Seasonal climate prediction: a new source of information for the management of wind energy resources

Verónica Torralba *

Barcelona Supercomputing Center (BSC), Barcelona, Spain

Francisco J. Doblas-Reyes

Barcelona Supercomputing Center (BSC), Barcelona, Spain

Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Dave MacLeod

Atmospheric, Oceanic and Planetary Physcis, Deptartment of Physics, University of Oxford, Oxford, UK

Isadora Christel

Barcelona Supercomputing Center (BSC), Barcelona, Spain

Melanie Davis

Institut Català de Ciències del Clima, Barcelona (IC3), Barcelona, Spain

**Corresponding author address:* Verónica Torralba, Earth Sciences Department. Barcelona Supercomputing Center (BSC), C/Jordi Girona, 29, 08034 Barcelona, Spain.

E-mail: veronica.torralba@bsc.es

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. Traditionally, wind energy users employed a simple approach based on an estimate of a retrospective climatology. 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 similar statistical properties 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 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 measure the bias adjustment techniques are indispensable to produce statistically consistent and reliable predictions, particularly the calibration method. The forecast quality assessment shows that calibration is a fundamental requirement for a high-quality climate service.

1. Introduction

The demand for renewable energy sources as an alternative to fossil-fuel sources has increased due to 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; Solomon 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 the energy efficiency and replace the fossil fuels by the use of renewable energies (Lane 2016). Wind energy is the cheapest option for the new sources of power generating capacity and the second leading renewable energy source worldwide, only exceeded by hydropower in terms of installed capacity (Pryor and Barthelmie 2010; Santos et al. 2015). In recent years, wind power installed capacity has experienced a 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 wind power generation processes 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. 2013), 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 (from hours to a few days) time scales because near-surface winds and thus wind energy production, strongly depend on the meso- 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. However this information is not readily available to the relevant users, who have to rely on past information based on observations, and this is often only available as short time series. The need of climate information 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).

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

It has been shown recently that climate predictions are capable to provide additional value for wind energy applications, especially 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 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 made with a better anticipation of market changes. This framework will favor greater penetration of the renewable electricity into the markets.

While the scenario described is of great interest to the renewable energy community, little progress had been made in practice. However in recent years the skill of the climate predictions has significantly improved (Doblas-Reyes et al. 2013). For instance, seasonal prediction systems (i.e. those providing information for periods ranging from one month to slightly longer than one year into the future) are now providing skillful forecasts for extra-tropical 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 wind energy decision making as illustrated for different energy sources (De Felice et al. 2015; Garcia-Morales and Dubus 2007). Currently however there are very few instances of the application of seasonal predictions in the wind energy industry. Improved climate information which includes seasonal forecasts may change this, 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 if 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 post-processed to have similar statistical features as 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). The recent FP7 European projects on climate services EUPORIAS [1] and SPECS [2] have tried to address these challenges and support the development of sectorial climate services in Europe through the involvement of stakeholders in the definition of effective ways to develop climate information.

[1] http://www.euporias.eu/

[2] http://www.specs-fp7.eu/

This paper raises the limits associated with current seasonal prediction systems for their use in wind energy applications. It focuses on the description of appropriate bias adjustment techniques to overcome some of these limits and promote the use of the climate prediction information in those occasions in which it can provide greater accuracy than current approaches. The methodology described recognizes that end users must be provided with information about the prediction uncertainty (Alessandrini et al. 2013), so that 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 data sets and describes one of the most widely used seasonal prediction systems and its limitations. Section 3 describes appropriate bias adjustment techniques and introduces forecast quality assessment measures and explains their relevance in a user context. Section 4 presents the impact of the bias adjustments over the wind speed seasonal forecasts including an analysis of the changes in the statistical properties of the post-processed predictions. Finally, 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 the 10-m wind speed forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) System 4 (System 4) operational seasonal prediction system (Molteni et al. 2011), which is based on a global climate model, with coupled atmospheric and oceanic components. System 4 comprises of the 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, which is coupled to the ocean model

NEMO (Nucleus for European Modeling of the Ocean) version 3.0. The ocean model uses a grid with horizontal resolution of around 1° in the extratropics with equatorial refinement and 42 levels in the vertical. The atmosphere and ocean are coupled using a version of the OASIS3 (Ocean Atmosphere Sea Ice Soil) coupler developed at the CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique).

System 4 is run in ensemble prediction mode. Ensemble predictions are a way to deal with uncertainties in the climate system, in particular those associated with the imperfections of the initial conditions and in 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 sub-grid physical processes in the atmosphere. This allows the prediction of the forecast uncertainty (measured by the ensemble dispersion), along with the prediction itself. The simulations are performed for up to seven months into the future.

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

The analysis in this paper focuses on the boreal winter as the winter season has larger wind speed variability in the Northern Hemisphere (Archer and Jacobson 2013).. In addition, the analysis of the seasonal predictions of wind speed in winter can be relevant due to the higher variability of wind power supply in that particular season (Bett and Thornton 2015). This illustrates the potential of seasonal predictions for end users as they potentially have more impact where the inter-annual variability is the largest, although other seasons have also been analyzed (Fig. S1) and the conclusions apply equally. The predictions considered here are those issued on the 1st of November, for which three-month statistics for the December-January-February (DJF, also known as one-month lead seasonal forecast) period 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 when 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 Integrated Forecast System (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 data set. 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 with the lack of long enough historical data needed have also promoted the use of reanalyses by the wind industry (Rose and Apt 2015).

For this reason, and being aware that reanalysis estimates could often be far from point observed values, the reanalysis has been used as the best available estimate of wind speed. The choice of reanalysis is arbitrary and the conclusions are equally valid when using other reanalysis, both global or regional. Further work is needed to assess the seasonal predictions for specific wind farm locations.

3. Methodology

a. Data Processing

The wind speed forecasts are affected by biases resulting from the inability numerically reproduce all the relevant processes 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 additional difficulties appear such as the lack of forecast reliability (Pinson 2012), which 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. The bias adjustment of the wind speed has been identified as a requirement of the wind energy sector to fulfill acceptable reliability requirements to be used in their decision-making processes (Alessandrini et al. 2013).

