Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 40 (2013) 67 – 76

Energy Procedia

European Geosciences Union General Assembly 2013, EGU

Division Energy, Resources & the Environment, ERE

On the application of Principal Component Analysis for accurate statistical-dynamical downscaling of wind fields

Roberto Chávez-Arroyo[a,b], Sergio Lozano-Galiana[b], Javier Sanz-Rodrigo[b], Oliver Probst[a]*

[a]*Physics Department, Tecnológico de Monterrey, Eugenio Garza Sada 2501, Monterrey 64849, Mexico*
[b] *National Renewable Energy Centre (CENER), Ciudad de la Innovación 7, Sarriguren 31621, Spain*

Abstract

A new methodology for the accurate statistical-dynamical downscaling of surface wind fields for long-term periods is described in this work. This new method is based on stratified sampling of long-term mean Sea Level Pressure fields combined with Principal Component Analysis for the determination of the most representative synthetic year. Validation is performed with 9 years of dynamically downscaled wind fields for the Iberian Peninsula obtained with the mesoscale model SKIRON. The results show that compared to the traditional method of dynamically downscaling random annual periods, the error in the predicted average wind speed is reduced by almost 30%.

Keywords: statistical-dynamical downscaling; long-term wind resource assessment; principal component analysis; mean sea level pressure

1. Introduction

Methods providing an accurate estimation of the long-term wind resource are important for a range of applications including wind resource assessment, where an accurate knowledge of the long-term wind

* Corresponding author. Tel.: +52-81-8358-2000
E-mail address: oprobst@itesm.mx

doi:10.1016/j.egypro.2013.08.009

resource at mesoscale levels is required both in the early prospection and project development phases, as well as for financing and insurance. Since extended high-quality wind measurement records as suitable references are not always available in many parts of the world, mesoscale modelling or downscaling tools emerge as a natural alternative. The basic idea is the construction of regional wind maps derived from long-term large-scale data sets such as global reanalysis data sets. Three main approaches can be distinguished: (1) Dynamical , (2) statistical, and (3) statistical-dynamical.

In the dynamical method (e.g. [1-4]) a large-scale data set is downscaled to the regional level by using a dynamical computational model of the atmosphere such as WRF-ARW [5] or SKIRON [6] based on temporal and spatial initialization by the global model. Available long-term global data sets are thereby translated into corresponding regional maps from which time series for the assessment of long-term averages and fluctuations can be extracted. Although the dynamical downscaling of either global or otherwise coarse numerical weather data is the most physically consistent and in principle most accurate method to simulate the wind characteristics at a regional level, the computational cost of downscaling a long-term global data set to the regional level is considerable, and the computing time achievable at standard facilities is often not acceptable. An alternative approach are statistical procedures (e.g. [7-10]) which attempt to establish statistical relationships between large-scale and mesoscale phenomena and use these relationships for the long-term prospection at the site of interest. A hybrid class of methods combining the strengths of both preceding methods at a reduced computational cost are statistical-dynamical methods [11-25]. The underlying idea in most statistical-dynamical downscaling (SDD) methods is the assumption that the climate of a region can be classified into several typical weather situations which should be sufficiently characteristic of the long-term behaviour of the phenomena under study. Afterwards, each group is integrated with a high-resolution numerical weather model to predict the local climate features [15].

Although several authors have described SDD methods, based on either stationary classification schemes [11-20] or on the classification of weather episodes of different duration using linear and non-linear methods [21-26], it is worth noting that comparisons against purely dynamical procedures are scarce, in particular for the wind statistics important for wind power engineering such as the seasonal and daily wind profiles, wind frequency roses and average wind speed.

SDD methodologies based on the stationary classification of synoptic-scale weather regimes are generally physically meaningful and often lead to a reasonably small number of classes. There are however several important downsides limiting their usefulness for wind prospecting purposes. First of all, the statistical treatment of the input data leads to a loss of time-dependent phenomena. Secondly, the simulation of daily or seasonal phenomena, required for a meaningful validation, is difficult because of the large number of classes required for such a distinction. An alternative are SDD methods based on algorithmic approaches that look for dynamical or quasi-dynamical weather episodes; such approaches allow for the simulation of entire weather cycles, and their predictions are easier to compare with measurements. Fuentes and Heimann [20] selected particular multi-day mesoscale episodes by hierarchical cluster analysis in combination with a spatio-temporal distance measure to aggregate consecutive dates with similar synoptic patters described by the first Principal Components of the geopotential at 500hPa. They then dynamically downscaled those representative episodes selected by the k-means cluster algorithm applied to the episodes with the same temporal length. Hagemann [21] proposed the use of Self-Organizing Maps (SOM) to select a representative continuous set of 365 days which were afterwards downscaled dynamically. The selection was based on the evaluation of those 365-days sets that best matched the frequency of occurrence of each SOM node (clusters) compared to the long-term clusters' frequencies. Similarly, Hahmann et al. [22] also proposed the application of SOM for the generation of wind speed atlases by using SOM as a synoptic classification tool to select representative samples of large scale wind forcing. Rife et al. [23] also identified representative periods

