An analogous downscaling methodology for predicting daily mean wind speed and daily gust wind speed over Iberia: applications for extremes

García, Amelia
Escuela Universitaria de Informática. Universidad de Valladolid, Dpto. Matemática Aplicada. Pza. Sta. Eulalia 9-11. 40005 Segovia. Spain. E-mail: amegar@eii.uva.es

Pascual, Álvaro
Facultad de CC Físicas. Universidad Complutense de Madrid. Dpto. Astrofísica y CC. de la Atmósfera. Ciudad Universitaria s/n. 28040 Madrid. Spain. E-mail: a.depascual@fis.ucm.es.

Martín, María Luisa
Escuela Universitaria de Informática. Universidad de Valladolid, Dpto. Matemática Aplicada. Pza. Sta. Eulalia 9-11. 40005 Segovia. Spain. E-mail: mlmartin@eii.uva.es

Sebastián, Luis Ignacio
Escuela Universitaria de Informática. Universidad de Valladolid, Dpto. Informática. Pza. Sta. Eulalia 9-11. 40005 Segovia. Spain. E-mail: lisebastian@infor.uva.es

Valero, Francisco
Facultad de CC Físicas. Universidad Complutense de Madrid. Dpto. Astrofísica y CC. de la Atmósfera. Ciudad Universitaria s/n. 28040 Madrid. Spain. E-mail: valero@fis.ucm.es.

Fernández, Miguel Alejandro
Escuela de Ingenierías Industriales. Universidad de Valladolid. Dpto. Estadística e Investigación Operativa. Paseo del Cauce, 59. 47011 Valladolid. Spain. *E-mail:* miguelaf@eio.uva.es

1. Introduction

From last decades, the world is seeing a progressive increase of interest on using sustainable and clean renewable energies. In this framework, wind energy is seen today as a proven useful technology. The knowledge of atmospheric patterns, mainly those configurations dealing with atmospheric patterns conducive to risky meteorological situations related to extreme wind events, is especially important for wind energy applications. In the present study, from a weather pattern classification obtained by multivariate methods, forecasts of wind speed and wind gust fields are analyzed by means of an analogous downscaling methodology. Wind speed and wind gust results, obtained from similar atmospheric analog patterns, are compared by means of several deterministic statistical and probabilistic tools.

2. Data and Methods

Several data sets have been used in this study: daily sea level pressure (SLP) at 12:00 UTC over the North Atlantic Ocean and Europe, daily observational mean wind speed (MWS) and daily observational wind gusts (WGU) for several stations distributed over Iberia. The datasets have been seasonally grouped: Winter (DJF), Spring (MAM), Summer (JJA) and Autumn (SON). SLP data are a product of the ERA40 Reanalysis (Gibson et al., 1997). Wind data used in this paper consists of time series of MWS from 23 sites over the Iberian Peninsula (station locations will be displayed in Fig. 1), covering the 1980-2001 period. From the daily data, daily mean wind speed field has been obtained over the 22-year period. On the other hand, the WGU data consists of 73 time series (station locations will be displayed in Fig. 1) covering the same time period.

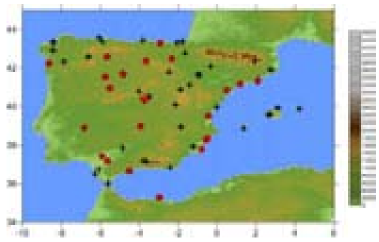


Figure 1: Wind speed stations (red) and wind gust stations (black) for the Iberian Peninsula with its orography detailed. The x-axis corresponds to longitude, positive (negative) for East (West).

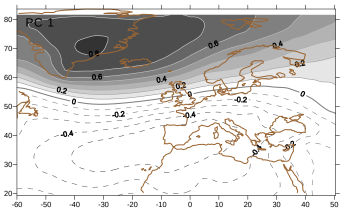
In order to extract the general behaviour hidden in the data set, a principal component analysis (PCA) was applied to the SLP database (Preisendorfer, 1988). The S-mode decomposition enables patterns to be identified that can be attributed to specific physical processes by statistical assessment. The new uncorrelated variables are called principal components (PCs) and consist on linear combinations of the original variables derived from the diagonalization of the covariance/correlation matrix. The coefficients of the linear combinations represent the weight of the original variables in the PCs and they are named loadings or PC patterns. The projection of the original series onto each eigenvector gives as result the time-dependence coefficient named scores or PC time series. On the other hand, a Wavelet Multiresolution Analysis (WMA) of the PC time series has been carried out. The wavelet transform technique was introduced and formulated by Morlet et al. (1982) and Grossmand and Morlet (1984). Wavelet transforms have been applied successfully to different studies of meteorological and climatological time series in order to understand their temporal scales of variability (Gamage and Blumen, 1993; Morata et al., 2008), highlighting the advantages of the technique compared with Fourier transform analysis. A more traditional correlation analyses based on spectral analyses and Mann-Kendall tests have also been applied to the PC time series (Goossens and Berger, 1986) as well as the wavelet analysis.

