AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2017JD026533

**Key Points:**
- Uncertainty of an ensemble prediction system was estimated
- Ensemble spread was used as an estimator of model predictability
- Several tools were developed to facilitate decision making

**Correspondence to:**
S. Fernández-González,
sefern04@ucm.es

**Citation:**
Fernández-González, S., M. L. Martín, A. Merino, J. L. Sánchez, and F. Valero (2017), Uncertainty quantification and predictability of wind speed over the Iberian Peninsula, *J. Geophys. Res. Atmos.*, *122*, 3877–3890, doi:10.1002/2017JD026533.

Received 23 JAN 2017
Accepted 27 MAR 2017
Accepted article online 29 MAR 2017
Published online 12 APR 2017



# Uncertainty quantification and predictability of wind speed over the Iberian Peninsula

**S. Fernández-González[1], M. L. Martín[2], A. Merino[3], J. L. Sánchez[3], and F. Valero[1]**

[1]Department of Earth Physics, Astronomy and Astrophysics II, Faculty of Physics, Complutense University of Madrid, Madrid, Spain, [2]Department of Applied Mathematics, Faculty of Computer Engineering, University of Valladolid, Segovia, Spain, [3]Atmospheric Physics Group, IMA, University of León, León, Spain

**Abstract** During recent decades, the use of probabilistic forecasting methods has increased markedly. However, these predictions still need improvement in uncertainty quantification and predictability analysis. For this reason, the main aim of this paper is to develop tools for quantifying uncertainty and predictability of wind speed over the Iberian Peninsula. To achieve this goal, several spread indexes extracted from an ensemble prediction system are defined in this paper. Subsequently, these indexes were evaluated with the aim of selecting the most appropriate for the characterization of uncertainty associated to the forecasting. Selection is based on comparison of the average magnitude of ensemble spread (ES) and mean absolute percentage error (MAPE). MAPE is estimated by comparing the ensemble mean with wind speed values from different databases. Later, correlation between MAPE and ES was evaluated. Furthermore, probability distribution functions (PDFs) of spread indexes are analyzed to select the index with greater similarity to MAPE PDFs. Then, the spread index selected as optimal is used to carry out a spatiotemporal analysis of model uncertainty in wind forecasting. The results indicate that mountainous regions and the Mediterranean coast are characterized by strong uncertainty, and the spread increases more rapidly in areas affected by strong winds. Finally, a predictability index is proposed for obtaining a tool capable of providing information on whether the predictability is higher or lower than average. The applications developed may be useful in the forecasting of wind potential several days in advance, with substantial importance for estimating wind energy production.

## 1. Introduction

Worldwide installed wind power capacity has risen from 3 to 200 GW over the past 20 years. However, a reduction of wind energy costs is necessary to be more competitive relative to traditional energy sources. To achieve this objective, accurate wind power forecasts are needed for adequate integration into the electric grid system [*Draxl et al.*, 2014]. In contrast with other energy sources, wind potential is limited by strong spatiotemporal variations. As a result, wind energy production fluctuates more than those of traditional energies (e.g., fossil fuel sources). Therefore, wind power forecasting requires uncertainty quantification. In this regard, numerical weather prediction models have been used [*Ellis et al.*, 2015].

In recent decades, numerical weather prediction models have experienced great improvement, largely owing to the advancement of computing. Until the late twentieth century, weather forecasts were developed from a deterministic approach. In this method, from a particular combination of physical parameterizations and specific initial conditions considered to be optimal, a weather forecast was obtained, which was taken as the most accurate estimate [*Gneiting and Raftery*, 2005].

During recent years, efforts have focused on improving probabilistic techniques in order to consider forecast uncertainties [*Thorarinsdottir and Gneiting*, 2010; *Messner et al.*, 2014; *Scheuerer and Möller*, 2015]. Probabilistic forecasting can be achieved by considering initial values, boundary conditions, and/or model-related uncertainty [*Toth*, 2001]. The most common technique involves the perturbation of initial conditions, with the aim of estimating uncertainty related to model assimilation [*Hamill et al.*, 2000]. This technique is based on the assertion that the main source of model errors at midlatitudes is the unstable growth of initial errors, as opposed to model deficiencies [*Reynolds et al.*, 1994]. This technique has been used by the European Centre for Medium-Range Weather Forecasts (ECMWF) [see *Molteni et al.*, 1996] and National Centers for Environmental Prediction [*Toth and Kalnay*, 1997] for their global forecast models.

Estimates of the uncertainties associated with prediction can be attained by using ensemble prediction systems (EPS). Thus, the development of probabilistic approaches, in which forecasting is complemented by uncertainty bands, has advanced in recent decades [*Keune et al.*, 2014]. EPS have proved very useful for increasing confidence in predictions, particularly in extreme weather situations such as heavy rainfall or strong winds. However, EPS tend to produce underdispersion and overconfidence, and therefore unreliable forecasts [*Bowler et al.*, 2009; *Ho et al.*, 2013]. This can lead to underestimation of the risks associated with extreme events.

Given the probabilistic nature of EPS, it is possible to define different moments of the generated prediction distribution. One of the most common tools to quantify uncertainty associated with forecasting is to use the spread between ensemble members [*Grimit and Mass*, 2007; *Hopson*, 2014; *Fortin et al.*, 2014]. In this way, establishing a relationship between greater ensemble spread (ES) and more uncertainty can facilitate decision making.

To characterize ES, different statistical parameters have been used. The most common is the standard deviation, but variance, mean absolute deviation, and others have also been used [*Dey et al.*, 2014; *Keune et al.*, 2014; *Satterfield and Bishop*, 2014]. As long as the ES is strongly correlated with estimated prediction error, the spread can be used as an estimator of model predictability. Thus, less ES means greater predictability [*Saetra et al.*, 2004; *Jerez et al.*, 2013]. Regardless of the parameter selected, it must be related to the prediction error to be effective in the estimation of uncertainty [*Grimit and Mass*, 2002].

Several methods have been proposed to quantify model predictability, such as comparing the ratio between root-mean-square error (RMSE) of the forecast and climatology [*Griffies and Bryan*, 1997]. Concerning ensemble forecasts, predictability can be evaluated from the relationship between standard deviation of forecast errors and ensemble variance [*Wang and Bishop*, 2003]. Another option consists in relating predictability to ES, such that large (small) spread leads to low (high) predictability [*McMurdie and Ancell*, 2014]. *Schneider and Griffies* [1999] defined a new metric, called predictive power, which estimates forecast uncertainty by comparing a climatological probability distribution function (PDF) with those of the ensemble forecast. In this approach, when the forecast PDF is wider than its climatological PDF, strong uncertainty is expected, and vice versa.

Within the New European Wind Atlas (NEWA) project framework, several methods have been analyzed to quantify uncertainty associated with wind speed forecasts by several atmospheric probabilistic models. In the present work, the main aim was to find an optimal spread index that is able to summarize information provided by different EPS members and to estimate predictability of the forecast. To achieve this goal, ECMWF EPS forecasts are selected for an EPS database. Then, comparisons between EPS operational forecasts and two different databases are made. Several spread indexes, based on different percentile ranges of the ES, are defined. In order to select the best spread index among all those defined in this paper, we have calculated the correlation between the observed error and each of the spread indexes. Furthermore, we have also compared the PDFs of the mean absolute percentage error (MAPE) and the dispersion indexes for having different parameters with which to justify the choice of the best spread index. In addition, a spatiotemporal analysis of the spread is done, which can be very useful in the detection of areas with variable sensitivity to prediction errors and for examination of the spread increase rate with forecast lead time. Finally, a predictability index (PI) is developed for evaluating forecast reliability over a specific region and lead time. As a result, a method for estimating wind-resource uncertainties is developed by relating the spread index magnitude for a particular forecast with its average value from the preceding 30 days.

The paper is organized as follows. Section 2 includes a description of the databases used. The methodology for obtaining the spread indexes and PI is presented in section 3. In section 4, results and a discussion are presented. Finally, concluding remarks are provided in section 5.

## 2. Database

### 2.1. ECMWF EPS

Operational 50-member forecasts of the ECMWF EPS at T639 L62 resolution were used in our research. These data were interpolated to a horizontal grid with a resolution of 0.25° × 0.25°. Ensemble data were obtained via The Observing System Research and Predictability Experiment and Interactive Grand Global Ensemble


**Figure 1.** Orography of study area.

database [*Bougeault et al.*, 2010]. The ECMWF creates an ensemble of initial conditions by adding and subtracting potential errors (known as singular vectors) to and from a global analysis [*Buizza*, 1997].