and based their selection on testing the similarity between the long-term period and a random sample drawn from a large number of reduced daily sets. The similarity metric was based on wind speed and direction distributions obtained from the first native vertical level of the MERRA global reanalysis data set for the point that best correlated with the measurements. In a similar approach Tammelin et al. [24] dynamically downscaled 4 years of different monthly periods requiring the geostrophic wind components (winds at 850hPa) of each 48-month period to be as representative as possible of the wind speed and direction distributions of the reference 19-year long-term period. In a recent work, Martínez et al. [25] use Principal Component Analysis to determine the Empirical Orthogonal Functions (EOF) corresponding to the long-term large-scale climate and downscale them to the regional scale. By applying the Principal Component time series of the large-scale data set to the downscaled version of the corresponding EOF they were able to extract time series which could be compared to measured wind speed records.

In the present work, a new methodology is proposed which is based on the determination of an optimal set of 365 daily episodes which best represents the long-term large-scale climate in a region of interest. Maps of the mean sea level pressure (SLP) field from the NCEP-DOE global reanalysis II project [27] are used as a proxy for the large-scale long-term climate. Once this optimal or representative 365-day large-scale climate set has been determined, it can be used for dynamical downscaling for the creation of regional surface wind field data sets. The selection of this period is based on the following steps: (1) Creation of a large number of monthly data sets by randomly sampling the long-term data base using a technique similar to the one by Rife et al. [23]. This approach is also known as *stratified sampling*. (2) The Empirical Orthogonal Functions (EOF) both for the long-term fields and the random 365-day sample are calculated and compared to each other by calculating the Euclidean distance between the two EOF sets. The optimal or representative sample is chosen by the requirement that the distance between the EOF sets should be minimal. (3) The results are validated by comparing the wind fields obtained from dynamically downscaling 9 years of NCEP/GFS (National Centers for Environmental Protection / Global Forecast System) global climate data for the Iberian peninsula using the SKIRON model [6, 27] against the wind fields extracted from the SKIRON results for the representative period determined in step (2). Since SKIRON runs were conducted on a daily basis (36-hour runs, discarding the initial 12 hours of each period to account for spin-up) this extraction procedure is equivalent to downscaling the 365-day representative period. Wind statistics relevant for wind resource assessment defined in section 2.3 are used for discussion.

2. Methods and Data

2.1. Selection of the most representative period

As mentioned above, the new method presented in this work belongs to the field of statistical-dynamical downscaling of large-scale climate data. As a suitable variable representing the synoptic meteorology we use the mean sea level pressure (SLP) field from the R-2 which has been successfully related to the wind speed variability in literature [28, 29]. Following the approach proposed by Rife et al. [23] a large number N of samples, typically N=100,000, composed of 365-day daily time series are created by random sampling the long-term SLP data set, ensuring that each sample contains 28|30|31-days per month; this is referred to as a stratified sampling technique. Then, each month of the random sample is compared to the corresponding average month of the long-term period by using the procedure described below.

The monthly (j-th month) anomalies for the long-term period and i-th sample set, X^j and Y^{ij}, respectively, are calculated by subtracting both the time average and spatial average from the sea level pressure (SLP) field (Lui et al. [31])

$$X(t,s)=\begin{pmatrix} x_{1,1} & \cdots & x_{1,S} \\ \vdots & \ddots & \vdots \\ x_{T,1} & \cdots & x_{T,S} \end{pmatrix} = SLP(t,s)-\overline{SLP}(s)-\overline{\overline{SLP}}(t) \tag{1}$$

