forecasting

MDPI

Article

Forecasting Wind Speed Using Climate Variables

Rafael Araujo Couto, Paula Medina Maçaira Louro * and Fernando Luiz Cyrino Oliveira

Department of Industrial Engineering, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro 22541-041, Brazil; rafael.arcouto@puc-rio.br (R.A.C.); cyrino@puc-rio.br (F.L.C.O.)

* Correspondence: paulamacaira@puc-rio.br

Abstract: Wind energy in Brazil has been steadily growing, influenced significantly by climate change. To enhance wind energy generation, it is essential to incorporate external climatic variables into wind speed modeling to reduce uncertainties. Periodic Autoregressive Models with Exogenous Variables (PARX), which include the exogenous variable ENSO, are effective for this purpose. This study modeled wind speed series in Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Rio Grande do Sul, and Santa Catarina, considering the spatial correlation between these states through PARX-Cov modeling. Additionally, the correlation with ENSO indicators was used for out-of-sample prediction of climatic variables, aiding in wind speed scenario simulation. The proposed PARX and PARX-Cov models outperformed the current model used in the Brazilian electric sector for simulating future wind speed series. Specifically, the PARX-Cov model with the Cumulative ONI index is most suitable for Pernambuco, Rio Grande do Sul, and Santa Catarina, while the PARX-Cov with the SOI index is more appropriate for Rio Grande do Norte. For Alagoas and Sergipe, the PARX with the Cumulative ONI index is the best fit, and the PARX with the Cumulative Niño 4 index is most suitable for Paraíba.

Keywords: wind speed; PARX; covariance; ENSO

check for updates

Academic Editors: Sonia Leva and Mohamed Louzazni

Received: 29 January 2025
Revised: 24 February 2025
Accepted: 4 March 2025
Published: 11 March 2025

Citation: Couto, R.A.; Maçaira Louro, P.M.; Cyrino Oliveira, F.L. Forecasting Wind Speed Using Climate Variables. *Forecasting* **2025**, *7*, 13. https://doi.org/10.3390/forecast7010013

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Electricity generation in Brazil is predominantly renewable, with more than 80% of the total generation capacity coming from renewable sources, primarily hydroelectric power, which constitutes over 65% of the country's energy matrix [1]. However, during droughts, which can severely impact water reservoirs, thermal power plants are needed to compensate for the shortfall, operating continuously at maximum capacity [2]. To sustainably address this challenge, it is crucial to diversify into other renewable sources, such as wind energy, which can complement hydroelectric power [3,4].

Accurate wind speed modeling and forecasting are crucial for effectively planning, operating, and monitoring electrical systems, especially in a complex grid like Brazil's. Pinson [5] highlights the importance of addressing the stochastic nature of renewable energy generation to enhance the robustness and reliability of energy systems. Effective stochastic modeling is vital for informed decision-making in both public and private sectors [6].

Currently, the Brazilian electricity sector employs a Periodic Autoregressive (PAR) model [7] to generate synthetic wind speed series correlated with hydroelectric reservoir inflows, based on the work of Maceira et al. [3]. While this model provides a foundation, it primarily considers wind speed in isolation and assumes that wind series are stationary, linear, and follow a normal distribution. Furthermore, it does not incorporate exogenous variables that could influence wind regimes and energy production [7].

To improve the accuracy of wind energy forecasts, it is essential to integrate current climate variables, as wind energy generation is significantly affected by climate conditions, which can impact the availability and production of wind resources in Brazil [8]. Incorporating climate variables into wind speed modeling can reduce forecasting uncertainties [9,10]. While common climate variables include pressure, temperature, and precipitation, the El Niño-Southern Oscillation (ENSO) phenomenon also shows a strong relationship with wind speed [11,12]. Notably, ENSO's influence on wind speed has been documented in Brazil and globally [13–15], supporting its use in forecasting models.

Moreover, Maçaira et al. [16] have systematically reviewed various forecasting techniques, identifying regression models, neural networks, AutoRegressive Integrated Moving Average with Exogenous Variables (ARIMAX), support vector machines (SVMs), and structural models as commonly used in studies incorporating exogenous variables. A more recent study by Pessanha et al. [17] explored dynamic models combined with Bayesian approaches for generating wind speed time series, further advancing the field of wind speed forecasting.