The entire EPS domain was restricted to a regular grid with a horizontal resolution of 0.25° (both latitude and longitude) covering the entire Iberian Peninsula (35.6°N, 14.2°W; 45°N, 6°E), composed of 5406 grid points (Figure 1). From this grid, we extracted meteorological fields of *u* and *v* wind components at 10 m above ground level, from which wind speed at 10 m was calculated (m $s^{-1}$). This variable was obtained for the 50 members of the ECMWF EPS, in order to calculate the ensemble mean and ES. Additionally, to compare this database with that of the Deutscher Wetterdienst (DWD), daily mean wind speed forecast by the ECMWF model were also derived for the selected domain.

Operational forecasts (produced twice daily) from 1 January to 31 December 2010 were analyzed. The forecast horizon was limited to 144 h (focusing on the first 6 days), with a temporal resolution of 12 h. Therefore, we used simulations with 13 time steps (+0, +12, +24, +36, +48, +60, +72, +84, +96, +108, +120, +132, and +144 h).

### 2.2. DWD Data Set

The new high-resolution DWD data set, which was developed for regional model validation, consists of a gridded daily-mean wind speed at 10 m above ground (data only available over land surfaces). It covers most of Europe, including the entire Iberian Peninsula. This database is freely available online from the DWD Climate Data Centre, under the framework of the DecReg/Miklip project. The horizontal resolution of this database is ~0.044° (5 × 5 km) but was interpolated to a 0.25° × 0.25° grid for comparison with the ECMWF EPS.

The grid was obtained by spatial interpolation of meteorological station data in Europe. The interpolation is based on the regression kriging method developed by *Krähenmann et al.* [2011] but also involves predictor fields of exposure, roughness length, and coastal distance. Interpolation uncertainty of the database was calculated by means of the kriging variance and regression uncertainties. In addition, cross validation was performed to assess the quality of the final daily gridded data. An adequate quality control was applied to this database to remove erroneous data and inhomogeneities. Further information on this database is in *Brinckmann et al.* [2015].

### 2.3. ERA-Interim Reanalysis

Because the DWD database has no wind data over the sea surface, it was decided to use an additional database: the ERA-Interim reanalysis. ERA-Interim is the latest global atmospheric reanalysis produced by the ECMWF. It is generated by using a sequential data assimilation scheme. This involves the computation of variational analysis of basic upper air atmospheric fields (temperature, wind, humidity, and surface pressure), followed by separate analyses of near-surface parameters (2 m temperature and humidity), soil moisture and temperature, snow, and ocean waves. These data were used to initialize a short-range model

forecast. Use of the model equations makes it possible to extrapolate information from locally observed parameters to unobserved regions in a physically meaningful way. Further explanation of this database is in *Dee et al.* [2011]. The ERA-Interim database was obtained on a regular grid with 0.25° spacing.

To compare with data of the ECMWF EPS, the same grid and spatial and temporal resolutions were selected for the year 2010. Moreover, the same meteorological fields were extracted, i.e., *u* and *v* wind components (m $s^{-1}$) at 10 m above the surface. This database has been used for verification purposes [*Newman et al.*, 2015] and comparison with ECMWF EPS forecasts [*Komaromi and Majumdar*, 2014].

## 3. Methodology

The spread is a parameter widely used to quantify the uncertainty associated with EPS predictions [*Buizza et al.*, 1999; *Saetra et al.*, 2004; *Grimit and Mass*, 2007; *Fortin et al.*, 2014]. Currently, there are many different methods for selecting a spread index, using distinct mathematical parameters [*Dey et al.*, 2014; *Satterfield and Bishop*, 2014]. In fact, it is possible that a spread index is optimal for a particular model and/or region and a specific meteorological field but does not accommodate other purposes [*Hopson*, 2014].

*Chai and Draxler* [2014] recommended the use of mean absolute error instead of RMSE for validation purposes since the latter is not a good parameter to estimate average model performance. In our study, an error indicator in percentage was used to compare regions with different wind potential: the mean absolute percentage error (MAPE). Thus, the MAPE between an ensemble mean and DWD and ERA-Interim databases was calculated for selecting the optimal spread index. This measure of forecast accuracy has been compared with other estimators with satisfactory results [*Chun et al.*, 2016]. The MAPE is calculated by

$$\mathrm{MAPE} = \frac{100}{N}\sum_{i=1}^{n}\left|\frac{O_i - P_i}{O_i}\right|$$

where *O* indicates the observed values, *P* refers to predicted values, and *i* denotes each of grid points and temporal steps.

For MAPE estimation, both the ensemble mean and median were tested. Finally, the ensemble mean was selected because of the slightly larger MAPE obtained with use of the ensemble median. The median reduces the influence of extremely large and/or small values (outliers), which can be detrimental during extreme wind episodes. However, information provided by extreme wind values estimated by any ensemble member is contained within the ensemble mean. We did not use the ensemble mode in the verification because it is more useful for nonnumerical data, such as precipitation type [*Wilks*, 1995].

We rejected standard deviation for use as a spread index, because this statistical measure is more convenient for variables that fit a normal distribution [*Tracton and Kalnay*, 1993]. Wind data are expected to fit a distribution positively skewed with a long tail to the right, such as the Weibull distribution [*Koh et al.*, 2011]. For this reason, the use of spread indexes based on percentile ranking of data may be useful.

The total spread index ($SI_{total}$) is defined as the maximum difference between any pair of ensemble simulations [*Jerez et al.*, 2013]. Here several symmetric percentile ranges ($q_{xx}$) were predefined for determining which was the most similar to the magnitude and probability distribution function of the estimated relative error. In addition, two nonsymmetric indexes ($SI_{85-25}$ and $SI_{95-25}$) have been proposed in order to compare their efficiency against symmetric indexes. To compare regions of variable wind potential, the spread indexes were normalized for obtaining spread values as percentages. As a result, the following spread indexes were defined.

$$SI_{total} = \left(\frac{\text{ensemble maximum} - \text{ensemble minimum}}{\text{ensemble mean}}\right) \times 100$$

$$SI_{95-05} = \left(\frac{q_{95} - q_{05}}{\text{ensemble mean}}\right) \times 100 \qquad SI_{90-10} = \left(\frac{q_{90} - q_{10}}{\text{ensemble mean}}\right) \times 100$$

$$SI_{85-15} = \left(\frac{q_{85} - q_{15}}{\text{ensemble mean}}\right) \times 100 \qquad SI_{80-20} = \left(\frac{q_{80} - q_{20}}{\text{ensemble mean}}\right) \times 100$$

$$SI_{75-25} = \left(\frac{q_{75} - q_{25}}{\text{ensemble mean}}\right) \times 100 \qquad SI_{70-30} = \left(\frac{q_{70} - q_{30}}{\text{ensemble mean}}\right) \times 100$$

$$SI_{85-25} = \left(\frac{q_{85} - q_{25}}{\text{ensemble mean}}\right) \times 100 \qquad SI_{95-25} = \left(\frac{q_{95} - q_{25}}{\text{ensemble mean}}\right) \times 100$$

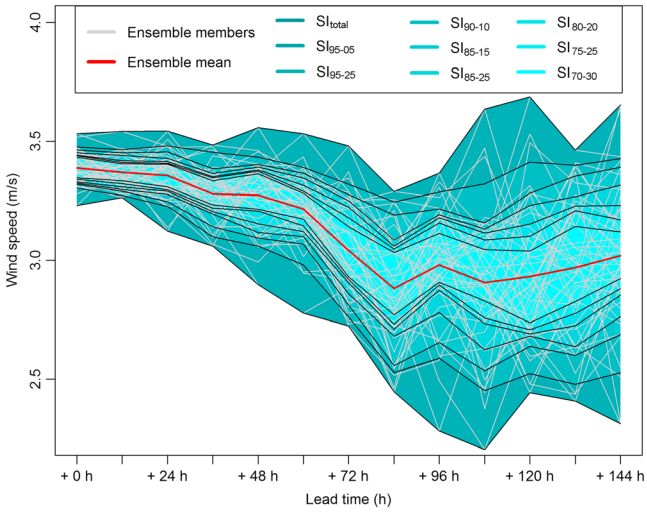


**Figure 2.** Example of ECMWF EPS forecast for specific grid point. Fifty ensemble members are represented by grey lines, ensemble mean by a red line, and distinct spread indexes are shaded.

Then, correlation coefficient ($r$) was used for evaluating the spatial correlation between the different spread indexes and the MAPE. This methodology has been used previously for similar purposes with satisfactory results [*Grimit and Mass*, 2002].