Where T is the time and S the spatial dimension of the field $X(t, s)$; X stands for either X^j or Y^{ij}. Principal Component Analysis (PCA) is applied to both X^j and Y^{ij}, yielding an orthonormal base, commonly known as the set of Empirical Orthogonal Functions (EOFs), for each field. We will refer to these bases as $\{e_{i,m}{}^X\}$ and $\{e_{ij,m}{}^Y\}$, respectively. If $\{e_m\}$ is referred to as either set of EOFs, then the field X can be expressed as

$$X(t,s)=\sum_{m=1}^{P}\alpha_m(t)e_m(s) \tag{2}$$

Where the coefficients or amplitudes $\alpha_m(s)$ are usually referred to as the Principal Components (PC) of the field X. The basis $\{e_m\}$ is formally obtained from the eigenvalue problem of the covariance matrix Σ of X, and the corresponding eigenvalues λ_m represent the variance associated with each EOF e_m.

The idea behind the new method proposed in this work is to use the similarity between the sets $\{e_{i,m}{}^X\}$ and $\{e_{ij,m}{}^Y\}$ as a criterion of the representativeness of the sample Y^{ij} with respect to the long-term field X^j for a given month j. While in principle the full sets could be used for comparison, most of the variation of the field is contained in the first few eigenvalues of the covariance matrix Σ. This allows to significantly increase the computational efficiency of the procedure by limiting the comparison to the first M EOFs which capture a certain fraction of the total variance. In our case, we found 70% to be a convenient number, i.e.

$$\sum_{m=1}^{M}\lambda_m \cong 0.7\sum_{m=1}^{S}\lambda_m = 0.7\mathrm{Tr}(\Sigma) \tag{3}$$

Where S, as before, is the spatial dimension of X and therefore the dimension of the covariance matrix. In the second equality we have expressed the fact that the trace of a (square) matrix is equal to the sum of its eigenvalues and invariant under changes of basis. Consequently, the total variance can be calculated without having to calculate the full set of EOFs.

Once the reduced sets of EOFs have been determined, the sample R_j which is most representative of the long-term base is determined as the monthly sample that minimizes the Euclidean distance δ between the bases $\{e_{j,m}{}^X\}'$ and $\{e_{ij,m}{}^Y\}'$, where the apostrophe (') indicates that only the first M EOFs of each original base have been considered:

$$\delta_{ij}=\sqrt{\sum_{m=1}^{M}\left\|e_{j,m}^{X}-e_{ij,m}^{Y}\right\|^2} \tag{4}$$

$$R_j = X^{ij} \mid \delta_{ij}=\inf_i\left\{\delta_{ij}\right\} \tag{5}$$

An example of the results from the PCA decomposition method applied to the SLP fields is depicted in Figure 1 where the first two eigenvectors from the 33-years SLP anomalies for December are shown. These EOFs represent two of the most important SLP archetypes over Europe since their variance accounts for more than 40% of the total variance and from which we can find the well-known pattern

associated with the North Atlantic Oscillation [32] in the first EOF and the Euro-Asian configuration for the second EOF [33].

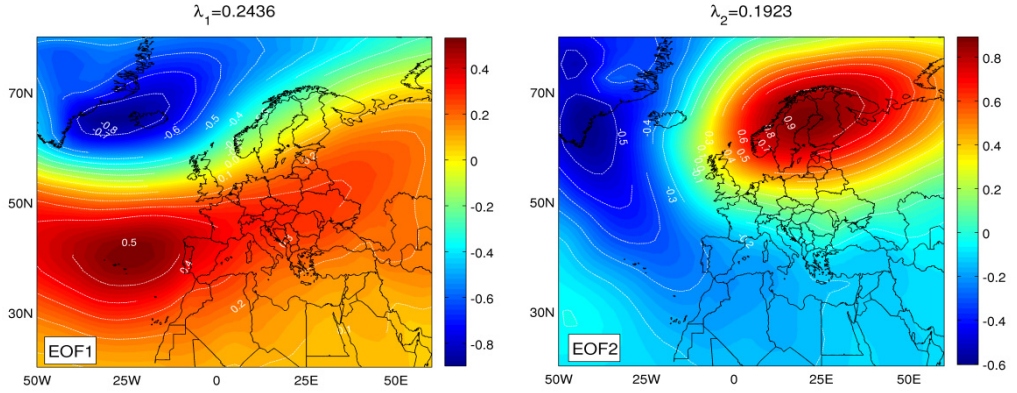


Fig. 1. First two EOFs from the 33-years SLP December anomalies with their corresponding eigenvalues indicated at the top.