This study introduces an advanced forecasting approach by extending the existing Periodic Autoregressive (PAR) model to incorporate Periodic Autoregressive models with Exogenous Variables (PARX) [18,19]. This innovative framework builds upon previous successful applications, such as Maçaira works on streamflow forecasting that uses a PAR that considers exogenous variables [20,21] and ARX models used on the wind and electricity [22,23]. By integrating climate variables like ENSO into the PARX framework, the proposed approach significantly enhances the modeling and forecasting capabilities for wind speed, offering substantial improvements over current methods used in the Brazilian energy sector.

Furthermore, to enhance the understanding of wind speed patterns, variability, and trends, this study considers the spatial covariance between wind speeds across different states in Brazil. Previous research, such as that by Duran et al. [22], demonstrated that aggregating forecasts from multiple wind farms can improve prediction accuracy. Additionally, Iung et al. [24] conducted a comprehensive literature review highlighting various methods to quantify temporal dependence in renewable energy modeling, underscoring the importance of advanced forecasting techniques.

The primary objective of this research is to develop a methodology for forecasting wind speed, aiming to improve the accuracy of wind speed predictions and, consequently, wind power. Specifically, the study seeks to achieve the following secondary objectives: (i) integrate an exogenous variable into the PAR model by employing the Periodic Autoregressive model with Exogenous Variables (PARX); (ii) consider the covariance between wind regimes across states in each Brazilian region to enhance modeling precision; (iii) account for the correlation between ENSO phenomenon indicators to facilitate out-of-sample forecasting of climatic variables; (iv) utilize these forecasts to simulate wind speed scenarios.

This work is structured into four sections. Following this introduction, Section 2 presents the applied methodology. Section 3 analyzes the results and discusses their implications. Finally, Section 4 outlines the conclusions.

2. Methodology

To summarize, the steps required to achieve the objectives of this research are outlined in the methodological framework shown in Figure 1. Each step is explained in detail throughout this section. The process is divided into three main groups: Pre-Processing, Modeling, and Post-Processing.

The first group, Pre-Processing, addresses the wind speed and ENSO datasets, along with an in-depth discussion on forecasting Climate Variables and their Extrapolation. Next,

Modeling presents the statistical approaches explored in this study, starting with the PAR benchmark and progressing to the three proposed enhancements: PAR-Cov, PARX, and PARX-Cov. Finally, Post-Processing introduces the Performance Metrics RMSE, MAE, and R^2, followed by seven Fitting and Forecasting Windows, two fundamental subsections that lay the foundation for the next key steps. These include evaluating the proposed models through the Stochastic Simulation of Wind Speed Scenarios and Wind Speed Forecasting.

