Multiscale entropy analysis of wind speed dynamics in Petrolina, Northeast Brazil

Analise da dinâmica do vento em Petrolina, Nordeste do Brasil, utilizando Multiscale Entropy

Análisis de la dinámica del viento en Petrolina, Noreste de Brasil, utilizando Multiscale Entropy

Abstract

Purpose: In this paper, we analyzed the intra-annual variability of complexity of wind dynamics in Petrolina, Brazil and its relation with the wind potential. Methodology: We applied the Multiscale Sample Entropy (MSE) method on wind speed temporal series for each month of 2010. The data are recorded every 10 min at 50m height. Results: The results showed higher entropy values at higher temporal scales indicating that wind speed fluctuations are less regular and less predictable when wind speed is observed at lower temporal frequency. For all months, average wind speed is above a cut in level 3.5 ms⁻¹, the speed at which turbines start operating and producing electricity, indicating that the location of Petrolina is promising for wind energy generation. We also found that the wind speed is positively correlated with entropy values for all months when recorded at 10min frequency and between August and December when recorded at 1 h frequency. Conclusion: In these periods wind speed temporal fluctuations are more irregular, which is considered as unfavorable condition for the operation of wind turbines, leading to lower efficiency in the capture of wind energy for electricity production.

Keywords: Wind speed; Multiscale entropy; Wind energy.

Resumo

Objetivo: Neste artigo analisamos a variabilidade intra-anual da complexidade da dinâmica do vento em Petrolina, Brasil e sua relação com o potencial eólico. Metodologia: Aplicamos o método Multiscale Sample Entropy (MSE) nas séries temporais da velocidade do vento para cada mês de 2010. Os dados são registrados a cada 10 minutos a 50m de altura. Resultados: Os resultados mostraram os valores de entropia mais altos em escales temporais maiores, indicando que as flutuações da velocidade do vento são menos regulares e menos previsíveis quando a velocidade do vento é observada em frequência temporal mais baixa. Em todos os meses a velocidade média do vento está acima de 3,5 ms⁻¹, velocidade na qual as turbinas começam a operar e produzir eletricidade, indicando que a localização de Petrolina é promissora para a geração de energia eólica. Também descobrimos que a velocidade do vento está positivamente correlacionada com os valores de entropia para todos os meses quando dados são registrados na frequência de 10min e entre agosto e dezembro quando são registrados na frequência de 1 h. Conclusão: Nestes períodos as flutuações temporais da velocidade do vento são mais irrregulares, o que é considerado condição desfavorável para o funcionamento dos aerogeradores, levando a uma menor eficiência na captação de energia eólica para a produção de eletricidade.

Palavras-chave: Velocidade do vento; Multiscale entropy; Energia eólica.

Resumen

Objetivo: En este trabajo analizamos la variabilidad intra-anual de la complejidad de la dinámica del viento en Petrolina, Brasil y su relación con el potencial eólico. Metodología: Se aplicó el método de Multiscale Sample Entropy Entropy (MSE) en series temporales de velocidad del viento para cada mes de 2010. Los datos se registran cada 10 min a 50 m de altura. Resultados: Los resultados mostraron valores de entropia más altos en escalas temporales más altos, lo que indica que las fluctuaciones de la velocidad del viento son menos regulares y menos predecibles cuando la velocidad del viento se observa a una frecuencia temporal más baja. Para todos los meses, la velocidad media del viento está por encima de 3,5 ms⁻¹, la velocidad a la que las turbinas comienzan a funcionar y a producir electricidad, lo que indica que la ubicación de Petrolina es prometedora para la generación de energía eólica. También encontramos que la velocidad del viento se correlaciona positivamente con los valores de entropia para todos los meses cuando se registra a una frecuencia de 10 min y entre agosto y diciembre cuando se registra una frecuencia de 1 h. Conclusión: En estos
períodos las fluctuaciones temporales de la velocidad del viento son más irregulares, lo que se considera una condición desfavorable para el funcionamiento de los aerogeneradores, lo que conlleva una menor eficiencia en la captación de energía eólica para la producción de electricidad.

