

**Comparação dos dados da velocidade do vento no Nordeste do Brasil da ERA-40 e Instituto Nacional de Meteorologia (INMET) utilizando medidas de entropia**

**Comparison of wind speed data in Northeast Brazil of ERA-40 and National Institute of Meteorology (INMET) using entropy measurements**

**Comparación de datos de velocidad del viento sobre en el noreste de Brasil de ERA-40 e Instituto Nacional de Meteorología (INMET) utilizando mediciones de entropía**

Recebido: 03/06/2020 | Revisado: 29/06/2020 | Aceito: 01/07/2020 | Publicado: 12/07/2020

**Lêda Valéria Ramos Santana**

ORCID: <https://orcid.org/0000-0002-5704-8756>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [ledaestatistica@gmail.com](mailto:ledaestatistica@gmail.com)

**Tatijana Stosic**

ORCID: <https://orcid.org/0000-0002-5691-945X>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [tastosic@gmail.com](mailto:tastosic@gmail.com)

**Tiago Alessandro Espínola Ferreira**

ORCID: <https://orcid.org/0000-0002-2131-9825>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [taef.first@gmail.com](mailto:taef.first@gmail.com)

**Antonio Samuel Alves da Silva**

ORCID: <https://orcid.org/0000-0002-8759-0036>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: [samuelmthematical@gmail.com](mailto:samuelmthematical@gmail.com)

**Resumo**

Séries históricas de velocidade do vento oriundas das bases de dados do Instituto Nacional de Meteorologia (INMET) e ECMWF Re-Analyses (ERA-40) foram analisadas, com a finalidade de quantificar o grau de regularidade da série temporal e o grau de similaridade entre as bases de dados de estações convencionais (INMET) e de reanálise (ERA-40), utilizando os métodos Sample Entropy e cross-Sample Entropy da teoria da informação. Devido à falta de informações na base do INMET, as análises foram realizadas no período de oito anos de dados simultâneos

(1993 a 2000) para a base de dados do INMET e da ERA-40, durante às 00h, 12h e na série Completa/Total (série original). Os resultados mostram que os maiores registros de velocidade do vento para diferentes séries encontram-se no Norte das quatro sub-regiões do NE. A Sample Entropy, apresentou maior regularidade da velocidade do vento no Meio Norte, área onde a velocidade do vento é menor, apresentando melhor previsibilidade nesta área. A cross-Sample Entropy mostrou uma sincronização moderada das séries do INMET e ERA-40, indicando uma superestimação ou subestimação dos dados da ERA-40 em relação aos dados do INMET.

**Palavras-chave:** Nordeste; Velocidade do vento; Sample entropy; Cross-sample entropy.

### **Abstract**

Historical wind speed series from the databases of the National Institute of Meteorology (INMET) and ECMWF Re-Analyzes (ERA-40) were analyzed in order to quantify the degree of regularity of the time series and the degree of similarity between the databases of conventional stations (INMET) and reanalysis (ERA-40), using the Sample Entropy and cross-Sample Entropy methods of information theory. Due to the lack of information in the INMET database, the analyzes were carried out over a period of eight years of simultaneous data (1993 to 2000) for the INMET and ERA-40 database, during 00h, 12h and in the Complete / Total (original series). The results show that the largest wind speed records for different series are found in the North of the four sub-regions of the NE. Sample Entropy showed greater regularity of wind speed in the Middle North, an area where wind speed is lower, with better predictability in this area. The cross-Sample Entropy showed a moderate synchronization of the INMET and ERA-40 series, indicating an overestimation or underestimation of the ERA-40 data in relation to the INMET data.

**Keywords:** Northeast; Wind speed; Sample entropy; Cross-sample entropy.

### **Resumen**

Se analizaron series históricas de la velocidad del viento de las bases de datos del Instituto Nacional de Meteorología (INMET) y Re-Análisis de ECMWF (ERA-40) para cuantificar el grado de regularidad de las series de tiempo y el grado de similitud entre las bases de datos de estaciones convencionales (INMET) y reanálisis (ERA-40), utilizando los métodos de teoría de la información de Entropía de muestra y Entropía de muestra cruzada. Debido a la falta de información en la base de datos INMET, los análisis se llevaron a cabo durante un período de ocho años de datos simultáneos (1993 a 2000) para la base de datos INMET y ERA-40, durante 00h, 12h y en el Completo / Total (serie original). Los resultados muestran que los mayores

registros de velocidad del viento para diferentes series se encuentran en el norte de las cuatro subregiones del NE. La muestra de entropía mostró una mayor regularidad de la velocidad del viento en el Medio Norte, un área donde la velocidad del viento es más baja, mostrando una mejor previsibilidad en esta área. La entropía de muestras cruzadas mostró una sincronización moderada de las series INMET y ERA-40, lo que indica una sobreestimación o subestimación de los datos de ERA-40 en relación con los datos de INMET.

**Palabras clave:** Noreste; Velocidad del viento; Entropía de muestra; Entropía de muestra cruzada.

## 1. Introdução

A matriz elétrica brasileira é composta por diversas fontes renováveis (biomassa, hídrica, solar e eólica) e não renováveis (fóssil, nuclear). Embora a maior produção seja das hidroelétricas, o fornecimento da energia eólica tem sido ampliado e representa atualmente 9,7% da matriz energética brasileira, com previsão de capacidade instalada de 23,5 MW (Megawatts) ao final de 2025 (ABEEólica, 2020).

Atualmente o Brasil possui 624 usinas eólicas instaladas com capacidade de 15,6 GW. O Nordeste (NE) abriga 80% dos parques eólicos do país, com o Rio Grande do Norte responsável por 4319,5 MW de capacidade instalada, acompanhado da Bahia com 4074,4 MW e do Ceará com 2045,5 MW (ABEEólica, 2020; GWEC, 2019). Além de, deter a vantagem sobre o aspecto da complementaridade sazonal do regime de ventos com os fluxos hidrológicos nos reservatórios hidrelétricos, uma vez que em período de estiagem quando os níveis dos reservatórios baixam, a velocidade do vento é maior (Amarante, 2001).

Diante da capacidade de geração de energia eólica do país, pesquisadores têm explorado a região NE do Brasil com a finalidade de obter informações sobre o comportamento do vento local e assim poder contribuir em estudos que viabiliza a configuração e instalação de parques eólicos na região (Lira et al., 2017; Camelo et al., 2016; Carneiro & Carvalho, 2015; Santana et al., 2015).

As pesquisas são fundamentadas em diferentes fontes de dados. Algumas fontes de dados como as convencionais, apresentam desvantagens por apresentarem ausência de informação em seus registros. Para atender limitações como esta, centros de estudos climáticos desenvolveram modelos de circulação global que fornecem informações contínuas para diversas variáveis climáticas a nível mundial, denominando-o de reanálise, representados por registros regulares dispostos em uma grade de forma uniforme (ECMWF, 2017).