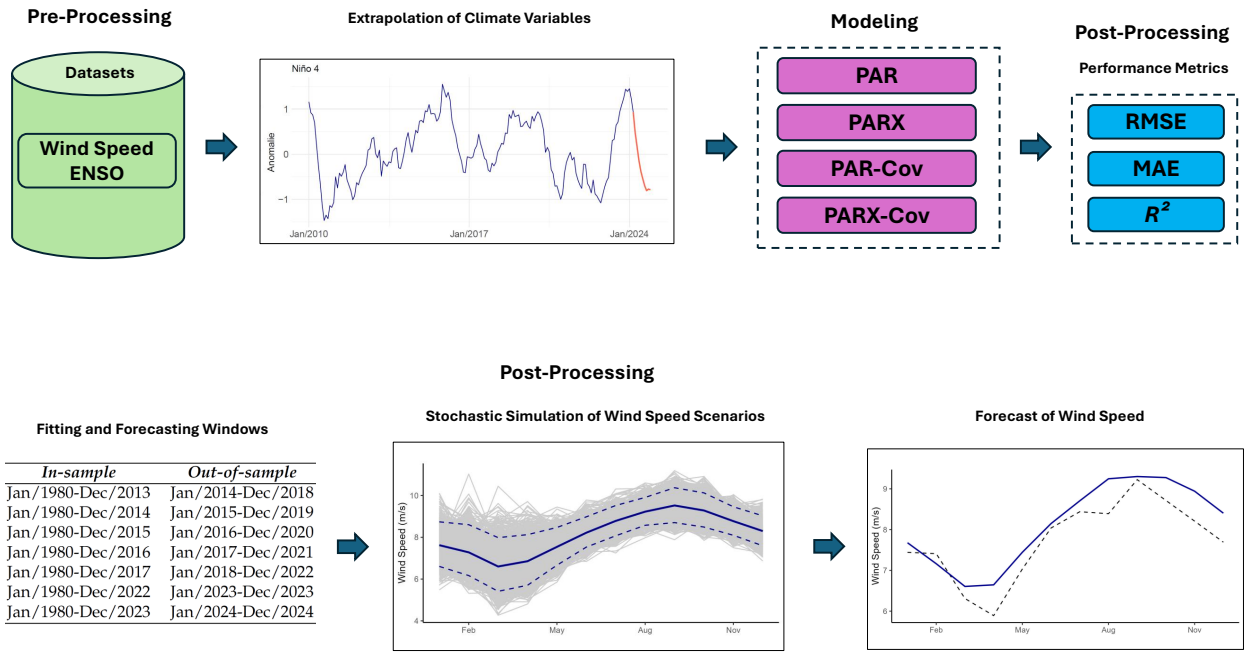


Figure 1. Methodological framework.

2.1. Pre-Processing

2.1.1. Datasets

For the analyses conducted in this study, five states from the northeastern region were selected: Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, and Sergipe, along with two states from the southern region: Rio Grande do Sul and Santa Catarina. These states feature coastal areas with high wind power generation in Brazil, as highlighted in Figures 2 and 3 in green for the northeastern states and yellow for the southern states, along with the wind potential for Brazil based on the Global Wind Atlas [25]; the redder it is, the greater the wind potential.

The MERRA-2 dataset is one of the most widely used reanalysis datasets in the literature for obtaining wind speed time series [26–28]. In this context, the data used in this study were sourced from MERRA-2. Specifically, the data for the regions under study were collected using an automated script connected to the `Renewables.ninja` website [29], covering the period from January 1980 to December 2023 [30,31].

`Renewables.ninja` provides hourly data, and the script transforms these data into monthly aggregates. The coordinates for data collection were selected based on the Global Wind Atlas [25] once more (Figure 3). In each state, three points at an altitude of 100 m above the surface were selected, which exhibited the highest wind speeds possible inside the state, not being far from each other (<100 km), to ensure that the wind regimes were similar. Subsequently, the average of the time series of these three coordinates was calculated and used to represent the historical series of each state. It is worth noting that the choice of multiple points, rather than a single one, aimed to avoid data bias and provide better representativeness by covering a larger area.

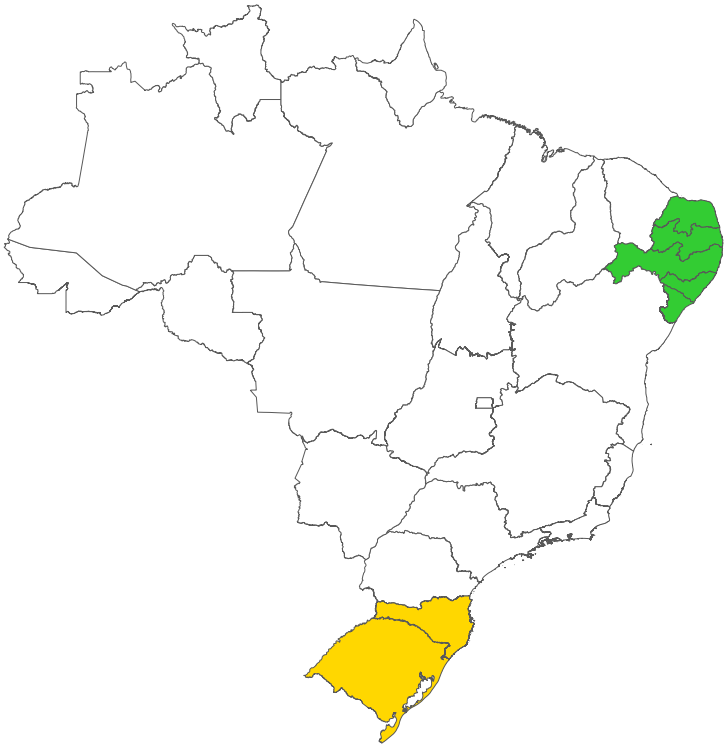


Figure 2. Selected states, with the northeast in green and the south in yellow.



Figure 3. Brazil's wind potential [25].

Data on ENSO anomalies, where El Niño refers to the warming of Pacific Ocean waters and La Niña to cooling, were divided into historical and forecast groups. Historical data were directly obtained from the Climate Prediction Center (CPC) of the National Oceanic and Atmospheric Administration (NOAA) [32,33], covering the period from 1931 to March 2024, with the initial date varying between ENSO indices. This dataset created new variables: the first variable identifies ENSO periods classified as El Niño, La Niña,

or Neutral; the others represent cumulative indices over time. This dataset is relevant for investigating trends in cumulative indices, which can indicate variations in sea pressure and temperature.