This paper illustrates the relative merits of different techniques for the 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

The simple bias correction is based on the assumption that both the reference and predicted distributions of seasonal wind speed, are well approximated by a Gaussian (normal) distribution. The adjustment creates predictions with the same mean and standard deviation as the reference data set. This is a zero-order approach for the 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:

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

Seasonal mean anomalies are calculated by subtracting the ensemble mean of the seasonal averages ($\bar{x}$) from the seasonal average of each forecast (x_{ij}) for each year i and for each 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 data set (σ_{ref}) to the interannual standard deviation of the ensemble members (σ_e), and adding the climatology of the reference data set ($\bar{o}$). This is done for each grid cell separately, resulting in a new wind speed forecast ensemble, with 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 interannual variance equivalent to that of a reference data set in a similar way to the bias correction method, but at the same time ensuring an increased reliability of the probability predictions. Here we apply the variance inflation technique (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 with 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:

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

$$\beta = \sqrt{1-\rho^2}\frac{\sigma_{ref}}{\sigma_e}. \tag{4}$$

The σ_{em} is the standard deviation of the ensemble mean (the time series of xi), 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 data set. The α and β coefficients are found under two constraints. The former is that the standard deviation of the inflated prediction is the same as that for the reference and the latter is that the predictable signal after the inflation is made equal to the correlation of the ensemble mean with the reference data set.

b. Forecast quality assessment

Seasonal forecast systems, as in any other forecasting process, have to be systematically compared to 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 to the prediction problem because a prediction has no value without an estimate of its quality based on past performance (Doblas-Reyes et al. 2013). Moreover the quantification of the uncertainty is one of the most crucial aspects for the successful development of wind industry and the 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, insufficient number of forecast cases, and imperfect reference values due to 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.

The reader should 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 where 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 predictions system as well as the impact of the two bias adjustment techniques over the 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 where 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

The skill estimates based on the performance of the system in the past, may guide users about the expected performance of the future forecasts (Weisheimer and Palmer 2014), always with the caveat that the predictability of the climate system might change over time. Skill scores are a tool for end users to develop alternative strategies to their baseline information to minimize the risk and to perform an 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 data set 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.

However the bias adjustment techniques that are illustrated in this paper (defined in section 3.a) 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 post-processed forecast changes relative to the correlation computed directly with the uncorrected forecasts.

A comprehensive measure of the predictive skill for the 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, that in our case has been taken as the climatology (made of all 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 (RPS), a measure of the squared distance between the forecast and the reference cumulative probabilities. In the present case the RPSS has been computed based on categorical forecasts for terciles. Three equi-probable events associated with the two terciles of the climatological distribution of the reference: wind speed exceeding the upper tercile (above normal category), not exceeding the lower tercile (below normal category) and values 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 data set in the verification time series can fall in any of the three categories with probability determined by the probability density function (PDF) for the target season.

The continuous ranked probability skill score (CRPSS) is a commonly used probabilistic skill score (Jolliffe and Stephenson 2012) that has been used to evaluate the predictive skill of the full probability distribution. It is based on the continuous ranked probability score (CRPS), 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, measuring the performance of a forecast relative to the climatology.

The RPSS and CRPSS range between 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, up 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 favours predictions with ensemble members that perform as if they have been sampled from the same distribution than the reference dataset. The fair version of the RPSS and CRPSS have been used in order 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 and the basic scores are small as has been shown for the RPSS in the supplementary material (Fig. S2).

2) Reliability

Reliability analysis of prediction systems remains as a prime concern for the wind energy sector, as for any user of probability predictions, due to 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 know if the forecast is assumed to be reliable and then it is expected to be flat. However, some deviations from uniformity can appear for reliable forecasts due to randomness. The rank histograms have been displayed on probability paper (Bröcker 2008). In the y-axis rank histograms display cumulative probabilities instead of the traditional observed frequency which indicate how probable that observed frequency would be if the prediction was reliable. This information is useful to identify if the deviations from a 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 percent simultaneous confidence intervals have been represented.

Rank histograms illustrate if the ensemble members and the verifying observation come from the same probability distribution, in which case the forecasts are statistically consistent then no calibration of the ensemble is needed. This happens when the rank histogram is flat (as if coming from a uniform distribution). However, because of sampling variations the histograms are almost never flat. To assess if the deviations from flatness are attributed to chance or deficiencies in the forecasts, goodness-of-fit test statistics are computed: 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/over-dispersion (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 draw a line as closely 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 due to the limited samples typical of seasonal forecast systems (Jolliffe and Stephenson 2012). To deal with this problem we have inclded 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 (ten 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 sharpness diagram. Sharpness gives an indication of the variation in forecast probabilities issued by the prediction system, independently 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, where one assesses the features of both the forecast system and the statistical model that transforms the ensemble into probabilities.

4. Results

Total wind power installed indicates the wind power capacity available in each wind farm. It has been represented in Fig. 2 to identify which are the most important locations from a wind energy user point of view. To illustrate the performance of the seasonal predictions two key regions for the wind energy sector because wind farms are located there have been selected. For the selection of the regions we have also taken into account the potential skill available in such regions (Fig. S1). The first region is in Canada [longitude:112.5°-113.2°W and latitude: 50.3°-51.0°N]. This country is an important player in terms of energy resources (Vaillancourt et al. 2014) and 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 new installed capacity (Canadian Wind Energy Association 2015) that year. The North Sea region [longitude: 9.8°-10.6°E and latitude: 58.0°-58.7°] is the second region considered. It is the most important region for offshore energy activities in Europe due to 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).

Fig. 3 displays the predictions for the uncorrected, bias corrected and calibrated sets for these two regions. The effect of the bias adjustment over the predictions is that when the corrections are applied, the hindcasts (grey dots) show similar mean and variance to the reference data set (black dots). 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 the 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, variance) of the predicted variables with respect to those in the observations as well as 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 forecast quality measure, 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. In order to explore how the size of the region affects the forecast quality we have estimated the forecast quality for larger regions (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 non-trivial task because the bias adjustment techniques necessary in seasonal forecasting require long-enough observational references that are not readily available.

