

## **Avaliação da Velocidade do Vento Utilizando Hindcast Desenvolvido pela Climatempo para a Aplicação dos Recursos Eólicos no Brasil**

Ismael Guidson Farias de Freitas<sup>1</sup>, Vitor Hugo de Almeida Marrafon<sup>1</sup>, Pedro Regoto de Souza<sup>1</sup>, Vitor Vaz Hassan<sup>1</sup>, Renata Genova Martins<sup>1</sup>, Luciano Ritter Nolasco Júnior<sup>1</sup>, Lara Amanda Marques Vieira<sup>1</sup>, Rafael Bitencourt Benassi<sup>1</sup>

1, Agência Brasileira de Meteorologia Ltda - Climatempo

ismael.freitas@climatempo.com.br, vitor.marrafon@climatempo.com.br,  
pedro.regoto@climatempo.com.br, vitorhassan@climatempo.com.br,  
renata.martins@climatempo.com.br, luciano.ritter@climatempo.com.br,  
lara.marques@climatempo.com.br, rafael.benassi@climatempo.com.br

### **RESUMO**

A ausência de dados meteorológicos no Brasil é um problema recorrente na análise e prospecção de novas áreas para construção de parques eólicos. Portanto, este estudo tem como objetivo validar os dados de velocidade de vento simulados pelo hindcast Climatempo através do modelo Weather Research & Forecasting Model (WRF), usando medições de estações anemométricas e dados da reanálise do ERA5 em locais que possuem parques eólicos no Brasil. Este tipo de avaliação torna-se importante, uma vez que estes dados podem auxiliar na identificação de locais propícios à instalação de novos parques eólicos. Desta forma, o hindcast Climatempo foi configurado com resolução espacial de 9.5 km e comparado com a reanálise ERA5 que possui resolução espacial de 27 km, ambos com resolução temporal de 1 h. Na validação, foram selecionadas 4 estações anemométricas do portal de dados abertos do ONS (Operador Nacional do Sistema Elétrico). Para avaliação estatísticas dos dados foram utilizados as métricas de erro médio (Bias), raiz do erro quadrático médio (RMSE) e correlação de Pearson (R). Os resultados mostraram que o hindcast Climatempo possui erros de Bias e RMSE inferiores em relação ao ERA5, abaixo de -0.89 m/s e 3.37 m/s, respectivamente. Tais resultados, indicam que os dados do hindcast Climatempo possuem bom desempenho em representar a velocidade do vento e podem ser aplicados em estudos com foco na avaliação de potencial eólico em diferentes regiões do Brasil.

**Palavras-chaves:** Potencial eólico, Hindcast Climatempo, Base de dados eólico, Velocidade do vento

### **ABSTRACT**

The lack of meteorological data in Brazil is a recurring problem in the analysis and prospecting of new areas for the construction of wind farms. Therefore, this study aims to validate the wind speed data simulated by the Climatempo hindcast through the Weather Research & Forecasting Model (WRF) model, using measurements from anemometric stations and data from the ERA5 reanalysis in locations that have wind farms in Brazil. This type of assessment is important because it can help identify suitable locations for the installation of new wind farms. Thus, the Climatempo hindcast was configured with a spatial resolution of 9.5 km and compared with the ERA5 reanalysis, which has a spatial resolution of 27 km, both with a temporal resolution of 1 hour. In the validation process, four anemometric stations were selected from the open data portal of the ONS (Operador Nacional do Sistema Elétrico). Statistical evaluation of the data was done using mean error (Bias), root mean square error (RMSE), and Pearson correlation (R). The results showed that the Climatempo hindcast has lower Bias and RMSE errors compared to ERA5, with values below -0.89 m/s and 3.37 m/s, respectively. These results indicate that the data from the Climatempo

hindcast perform well in representing wind speed and can be applied in studies focused on assessing wind potential in different regions of Brazil.

**Keywords:** Wind potential; Climatempo hindcast; Wind data database, Wind speed

## 1. INTRODUÇÃO

O Brasil possui uma vasta extensão territorial e possui um imenso potencial para aproveitamento de recursos energéticos renováveis, como exemplo cita-se a energia eólica que, nos últimos anos, tem se destacado como uma importante fonte de energia limpa no país [1]. A critério de conhecimento, segundo dados do último INFOVENTO da Associação Brasileira de Energia Eólica (ABEEólica), a geração de energia proveniente de parques eólicos em operação, apenas em 2022, produziu 3.843 GWh [2]. Cabe destacar que, a exploração eólica está se expandindo cada vez mais em áreas continentais com maiores velocidades de vento e que conseqüentemente possuem alto potencial de geração [3 e 4].

No Brasil, a utilização de energia eólica vem exercendo grande importância para o sistema elétrico, pois representa 15.5% do total da matriz energética, com 26 GW de capacidade instalada distribuídos em 916 parques eólicos com 10.178 aerogeradores. Estima-se que até 2028, a capacidade instalada atingirá 44,78 GW, considerando os parques em fase de construção e os leilões finalizados [2]. Como esse recurso natural é abundante em diversas regiões do Brasil [4 e 5], ainda é importante realizar estudos que identifiquem locais propícios para a instalação de novos parques eólicos visando aproveitar ao máximo a capacidade energética disponível para atender a demanda de diferentes setores da economia [1].