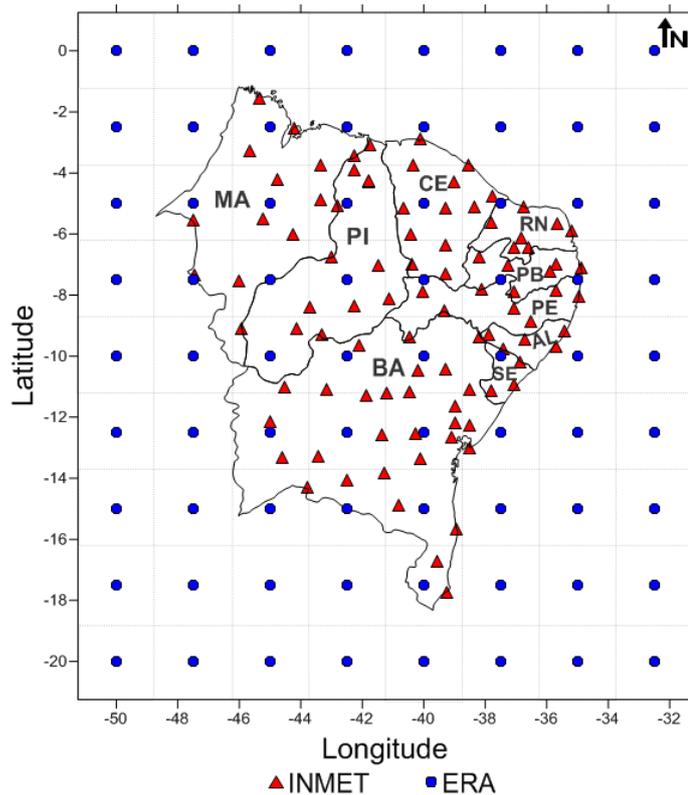
Como as bases de dados têm estações/grade sobre a mesma região, um ponto a ser explorado é se ambas conseguem reproduzir a mesma informação. As estatísticas clássicas utilizadas como esse propósito são, o coeficiente de correlação, erro quadrático médio, índice de concordância e coeficiente de desempenho (Witzler et al., 2016). Contudo, o uso da medida de entropia cross-Sample Entropy, introduzida por Richman & Moorman (2000), utilizada em diversas áreas como processos fisiológicos (Valenza et al., 2014), processos hidrológicos (Chou, 2014) e dados financeiros (Reddy & Sebastin, 2006), surge como uma nova abordagem, em particular, para o estudo de séries da velocidade do vento que pertencem a base de dados provenientes de diferentes instituições, permitindo verificar a similaridade entre as séries. Assim como, o uso da Sample Entropy para verificar a regularidade das séries sobre o NE.

O objetivo do trabalho é utilizar métodos de entropia para medir o grau de similaridade entre os dados disponibilizados pela ERA-40 e INMET, assim como, identificar quais regiões do Nordeste do Brasil apresentam maior/menor previsibilidade da velocidade do vento.

## **2. Metodologia**

O presente estudo é uma pesquisa de natureza quantitativa (Pereira et al., 2018). Os dados analisados estão dispostos sobre o Nordeste brasileiro, localizado entre os meridianos  $48^{\circ} 05' W$  a  $35^{\circ} 02' W$  e paralelos  $1^{\circ} S$  a  $18^{\circ} 05' S$ . Na Figura 1, observa-se os limites geográficos, bem como a disposição das coordenadas dos dados do INMET composto por 94 estações, e da reanálise ERA-40 que possuem 72 pontos de grade.

**Figura 1.** Estações do INMET (triângulos) e grades da ERA-40 (círculos) sobre a região NE.



Fonte: Elaborado pelos autores (2020).

Na Figura 1, o ponto azul no centro de cada célula/grade, representam as informações da reanálise ERA-40 geradas por modelos de assimilação de dados. As estações do INMET proveniente de estações convencionais são caracterizadas por triângulos vermelhos dispostos de acordo com a respectivas coordenadas geográficas. Ambos os bancos de dados foram analisados em três séries, às 00 e 12h onde são consideradas os registros para os respectivos horários em cada série da velocidade do vento e, a denominada série Completa/Total onde não há discriminação do período.

Os dados da ERA-40 são registros realizados em  $m/s^{-1}$  as, 00, 06, 12, 18 UTC, a 10 m do solo dispostos em uma grade com resolução  $2,5^{\circ} \times 2,5^{\circ}$  (lat/long) que equivale a cerca de  $275 \text{ km}^2$ , em formato \*.GRIB, no período de 45 anos (1957-2002), disponível no endereço eletrônico <<http://apps.ecmwf.int/datasets/data/era40-daily/levtype=sfc/>>. Dentre as diversas variáveis disponibilizadas, foram selecionadas as componentes zonal (u) e meridional (v), representando a direção do vento no sentido Oeste-Leste e Norte-Sul, respectivamente. Com base nas componentes u e v, pode-se obter velocidade escalar do vento por meio da seguinte expressão:

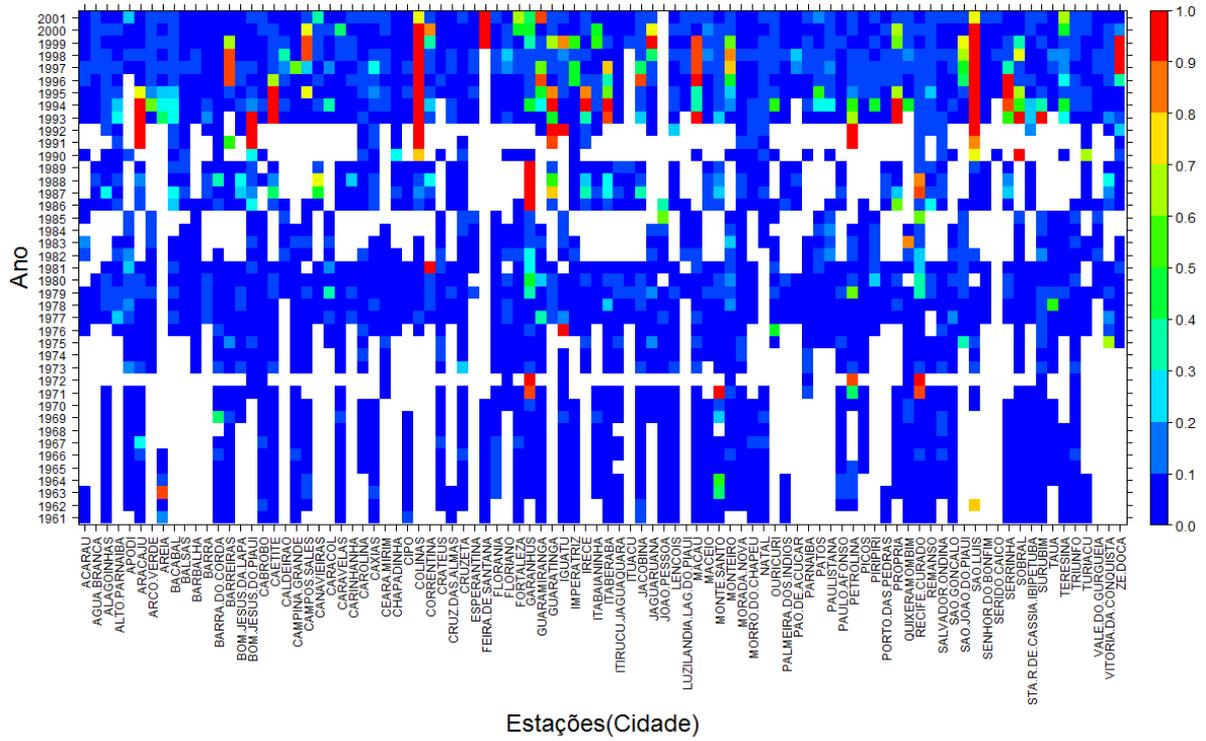
$$s = \sqrt{u^2 + v^2}. \quad (1)$$

As coordenadas da ERA-40 foram selecionadas considerando se existia alguma coordenada da base de dados do INMET dentro da grade da ERA-40. Deste modo, primeiro foram escolhidas 29 coordenadas entre as 72 coordenadas disponíveis.