The forecast system considered allows estimating the global forecast quality of the different sets of predictions. The fair RPSS maps for the uncorrected, bias corrected and calibrated wind speed are shown in Fig. 4. The uncorrected predictions (Fig. 4 (a)) display very low scores all around the world. The highest values are found in tropical regions, in particular in some regions of North East of South America and North Western Africa. This maximum can be explained because the largest predictability at seasonal timescales is attributed to anomalies in the tropical sea surface temperatures (SST) 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.

Fig. 4 (b) and (c) show that the fair RPSS increases globally when bias adjustment is applied. This kind of assessments are widely available for variables like temperature and precipitation, but are not available for wind speed. The skill improvement has been quantified in the Fig. 4 (d) and (e), which indicate that the skill scores for the bias adjusted predictions increase more than 1 relative to the uncorrected ones. The fair RPSS maps (Fig. 4 (b) and (c)) for the post-processed 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 as the North Sea or Scandinavia display positive values. Wind speed predictions show the highest skill in Northern Europe, while in Southern Europe negative RPSS values is found. This is in agreement with previous work (e.g. Weisheimer et al. 2011) indicating that seasonal dynamical predictions have limited forecast quality over Europe.

The skill improvement is also present in South-eastern Asia, central United States or North-eastern 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 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 the Fig. S3 and S4. The correlation of the uncorrected forecasts is always higher due to the cross-validation leakage mentioned above. It is noticeable that the correlation spatial distribution in the calibrated hindcasts is noisier than the two other types of forecasts considered. This 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 (Fig. 5 (a) and (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 (Figure 5 (a)) where all the observations are exceeded by the majority of the ensemble members, leaving the highest rank categories almost empty. The bias corrected and calibrated forecasts show more homogeneously populated ranks indicating that the reliability of the ensemble improves when the bias adjustment is applied. However, 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. 5 (e)), the rank 50 shows a very large value that might indicate that the ensemble overestimates the true uncertainty range.

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

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 under-dispersive, as evidenced by the high JP convex test. The high values of the JP slope show that the forecasts are also affected by biases. The uncorrected forecasts in the North Sea have also biases and are under-dispersive, although the statistical tests have smaller values than those in Canada. The results are statistically significant, with the p-values being virtually zero.

The tests applied to the simple bias corrected and the calibrated forecasts indicate that the deviation from flatness is minimised when the bias adjustment is applied. The Pearson χ^2 for the calibrated data in the Canada region has higher values than the bias corrected ones (p-value 0.01), while the JP tests provide no evidence of departures from flatness with p-values higher than 0.01. Consequently this result shows that the biases and the under-dispersion in the raw ensemble are corrected, and the deviations from uniformity are independent of these specific problems. Making sure that the ensemble is well calibrated, which 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 analyse the impact of the bias adjustment on reliability, reliability diagrams (Fig. 6) allow the comparison between the observed frequencies with 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 a 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 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 overconfident. For the uncorrected forecasts in Canada (Fig. 6 (a)), 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 the observed below-normal wind speeds. The reliability diagram for the uncorrrected predictions in the North Sea (Fig. 6 (b)) shows only a narrow set of probabilities issued, with values ranging from 0.1 to 0.5 for the above (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 falls outside the consistency bars. This illustrates the poor reliability for that event in the North Sea when the predictions are uncorrected.

The reliability curves of the bias corrected predictions (Fig. 6 (c,d)) show similar features to 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 (Fig. 6 (e,f)) have reliability diagrams with their points lying closer to the diagonal than found for the uncorrected and bias corrected predictions. This 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. In the North Sea (Fig. 6 (f)) the slope of the curve for the normal category (orange line) (Fig. 6 (f)) becomes horizontal suggesting that the system can not 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 those 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 (Fig. 6 (e,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 predict often 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. However, some applications in the energy sector of this type of forecasts have been recently identified. 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 of 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. However, dynamical seasonal predictions 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 of the observational reference and hence to be applicable by the users. Concerning the bias adjustment, the simple bias correction and the calibration methods produce predictions with statistical properties that allow their actual application. The most important gain in forecast quality for the seasonal predictions comes through the increase in their skill and reliability, the latter 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 forecast quality measure 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 calibration methods produce an improvement in the consistency of the ensemble. Besides, the reliability diagrams demonstrate that the calibration method, which also corrects the deficiencies in the ensemble spread, provides more reliable predictions than the simple bias correction technique. Improvements in reliability are fundamental from a user perspective because it 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 the seasonal predictions to be usable. The development of these strategies is part of a recent initiative undertaken by the climate community where 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 their 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. The 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 of a bias adjustment process will be unavoidable. Finally, there are simple ways to convert the wind speed into energy density that will be explored from the seasonal prediction point of view, while 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 to allow wind industry user easily explore probabilistic predictions. An example of a prototype of interactive platform that incorporates bias-adjusted predictions can be found at Project Ukko[3] interface designed in the framework of

[3] http://www.project-ukko.net

the EUPORIAS project. In addition, the New European Wind Atlas (NEWA[4]), which is currently in development will provide access to skill evaluations of climate predictions. Further interactions between the climate science community and renewable energy community are also indispensable to quantify the actual economic value of climate predictions and evaluate the predictions performance in the past. This is a necessary step to demonstrate to energy stakeholders the saliency of climate predictions outcomes.

Acknowledgments. The authors acknowledge funding support from the RESILIENCE (CGL2013-41055-R) project, funded by the Spanish Ministerio de Economía y Competitividad (MINECO) and the FP7 EUPORIAS (GA 308291) and SPECS (GA 308378) projects. Special thanks to Nube Gonzalez-Reviriego and Albert Soret for helpful comments and discussion. We also acknowledge the COPERNICUS action CLIM4ENERGY-Climate for Energy (C3S 441 Lot 2) and the New European Wind Atlas (NEWA) project funded from ERA-NET Plus, topic FP7-ENERGY.2013.10.1.2. We acknowledge the s2dverification[5] and SpecsVerification[6] R-based packages. Finally we would like to thank Pierre-Antoine Bretonnière, Oriol Mula and Nicolau Manubens for their technical support at different stages of this project.

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. *Applied 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. *Applied Geography*, **45**, 119–130, doi:10.1016/S0031-8914(53)80099-6.