**Palabras clave:** Velocidad del viento; *Multiscale entropy*; Energía eólica.

1. Introduction

Among renewable energy sources wind power is one with most rapidly growing rate over the last decades because of its high efficiency and low pollution. As a wind producer Brazil ranks in the sixth place in the world, behind China, USA, Germany, India and France. At the end of 2019 Brazilian production was 15.45 GW representing 2.5% of the total worldwide onshore capacity (GWEC, 2019). It is projected to grow to 8.0%, in 2020, where half of the total capacity will be installed in the Northeastern region (Witzler, Ramos, Camargo, & Guarnier, 2016). This remarkable growth is the result of the government Program for Incentive of Alternative Electric Energy Sources (Programa de Incentivo às Fontes Alternativas de Energia Elétrica - Proinfa), which was created in 2002 to stimulate the electricity generation from wind power, biomass, and small hydroelectric plants (Dutra & Szklo, 2008). During the dry season in Northeast of Brazil, the temporal variation of the wind potential shows complementarity with the flows of the São Francisco river (Faria, Justino, & Monteiro, 2011) and wind electricity generation is extremely important for electricity supply in this region which often experiences prolonged drought periods, when hydroelectric plants decrease the electricity production due to very low flows and reservoirs levels (Cavalcante, Vieira, Campos, Brandini, & Medeiros, 2020). However, due to intermittency and high spatio-temporal variability of wind speed, large scale integration of wind power into electricity grid is still challenging task (Behera, Sahoo, & Pati, 2015).

The evaluation of the wind potential at certain location requires a detailed statistical analysis of the wind speed and its frequency distribution at different heights and different periods during the year (de Araujo Lima & Bezerra Filho, 2010; Tar, 2008). However, the knowledge of the temporal organization (complexity) of wind speed can provide valuable information about underlying stochastic processes that cannot be assessed through traditional statistical analysis. This knowledge is important for the planning of wind energy production and for evaluation of predictive models for wind speed and wind power.

During the last decades various concepts and methods were developed to analyze complexity of temporal series, among which fractal and multifractal methods, methods based on information theory and complex networks were used to analyze wind speed data from different parts of the world (Koçak, 2009; Laib, Golay, Telesca, & Kanevski, 2018; Laib, Guignard, Kanevski, & Telesca, 2019; Q. Li & Zuntao, 2014). Among these methods, approaches that rely on entropy are particularly interesting due to their simplicity, lack of intensive computations, and robustness for short non-stationary and noisy data.

In this work we evaluate the complexity of wind speed dynamics using Multiscale Entropy method which provides information about time series regularity for multiple temporal scales (Costa, Goldberger, & Peng, 2002). We analyze high frequency data recorded in the city of Petrolina, which is considered one of most promising locations for wind energy production in Pernambuco, Brazil (Silva et al., 2002).

2. Methodology

The present study is a quantitative research (Pereira, Shitsuka, Parreira, & Shitsuka, 2018) that evaluates the complexity of wind speed dynamics using Multiscale Entropy method.

2.1 Data

The data used in this work are wind speed temporal series recorded at SONDA station (SONDA–Sistema Nacional de Organização de Dados Ambientais) in Petrolina, Brazil (09 ° 04’ 08” S latitude, 40 ° 19’ 11” O longitude, and altitude of 387m). The data was obtained from INPE- Instituto Nacional de Estudos Espaciais and are available at the electronic address
The series are composed of 10-minute observations recorded at height of 50m in the year 2010 (~ 52000 data points) as shown in Figure 1.

**Figure 1.** Original wind speed series recorded at the Petrolina station.

Note: Figure 1 shows wind speed (m/s) in each month of 2010 in Petrolina. Source: Author.