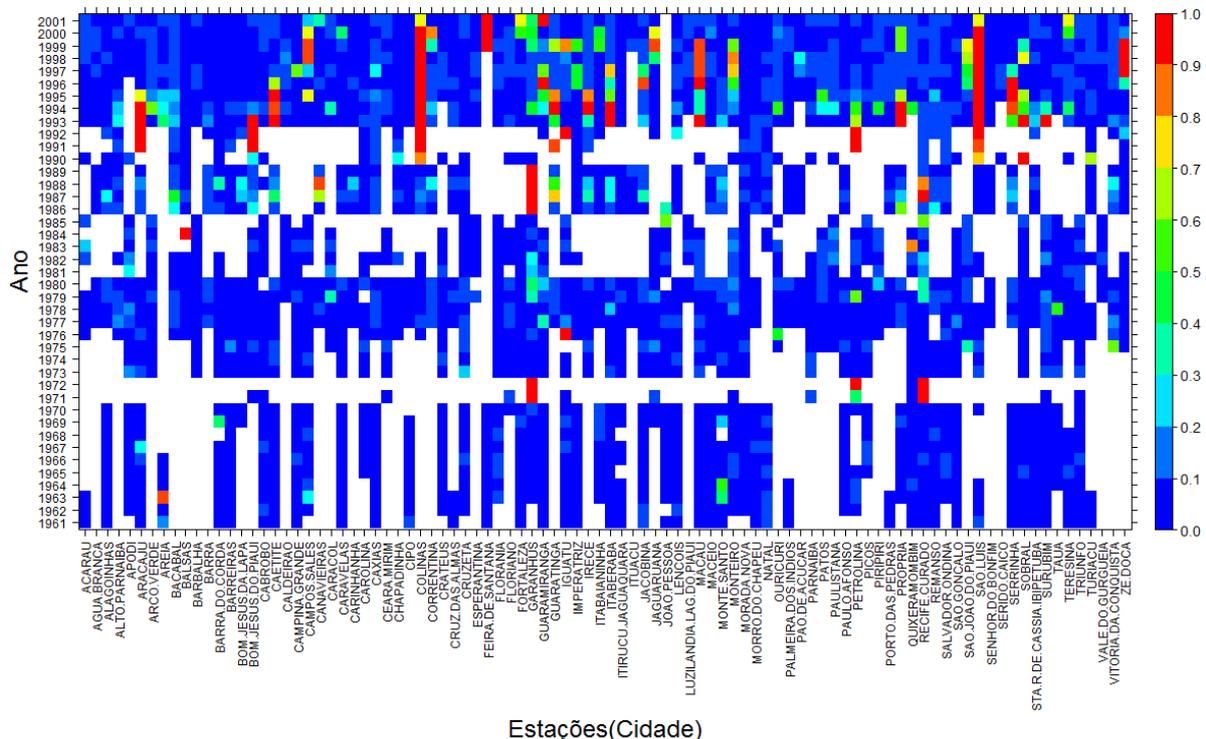
Os dados do INMET são registros de estações convencionais, coletadas manualmente, pertencentes ao Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP) que apoia atividades de ensino e pesquisa com registros horários às 00, 12 e 18 UTC, a 10 m do solo, para a variável velocidade do vento na região NE disponíveis em [www.inmet.gov.br/portal/index.php?r=estacoes/estacoesconvencionais](http://www.inmet.gov.br/portal/index.php?r=estacoes/estacoesconvencionais).

A Figura 2, representa o percentual de dados ausentes no banco de dados do INMET por estação e ano para 00h (Figura 2 (a)) e 12h (Figura 2 (b)).

**Figura 2.** Estações do INMET com percentual de dados ausentes às 00h (a) e 12h (b) no período de 1961 a 2001.



Estações(Cidade)  
(a)



Estações(Cidade)  
(b)

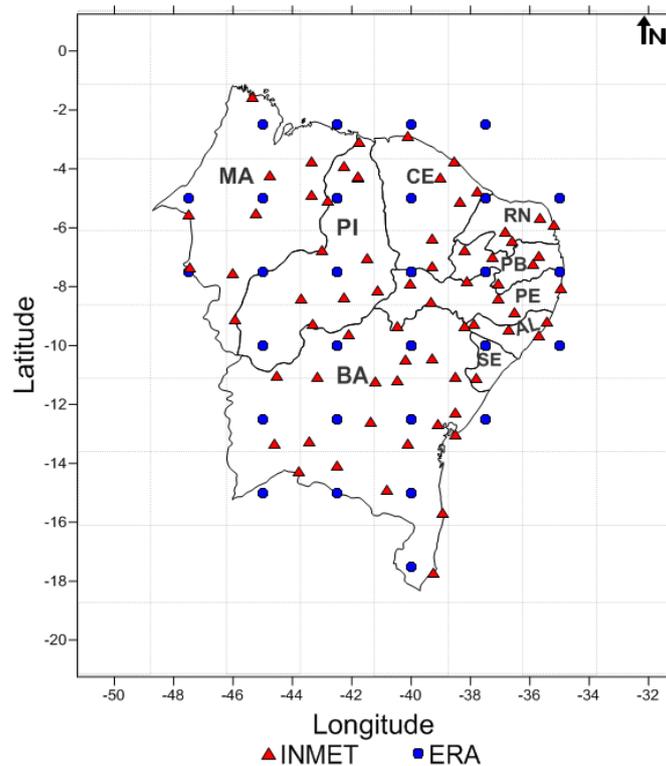
Fonte: Elaborado pelos autores (2020).

Na Figura 2 são apresentadas as estações do INMET com percentual de dados ausentes às 00h (a) e 12h (b) no período de 1961 a 2001. A escala varia entre 0 e 1, indicando o percentual de dados ausentes nos registros da base de dados. Os valores iguais a zero denotam que não há falhas nos registros, caso contrário, os valores indicam o respectivo percentual de dados ausentes nos registros da estação. Os locais em branco informam que a estação não existia ou foram desativadas.

Observa-se, ainda nas Figura 2 (a) e (b), que há estações com períodos sem registros e outros com quantidade reduzida. Devido a deficiência apresentada nos registros do INMET, foram selecionadas as estações que contém no máximo 30% de registros ausentes (Stüker et al., 2016; Schmidt, Cancelli e Junior, 2016). Portanto, ocorreu uma reformulação na base de dados, onde os dados do INMET ficaram composto por 69 estações no período de 1/1/1993 a 31/12/2000, correspondendo a 8 anos de observações às 00h e às 12h. Algumas lacunas persistiram, para preencher os registros ausentes nas 69 estações selecionadas foi utilizado o método de interpolação IDW ( $p = 2$ ). As estimativas foram realizadas considerando as informações ausentes para cada hora de cada dia de cada mês e ano, no software R (Team, 2006), preenchendo toda base de dados.

Os dados da ERA-40 foram reduzidos ao mesmo intervalo de tempo dos dados do INMET. Conseqüentemente, foram verificadas as estações do INMET que estavam dentro da grade que corresponde às coordenadas da ERA-40 (Figura 3). Sendo que cada ponto corresponde ao centro de uma grade da ERA-40 e, os triângulos as coordenadas das estações convencionais do INMET. No caso, em que há mais de uma estação dentro de uma grade da ERA-40, as estações do INMET ficaram representada por um único valor, obtido através da média das estações.

**Figura 3.** Estações do INMET (triângulos) e grades da ERA-40 (círculo) selecionadas para análise no período de 1993-2000.



Fonte: Elaborado pelos autores (2020).

### Sample Entropy

A regularidade das séries da ERA-40 e do INMET foram analisadas aplicando o método Sample Entropy. Os parâmetros foram definidos como em Pincus (1991), com nível de tolerância  $r = 20\%$  e tamanho dos vetores  $m = 2$ . Para cada série foram construídos vetores de tamanho dois e, definidas as distâncias entre os elementos do vetor. Em seguida obteve-se a probabilidade de vetores  $m$  e  $m+1$  pontos similares, com uma tolerância de  $20\% \cdot \text{desvio padrão}$ . Por fim, verificou-se o índice de padrões similares para as séries de vetores com tamanho  $m$  e  $m + 1$  pontos. As análises da Sample Entropy foram executadas no software R (Team, 2006).

De acordo com Richman & Moorman (2000), seja  $x(1), \dots, x(N)$  uma série temporal de tamanho  $N$ , o algoritmo da *SampEn*( $m, r, N$ ) é descrito como:

1. Construção de  $N + m - 1$  vetores  $x(1), x(2), \dots, x(N + m - 1)$ , definindo  $x(i)$  de acordo com a equação abaixo,

$$x(i) = [x(i), x(i + 1), \dots, x(i + m - 1)], \quad i = 1, \dots, N - m + 1 \quad (2)$$

em que  $m$  é o comprimento do vetor.