Forecast data were obtained from the International Research Institute (IRI), affiliated with NOAA [34]. The IRI provides multiple models with a forecast horizon of up to nine months for the Oceanic Niño Index (ONI). Two forecast periods were selected: the first from April 2023 to December 2023, intended to improve model fitting and prediction compared to relying solely on observed values; and the second from April 2024 to December 2024, aimed at forecasting future out-of-sample scenarios [34]. The selection of these forecast periods was guided by two main considerations. First, the study was conducted around April, making it a natural starting point. Second, since the IRI provides forecasts with a horizon of up to nine months, the period from April to December was chosen for 2023 and 2024 to maximize the forecast window within a single calendar year.

2.1.2. Extrapolation of Climate Variables

Official forecasts of the ENSO phase probabilities by the CPC are based on a consensus among their meteorologists and the IRI. They are grounded in observational and predictive information from the beginning and previous months, meaning they incorporate an analysis of various model outputs and human judgment. Models applied by the IRI generate forecasts of ENSO anomalies and are divided into dynamic and statistical groups in addition to their ensemble mean. The NOAA CPC consolidated model averages certain models [34].

In Figure 4, it can be seen that the forecasts suggest a transition from El Niño to La Niña in 2024, highlighted by the phase probabilities.

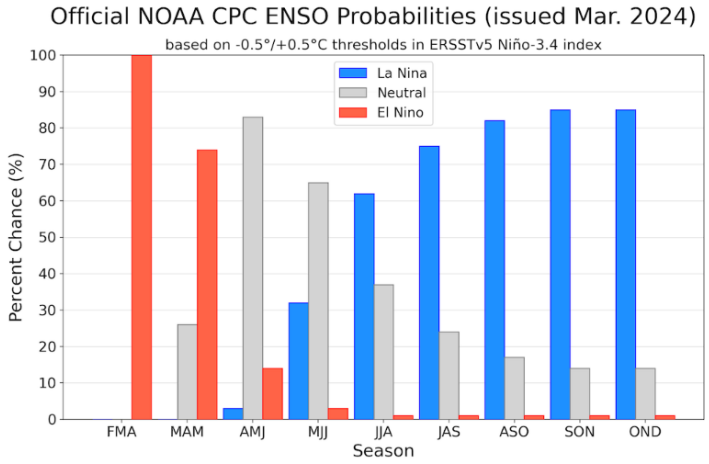


Figure 4. Forecasts of phase probabilities of ENSO [34].

The anomalies also reflect a sharp decline in the ONI index anomalies in Figure 5. The average of the anomalies from all ONI index models is also provided and was used in this work. To obtain the forecast of anomalies for other indices not provided by the IRI, a linear regression was applied [35,36].

