WIND ENERGY
Wind Energ. 2008; **11**:483–502
Published online 26 March 2008 in Wiley Interscience
(www.interscience.wiley.com) DOI: 10.1002/we.270

The Influence of the Weibull Assumption in Monthly Wind Energy Estimation

E. García-Bustamante, Departamento de Energías Renovables, CIEMAT, 28040, Madrid, Spain and Departamento de Astrofísica y CC. de la Atmósfera, Facultad CC., Físicas, Universidad Complutense de Madrid, 28040, Madrid, Spain
J. F. González-Rouco*, Departamento de Astrofísica y CC. de la Atmósfera, Facultad CC., Físicas, Universidad Complutense de Madrid, 28040, Madrid, Spain
P. A. Jiménez, Departamento de Energías Renovables, CIEMAT, 28040, Madrid, Spain and Departamento de Astrofísica y CC. de la Atmósfera, Facultad CC., Físicas, Universidad Complutense de Madrid, 28040, Madrid, Spain
J. Navarro, Departamento de Energías Renovables, CIEMAT, 28040, Madrid, Spain
J. P. Montávez, Departamento de Física, Universidad de Murcia, 30071, Murcia, Spain

Key words:
wind speed;
wind energy;
Weibull;
power curve;
χ^2 test

An estimation of the monthly wind energy output for the period 1999–2003 at five wind farms in northeastern Spain was evaluated. The methodology involved the calculation of wind speed histograms and the observed average wind power versus wind relation obtained from hourly data. The energy estimation was based on the cumulated contribution of the wind power from each wind speed interval. The impact of the Weibull distribution assumption as a substitute of the actual histogram in the wind energy estimation was evaluated.

Results reveal that the use of a Weibull probability distribution has a moderate impact in the energy calculation as the largest estimation errors are, on average, no larger than 10% of the total monthly energy produced. However, the evaluation of the goodness of fit through the χ^2 statistics shows that the Weibull assumption is not strictly substantiated for most of the sites. This apparent discrepancy is based on the partial cancellation of the positive and negative departures of the Weibull fitted and the actual wind frequency distributions.

Further investigation of the relation between the χ^2 and the error contribution exposes a tendency of the Weibull distribution to underestimate (overestimate) the observed histograms in the lower and upper (intermediate) wind speed intervals. This fact, together with the larger wind power weight over the highest winds, results in a systematic total wind energy underestimation. Copyright © 2008 John Wiley & Sons, Ltd.

Received 20 June 2007; Revised 28 January 2008; Accepted 30 January 2008

Introduction

Through the last decades, renewable energies have been progressively established as a competitive and feasible energy resource that can be used as an alternative to more problematic energy sources such as oil or carbon-based or nuclear technologies.[1] Within this context, wind energy has undergone considerable development in

*Correspondence to: J. F. González-Rouco. Facultad de CC. Físicas, Universidad Complutense de Madrid, Ciudad Universitaria, 28040, Madrid, Spain.
E-mail: fidelgr@fis.ucm.es

many regions of the world.[2–5] Wind energy assessment is therefore both a topic of scientific interest and an issue of relevance with ecological, economic and political implications for society. Efforts have been oriented to the analysis and understanding of wind variability and its relation to wind energy production, as well as the determination and evaluation of potential local and regional wind energy resources.

There has been an emphasis in improving the understanding of wind regimes and its variability at different timescales in order to assess the suitability of operating conditions, both to ensure quality in the generation of the electricity supply to the network and guidance in the long-term management of large- and small-scale wind farms. The use and analysis of wind probability distributions have been one of the necessary lines of procedure due to its involvement in the computation of total available wind energy and wind energy distribution (wind power density). Numerous studies have paid attention to wind speed probability distributions in relation to a variety of topics such as wind energy development and sustainability,[6–8] analysis of wind loads,[9,10] studies of wave fields and surface roughness,[11–13] etc. The problems associated with these studies relate to an understanding of the wind and often, the wind power distributions. With this aim, a variety of theoretical probability distributions have been explored according to specific regions and site characteristics: Wentink[14] investigated some methods of fitting the Planck distribution to explore the wind power potential in Alaska; both Baynes[15] and Celik[16] experimented with the Rayleigh distribution, the former studied extreme winds, which could affect structure design and the latter evaluated the suitability of this function to estimate wind energy hourly wind speed time series; Sherlock[17] used the Pearson Type III (gamma) family to describe wind speed distributions; García *et al*.[18] employed the Lognormal and Weibull distributions to fit hourly wind speed data and to estimate wind energy production in Navarra (Spain), comparing it with experimental wind power production; Li and Li[19] and Ramírez and Carta[8] recently analysed the MEP-type distribution functions (a family of exponential probability distributions derived from the maximum entropy principle) and compared them with the Weibull distribution in its ability to reproduce wind and wind power density. Of the above, the most frequently referred probability function is the Weibull distribution.[20–24]

The Weibull distribution has been employed using a number of different perspectives: on one hand, its suitability to reproduce some aspects of the wind speed frequency distribution;[25–27] on the other hand, the use of the Weibull function has been included in techniques, which provide estimations of the available wind power[28–30] and, from this perspective, it has been used directly to assess the suitability of specific locations to generate wind energy.[31,32]

Within the context described above, wind energy density is calculated through two main approaches: either using the parameters derived from the fitted theoretical distributions[6,30,33] or through the calculation of the wind power density contribution from the particular range of wind speeds at a specific site. The following is concerned with the second approach in which the wind energy estimation is obtained through the product of the power corresponding to each wind value by the probability of each wind speed.[16,28,34] In practice, the relationship between wind power and wind is expressed in terms of a transfer function which relates both variables. This can be either a theoretical power curve[35–37] provided by the manufacturer, or some specifically interpolated or fitted curve representing power versus wind that expresses the effective relationship present in the actual data set.[38] As for the probability values, these can either be represented by the empirical frequency histogram[29] constructed from the wind speed data set or some theoretical distribution, which fits the wind speed statistical properties at the site.[31,32,39,40] In general, the errors in this approach will stem both from the approximations made to express the wind power versus wind relationship and from those relative to the assumption of a given theoretical frequency distribution. Herein, the focus is on the latter source of error and efforts are made to illustrate and quantify the contribution to the error of the probability terms in estimating total wind energy. For this purpose, monthly observations of wind energy and estimations are compared. The analysis assumes an *a priori* selection of the Weibull function as a candidate to fit the experimental histograms. This choice was made on the basis of its extended use and on its interesting properties and applications;[18,29,30] also, it was based on the fact that previous studies have found[18] this distribution appropriate for specific sites in this area of Spain. However, the analysis shows that a Weibull distribution is not adequate for all sites and time steps. The impact that the quality of the fit of Weibull statistics to the observed values has on the final estimation of wind energy will be discussed. In doing so, an overall evaluation of the total error obtained using this

DOI: 10.1002/we

approach (under the assumption of the Weibull statistics) is also provided. The study uses data from five wind farms located in Navarra (northeast of the Iberian Peninsula).

The data section presents the wind speed and wind power data set. The methodology section describes the approach employed in the estimation of monthly wind energy as well as general aspects of the Weibull distribution, its empirical fit and the evaluation of the goodness of fit. The results are presented in the analysis of results section. Finally, conclusions are presented in the last section.

Data

In this study, wind speed (w) and wind power production data from five wind farms in the Comunidad Foral de Navarra (CFN) are used. This is a region in the northeast of Spain that has quickly reached a high level of significance in the development of wind energy resources/facilities (Figure 1). The geographical location of the five sites considered, the wind sensor heights and dates with available observations are summarized in Table I.

Hourly wind velocity measurements were collected by anemometers placed in meteorological masts at each wind farm (see Table I for details). For the particular case of the wind farm at Leoz, wind data were additionally available at each wind turbine at the hub height, this varying between 40 and 45 m. Thus, in Leoz, all the 30 wind speed time series (one series per wind turbine) were spatially averaged obtaining a single time series representative for the whole wind farm. Wind speed mean values ($\overline{w}$) and deviations (S_w) for each site are shown in Table II. All wind farms have similar $\overline{w}$ values, ranging from 7.4 m s^{-1} San Martín to 7.9 m s^{-1} in Alaiz and Leoz, except for El Perdón, which is the windiest location (9.2 m s^{-1}). Its deviation is also the largest one (5.2 m s^{-1}). S_w values in the rest of the locations vary from 3.8 to 4.8 m s^{-1}.