2. Definição da distância  $d$  entre os vetores  $x(i)$  e  $x(j)$ , como a máxima diferença absoluta entre seus elementos escalar correspondente, em que:

$$d[x(i), x(j)] = \max_{k=0, m, m-1} [|x(i+k) - x(j+k)|] \quad (3)$$

3. Para cada  $i = 1, \dots, N - m$ , calcula a probabilidade da ocorrência do padrão  $i$ , de tamanho  $m$ , dada uma tolerância  $r$  de combinação, para  $x_m(i) \neq x_m(j)$  garantindo que não haverá autocontagem.

$$B_i^m(r) = \frac{\text{número de } d \leq r}{N - m - 1} \quad (4)$$

$$A_i^{m+1}(r) = \frac{\text{número de } d \leq r}{N - m - 1} \quad (5)$$

4. Calcula-se a probabilidade que duas sequências são similares para  $m$  pontos ( $B^m(r)$ ) e  $m + 1$  pontos ( $A^{m+1}(r)$ ).

$$B^m(r) = \frac{1}{N - m} \left( \sum_{i=1}^{N-m} B_i^m(r) \right) \quad (6)$$

$$A^{m+1}(r) = \frac{1}{N - m} \left( \sum_{i=1}^{N-m} A_i^{m+1}(r) \right) \quad (7)$$

5. Por fim, calcula-se o índice Sample Entropy:

$$\text{SampEn}(m, r, N) = -\ln \left( \frac{A^{m+1}(r)}{B^m(r)} \right) \quad (8)$$

### Cross Sample Entropy

A técnica Cross-Sample Entropy (cross-Sample Entropy) baseada no método Sample Entropy apresenta maior consistência que a Cross-Approximate Entropy (crossApEN), é utilizada para medir a similaridade ou grau de sincronia entre duas séries temporais. A cross-Sample Entropy pode assumir valores altos correspondendo a séries altamente assíncronas e,

valores baixos equivalendo à sincronia entre as séries, os valores podem variar entre 0 (baixa entropia) e 2 (alta entropia), valores maiores que 2 não fornecem qualquer informação sobre a série (Richman & Moorman, 2000). As análises da cross-Sample Entropy foram processadas em linguagem C. O algoritmo para o cálculo da (cross-Sample Entropy) é apresentado a seguir:

Considere  $u = (u(1), u(2), \dots, u(N))$  e  $v = (v(1), v(2), \dots, v(N))$  serem duas séries de tamanho  $N$ . Fixando o tamanho dos vetores a serem comparados,  $m$ , e o nível de tolerância para aceitar similaridade,  $r$ .

1. Forme sequência de vetores para  $u$  e  $v$ , respectivamente.

$$x_m(i) = u(i), u(i + 1), \dots, u(i + m - 1), 1 \leq i \leq N - m \quad (9)$$

$$y_m(j) = v(j), v(j + 1), \dots, v(j + m - 1), 1 \leq j \leq N - m \quad (10)$$

2. Para cada  $i \leq N - m$  obtenha o número de vetores  $y_m(j)$  e  $y_{m+1}(j)$  dentro do nível de tolerância  $r$  do vetor  $x_m(i)$  e  $x_{m+1}(i)$ , respectivamente.

$$B_i^m(r)(v||u) = \frac{\text{número } 1 \leq j \leq N - m \text{ para } d[x_m(i), y_m(j)] \leq r}{N - m} \quad (11)$$

$$A_i^{m+1}(r)(v||u) = \frac{\text{número } 1 \leq j \leq N - m \text{ para } d[x_{m+1}(i), y_{m+1}(j)] \leq r}{N - m} \quad (12)$$

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

3. Calcula-se o número de pares de vetores das duas séries que estão dentro do nível de tolerância  $r$ .

$$B^m(r)(v||u) = \frac{(\sum_{i=1}^{N-m} B_i^m(r)(v||u))}{N - m} \quad (13)$$

$$A^{m+1}(r)(v||u) = \frac{(\sum_{i=1}^{N-m} A_i^{m+1}(r)(v||u))}{N - m} \quad (14)$$

4. Finalmente,

$$\text{crossSampEn}(m, r, N)(v||u) = \frac{A^{m+1}(r)(v||u)}{B^m(r)(v||u)} \quad (15)$$

### Inverse Distance Weighting

O Inverse Distance Weighting (IDW) é um método de interpolação proposto por Shepard (1968), que combina a ideia de proximidade, supondo que pontos mais próximos são mais semelhantes do que aqueles que estão mais distantes, ou seja, os valores medidos mais próximos ao local de previsão terão maior influência sobre o valor previsto, admitindo uma influência que diminuirá com a distância (Luo; Taylor; Parker, 2008). O IDW foi utilizado para espacialização dos resultados obtidos sobre o NE, corroborando nas análises. Os mapas foram confeccionados no software R (Team, 2006). A expressão matemática do IDW é dada pela equação,

$$\hat{Z}_{s_0} = \sum_{i=1}^N \lambda_i Z_{(s_i)} \quad (16)$$

em que  $\hat{Z}_{s_0}$  é o valor estimado para cada local  $s_0$ ,  $N$  é o número de pontos observados,  $Z_{(s_i)}$  é o valor observado e  $\lambda_i$  é o peso atribuídos a  $i$  – ésima localidade. O peso é determinado pelas equações:

$$\lambda_i = \frac{d_{i0}^{-p}}{\sum_{i=1}^N d_{i0}^{-p}} \quad (17)$$

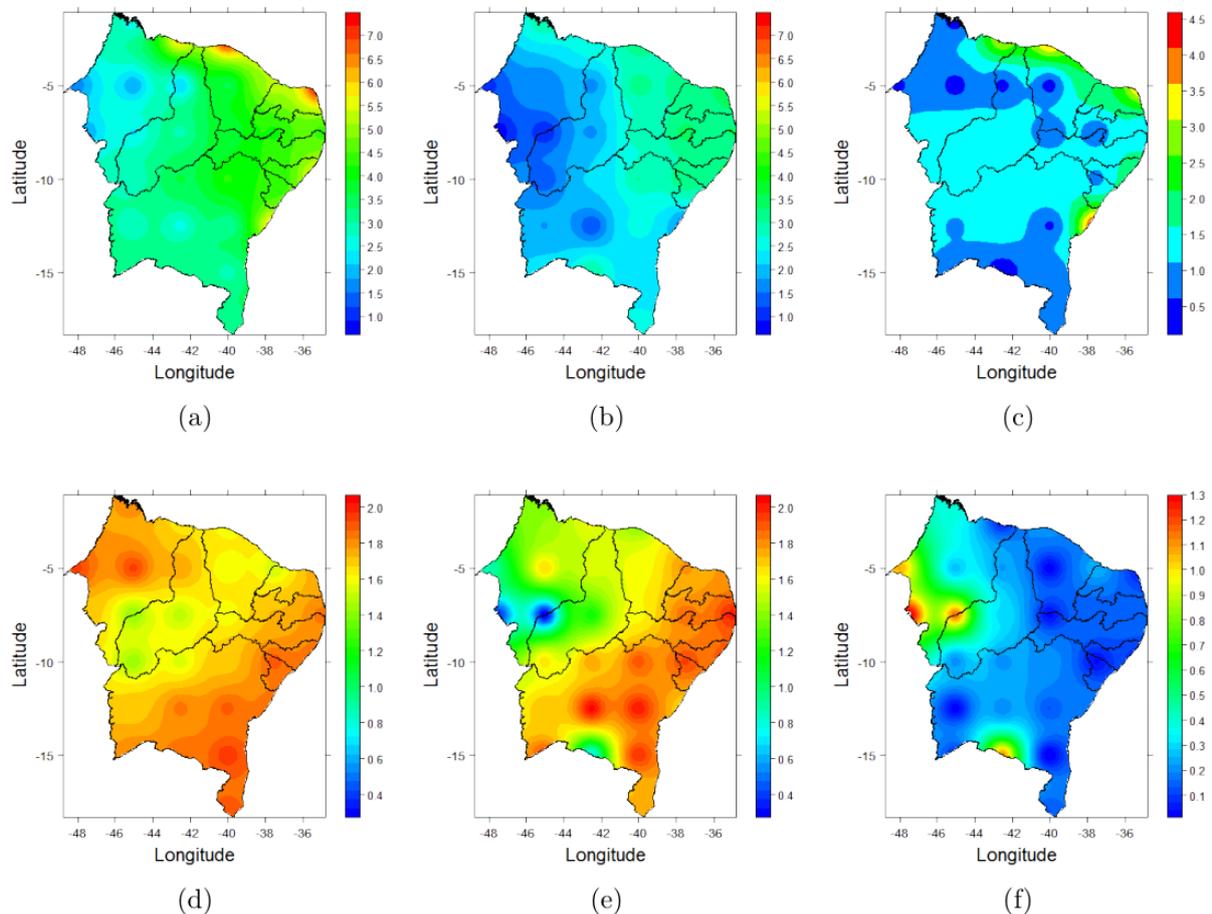
$$\sum_{i=1}^N \lambda_i = 1. \quad (18)$$

em que  $d_{i0}$  é a distância entre o local de previsão  $s_0$  e os locais medidos  $s_i$ . Os pesos atribuídos aos dados terão soma igual a 1.