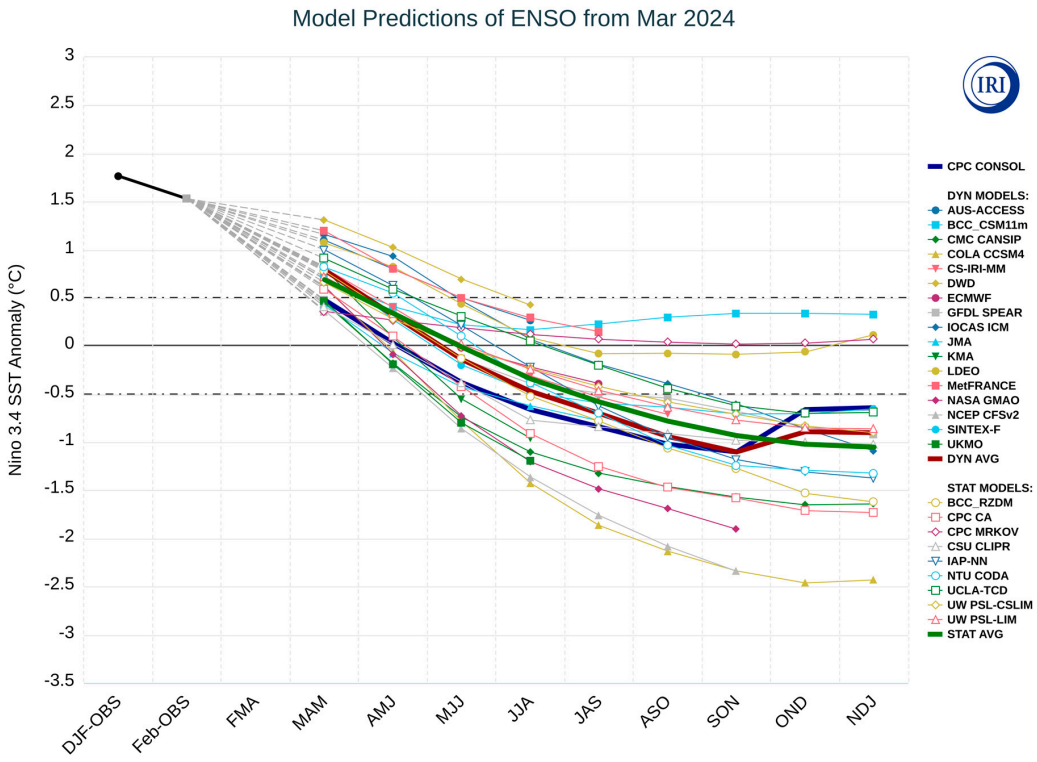


Figure 5. Forecasts of anomalies of ENSO [34].

2.2. Modeling

2.2.1. Periodic Autoregressive Model (PAR)

According to Hipel and McLeod [37], the PAR model is an approach used for modeling seasonal time series. When fitting a PAR model to a seasonal series, an individual autoregressive (AR) model is applied to each recurring period of the seasonality. For example, in a monthly seasonal series, the PAR model is configured so that each month has its own AR model, allowing for a more precise capture of specific variations within each period over time. PAR is also denoted as PAR(p), where p represents the order of the model.

Following the notation commonly used when referring to the PAR model, let Z be a series with S periods and N number of years; then, $Z = \left[z_{(1,1)}, z_{(1,2)}, \ldots, z_{(1,S)}, \ldots, z_{(N,S)}\right]$. The PAR model of series Z in period m is mathematically described by Equation (1).

$$\left(\frac{z_{(t,m)} - \mu_m}{\sigma_m}\right) = \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t,m-i)} - \mu_{m-i}}{\sigma_{m-i}}\right) + a_{t,m}, \quad (1)$$

where μ_m is the mean of period m, σ_m is the standard deviation of period m, $\varphi_i^{(m)}$ is the i-th autoregressive coefficient of period m, p_m is the order of the autoregressive operator of period m, and $a_{t,m}$ is the series of independent noises with mean 0 and standard deviation σ_m^a. In the specific case of January ($m = 1$), the model is applied to December of the previous year, i.e., to the period $(t-1, m-1)$, where it is assumed that December is represented by $m = 0 = 12$.

There are essentially two approaches to determining the autoregressive order for each month: one based on the analysis of the Periodic Autocorrelation Function (PeACF) and the Periodic Partial Autocorrelation Function (PePACF) [38], and another using information criteria, such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), which are penalization methods for model selection and comparative measures of goodness of fit, based on the number of parameters and likelihood [39].

A PeACF for period m and *lag* k can be calculated as in Equation (2).

$$\rho_k^{(m)} = \frac{\gamma_k^{(m)}}{\sigma_m \sigma_{m-k}}, \tag{2}$$

where

$$\gamma_k^{(m)} = \frac{1}{N} \sum_{i=1}^{N} (z_{(i,m)-\mu_m})(z_{(i,m-k)-\mu_{m-k}}). \tag{3}$$

To select the autoregressive order for each month, the Bayesian Information Criterion (BIC) is used, as it is effective for simpler models and selection within a group. The order with the lowest BIC is chosen for each period in the wind speed series of the PAR model [40,41].

$$BIC = \ln(n)k - 2\ln(\hat{L}). \tag{4}$$

After determining the model's order, it is necessary to estimate the parameters $\varphi_i^{(m)}$. Let $\beta_m = (\varphi_1^{(m)}, \ldots, \varphi_{p_m}^{(m)})$ be the vector of autoregressive parameters for period m. An asymptotically efficient estimator, $\hat{\beta}_m$, can be obtained by solving Equation (5) using Ordinary Least Squares (OLS) [42].

$$\gamma_l^{(m)} = \sum_{i=1}^{p_m} \hat{\varphi}_i^{(m)} \gamma_{l-i}^{(m-i)}, l = 1, \ldots, p_m. \tag{5}$$

2.2.2. Periodic Autoregressive Model with Exogenous Variables (PARX)

According to Ursu and Pereau [18] and Silveira et al. [19], the Periodic Autoregressive model with Exogenous Variables (PARX) is an extension of the PAR model. In addition to the seasonal autoregressive structure, it incorporates an additional explanatory variable denoted by X. This auxiliary variable, X, allows the model to consider and capture the effects and influences of this variable on the seasonal time series, providing a more comprehensive analysis and potentially improving forecasting capabilities by accounting for external factors that impact the seasonality of the series.