In order to compare the new method to standard wind industry practices for the estimation of the long-term wind resource (as described by Rife et al. [23]) a reference method was also implemented, where the representative period is obtained as a random sample drawn from the long-term period. In this approach each calendar day of the representative year is determined from randomly selecting among the corresponding set of repetitions of that calendar day within the long-term period, in this case for a 33-year period. In the following we will refer to this method as the *traditional* one (TRAD).

2.2. Validation procedure –Iberian Peninsula case study.

The performance of the proposed statistical-dynamical downscaling method is assessed by the comparison of the wind fields obtained with the regional weather model SKIRON for a domain covering the Iberian Peninsula for the selected (representative) 365-days sample and the full 9-year period modeled; the model domain is shown in Figure 2 (dotted area). The SKIRON model is currently developed and supported by the Atmospheric Modeling and Weather Forecasting Group (AM&WFG) at the University of Athens (NKUA) [28]. This regional model is based on the Eta/NCEP model which in turn is built over the eta coordinates system in the vertical and over the semi-staggered Arakawa E grid in the horizontal [6].

The SKIRON wind fields are obtained by dynamically downscaling the 12UTC cycle analysis from the NCEP/GFS data down to a horizontal resolution of 0.05° x 0.05° with 45 eta levels in the vertical. Additionally, daily SST (Sea Surface Temperature) information is used for model initialization. The temporal horizon for the SKIRON runs is 36 hours; the first 12 hours are discarded to avoid the spin-up of the model (see Gaston et al. [3] for more details). The results from the downscaling runs are post-processed in order to obtain the wind values at 80m height by performing a power law interpolation from the eta levels.

For the PCA methodology a wide enough domain from the R-2 SLP data that includes Europe, northern Africa and part of Middle East is used in order to account for the synoptic patterns affecting the Peninsular region (Figure 2). The SKIRON domain is limited to an area slightly larger than the Iberian Peninsula with the purpose of leaving a buffer zone than can be used for the transition needed to accommodate the dynamical and physical differences between the GFS data used for initialization and SKIRON.

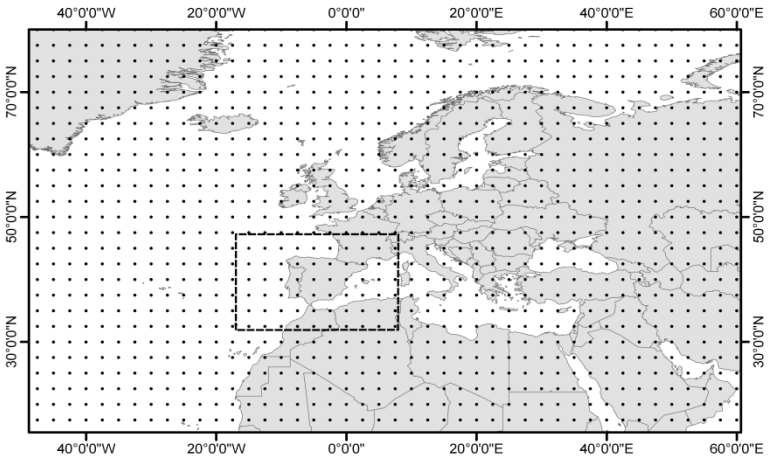


Fig. 2. European window used for the PCA decomposition of the SLP fields from the R-2 data set, where grid points are represented by black dots. The dotted black rectangle represents the domain covered by the downscaling performed with the SKIRON model.

2.3. Validation metrics.

Several metrics are used to compare the long-term (9-year) time series and the time series generated from the 365 selected days (representative sample R) for each point of the domain simulated with SKIRON. These metrics are inspired by wind statistics that are meaningful for wind resource assessment studies. The following metrics have been used in this work are described below:

The first metric is the average wind speed difference:

$$\varepsilon_S = \left\langle S^{LT} - S^{RP} \right\rangle, \tag{6}$$

where $S^{LT}(x, y)$ and $S^{RP}(x, y)$ are the wind speed maps (80m) for the long-term and the representative period, respectively. The second, third, and four metrics are the mean absolute error (MAE) and the correlation coefficients (CC) for hourly and seasonal profiles, respectively:

$$\left|\varepsilon_T\right| = \frac{1}{T}\sum_{p=1}^{T}\left|S_p{}^{LT} - S_p{}^{RP}\right| \tag{7}$$

$$\rho_T = \frac{\mathrm{cov}\left(S_T{}^{LT}, S_T{}^{RP}\right)}{\sigma(S_T{}^{LT})\sigma(S_T{}^{RP})} \tag{8}$$