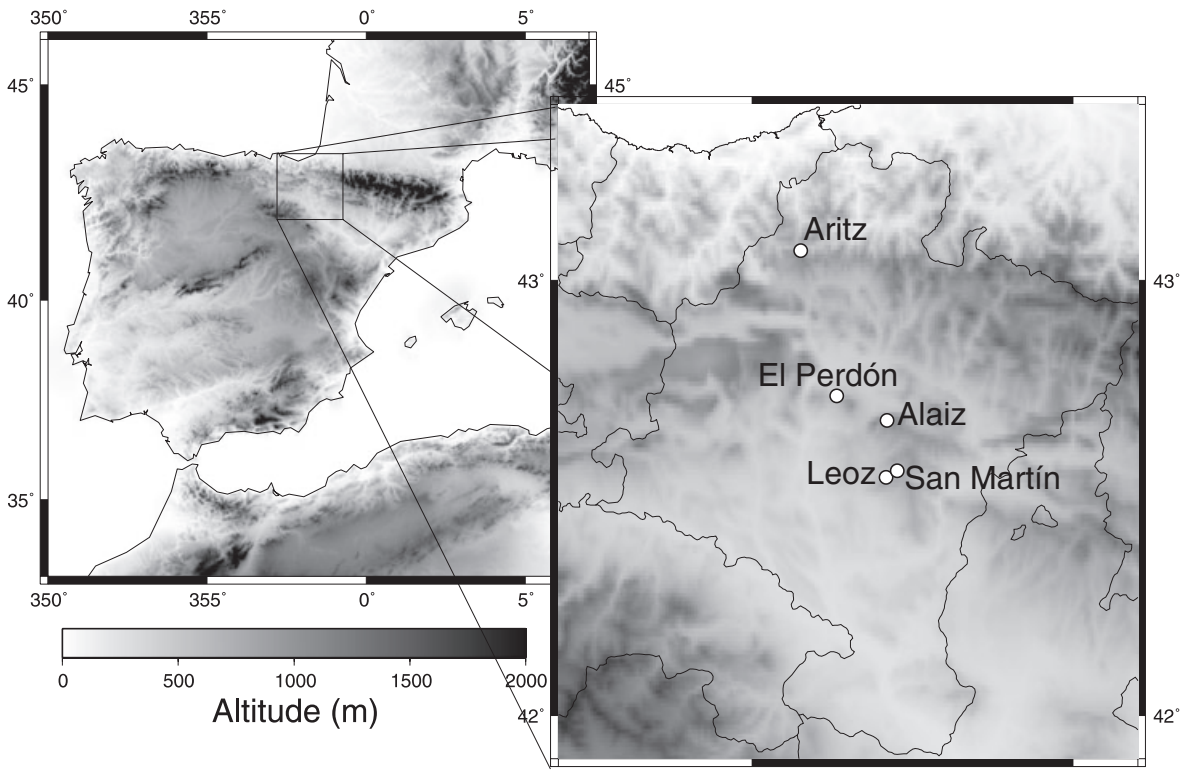

Figure 1. Orography characteristics (shaded) and wind farms locations (circles)

DOI: 10.1002/we

Table I. Geographical details: height of sensor, dates of available observations, turbine models and number of turbines of each type for the five wind farms considered in this study

	Lat (ºN)	Lon (ºW)	Sensor height (m)	No months (yr/mon)	Turbines model (nº)
Alaiz	42.679	1.579	30	42 (99/12–03/05)	5 (39)
Aritz	43.067	1.847	40	41 (00/01–03/05)	3 (32)
El Perdón	42.736	1.735	40	47 (99/06–03/05)	1 (26), 2 (11), 3 (3)
Leoz	42.549	1.582	40–45	32 (00/11–03/06)	3 (16), 4 (13), 5 (1)
San Martín	42.565	1.548	30	22 (01/03–02/12)	3 (21), 4 (19)

The codes for the turbine model are 1 for G39-500, 2 for G42-500, 3 for G42-600, 4 for G44-600 and 5 for G47-660. See also Figure 1.

Table II. Monthly mean wind ($\overline{w}$) and wind energy ($\overline{wE}$), mean Weibull parameters ($\overline{k}$, $\overline{c}$), monthly mean estimations using the observed histograms (wE_{ref}) and estimations using the Weibull histograms (wE_{weib}) and their respective standard deviations (S_w, S_{wE}, S_k, S_c, $S_{wE_{\mathrm{ref}}}$, $S_{wE_{\mathrm{weib}}}$) at each wind farm

	$\overline{w}$ (S_w) (m s^{-1})	$\overline{wE}$ (S_{wE}) (GWh)	$\overline{k} \pm S_k$	$\overline{c} \pm S_c$ (m s^{-1})	$\overline{wE_{\mathrm{ref}}}$ ($S_{wE\mathrm{ref}}$) (GWh)	$\overline{wE_{\mathrm{weib}}}$ ($S_{wE\mathrm{weib}}$) (GWh)
Alaiz	7.9 (4.2)	3.4 (0.8)	2.0 ± 0.3	9.0 ± 1.3	3.4 (0.8)	3.2 (0.8)
Aritz	7.7 (4.8)	2.1 (0.8)	2.1 ± 0.5	9.4 ± 2.1	2.0 (0.8)	2.0 (0.8)
El Perdón	9.2 (5.2)	3.2 (0.7)	2.4 ± 0.4	11.4 ± 1.5	3.2 (0.6)	3.1 (0.5)
Leoz	7.9 (4.1)	2.2 (0.5)	2.1 ± 0.2	8.9 ± 1.0	2.2 (0.5)	2.0 (0.4)
San Martín	7.4 (3.8)	2.9 (0.7)	2.5 ± 0.2	9.1 ± 1.2	2.8 (0.7)	2.8 (0.7)

Figure 2 illustrates the variability of the monthly frequency distribution of wind at each location. The frequency levels of occurrence (points) were calculated from the hourly observations and depicted for each month over a fixed array of 2 m s^{-1} width wind speed intervals common to all months for each site. The range of variability of the monthly frequency of wind occurrence changes from site to site, subject to the particularities of the annual cycle and interannual variability. For the purpose of illustration, the wind speed annual cycle is represented in Figure 3 at each site. It is possible to observe that summer months are less windy than winter ones, except for El Perdón, which shows a less pronounced annual cycle with the minimum of the curve slightly displaced to the autumn months.

An estimation of a mean histogram is provided in Figure 2 through the average of all monthly frequency values for each wind interval. The overall tendency to present positive skewness (values larger than 0)[28] is evident from the average histograms. The most positively skewed distributions are found at Aritz and Alaiz with some monthly values that reach 1.57 and 0.96, respectively. Some examples of negatively skewed histograms are also found at some locations, with maximum values reaching –0.93 and –0.3 at El Perdón and Leoz, respectively. Most months show platykurtic (i.e. flatter than normal; kurtosis lower than 3) distributions with values typically between 2 and 3; minimum values of 1.8 are reached at El Perdón and Leoz. Leptykurtic distributions (i.e. more peaked than normal; kurtosis larger than 3), though less frequent, can also be found mainly at Aritz and San Martín, where maximum values reach 7.01 and 4.07, respectively. The values obtained in this study for the symmetry (skewness) of histograms can be considered moderate if compared, for instance, with those obtained by Torres *et al.*[41] also for Navarra locations. There, the wind speed data were first divided into sectors according to wind direction and an extended sample was obtained, resulting in larger skewness values than the ones obtained in the present work. In the case of kurtosis or the *peakedness*, they also suggested a typical flattening of the distribution when considering the dominant direction of the wind speed. Additionally, they found some cases with larger values than any of those found in the present study (some cases reached kurtosis values of 10).

DOI: 10.1002/we

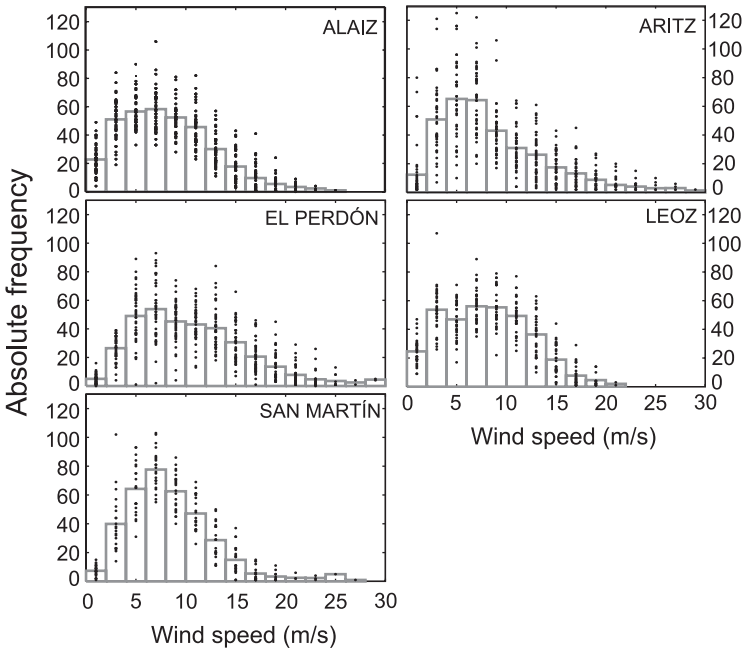

Figure 2. Wind speed monthly histograms at each site. The monthly frequency levels for each representative wind value are indicated by points. Bars represent the average histogram for all available months. See text for details

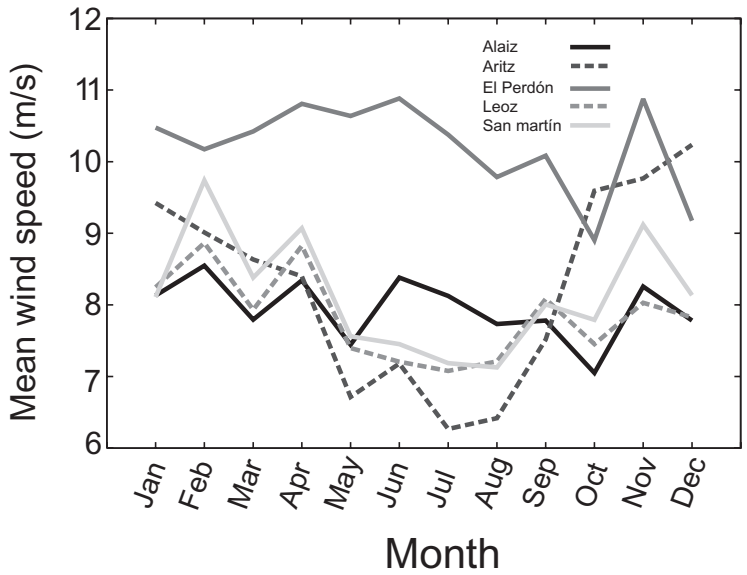

Figure 3. Long-term monthly mean wind speed for the period June 1999 to June 2006

A wind power production time series at each wind farm was obtained by spatially averaging the hourly power outputs from every wind turbine within the wind farm. Some of the wind farms incorporate more than one type of turbine model with different technical properties, which also imply different power production. This is the case of El Perdón, Leoz and San Martín wind farms. El Perdón combines three types of turbines: G39-500, G42-500 and G42-600. The nomenclature GXX-YYY is related to the blade diameter in meters (XX) and

DOI: 10.1002/we

the *rated* power (the power output under nominal operating conditions which are determined by the optimum rotor speed and the installed rated generator power[42]) in kilowatt (YYY). In the case of Leoz, three types are installed: G42-600, G44-600 and G47-660. Finally, San Martín makes use of the types G42-600 and G44-600. The types of machine described are angle pitch regulated.[42] For cases in which several turbine models were available, wind power time series were calculated for each model subgroup. In the remaining wind farms, only one type of model was available at each site, namely G47-660 and G42-600 for Alaiz and Aritz, respectively. The number and types of wind turbines installed at each wind farm are showed in Table I.