Let Z be the previously defined periodic series and X be the exogenous variable in the modeling of Z, with the same number of observations ($N \times S$) and periodicity (S) as Z. The PARX for the dependent variable Z and the exogenous variable X can be mathematically expressed as follows [18,19]:

$$\left(\frac{z_{(t,m)} - \mu_m}{\sigma_m}\right) = \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t,m-i)} - \mu_{m-i}}{\sigma_{m-i}}\right) + \sum_{j=0}^{v_m} \theta_j^{(m)} \left(\frac{x_{(t,m-j)} - \mu_{m-j}^{(x)}}{\sigma_{m-j}^{(x)}}\right) + a_{t,m}, \tag{6}$$

where μ_m is the mean of the dependent variable Z for the period m, σ_m is the standard deviation of the dependent variable Z for the period m, $\varphi_i^{(m)}$ is the i-th autoregressive coefficient of the dependent variable Z for period m, and p_m is the order of the autoregressive operator of the dependent variable Z for period m. $\mu_m^{(x)}$ is the mean of the independent variable X for period m, $\sigma_m^{(x)}$ is the standard deviation of X for period m, $\theta_j^{(m)}$ is the j-th autoregressive coefficient of the exogenous variable X for period m, v_m is the order of the autoregressive operator of the exogenous variable X for the period m, and $a_{t,m}$ is the series of independent noises with mean 0 and standard deviation σ_m^a. In the particular case of January ($m = 1$), a similar approach to the PAR model is applied. The model utilizes December of the previous year, referring to the instant $(t-1, m-1)$, where it is considered that the period of December is represented by $m = 0 = 12$.

The BIC criterion is employed again to determine the autoregressive orders of the dependent variable and the exogenous variable for each period (p_m, v_m). In the context of the PARX model, the obtained BIC is associated with the set (p_m, v_m) for each period. In

other words, the set of parameters that results in the lowest BIC value is selected as the most suitable for the model [40,41].

The parameter estimation for the model, similar to the PAR model, is performed via OLS [18]. Let $Y_{ns+m} = \left(\frac{z_{(t,m)}-\mu_m}{\sigma_m}\right)$ and $X_{ns+m} = \left(\frac{x_{(t,m-j)}-\mu^{(x)}_{m-j}}{\sigma^{(x)}_{m-j}}\right)$, where $n = 0,\ldots,N-1$ and $m = 1,\ldots,s$, with size Ns. Let $\mathbf{w}_m = [Y_m, Y_{m+s},\ldots,Y_{(N-1)s+m}]$ and $\mathbf{a}_m = [a_m, a_{m+s},\ldots,a_{(N-1)s+m}]$ be vectors of dimension $(N\times 1)$, with T being the transpose operator, and $\mathbf{W}_m = [\mathbf{Y}_m, \mathbf{X}_m]$ the matrix with dimension $N\times(p_m+1+v_m)$, where $\mathbf{Y}_m$ and $\mathbf{X}_m$ are described by

$$\mathbf{Y}_m = \begin{bmatrix} Y_{m-1} & Y_{m-2} & \cdots & Y_{m-p_m} \\ Y_{s+m-1} & Y_{s+m-2} & \cdots & Y_{s+m-p_m} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{(N-1)s+m-1} & Y_{(N-1)s+m-2} & \cdots & Y_{(N-1)s+m-p_m} \end{bmatrix}; \tag{7}$$

$$\mathbf{X}_m = \begin{bmatrix} X_m & X_{m-1} & \cdots & X_{m-v_m} \\ X_{s+m} & X_{s+m-1} & \cdots & X_{s+m-v_m} \\ \vdots & \vdots & \ddots & \vdots \\ X_{(N-1)s+m} & X_{(N-1)s+m-1} & \cdots & X_{(N-1)s+m-v_m} \end{bmatrix}. \tag{8}$$