Where $T = 12$ (months) or $T = 24$ (hours) for the seasonal and hourly wind speed averages and correlation coefficients, respectively. The last metric considered in this study is the weighted compound frequency histogram defined as:

$$D = \sum_{b=1}^{nbins} w_b \left| \frac{f_b{}^{LT} - f_b{}^{RP}}{f_b{}^{LT}} \right|, \tag{9}$$

Where f_b^{LT} and f_b^{RP} are the relative frequencies for the long-term and representative period computed in a

total number of bins given by *nbins* for the wind speed and the w_b is a weighting factor which was defined as the long-term frequency i.e. $w_b = f_b^{LT}$. The aforementioned validation procedure was carried out for the PCA method and then compared with the results obtained with the *traditional* method described above.

3. Results

An example of the results of the PCA methodology in terms of the SLP patterns can be observed in Figure 3 where the leading EOF for May is depicted for a) the period 2003-2012 as well as for b) the final selection made by the PCA methodology and c) for a randomly selected period (industry standard method). Despite the expected differences, it can be noticed that the di-pole system between Greenland and the Azores Islands present in the long-term fields is well reproduced with the novel method proposed in this work, but degenerated by the traditional method. This difference is illustrated by a comparison of the Euclidean distance (equation 4) between the long-term and the sample set; this distance is 1.42 for the traditional compared to 0.5 for the new PCA-based method.

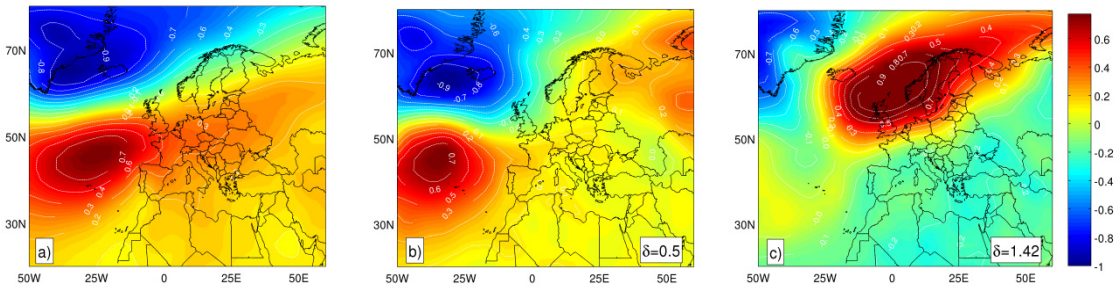


Fig. 3. First EOF for the SLP May anomalies for a) the long-term period, b) PCA the representative period selected by the new PCA-based method, c) the period selected by the industry-standard method (TRAD) . The associated Euclidean distance to the long-term EOF is shown in the lower right corner of figs b) and c).

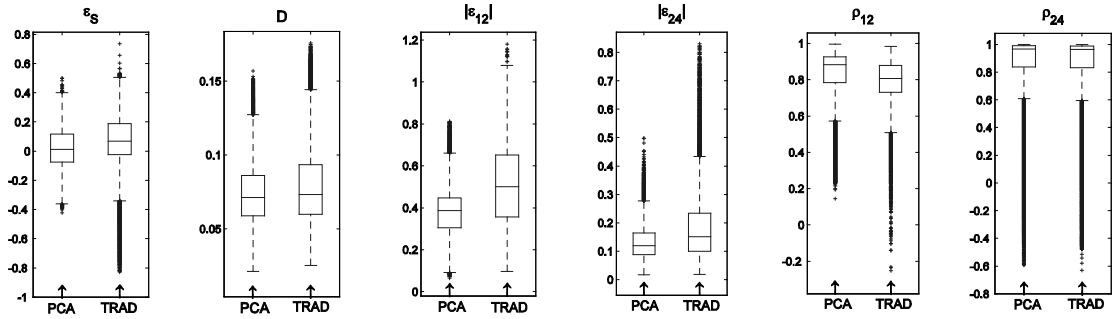


Fig. 4. Resulting metrics (equations 6 - 9) from the comparison between periods selected by the PCA method and the traditional method (TRAD) for the different wind-based statistics described in section 2.3

The influence of this synoptic system on the surface winds is shown in the boxplots of Figure 4 which illustrate the overall results obtained for the 99051 grid points in the SKIRON domain for the metrics defined in section 2.3. In all cases, except for the correlation coefficients, the lower the values of the statistics the better the prediction of the long-term wind resource. The more significant improvements can be observed for the difference in average wind speed ε_s and the monthly and daily wind cycle MAE ($|\varepsilon_{12}|$ and $|\varepsilon_{24}|$) in which not only the dispersion is reduced but also the median of all points is drastically reduced by ~80% , ~25% and ~20% for ε_s, $|\varepsilon_{12}|$ and $|\varepsilon_{24}|$) respectively.