Finally, a predictability index (PI) was developed for estimation of forecast reliability for a specific region and lead time. Several PIs have been defined in recent years as a "relative measure of predictability" [*Toth et al.*, 2001], which can identify forecasts with above or below average uncertainty. In this paper, the PI is defined as subtraction of the average value over the previous 30 days by the spread index value for a particular day. As seen in section 4, the spread index selected is obtained from the interquartile range ($SI_{75-25}$), so this index is used for estimation of the PI. Values of the PI are calculated independently for each grid point and lead time. PI values $>0$ refer to less spread than the average, so the forecast meeting that criterion is more reliable than usual. In contrast, PI values $<0$ indicate considerable uncertainty in the forecast.

$$\mathrm{PI} = \mathrm{SI}_{75-25}\ (\mathrm{average}) - \mathrm{SI}_{75-25}\ (\mathrm{forecast})$$

It should be noted that for validation purposes, only wind speed values over 1.5 m $s^{-1}$ were considered in order not to alter the estimation of MAPE. This is connected to the fact that low-wind episodes can be related to high percentage errors, causing incoherent extremely high MAPE values. Therefore, wind speed values lower than 1.5 m $s^{-1}$ (i.e., levels 0 and 1 of the Beaufort scale) are not considered in the validation. It was checked that an underestimation of the forecast error was not induced by excluding these points since MAPE values were almost identical after excluding wind speed values over 1.5 m $s^{-1}$.

## 4. Results and Discussion

### 4.1. Selection of an Optimal Spread Index

This section describes results of optimal spread index selection for the EPS over the Iberian Peninsula. Figure 2 shows a forecast of the ECMWF EPS at a specific grid point in the study domain. This ensemble is composed of 50 members (grey lines) generated by the perturbation of initial conditions, as explained in section 2. Then, the ensemble mean is estimated as the average wind speed of the 50 ensemble members (red line). Several percentile ranges of ES (described in section 3) are shaded in Figure 2. Thus, the control forecast (ensemble mean) is complemented by uncertainty bands.

*Ancell* [2013] claimed that the best representation of total spread is to consider all the ensemble members, because all variability of the EPS is taken into account. However, outliers sometimes appear, causing unrealistic values. It may be desirable to filter these out to avert disproportionate uncertainty estimates. This supports the use of ES percentile ranges.

Selection of the most suitable percentile range may vary with the meteorological information required by the end user. For instance, maximum values of the spread can be useful in the estimation of extreme wind episodes. This issue is examined in section 4.3. However, we selected a balanced spread index, trying to minimize the difference between the MAPE estimated by the ensemble mean and the spread index.

**Table 1.** Correlation Coefficient (*r*) Obtained With Different Spread Indexes

| | $SI_{total}$ | $SI_{95-25}$ | $SI_{95-05}$ | $SI_{85-25}$ | $SI_{90-10}$ | $SI_{85-15}$ | $SI_{80-20}$ | $SI_{75-25}$ | $SI_{70-30}$ |
|---|---|---|---|---|---|---|---|---|---|
| *r* | 0.241 | 0.242 | 0.247 | 0.249 | 0.250 | 0.251 | 0.252 | 0.253 | 0.253 |

The first and most important requirement of an optimal spread index is the existence of a significant correlation between the spread index and the MAPE. For this reason, correlation coefficient (*r*) has been estimated for each spread index. The values of *r* have been calculated for each day, being the mean values of the whole period exposed in Table 1. These values are in the same order of magnitude of those calculated by *Grimit and Mass* [2002] during nonextreme spread episodes.

Although the *r* values are relatively low, they are statistically significant for all the spread indexes due to the huge number of degrees of freedom in this research (5406 grid points). For instance, the *P* value obtained for the $SI_{95-25}$ is equal to $2.2 \times 10^{-16}$, markedly significant for a 0.05 significance level. This fact means that the null hypothesis (the two variables are uncorrelated) can be rejected. In other words, there is evidence that the variables are significantly related, and therefore, the first requirement is achieved.

When comparing the correlation calculated with each spread index, it can be concluded that it increases by removing the outliers, since the correlation improves in the spread indexes in which highest and lowest percentiles have been removed ($SI_{75-25}$ and $SI_{70-30}$). The use of nonsymmetric indexes ($SI_{85-25}$ and $SI_{95-25}$) does not seem to improve correlation to MAPE.

According to *Buizza et al.* [1999], mean error of the control forecast (ensemble mean in this case) must be similar to the average spread of the index. Therefore, one of the requirements of the optimal spread index is that its magnitude must be similar to that of the MAPE. The goal is such that the SI can efficiently capture the main part of the model error [*Kieu et al.*, 2014]. ES has been compared with forecast error to establish a connection between ES and predictability [*Hamill and Colucci*, 1998; *Hou et al.*, 2001]. *Kolczynski et al.* [2011] pointed out that an ensemble is perfectly calibrated if the magnitude of the spread index and ensemble mean error has a 1:1 ratio.

The histogram of Figure 3 shows average values for the entire domain and lead times of the various spread indexes analyzed. The red and black horizontal lines indicate the MAPE obtained from the DWD and ERA-Interim databases, respectively. It is clearly seen that the spread index most similar to the MAPE magnitude (for both ERA-Interim and DWD databases) is $SI_{75-25}$. Therefore, this would be the most appropriate index based on comparison of the error magnitude with the dispersion index. The magnitude of MAPE is lower in nonsymmetric indexes ($SI_{85-25}$ and $SI_{95-25}$) than its respective symmetric indexes ($SI_{85-15}$ and $SI_{95-05}$) but higher than those of $SI_{75-25}$.

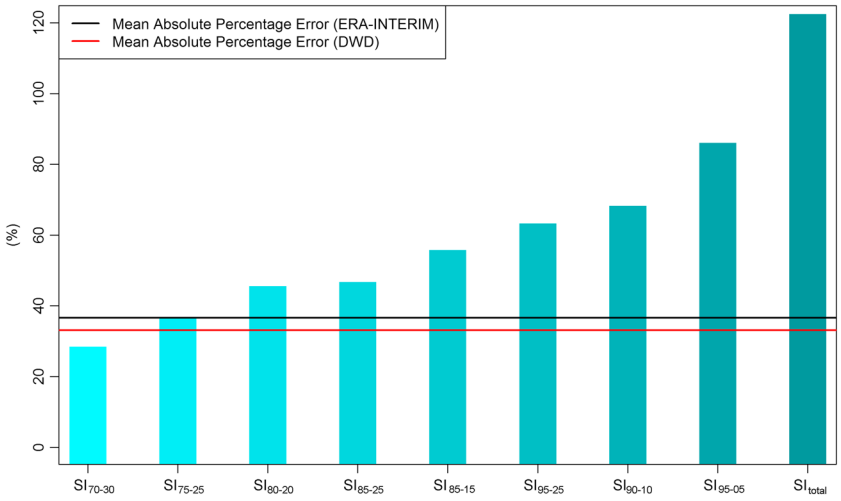


**Figure 3.** Average values (%) of several spread indexes and MAPE during 2010.

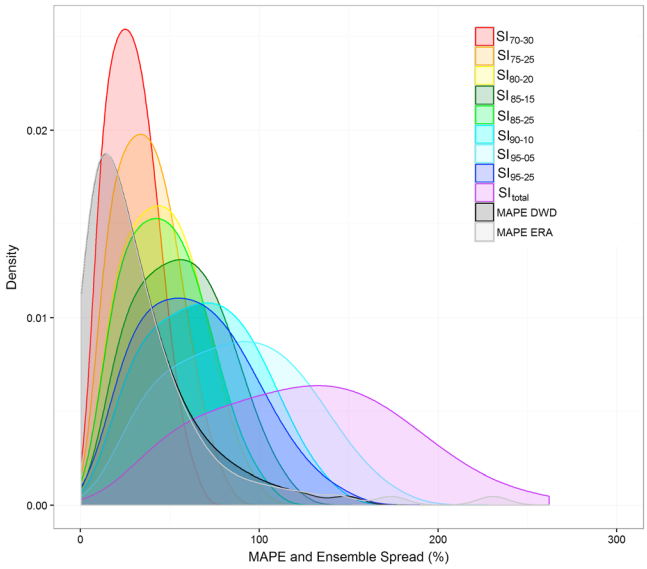


**Figure 4.** Comparison of spread indexes and MAPE PDFs.

By choosing the dispersion index that most closely matches the magnitude of the ensemble mean error, it is very likely that the spread magnitude is indicative of the magnitude of forecast error. According to *Grimit and Mass* [2002], this implies that large-spread episodes are less predictable; on the contrary, if the spread is smaller than usual, a more reliable forecast is expected. This matter is discussed further in section 4.3.