An alternative strategy to using a single wind and wind power series at each farm would involve the use of observations obtained at each mast. Such a procedure could introduce improvements in the estimation of wind power as a more realistic wind–wind power relation would be expressed for each mast location. However, rather than refining as much as possible our estimates, the purpose of this analysis was to explore the impacts of the Weibull assumption in the calculation of monthly wind energy. The simplification adopted herein through the use of a single wind power time series within the farm is supported by the high similarity of wind power production from mast to mast. In turn, the use of a single wind time series is somewhat forced upon us by the unavailability of turbine specific wind observations in most wind farms.

Nevertheless, the use of a pair of average time series as representative of the whole wind farm introduces some perturbations in the wind power versus wind relation at hourly timescales. Figure 4 illustrates this by showing, for each wind farm, the theoretical power curve together with the dispersion diagram of the hourly wind–wind power pairs. For reasons of simplicity, in the cases of wind farms with several wind turbine models (El Perdón, Leoz and San Martín), diagrams include only hourly data from one model of turbine. The hourly pairs and the theoretical power curve shown in Figure 4 correspond to the model G42-600 in all wind farms except for Alaiz (G47-660). The dispersion of points observed in Figure 4 reveals that the actual production can deviate considerably from the theoretical power curve, either distributing around it for intermediate or low wind speeds or typically below the *rated power* at high wind velocity values (e.g. Alaiz, Aritz or San Martín).

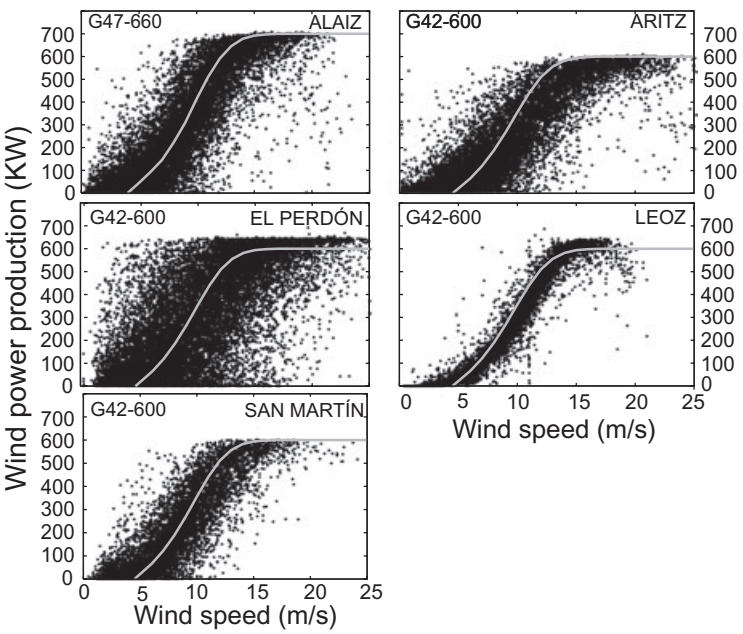

Figure 4. Dispersion diagrams of wind and wind power hourly values (points) obtained at each of the wind farms. Theoretical power curves are also shown (lines) for comparison. For simplicity, only one type of turbine model and the corresponding hourly data were arbitrarily selected in farms with multiple turbine types (G42-600 kW turbine model is used as example in all locations except for the case of Alaiz, where the model used is the G47-660 kW)

DOI: 10.1002/we

However, wind power values higher than the theoretical rated power can also be observed (El Perdón, Leoz). Thus, smaller than theoretically expected wind power values are present in most sites for high wind speeds and conversely larger than expected power outputs are achieved for some small wind velocities. These deviations could stem from many factors such as manipulation of the turbine parameters to test different operational conditions, shading effects of neighbouring wind masts, turbulence, etc. In addition, it is interesting to note the strong power reduction for some cases close to the *cutoff* wind value, particularly in Alaiz, Aritz and El Perdón, which is mainly due to the spatial averaging effect. Such patterns of behaviour are related to the operational control management (which in this case is angle pitch regulation) that forces some wind turbines to stop and thus, reduces the average wind power production relative to the expected values. This can happen for various reasons such as for technical assistance, breakdown or for instance, security reasons. In the case of the latter, a subset of wind turbines in the area of the highest intensity winds in the farm are switched off leading to a diminished average power production in the whole wind farm.[43]

Figure 5 illustrates the variability in the monthly wind–wind power relation produced by the perturbations discussed in Figure 4. Monthly *effective power curves* (EPCs) are obtained from the hourly values through the average of wind power measurements within each wind speed interval. As in the case of Figure 4, to obtain the EPCs shown in this figure, data corresponding only to one model of wind turbine have been used (G47-660 in Alaiz and G42-600 in the rest of wind farms). Aritz and El Perdón are the sites showing more month to month variability, while Leoz and San Martín show less scattered wind–wind power pairs, a feature compatible with the standard deviation values in Table II. In the case of Leoz, the smallest dispersion of the EPCs can be influenced by the fact of having wind measurements at each wind turbine. Therefore, the reduced variability of the EPCs at this site suggests that the average of mast specific wind values provides a better representation of the wind–wind power relation for the whole wind farm than using observations from a single meteorological station as in the case of the other farms.

The total monthly mean wE ($\overline{wE}$) and its corresponding standard deviation (S_m) are indicated in the second column of Table II. The final total production depends on the balance between the number of wind turbines available (see Table I) and the rated power. The maximum monthly energy production is achieved in Alaiz,

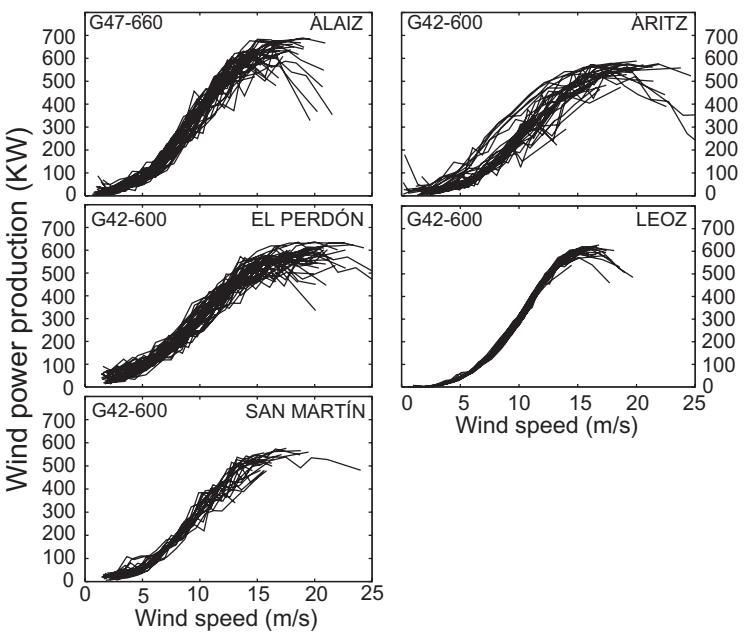

Figure 5. Effective power curves at the different wind farms. Each curve depicts the observed wind power versus wind relation for each month within the data set. The data used to represent the EPCs correspond to the turbine model G47-660 kW in Alaiz, while for the other sites, the model G42-600 kW is employed

DOI: 10.1002/we

where the most powerful turbines of the type G47-660, are installed. Aritz is the location with less energy production, stemming both from a lower rated power and fewer turbines.

For the calculations shown below, the final values of monthly energy production were obtained using a subset of 350 randomly selected pairs of hourly wind speed and wind power production observations within each month. This was done to exclude the potential disturbing influence of missing data in the comparison of observed and estimated energy values. The 350 pairs were chosen with the condition of being uniformly distributed through the month and containing both the wind and wind power values. This issue will be further discussed along with methodological aspects in the next section.

Methodology

The approach used to estimate monthly *wE* production is based on the use of the hourly frequency distribution of wind velocity within each month and an estimate of the power production dependence on wind, based, for instance, on the theoretical power curve or some substitute of it obtained from the observations.[16,29,38]

An estimation of the total monthly wind energy (*wE*) production can be obtained from summing all the power contributions from the various wind classes considering their frequency of occurrence and scaling the sum to the total number of time steps and turbines in the farm:

$$wE = \Delta t N N_t \sum_{i=1}^{n} p_{out}(w_i) f(w_i) \tag{1}$$

where $f(w_i)$ is the expected frequency for each of the n wind speed class intervals represented by w_i and $p_{out}(w_i)$ is a transfer function for the wind power versus wind relationship that provides an estimation of the power production for a given w_i; Δt is the time resolution of the input wind speed data (1 h), N is the total number of hours per month and N_t is the number of wind turbines. In those wind farms with several types of turbines installed, Equation (1) is computed for each group of wind turbines, with different $p_{out}(w_i)$ and N_t and the results are summed to provide a single estimation for the whole wind farm. Consequently, this method produces a cumulative contribution of the frequency of wind in each interval to the total wind power production by multiplying the frequency terms by the corresponding power contribution of each bin, $p_{out}(w_i)$. Many practical cases involve situations in which *wE* must be estimated without knowing the precise wind frequency distribution, $f(w_i)$, or the exact wind power versus wind relation, $p_{out}(w_i)$. In such cases, two potential sources of error are involved in the *wE* estimation through Equation (1): one stems from the use of an approximation for $p_{out}(w_i)$, for instance if the theoretical/manufacturer power curve is used;[32,44] the other is associated with the use of a theoretical distribution for $f(w_i)$.[8,30,45]

As previously mentioned, the use of an approximation for $p_{out}(w_i)$ introduces an error related to the fact that the power curve itself is an approach to the more complex actual wind–wind power relation (see Figure 4), which also varies with the time (e.g. Figure 5). Due to factors, such as turbulence, the effect of wind velocity reduction caused by the shadow between turbines, the variability of wind direction, etc., the actual power production does not coincide exactly with what theoretical power curves predict. Furthermore, the use of a global (spatially averaged in this work) power curve also introduces perturbations with respect to the expected power production since the wind power generation from each individual wind turbine within the site is not identical to the rest.[37] The use of a theoretical probability distribution instead of the observed wind velocity frequency involves an additional source of error when estimating *wE* in Equation (1) as much as the specific wind speed frequency distribution of a certain month deviates from the shape of the assumed theoretical probability distribution function.