Let

$$\beta_m = \left(\varphi^{(m)}, \theta^{(m)}\right)^T \tag{9}$$

be the parametric vector, where

$$\varphi^{(m)} = \left(\varphi_1^{(m)},\ldots,\varphi_{p_m}^{(m)}\right)^T;\ \theta^{(m)} = \left(\theta_1^{(m)},\ldots,\theta_{v_m}^{(m)}\right)^T. \tag{10}$$

Given that Equation (6) is a linear model, it can be written in the form of a regression model:

$$w_m = W_m\beta_m + a_m,\ m = 1,\ldots,s. \tag{11}$$

The covariance matrix of the random vector $\mathbf{a}_m$ is $\sigma_m^2\mathbf{I}_N$, where $\mathbf{I}_N$ is the identity matrix of size N. The ordinary least squares estimator of β_m is obtained by minimizing Equation (12).

$$S(\beta) = \sum_{m=1}^{s}\mathbf{a}_m^T\mathbf{a}_m = \sum_{n=0}^{N-1}\sum_{m=1}^{s}\left(Y_{ns+m} - \sum_{i=1}^{p_m}\varphi_i^{(m)}Y_{ns+m-i} - \sum_{i=1}^{v_m}\theta_j^{(m)}X_{ns+m-j}\right)^2. \tag{12}$$

Finally, applying the difference operator to Equation (12) yields the least squares estimators $\hat{\beta}_m = \left(\varphi^{(m)}, \theta^{(m)}\right)^T$.

$$\hat{\beta}_m = \left\{W_m^TW_m\right\}^{-1}W_m^Tw_m. \tag{13}$$

2.2.3. Covariance (PAR-Cov and PARX-Cov)

Given the modeling of the PAR and PARX models above, the correlation acts on their residuals. The concept of covariance is introduced among the states in each Brazilian region, aiming to enhance modeling accuracy and simulation. This methodology assesses the relationship between wind speeds at different points within a given space, specifically in the Brazilian states under study. Moreover, considering covariance aims to improve the understanding of wind speed data by identifying patterns, variability, and trends [43,44].

This methodology can be approached through the following steps:

1. Calculation of the covariance matrix Σ.
2. Spectral decomposition.
3. Multivariate normal distribution.

Thus, two new models are developed: the PAR-Covariance (PAR-Cov) model and the PARX-Covariance (PARX-Cov) model.

2.3. Post-Processing

2.3.1. Performance Metrics

This study employs three widely used performance evaluation metrics to assess the accuracy of wind models in Brazil [45,46] and in the world [41,47]: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), both measures of error, and the Coefficient of Determination (R^2), a measure of model fitting, given by

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - f_t)^2}; \text{MAE} = \frac{1}{T}\sum_{t=1}^{T}|y_t - f_t|; R^2 = 1 - \frac{\sum_{t=1}^{T}(y_t - f_t)^2}{\sum_{t=1}^{T}(y_t - \bar{y})^2}, \quad (14)$$

where y_t represents the observed wind speed at time t, f_t is the wind speed predicted by the model at the same time, $\bar{y}$ denotes the observed mean, and T is the forecast horizon length.

All three metrics were given equal weight, and the best models were identified as performing best across most metrics.

2.3.2. Fitting and Forecasting Windows

The proposed models were evaluated based on their forecasting accuracy, generating wind speed scenarios using the selected model. To assess the performance of each model, the dataset was divided into seven fitting and forecasting windows, as detailed in Table 1.

Table 1. Fitting and forecasting windows.

Windows	In-Sample	Out-of-Sample
1	Jan/1980–Dec/2013	Jan/2014–Dec/2018
2	Jan/1980–Dec/2014	Jan/2015–Dec/2019
3	Jan/1980–Dec/2015	Jan/2016–Dec/2020
4	Jan/1980–Dec/2016	Jan/2017–Dec/2021
5	Jan/1980–Dec/2017	Jan/2018–Dec/2022
6	Jan/1980–Dec/2022	Jan/2023–Dec/2023
7	Jan/1980–Dec/2023	Jan/2024–Dec/2024

In Window 1, the selection of parameters $p(v)$ and $m(v)$ was determined for the PARX and PARX-Cov models. For each state and each parameter combination, the following steps were executed, considering exclusively the first window:

(i) Fit the wind speed series for the in-sample period;
(ii) Simulate scenarios of the out-of-sample period (using the observed values of climatic variables from the in-sample period);
(iii) Compare the forecast values, calculated as the average of the scenarios, with the observed value;
(iv) Record the errors obtained in the validation set.

