl. Energy*, **96**, 422–430, doi:10.1016/j.apenergy.2012.02.076.

Clark, R. T., P. E. Bett, H. E. Thornton, and A. A. Scaife, 2017: Skilful seasonal predictions for the European energy industry. *Environ. Res. Lett.*, **12**, 024002, doi:10.1088/1748-9326/aa57ab.

Coelho, C. A. S., and S. M. S. Costa, 2010: Challenges for integrating seasonal climate forecasts in user applications. *Curr. Opin. Environ. Sustainability*, **2**, 317–325, doi:10.1016/j.cosust.2010.09.002.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.

De Felice, M., A. Alessandri, and F. Catalano, 2015: Seasonal climate forecasts for medium-term electricity demand forecasting. *Appl. Energy*, **137**, 435–444, doi:10.1016/j.apenergy.2014.10.030.

Doblas-Reyes, F. J., R. Hagedorn, and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting—II. Calibration and combination. *Tellus*, **57A**, 234–252, doi:10.3402/tellusa.v57i3.14658.

——, J. García-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, 2013: Seasonal climate predictability and forecasting: Status and prospects. *Wiley Interdiscip. Rev.: Climate Change*, **4**, 245–268, doi:10.1002/wcc.217.

Dunstone, N., D. Smith, A. Scaife, L. Hermanson, R. Eade, N. Robinson, M. Andrews, and J. Knight, 2016: Skilful predictions of the winter North Atlantic Oscillation one year ahead. *Nat. Geosci.*, **9**, 809–814, doi:10.1038/ngeo2824.

Elmore, K. L., 2005: Alternatives to the chi-square test for evaluating rank histograms from ensemble forecasts. *Wea. Forecasting*, **20**, 789–795, doi:10.1175/WAF884.1.

Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. *J. Appl. Meteor.*, **8**, 985–987, doi:10.1175/1520-0450(1969)008<0985:ASSFPF>2.0.CO;2.

Ferro, C. A. T., 2014: Fair scores for ensemble forecasts. *Quart. J. Roy. Meteor. Soc.*, **140**, 1917–1923, doi:10.1002/qj.2270.

Fricker, T. E., C. A. T. Ferro, and D. B. Stephenson, 2013: Three recommendations for evaluating climate predictions. *Meteor. Appl.*, **20**, 246–255, doi:10.1002/met.1409.

Füss, R., S. Mahringer, and M. Prokopczuk, 2015: Electricity derivatives pricing with forward-looking information. *J. Econ. Dyn. Control*, **58**, 34–57, doi:10.1016/j.jedc.2015.05.016.

García-Bustamante, E., J. F. González-Rouco, P. A. Jiménez, J. Navarro, and J. P. Montávez, 2009: A comparison of methodologies for monthly wind energy estimation. *Wind Energy*, **12**, 640–659, doi:10.1002/we.315.

Garcia-Morales, M. B., and L. Dubus, 2007: Forecasting precipitation for hydroelectric power management: How to exploit GCM's seasonal ensemble forecasts. *Int. J. Climatol.*, **27**, 1691, doi:10.1002/joc.1608.

Gershunov, A., and D. R. Cayan, 2003: Heavy daily precipitation frequency over the contiguous United States: Sources of climatic variability and seasonal predictability. *J. Climate*, **16**, 2752–2765, doi:10.1175/1520-0442(2003)016<2752:HDPFOT>2.0.CO;2.

Graff, M., R. Peña, A. Medina, and H. J. Escalante, 2014: Wind speed forecasting using a portfolio of forecasters. *Renewable Energy*, **68**, 550–559, doi:10.1016/j.renene.2014.02.041.

Hamlington, B. D., P. E. Hamlington, S. G. Collins, S. R. Alexander, and K.-Y. Kim, 2015: Effects of climate oscillations on wind resource variability in the United States. *Geophys. Res. Lett.*, **42**, 145–152, doi:10.1002/2014GL062370.

Hueging, H., R. Haas, K. Born, D. Jacob, and J. G. Pinto, 2013: Regional changes in wind energy potential over Europe using regional climate model ensemble projections. *J. Appl. Meteor. Climatol.*, **52**, 903–917, doi:10.1175/JAMC-D-12-086.1.

IPCC, 2007: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.

Jolliffe, I. T., and C. Primo, 2008: Evaluating rank histograms using decompositions of the chi-square test statistic. *Mon. Wea. Rev.*, **136**, 2133–2139, doi:10.1175/2007MWR2219.1.

——, and D. B. Stephenson, Eds., 2012: *Forecast Verification: A Practitioner's Guide in Atmospheric Science.* John Wiley and Sons, 292 pp.

Kirtman, B., and A. Pirani, 2009: The state of the art of seasonal prediction: Outcomes and recommendations from the First World Climate Research Program Workshop on Seasonal Prediction. *Bull. Amer. Meteor. Soc.*, **90**, 455–458, doi:10.1175/2008BAMS2707.1.

Lane, J.-E., 2016: Implementation success or failure: The COP21 agreement for the 21rest century. *J. Multidiscip. Eng. Sci. Technol.*, **2**, 158–166. [Available online at http://journalijcms.com/sites/default/files/issue-files/0076.pdf.]

Leung, L. R., A. F. Hamlet, D. P. Lettenmaier, and A. Kumar, 1999: Simulations of the ENSO hydroclimate signals in the Pacific Northwest Columbia River basin. *Bull. Amer. Meteor. Soc.*, **80**, 2313–2329, doi:10.1175/1520-0477(1999)080<2313:SOTEHS>2.0.CO;2.