This study explored the second source of error through evaluating the extent to which the assumption of a given probability distribution, $f(w_i)$, can contribute to the error in the final monthly *wE* estimation. A complete evaluation of errors would also involve analysing the impacts of selecting a certain $p_{out}(w_i)$ power curve. Though some of the results obtained herein do have implications concerning the effect of the $p_{out}(w_i)$, this source of error will be subject of attention elsewhere.

DOI: 10.1002/we

The Weibull probability distribution has been selected for this study as it is the most frequent probability density function in wind and wind energy studies.[36,39,46–49] In the context of wind energy, the use of the Weibull probability function is well extended because it generally provides a good fit to the wind speed distribution,[25,50] replicates its skewness well[28] and provides a suitable estimation of the cube wind speed for wind power analyses.[29] This distribution was successfully applied to wind speed measurements at several meteorological stations located in CFN, the study area employed in this analysis.[18] A limitation of the Weibull density function is that it can not estimate with accuracy the probability of calms or very low wind velocities.[33] However, this aspect does not produce a considerable impact on the final wE estimation as, for the lowest wind intervals, there is no relevant contribution to the wind power production, i.e. the corresponding $p_{\text{out}}(w_{\text{i}})$ is negligible in comparison to the power associated with the largest wind velocities, see for instance the values in Figures 2 and 5 for low wind speeds. For wind velocities below 4 m s^{-1} the production is 0 kW and for wind speeds smaller than 6 m s^{-1}, the power output is under 100 kW. Though the selection of a particular probability distribution such as the Weibull can inevitably narrow the final conclusions to the specific case studies in which this distribution is used, the following sections will arrive at some implications which attain a broader focus.

The evaluation of the error contribution from the probability terms $f(w_{\text{i}})$ can be easily done through establishing a maximum benchmark level of predictability with respect to which other errors can be defined. Of all possibilities, the most accurate estimation of wE with Equation (1), would be the one obtained by using the observed frequency histogram in each specific month as $f(w_{\text{i}})$ and the observed wind power versus wind relation as $p_{\text{out}}(w_{\text{i}})$, within each particular month. The variability in the actual monthly frequency distributions and effective power curves is shown in Figures 2 and 5. Such an estimation serves as the most accurate estimation possible and therefore as a benchmark of the upper limit of predictability that can be obtained using Equation (1). This estimate will be denoted as wE_{ref}. The error made by this estimation is obtained by calculating the difference between wE_{ref} and the observed wE (wE_{obs}), and will be denoted as

$$\xi_1 = wE_{\text{obs}} - wE_{\text{ref}} \tag{2}$$

This quantity provides information about the error in the method described by Equation (1), independent on any assumption made on $p_{\text{out}}(w_{\text{i}})$ and $f(w_{\text{i}})$. This error is derived from the extent to which the EPC is a bad representation of the wind power versus wind relation in the different masts and from the approximation, which implies considering a frequency histogram. If, in a limit case, the wind power versus wind transfer function was perfect when using point values, the estimated and observed wE would converge.

The second control estimation calculated also assumed the corresponding EPC for each month as $p_{\text{out}}(w_{\text{i}})$, but in this case, the theoretical Weibull frequencies were employed as $f(w_{\text{i}})$ in Equation (1). This estimation is denoted as wE_{weib}. As wE_{ref} and wE_{weib} only differ in the use of the Weibull expected frequencies in one case and the observed ones in the other, the contribution to the error of the frequency terms, under the assumption of a Weibull distribution can be studied from the difference between both previous estimations

$$\begin{aligned}\xi_2 &= wE_{\text{ref}} - wE_{\text{weib}} \\ &= \Delta t N N_{\text{t}} \sum_{i=1}^{n} p_{\text{out}}(w_{\text{i}})\, f_{\text{obs}}(w_{\text{i}}) - \Delta t N N_{\text{t}} \sum_{i=1}^{n} p_{\text{out}}(w_{\text{i}})\, f_{\text{weib}}(w_{\text{i}}) \\ &= \Delta t N N_{\text{t}} \sum_{i=1}^{n} p_{\text{out}}(w_{\text{i}})[f_{\text{obs}}(w_{\text{i}}) - f_{\text{weib}}(w_{\text{i}})]\end{aligned} \tag{3}$$

This should help to understand qualitatively and quantitatively the effect that discrepancies between the observed and the expected Weibull frequencies produce in the monthly wE estimation at the locations considered. A third error that addresses the combined effect of both kinds of error described above is provided by

$$\xi_3 = wE_{\text{obs}} - wE_{\text{weib}} \tag{4}$$

This magnitude incorporates the effects of assuming a certain probability function to the method errors in ξ_1.

DOI: 10.1002/we

The Weibull probability distribution

The Weibull probability density function,[51,52]

$$f(w) = \frac{k}{c}\left(\frac{w}{c}\right)^{k-1} exp\left[-\left(\frac{w}{c}\right)^{k}\right] \tag{5}$$

where w is the wind speed, is defined by two parameters, k and c. k is the dimensionless *shape parameter* related to the variability of wind and therefore, provides an approximation of the flatness of the distribution. c is the *scale parameter*, with metre per second dimensions, and is related to the mean value of the distribution.[24] In this work, both parameters defining the Weibull function are calculated using the *metho*d *of moments*, which provides a suitable estimation of the parameters[53] and the best result for the higher wind speed values in the distribution according to Tuller and Brett.[21] This method is based on the calculation of the first n sample moments and the use of them as estimators of population parameters of the distribution. This method performs well in comparison with other methods.[44,54] The mean Weibull parameters, calculated over the whole period of data, and their corresponding standard deviations are summarized in Table II (columns 3–4). The overall values of the shape parameter distribute close to the value $k = 2$, typical in practice[34], a limit value in which the Weibull distribution reverts to the Rayleigh distribution. The deviation of k is larger at Aritz and El Perdón, an unsurprising feature in view of the variability of the monthly distributions (points) in Figure 2 and the wind speed standard deviation values in Table II. The scale parameter (c) provides larger values for Aritz and El Perdón, the sites with maximum wind speed; deviations for this parameter are also larger at these sites.

The goodness of the fit of the Weibull adjusted distributions to the observed data was established in terms of the χ^2 statistic. This allows to test the null hypothesis that the actual sampled data follow a Weibull function.[55,26] The test statistic is defined as

$$\chi^2 = \sum_{i=1}^{n} \frac{[O(w_i) - E(w_i)]^2}{E(w_i)} = N \sum_{i=1}^{n} \frac{[f_{obs}(w_i) - f_{weib}(w_i)]^2}{f_{weib}(w_i)} \tag{6}$$

with $O(w_i)$ and $f_{obs}(w_i)$ being the absolute and relative observed frequencies for interval i, $E(w_i)$ and $f_{weib}(w_i)$ the absolute and relative expected Weibull frequencies for the same interval and N the total number of observations. The presence of the expected frequency in the denominator in Equation (6) attempts to weight the squared differences between frequencies eliminating the dependence on the specific shape of the distribution. That is, with the expected Weibull frequency in the denominator all intervals have the same weight in the calculation of differences. χ^2 will be used in the next section to analyse the goodness of fit and to illustrate its relationship with the error in the total monthly wE estimation.

One further comment is needed to address the sensitivity of the χ^2 test to the number of intervals and the existence of missing data. χ^2 is sensitive to the choice of the number of intervals to calculate histograms and differences in Equation (6). Also, Equation (6) is sensitive to small frequency numbers in $f_{weib}(w_i)$, which are often related to the handling of the wind intervals at the tails of the distribution. In order to mitigate this problem, the same conditions were imposed to each monthly histogram, that is, the number of intervals were fixed and the tails of the distribution were forced to contain 3% of the data.

The presence of a varying number of missing data in each month would also affect the results in Equation (6). This could make it difficult to discriminate whether changes in χ^2 correspond to a feature of the methodology or to the varying number of observations in each monthly sample, which could plausibly influence the resulting values obtained in Equation (6). Additionally, final monthly estimations of wE in Equation (1) could be affected by the lack of data at certain months, which could lead to an inaccurate result when computing contributions from the available hourly data. This possible perturbation was avoided by randomly selecting a fixed number of 350 h (ca. 50% of observations in each month) from the available wind–wind power pairs for each month (see data section). This number represents a balanced decision between using months with a very low number of observations, which might not adequately represent adequately the monthly means and totals and alternatively, selecting only months with a high number of measurements, which would lead to a scarceness of monthly records in the original data set.

DOI: 10.1002/we

Analysis of Results

This section compares the observed and calculated monthly wind energies at the five sites and analyses the implications of using a Weibull probability function to model wind behaviour. The section on the error in *wE* estimation presents the errors associated with the use of Equation (1) and evaluates the impact of using the Weibull probability distribution on the energy estimations in terms of the behaviour of the errors at each site along the period of observation. The section on the goodness of fit explores the relation of the error terms with the quality of the fit through illustrating and discussing systematic biases in the Weibull fit at the various sites.