Within the framework of the European Project SafeWind, several works based on multivariate methodologies have being developing for obtaining atmospheric situations analog to a situation associated with extreme winds (Martin et al., 2010; Pascual et al., 2010). Searching analogs is based on the analysis of atmospheric situations, finding similar atmospheric situations. Using the analog methodology in this paper, from an "input" atmospheric situation, several similar weather situations are found in diverse dates, so finding and generating historic situations with high meteorological similitude. Thus, diverse wind fields of such dates are selected and from them, a forecast wind field can be offered in a study area. In the algorithm process of any analog model it is needed a weighting function (Cofiño, 2004) that considers the similitude of a situation to the past situations. In the present work, several distance functions have been used and proposed (Pascual et al., 2010) based on the use of different multivariate methodologies, allowing reducing the dimensionality in finding analogs and maximizing the explaining variance in the data bases. Using all these methods, similar atmospheric situations to any particular input situation to be modeled will be determined and, thus several and different wind fields come from the observational MWS and WGU will be obtained, trying to characterize and predict the wind field over Iberia.

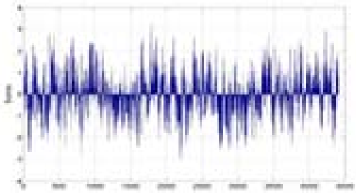
3. Results

PC spatial and time results

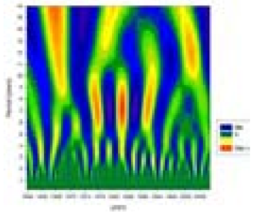
For reasons of brevity, the explanations about the obtained PC will be reduced to the first component of the winter season although the analog methodology is finally applied to fifteen PCs (accounting for about 76% of the total variance) associated with the four seasons. The leading PC accounts for about 21% of the total variance for the SLP field. The first SLP spatial pattern (Fig. 2a) is characterized by a north-south dipole configuration with high correlation values. One of the two centres is located centred over the Iberian Peninsula while the other one is situated eastward to Greenland. The spatial pattern shows a meridionally gradient similar to the well known North Atlantic Oscillation, NAO, teleconnection pattern (Barnston and Livezey, 1987). The reinforcement of high pressure in midlatitudes promotes extreme westerly wind across the North Atlantic Ocean, below normal temperatures in western Greenland and above normal temperatures over the eastern area of the United States and northern Europe.



a)



b)



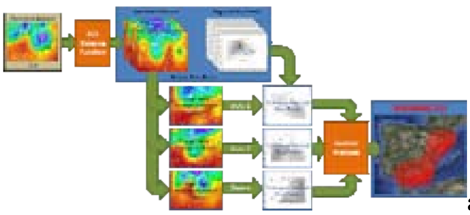
c)

Figure 2: (a) The leading PCA pattern of SLP in wintertime. Contours indicate correlation values. (b) Time series of PC corresponding with the first mode of SLP. (c) The wavelet power spectrum of the time series corresponding to the first PC of SLP. The y-axis represents the variability scale (years) and the x-axis corresponds to the time period.

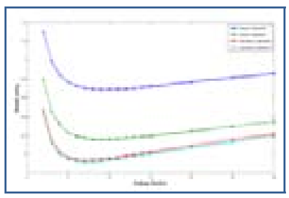
The time variability of the first mode is characterized by a marked interannual character (Fig. 2b). No significant trends are found in the time series. The power spectrum (Fig. 2c), displayed as a function of period and time, associated with the first PC time series of the SLP field is mainly characterized by scales evolving between 5 and 10 years, throughout the whole time period (1958 to 2007), showing power spectrum intensity mainly concentrated between 1974-1994, with a noticeable evolution of maximum/minimum nucleus. There are energetic oscillations, mainly negative nuclei, in scales below 3 years, indicating low intra-year and inter-year variability at such periods.

Analog results

Taking into account the obtained PC patterns in the four seasons, the analog methodology has been used to find comparable atmospheric situations (historic SLP fields) in each season, similar and related to an atmospheric situation (input variable). For each season, an input field score (SLP) is compared with fifteen scores of the fifteen obtained PC of the SLP field (Fig. 3a) in order to find the most similar scores throughout the historic scores time record (in our case, 1958-2007). Once the close scores have been obtained, their corresponding dates present associated MWS and WGU fields that finally give a predicted Iberian wind field.



a)



b)

Figure 3: (a) Diagram showing the steps of the used analog method. (b) The rmse associated to the estimation of SLP by applying different metrics in the distance function.