[4]http://euwindatlas.eu/

[5]http://cran.r-project.org/web/packages/s2dverification

[6]http://cran.r-project.org/web/packages/SpecsVerfication

Barnston, A. G., 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? *Bulletin of the American Meteorological Society*, **93 (5)**, 631–651, doi:10.1175/BAMS-D-11-00111.1.

Barnston, A. G., and H. M. van den Dool, 1993: A degeneracy in cross-validated skill in regression-based forecasts. *Journal of Climate*, **6 (5)**, 963–977, doi:10.1175/1520-0442(1993)006<0963:ADICVS>2.0.CO;2.

Bett, P. E., and H. E. Thornton, 2015: 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 in Geophysics*, **15 (4)**, 661–673, doi:10.5194/npg-15-661-2008.

Bröcker, J., and L. A. Smith, 2007: Scoring probabilistic forecasts: The importance of being proper. *Weather and Forecasting*, **22 (2)**, 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 Management*, **6**, 1–5, doi:10.1016/j.crm.2014.10.002.

Canadian Wind Energy Association, 2015: Installed Capacity. URL http://canwea.ca/wind-energy/installed-capacity/.

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. *Applied 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. *Environmental Research Letters*, **12 (2)**, 024 002.

Coelho, C. A. S., and S. M. S. Costa, 2010: Challenges for integrating seasonal climate forecasts in user applications. *Current Opinion in Environmental Sustainability*, **2 (5)**, 317–325, doi:10.1175/JCLI3801.1.

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

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597, doi:10.1002/qj.828.

Doblas-Reyes, F. J., J. García-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, 2013: Seasonal climate predictability and forecasting: status and prospects. *Wiley Interdisciplinary Reviews: Climate Change*, **4 (4)**, 245–268, doi:10.1002/wcc.217.

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 A*, **57 (3)**, 234–252, doi:10.1111/j.1600-0870.2005.00104.x.

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. *Nature Geoscience*.

Elmore, K. L., 2005: Alternatives to the chi-square test for evaluating rank histograms from ensemble forecasts. *Weather and forecasting*, **20 (5)**, 789–795, doi:10.1175/WAF884.1.

Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. *Journal of Applied Meteorology*, **8 (6)**, 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. *Quarterly Journal of the Royal Meteorological Society*, **140 (683)**, 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. *Meteorological Applications*, **20 (2)**, 246–255, doi:10.1016/j.renene.2011.05.033.

Füss, R., S. Mahringer, and M. Prokopczuk, 2013: Electricity derivatives pricing with forward-looking information. *University of St. Gallen, School of Finance Research Paper*, doi:10.2139/ssrn.2219855.

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 (7)**, 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. *International Journal of Climatology*, **27 (12)**, 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. *Journal of Climate*, **16 (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. *Geophysical Research Letters*, **42 (1)**, 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. *Journal of Applied Meteorology and Climatology*, **52 (4)**, 903–917.

Jolliffe, I. T., and C. Primo, 2008: Evaluating rank histograms using decompositions of the chi-square test statistic. *Monthly Weather Review*, **136 (6)**, 2133–2139, doi:10.1175/2007MWR2219.1.

Jolliffe, I. T., and D. B. Stephenson, 2012: *Forecast verification: a practitioner's guide in atmospheric science*. John Wiley & Sons.

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. *Bulletin of the American Meteorological Society*, **90 (4)**, 455–458, doi:10.1175/2008BAMS2707.1.

Lane, J.-e., 2016: Implementation success or failure: the cop21 agreement for the 21rest century. *Journal of Multidisciplinary Engineering Science and Technology*, **2**, 158–166.

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. *Bulletin of the American Meteorological Society*, **80 (11)**, 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. *Monthly Weather Review*, **142 (8)**, 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)*. European Centre for Medium-Range Weather Forecasts.

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

Pinson, P., 2012: Adaptive calibration of (u, v)-wind ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, **138 (666)**, 1273–1284, doi:10.1002/qj.1873.

Pinson, P., 2013: Wind energy: Forecasting challenges for its operational management. *Statistical Science*, **28 (4)**, 564–585, doi:10.1002/qj.1873.

Pinson, P., H. A. Nielsen, H. Madsen, and G. Kariniotakis, 2009: Skill forecasting from ensemble predictions of wind power. *Applied Energy*, **86 (7-8)**, 1326–1334, doi:10.1016/j.apenergy.2008.10.009.

Pinson, P., and J. Tastu, 2013: Discrimination ability of the Energy score. Tech. rep., Technical University of Denmark.

Pryor, S. C., and R. J. Barthelmie, 2010: Climate change impacts on wind energy: A review. *Renewable and sustainable energy reviews*, **14 (1)**, 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 US seasonal forecast skill. *Journal of climate*, **19 (13)**, 3279–3293, doi:10.1175/JCLI3789.1.

Renewable Energy Policy Network for the 21st Century, 2015: Renewables 2015 Global Status Report. URL http://www.ren21.net/status-of-renewables/global-status-report/{\}.

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. *International Journal of Climatology*, **35 (2)**, 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. *Geophysical Research Letters*, **41 (7)**, 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. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **369 (1956)**, 4751–4767, doi:10.1098/rsta.2011.0161.

Solomon, S., 2007: *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*, Vol. 4. Cambridge University Press.

Troccoli, A., 2010: Seasonal climate forecasting. *Meteorological Applications*, **17 (3)**, 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. *Applied 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. *Nature communications*, **5**, doi:10.1038/ncomms4196.

Von Storch, H., and F. W. Zwiers, 2001: *Statistical analysis in climate research*. Cambridge university press.

Weisheimer, A., and T. N. Palmer, 2014: On the reliability of seasonal climate forecasts. *Journal of The Royal Society Interface*, **11 (96)**, 20131 162, doi:10.1098/rsif.2013.1162.

Weisheimer, A., T. N. Palmer, and F. J. Doblas-Reyes, 2011: Assessment of representations of model uncertainty in monthly and seasonal forecast ensembles. *Geophysical Research Letters*, **38 (16)**, doi:10.1029/2011GL048123.

Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences*, Vol. 100. Academic press.

World Wind Energy Association, 2015: WWEA Bulletin Issue 1 – 2015. URL http://www.wwindea.org/wwea-bulletin-issue-1-2015/{\}.