Errors in wE estimation

The time evolution of the observed total monthly wind energy (wE_{obs}) is compared with the resulting wE_{ref} and wE_{weib} estimations in Figure 6 for all wind farms. The mean values of the two estimated series (wE_{ref} and wE_{weib}) are shown in Table II (columns 5 and 6) for all locations. Both estimations are close to the monthly mean observed value (second column in Table II). Figure 6 shows that there is a general agreement between measurements and estimations in all the cases analysed. The estimations using a realistic power curve and actual monthly frequency histograms (wE_{ref}) are difficult to distinguish from the observational values (wE_{obs}). Substituting the frequency histograms by the Weibull approximation introduces a slight error that becomes obvious only at Leoz, i.e. the largest error is found in Leoz for the wE_{weib} estimation, which on average is no greater than 227 MWh (10.3%). It is interesting to note that wE_{obs} displays a considerable amount of intra-annual and interannual variability, which wE_{ref} and wE_{weib} are able to capture. Such variability in energy production is likely related to changes in atmospheric circulation and will be dealt with elsewhere. As expected from the wind energy production means and standard deviations shown in Table II, Alaiz and El Perdón are the sites reaching higher energy productions in Figure 6.

The monthly evolution of the three types of error, ξ_1, ξ_2 and ξ_3 in the five wind farms, is presented in Figure 7 and allows for some further insight into the error behaviour. The errors are expressed in relative terms and

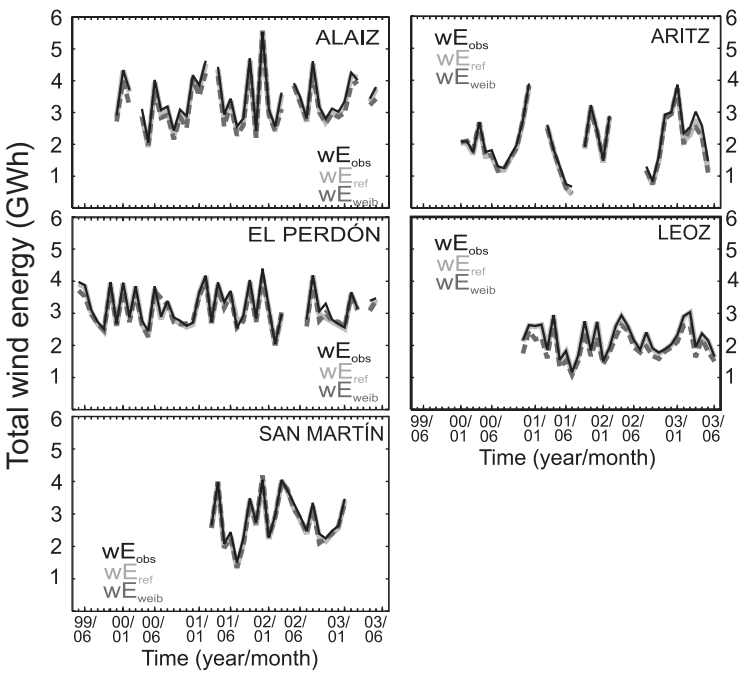

Figure 6. Total monthly observed wind energy (wE_{obs}) and estimations (wE_{ref} and wE_{weib}) at each site

DOI: 10.1002/we

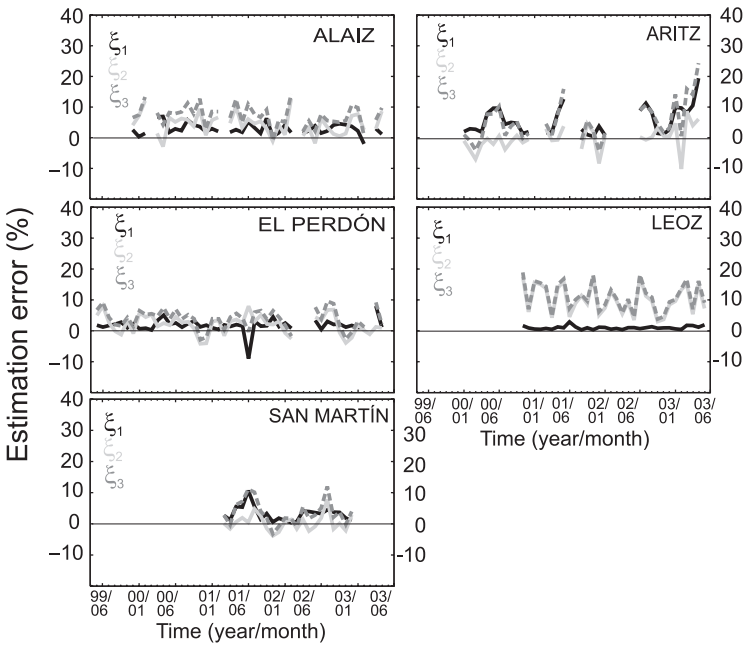

Figure 7. Time series of monthly errors (ξ_1, ξ_2, ξ_3) in wind energy estimation at each site

Table III. Averaged relative/total errors in monthly *w*E estimation at each wind farm (columns 1–3), average χ^2 values (column 4) and correlations (r) between ξ_2 and χ^2 time series at each wind farm (column 5)

	$\overline{\xi_1}$ (%/GWh)	$\overline{\xi_2}$ (%/GWh)	$\overline{\xi_3}$ (%/GWh)	$\overline{\chi_2}$	$r\,(\xi_2, \chi^2)$	$r\,(\xi_2^-, \chi^2)$	$r\,(\xi_2^+, \chi^2)$
Alaiz	2.6/0.1	4.5/0.2	7.2/0.2	34.8	0.3	–0.5	0.5
Aritz	5.0/0.1	$0.2/4.10^{-3}$	5.2/0.1	48.4	0.2	–0.5	0.6
El Perdón	$1.7/5.10^{-2}$	2.4/0.1	4.2/0.1	36.8	0.5	–0.5	0.6
Leoz	$0.9/2.10^{-2}$	10.3/0.2	11.3/0.2	55.0	0.6	–0.6	0.7
San Martín	2.9/0.1	0.3/0.01	3.0/0.1	29.2	–0.2	0.2	–0.2

Columns 6 and 7 show correlations between errors and χ^2 for the wind intervals where the contribution to $Np_{out}(w_i)[f_{obs}(w_i) - f_{weib}(w_i)]$ is negative (ξ_2^-) and positive (ξ_2^+) (see text for details).

are the result of dividing the average total error by wE_{obs} in the case of ξ_1 and ξ_3 and by wE_{ref} in the case of ξ_2. It can be observed from Figure 7 that most errors are positive, thus suggesting some bias to underestimate wE_{obs}. Celik[29] and Biswas *et al.*[34] reported results showing a tendency of the Weibull energy estimation to underestimate the reference energy in similar approaches. A summary with the averaged errors for the whole period at each site is given in Table III. The smallest mean error corresponds to ξ_1 ($wE_{obs} - wE_{ref}$) at the three of sites (Alaiz, El Perdón and Leoz) and to ξ_2 ($wE_{ref} - wE_{weib}$) at Aritz and San Martín. Thus, the substitution of the empirical histogram by a fitted distribution does not necessarily lead to larger errors in the energy estimation in spite of introducing additional errors in the probability representation. This feature is also apparent in the time evolution of errors in Figure 7. Though this might appear counter intuitive at first sight, it stems simply from Equation (1) and will be discussed in the next section.

The largest impact of introducing the Weibull approximations is found at Leoz, where mean relative differences between wE_{ref} and wE_{weib} amount to 10.3% (see Table III). These larger differences are also apparent in Figure 7. Interestingly, the ξ_1 methodological errors at Leoz are considerably smaller than at the other sites, a

DOI: 10.1002/we

feature that probably relates to the lower amount of variability in the wind power versus wind relation at this location (see Figures 4 and 5), where a spatially averaged wind time series was computed from every wind turbine.

A last comment can be made concerning the behaviour of the ξ_3 errors that basically accumulate the ξ_1 errors inherent to the method and ξ_2 related to incorporating a theoretical distribution function as a substitute for the observed histogram. As a result, changes with time in ξ_3 (Figure 7) are coherent with those in ξ_1 and ξ_2, depending on which one has the largest contribution to the error.

Goodness of fit

The ξ_2 errors shown in Figure 7 represent a relatively small part of the variability in wE_{obs} except for Leoz, where they stand out amounting to about 10% on average. Even if the impact of *a priori* introducing the Weibull distribution assumption is relatively small, the associated errors still represent a considerable amount of energy, and it is interesting to elucidate the reasons for such errors and the potential contribution of the goodness of the fit to them. In addition, it cannot be ruled out that, for other sites not considered here, such contributions may be larger, thus, it is relevant to understand the role of the frequency terms in Equation (3). For this reason, the error contribution of the Weibull distribution was evaluated through an exploratory analysis of the relation between ξ_2 and the monthly evolution of the goodness of the fit by means of the χ^2 statistic.