Dados observacionais de velocidade do vento em diferentes níveis são amplamente utilizados para analisar e definir as características de áreas propícias para a geração de energia eólica [6]. Comumente estes dados são obtidos através de estações anemométricas automáticas. No entanto, os principais problemas que ainda dificultam os estudos relacionados a dados anemométricos é a falta de medição de vento em diversas áreas do Brasil, devido aos desafios técnicos e altos custos associados à realização de projetos de medição [7]. De fato, ao analisar a distribuição das estações de medição de velocidade do vento, é comum observar uma concentração maior dessas estações próximas a grandes centros urbanos.

Diante da ausência de dados in situ, outras fontes de dados têm exercido um papel fundamental na caracterização e monitoramento de velocidade de vento e esforços vêm sendo realizados para contornar a falta de dados observacionais [1, 8, 9, 10 e 11]. Em âmbito nacional, são realizados estudos comparando diferentes fontes de dados (MERRA, ERA5, entre outras) para avaliar os recursos eólicos e a qualidade desse dado com observações de superfície [8, 11 e 12]. À vista disso, estudos concluíram que existe vantagem na utilização desses dados, principalmente de modelos de alta resolução, pois eles possuem uma boa resolução espacial e temporal e ausência de falhas na série. Um exemplo deste dado é o hindcast, que consiste em um sistema de análise retrospectiva criado com simulações numéricas usadas para determinar as condições passadas de forma a obter informações mais próximas às observações sobre áreas e períodos, onde dados observados não estão presentes.

Nesse contexto, a Climatempo tem desempenhado um papel fundamental na avaliação da velocidade do vento, fornecendo informações precisas para o setor de energia eólica brasileiro. Através do desenvolvimento do hindcast, a Climatempo visou contribuir para a avaliação dos recursos energéticos do Brasil, possibilitando a identificação de áreas com maior potencial para a geração de energia eólica, permitindo uma utilização mais eficiente e sustentável dos recursos naturais do País.

Diante deste contexto, o objetivo deste trabalho é validar os dados do hindcast desenvolvido pela Climatempo utilizando o modelo WRF4.3.3, comparando com dados do ERA5 e medições de superfície para um determinado período em que os dados observados estavam disponíveis (ano de 2022).

## 2. DADOS E METODOLOGIA

### 2.1. DADOS

#### 2.1.1 REANÁLISES FNL

Para realização das simulações do hindcast foi utilizado o modelo WRF4.3.3, e como condições iniciais e de contorno os dados da reanálise FNL do National Centers for Environmental Prediction (NCEP) com resolução de  $0,25^\circ \times 0,25^\circ$  de latitude-longitude para todo o globo disponibilizados a cada 6 h. Estes dados representam uma análise global operacional do NCEP FNL, também conhecidos como NCEP final, onde são produzidos a partir do Global Data Assimilation System (GDAS). O FNL utiliza o mesmo sistema

de assimilação e previsão de dados do GFS-NCEP (Global Forecast System), porém são atrasados para que mais dados observacionais possam ser usados, dessa forma, o GFS para a realização de previsão usa o FNL do ciclo anterior de 6 horas como parte de sua inicialização [13].

### 2.1.2 REANÁLISES ERA5

Em diversos estudos que envolvem velocidade do vento, têm sido utilizados dados de reanálises do ERA5, dessa forma, utilizamos a reanálises para verificar também o desempenho do hindcast em relação ao ERA5.

As reanálises ERA-5 do ECMWF (European Centre for Medium-Range Weather Forecasts) são obtidas a partir da API (Application Programming Interface) da CDS (Climate Data Store), conforme detalhado por Hersbach e Dee (2016) [14].

Nesse estudo foram utilizados dados da componente zonal e meridional do vento nas alturas de 100 m de altura em relação à superfície para extrair informações de velocidade no ano de 2022, vale ressaltar, que os dados para 80 metros foram extrapolados usando a lei da potência [15]. O ERA5 fornece dados em intervalos de 1 hora com resolução horizontal de  $0,25^\circ \times 0,25^\circ$  de latitude e longitude (27 km).

### 2.1.3 VALIDAÇÃO: DADOS DE VELOCIDADE DO VENTO ONS

Para avaliar o desempenho das simulações do hindcast de forma quantitativa, foram selecionadas 4 estações anemométricas para realização das comparações entre as bases. Essas estações fazem parte do Portal de dados Operador Nacional do Sistema Elétrico (ONS) que visa facilitar, melhorar, democratizar o acesso e consumo desse tipo de informação por parte de todos os seus usuários e públicos, disponibilizando diversos conjuntos de dados relacionados ao setor elétrico brasileiro, e assim garantindo a disponibilidade de acesso, reuso e redistribuição [16].

Os dados estão disponíveis no portal de dados abertos do ONS, onde foram baixados os valores de velocidade do vento com resolução temporal de 30 minutos medidos nas alturas de 80 metros em locais que possuem parques eólicos no Brasil, dessa forma, foi utilizado também a lei da potência para extrapolação [15]. A comparação entre os dados do hindcast e observados foi realizada de janeiro de 2022 a dezembro de 2022. As informações das estações com os nomes dos locais e posição estão listados na Tabela 1.