Another issue to keep in mind is that the spread index PDF should be similar to those of the prediction error [*Buizza et al.*, 1999]. Similarly, *Keune et al.* [2014] justified the analysis of the probability distributions to identify those statistics that characterize the distribution. From a statistical standpoint, an index is a numerical indicator that should summarize the information stored in a data set [*Jerez et al.*, 2013]. Thus, the spread index will provide significant information about the probability distribution estimated by a probabilistic system. According to *Zhu et al.* [2002], the PDF based on an ensemble can be centered closer to truth than the distribution of a single high-resolution forecast due to its capability of filtering nonlinear errors, especially at longer lead times. Therefore, we considered convenient to introduce a third requirement that consist in comparing the PDFs of MAPE and spread indexes. In this way, the choice of the optimal spread index will be more robust.

Figure 4 shows the PDF of preselected spread indexes, together with PDFs of both estimated MAPEs (those obtained by comparing the ensemble mean with the DWD and ERA-Interim databases). Both the MAPE and spread index PDFs are characterized by positive skewness. That implies that PDFs have their maximum frequency (their mode) to the left of the median, whereas the arithmetic mean is located further right. The PDFs are also characterized by a long tail at their right. Although the MAPEs obtained from both databases are slightly shifted to the left, the spread index with the greatest similarity is $SI_{75-25}$. The shape of nonsymmetric spread indexes PDFs is more similar to those of MAPE PDFs, although their maximum density is lower than desirable. As seen in Figure 4, maximum density of the $SI_{75-25}$ PDF has the same magnitude for both MAPE PDFs.

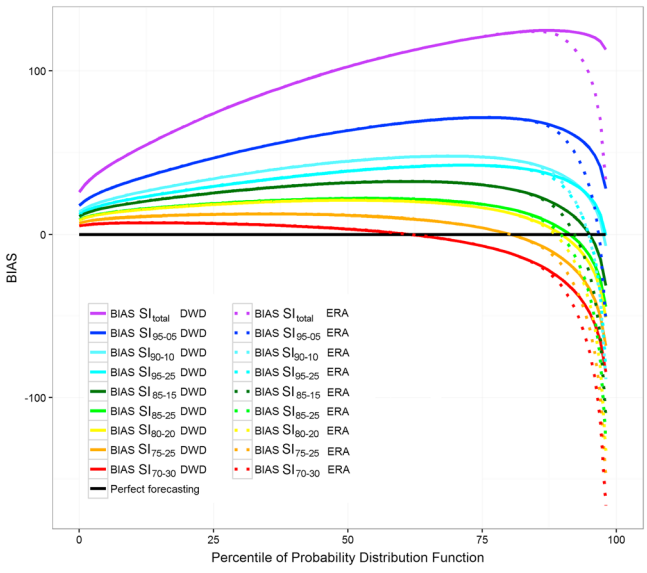


**Figure 5.** Bias of distinct spread index PDFs regarding MAPE PDF (for both ERA-Interim and DWD data sets).

With the goal of clearly visualizing deviation from the expected value, the bias was calculated by subtracting values of the MAPE (for both DWD and ERA-Interim) from those of the various spread indexes for each PDF percentile. As shown in Figure 5, bias values were very large when considering the PDF of $SI_{total}$. This

**Table 2.** Mean Bias Values Obtained With Different Spread Indexes

| (%) | $SI_{total}$ | $SI_{95-05}$ | $SI_{90-10}$ | $SI_{95-25}$ | $SI_{85-15}$ | $SI_{85-25}$ | $SI_{80-20}$ | $SI_{75-25}$ | $SI_{70-30}$ |
|---|---|---|---|---|---|---|---|---|---|
| BIAS | 92.59 | 54.66 | 37.31 | 32.87 | 26.02 | 18.71 | 18.05 | 12.97 | 11.49 |

means that the spread is much greater than the average error, so there is great overdispersion. Therefore, removing outliers of the EPS by selecting a specific percentile range of the ensemble members appears reasonable. As lower percentile ranges are selected, the overdispersion decreases, and there is even underdispersion for the higher PDF percentiles.

Because the results of using the DWD and ERA-Interim databases are very similar, bias averages for both databases were derived (Table 2). The best results were obtained by using $SI_{70-30}$. Nevertheless, this spread index is characterized by underdispersion from the 62nd percentile, which could cause nondetection of extreme wind events. Therefore, it is valuable to use a more balanced index. This is the case for $SI_{75-25}$, which does not show underestimation until above the 75th percentile, with a bias very similar to the previous case. However, the elimination of the ensemble's highest percentiles may lead to the underestimation of strong winds. For this reason, the use of $SI_{80-20}$, $SI_{85-15}$, $SI_{90-10}$, $SI_{85-25}$, and $SI_{95-25}$ may be advisable if the aim is focused on strong wind episodes, because these indexes can retain information about such episodes. However, the main disadvantage of these indexes is that they have overdispersion during the episodes located below the 75th percentile of wind speed climatology. The use of $SI_{95-05}$ and $SI_{total}$ is discouraged because of large bias.

Previous works [*Quine*, 2000; *Lun and Lam*, 2000] have demonstrated that any wind speed data set (regardless of location) can be fit by a Weibull distribution. To check if the spread indexes defined in this paper fit to a Weibull distribution, we constructed Figure 6, which shows a qqplot of quantiles from a Weibull distribution on the *y* axis and quantiles from $SI_{75-25}$ on the *x* axis. In this figure, the expected Weibull distribution is marked by a red line, and the estimated $SI_{75-25}$ distribution is represented by black dots. As seen, the adjustment is satisfactory over the entire distribution.

Although the qqplot allows visual determination that the $SI_{75-25}$ PDF satisfactorily fits the Weibull distribution, the Kolmogorov-Smirnov test was also applied. This test has been used for determining the optimal PDF for meteorological data sets [*Pavia and O'Brien*, 1986], demonstrating its utility for estimating the goodness of distribution fit [*Guo et al.*, 2016]. Considering a significance level of 1%, the Kolmogorov-Smirnov test indicates that the null hypothesis (that $SI_{75-25}$ data follow a Weibull distribution) cannot be rejected. Therefore, it is concluded that the $SI_{75-25}$ data adequately fit that distribution.

In short, the $SI_{75-25}$ is considered optimal for the purposes of the present research because of the following:

1. A significant correlation between MAPE and spread indexes was obtained, being the $SI_{75-25}$ and $SI_{70-30}$ the spread indexes with a higher correlation.
2. The MAPE has approximately the same magnitude as the average $SI_{75-25}$.
3. The bias is more balanced than any other indexes analyzed.
4. Therefore, $SI_{75-25}$ is the index used in the following sections to analyze the spatiotemporal behavior of uncertainty and predictability.

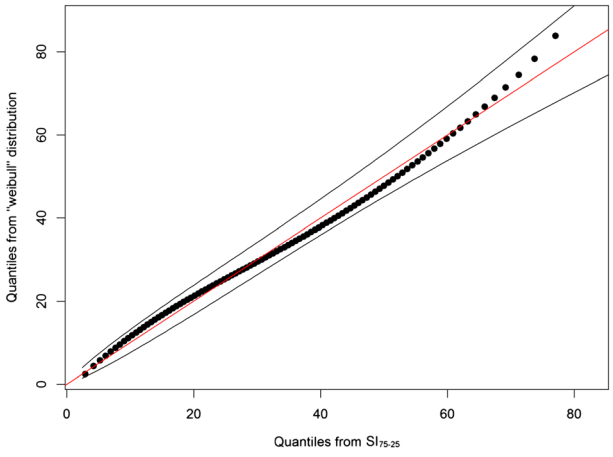


**Figure 6.** Fit of $SI_{75-25}$ PDF to Weibull distribution. The black lines denote the confidence interval for a satisfactory fit to the Weibull distribution.