### 3. Resultados e Discussão

Foram confeccionados mapas das médias e Sample Entropy da velocidade do vento, e das diferenças entre as médias e Sample Entropy da ERA-40 e INMET para séries Completa, 00h e 12h. Percebe-se que a média da ERA-40 (Figura 4 (a)) é maior que a do INMET (Figura 4 (b)) em todo o NE, com diferença de até 3,5 ms<sup>-1</sup> no Norte das quatro sub-regiões (Figura 4 (c)), onde também encontram-se as maiores médias de velocidade do vento.

**Figura 4.** Média da velocidade do vento para série Completa da ERA-40 (a), INMET (b) e a diferença (c) entre ambas. Sample Entropy da ERA-40 (d), INMET (e) e a diferença (f) entre as entropias sobre o NE brasileiro.



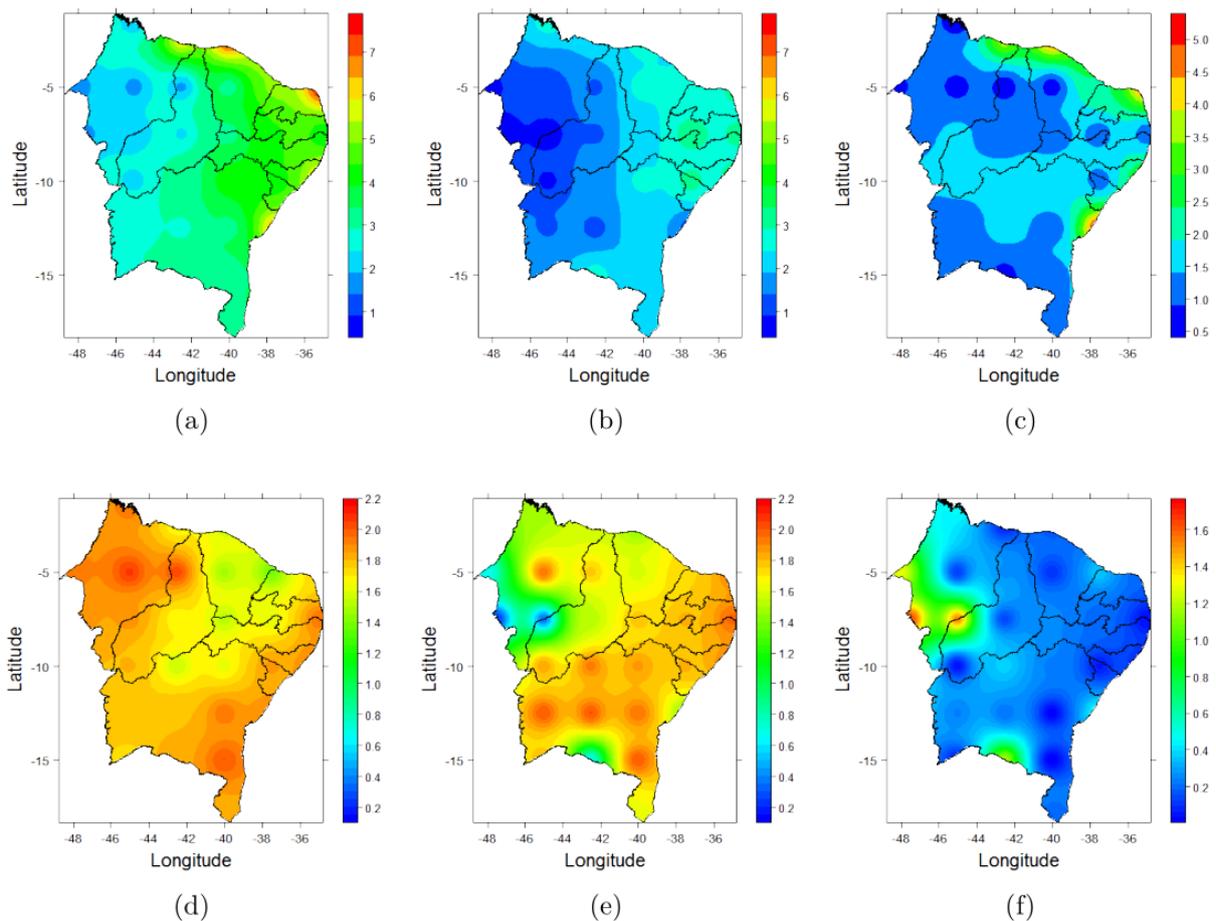
Fonte: Elaborado pelos autores (2020).

A Sample Entropy para ERA-40 (Figura 4 (d)) e INMET (Figura 4 (e)) sobre a Zona da Mata, Agreste e Sertão tem padrão semelhante com valores alto em locais onde a média da velocidade do vento são maior, indicando que a velocidade do vento é pouco previsível nestas regiões. Enquanto no Meio Norte o comportamento entre ambas é distinto, fato que é refletido na diferença da Sample Entropy (Figura 4 (f)). Vale ressaltar, que as áreas como o Norte das quatro sub-regiões, que tem maior registro de velocidade do vento apresentam as menores diferenças de Sample Entropy (Figura 4 (f)) com desempenho moderado, tal comportamento é interessante, uma vez que, nos estudos sobre a instalação de turbinas eólicas um dos fatores a ser considerado é a regularidade da velocidade do vento.

Na Figura 5, observa-se o comportamento da média e da Sample Entropy para a velocidade do vento às 00h. Verifica-se que a média da ERA-40 (Figura 5 (a)) distingue-se dos

valores das médias do INMET (Figura 5 (b)), com menor diferença (Figura 5 (c)) nas áreas de menor média de velocidade do vento.

**Figura 5.** Média da velocidade do vento às 00h para ERA-40 (a), INMET (b) e a diferença (c) entre ambas. Sample Entropy da ERA-40 (d), INMET (e) e a diferença (f) entre as entropias sobre o NE brasileiro.