**Tabela 1** - Localização das estações anemométricas disponíveis no portal de Dados Abertos do ONS utilizadas nas avaliações.

Nome	Latitude (Sul)	Longitude (Leste)
Coxilha Seca	-30.8780	-55.6701
Serra do Espinhaço	-14.4830	-42.5842
Atlântica IV	-30.2995	-50.3278
Ventos de Campo Formoso II	-10.5338	-40.6270

## 2.2 HINDCAST

O hindcast foi desenvolvido utilizando o Modelo WRF (Weather Research & Forecasting Model) versão 4.3.3 com níveis verticais ETA [17]. Estes níveis são uma coordenada vertical para modelos atmosféricos definidos com uma representação de topografia semelhante a um degrau. Estes níveis foram selecionados para aumentar (diminuir) a resolução em altitudes específicas da atmosfera e também para fornecer melhor estabilidade do modelo, uma vez que terrenos mais acidentados podem induzir o cisalhamento do vento.

O primeiro critério para configuração do hindcast foi selecionar os níveis ETA (maior resolução na área estudada), selecionando 47 níveis, seguindo um perfil de tangente hiperbólica, com quase metade acima de 800 hPa, e um quarto deles acima de 940 hPa, dessa forma, cria níveis em torno do Alturas alvo (ou muito próximas a ela), como também melhora os resultados em convecção de pequena escala, e ventos topográficos gerados. A resolução espacial foi de 9.5 km de latitude e longitude e resolução horária.

A área de integração do modelo está compreendida entre as latitudes de 21.92°N – 52.94°S e longitude de -31.38°W – -88.64.11°W, onde engloba toda a região da América do Sul. (Figura 1).

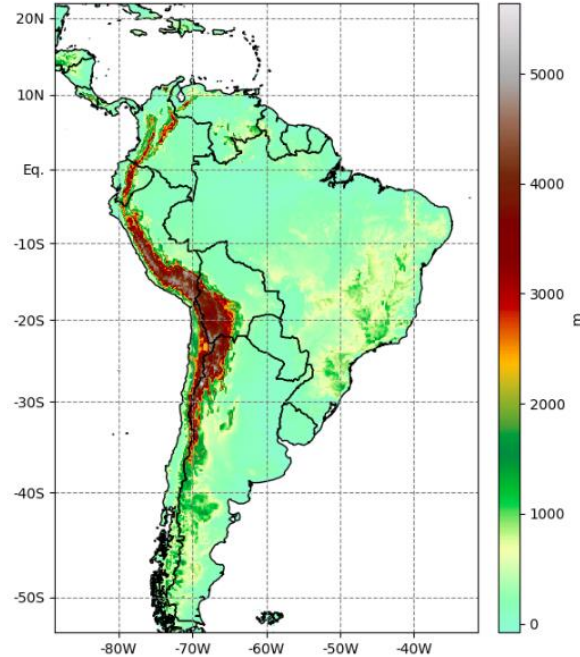


Figura 1 - Domínio da área de integração do hindcast utilizado pelo WRF. A figura representa também a topografia (m: metros)

## 2.2 VALIDAÇÃO ESTATÍSTICA

Para verificar a destreza do modelo quantitativamente, foram calculados indicadores estatísticos como o Erro Médio (BIAS), a Raiz do Erro Quadrático Médio (RMSE) e a Correlação de Pearson ( $r$ ), e apresentado Box-plot para períodos sazonais (DJF- Dezembro, Janeiro e Fevereiro, MAM- Março, Abril e Maio, JJA- Junho, Julho e Agosto, e SON- Setembro, Outubro e Novembro), conforme descrito abaixo [18].

O BIAS, considera uma determinada variável ( $X$ ) com um sub-índice ( $M$ ), mostrando o valor simulado ( $X_M$ ) e o valor observado ( $X_O$ ), sendo possível identificar se os dados observados estão sendo subestimado ou superestimado pelo hindcast. Assim, o BIAS para a variável de interesse ( $X$ ), foi calculado da seguinte forma (Equação 1):

$$bias = \frac{1}{N} \sum_{i=1}^N (X_M - X_O) \quad (01)$$

O RMSE mede a dimensão real dos erros, calculado de acordo com a Equação 2:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_M - X_O)^2} \quad (02)$$

Foi calculado também o coeficiente de correlação de Pearson, que leva em consideração a fase e a diferença entre duas variáveis (simuladas e observadas), possibilitando comparar os índices de velocidade de vento, simulados pelo hindcast e observados pelas estações anemométricas. Descrito através da Equação 3:

$$R = \frac{\sum_{t=1}^N (X_M - X'_M)(X_O - X'_O)}{\sqrt{\sum_{i=1}^N (X_M - X'_M)^2 \sum_{t=1}^N (X_O - X'_O)^2}} \quad (3)$$

Onde  $X'$  indica a média.

### 3. RESULTADOS E DISCUSSÕES

Vale destacar que estes resultados ainda são preliminares e foi avaliado o ano de 2022. Esta análise preliminar de apenas um ano é devido à falta de dados observados para os demais anos simulados. Dessa forma, como o foco são as regiões eólicas, selecionamos 4 locais que possuem parques eólicos e estações anemométricas disponíveis para a validação.