### 2.2 Sample Entropy

Sample entropy (SampEn) was introduced by Richman and Moorman, to evaluate the complexity of short non-stationary signals by examining time series for similar epochs, where more frequent and similar epochs (i.e., increased regularity in the time series) lead to lower values of SampEn, while less frequent and similar epochs (decreased regularity) lead to higher values of SampEn (Richman & Moorman, 2000). SampEn \((m, r, N)\) is defined as the negative natural logarithm of the conditional probability that two sequences (of consecutive data within the time series) which are similar for \(m\) points remain similar at the next point, where self-matches are not included. In calculating the probability,

Sample entropy algorithm is described as follows (Silva et al., 2002):
For a time, the series of length $N$, $u(j), j = 1, ..., N$, we form $N - m + 1$ vectors of length $m$, $x_m(i), i = 1, ..., N - m + 1$ where $x_m(i) = \{u(i + k) : k = 0, ..., m - 1\}$

The distance between vectors $x_m(i)$ and $x_m(j)$ is defined as

$$d[x_m(i), x_m(j)] = \max\{u(i + k) - u(j + k) : k = 0, ..., m - 1\}$$

Next we count the number $B_i$ of vectors $x_m(j)$ such that $d[x_m(i), x_m(j)] \leq r$ where $r$ is a tolerance level ($r \equiv r\sigma$, $\sigma$ - standard deviation of $u(i), i = 1, ..., N, j = 1, ..., N - m$ and $j \neq i$) to exclude self-matches.

We then define

$$B^m_i(r) = \frac{B_i}{N - m + 1} \quad \text{and} \quad B^m(r) = \frac{\sum_{i=1}^{N-m} B^m_i(r)}{N - m}$$

where $B^m(r)$ is the probability that two vectors will match for $m$ points.

We repeat steps i-iv for vectors of length $m + 1$ and count the number $A_i$ of vectors $x_{m+1}(j)$ which are within tolerance level $r$ of $x_{m+1}(i)$ and again we exclude self-matches. We define

$$A^m_i(r) = \frac{A_i}{N - m + 1} \quad \text{and} \quad A^m(r) = \frac{\sum_{i=1}^{N-m} A^m_i(r)}{N - m}$$

where $A^m(r)$ is the probability that two vectors will match for $m + 1$ points.

Sample entropy (SampEn) is defined as

$$S_E(m, r, N) = \lim_{N \to \infty} \left[- \ln \frac{A^m(r)}{B^m(r)}\right]$$

which is estimated by the statistics

$$S_E(m, r, N) = - \ln \frac{A^m(r)}{B^m(r)}$$

Sample entropy method was used to analyze data in physiology, geophysics, engineering and hydrology (Balasis et al., 2009; Chou, 2014; Kumar, Pachori, & Acharya, 2017; Ni, Feng, Wang, Yang, & Wang, 2017; Santana, Stosic, Ferreira, & Silva, 2020a, 2020b).
2.3 Multiscale Sample Entropy

Multiscale sample entropy method (MSE) was introduced by Costa et al., as a generalization of Sample entropy method (Costa et al., 2002; Richman & Moorman, 2000). MSE takes into account the multiple time scales by calculating sample entropy for consecutive coarse-grained time series \[ x'_{j}(i) \], \( i = 1, \ldots, N / \tau \) determined by the scale factor \( \tau \):

\[
x'_{j}(i) = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j \tau} x(i) \quad \text{where} \quad x(i), i = 1, \ldots, N \quad \text{is original time series.}
\]

It was shown that uncorrelated signals (white noise) have, for larger scales \( \tau \), lower MSE values than correlated noise, indicating that MSE is more efficient for quantifying complexity in short and noisy time series, than traditional entropy methods that evaluate pattern repetition on single temporal scale. MSE method was used in analyzing physiological signals, hydrological processes, geophysical signals, and financial time series (Courtiol et al., 2016; Gamboa, Marques, & Stosic, 2019; Guzman-Vargas, Ramírez-Rojas, & Angulo-Brown, 2008; Zhou, Zhang, Li, & Chen, 2012).

3. Results and Discussion

The results of the descriptive statistics are presented in Table 1. For most wind turbines, the range of cut-in wind speed (the speed at which the turbine starts producing the energy) is 3.5–4.5 m/s (Ayodele & Ogunjuyigbe, 2016). It is seen from Table 1 that the average wind speed is above cut in level during the whole year confirming that the location of Petrolina is promising for wind energy generation. The average wind speed is the highest (with lowest variation) in September indicating the most favorable period of year for energy generation.