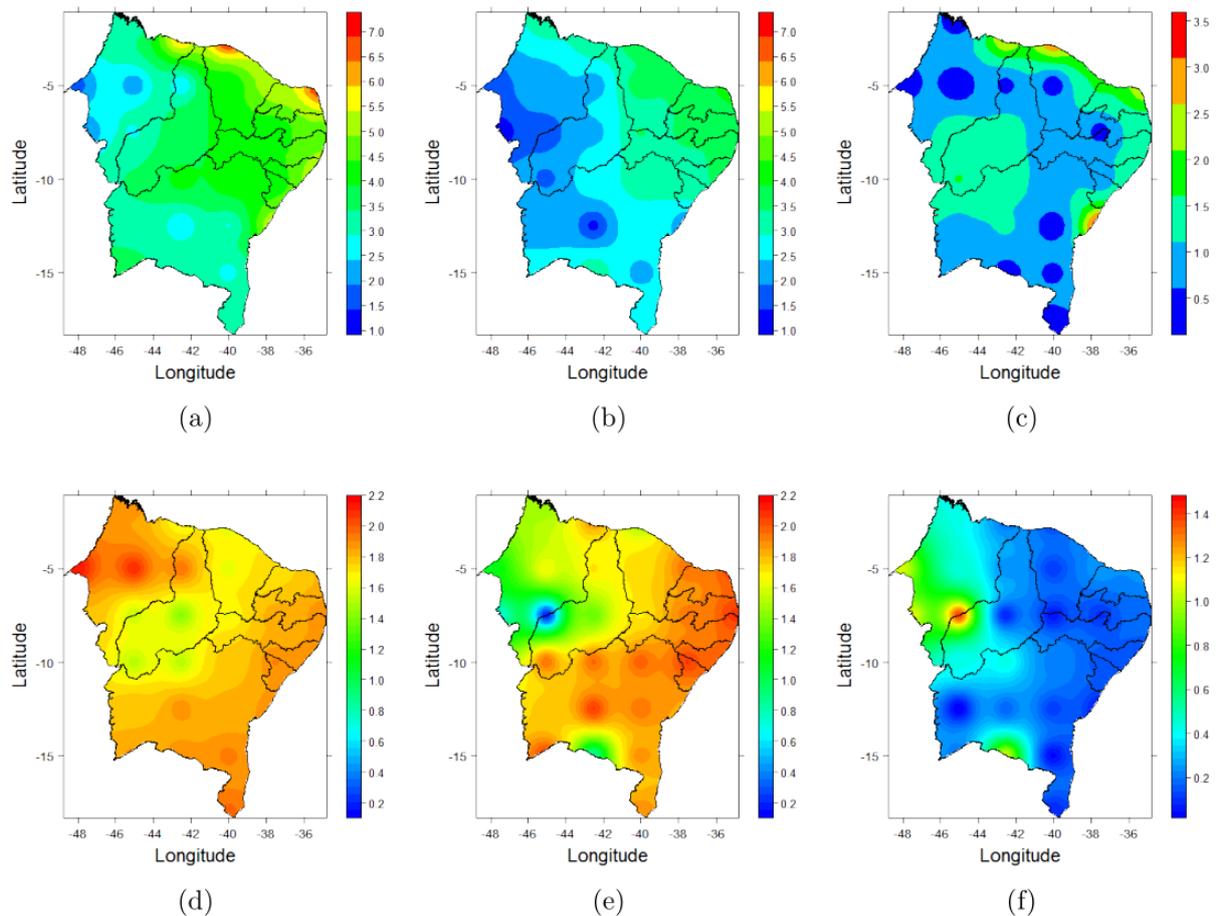
Fonte: Elaborado pelos autores (2020).

Observando a Sample Entropy para ERA-40 (Figura 5 (d)) a regularidade é moderada no Sertão entre estado do CE e RN e, as demais localidades apresentam-se mais irregulares. Para o INMET (Figura 5 (e)) o comportamento no Meio Norte é mais previsível, com ênfase na divisa do Sul do MA e PI, área que também apresenta a maior diferença (Figura 5 (f)) da regularidade entre as bases de dados. Portanto, têm-se que as áreas com menor diferença são também as de maior média da velocidade do vento, logo, são mais previsíveis.

Durante as 12h (Figura 6) a média da velocidade do vento é mais intensa no Norte das quatro sub-regiões tanto para ERA-40 (Figura 6 (a)) quanto para o INMET (Figura 6 (b)). Tal

conduta, pode estar relacionada as brisas marítimas que contribuem no aumento da intensidade dos ventos sobre NE (Varejão-Silva, 2006; Amarante et al., 2001).

**Figura 6.** Média da velocidade do vento às 12h para ERA-40 (a), INMET (b) e a diferença (c) entre ambas. Sample Entropy da ERA-40 (d), INMET (e) e a diferença (f) entre as entropias sobre o NE brasileiro.

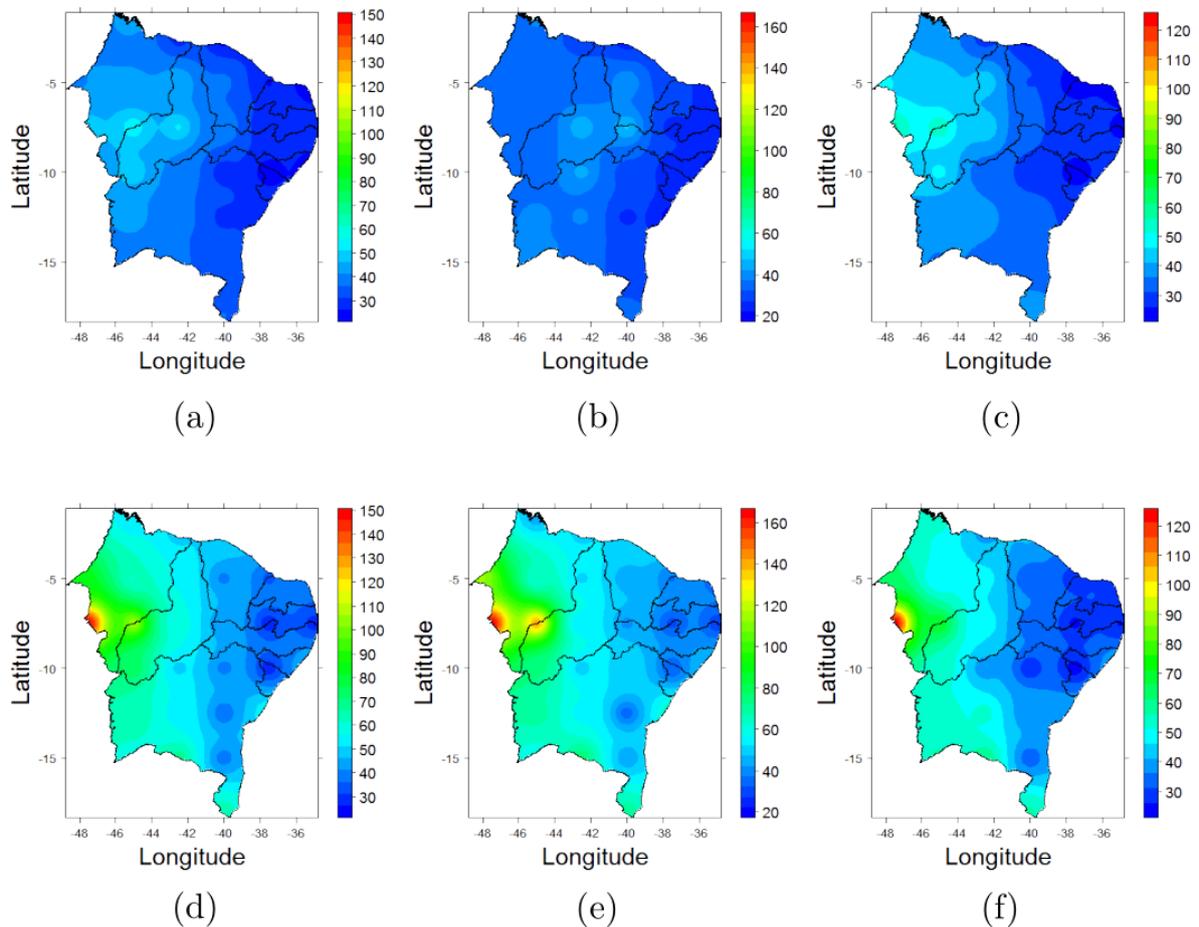


Fonte: Elaborado pelos autores (2020).

A regularidade da série às 12h é similar a das 00h (Figura 5), com menor previsibilidade na maior parte do NE. Contudo, as médias da velocidade do vento são mais regulares para o INMET (Figura 6 (e)). Ressalta-se a maior diferença (Figura 6 (f)) na porção Sul do Meio Norte, onde estão as menores médias da velocidade do vento. De acordo com Lira et al. (2017), essa região dispõe de propriedades propícias para produção de energia eólica. Além de, apresentar melhor regularidade da velocidade do vento como pode ser notado na Figura 6 (e), com maior previsibilidade que as demais regiões.

Na Figura 7, são verificados o coeficiente de variação (CV) da velocidade do vento da ERA-40 e do INMET.

**Figura 7.** Espacialização do coeficiente de variação da ERA-40 ((a), (b), (c)) e do INMET ((d), (e), (f),) para séries completa ((a), (d)), 00h ((b), (e)), 12h ((c), (f)) sobre o NE brasileiro.



Fonte: Elaborado pelos autores (2020).