As Figuras 2 e 3, mostram box-plot da distribuição de dados do hindcast, ERA5 e estações anemométricas (lado esquerdo), e distribuição de resíduos (lado direito) para 80 e 100 metros, respectivamente. De maneira geral, observa-se que os dados do hindcast são capazes de representar a variabilidade dos dados observados de velocidade do vento nas estações avaliadas, apresentando forte semelhança entre as bases para a maioria das estações, fica evidente ao verificar a distribuição de resíduos, onde ao comparar com ERA5 verifica-se que o hindcast apresenta erros mais próximos de 0 m/s, indicando que o hindcast possui melhor capacidade de representar a velocidade do vento em 80m e 100m.

Ainda com relação às Figuras 2 e 3, é possível notar que o ERA5 possui tendência a subestimar a velocidade do vento em aproximadamente 15 m/s, enquanto o hindcast possui uma tendência maior em superestimar, porém com menos intensidade, principalmente nas estações Coxilha Seca, Serra do Espinhaço e Ventos de Campo Formoso II, durante os trimestres DJF, MAM, JJA e SON. Uma exceção é a estação Atlântica IV, onde o ERA5 apresentou o melhor desempenho em todos os trimestres.

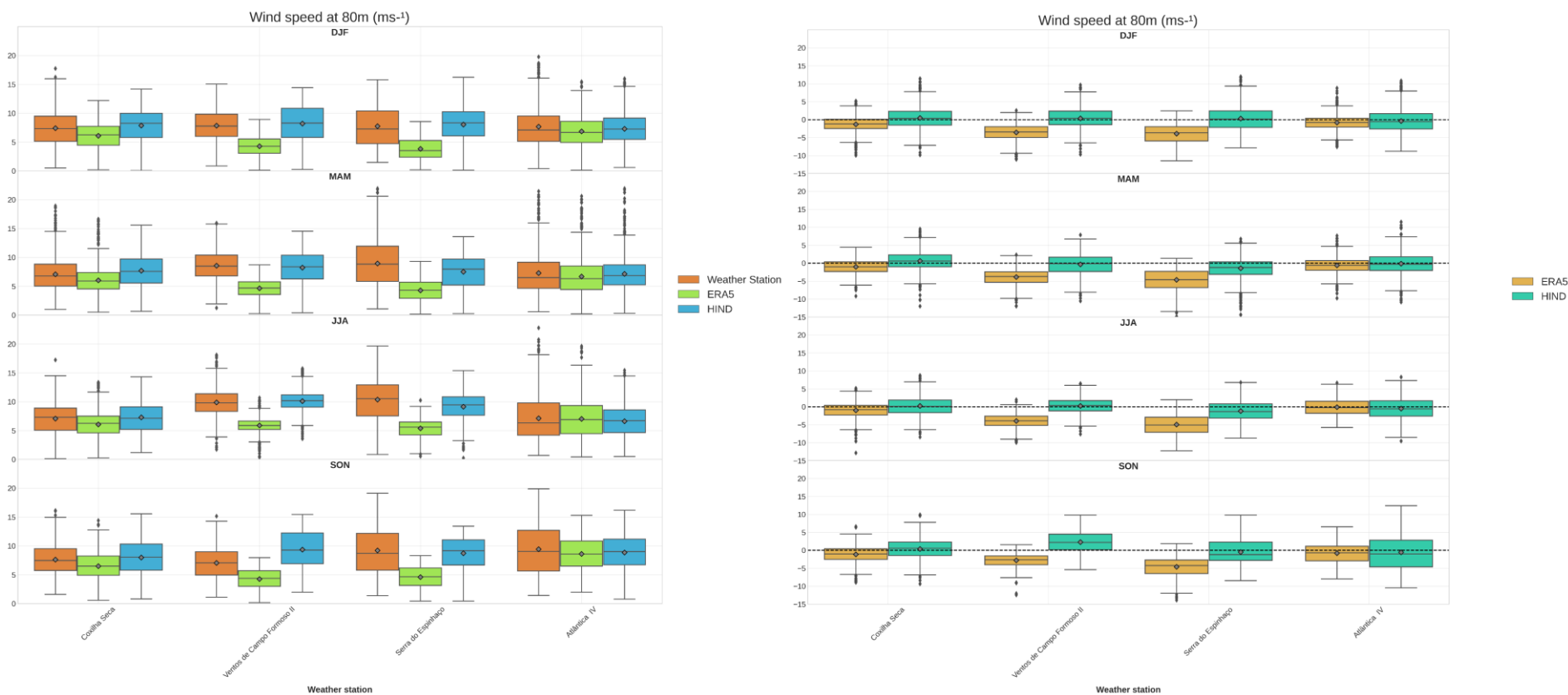


Figura 2 - Box-plot da distribuição de dados do hindcast, ERA5 e estações anemométricas (lado esquerdo), e distribuição de resíduos (lado direito) para velocidade do vento a 80m.

