Wind Speed Evaluation of MERRA-2, ERA-Interim and ERA-5 Reanalysis Data at a Wind Farm Located in Brazil

Jhoseny Souza Santos[1], Yoshiaki Sakagami[1], Reinaldo Haas[1], Júlio César Passos[1], Mônica Nassar Machuca[1], William Corrêa Radünz[1], Eduardo Dias[2] and Mayara Miqueletti Lima[2]

[1] Federal University of Santa Catarina, Florianópolis (Brazil)

[2] Atlantic Energias Renováveis S.A., Curitiba (Brazil)

Abstract

Climatological normal is a powerful tool to understand the behavior of meteorological variables like wind speed, and it is typically employed in wind resource assessments. This paper compares three different sources of long-term wind speed data at a wind farm, located in the Northeast of Brazil, with meteorological mast observations at the site from 2009 to 2018. The main goal is to determine which one better describes the wind climate at the wind farm. All long-term sources, MERRA-2, ERA-Interim and ERA-5, showed satisfactory correlations with meteorological mast data for monthly means, which were 0.91, 0.94 and 0.95, respectively. Even though the wind speed observations were systematically underestimated, all reanalysis databases represented well the seasonal variability, but were not able to capture extreme winds. Although negative bias was reduced using a linear regression, it remained in the months of austral winter in all datasets. The correlation coefficient was not improved by the vertical correction of ERA-Interim wind speed, therefore the dataset from 1000 hPa represents satisfactorily the mean wind speed behavior at the target site.

Keywords: reanalysis, wind speed, correlation, wind farm

1. Introduction

According to the Global Wind Energy Council (GWEC, 2019), Brazil kept the 8th position in the cumulative installed capacity globally, which accounts for 14.7 GW shared between 12 states. Wind energy reached the second position in the Brazilian electricity matrix, which represents 9.0% of all electricity produced, most of which originated from wind farms located in Northeast Brazil (NEB). This significant share of wind power from NEB in the electricity matrix is partly explained by the atmospheric circulation (ABEEOLICA, 2019).

The unequal heating of the surface produces three cells called Hadley, Ferrel and Polar. The first one is located between the 0° and 30° of latitude and has the following mechanism: the warm air rises near the Equator and is redirected towards the poles near the tropopause. Around 30° (N/S), it descends because of the cooling. Due to the horizontal pressure gradient, the air flows back to the Equador to close the circulation. This equatorward motion is deflected to the left (right) in the Southern (Northern) Hemisphere due to the Coriolis force (AHRENS and HENSON, 2017).

The deflected wind defines the southeastern trade winds of the Southern Hemisphere which are strong and constant in both speed and direction in the ocean (WALLACE and HOBBS, 2006). On a seasonal scale, the strongest winds are verified during the austral spring and winter which match with the dry season of the region. Besides these global and seasonal scales, some wind farms located nearby elevated regions, known as a plateau, in the countryside of Bahia (Brazil) can be benefited by the mountain-valley circulation (SCHUBERT, 2013).

Regarding interannual oscillations, being able to estimate the long-term wind speed is essential for the wind industry to warrant their long term contracts, which are of 20 years in Brazil's energy auctions (Brazil, No. 10.438/02). Reanalysis data are typically used to investigate the long-term wind climate at a given site via an analysis entitled climatological normal. The World Meteorological Organization (WMO) defines climatological normal as the mean climate of a 30-year period, but recent studies showed that the length of this period can change according to site specific conditions (WMO-No. 100, 2011; WMO/TD-No. 1377).

doi:10.18086/swc.2019.45.10 Available at http://proceedings.ises.org

Parker (2016) defines reanalysis as a hybrid dataset which uses both data observation and forecasting models. The forecasting models, also known as Numerical Weather Prediction, provide an initial estimation of the atmospheric state and then the information from the observations, which can be from meteorological stations, airplanes and satellites, are assimilated to improve the forecasts. Hence, this data assimilation is a crucial part of the past long-term analysis construction.

Currently, the energy industry is studying the correlation of their datasets with reanalysis to obtain a reconstructed 20-year dataset for their site, as demonstrated in Nunes et al. (2019). Recently, Kim et al. (2018) obtained a correlation coefficient (R^2) of 0.61 between Modern Era Retrospective-analysis for Research and Applications - Version 2 (MERRA-2) and wind speeds measured at 100 m with an onshore meteorological tower located in South Korea. On the other hand, at three other onshore towers (30 m height) the correlations varied between 0.26 and 0.53. Regarding ERA-Interim, the best R^2 value found was 0.58 and the poorest correlation was 0.25, both for different towers with 30 m height. Other works as Zhang and Bao (2012), Liléo and Petrik (2011) and Kiss et al. (2011) also compared reanalysis datasets with *in situ* measurements for different regions.

In Brazil, Stuker et al. (2016) evaluated ERA-Interim's performance at 33 meteorological masts equipped with cup anemometers 10 m above the ground in the South of Brazil. The mean value of the Pearson Correlation Coefficient (R) was 0.70. Micalichen and Dias (2018) compared measured wind speed data from 17 meteorological stations over Minas Gerais with CFSR and CFSV2 reanalysis data and obtained a poorer correlation (0.39).

For the NEB, specifically Ceará, Vieira et al. (2006) correlated monthly mean wind speed of a meteorological mast with NCEP/NCAR reanalysis and the greatest values found were 0.78, 0.74 and 0.78 for March/1981, November/1978 and October/1977, respectively. Both Micalichen and Dias (2018) and Vieira et al. (2006) compared reanalysis data to meteorological station observations at 10 m height.