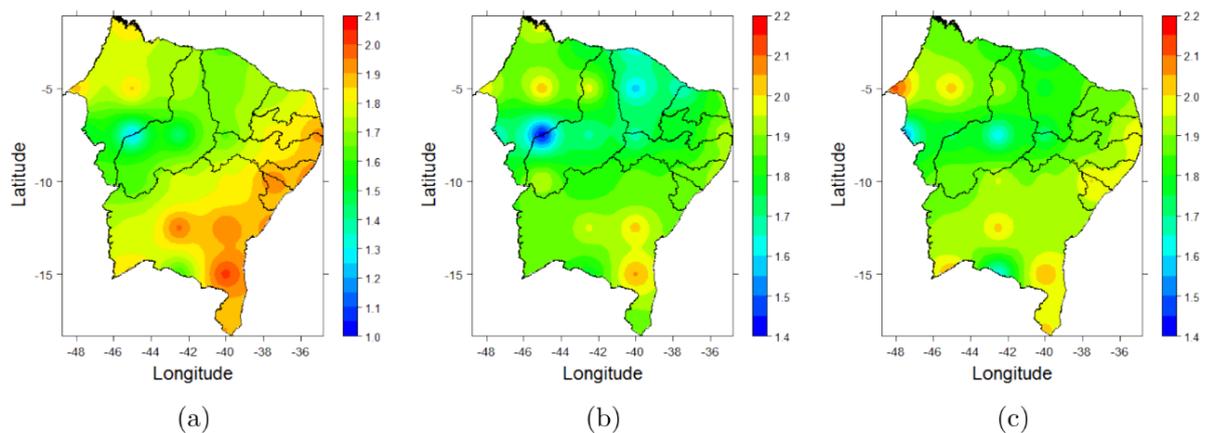
Na Figura 7, são verificados o coeficiente de variação (CV) da velocidade do vento da ERA-40 (Figura 7 (a), (b), (c)) e do INMET (Figura 7 (d), (e), (f)) para séries completa (Figura 7, (a), (d)), 00h (Figura 7 (b), (e)) e 12h (Figura 7 (c), (f)) a variabilidade da velocidade do vento é maior no Oeste do NE abrangendo as sub-regiões do Meio Norte e parte do Sertão para as diferentes base de dados. Observando que a escala altera de acordo com a série que está sendo analisada. Percebe-se também que para diferentes séries do INMET (Figura 7 (d), (e), (f)) há maior variabilidade principalmente no Meio Norte, área de menor registro de média da velocidade do vento com CV em torno de 140%, portanto, o coeficiente de variação é muito alto (Silva et al., 2012).

O grau de similaridade entre as séries do INMET e da ERA-40 foi avaliado através da cross-Sample Entropy. Os parâmetros foram determinados como  $m = 2$  e  $r = 20\%$ , definidos como em Pincus (1991). O algoritmo da cross-Sample Entropy investigou quantos vetores do

INMET e da ERA-40 ocorrem dentro de um intervalo estatisticamente significativo simultaneamente e, fornece um índice de similaridade entre ambas.

A Figura 8, representa a cross-Sample Entropy da ERA-40 com relação às estações convencionais do INMET.

**Figura 8.** Cross-Sample Entropy para série da velocidade do vento horária no NE brasileiro. Série completa (a), 00h (b) e 12h (c).



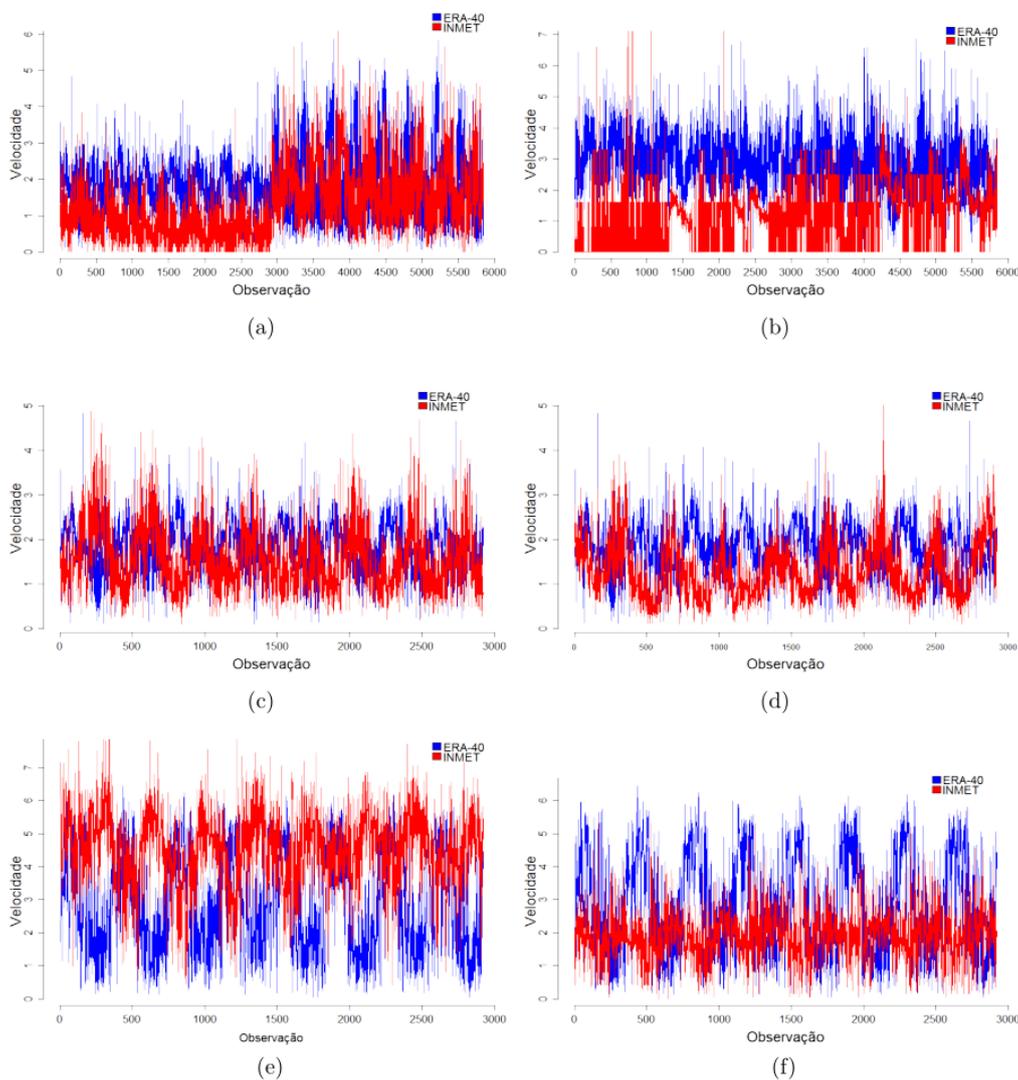
Fonte: Elaborado pelos autores (2020).

Observa-se que para a série completa (Figura 8 (a)), a similaridade é moderada no Meio Norte, com valores da cross-Sample Entropy entre 1.4 e 1.85, com ênfase na divisa do Sul do MA com o PI. Essa característica também pode ser observada durante as 00h (Figura 8(b)). Às 12h (Figura 8(c)), as séries já não possuem muitas semelhanças indicando assim que os dados da ERA-40 podem estar superestimando ou subestimando as observações do INMET. Desempenho semelhante foi encontrado com dados de reanálise no Rio Grande do Sul (Stüker et al., 2016).

Schmidt et al. (2016), ao analisar dados de estações automáticas do INMET, da reanálise ECMWF e NCAR, identificou que os dados da ECMWF reproduzem melhor os dados quando comparado ao NCAR, ressalta-se portanto, que foi necessário inserir um fator calibração nas médias da velocidade do vento do ECMWF para que tal resultado fosse obtido.

A Figura 9 indica que a ERA-40 superestima ou subestima os registros do INMET por meio das séries temporais da velocidade do vento que apresentam o menor e maior valor de cross-Sample Entropy.

**Figura 9.** Séries temporais da velocidade do vento para as séries Completa (a, b), 00h (c, d) e 12h (e, f), que obtiveram o valor mínimo (a, c, e) e máximo (b, d, f) na Cross Sample Entropy.



Fonte: Elaborado pelos autores (2020).