### 4.2. Spatiotemporal Analysis of Spread

Once the $SI_{75-25}$ was selected as optimal, a spatiotemporal analysis was done for the data of 2010. First, evolution of the $SI_{75-25}$ PDF

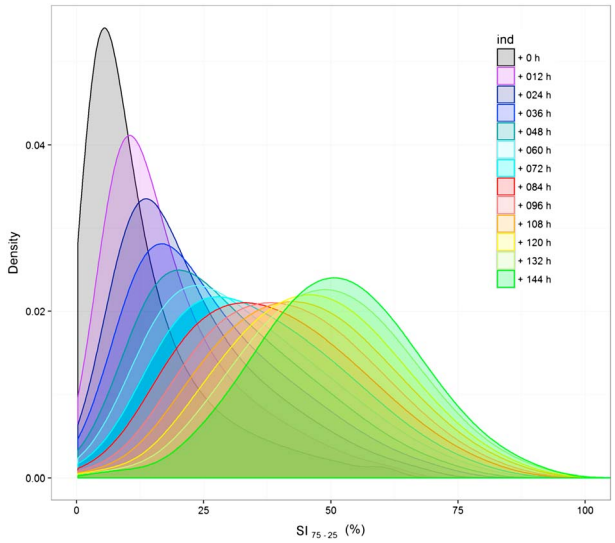


**Figure 7.** Evolution of $SI_{75-25}$ with forecast lead time.

throughout the forecast horizon was developed. Any weather prediction, regardless of how it is generated, has a horizon beyond which it is not reliable. Atmospheric predictability (i.e., the study of the extent to which the atmosphere is or is not predictable) has been an open question for decades, which has fostered the development of various methods of weather forecast generation [*Pellerin et al.*, 2003; *García-Moya et al.*, 2011].

Figure 7 illustrates the PDF of each forecast lead time. The first lead times are characterized by asymmetric distributions. These PDFs are positively skewed, which is typical of a Weibull distribution. They have a long tail to the right, with a maximum frequency (mode) to the left of the median and mean. However, PDFs of the last lead times tend to approximate to a normal distribution. Kurtosis is greater in the first prediction intervals. The kurtosis decreases with forecast horizon, which means that there is more dispersion in the last steps of the forecast. This reflects that uncertainty increases with lead time. This brief analysis suggests that it is not useful to extend the forecast horizon to more than 144 h, because the uncertainty increases markedly and the PDF loses its similarity to those of the wind speed databases. However, a more detailed analysis is required to accurately determine the maximum forecast horizon.

Subsequently, we analyzed the temporal evolution of $SI_{75-25}$ to measure the spread rate of increase with lead time. As expected, it was found that the longer the lead time, the greater the spread. Then, we estimated the correlation between the spread increase rate and the lead time, obtaining a Pearson's correlation coefficient of 0.99 (larger than the 0.01 significance level). Therefore, the null hypothesis that the two variables are not correlated can be rejected, and the temporal evolution of $SI_{75-25}$ can be fit to a first-degree polynomial ($y = \alpha + \beta x$). The equation of that polynomial was estimated individually at each grid point of the domain. In that way, we detected areas with a greater (less) spread increase rate and, consequently, more (less) uncertainty.

Intercept ($\alpha$) values of the first-degree polynomial equation are shown in Figure 8a. Small $\alpha$ values stand out over the Atlantic Ocean. Meteorological data there are scarce, which can be one reason for the small initial dispersion among members of the ECMWF EPS. Large $\alpha$ values are found near complex terrain in Iberia, such

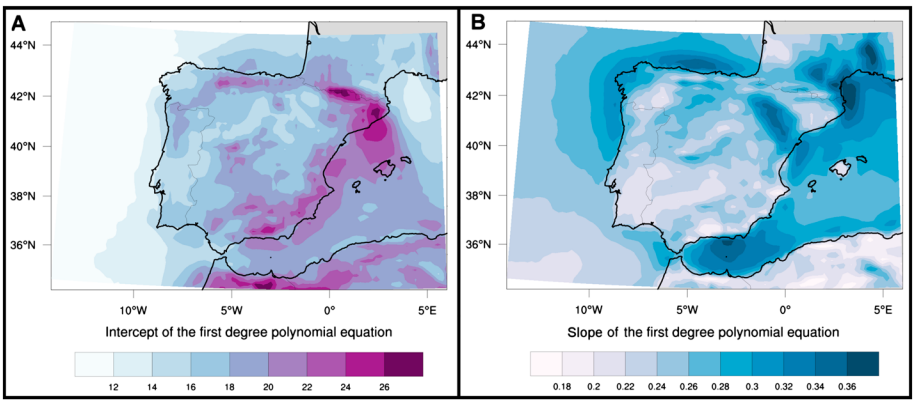


**Figure 8.** (a) Intercept and (b) slope of first-degree polynomial equation.

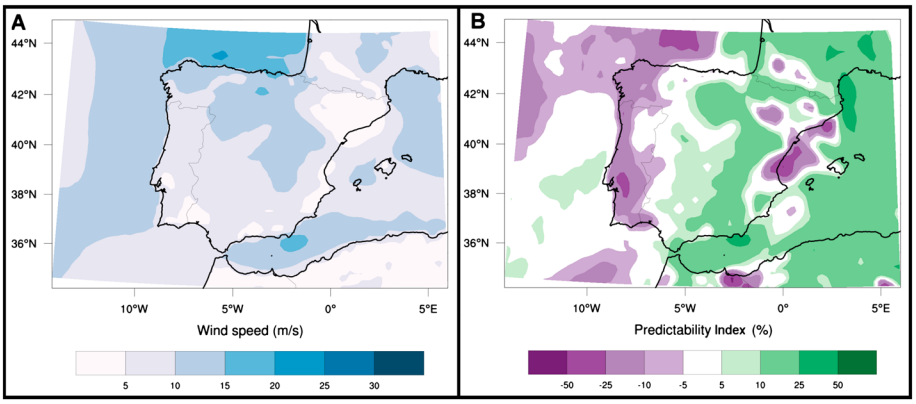


**Figure 9.** (a) Ensemble mean wind speed (m $s^{-1}$) forecast by ECMWF EPS. (b) PI (in percentage) is also shown. Both panels correspond to simulation initialized at 00:00 UTC on 26 February 2010 (lead time + 48 h, or 00:00 UTC on 28 February 2010).

as the Pyrenees, Cantabrian range, and the Atlas Mountains. The $\alpha$ values are also large on the Mediterranean coast, where sea breezes caused by land-sea thermal contrast are common. Large $\alpha$ values are also evident over northeastern Iberia and the Balearic Islands, which are affected by a northern or northeastern wind called Tramontana [*Font*, 2000; *Martin et al.*, 2011].

Figure 8b displays slope $\beta$, a measure of the spread increase rate with forecast lead time. In this case, the spread growth was greater in areas where strong winds are frequent. Thus, large slope values are seen on the northwestern Iberian coast and over the Cantabrian Sea, where strong winds associated with Atlantic storms are very common. Similarly, large values of $\beta$ are over the Ebro Valley, Gulf of Lion, and Alboran Sea, regions associated with strong winds known as Cierzo (northwesterly wind), Mistral (northwesterly), and Levante (easterly), respectively, which are caused by orographic channelling [*García*, 1985; *Font*, 2000].

### 4.3. Ability of ECMWF EPS in the Forecast of Extreme Winds. A Case Study

A particular case study was selected to develop tools focused on the prediction of extreme wind episodes. During 26–28 February 2010, extratropical cyclone Xynthia affected the Iberian Peninsula, France, and central Europe. The explosive development of this storm was related to anomalously warm sea surface temperatures in the North Atlantic. During these days, a minimum sea level pressure of 968 hPa and wind gusts $>$50 m $s^{-1}$ were recorded, and there were 65 casualties and more than a billion Euros in economic losses [*Fink et al.*, 2012]. For these reasons, the storm was considered suitable for studying an extreme wind episode.

At midday on 26 February 2010, there was a cyclone with minimum sea level pressure 995 hPa over the subtropical North Atlantic, further south of where Atlantic storms are usually located. During the subsequent 24 h, the cyclone underwent a striking intensification (minimum central pressure 975 hPa) while it moved toward the west coast of the Iberian Peninsula. The next day (28 February), the center of Xynthia was over the Bay of Biscay, very close to the French Atlantic coast. There, the cyclone reached its minimum sea level pressure. Around midnight of 27–28 February, a cold front associated with Xynthia crossed the Iberian Peninsula, causing strong wind gusts, especially along its north coast. The rapid cyclone intensification was favored by diabatic processes and the location of the jet stream, which was vertically aligned with the trajectory of Xynthia. More information about this episode can be found in *Liberato et al.* [2013].

In the following sections, the ECMWF EPS forecast on 26 February 2010 (initialized at 00:00; all times UTC) is analyzed. Figures 9 and 10 correspond to the +48 h lead time of this simulation; i.e., the forecast time is 00:00 on 28 February.