The studied literature reveals the highly site-specific performance of various sources of reanalysis data when compared to in situ observations. Therefore, the proposal of this work is to compare MERRA-2, ERA-Interim and ERA-5 wind speeds at a wind farm with complex terrain located in the NEB region from 2009 to 2018 with field observations of a meteorological mast at 78 m height, which is the hub height of the target site turbines, and thereafter determine whether they may be used to calculate the climatological normal.

2. Data and Methodology

2.1 Reanalysis Data

Three reanalysis data were used in this study: MERRA-2, produced by NASA, ERA-Interim and ERA-5, developed by European Centre for Medium-Range Weather Forecasts (ECMWF). MERRA-2 has temporal and grid resolutions of 1 hour and 0.5° x 0.625°, respectively. The zonal and meridional wind components from MERRA-2 were processed at 50 m above ground (Gelaro et al., 2017).

Unlike MERRA-2, ERA-Interim has a 6-hour resolution and a regular grid of 0.75° x 0.75° (Dee et al., 2011). ERA-Interim dataset was obtained at the 1000 hPa pressure level, whereas this reanalysis does not have the wind speed measured at a fixed height. ERA-5, the most recent reanalysis has a temporal resolution of 1 hour and spatial grid of 0.25° x 0.25° (Copernicus Climate Change Service, 2019) for wind speeds at 100 m above ground.

In this study, we used data from the node closest to met mast coordinate for both datasets. The linear distances associated were of 10050 m for MERRA-2 and 14400 m for ERA-Interim and ERA-5.

2.2 Meteorological Mast Data

For comparison purposes, the wind speed observations at Morrinhos wind farm, which is owned by the wind farm developer and operator Atlantic Energias Renováveis S.A., were recorded each 10-min using a cup anemometer at 78 m above ground level. The meteorological mast used in this study is located in the northern region of Bahia, as shown in Figure 1. Field observations of wind speed were also processed to filter outliers. Subsequently, all data were processed into hourly, 6-hours and monthly means and then normalized using the maximum value method due to data confidentiality issues.

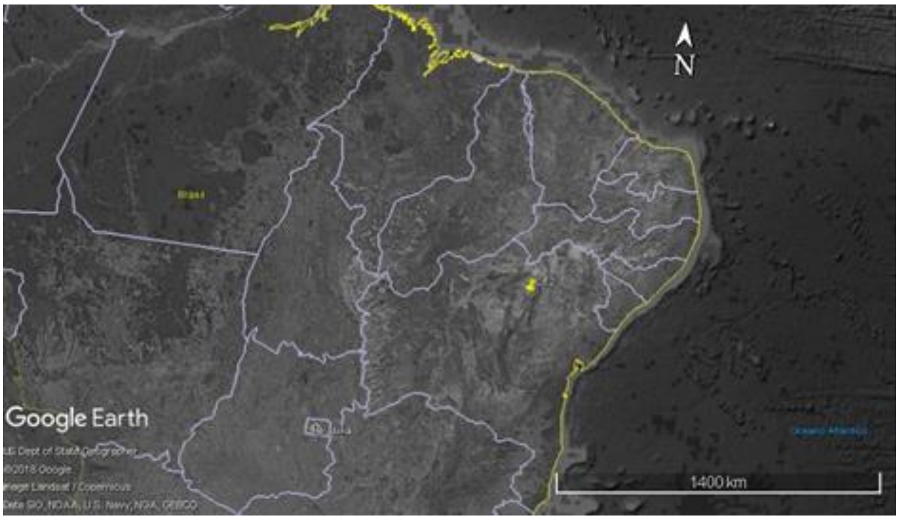


Figure 1: Morrinhos wind farm located at the yellow pin. Source: Google Earth (2019).

2.3 MCP Method

In many regions, there is a lack of reliable and long meteorological datasets, which may enhance wind resource assessment uncertainty. The alternative commonly used by industry is a method called Measure-Correlate-Predict (MCP), in which a short dataset of wind speed at the target site can be correlated with a longer dataset at another reference site to produce long-term time series at the target site. There are some studies, such as Carta et al. (2013) and Thøgersen et al. (2007), that compare different MCP methods. A linear regression approach was chosen for this study because it usually produces good results, as shown by Thøgersen et al. (2007) and Hanslian (2015). The wind speeds are linearly related by the following expression:

$$y = A * x + B \qquad \text{(eq. 1)}$$

where y and x are the meteorological mast and reanalysis wind speeds, respectively. For this study, both observations and reanalysis data encompass the period between November 2009 to December 2018.

2.4 Metric coefficients: Correlation and Bias

To verify how the reanalysis data and the meteorological mast data correlate, the Pearson Correlation Coefficient (R) was calculated using Equation 2 and for the bias calculation Equation 3 was used. Those are given by:

$$R = \frac{Cov(x,y)}{\sqrt{var(x)*var(y)}} \qquad \text{(eq. 2)}$$

$$Bias = \frac{1}{N}\sum_{N}^{i}(x-y) \qquad \text{(eq. 3)}$$

In all the above equations, x, y and N represent the reanalysis data, the meteorological mast data and the length of the dataset, respectively.