Lynch, K. J., D. J. Brayshaw, and A. Charlton-Perez, 2014: Verification of European subseasonal wind speed forecasts. *Mon. Wea. Rev.*, **142**, 2978–2990, doi:10.1175/MWR-D-13-00341.1.

Mason, S. J., and O. Baddour, 2008: Statistical modelling. *Seasonal Climate: Forecasting and Managing Risk*, Springer, 163–201.

Molteni, F., and Coauthors, 2011: The new ECMWF seasonal forecast system (system 4). ECMWF Tech. Memo. 656, 51 pp. [Available online at http://www.ecmwf.int/sites/default/files/elibrary/2011/11209-new-ecmwf-seasonal-forecast-system-system-4.pdf.]

Najafi, A., H. Falaghi, J. Contreras, and M. Ramezani, 2016: Medium-term energy hub management subject to electricity price and wind uncertainty. *Appl. Energy*, **168**, 418–433, doi:10.1016/j.apenergy.2016.01.074.

Pinson, P., 2012: Adaptive calibration of (*u*, *v*)-wind ensemble forecasts. *Quart. J. Roy. Meteor. Soc.*, **138**, 1273–1284, doi:10.1002/qj.1873.

——, 2013: Wind energy: Forecasting challenges for its operational management. *Stat. Sci.*, **28**, 564–585, doi:10.1214/13-STS445.

——, and J. Tastu, 2013: Discrimination ability of the energy score. Technical University of Denmark Tech. Rep., 17 pp. [Available online at http://orbit.dtu.dk/files/56966842/tr13_15_Pinson_Tastu.pdf.]

——, H. A. Nielsen, H. Madsen, and G. Kariniotakis, 2009: Skill forecasting from ensemble predictions of wind power. *Appl. Energy*, **86**, 1326–1334, doi:10.1016/j.apenergy.2008.10.009.

Pryor, S. C., and R. J. Barthelmie, 2010: Climate change impacts on wind energy: A review. *Renewable Sustainable Energy Rev.*, **14**, 430–437, doi:10.1016/j.rser.2009.07.028.

Quan, X., M. Hoerling, J. Whitaker, G. Bates, and T. Xu, 2006: Diagnosing sources of U.S. seasonal forecast skill. *J. Climate*, **19**, 3279–3293, doi:10.1175/JCLI3789.1.

Renewable Energy Policy Network for the 21st Century, 2015: Renewables 2015 global status report: Key findings 2015. REN21 Tech. Rep., 33 pp. [Available online at http://www.ren21.net/wp-content/uploads/2015/07/GSR2015_KeyFindings_lowres.pdf.]

Reyers, M., J. G. Pinto, and J. Moemken, 2015: Statistical–dynamical downscaling for wind energy potentials: Evaluation

and applications to decadal hindcasts and climate change projections. *Int. J. Climatol.*, **35**, 229–244, doi:10.1002/joc.3975.

Rose, S., and J. Apt, 2015: What can reanalysis data tell us about wind power? *Renewable Energy*, **83**, 963–969, doi:10.1016/j.renene.2015.05.027.

Santos, J., C. Rochinha, M. Liberato, M. Reyers, and J. Pinto, 2015: Projected changes in wind energy potentials over Iberia. *Renewable Energy*, **75**, 68–80, doi:10.1016/j.renene.2014.09.026.

Scaife, A. A., and Coauthors, 2014: Skillful long-range prediction of European and North American winters. *Geophys. Res. Lett.*, **41**, 2514–2519, doi:10.1002/2014GL059637.

Schillings, C., T. Wanderer, L. Cameron, J. T. van der Wal, J. Jacquemin, and K. Veum, 2012: A decision support system for assessing offshore wind energy potential in the North Sea. *Energy Policy*, **49**, 541–551, doi:10.1016/j.enpol.2012.06.056.

Slingo, J., and T. Palmer, 2011: Uncertainty in weather and climate prediction. *Philos. Trans. Roy. Soc. London*, **A369**, 4751–4767, doi:10.1098/rsta.2011.0161.

Troccoli, A., 2010: Seasonal climate forecasting. *Meteor. Appl.*, **17**, 251–268, doi:10.1002/met.184.

Vaillancourt, K., and Coauthors, 2014: A Canadian 2050 energy outlook: Analysis with the multi-regional model TIMES-Canada. *Appl. Energy*, **132**, 56–65, doi:10.1016/j.apenergy.2014.06.072.

Vautard, R., F. Thais, I. Tobin, F.-M. Bréon, J.-G. D. de Lavergne, A. Colette, P. Yiou, and P. M. Ruti, 2014: Regional climate model simulations indicate limited climatic impacts by operational and planned European wind farms. *Nat. Commun.*, **5**, doi:10.1038/ncomms4196.

von Storch, H., and F. W. Zwiers, 2001: *Statistical Analysis in Climate Research*. Cambridge University Press, 496 pp.

Weisheimer, A., and T. N. Palmer, 2014: On the reliability of seasonal climate forecasts. *J. Roy. Soc. Interface*, **11**, 20131162, doi:10.1098/rsif.2013.1162.

——, ——, and F. J. Doblas-Reyes, 2011: Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophys. Res. Lett.*, **38**, L16703, doi:10.1029/2011GL048123.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd. ed. Elsevier, 676 pp.

World Wind Energy Association, 2015: New record in worldwide wind installations. *WWEA Bull.*, No. 1, 4–5. [Available online at http://www.wwindea.org/wwea-bulletin-issue-1-2015/.]