Table III (column 4) shows the average of all monthly χ^2 values at each site and Figure 8 compares the monthly evolution of ξ_2 and χ^2 along the period of observation. The $\chi^2_{0.01,14}$ critical value, beyond which the null hypothesis is rejected for a 0.01% significance level and 14 degrees of freedom (the analysis uses 17 intervals for every month and two fixed parameters, k and c), is also shown. $\overline{\chi_2}$ averages in Table III show the lowest values for San Martín and the largest ones for Leoz in agreement with the relatively small ξ_2 values in Table III and Figure 7 for both sites. It is interesting to note though, that Aritz shows a very low ξ_2, while it displays a comparatively large $\overline{\chi_2}$. This suggest there is not a close correspondence between smaller (larger) ξ_2 and smaller (larger) χ^2. Figure 8 allows for the assessment of this relation at a monthly resolution, studying the time series of χ^2 in comparison with ξ_2. The amount of months that satisfy the null hypothesis is variable from site

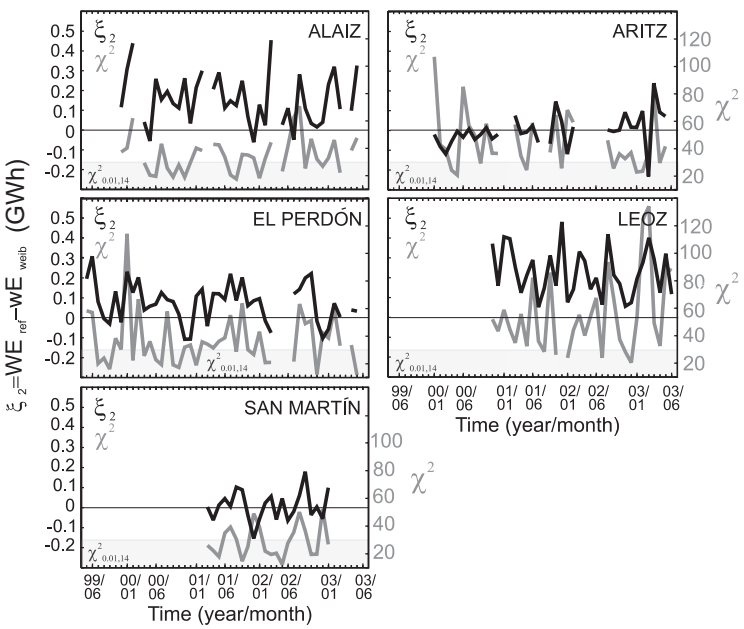

Figure 8. Time series of ξ_2 errors and monthly values of χ^2 at each site

DOI: 10.1002/we

to site. Though most of the months with available data at all sites register relatively low χ^2, close to or lower than the critical value, the percentage of months strictly respecting the threshold level is smaller, e.g. at Alaiz, El Perdón and San Martín, 29, 33 and 60% of the months (respectively) show χ^2 values that strictly suggest a Weibull distribution and at Aritz and Leoz lower percentages of 2 and 14% are registered. Thus, according to this test, the initial assumption of using a Weibull distribution is not broadly met at all sites and time steps. To a large extent, San Martín can be considered Weibull distributed or very close to the Weibull distribution in most months, but this assumption is only partially substantiated at the other sites. In spite of this, the replication of the energy production is reasonable in Figure 2 as previously illustrated. As it was argued before, the poorer quality of the fit along the period of observation at Leoz is consistent with the high ξ_2. However, this link between larger (smaller) errors and worse (better) fit is not evident at every site and time step.

The temporal variability in χ^2 and ξ_2 at Leoz reveals coherent changes, which lead to a correlation of 0.6 linking larger (smaller) distribution errors to larger (smaller) energy errors. Correlations for the other sites are shown in Table III (column 5) with values of 0.3 at Alaiz and 0.5 at El Perdón and negligible values (0.2) at the remaining sites. Thus, some of the sites show some consistent temporal link between χ^2 and ξ_2, however, as it was argued on the basis of Table III above, larger (smaller) ξ_2 do not regularly coincide with larger (smaller) ξ_2 values, suggesting that a closer Weibull fit does not guarantee smaller ξ_2 errors. For instance, if the time series at Aritz is considered, May and June of year 2000 show respectively a low and a high χ^2 value that meets in the first case the critical level and considerably departs from it in the second, while at the same time similar ξ_2 values are attained for both months; such type of events are not infrequent through the time series in Figure 8. The reasons for this behaviour is discussed on the basis of the error contributions by the $[f_{\text{obs}}(w_i) - f_{\text{weib}}(w_i)]$ terms in Equation (3).

Figure 9 illustrates the quality of the Weibull fit for various χ^2 levels and provides some insight into their contribution to Equation 3. Figure 9 presents two examples of histograms and associated Weibull fits for low and relatively large χ^2 values at Alaiz and Leoz. The selected months were arbitrarily chosen to be December 2001 and July 2002. For December 2001, deviations of the observed histogram from the Weibull shape produce χ^2 values of 19 (Alaiz) and 22 (Leoz). Larger deviations from the Weibull shape in July 2002 generate χ^2 values of 72 (Alaiz) and 95 (Leoz).

The larger differences contributing to the error in July are due to the Weibull function overestimating the observed histogram in the intermediate range of wind values and underestimating it in the lower and upper ranges. These discrepancies between the observed and adjusted distribution seem to often present a similar

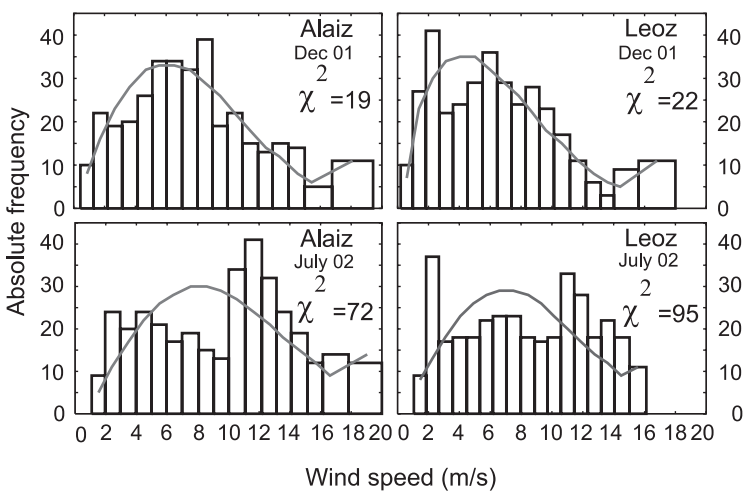

Figure 9. Absolute frequency histograms (bars) and adjusted Weibull probability functions (scaled to absolute frequencies) for December 2001 and July 2001 at Alaiz and Leoz. χ^2 values are indicated for each case

DOI: 10.1002/we

pattern of over-(under-) estimation in the intermediate (lower and upper) wind ranges, i.e. negative (positive) values of the $[f_{obs}(w_i) - f_{weib}(w_i)]$ terms in Equation (3). In order to illustrate this deformation in the observed histogram, a dispersion diagram in which points represent the differences in absolute frequencies, $N\,[f_{obs}(w_i) - f_{weib}(w_i)]$ (observed minus expected absolute frequencies), for all intervals, in every month within the specific wind farm, is represented in Figure 10, together with the line representing the average of all points. It can be appreciated that there is a tendency to under-(over-) estimate the Weibull frequency values in the lower and upper (middle) range of wind speeds, generating positive (negative) results in $N\,[f_{obs}(w_i) - f_{weib}(w_i)]$. This produces therefore positive (negative) contributions in Equation (3) as it is the difference $f_{obs}(w_i) - f_{weib}(w_i)$ that imposes the sign to the error contribution in $\xi_2 = wE_{ref} - wE_{weib}$.

This underestimation (overestimation) is more pronounced at Leoz and it is less noticeable at San Martín and Aritz. At Alaiz and El Perdón such deformation is also apparent in the experimental histogram. The error contributions of different sign tend, therefore, to be systematic in Figure 10 and can be partially averaged out in Equation (3). Thus in spite of having relatively large differences between the theoretical and actual frequency histogram, the impact of them can be at least partially balanced out by their sign, producing relatively small contributions in Figures 6 and 7. Additionally, the differences between both distributions are weighted by the factor, $p_{out}(w_i)$ in Equation (3). The most significant production of wE corresponds to the upper intervals of the histogram, that is, to the higher wind speed values, where the power generation is nearly the *rated power* (Figure 4). Thus, the weight applied by the factor $p_{out}(w_i)$ is considerably greater for these intervals than for the rest. This makes their contribution to the errors larger. A dispersion diagram representing all the products, $Np_{out}(w_i)[f_{obs}(w_i) - f_{weib}(w_i)]$ is presented in Figure 11. It shows that these products are negligible in the lower range of wind speeds as the energy produced here is also small. In the upper (intermediate) range of wind values, the net contribution of the products in Equation (3) becomes positive (negative) as shown in Figure 10. As the contribution to the error is larger for the upper range of wind speeds, the dominant effect is a positive error in $\xi_2 = wE_{ref} - wE_{weib}$ and thus an underestimation of the actual energy production as observed in

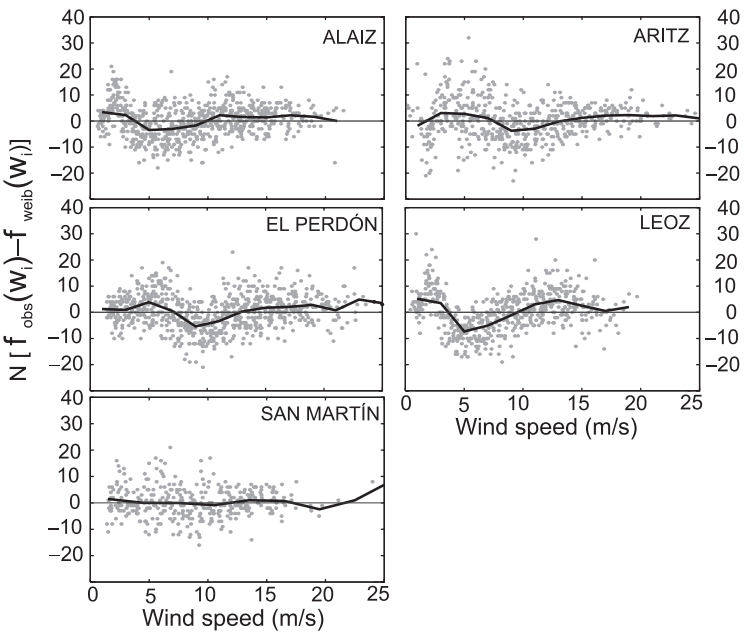

Figure 10. Differences $N\,[f_{obs}(w_i) - f_{weib}(w_i)]$ (points) in absolute frequency histograms between observed and Weibull adjusted probability functions (scaled to absolute frequencies). Differences are overlaid for each monthly histogram along the whole period of observations at each site. Averages of differences are calculated for wind speed intervals of 2 m s^{-1} and depicted with lines for each site