2.5 Periodogram using Fast Fourier Transform

The analysis of a meteorological variable using a decomposition of the time series into several frequencies is commonly used to detect strong oscillations. There are some studies, such as Van der Hoven (1957) and Hwang (1970), that applied the Power Spectral Density (PSD) in wind speed data near the surface to analyze the contribution of eddy kinetic energy. The proposal of this study is to apply the same tool but in a different timescale. The spectral analysis method consists of employing a Fast Fourier Transform (FFT), which takes the Discrete Fourier Transform (DFT) and is calculated by equation 4, to generate a periodogram defined at N/2 + 1 frequencies, where N is the length of the data, as follows:

$$Z_t(k) = \sum_{J=1}^{N} X(j) W_N^{(j-1)(k-1)} \qquad \text{(eq. 4)}$$

$$W_n = e^{(-2\pi i/N)} \qquad \text{(eq. 5)}$$

In equation 4, k is defined as an interval [1,(N/2)+1]. The PSD values are calculated by equation 6, excluding the zero frequency.

$$I(\omega) = 2 * \frac{1}{Fs*N} * |Z_t(k)|^2 \qquad \text{(eq. 6)}$$

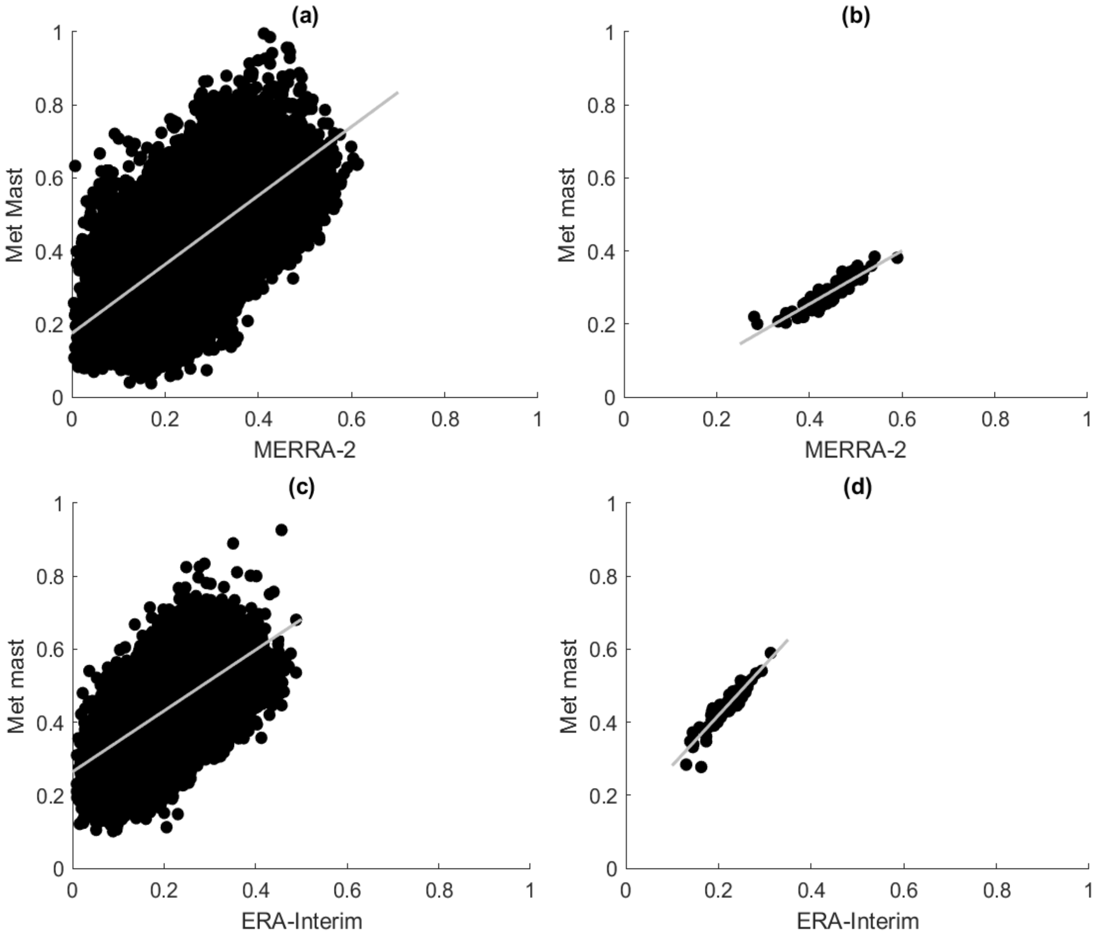
The factor 2 exists to conserve the total power and Fs is the measured data frequency. Further information about the mathematical equations included in this method can be found in Numerical Recipes (1986).

3. Results

3.1 Correlation

Figure 2 shows the scatter plots and linear regressions between each reanalysis datasets versus observations, while Table 1 shows the Pearson Correlation Coefficient. For hourly means, Figures 2a, 2c and 2e show a large dispersion around the straight line, which reflects the R values around 0.5, but for monthly means, the correlation values increase significantly, reaching 0.95 for ERA-5, 0.94 for ERA-Interim and 0.91 for MERRA-2 (Table 1).

It can be seen that for lower measured wind speeds the data are not easily approximated by a line segment (Figures 2b, 2d and 2f). This effect may be partially attributed to the wind speeds being more sensitive to local (microscale) conditions for lower wind speeds, whereas ERA-5, ERA-Interim and MERRA-2 displays better correlation during more intense synoptic activity.