These differences are illustrated in Figure 5 for a single point time series within the SKIRON domain (coordinates [42.65 , -7.75]) in which the wind speed frequency is shown together with their corresponding Weibull probability density function curves (Figure 5a) for the long-term, the new PCA-based method and the traditional approach (TRAD); the new method clearly provides an improved prediction of the long-term Weibull parameters. Figure 5 (b) shows the monthly and hourly wind profiles for these three data sets; in both cases the PCA-based method provides an improved prediction with a noteworthy improvement in the daily profile which is very relevant for time-of-day generation of wind farms.

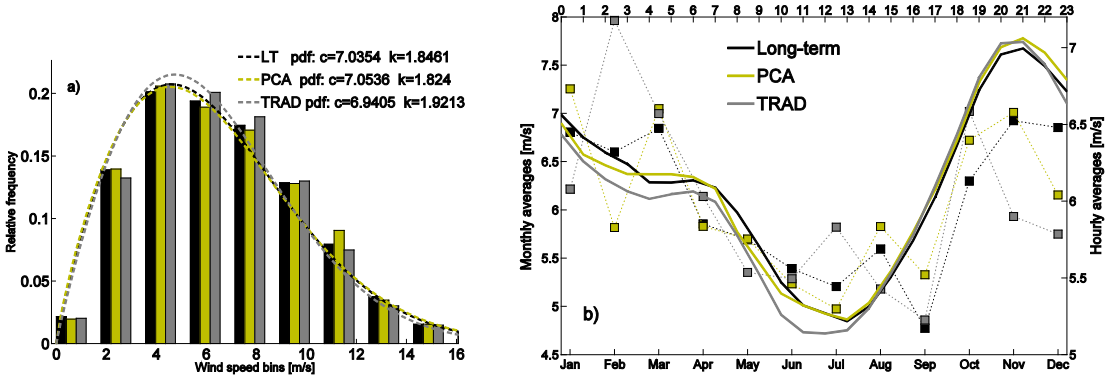
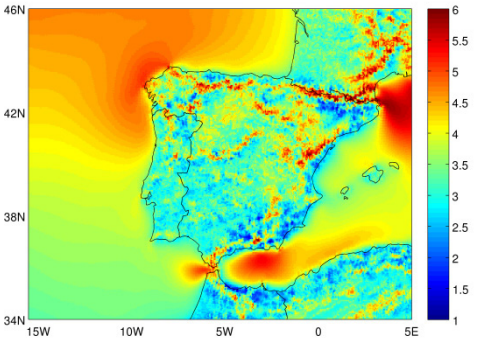


Fig. 5.Wind statistics for the virtual time series generated for one SKIRON point with a) the wind speed frequency plot together with their Weibull distribution curves (dotted lines) where the corresponding scale ("c") and shape ("k") factors are shown in the legends. The right- hand figure b) shows the monthly (dotted lines) and the hourly (solid lines) profiles for the three time series.



Regarding the spatial distribution of the errors, it can be observed from Figure 7 that the new PCA-based method produces a far more homogeneous error map than the industry-standard approach. Higher wind error statistics are related to areas with higher wind variability as illustrated in Figure 6. This coincidence was expected since extra variability is highly related to both windy areas and sites with strong influence of local phenomena.

Fig. 6. Standard deviation (m/s) for the SKIRON 9-years wind map.