#### 4.3.1. Uncertainty Quantification: Predictability Index (PI)

We developed a method to quantify forecast uncertainty. The ensemble mean was used to estimate wind speed across the study area at 00:00 on 28 February 2010. Previous works [*Hamill and Colucci*, 1998; *Stensrud et al.*, 1999] have demonstrated that the ensemble mean tends to be a more accurate forecast than any individual member of the ensemble. This is due to ensemble mean that provides a nonlinear filtering that removes part of the growing errors [*Toth and Kalnay*, 1993].

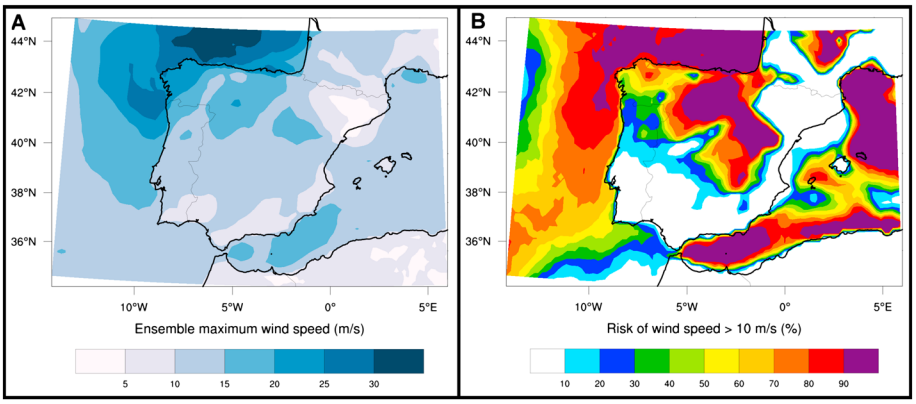


**Figure 10.** (a) Maximum wind speed (m $s^{-1}$) forecast by any member of ECMWF EPS. (b) Risk (in percentage) of sustained wind speed >10 m $s^{-1}$. Both panels correspond to simulation initialized at 00:00 UTC on 26 February 2010 (lead time + 48, or 00:00 UTC on 28 February 2010).

Figure 9a displays wind speed estimated by the ensemble mean. The forecast based on that mean estimates sustained winds stronger than 20 m $s^{-1}$ over the Cantabrian Sea, associated with the path of Xynthia near northwestern Iberia. Sustained winds stronger than 10 m $s^{-1}$ are forecast on the northern plateau, Iberian and Central Ranges, and, in general, the northern third of the peninsula (except in the northeast, where the Pyrenees blocks northern air mass advection).

According to *Toth et al.* [2001], the ensemble mean will more likely verify when the EPS spread is small. The spread of models is evaluated by diverse plots and diagrams (e.g., rank histograms and spread error plots), but these are usually obtained from areal averages that exclude the spatial variability of uncertainties. To avoid this shortcoming, the spread index was calculated at every grid point to display the PI geographically. As explained in section 3, the PI is calculated by subtracting the $SI_{75-25}$ for a particular forecast from the $SI_{75-25}$ average for each grid point and lead time during the previous 30 forecasts. According to *Toth et al.* [1998], the preceding 30 day average is a useful period for the comparison of spread. The ES is correlated with forecast uncertainty [*Lee et al.*, 2012], allowing its use for measurement of predictability in probabilistic forecasts [*McMurdie and Ancell*, 2014].

Figure 9b shows PI calculated at 00:00 on 28 February 2010. In this figure, the green colors indicate $SI_{75-25}$ values smaller than normal. Thus, reliability of the forecast is greater in green areas, because its ensemble members agree more than usual; therefore, the predictability is greater and the model is more reliable. In contrast, the purple colors indicate more spread than usual, so forecast uncertainty in such regions is greater and the predictability therefore is less. This is in accord with *Buizza et al.* [1999], who claimed that small spread around the control (ensemble mean in this case) indicates a skillful control forecast, at least in midlatitudes where model physics are accurate [*Buizza*, 1997].

As seen in Figure 9b, predictability was very poor along the western Iberian coast and, in particular, over the Cantabrian Sea, where the center of extratropical cyclone Xynthia was located at the time. Therefore, there was huge spread between the various members of the ECMWF EPS, so the wind speed forecast using the ensemble mean may not be very accurate. In this kind of situation, additional information is required. For this reason, specific tools for forecasting extreme wind episodes were developed, as described in the next section.

#### 4.3.2. Probabilistic Forecasting of Extreme Wind Episodes

Strong winds can pose a risk to human activities [*Staid et al.*, 2015]. For this reason, an analysis of the maximum sustained wind forecast by any member of the ECMWF EPS for +48 h lead time (00:00 on 28 February 2010) is shown in Figure 10a. In previous sections, it was decided to exclude ensemble members above the 75th percentile and below the 25th percentile for estimation of model predictability. However, underestimation of extreme wind episodes is very likely when only considering information provided by ensemble mean as is the case of Figure 9a. Therefore, it is useful to analyze information provided by ensemble members with larger wind speed values, at least for forecasting of extreme wind events with a long return period (as in the case of the episode treated in this section).

As shown in Figure 10a, when representing maximum wind speed estimated by any member of the ensemble, the values obtained are much larger than those calculated by the ensemble mean (Figure 9a). In fact, sustained winds $>30$ m $s^{-1}$ are reached over the Cantabrian Sea (~50% stronger than in the ensemble mean case), which are more similar to actual maxima registered during the episode [*Liberato et al.*, 2013].

Figure 10b shows a tool for estimating risk of wind speeds over a certain threshold (in this case, it was decided to select a threshold of 10 m $s^{-1}$). This application is known as probability of exceedance and is calculated by counting the number of ensemble members exceeding a certain threshold. Then, that number is divided by the total number of ensemble members, which is finally multiplied by 100 for estimates of the risk in percentage.

In this case study, higher-risk values of wind speed $>10$ m $s^{-1}$ were found in the zone most affected by the Xynthia cyclone track, i.e., the northwestern Iberian coast and Cantabrian Sea. These areas exhibited less predictability, with great spread between EPS members (Figure 9b). However, other areas of wind speed $>10$ m $s^{-1}$ (such as the Alboran Sea) predicted by the EPS (Figure 10b) displayed strong predictability in Figure 9b. Locations where the difference between ensemble mean (Figure 9a) and ensemble maximum (Figure 10a) was large (e.g., Cantabrian Sea) are characterized by poor predictability (Figure 9b). However, forecast uncertainty is low where the ensemble mean is similar to the ensemble maximum (e.g., Alboran Sea).

Information provided by combination of the above products can be very useful in the decision making of civil protection authorities prior to episodes posing potential danger to human activities. Thus, by knowing the wind strength threshold that can pose a threat, authorities can establish warnings and even restrict certain activities when weather models predict a strong probability of threshold exceedance. In this way, the prediction disposes of probabilistic information about the risk of occurrence of strong sustained winds. Similar applications have been successfully used in the forecasting of other meteorological risks, such as snowfall [*Fernández-González et al.*, 2015].

## 5. Concluding Remarks

Probabilistic models can output different scenarios of weather evolution. When several ensemble members agree on a particular forecast pattern, then such a pattern is very likely to verify. This is related to the fact that the ensemble mean is more accurate than any individual ensemble member, which is one of the advantages of probabilistic forecasts. This can be very useful for minimizing failures from using only a single forecast. However, estimation of the likelihood of each scenario is very difficult to determine. Furthermore, uncertainty and predictability quantification is still a challenge in modern meteorology. In this respect, we have reached the following conclusions.

1. The ensemble mean is probably the most accurate indicator of a probabilistic model, but it is useful to complement it with an estimate of the uncertainty associated with the forecast.
2. The interquartile range of the ES ($SI_{75-25}$) was selected as the most balanced spread index, owing to its correlation with the MAPE and its correspondence with the PDF of surface wind databases.
3. The $SI_{75-25}$ rate of increase had an approximately linear behavior, so it was possible to adjust the increase of the spread with forecast lead time to a first-degree polynomial equation. By calculating the values of intercept and slope at each grid point of the domain, we detected areas where the spread was greater at the beginning of the forecast (mainly orographic barriers and Mediterranean coast) and where the spread increment was more pronounced (Ebro Valley, Cantabrian Sea, northwest coast of the Iberian Peninsula, Alboran Sea, and Gulf of Lion).
4. The PI was defined by subtracting the $SI_{75-25}$ of a particular forecast from the $SI_{75-25}$ registered during the previous 30 days. Thus, during a specific weather event, it can be estimated whether the model has more ($PI > 0$) or less ($PI < 0$) predictability than usual.
5. The maximum wind speed estimated by any ensemble member at each grid point is very useful for the forecasting of strong wind episodes. This information is complemented by the risk of occurrence (in percentage) of sustained wind gusts stronger than a certain threshold.
6. The tools developed herein can be very useful as predictors during the decision making process, helping to optimize wind energy production and minimizing risks associated with extreme wind episodes.