LIST OF TABLES

	Uncorrected		Simple bias corrected		Calibrated	
	Canada	North Sea	Canada	North Sea	Canada	North Sea
Pearson χ^2	2080.37	148.25	36.12	57.25	77.56	68.62
p value	0.00	2.01×10^{-11}	0.94	0.25	0.01	0.05
JP slope	291.69	38.46	0.17	0.1	0.03	0.12
p value	2.12×10^{-65}	0.56×10^{-11}	0.68	0.75	0.87	0.72
JP convex	319.69	11.32	1.15	1.77	0.73	3.07
p value	1.69×10^{-71}	0.76×10^{-04}	0.28	0.18	0.39	0.06

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 1981-2012.

LIST OF FIGURES

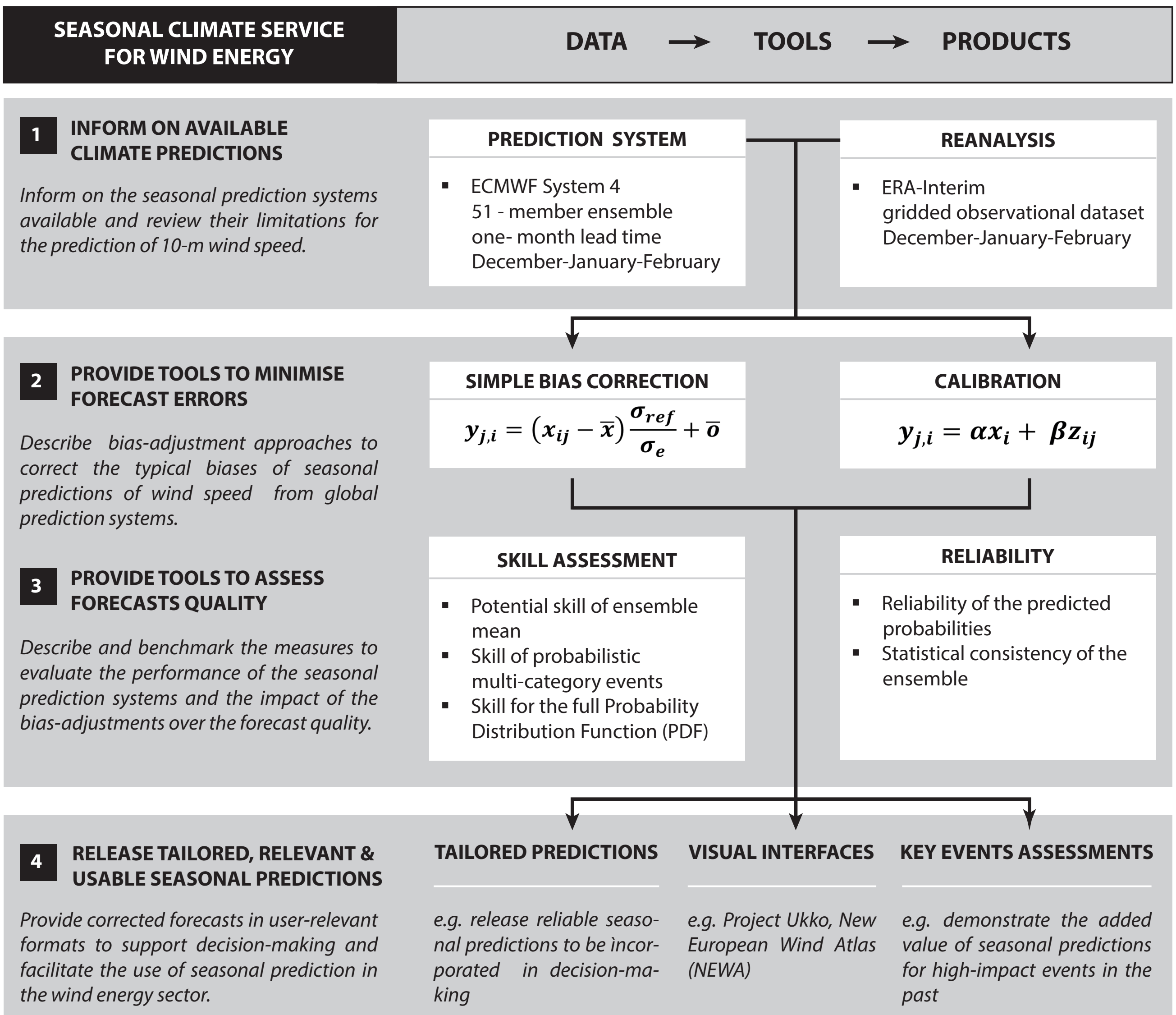

FIG. 1. Main steps for the development of a climate service for the wind energy sector based on seasonal climate predictions. Steps 2 and 3 in the diagram outline the main challenges addressed in this paper.