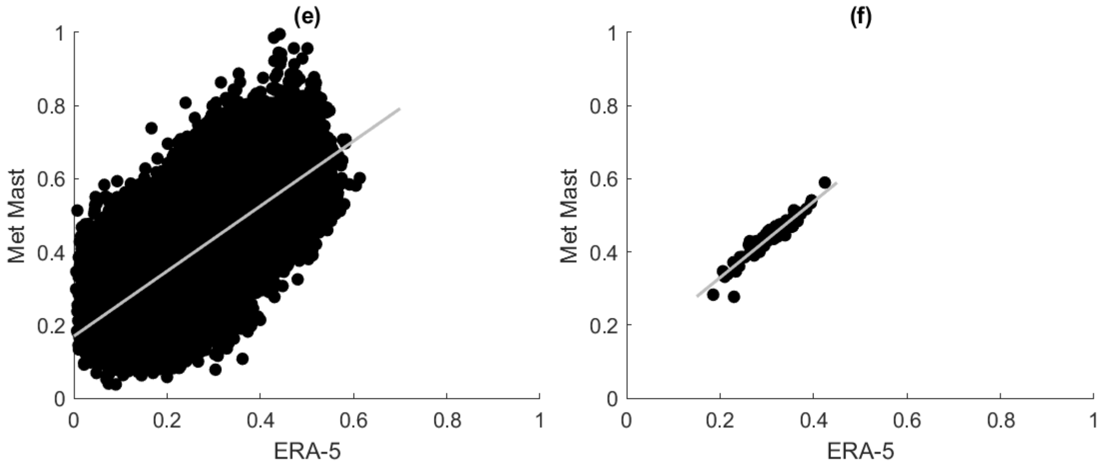


Figure 2: Scatter graphs for MERRA-2, ERA-Interim and ERA-5. Left-side graphs are the scattering from hourly means and right-side graphs are from monthly means.

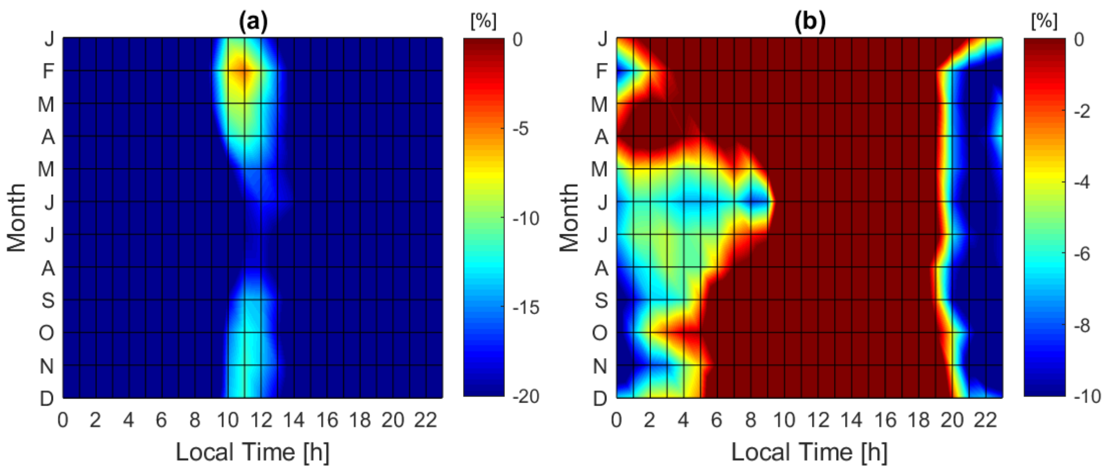
Table 1: Pearson Correlation Coefficient for hourly and monthly means.

Reanalysis	R (Hourly mean)	R (Monthly mean)
MERRA-2	0.59	0.91
ERA-Interim	0.53	0.94
ERA-5	0.64	0.95

3.2 Diurnal Biases

As can be seen in the six heatmaps of Figure 3, there are negative biases in all reanalysis datasets before the MCP application, which is strongest during the night period. After MCP the bias was significantly reduced, reaching zero from 9 am to 19 pm for MERRA-2. For ERA-Interim, before the MCP underestimated the wind speed more than 20%, and afterwards it was reduced to almost zero. Regarding the seasonal biases, the MCP was not so effective during the austral fall and winter, as shown in Figures 3b and 3c.

ERA-5 had the lowest bias of all three reanalysis during the entire year and for all day period, with the highest value around 5% from 4 pm to 8 am period (Figure 3e). It is interesting to note that after the MCP application, the bias was reduced to zero from 2 am to 6 pm for all months, but for the period of 7 pm to 1 am that displayed an increase of the bias, varying from 5% to over 10% (Figure 3f). Although this nighttime pattern before the MCP was observed in both MERRA-2 and ERA-Interim analyses, this problem did not apper when using ERA-5 (Figure 3e).