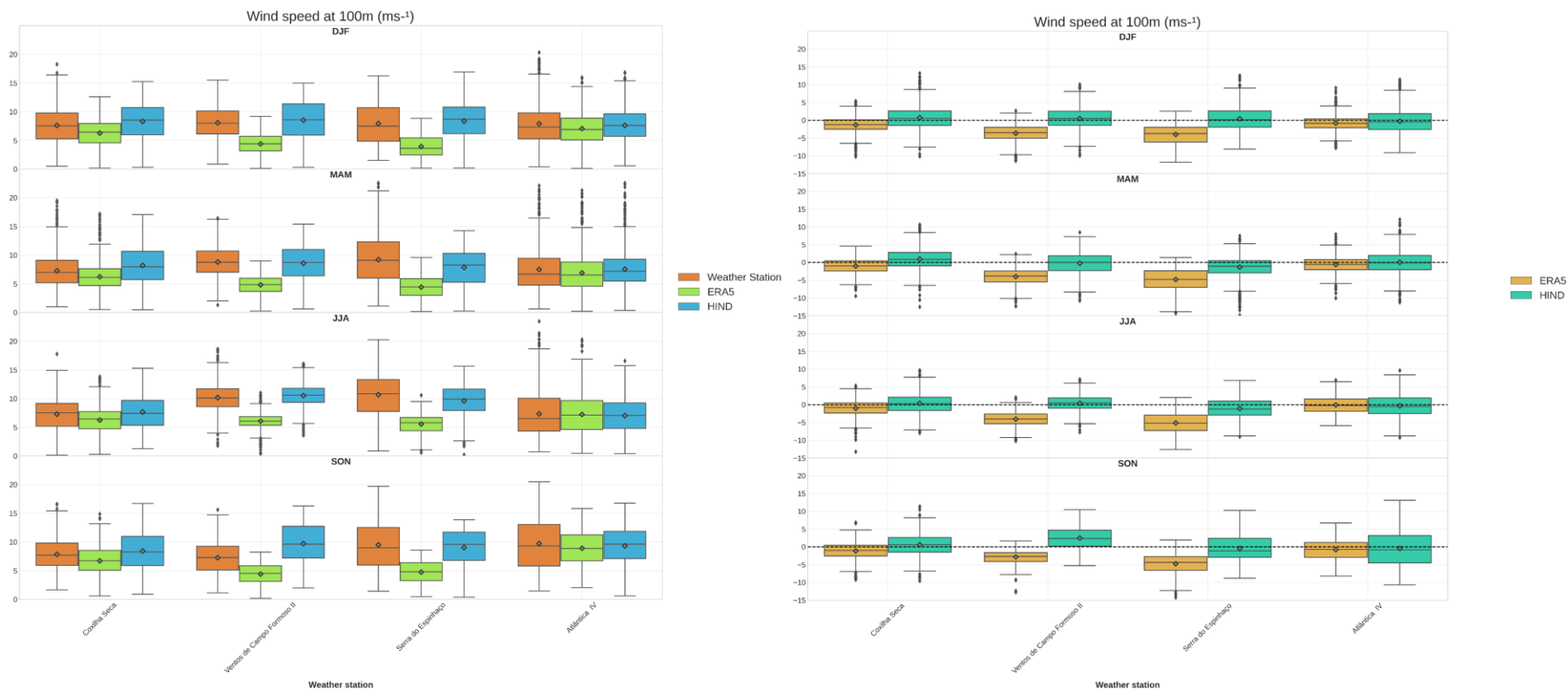


Figura 3 - Box-plot da distribuição de dados do hindcast, ERA5 e estações anemométricas (lado esquerdo), e distribuição de resíduos (lado direito) para velocidade do vento a 100m.

As figuras 4 e 5, mostram o ciclo diário de todo o período dos dados do ERA5 e hindcast na altura de 80m e 100m, respectivamente, avaliando as métricas de Biais, RMSE e Correlação. De acordo com os valores de Biais, observa-se nas alturas de 80m e 100m que o ERA5 tende a subestimar a velocidade do vento em todas as estações em aproximadamente 0 e -5.5 m/s, enquanto o hindcast apresentou valores de Biais entre 2.5 e -1.5 m/s, sem grandes tendências de subestimar/superestimar a velocidade do vento. Ainda com relação ao erro, o hindcast apresentou valores sempre abaixo de 4 m/s em todos os horários e estações, enquanto o ERA5 apresentou RMSE mais alto, como no caso das estações de Serra do Espinhaço e Ventos de Campo Formoso II que apresentaram RMSE de aproximadamente 7 m/s. Com relação à correlação, o ERA5 apresentou melhor desempenho, porém o hindcast ainda possui valores de correlação significativos acima de 0.5.