As it is abovementioned, in the algorithm process of any analog model, functions weighing the similitude of a situation to the past situations have to be used. In the present work, several metrics based on different distances between the input score and the scores obtained from the PCA have been used. The number of analogs used to finally predict the wind field (Fig. 3b) has also been taken into account (Pascual et al., 2010). To do this, a decomposition of the input fields on the vectorial space obtained from the multivariate analysis of the historic atmospheric situations is made. The root mean squared error (rmse) associated to the estimation of the SLP by using different metrics in searching analogs to an input atmospheric situation have been derived. For wintertime, different color lines (Fig. 3b) show the errors obtained by applying different metrics in the distance function. It is clear that there are differences in the results associated to the diverse metrics. The final metric used to validate the SLP field have been the metric with smaller errors (light blue line), based on statistically adjusted scores. Thus, similar atmospheric situations to any particular situation to be modelled have been determined and, thus several and different wind fields come from both the observational MWS and WGU have been obtained and averaged, characterizing the wind field over Iberia and giving a forecasted wind field.

From the deterministic point of view, several statistical tools are used to show and to compare the results of the MWS and the WGU. On the other hand, the probabilistic performance of a wind forecast can be evaluated following the difference between a forecast probability distribution and the observed probability distribution. Regarding the probabilistic approach, the probabilistic point of view is shown by means of the Brier Skill Score, BSS,

$BSS = 1 - \frac{BS}{BS_{ref}}$ where BS indicates the Brier Score, $BS = \frac{1}{n}\sum_{i=1}^{n}\left(p_i - o_i\right)^2$

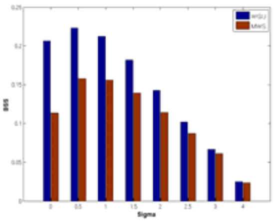
being p_i and o_i the predicted and observed probability of a value and n, the time. The BS_{ref} used in this work corresponds to the climatological value (the average value in 30 years).

Deterministic results are shown in Table 1. The bias, rmse and correlation values are derived between the areal averaged observational and forecast data for both MWS and WGU. The values point to significant relationships, that is, the input and the output data are strongly related. Therefore, the input in the model (SLP) involves enough information in the output (SLP) fields, and consequently in the wind results. The bias presents very small value for the two cases and the correlations are quite high, highlighting the existent relationships between the large-scale atmospheric field and the regional winds. Although the rmse values are in general small, it is clear that the results of MWS and WGU are not comparable between them because of the different variability range of each variable.

	MWS	**WGU**
bias (m/s)	-0.1	-0.2
rmse (m/s)	1.4	3.0
r	0.6	0.7

Table 1: Bias, rmse and Correlation (r) between the observational and the forecast data for MWS and WGU.

Several probabilistic verification results are also shown in contrast to the deterministic verification. The BSS is derived taking into account several thresholds, in function of the values of the standard deviation, σ, of the original data, MWS and WGU. The BSS for several σ's shows an asymmetric shape (Fig. 4a), being the best result associated with the value of σ = 0.5 in both cases although the WGU result is better than the MWS one. The model shows the best results very close to the observational MWS and WGU means; for such threshold, forecasts are better than the climatologic values. For extreme gust values, associated with the BSS ≈ 0.12 for σ ≥ 2, a gust range of 16 - 26 ms^{-1} is recorded (not shown), locating the high values on the Ebro Valley and on the Gibraltar Strait (≈ 93 kmh^{-1}). It is worth to note that for extreme values, MWS and WGU, show similar behaviour.


a)

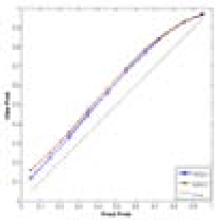
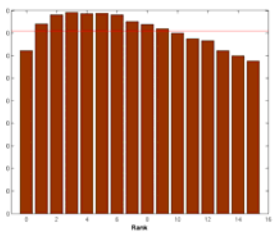
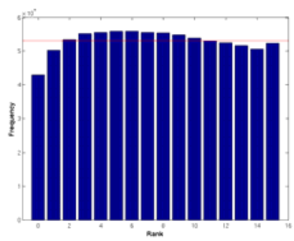

b)

Figure 4: (a) Illustration of BSS values for different thresholds, σ's. (b) Reliability diagram.

The reliability curve (Fig. 4b) is the relationship between forecast probability and observed frequency. Therefore, the best curve is a line along the 45º diagonal. The reliability curves deviate from the best line, and the model forecast probability is smaller than observed frequency, underestimating the observational wind frequencies in both cases. Again, the WGU behaviour is quite better than the MWS one.