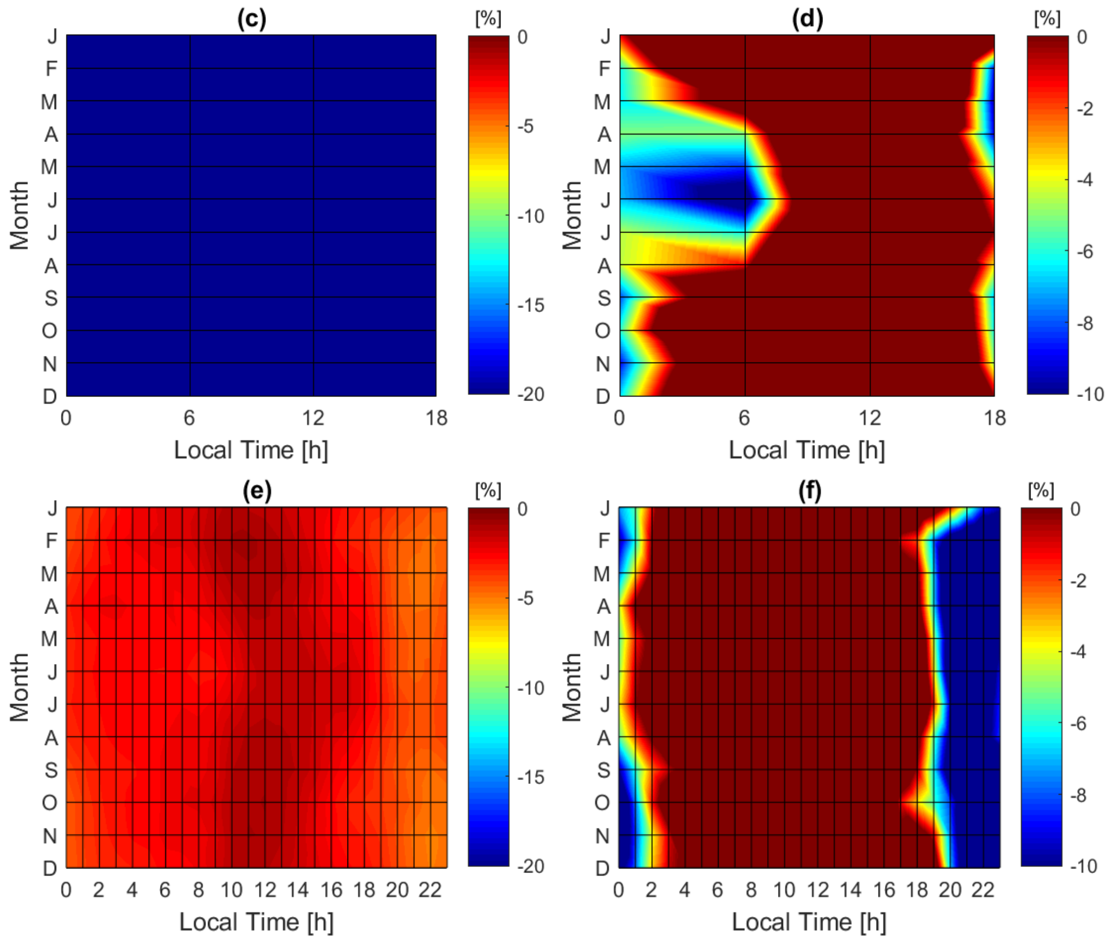
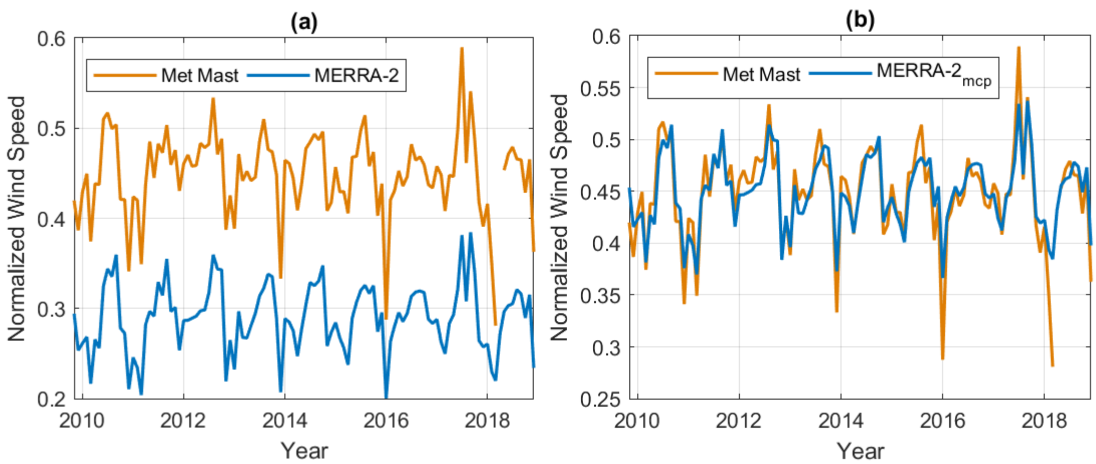


Figure 3: Bias from MERRA-2, ERA-Interim and ERA-5. Left-side graphs are the original dataset from reanalysis and figures from the right-side graphs are from the dataset with MCP.

3.3 Time Series and Signals Comparison

Figures 4a, 4c and 4e show that observed wind speeds are systematically underestimated by both reanalysis data, but can, nonetheless, characterize the seasonal variations reasonably and reproduce abrupt peaks and trends in the time series. Additionally, the MCP was applied to verify the efficiency of this method. Although, for all reanalyses (Figures 4b, 4d and 4f) low wind speeds cannot be predicted as well as the high wind speeds, ERA-5 produced a better fit with observations.