Na Figura 9, os resultados obtidos para as séries Completa (Figura 9 (a), (b)), às 00h (Figura 9 (c)), (d) e as 12h (Figura 9 (e), (f)), corroboram para os resultados obtidos na Figura 8. Por meio do comportamento das séries é possível notar que para o menor valor de cross-Sample Entropy, as observações do INMET e da ERA-40 têm desempenho similares (Figura 9 (a), (c), (e)). Enquanto para o maior valor da cross-Sample Entropy as séries têm performance diferente (Figura 9 (b), (d), (f)). Durante as 12h (Figura 9), as diferenças são mais evidentes tanto para o valor mínimo (Figura 9 (e)), quanto para o valor máximo (Figura 9 (f)), sendo que nesse horário os registros apresentam maior velocidade do vento.

#### **4. Considerações Finais**

Os resultados mostraram que em todas as séries os valores da Sample Entropy e da média foram maiores para ERA-40 do que para os INMET. No Sul do Meio Norte em todas as escalas, a Sample Entropy apresenta regularidade moderada, com ênfase para o INMET, embora, as médias da velocidade do vento nesta sub-região sejam menor. Por meio da Sample Entropy também foi possível observar que, locais com menor média da velocidade do vento tem maior previsibilidade. Quanto às diferenças, o Norte das quatro sub-regiões do NE detém a maior diferença nas médias para as três séries, enquanto a Sample Entropy tem menor diferença nesta localidade. Desta forma, a velocidade do vento apresenta regularidade moderada nas áreas com maior diferença na média da velocidade do vento e, portanto, são mais previsíveis quando comparada às demais áreas.

A comparação entre os dados da ERA-40 e do INMET através da cross-Sample Entropy apontam que as bases de dados são pouco similares em grande parte do NE principalmente para a série Total/Completa. Desta forma, observa-se uma superestimação ou subestimação nos dados da ERA-40 com relação às observações das estações convencionais do INMET. Contudo, como os dados analisados estão a 10m do solo, o desempenho das estações da ERA-40 e do INMET com relação a velocidade do vento mínima necessária para explorar o potencial eólica da região, alcança no Norte do NE média de velocidade do vento entre 7 e 8 ms<sup>-1</sup>, que de acordo com o Atlas de Energia Eólica é o mínimo desejado para sistemas eólicos de grande porte à uma altura 50 metros.

Como trabalho futuro, pode ser utilizado as diferenças obtidas entre as médias do INMET e ERA-40 como um possível fator de calibração e, verificar o comportamento das mesmas. Assim como, realizar a análise das células da grade da ERA-40 com uma resolução menor.

#### **Agradecimentos**

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

#### **Referências**

ABEEólica. (2020). Associação Brasileira de Energia Eólica. Acesso em 15 maio 2020, em: [http://abeeolica.org.br/wp-content/uploads/2020/04/Infovento-15\\_PT.pdf](http://abeeolica.org.br/wp-content/uploads/2020/04/Infovento-15_PT.pdf).

- Amarante, O. A., Brower, M., Zack, J., & Sá, A. L. (2001). Atlas do potencial eólico brasileiro. Brasília, DF. CD-ROM.
- Camelo, H. do N., Lucio, P. S., Gomes, O. M., & Leal Junior, J. B. V. (2016). Predição de velocidade do vento em municípios do Nordeste brasileiro através de regressão linear e não linear para fins de geração eólica. *Revista Brasileira de Geografia Física*, 9(03), 927-939.
- Carneiro, T. C., & de Carvalho, P. C. M. (2015). Caracterização de potencial eólico: estudo de caso para Maracanaú (CE), Petrolina (PE) e Parnaíba (PI). *Revista Brasileira de Energia Solar*, 6(1). Acesso em 29 abril 2020. Disponível em: <https://rbens.emnuvens.com.br/rbens/article/view/122/122>.
- Chou, C. M. (2014). Complexity analysis of rainfall and runoff time series based on sample entropy in different temporal scales. *Stochastic Environmental Research And Risk Assessment*, 28(6), 1401-1408. doi: <https://doi.org/10.1007/s00477-014-0859-6>.
- ECMWF. European Centre for Medium-Range Weather Forecasts. 2017. Disponível em: <https://www.ecmwf.int/>.
- GWEC, G. W. P. C. (2019). Global Wind Energy Outlook. Brussels.
- Lira, M. A. T., Neto, J. M. M., Loiola, J. V. L. d., Silva, E. M. d., & Alves, J. M. B. (2017). Caracterização do regime de ventos no Piauí para o aproveitamento de energia eólica. *Revista Brasileira de Meteorologia*, 32(1), 77–88. doi: <http://dx.doi.org/10.1590/0102-778632120150712>.
- Luo, W., Taylor, M. C., & Parker, S. R. (2008). A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 28(7), 947-959. doi: <https://doi.org/10.1002/joc.1583>.
- Pereira A. S. et al. (2018). *Metodologia da pesquisa científica*. [e-book]. Santa Maria. Ed. UAB/NTE/UFSM. Disponível em:

[https://repositorio.ufsm.br/bitstream/handle/1/15824/Lic\\_Computacao\\_Metodologia-Pesquisa-Cientifica.pdf?sequence=1](https://repositorio.ufsm.br/bitstream/handle/1/15824/Lic_Computacao_Metodologia-Pesquisa-Cientifica.pdf?sequence=1).

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88(6), 2297–2301. doi: <https://doi.org/10.1073/pnas.88.6.2297>.

Reddy, Y. V., & Sebastin, A. (2006, December). Parameters for estimation of entropy to study price manipulation in stock market. In *10th Capital Markets Conference, Indian Institute of Capital Markets Paper*. doi: <http://dx.doi.org/10.2139/ssrn.962329>.

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6), H2039–H2049. doi: <https://doi.org/10.1152/ajpheart.2000.278.6.H2039>.

Santana, L. V. R., Stosic, T., Dezotti, C. H., De Albuquerque Moura, G. B., De Araújo, L. H. G. D., & da Silva, A. S. A. (2015). Spatial analyses of wind speed in the North-Brazil with data from ERA-40. *Revista Brasileira de Biometria*, 33(3), 414–432. Acesso em 30 maio 2020. Disponível em: <http://www.biometria.ufla.br/index.php/BBJ/article/view/24>.

Schmidt, J., Cancelli, R., & Junior, A. O. P. (2016). The effect of windpower on long-term variability of combined hydro-wind resources: The case of Brazil. *Renewable and Sustainable Energy Reviews*, 55, 131–141. doi: <https://doi.org/10.1016/j.rser.2015.10.159>.

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM national conference*, 517–524. doi: <https://doi.org/10.1145/800186.810616>.

Silva, Vicente de Paulo Rodrigues da, Pereira, Emerson Ricardo Rodrigues, & Almeida, Rafaela Silveira Rodrigues. (2012). Estudo da variabilidade anual e intra-anual da precipitação na região Nordeste do Brasil. *Revista Brasileira de Meteorologia*, 27(2), 163–172. <https://doi.org/10.1590/S0102-77862012000200005>.

Stüker, E., Schuster, C. H., Schuster, J. J., Santos, D. C., Medeiros, L. E., Costa, F. D., ... & Puhales, F. S. (2016) Comparison of wind data of ERA-Interim reanalysis and cfsr with the data from automatic inmet stations Rio Grande Do Sul. *Ciência e Natura*, 38(IX WORKSHOP), 284. Doi: 10.5902/2179460X20233.

Team, R. C. (2000). A language and environment for statistical computing. R Foundation for Statistical Computing.

Varejão-Silva, M. A. (2006). *Meteorologia e climatologia*. Recife, PE.

Witzler, L. T., Ramos, D. S., Camargo, L. A. S., & Guarnier, E. (2016, June). Reconstruction of wind generation historical series aiming at the analysis of energy complementarity: Methodology and applications. In *2016 13th International Conference on the European Energy Market (EEM)*, 1-6. IEEE. Doi: 10.1109/EEM.2016.7521324.

#### **Porcentagem de contribuição de cada autor no manuscrito**

Lêda Valéria Ramos Santana – 25%

Tatijana Stosic – 25%

Tiago Alessandro Espínola Ferreira – 25%

Antonio Samuel Alves da Silva – 25%