4. Conclusions

A new approach for the statistical-dynamical downscaling for the generation of accurate long-term wind maps is presented in this work. The new method is based on the construction of a representative period of 365 daily episodes which best matches the long-term climate in a region of interest. The selection of this period is achieved through a method based on Empirical Orthogonal Function from a large number of sample candidates generated through a random process. As any statistical-dynamical downscaling method the new approach offers a very significant reduction of computing time and power over the direct downscaling of the long-term wind climate, while providing an accurate prediction not only of the average wind speed but also of wind velocity parameters relevant to wind resource assessment, including wind speed distributions and daily and monthly wind speed profiles. Compared to the industry-standard approach for the generation of long-term wind maps based on the downscaling of a

randomly selected annual period very significant improvements in prediction accuracy have been achieved. A particularly noteworthy improvement of this methodology is a better representation of the temporal variability with typical MAE values of 0.4m/s for the monthly and 0.15 m/s for the hourly profiles, which is around 30% better than the industry standard method and which can be considered small when compared to other important sources of uncertainties such as those related to the NWP physics.

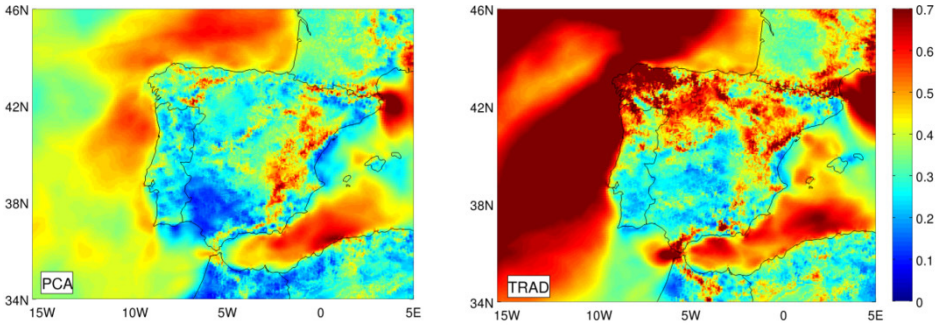


Fig. 7. Spatial distribution of the Mean Absolute Errors in the monthly profiles of the proposed Principal Component Analysis method (left) and the traditional method (right)

Acknowledgements

The financial support of Tecnológico de Monterrey through funding of the research chair for wind energy (contract number CAT158) as well as the support from the Mexican Council for Science and Technology (CONACyT) is gratefully acknowledged. A late stage of this project also benefitted from financial aid from the Marie Curie training program through the WAUDIT project at CENER.

References

[1] Larsén XG, Mann J, Berg J, Göttel H, Jacob D. Wind climate from the regional climate model REMO. Wind Energy. 2010; 13:279–96.

[2] Al-Yahyai S, Charabi Y, Gastli A. Review of the use of Numerical Weather Prediction (NWP) Models for wind energy assessment. Renew. Sust. Energ. Rev. 2010;14(9):3192–8.

[3] Gastón M, Pascal E, Frías L, Martí I, Irigoyen U, Cantero E, et al. Wind resources map of Spain at mesoscale. Methodology and validation on. EWEC conference proceedings. Brussels, Belgium. 2008.

[4] Hahmann AN, Rostkier-Edelstein D, Warner TT, Vandenberghe F, Liu Y, Babarsky R, et al. A Reanalysis System for the Generation of Mesoscale Climatographies. *J. Appl. Meteorol.* 2010; 49(5):954–72.

[5] Skamarock WC, Klemp JB, Gill DO, Barker DM, Duda MG, Wang W, et al. A Description of the Advanced Research WRF Version 3. Atmospheric Research. Boulder, CO; 2008.

[6] Kallos G, Nickovic S, Papadopoulos A, Jovic D, Kakaliagou O, Misirlis N, et al. The regional weather forecasting system SKIRON: an overview. *Proceedings of the International Symposium on Regional Weather Prediction on Parallel Computer Environments* 1997. pp. 109–122. Athens,

[7] Zorita E, von Storch H. The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods. *J. Climate.* 1999; 12(8):2474–89.

[8] Wilby RL, Wigley TML. Downscaling general circulation model output: a review of methods and limitations. *Prog. Phys. Geog.* 1997; 21(4):530–48

[9] Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO. Guidelines for Use of Climate Scenarios Developed from

Statistical Downscaling Methods. *Supporting material of the IPCC*. 2004; p 1–27.

[10] Pryor SC. Empirical downscaling of wind speed probability distributions. *J. Geophys. Res.* 2005; 110 p. D19109.1-D19109.12

[11] Wippermann F, Gross G. On the construction of orographically influenced wind roses for given distributions of the large-scale wind. *Beitr. Phys. Atmos*. 1981; 54, 492-501.