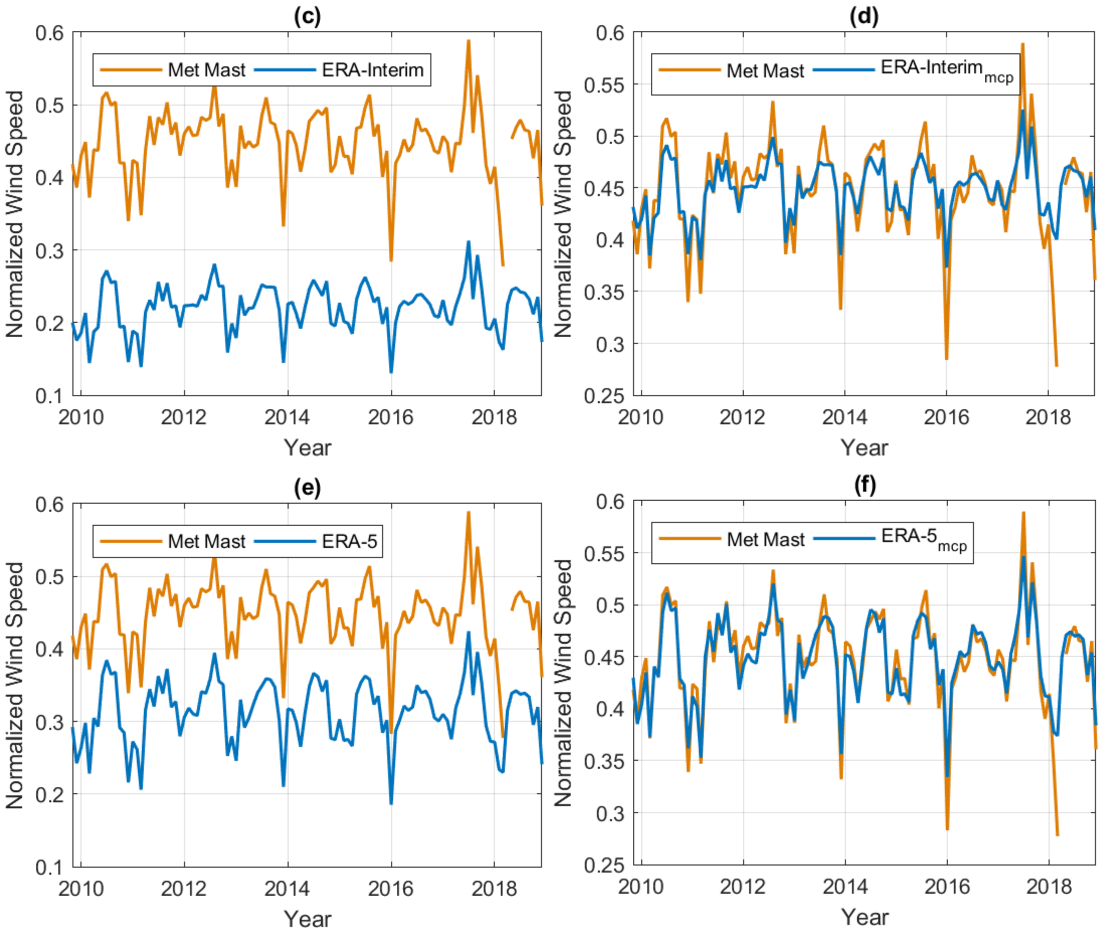


Figure 4: Wind speed time series comparison for MERRA-2, ERA-Interim and ERA-5. Left-side figures represent the monthly means time series from the original dataset and right-side figures are the monthly means time series from the dataset with MCP applied.

A periodogram of Power Spectral Density (PSD) was generated using FFT to verify the capability of identifying cycles present in wind speed measurements and the reanalysis dataset. Figures 5a, 5b and 5c display the Power Spectral Density [W] of the wind speed where the frequency is depicted along the abscissa on a logarithmic scale.For all periodograms, the reanalysis data are processed into an hourly means for MERRA-2 and ERA-5, and 6-hour means for ERA-Interim.

The strongest signal arises from the diurnal cycle (10^0) and is captured by all datasets. The measurements of hourly wind speed show two strong cycles with 6.9 and 27.8 days of duration which could not be detected by any of the three reanalyses. The last visible cycle is the annual cycle with 370.4 days/cycle from the dataset measured at the site and 365.2 days/cycle for MERRA-2 and ERA-5 datasets.

The periodogram generated by ERA-Interim reanalysis is poorer than the ones generated by MERRA-2 and ERA-5, as these two datasets are processed in hourly means. On the other hand, the ERA-Interim periodogram could detect precisely the annual cycle of 365.2 days, while the dataset from the meteorological mast (using 6-hour mean as well) registered an annual cycle of 365.3 days.