It should be added that wind speed predictions are very sensitive to model physics, such as the planetary boundary layer processes [*Yang et al.*, 2016]. As a result of ECMWF EPS does not consider model deficiencies; sometimes, the ES may not reflect total forecast uncertainty [*Toth et al.*, 2001]. Therefore, future studies will focus on creating an ensemble that combines uncertainty in initial conditions and model physics. Furthermore, improvement in forecasting wind fields by the use of a mesoscale model should be examined, in order to study any enhancement of wind strength by Iberian Peninsula orography. In addition, the methods developed in the present work will be tested not only by surface wind but also by winds at heights nearer the hub of wind turbines (~100 m above ground level). Otherwise, the possibility of rescaling spread measures with too large bias will be considered. Finally, the use of a proper score [e.g., *Christensen et al.*, 2014] will be considered for evaluating how well forecast uncertainty is represented by an ensemble.

**Acknowledgments**
This work was partially supported by research projects METEORISK (RTC-2014-1872-5), CGL2011-25327, AYA2011-29967-C05-02, PCIN-2014-013-C07-04 (UE ERA-NET Plus NEWA Project), ESP2013-47816-C4-4-P, CGL2010-15930, and CGL2016-78702, and by the Instituto de Matemática Interdisciplinar (IMI) of the Universidad Complutense. Special thanks are due to Roberto Weigand, Steven Hunter, and Analisa Weston. The authors also thank the Deutscher Wetterdienst (DWD) and European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the gridded daily mean near-surface (10 m) wind speed for Europe (DWD), EPS, and ERA-Interim databases (ECMWF). To request the data, please contact S. Fernández-González (sefern04@ucm.es).

## References

Ancell, B. C. (2013), Nonlinear characteristics of ensemble perturbation evolution and their application to forecasting high-impact events, *Weather Forecasting*, *28*(6), 1353–1365, doi:10.1175/WAF-D-12-00090.1.

Bougeault, P., et al. (2010), The THORPEX Interactive Grand Global Ensemble, *Bull. Am. Meteorol. Soc.*, *91*, 1059–1072, doi:10.1175/2010BAMS2853.1.

Bowler, N. E., A. Arribas, S. E. Beare, K. R. Mylne, and G. J. Shutts (2009), The local ETKF and SKEB: Upgrades to the MOGREPS short- range ensemble prediction system, *Q. J. R. Meteorol. Soc.*, *135*, 767–776, doi:10.1002/qj.394.

Brinckmann, S., S. Krähenmann, and P. Bissolli (2015), High-resolution daily gridded datasets of air temperature and wind speed for Europe, *Earth Syst. Sci. Data Discuss.*, *8*, 649–702, doi:10.5194/essdd-8-649-2015.

Buizza, R. (1997), Potential forecast skill of ensemble prediction, and spread and skill distributions of the ECMWF ensemble prediction system, *Mon. Weather Rev.*, *125*, 99–119.

Buizza, R., M. Miller, and T. Palmer (1999), Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, *Q. J. R. Meteorol. Soc.*, *125*, 2887–2908.

Chai, T., and R. R. Draxler (2014), Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature, *Geosci. Model Dev.*, *7*, 1247–1250, doi:10.5194/gmd-7-1247-2014.

Chun, J. A., S. Kim, W.-S. Lee, S. M. Oh, and H. Lee (2016), Assessment of multimodel ensemble seasonal hindcasts for satellite-based rice yield prediction, *J. Agric. Meteorol.*, *72*(3–4), 107–115, doi:10.2480/agrmet.D-15-00019.

Christensen, H. M., I. M. Moroz, and T. N. Palmer (2014), Evaluation of ensemble forecast uncertainty using a new proper score: Application to medium-range and seasonal forecasts, *Q. J. R. Meteorol. Soc.*, *141*, 538–549, doi:10.1002/qj.2375.

Dee, D. P., et al. (2011), The ERA-interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), 553–597, doi:10.1002/qj.828.

Dey, S. A., G. Leoncini, N. M. Roberts, R. S. Plant, and S. Migliorini (2014), A spatial view of ensemble spread in convection permitting ensembles, *Mon. Weather Rev.*, *142*, 4091–4107, doi:10.1175/MWR-D-14-00172.1.

Draxl, C., A. N. Hahmann, A. Peña, and G. Giebel (2014), Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes, *Wind Energy*, *17*(1), 39–55, doi:10.1002/we.1555.

Ellis, N., R. Davy, and A. Troccoli (2015), Predicting wind power variability events using different statistical methods driven by regional atmospheric model output, *Wind Energy*, *18*(9), 1611–1628, doi:10.1002/we.1779.

Fernández-González, S., F. Valero, J. L. Sánchez, E. Gascón, L. López, E. García-Ortega, and A. Merino (2015), Numerical simulations of snowfall events: Sensitivity analysis of physical parameterizations, *J. Geophys. Res. Atmos.*, *120*, 10,130–10,148, doi:10.1002/2015JD023793.

Fink A. H., S. Pohle, P. Knippertz, and J. Pinto (2012), Diagnosing the influence of diabatic processes on the explosive deepening of extratropical cyclones over the North Atlantic, *Geophys. Res. Lett.*, *39*, L07803, doi:10.1029/2012GL051025.

Font, I. (2000), *Climatología de España y Portugal*, vol. 422, pp. 116–128, Ediciones Universidad de Salamanca, Spain.

Fortin, V., M. Abaza, F. Anctil, and R. Turcotte (2014), Why should ensemble spread match the RMSE of the ensemble mean?, *J. Hydrometeorol.*, *15*, 1708–1713, doi:10.1175/JHM-D-14-0008.1.

García, L. (1985), La predicción del tiempo en el Valle del Ebro, Technical Report Series A, No 38, INM, Madrid.

García-Moya, J. A., A. Callado, P. Escribà, C. Santos, D. Santos-Muñoz, and J. Simarro (2011), Predictability of short-range forecasting: A multimodel approach, *Tellus A*, *63*, 550–563.

Gneiting, T., and A. E. Raftery (2005), Weather forecasting with ensemble methods, *Science*, *310*(5746), 248–249, doi:10.1126/science.1115255.

Griffies, S. M., and K. Bryan (1997), A predictability study of simulated North Atlantic multidecadal variability, *Clim. Dyn.*, *13*, 459–488, doi:10.1007/s003820050177.

Grimit, E. P., and C. F. Mass (2002), Initial results of a mesoscale short-range ensemble forecasting system over the Pacific northwest, *Weather Forecasting*, *17*(2), 192–205.

Grimit, E. P., and C. F. Mass (2007), Measuring the ensemble spread–error relationship with a probabilistic approach: Stochastic ensemble results, *Mon. Weather Rev.*, *135*, 203–221, doi:10.1175/MWR3262.1.

Guo, E., J. Zhang, H. Si, Z. Dong, T. Cao, and W. Lan (2016), Temporal and spatial characteristics of extreme precipitation events in the Midwest of Jilin Province based on multifractal detrended fluctuation analysis method and copula functions, *Theor. Appl. Climatol.*, 1–11, doi:10.1007/s00704-016-1909-4, in press.

Hamill, T. M., and S. J. Colucci (1998), Evaluation of eta–RSM ensemble probabilistic precipitation forecasts, *Mon. Weather Rev.*, *126*, 711–724.

Hamill, T. M., S. L. Mullen, C. Snyder, Z. Toth, and D. P. Baumhefner (2000), Ensemble forecasting in the short to medium range: Report from a workshop, *Bull. Am. Meteorol. Soc.*, *81*, 2653–2664.

Ho, C. K., E. Hawkins, L. Shaffrey, J. Bröcker, L. Hermanson, J. M. Murphy, D. M. Smith, and R. Eade (2013), Examining reliability of seasonal to decadal sea surface temperature forecasts: The role of ensemble dispersion, *Geophys. Res. Lett.*, *40*, 5770–5775, doi:10.1002/2013GL057630.

Hopson, T. M. (2014), Assessing the ensemble spread–error relationship, *Mon. Weather Rev.*, *142*, 1125–1142, doi:10.1175/MWR-D-12-00111.1.