Table 1. Descriptive statistics for wind speed (m/s) temporal series recorded in Petrolina during the year of 2010.

<table>
<thead>
<tr>
<th>Month</th>
<th>Average</th>
<th>Min</th>
<th>Max</th>
<th>SD</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>4.26</td>
<td>0.00</td>
<td>11.40</td>
<td>1.56</td>
<td>0.37</td>
</tr>
<tr>
<td>February</td>
<td>4.88</td>
<td>0.00</td>
<td>13.13</td>
<td>1.65</td>
<td>0.34</td>
</tr>
<tr>
<td>March</td>
<td>3.82</td>
<td>0.00</td>
<td>10.77</td>
<td>1.67</td>
<td>0.44</td>
</tr>
<tr>
<td>April</td>
<td>4.38</td>
<td>0.09</td>
<td>11.73</td>
<td>1.37</td>
<td>0.31</td>
</tr>
<tr>
<td>May</td>
<td>4.55</td>
<td>0.01</td>
<td>10.76</td>
<td>1.36</td>
<td>0.30</td>
</tr>
<tr>
<td>June</td>
<td>5.35</td>
<td>0.00</td>
<td>10.82</td>
<td>1.73</td>
<td>0.32</td>
</tr>
<tr>
<td>July</td>
<td>5.64</td>
<td>0.67</td>
<td>10.65</td>
<td>1.50</td>
<td>0.26</td>
</tr>
<tr>
<td>August</td>
<td>5.96</td>
<td>0.01</td>
<td>11.38</td>
<td>1.51</td>
<td>0.25</td>
</tr>
<tr>
<td>September</td>
<td>6.27</td>
<td>0.78</td>
<td>12.04</td>
<td>1.52</td>
<td>0.24</td>
</tr>
<tr>
<td>October</td>
<td>4.96</td>
<td>0.00</td>
<td>11.91</td>
<td>1.74</td>
<td>0.35</td>
</tr>
<tr>
<td>November</td>
<td>4.94</td>
<td>0.00</td>
<td>11.15</td>
<td>1.44</td>
<td>0.29</td>
</tr>
<tr>
<td>December</td>
<td>3.62</td>
<td>0.00</td>
<td>14.30</td>
<td>1.64</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Legend: Min – Minimum. Max – Maximum. SD – Standard Deviation. CV – Coefficient of Variation. Note: Table 1 shows wind speed (m/s) descriptive statistics for each month of 2010 in Petrolina. Source: Authors.

Table 2 presents the MSE values for \( \tau =1,2,3,4,5,6 \) where \( \tau =6 \) corresponds to calculating averages of 10 min wind speed values for consecutive 1-hour periods. It is seen from Table 2 that MSE increases with scale \( \tau \) indicating that wind dynamics is more irregular (less predictable) when observed on higher temporal scales. For most scales MSE values are lowest in October (\( \tau =2,3,4,6 \)) and highest in May (\( \tau =3,4,5,6 \)) indicating that the most regular (most predictable) wind regime at Petrolina location is in October while most irregular (least predictable) in May. For \( \tau =1 \) entropy value is the highest in September, month with the highest value of average speed. The power of the wind per unit area is given by \( P(v) = \frac{1}{2} \rho v^3 \), where \( P(W/m^2) \) is the power per unit area, \( \rho (kg/m^3) \) is the air density, and \( v (ms^{-1}) \) is the wind speed (Safari & Gasore, 2010). Considering average wind speed September is the moth with the most favorable conditions for energy generation. However, the dynamical
regularity of wind speed temporal series is the lowest in September (highest entropy values) which can be considered as the is unfavorable condition for the operation of wind turbines. In order to contribute to a better understanding of the intra-annual temporal variability of conditions for wind power generation in Petrolina, we analyzed the correlation between average wind speed and entropy values for each month.

**Table 2.** MSE values \((m=2, r=0.20, \tau=1,2,3,4,5,6)\) for wind speed temporal series recorded in Petrolina during the year 2010.