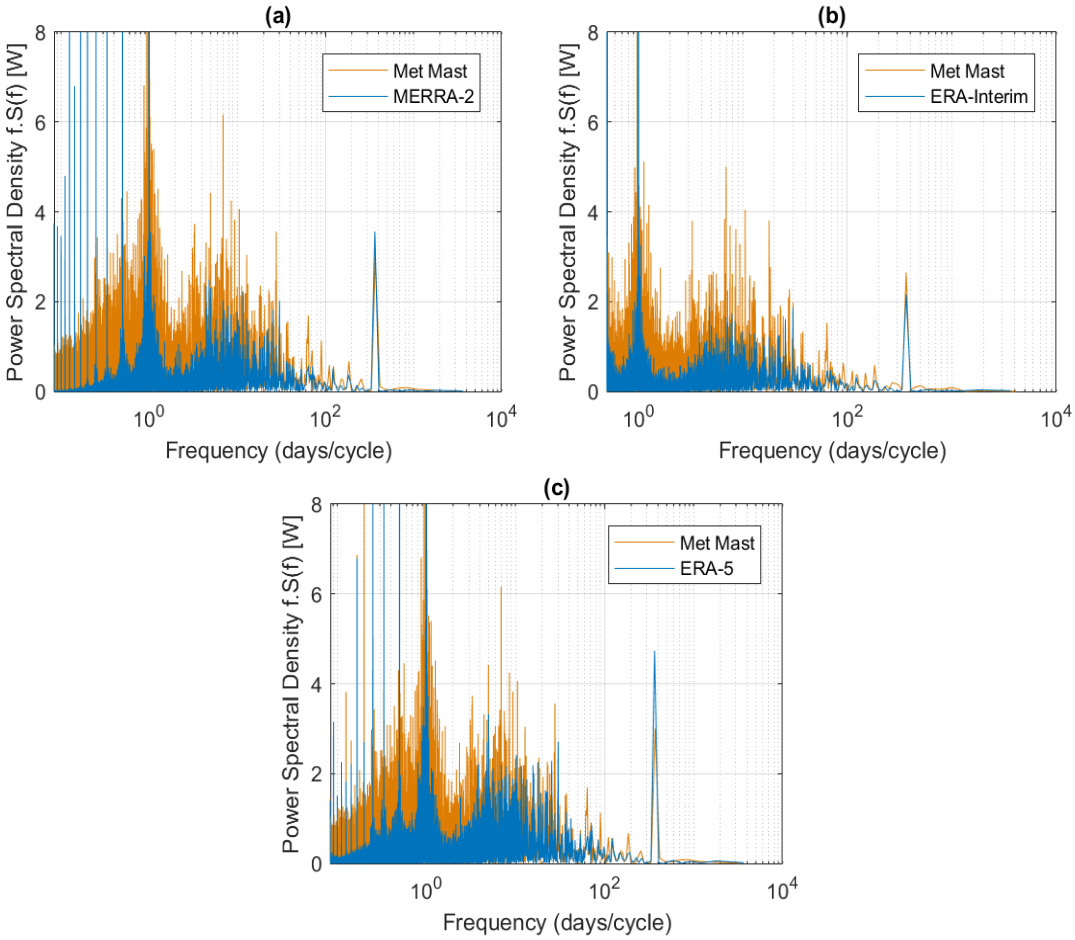


Figure 5: Periodogram using FFT for a) MERRA-2, b) ERA-Interim and c) ERA-5.

4. Concluding Remarks

This paper evaluated the correlation and spectral characteristics of wind speeds obtained from three different reanalysis datasets with in situ observations from a sonic anemometer installed in a meteorological mast. Even though all reanalyses presented a satisfactory performance regarding monthly mean wind speeds, the greatest correlation for a nine-year period (0.95) was obtained with ERA-5. It was visible that by applying the MCP method the biases can be corrected, on average during the day period, showing similar behaviour to that generated using wind measurements at the wind farm. Extreme wind speeds cannot be predicted using reanalysis data and only the annual and diurnal cycles can be detected by the reanalysis dataset using the periodogram analysis. Future studies can consider the use of Artificial Neural Network (ANN) in order to diminish the remaining bias verified in the fall and winter months.

5. Acknowledgments

The authors would like to thank Atlantic Energias Renováveis S.A. and CNPq for providing the meteorological dataset and the financial support for this research through and Research and Development project with LEPTEN (Laboratórios de Engenharia de Processos de Conversão e Tecnologia de Energia), UFSC (Universidade Federal de Santa Catarina) and IFSC (Instituto Federal de Santa Catarina).

6. References

ABEEÓLICA, 2019. Annual Report of Wind Production/Boletim Anual de Geração Eólica. Available in: http://abeeolica.org.br/wp-content/uploads/2019/05/Boletim-Anual_2018.pdf. Last access: 21/06/2019.

BRAZIL. Article 17, Law No. 10.438/02. Available in: http://www.planalto.gov.br/ccivil_03/LEIS/2002/L10438.htm. Last Access: 28/06/2019.