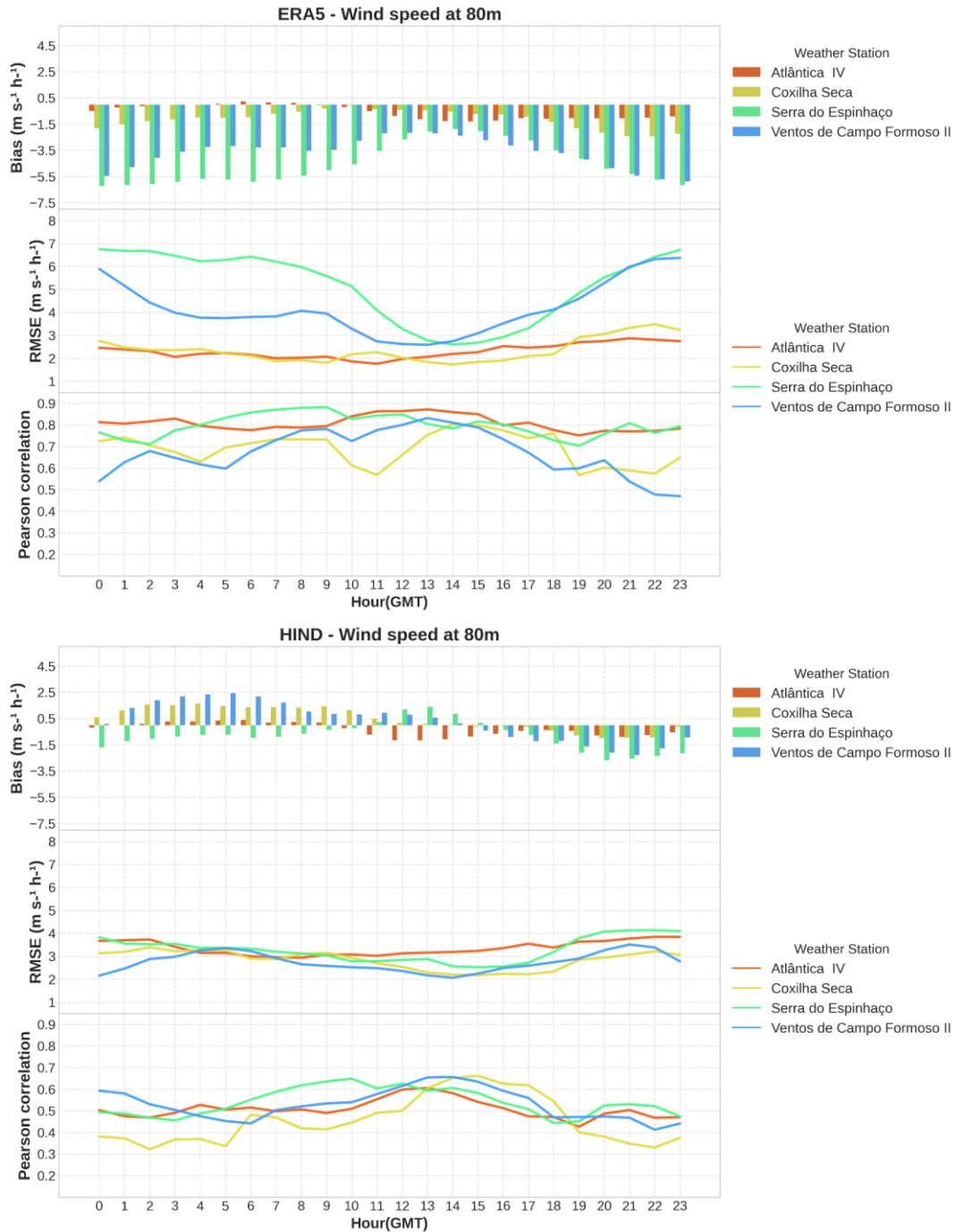


Figura 4 - Ciclo diário do desempenho dos dados do ERA5 e hindcast na altura de 80m considerando



todo o período disponível, avaliando as métricas de Biais, RMSE e Correlação.