<table>
<thead>
<tr>
<th>SampEn</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>1.159</td>
<td>1.259</td>
<td>1.341</td>
<td>1.399</td>
<td>1.443</td>
<td>1.487</td>
</tr>
<tr>
<td>Fev</td>
<td>1.202</td>
<td>1.296</td>
<td>1.294</td>
<td>1.379</td>
<td>1.375</td>
<td>1.393</td>
</tr>
<tr>
<td>Mar</td>
<td>1.171</td>
<td>1.317</td>
<td>1.386</td>
<td>1.426</td>
<td>1.536</td>
<td>1.550</td>
</tr>
<tr>
<td>Abr</td>
<td>1.180</td>
<td>1.320</td>
<td>1.388</td>
<td>1.458</td>
<td>1.490</td>
<td>1.492</td>
</tr>
<tr>
<td>Mai</td>
<td>1.224</td>
<td>1.341</td>
<td>1.422</td>
<td>1.493</td>
<td>1.559</td>
<td>1.599</td>
</tr>
<tr>
<td>Jun</td>
<td>1.116</td>
<td>1.235</td>
<td>1.291</td>
<td>1.345</td>
<td>1.343</td>
<td>1.374</td>
</tr>
<tr>
<td>Jul</td>
<td>1.205</td>
<td>1.306</td>
<td>1.412</td>
<td>1.459</td>
<td>1.488</td>
<td>1.516</td>
</tr>
<tr>
<td>Ago</td>
<td>1.278</td>
<td>1.340</td>
<td>1.358</td>
<td>1.389</td>
<td>1.443</td>
<td>1.473</td>
</tr>
<tr>
<td>Set</td>
<td>1.339</td>
<td>1.345</td>
<td>1.322</td>
<td>1.352</td>
<td>1.428</td>
<td>1.482</td>
</tr>
<tr>
<td>Out</td>
<td>1.155</td>
<td>1.231</td>
<td>1.250</td>
<td>1.334</td>
<td>1.364</td>
<td>1.349</td>
</tr>
<tr>
<td>Nov</td>
<td>1.285</td>
<td>1.350</td>
<td>1.378</td>
<td>1.449</td>
<td>1.507</td>
<td>1.509</td>
</tr>
<tr>
<td>Dez</td>
<td>1.131</td>
<td>1.221</td>
<td>1.287</td>
<td>1.375</td>
<td>1.427</td>
<td>1.447</td>
</tr>
</tbody>
</table>

Note: Table 2 shows wind speed (m/s) MSE values for scales 1 to 6 for each month of 2010 in Petrolina. Source: Authors.

The results for 10 min frequency \((\tau=1)\) and 1-hour frequency \((\tau=6)\) are shown on figure 2, where it can be observed that for \(\tau=1\) wind speed and entropy are positively correlated (Pearson coefficient P-0.64), while for \(\tau=6\) positive correlation is observed between August and December. Ahmed and Mandic analyzed vertical component and horizontal component (east-west) of wind velocity temporal series recorded with 3D anemometer and also found the increase of entropy with wind speed (Ahmed & Mandic, 2011). Li et al. 2011 analyzed the wind dynamics in the horizontal plane near the surface and also found that the entropy and wind speed are positively correlated (H. Li, Meng, Wang, & Zeng, 2011).
4. Conclusions

In this work, we studied intra annual complexity of the time series of wind speed in Petrolina using Multiscale Sample Entropy method (MSE). We found that the entropy values increase as the temporal scale increases, indicating less regular and less predictable wind speed series when recorded with lower temporal frequency.

Although for all month’s average wind speed is above cut in level 3.5 m s\(^{-1}\) at which most turbines start to produce electricity, dynamical regularity of wind speed temporal series revealed by MSE analysis provides an additional valuable information: for 10 min frequency wind speed and entropy are positively correlated (Pearson coefficient P=0.64), while for 1-hour positive correlation is observed between August and December. Higher entropy values indicate fewer regular dynamics, which is unfavorable condition for the operation of wind turbines and consequently there is less efficiency in the capture of wind energy for wind electricity production.

Future studies should be focused on large scale spatial analysis by including data from various locations that are considered as promising for wind power generation in Brazil.

Acknowledgments

The authors acknowledge support of Brazilian agency CNPq, through the following research grant: 304497/2019-3.
References