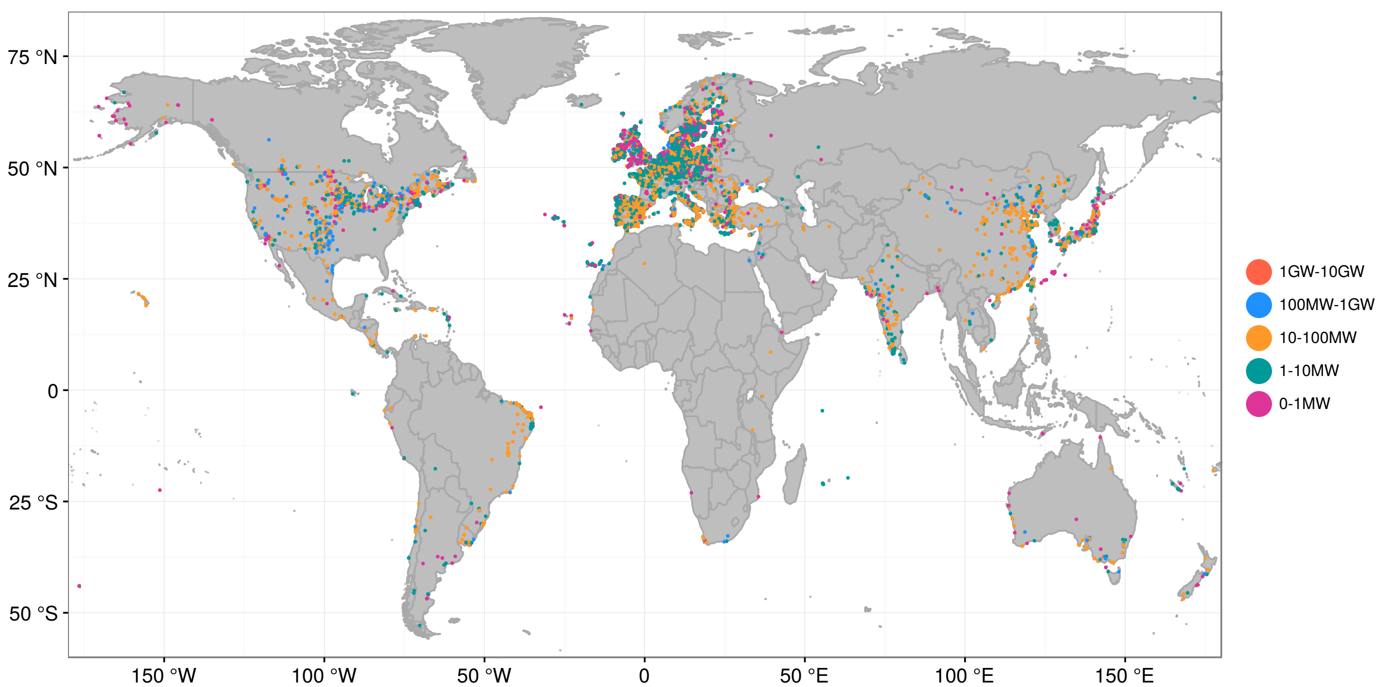

FIG. 2. Total installed wind power capacity for each individual wind farm (operational and under construction have been included) in 2015 (Source:www.thewindpower.net).

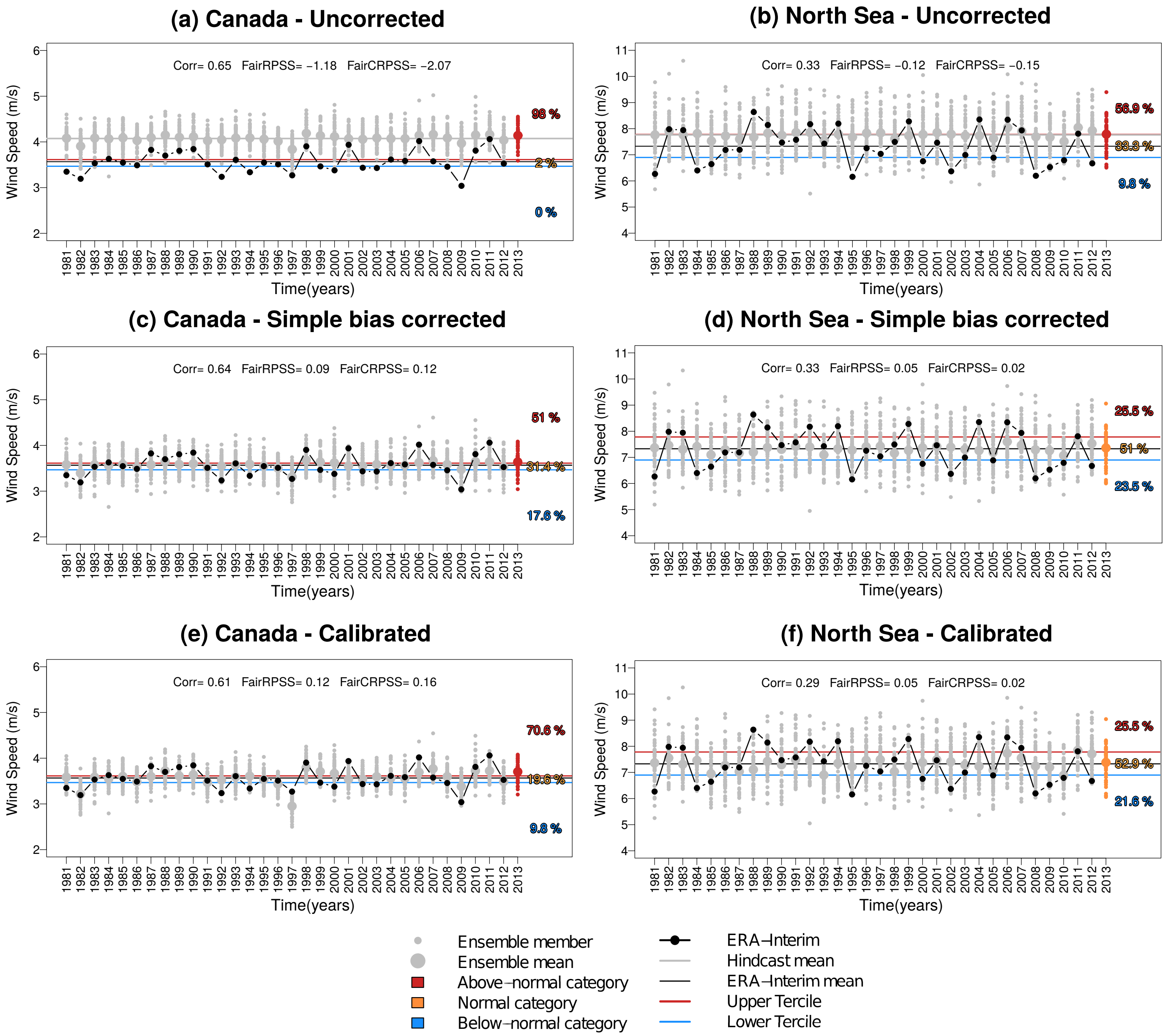

FIG. 3. Time series of 10-m wind speed from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on the first of November for the period of 1981-2013. The ensemble members of the hindcasts are represented as small grey dots and the ensemble mean is represented with a large grey dot for each start date. The grey horizontal line shows the mean of the hindcast in 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 are 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 in the 1981-2012 period. Correlation, RPSS and CRPSS are shown in the upper part of each panel.