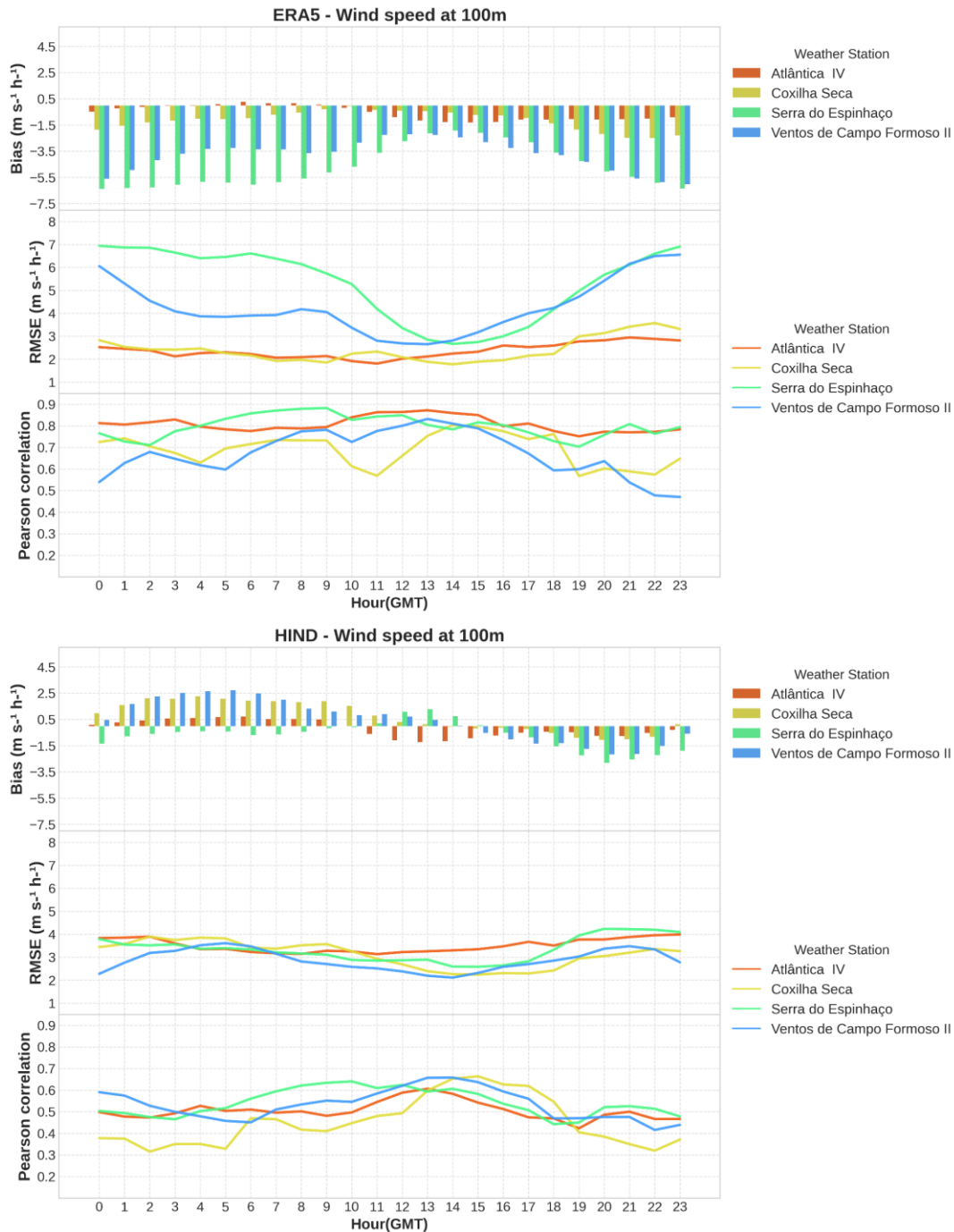


Figura 5 - Ciclo diário do desempenho dos dados do ERA5 e hindcast na altura de 100m considerando todo o período disponível, avaliando as métricas de Biais, RMSE e Correlação.

As Tabelas 2 e 3 apresentam valores estatísticos das métricas de Biais, RMSE e Correlação em todo período de análise para velocidade do vento em 80m e 100m, respectivamente. De acordo com as Tabelas, pode-se notar que o ERA5 apresenta erros maiores de Biais e RMSE, o que não acontece com os dados simulados do hindcast, indicando melhor desempenho do hindcast em simular os valores de velocidade do vento em 80m e 100m. Na Tabela 2 na altura de 80m, destaca-se as estações Serra do Espinhaço que reduziu o Biais -4.51 m/s (ERA5) para -0.89 m/s (hindcast), e Ventos de Campo Formoso II, onde o Biais era -3.64 m/s e passou 0.27 m/s.

Vale ressaltar que as mesmas estações apresentaram destaques na altura de 100m (Tabela 3), além disso observaram-se melhores valores de RMSE em todos os casos. Com relação aos valores de correlação, o ERA5 apresentou maiores valores em relação ao hindcast, porém com valores satisfatórios acima de 0.50.

Tabela 2 - Métricas estatísticas de Bias, RMSE e Correlação em todo período de análise para velocidade do vento em 80m.

Dados	Métricas	Estações Anemométricas			
		Coxilha Seca	Serra do Espinhaço	Atlântica IV	Ventos de Campo Formoso II
ERA5	BIAS	-1.19	-4.51	-0.43	-3.64
	RMSE	2.49	5.34	2.47	4.18
	R	0.7	0.72	0.76	0.65
HIND	BIAS	0.44	-0.89	-0.34	0.27
	RMSE	2.38	3.32	2.37	2.78
	R	0.52	0.60	0.51	0.55

Tabela 3 - Métricas estatísticas de Bias, RMSE e Correlação em todo período de análise para velocidade do vento em 100m.

Dados	Métricas	Estações Anemométricas			
		Coxilha Seca	Serra do Espinhaço	Atlântica IV	Ventos de Campo Formoso II
ERA5	BIAS	-1.2	-4.63	-0.43	-3.73
	RMSE	2.55	5.49	2.55	4.29
	R	0.7	0.72	0.76	0.65
HIND	BIAS	0.68	-0.71	-0.19	0.40
	RMSE	2.16	3.37	2.52	2.90
	R	0.55	0.60	0.50	0.52

#### 4. CONCLUSÕES

Neste estudo, foi avaliado o desempenho do hindcast desenvolvido pela Climatempo utilizando o modelo WRF4.3.3 na simulação de dados de velocidade do vento nas alturas de 80m e 100m para a utilização em avaliações de recursos eólicos em áreas onde não existem medições de superfície. O hindcast foi desenvolvido porque apresenta resolução espacial e temporal admissível para realizar essas análises em diferentes locais. O desenvolvimento desse tipo de dado vem sendo realizado em diferentes lugares no mundo e com diferentes aplicações, porém no Brasil dados com esse tipo de resolução ainda são inexplorados.