Ahrens, C.D.; Henson, R., 2017. Meteorology today: an introduction to weather, climate, and the environment. 12. ed. Boston, USA: Cengage Learning Inc. pp. 265.
Carta, J. A., Velázquez S., Cabrera P., 2013. A review of measure-correlate-predict (MCP) methods used to estimate long-term wind characteristics at a target site. Renewable & Sustainable Energy Reviews.
v. 27. pp. 362-400.
Copernicus Climate Change Service (C3S), 2017: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS). Available in: https://confluence.ecmwf.int/display/CKB/ERA5+data+documentation. Last access: 08/07/2019.
Dee, D. P. et al., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system, United Kingdom.
Gelaro, R. et al., 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Maryland.
Global Wind Energy Council, 2019. Global Wind Report 2018. Brussels, Belgium.
Hanslian, David, 2015. Matrix-analog measure-correlate-predict approach. The European Wind Energy Association (EWEA). Helsinki, Finland.
Hwang, H. J., 1970. Power Density Spectrum of surface wind speed on Palmyra Island, Vol 98, No. 1, pp. 70-74.
Kim, H-G et al., 2018. Comparative evaluation of the third-generation reanalysis data for wind resource assessment of the Southwestern Offshore in South Korea, Atmosphere Journal, Vol. 73, No. 9, Switzerland. DOI:10.3390/atmos9020073.
Kiss, P. et al., 2009. Comparison of wind power estimates from the ECMWF reanalyses with direct turbine measurements. Vol. 1, No 3. Journal of Renewable and Sustainable Energy. https://doi.org/10.1063/1.3153903
Liléo, S.; Petrik, O., 2011. Investigation on the use of NCEP/NCAR, MERRA and NCEP/CFSR reanalysis data in wind resource analysis. European Wind Energy Conference and Exhibition (EWEA). Brussels, Belgium.
Nunes et. al., 2019. Estudo de desempenho de séries de referência. In: Brazil Wind Power Conference and Exhibition. São Paulo.
Micalichen, M. L. M. de M.; Dias, N. L. da C., 2018. Análise comparativa da velocidade do vento e da temperatura do ar, entre dados gerados por reanálises meteorológicas e dados observacionais na região de Minas Gerais. Ciência e Natura, Vol. 40, pp. 20-25. Santa Maria, Brazil. DOI:10.5902/2179460X30467.
MathWorks, 2019. FFT. Available in: https://www.mathworks.com/help/matlab/ref/fft.html#buuutyt-6. Last Access: 11/07/2019.
MathWorks, 2019. Power Spectral Density Estimates Using FFT. Available in: https://www.mathworks.com/help/signal/ug/power-spectral-density-estimates-using-fft.html. Last Access: 11/07/2019.
Parker, W.S., 2016. Reanalyses and Observations: What's the Difference? Bulletin American Meteorological Society, pp. 1565-1572. DOI: 10.1175/BAMS-D-14-00226.1
Press, W. H, 1986. Numerical Recipes. Cambridge University Press, third ed., Washington, New York.
SCHUBERT, C., 2013. Atlas Eólico/Wind Power Atlas: Bahia. Salvador: SECTI/SEINFRA, CIMATEC/SENAI.
Stuker, E. et al., 2016. Comparação entre os dados de vento das reanálises meteorológicas ERA-Interim e CFSR com os dados das estações automáticas do INMET no Rio Grande do Sul. Ciência e Natura, Vol. 38, pp. 284-290, Santa Maria, Brasil. DOI:10.5902/2179460X20233
Thøgersen, M. L. et al., 2007. Measure-Correlate-Predict Methods: Case Studies and Software Implementation. European Wind Energy Conference & Exhibition (EWEC). Milan, Italy.
Van der Hoven, I., 1957. Power Spectrum of Horizontal Wind Speed in the Frequency Range from 0.0007 to 900 Cycles Per Hour. Journal of Meteorology, Vol. 14, No. 2, pp. 160-164.
Vieira, C. F. A. et al., 2006. Correlação entre dados de vento gerados no projeto reanalysis do ncep/ncar e observados em regiões do estado do Ceará. Revista Tecnologia. Vol 27, No 2, pp. 190-194. Fortaleza, Brasil.
Wallace, J. M.; Hobbs, P. V., 2006. Atmospheric Science: An Introductory Survey. 2nd. Edition. Washington, USA: Elsevier Inc. pp. 17.
World Meteorological Organization, 2007. The Role of Climatological Normals in Changing Climate (WMO-No. 1377), Geneva.
World Meteorological Organization, 2011. Guide to Climatological Practices (WMO-No. 100), Geneva.

Zhang, F.; Bao, X., 2012. Evaluation of NCEP–CFSR, NCEP–NCAR, ERA-Interim, and ERA-40 Reanalysis datasets against independent sounding observations over the Tibetan Plateau. Journal of Climate, Vol. 26, pp 206-214. DOI: 10.1175/JCLI-D-12-00056.1

Appendix

In order to verify the performance of wind speed dataset from ERA-Interim in the 1000 hPa level, a vertical correction was applied. For this, it was used the divergence dataset available in ECMWF site from November 2009 to December 2018 and then the values were divided by the gravity value at sea level (g = 9.81 m/s²) to obtain the real height of the data. To apply the vertical correction, the mean alpha coefficient was found using the power law (Equation 5).

$$U = Ur\left(\frac{Z}{Zr}\right)^{\alpha} \qquad \text{(eq. A1)}$$

Where Ur and Zr are the wind speed and the height at reference level, respectively, while U is the wind speed at hub height, which is 78 m. Figure A1a and A1b present the wind speed scatter graphs which show a negligible difference, 0.02 lower than those generated by the wind speed data directly from 1000 hPa level. In this analysis, the R values are 0.55 for the 6-hour mean and 0.92 for the monthly mean.

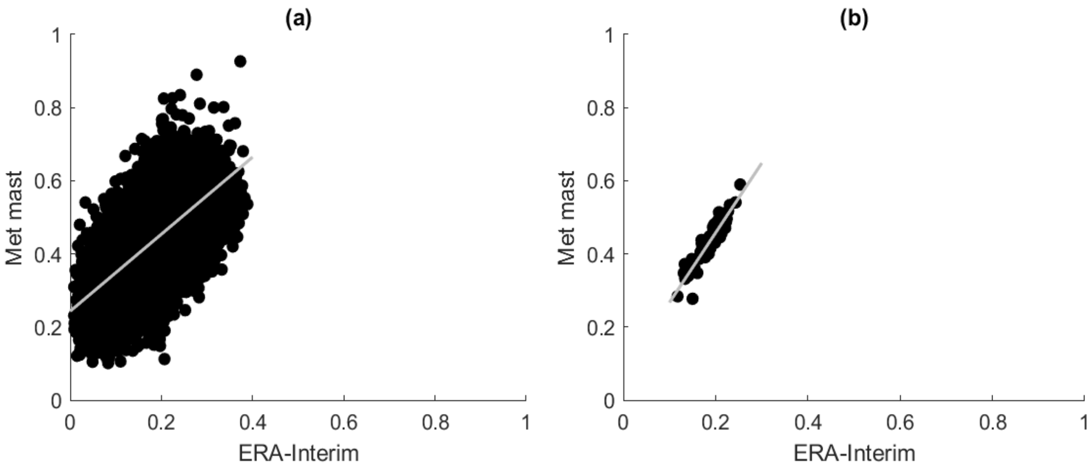


Figure A1: Scatter plot of wind speed with vertical correction from ERA-Interim (a) using hourly means and (b) using monthly means.