[12] Heimann D. Estimation of regional surface layer wind field characteristics using a three-layer mesoscale model. *Beitr. Phys. Atmosph.* 1986;59(4):518–37.

[13] Frey-Buness F, Heimann D, Sausen R. A Statistical-Dynamical Downscaling Procedure for Global Climate Simulations. *Theor. Appl. Climatol*. 1995;50:117–31.

[14] Fuentes U, Heimann D. Verification of statistical-dynamical downscaling in the Alpine region. *Climate Res*. 1996;7:151–68.

[15] Mengelkamp H. Statistical-dynamical downscaling of wind climatologies. *J. Wind Eng. Ind. Aerodyn*. 1997;67-68:449–57.

[16] Bergström H. Wind Statistics Mapping Using a Higher-order Closure Meso-scale Model. *EWEC conference proceedings* 2007.

[17] Mengelkamp H-T. Wind Climate Simulation over Complex Terrain and Wind Turbine Energy Output Estimation. *Theor. Appl. Climatol*. 1999 Aug 24;63(3-4):129–39.

[18] Frank HP, Landberg L. Modeling the wind climate of Ireland. *Bound. Lay. Meteorol*. 1997;359–78.

[19] Frank HP, Rathmann O, Mortensen NG, Landberg L. The Numerical Wind Atlas - the KAMM / W AsP Method. 2001 report available online at: http://www.risoe.dtu.dk/sitecore/content/Risoe_dk/Home/Knowledge_base/publications/Reports/ris-r-1252.aspx?sc_lang=en

[20] Fuentes U, Heimann D. An Improved Statistical-Dynamical Downscaling Scheme and its Application to the Alpine Precipitation Climatology. *Theor. Appl. Climatol*. 2000; 19;65(3-4):119–35.

[21] Hagemann K. Mesoscale Wind Atlas of South Africa. PhD thesis. University of Cape Town; 2008.

[22] Hahmann AN, Vincent CL, Division WE. Applications of Self Organizing Maps in Wind Energy Meteorology. *EMS Annual Meeting*. Zurich, Switzerland; 2010.

[23] Rife DL, Vanvyve E, Pinto JO, Monaghan AJ, Davis CA, Poulos GS. Selecting Representative Days for More Efficient Dynamical Climate Downscaling: Application to Wind Energy. *J. Appl. Meteor. & Climatol*. 2013;52:47–63.

[24] Tammelin B, Vihma T, Atlaskin E, Badger J, Fortelius C, Gregow H, et al. Production of the Finnish Wind Atlas. *Wind Energy*. 2013;16(2011):19–35.

[25] Martinez Y, Yu W, Lin H. A New Statistical–Dynamical Downscaling Procedure Based on EOF Analysis for Regional Time Series Generation. *J. Appl. Meteor. & Climatol*. 2013 Apr;52(4):935–52.

[26] Otero-Casal C, Miguez-Macho G, López P, Canoura F. Calculating a high-resolution Wind Resource Map with WRF using SOM clustering techniques. *12th WRF Users' Workshop*. 2011.

[27] Kanamitsu M, Ebisuzaki W, Woollen J, Yang S-K, Hnilo JJ, Fiorino M, et al. NCEP–DOE AMIP-II Reanalysis (R-2). *B. Am. Meteorol. Soc*. 2002 ;83(11):1631–43.

[28] Pytharoulis I, Katsafados P, Kallos G. The Weather Forecasting System SKIRON - Model configuration and setup. Athens, Greece; 2005.

[29] Conway D, Jones PD. The use of weather types and air flow indices for GCM downscaling. *J. Hydrol*. 1998 Dec;212-213:348–61.

[30] Davy RJ, Woods MJ, Russell CJ, Coppin P a. Statistical Downscaling of Wind Variability from Meteorological Fields. *Bound. Lay. Meteorol.* 2010;135(1):161–75.

[31] Liu Y, Weisberg R., He R. Sea Surface Temperature Patterns on the West Florida Shelf Using Growing Hierarchical Self-Organizing Maps. *J. Atmos. Oceanic Technol*. 2006;23:325–38.

[32] Hurrell JW, Van Loon H. Decadal variations in climate associated with the north atlantic oscillation. *Climatic Change*. 1997;36:301–26.

[33] Barnston AG and Livezey RE. Classification, Seasonality and Persistence of Low- Frequency Atmospheric Circulation Patterns. *Mon. Wea. Rev.* 1987; 115, 1083–1126