Hou, D., E. Kalnay, and K. K. Droegemeier (2001), Objective verification of the SAMEX '98 ensemble forecasts, *Mon. Weather Rev.*, *129*, 73–91.
Jerez, S., J. P. Montavez, J. J. Gomez-Navarro, R. Lorente-Plazas, J. A. García-Valero, and P. Jimenez-Guerrero (2013), A multi-physics ensemble of regional climate change projections over the Iberian Peninsula, *Clim. Dyn.*, *41*, 1749–1768, doi:10.1007/s00382-012-1551-5.
Keune, J., C. Ohlwein, and A. Hense (2014), Multivariate probabilistic analysis and predictability of medium-range ensemble weather forecasts, *Mon. Weather Rev.*, *142*, 4074–4090, doi:10.1175/MWR-D-14-00015.1.
Kieu, C., P. T. Minh, and H. T. Mai (2014), An application of the multi-physics ensemble Kalman filter to typhoon forecast, *Pure Appl. Geophys.*, *171*(7), 1473–1497, doi:10.1007/s00024-013-0681-y.
Koh, T.-Y., Y. S. Djamil, and C.-K. Teo (2011), Statistical dynamics of tropical wind in radiosonde data, *Atmos. Chem. Phys.*, *11*, 4177–4189, doi:10.5194/acp-11-4177-2011.
Kolczynski, W. C., D. R. Stauffer, S. E. Haupt, N. S. Altman, and A. Deng (2011), Investigation of ensemble variance as a measure of true forecast variance, *Mon. Weather Rev.*, *139*, 3954–3963, doi:10.1175/MWR-D-10-05081.1.
Komaromi, W. A., and S. J. Majumdar (2014), Ensemble-based error and predictability metrics associated with tropical cyclogenesis. Part I: Basinwide perspective, *Mon. Weather Rev.*, *142*(8), 2879–2898, doi:10.1175/MWR-D-13-00370.1.
Krähenmann, S., P. Bissolli, J. Rapp, and B. Ahrens (2011), Spatial gridding of daily maximum and minimum temperatures in Europe, *Meteorol. Atmos. Phys.*, *114*, 151–161, doi:10.1007/s00703-011-0160-x.
Lee, J. A., W. C. Kolczynski, T. C. Mccandless, and S. E. Haupt (2012), An objective methodology for configuring and down-selecting an NWP ensemble for low-level wind prediction, *Mon. Weather Rev.*, *140*(7), 2270–2286, doi:10.1175/MWR-D-11-00065.1.
Liberato, M. L. R., J. G. Pinto, R. M. Trigo, P. Ludwig, P. Ordóñez, D. Yuen, and I. F. Trigo (2013), Explosive development of winter storm Xynthia over the subtropical North Atlantic Ocean, *Nat. Hazards Earth Syst. Sci.*, *13*(9), 2239–2251, doi:10.5194/nhess-13-2239-2013.
Lun, I. Y. F., and J. C. Lam (2000), A study of Weibull parameters using long-term wind observations, *Renewable Energy*, *20*, 145–153.
Martin, M. L., F. Valero, A. Morata, M. Y. Luna, A. Pascual, and D. Santos-Muñoz (2011), Springtime coupled modes of regional wind in the Iberian peninsula and large-scale variability patterns, *Int. J. Climatol.*, *31*, 880–895, doi:10.1002/joc.2127.
Mcmurdie, L. A., and B. Ancell (2014), Predictability characteristics of landfalling cyclones along the North American West Coast, *Mon. Weather Rev.*, *142*(1), 301–319, doi:10.1175/MWR-D-13-00141.1.
Messner, J. W., G. J. Mayr, A. Zeileis, and D. Wilks (2014), Heteroscedastic extended logistic regression for postprocessing of ensemble guidance, *Mon. Weather Rev.*, *142*, 448–456, doi:10.1175/MWR-D-13-00271.1.
Molteni, F., R. Buizza, T. Palmer, and T. Petroliagis (1996), The ECMWF ensemble prediction system : Methodology and validation, *Q. J. R. Meteorol. Soc.*, *122*, 73–119.
Newman, K. M., C. S. Schwartz, Z. Liu, H. Shao, and X.-Y. Huang (2015), Evaluating forecast impact of assimilating microwave humidity sounder (MHS) radiances with a regional ensemble Kalman filter data assimilation system, *Weather Forecasting*, *30*(4), 964–983, doi:10.1175/WAF-D-14-00091.1.
Pavia, E. G., and J. J. O'Brien (1986), Weibull statistics of wind speed over the ocean, *J. Clim. Appl. Meteorol.*, *25*(10), 1324–1332.
Pellerin, G., F. Lefaivre, P. Houtekamer, and C. Girard (2003), Increasing the horizontal resolution of ensemble forecast at CMC, *Nonlinear Processes Geophys.*, *10*, 463–468.
Quine, C. P. (2000), Estimation of mean wind climate and probability of strong winds for wind risk assessment, *Forestry*, *73*, 247–258.
Reynolds, C. A., P. J. Webster, and E. Kalnay (1994), Random error growth in NMC's global forecasts, *Mon. Weather Rev.*, *122*(6), 1281–1305.
Saetra, O., H. Hersbach, J.-R. Bidlot, and D. S. Richardson (2004), Effects of observation errors on the statistics for ensemble spread and reliability, *Mon. Weather Rev.*, *132*, 1487–1501, doi:10.1175/1520-0493(2004)132<1487:EOOEOT>2.0.CO;2.
Satterfield, E. A., and C. H. Bishop (2014), Heteroscedastic ensemble postprocessing, *Mon. Weather Rev.*, *142*, 3484–3502, doi:10.1175/MWR-D-13-00286.1.
Scheuerer, M., and D. Möller (2015), Probabilistic wind speed forecasting on a grid based on ensemble model output statistics, *Ann. Appl. Stat.*, *9*, 1328–1349, doi:10.1214/15-AOAS843.
Schneider, T., and S. M. Griffies (1999), A conceptual framework for predictability studies, *J. Clim.*, *12*, 3133–3155, doi:10.1175/1520-0442(1999)012,3133:ACFFPS.2.0.CO;2.
Staid, A., P. Pinson, and S. D. Guikema (2015), Probabilistic maximum-value wind prediction for offshore environments, *Wind Energy*, *18*(10), 1725–1738, doi:10.1002/we.1787.
Stensrud, D. J., H. E. Brooks, J. Du, M. S. Tracton, and E. Rogers (1999), Using ensembles for short-range forecasting, *Mon. Weather Rev.*, *127*, 433–446.
Thorarinsdottir, T. L., and T. Gneiting (2010), Probabilistic forecasts of wind speed: Ensemble model ouput statistics by using heteroscedastic censored regression, *J. R. Stat. Soc. Ser. A*, *173*, 371–388, doi:10.1111/j.1467-985X.2009.00616.x.
Toth, Z. (2001), Ensemble forecasting in WRF, *Bull. Am. Meteorol. Soc.*, *82*(4), 695–697.
Toth, Z., and E. Kalnay (1993), Ensemble forecasting at NMC: The generation of perturbations, *Bull. Am. Meteorol. Soc.*, *74*, 2317–2330.
Toth, Z., and E. Kalnay (1997), Ensemble forecasting at NCEP: The breeding method, *Mon. Weather Rev.*, *125*, 3297–3318.
Toth, Z., Y. Zhu, T. Marchok, S. Tracton, and E. Kalnay (1998), Verification of the NCEP global ensemble forecasts, in *Preprints of the 12th Conference on Numerical Weather Prediction*, pp. 286–289, Phoenix, Ariz.
Toth, Z., Y. Zhu, and T. Marchok (2001), On the ability of ensembles to distinguish between forecasts with small and large uncertainty, *Weather Forecasting*, *16*, 436–477.
Tracton, M. S., and E. Kalnay (1993), Operational ensemble prediction at the National Meteorological Center: Practical aspects, *Weather Forecasting*, *8*, 379–398.
Wang, X. G., and C. H. Bishop (2003), A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes, *J. Atmos. Sci.*, *60*, 1140–1158, doi:10.1175/1520-0469(2003)060,1140:ACOBAE.2.0.CO;2.
Wilks, D. (1995), *Statistical Methods in the Atmospheric Sciences*, pp. 281–296, Academic Press, San Diego, Calif.
Yang, B., Y. Qian, L. K. Berg, P.-L. Ma, S. Wharton, V. Bulaevskaya, H. Yan, Z. Hou, and W. J. Shaw (2016), Sensitivity of turbine-height wind speeds to parameters in planetary boundary-layer and surface-layer schemes in the weather research and forecasting model, *Boundary Layer Meteorol.*, *62*, 117–142, doi:10.1007/s10546-016-0185-2.
Zhu, Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne (2002), The economic value of ensemble based weather forecasts, *Bull. Am. Meteorol. Soc.*, *83*, 73–83.