DOI: 10.1002/we

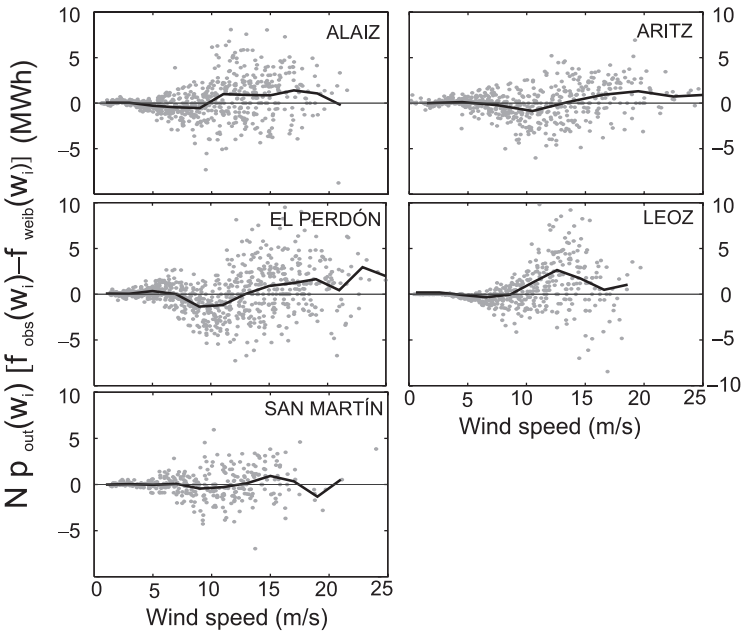

Figure 11. As in Figure 10 but for the terms $Np_{out}(w_i)$ $[f_{obs}(w_i) - f_{weib}(w_i)]$

Figures 6 and 7 and Table II (see section on the error in *wE* estimation). The underestimation is larger for the sites where the positive contribution of the upper range of wind speeds is larger in comparison with that of the intermediate speeds. This is the case for Leoz, Alaiz and El Perdón, the sites that displayed larger ξ_2 averages in Table III. Therefore, a larger level of underestimation of the monthly wE_{obs} is related to larger discrepancies between the observed and the Weibull modelled distribution for the higher wind speed bins.

The deformation of the wind speed histograms analysed in the previous paragraphs additionally allows for an understanding of why the time concordance between monthly χ^2 and ξ_2 values is hardly noticeable in Figure 8 for all wind farms, except for the case of Leoz. It could be expected that, as in the case of Leoz, the other locations would evidence links in the χ^2 and ξ_2 changes with time. The lack of this clear relation can be easily understood if it is recalled that the differences between observed and Weibull frequencies, $f_{obs}(w_i) - f_{weib}(w_i)$, take part in Equations (3) and (6) in a different way. While in their contribution to χ^2 is always positive, in ξ_2 positive or negative resulting differences can partially cancel their respective contributions. Therefore, it is possible that small ξ_2 values are associated with large χ^2 values. This fact disguises the expected relation between χ^2 and ξ_2. Alternatively, small χ^2 values would be indicative of small frequency differences, but can contribute to the final ξ_2 value differently, depending on the w range they are produced at and on the $p_{out}(w_i)$ factor that relatively amplifies or attenuates them.

The absence of concordance for χ^2 and ξ_2, i.e. the partial cancellation of error, would be lessened if the differences $f_{obs}(w_i) - f_{weib}(w_i)$ were mainly positive or mainly negative, that is, if there was a dominant sign in the intermediate, and preferably, in the upper intervals of wind speed, where the production of *wE* is larger. That is the case of Leoz wind farm (Figures 10 and 11) where the highest correlations were found (0.6) between χ^2 and ξ_2. Alternatively, the three bands in which Figures 10 and 11 conceptually divide the range of wind speed can be considered separately and compared with χ^2, the rationale being the fact that the underestimation and overestimation discrepancies between monthly observed and Weibull adjusted histograms take place in a related way, i.e. the larger underestimation in the lower and upper bands, the larger the overestimation in the intermediate band. This can be argued on the basis that the sum over all relative frequencies is 1, thus, if the probability of occurrence in the Weibull histogram is lower than the observed frequency at the intermediate frequencies, the probabilities in the lower and upper wind ranges must necessarily be larger. Therefore, the ξ_2

DOI: 10.1002/we

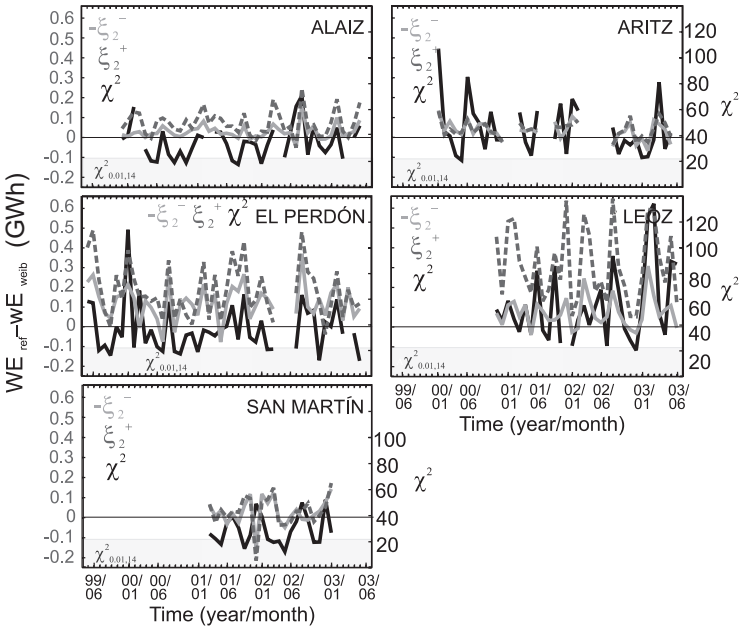

Figure 12. Time series of χ^2 and ξ_2 errors contributions to the low and upper (ξ_2^+) and the intermediate (ξ_2^-) wind speed intervals. Boundaries for the definitions of the three bands at each site were established from the intercepts of the error average lines in Figure 11 with the x axis. ξ_2^- is represented with positive sign ($-\xi_2^-$) to help perception of the visual agreement with the other lines

errors accumulated within the bands generating positive contributions (lower and upper) and that producing negative contributions (intermediate) in Equation (3) should behave coherently with time and display a more obvious relation respectively with χ^2 than the total ξ_2 added contributions in Equation (3). In order to illustrate this, the positive (negative) error contributions were computed using as boundaries the wind speed intercepts with the axis of the averages lines at each site in Figure 11. The resulting errors were denoted as ξ_2^+ (ξ_2^-) and plotted for comparison in Figure 12. These quantities verify that $\xi_2 = \xi_2^+ + \xi_2^-$. Their evolution in Figure 12 shows a much more consistent behaviour with changes in χ^2 as highlighted by the correlations values show in Table III (columns 6 and 7). The improved correlations with the different sign error contributions relative to those with the total ξ_2 (column 5 in Table III) highlight the fact that the lack of a clear association between χ^2 and ξ_2 in Figure 8 stems from the partial cancellation of errors of different sign (ξ_2^+ versusξ_2^-). In fact, Figure 12 allows for the discrimination of the dominant error contribution at each time step that justifies the occurrence of a large χ^2 value.

Conclusions

This work compared observed and estimated monthly wind energies. Observations were based on hourly wind velocity and wind power production measurements at five wind farms in the northeast of Spain (Figure 1).

As a first approach to the estimation of wind energy the best available information for the wind power versus wind relation and for the frequency histogram was considered for each month. This served as a means of quantification of the errors (ξ_1) inherent to the methodology. On average, a slight underestimation resulted which amounted to a maximum of 5.0% of the variability in one of the sites; individual peak error values seldom reached 10% of the target value.

A second approach incorporated the Weibull assumption, which contributed only slightly to underestimate target wind energies. Average error (ξ_2) biases ranged from 0.2% at Aritz to a maximum of 10.3% at Leoz.

DOI: 10.1002/we

Therefore, the assumption of the Weibull probability distribution did not seem to strongly impact the energy calculation. The reasons for this were not found to be based on a broad high quality of the Weibull fit but on a partial cancellation of error terms in the wind energy calculation.

An analysis of the goodness of the fit to the Weibull distribution based on the χ^2 statistic enabled an exploration of the relation between the quality of the wind speed fit and the discrepancies between observed and estimated wind energy. Such analysis revealed that the Weibull assumption could not be broadly substantiated at all the sites: the representativeness of the Weibull distribution varied from site to site and from month to month.

The relation observed between the χ^2 statistic and the estimated errors suggested that months that can not be strictly considered to be Weibull distributed, do not necessarily present higher errors in the wind energy estimation. The lack of an obvious relation between the error and the χ^2 statistic is due to partial cancellations of opposite sign contributions to the error in the calculation of the monthly energy which do not take place in the χ^2 computation.

The error contributions to the energy calculation and to the χ^2 statistic revealed that the Weibull distribution tends to overestimate the observed histograms in the intermediate wind speed range and to underestimate it in the lower and higher wind speed intervals. The effects of these departures from the observed frequency histogram tend to be averaged out due to the differences of sign. However, the product by the power terms in the calculation of wind energy weights more over the highest wind speeds, thus underestimation errors in this range tend to dominate and contribute to the systematic observed bias to underestimate total wind energy. It could be expected that sites/months with larger underestimations in the highest wind speed intervals disclose a much more evident relation between the energy estimation error and the performance of the Weibull probability distribution as it is the case of Leoz.