Para enfrentar o desafio de preencher grandes lacunas espaciais e temporais de dados, apresentamos o problema da ausência de dados e coletamos informações de velocidade de vento nas alturas de 80m e 100m na simulação do hindcast próximos de regiões que possuíam medições de superfície para validar, uma vez que os dados do hindcast estão na forma de ponto de grade. Dessa forma, os resultados indicaram que o hindcast possui bom desempenho na representação da velocidade de vento nos níveis avaliados, principalmente quando comparamos com outras bases comumente utilizadas pela comunidade científica, como o ERA5.

De maneira específica, os resultados da avaliação do hindcast apresentaram melhores valores de erro (RMSE e BIAS) em relação às observações e ao ERA5, com melhor desempenho nas estações da Serra do Espinhaço e Ventos de Campo Formoso II na altura de 80m com RMSE de 3.32 m/s e 2.78 m/s, respectivamente, e pior desempenho apenas na estação de Atlântica IV.

## REFERÊNCIAS

- [1] FREITAS, I. G. F. et al. Evaluation of Wind and Wave Estimates from CMEMS Reanalysis for Brazil's Offshore Energy Resource Assessment. *Wind*, v. 2, n. 3, p. 586-598, 2022.
- [2] ABEEólica - Associação Brasileira de Energia Eólica. INFOVENTO. Disponível online: <https://abeeolica.org.br/energia-eolica/dados-abeeolica/> (acessado em 10 Junho 2023).
- [3] Pryor, S.C.; Barthelmie, R.J. Comparison of potential power production at on and offshore sites. *Wind. Energy Int. J. Prog. Appl. Wind. Power Convers. Technol.* 2001, 4, 173–181.
- [4] VINHOZA, Amanda; SCHAEFFER, Roberto. Brazil's offshore wind energy potential assessment based on a Spatial Multi-Criteria Decision Analysis. *Renewable and Sustainable Energy Reviews*, v. 146, p. 111185, 2021.
- [5] LISBOA, Rodrigo C.; TEIXEIRA, Paulo RF; FORTES, Conceição Juana. Numerical evaluation of wave energy potential in the south of Brazil. *Energy*, v. 121, p. 176-184, 2017.
- [6] Pérez, C.C.; Greaves, D.; Iglesias, G. A review of combined wave and offshore wind energy. *Renew. Sustain. Energy Rev.* 2015, 42, 141–153.
- [7] GALVÃO, Luiz Claudio Ribeiro et al. Análise procedimental da energia contida no recurso do vento verificada através de tecnologias consolidadas de geração eólica. In: *Brazil Windpower 2019 Conference and Exhibition*. 2019. p. 1-12.
- [8] SANTOS, J. et al. Wind speed evaluation of MERRA-2, ERA-interim and ERA-5 reanalysis data at a wind farm located in Brazil. In: *Proceedings of the ISES Solar World Congress*. 2019. p. 1-10.
- [9] DE AQUINO FERREIRA, Saulo Custodio; OLIVEIRA, Fernando Luiz Cyrino; MAÇAIRA, Paula Medina. Validation of the representativeness of wind speed time series obtained from reanalysis data for Brazilian territory. *Energy*, v. 258, p. 124746, 2022.
- [10] GRUBER, Katharina et al. Towards global validation of wind power simulations: A multi-country assessment of wind power simulation from MERRA-2 and ERA-5 reanalyses bias-corrected with the global wind atlas. *Energy*, v. 238, p. 121520, 2022.

- [11] ARAÚJO, Caio Sérgio Pereira de et al. Evaluation of air temperature estimated by ERA5-Land reanalysis using surface data in Pernambuco, Brazil. *Environmental Monitoring and Assessment*, v. 194, n. 5, p. 381, 2022.
- [12] SIEFERT, Cesar Augusto Crovador et al. Avaliação de Séries de Velocidade do Vento de Produtos de Reanálises Climáticas para o Brasil. *Revista Brasileira de Meteorologia*, v. 36, p. 689-701, 2022.
- [13] CHEN, Biyan et al. A global assessment of precipitable water vapor derived from GNSS zenith tropospheric delays with ERA5, NCEP FNL, and NCEP GFS products. *Earth and Space Science*, v. 8, n. 8, p. e2021EA001796, 2021.
- [14] HERBACH, H.; DEE, D. J. E. N. ERA5 reanalysis is in production. *ECMWF newsletter*, v. 147, n. 7, p. 5-6, 2016.
- [15] MANWELL, J. F. et al. *Wind energy explained*. New York: Wiley, 2002, 569 p.
- [16] Dados Abertos - Restrição de Operação por Constrained-off de Usinas Eólicas - Detalhamento por Usina. Disponível online: <https://dados.ons.org.br/> (acessado em 02 Junho 2023).
- [17] SKAMAROCK, William C. et al. A description of the advanced research WRF model version 4. National Center for Atmospheric Research: Boulder, CO, USA, v. 145, n. 145, p. 550, 2019.
- [18] WILKS, Daniel S. *Statistical methods in the atmospheric sciences*. Academic press, 2011.