The rank histogram distribution shows the equally much verification in each analog forecast interval. In an ideal system, the Talagrand diagram should be flat with much verification regularly in each interval, meaning exchangeability between deterministic predictions and observed data. The distributions (Fig. 5a and b) are slightly ∩–shaped, indicating a large spread, with many observations falling outside the extremes of the analog estimations. Once more, the WGU data (Fig. 5b) present quite better behaviour than the MWS one (Fig. 5a).


a)


b)

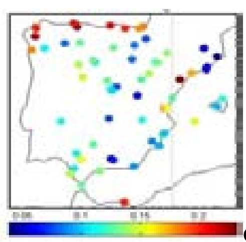

c)

Figure 5: Illustrations of the rank histogram of: (a) MWS; (b) WGU. (c) BSSs for σ = 0 over the stations shown in Fig. 1b.

The BBS is also derived for the mean value in each Iberian WGU stations and displayed in Fig. 5c. It can be noted how the analog model present the best results (> 0.2) over the Northern Iberia (Fig. 5c), while the remainder stations of Iberia show small values. It means that the atmospheric situations coming from the Atlantic Ocean are better to predict gusts in the

Northern Peninsula. On the contrary, the worse forecasts are located in the inner and eastern Iberian areas.

4. Conclusions

Daily MWS and WGU have been predicted from the principal components of large-scale SLP fields, after applying an analog method. The analogous method used in this paper compares a day circulation pattern with each one of the historical daily weather classifications which are obtained from the PCA. Several deterministic and probabilistic results have been shown.

Although the biases for the two variables show small values, the correlation is reasonably high, underlining the relationships between the large-scale atmospheric field and the regional wind field. Regarding the probabilistic approach, BSSs have been obtained taking into account several thresholds in function of the standard deviations of the original data.

The BSSs show the best results for WGU forecast values, being those ones very close to the observational mean. For extreme values, both variables show similar behaviour. The reliability curves deviate above from the best line, indicating that the model underestimates the observational wind gust frequencies for both WGU and MWS.

The rank histograms indicate large spread, with many observations falling outside the extremes of the analog estimations. Finally, the analog model has shown the best results over the Northern Iberia, where the atmospheric situations are not modified by local effects or orography.

Acknowledgements

This work has been partially supported by the research projects AYA2010-22039-C02-02, CGL2007-61328/CLI, UE Safewind G.A. No. 21374, AYA2009-14212-C05-05/ESP and VA025A10-2. The authors wish to thank the Agencia Estatal de Meteorología (AEMET) for providing the Iberian wind datasets and the European Centre for Weather Medium Forecast (ECWMF) for providing the ERA40 data.

REFERENCES

Barnston A.G. and R.E. Livezey. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Month. Wea. Rev*., 115, 1083-1126, 1987.

Cofiño, A. S. *Técnicas estadisticas y neuronales de agrupamiento adaptativo para la predicción probabilística de fenómenos meteorológicos locales. Aplicación en el corto plazo y en la predicción estacional.* Tesis doctoral, Universidad de Cantabria. 2004.

Gamage N, Blumen W. Comparative analysis of low-level cold fronts: Wavelets, Fourier, and empirical orthogonal function decomposition. *Month. Wea. Rev.,* 121: 2867-2878. 1993.

Gibson JK, Kallberg P, Uppala S, Hernandez A, Ñomura A, Serrano E. ERA description. *Technical Report Re-analysis Project Report Series 1*, ECMWF, Reading, UK1997.

Goossens C, Berger A. Annual and seasonal climatic variations over the Norther Hemisphere and Europe during the last century. *Annales Geophysicae* 4: 385-400. 1986.

Grossmand A, Morlet J. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.* 15: 732-736. 1984.

M. L. Martín, F. Valero, A. Pascual, A. Morata, and M. Y. Luna. Springtime connections between the large-scale sea level pressure field and gust wind speed over Iberia. *Nat. Hazards Earth Syst. Sci.,* 10, 1-13. (2010).

Morata A., Martín ML, Sotillo M, Valero F, Luna MY. Iberian autumn precipitation characterization through observed, simulated and reanalysed data. *Advances in Geosciences,* 16: 49-54. www.adv-geosci.net/16/49/2008/.2008.

Morlet GA, Fourgeau I, Giard D. Wave propagation and sampling theory. *Geophysics*, 47: 203-236. 1982.

Pascual, A., M. L. Martin, F. Valero, D. Santos-Muñoz, A. Morata and M. Y. Luna. Development of an analogous model for wind prediction using principal components. *SAFEWIND Workshop, Oldemburg*, Germany. 2010.

Preisendorfer RW. *Principal Component Analysis in Meteorology and Oceanography.* Elsevier Science Publishers BV: Amsterdam; 425. 1988.