Acknowledgements

The authors wish to thank the two anonymous reviewers for their comments and suggestions. Also, we are indebted to Acciona Energía and especially Manuel Calleja for providing the data employed in this work. A special recognition for many interesting and enriching comments to Dr. D. Bray, Dr. M. Montoya, Dr. L. von Bremen and Dr. E. Zorita. This work was done under cooperation between UCM and CIEMAT (covenant ref. no 2003/89) and Project CGL 2205-06966-C07/CLI. JFGR was funded by the Ramón y Cajal programme.

References

1. IEA (International Energy Agency). *World Energy Outlook* 2006. International Energy Agency: Paris, France, 2006.
2. Ackerman T, Soder L. An overview of wind energy-status 2002. *Renewable and Sustainable Energy Reviews* 2002; **6**: 67–128.
3. Jager-Waldau A, Ossenbrink H. Progress of electricity from biomass, wind and photovoltaics in the European Union. *Renewable and Sustainable Energy Reviews* 2004; **8**: 157–182.
4. Flowers LT, Dougherty PJ. Wind powering America: goals, approach perspectives and prospects. *Sample Journals* 2004; **40**: 44–46.
5. Kenisarin M, Karsli VM, Caglar M. Wind power engineering in the world and perspectives of its development in Turkey. *Renewable and Sustainable Energy Reviews* 2006; **10**: 341–369.
6. Jamil M, Parsa S, Majidi M. Wind power statistics and an evaluation of wind energy density. *Renewable Energy* 1995; **6**: 623–628.
7. Bechrakis DA, Deane JP, McKeogh EJ. Wind resource assessment of an area using short term data correlated to a long term data set. *Solar Energy* 2004; **76**: 725–732.
8. Ramírez P, Carta JA. The use of wind probability distributions derived from the maximum entropy principle in the analysis of wind energy. A case study. *Energy Conversion and Management* 2006; **47**: 2564–2577.
9. Davenport AG. Past, present and future of wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics* 2002; **90**: 1371–1380.
10. Holmes JD. Effective static load distributions in wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics* 2002; **90**: 91–109.

DOI: 10.1002/we

11. Johannessen K, Meling TS, Haver S. Joint distribution for wind and waves in the Northern Sea. *International Journal of Offshore and Polar Engineering* 2002; **31**: 450–460.
12. Caires S, Sterl A. 100-Year return value estimates for ocean wind speed and significant wave height from the ERA-40 data. *Journal of Climate* 2004; **18**: 1032–1048.
13. Beljaars ACM. On the memory of wind standard-deviation for upstream roughness. *Boundary-Layer Meteorology* 1987; **38**: 95–101.
14. Wentink Jr T. *Wind power potential of Alaska. Part I*. Sci. Rep. Geophysical Institute, University of Alaska: Fairbanks, Alaska, USA, 1974.
15. Baynes CJ. The statistics of strong winds for engineering applications. PhD Thesis, Res. Rep. BLWT-4-1974. Faculty of Engineering Science, University of Western Ontario, 1974.
16. Celik AN. On the distributional parameters used in assessment of the suitability of wind speed probability density functions. *Energy Conversion and Management* 2004; **45**: 1735–1747.
17. Sherlock RH. Analyzing wind for frequency and duration. *Meteorological Monographs American Meteorological Society* 1951; **4**: 72–79.
18. García A, Torres JL, Prieto E, de Francisco A. Fitting wind speed distributions: a case study. *Solar Energy* 1998; **62**: 139–144.
19. Li M, Li X. MEP-type distribution function: a better alternative to Weibull function for wind speed distribution. *Renewable Energy* 2005; **30**: 1221–1240.
20. Hennesey JP. Some aspects of wind power statistics. *Journal of Applied Meteorology* 1977; **16**: 119–128.
21. Tuller SE, Brett CA. The characteristics of wind velocity that favor the fitting of a Weibull distribution in wind speed analysis. *Journal of Climate and Applied Meteorology* 1984; **23**: 124–134.
22. Palutikof JP, Davies PM, Davies TD, Halliday JA. Impacts of spatial and temporal wind speed variability on wind energy output. *Journal of Climate and Applied Meteorology* 1987; **26**: 1124–1133.
23. Jaramillo OA, Borja, MA. Wind speed analysis in La Ventosa, Mexico: a bimodal probability distribution case. *Renewable Energy* 2004; **29**: 1613–1630.
24. Pérez IA, Sánchez ML and García MA. Weibull wind speed distribution: numerical considerations and use with sodar data. *Journal of Geophysical Research* 2007; **112**: D20112. DOI: 10.1029/2006JD008278.
25. Conradsen K, Nielsen LB. Review of Weibull statistics for estimation of wind speed. *Journal of Climate and Applied Meteorology* 1984; **23**: 1173–1183.
26. Dorvlo ASS. Estimating wind speed distribution. *Energy Conversion and Management* 2002; **43**: 2311–2318.
27. Ramírez P, Carta JA. Influence of the data sampling interval in the estimation of the parameters of the Weibull wind speed probability density distribution: a case study. *Energy Conversion and Management* 2005; **46**: 2419–2438.
28. Celik AN. Assessing the suitability of wind speed probability distribution functions on wind power density. *Renewable Energy* 2003; **28**: 1563–1574.
29. Celik AN. Energy output estimation for small-scale wind power generators using Weibull-representative wind data. *Journal of Wind and Industrial Aerodynamics* 2003; **91**: 693–707.
30. Pryor SC, Schoof JT. Empirical downscaling of wind speed probability distributions. *Journal of Geophysical Research* 2005; **110**: D19109.
31. Chang TJ, Wu YT, Hsu HY, Chu CR, Liao CM. Assessment of wind characteristics and wind turbine characteristics in Taiwan. *Renewable Energy* 2003; **28**: 851–871.
32. Akpinar EK, Akpinar S. An assessment on seasonal analysis of wind energy characteristics and wind turbine characteristics. *Energy Conversion and Management* 2005; **46**: 1848–1867.
33. Weisser D, Foxon TJ. Implications of seasonal and diurnal variations of wind velocity for power output estimation of a turbine: a case study of Grenada. *International Journal of Energy Research* 2003; **27**: 1165–1179.
34. Biswas S, Sraedhar BN, Singh YP. A simplified statistical technique for wind turbine energy output estimation. *Wind Engineering* 1990; **19**: 147–155.
35. Frandsen S, Antoniou I, Hansen JC, Kristensen L, Madsen H, Chaviaropoulos B, Douvikas D, Dahlberg JA, Derrick A, Dunbabin P, Hunter R, Ruffle R, Kanellopoulos D, Kapsalis G. Redefinition power curve for more accurate performance assessment of wind farms. *Wind Energy* 2000; **3**: 81–111.
36. Bivona S, Burlo R, Leone C. Hourly wind speed analysis in Sicily. *Renewable Energy* 2003; **28**: 1371–1382.
37. Noorgard P, Holttinen H. A multi-turbine power curve approach. *Proceedings of Nordic Wind Power Conference*, Chalmers University of Technology, 2004.
38. Celik AN. Weibull representative compressed wind speed data for energy and performance calculations of wind energy systems. *Renewable Energy* 2003; **44**: 3057–3072.
39. Mathew S, Pandey KP, Kumar A. Analysis of wind regimes for energy estimation. *Renewable Energy* 2002; **25**: 381–399.
40. Carta JA, Ramírez P. Use of finite mixture distributions models in the analysis of wind energy in the Canarian Archipielago. *Energy Conversion and Management* 2007; **48**: 281–291.

41. Torres JL, García A, Prieto E, De Francisco A. Characterization of wind speed data according to wind direction. *Solar Energy* 1992; **66**: 57–64.
42. Hau E. *Wind Turbines*: Fundamentals, Technologies, Applications, Economics. Springer: Berlin, 2006.
43. Aubrun S, Koppman R, Leitl B, Mollman-Coers M, Schaub A. Physical modeling of a complex forest area in a wind tunnel-comparison with field data. *Agricultural and Forest Meteorology* 2005; **129**: 121–135.
44. Seguro JV, Lambert TW. Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis. *Journal of Wind Engineering and Industrial Aerodynamics* 2000; **85**: 75–84.
45. Li M, Li X. Investigation of wind characteristics and assessment of wind energy potential for Waterloo region, Canada. *Energy Conversion and Management* 2005; **46**: 3014–3033.
46. Pavia EG, Orien JJ. Weibull statistics of wind speed over the ocean. *Journal of Climate and Applied Meteorology* 1986; **25**: 1324–1332.
47. Deaves DM, Lines IG. On the fitting of low mean wind speed data to the Weibull distribution. *Journal of Wind Engineering and Industrial Aerodynamics* 1997; **66**: 169–178.
48. Weisser D. A wind energy analysis of Grenada: an estimation using the Weibull density function. *Renewable Energy* 2003; **28**: 1803–1812.
49. Pryor SC, Schoof JT, Barthelmie R. Climate change impacts on wind speeds and wind energy density in northern Europe: empirical downscaling of multiple AOGCMs. *Climate Research* 2005; **29**: 183–198.
50. Lun IYF, Lam JC. A study of Weibull parameters using long-term wind observations. *Renewable Energy* 2000; **20**: 145–153.
51. Justus CG, Hargraves WR, Mikhail A, Graber D. Methods for estimating wind speed frequency distributions. *Journal of Applied Meteorology* 1977; **17**: 350–353.
52. Takle ES, Brown JM. Note on the use of Weibull statistics to characterize wind-speed data. *Journal of Applied Meteorology* 1978; **17**: 556–559.
53. Fawzan MA. Methods for estimating the parameters of the Weibull distribution. INTER STAT 2000. Available: http://interstat.statjournals.net/. (Accessed 9 May 2007).
54. Gove JH. Moment and maximum likelihood estimators for Weibull distributions under length-and-area-biased sampling. *Environmental and Ecological Statistics* 2003; **10**: 455–467.
55. Akpinar EK, Akpinar S. A statistical analysis of wind speed data used in installation of wind energy conversion systems. *Energy Conversion and Management* 2005; **46**: 515–532.

DOI: 10.1002/we