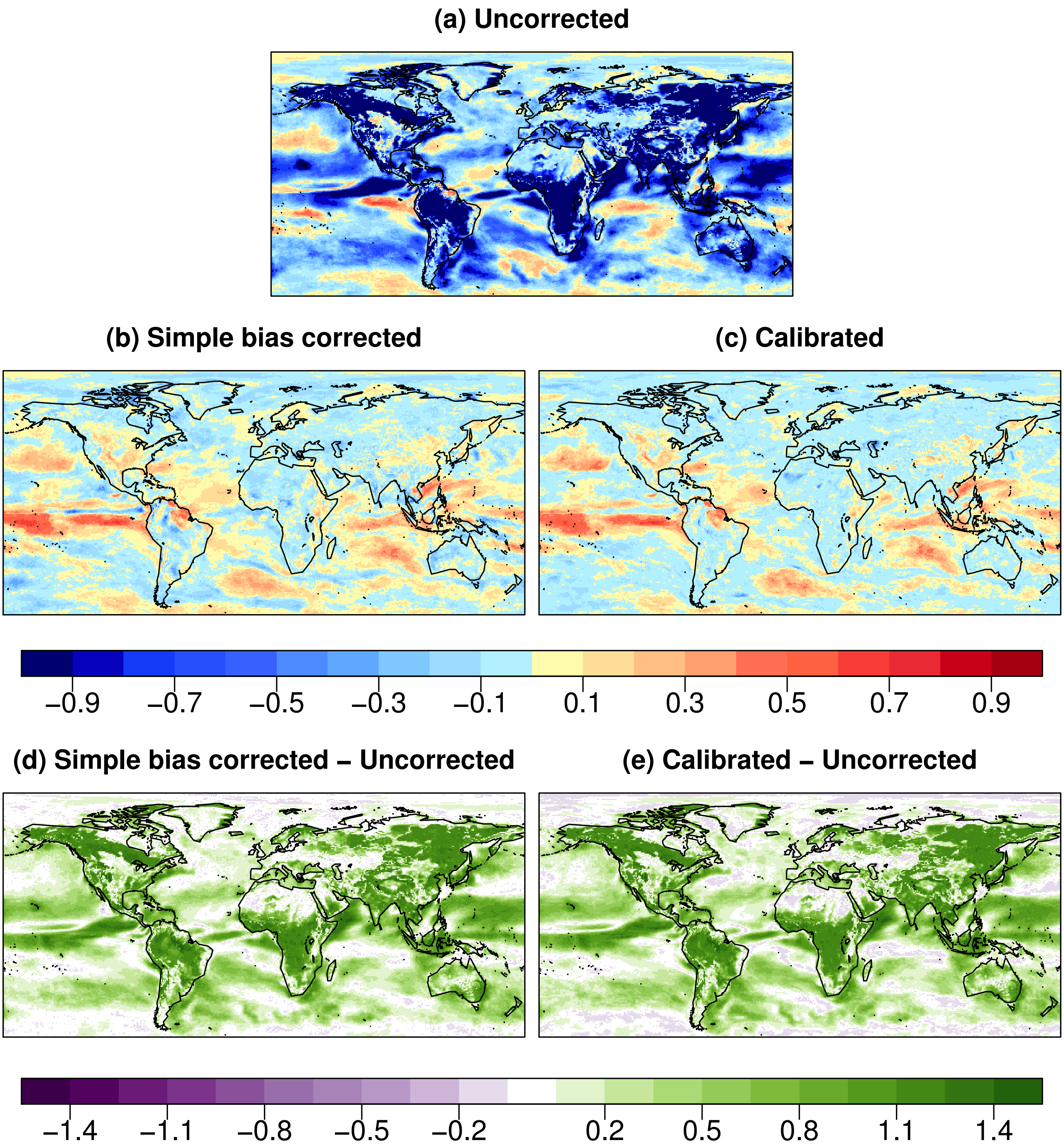

FIG. 4. Fair Ranked Probability Skill Score (RPSS) for tercile events of 10-m wind speed forecasts from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on the first of November for the period of 1981-2012.

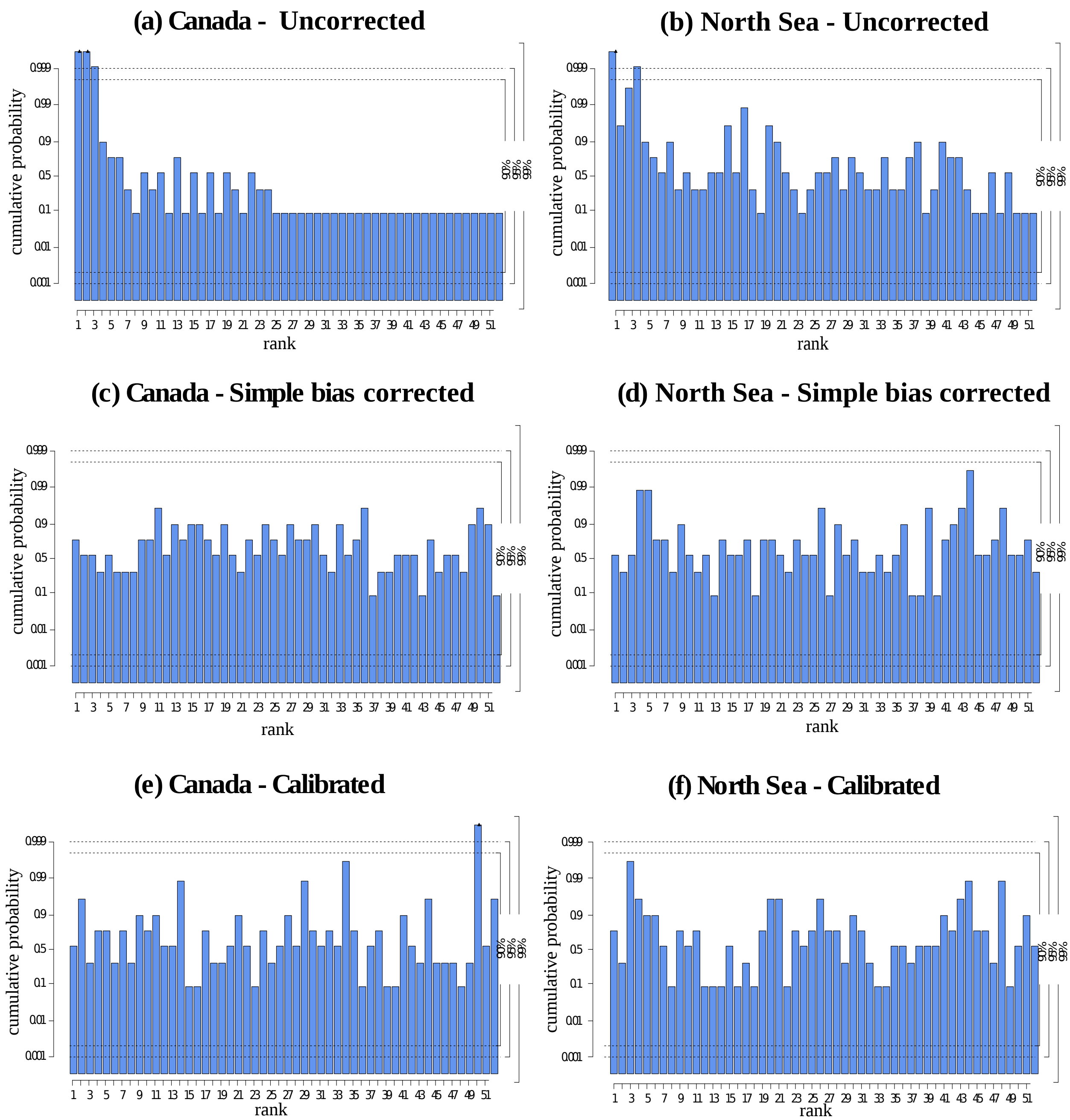

FIG. 5. Rank histograms of 10-m wind speeds forecasts from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on the first of November for the period of 1981-2012. These rank histograms have been represented on probability paper to show if the deviations from a 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 in the y-axis. On the right 90, 95 and 99 percent simultaneous confidence intervals are indicated. If all ranks were equally likely on average, approximately 90 percent of all rank histogram would be contained in the 90 percent confidence interval and approximately 10 percent of all rank histograms would have at least one bar that falls outside this interval.

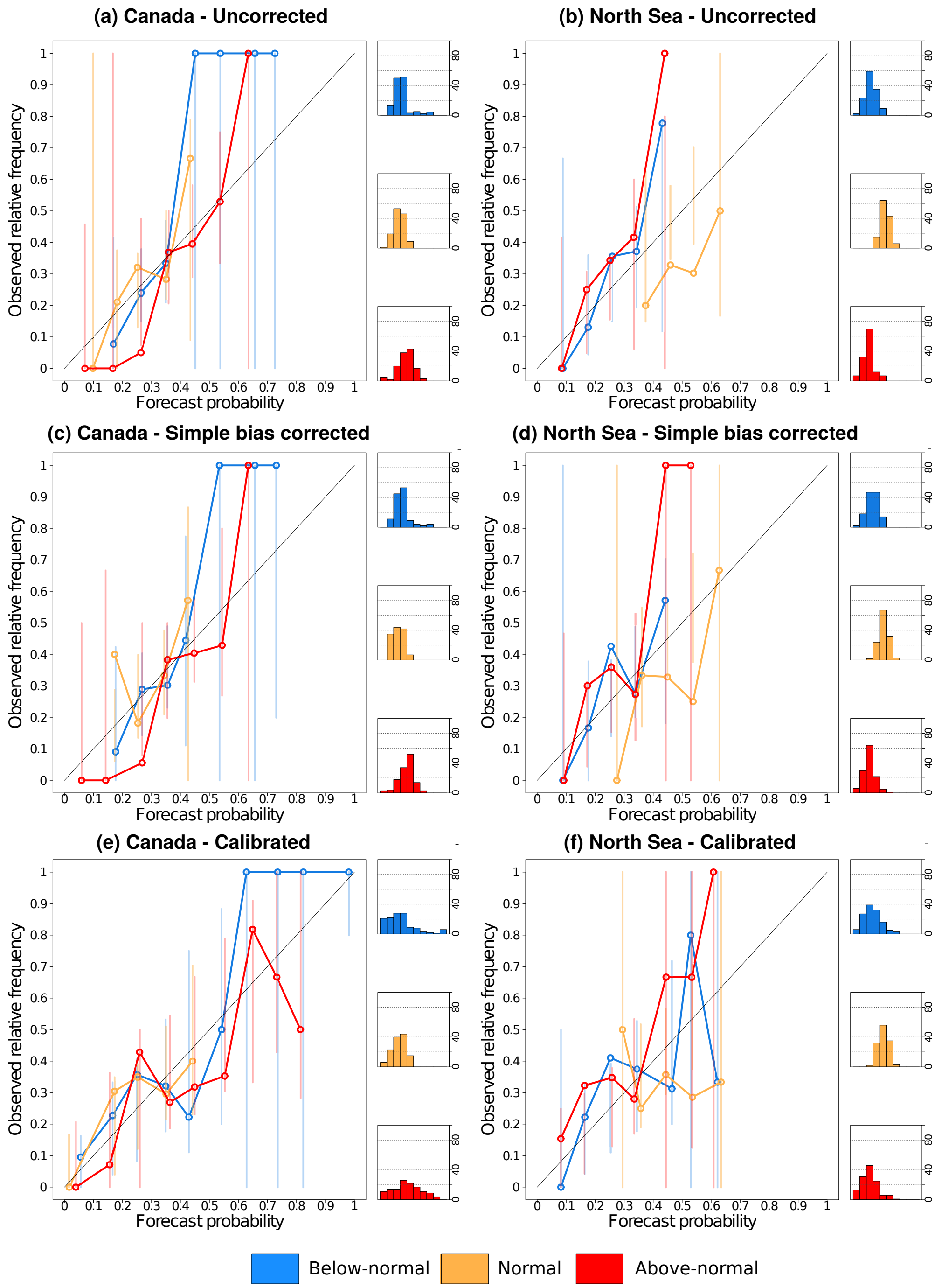

FIG. 6. Reliability diagrams of 10-m wind speeds forecasts from ECMWF System 4 and ERA-Interim reanalysis in winter (DJF). These predictions have been initialized on the first of November for the period of 1981-2012. Three events are represented: above-normal wind speeds (red line), normal wind speeds (orange) and below-normal wind speeds (blue). Right panels show 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.