The parameters selected for each state and model resulted in the smallest error. These parameters were applied to Windows 2 to 5 to calculate performance metrics, following the procedures from Windows 1 but focusing exclusively on the best combination identified initially. The error and adjustment values presented in this section were averaged across

the validation set of the four windows (2 to 5) to more robustly evaluate the predictive capability of the models using the RMSE, MAE, and R^2 metrics. Based on these results, the best models for each state were selected.

Once the best models had been identified, they were applied to Window 6 using the ENSO forecast for 2023, obtained similarly to the process used for the 2024 period. Finally, in Window 7, forecasts for future scenarios in 2024 were made, also utilizing out-of-sample ENSO data based on the previously identified best models.

2.3.3. Stochastic Simulation of Wind Speed Scenarios

After selecting the most appropriate models, synthetic scenarios for wind speed were created through stochastic simulation. Based on the original series, the goal was to reproduce stochastic behavior and generate new time series synthetically from one of the adjusted models, PAR, PAR-Cov, PARX, and PARX-Cov. These series wee distinct from the original historical data but equally plausible from a statistical perspective.

The established strategy for generating synthetic series involves fitting a three-parameter log-normal distribution to the monthly residuals ($a_{t,m}$) of the PAR model for two main reasons [7]. The first reason is to ensure that values are always positive, given the nature of wind data. The second reason stems from the strong asymmetry in the data (and residuals), making using a normal distribution impractical.

Firstly, Equation (1) of the PAR model is manipulated to isolate Z_t:

$$z_{(t,m)} = \mu_m + \sigma_m \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t,m-i)} - \mu_{m-i}}{\sigma_{m-i}} \right) + \sigma_m a_{t,m}. \tag{15}$$

Thus, to ensure that negative values of $z_{(t,m)}$ are not generated, the following holds:

$$a_{t,m} > -\frac{\mu_m}{\sigma_m} - \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t,m-i)} - \mu_{m-i}}{\sigma_{m-i}} \right), \tag{16}$$

$$a_{t,m} > \Delta. \tag{17}$$

Therefore, the variable Δ is a function of only the moments (mean and variance) of the period m and the autoregressive coefficients and is given by

$$\Delta = -\frac{\mu_m}{\sigma_m} - \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t,m-i)} - \mu_{m-i}}{\sigma_{m-i}} \right). \tag{18}$$

Defining μ_m^a and σ_m^a as the mean and standard deviation, respectively, of the residual series of period m ($a_{t,m}$), we have the following:

$$\xi_{t,m} \sim N(\mu_\xi, \sigma_\xi^2), \tag{19}$$

$$a_{t,m} = e^{\xi_{t,m}} + \Delta, \tag{20}$$

$$a_{t,m} \sim \text{LNormal}(\mu_\xi, \sigma_\xi^2, \Delta). \tag{21}$$

Since these are random noises,

$$a_{t,m} = e^{W\sigma_\xi + \mu_\xi} + \Delta. \tag{22}$$

The parameters μ_ξ and σ_ξ are estimated to preserve the moments of the residuals, as per Charbeneau [48] and reproduced by Pereira et al. [49].

$$\mu_{\xi} = \log\left(\frac{\sqrt{\sigma_m^a}}{\sqrt{\theta(\theta-1)}}\right), \tag{23}$$

$$\sigma_{\xi} = \sqrt{\log(\theta)}. \tag{24}$$

By manipulating the PARX model equations similarly, it is possible to determine its own isolated Z_t and Δ, as shown in the Equations ahead.

$$z_{(t,m)}^{PARX} = \mu_m + \sigma_m \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t,m-i)} - \mu_{m-i}}{\sigma_{m-i}}\right) + \sigma_m \sum_{j=0}^{v_m} \theta_j^{(m)} \left(\frac{x_{(t,m-j)} - \mu_{m-j}^{(x)}}{\sigma_{m-j}^{(x)}}\right) + \sigma_m a_{t,m}. \tag{25}$$

$$\Delta^{PARX} = -\frac{\mu_m}{\sigma_m} - \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t,m-i)} - \mu_{m-i}}{\sigma_{m-i}}\right) - \sum_{j=0}^{v_m} \theta_j^{(m)} \left(\frac{x_{(t,m-j)} - \mu_{m-j}^{(x)}}{\sigma_{m-j}^{(x)}}\right). \tag{26}$$

Furthermore, the Δ equations for the PAR-Cov and PARX-Cov models are similar to their respective models but with the incorporation of covariance.

The log-normal simulation method has residual nonlinearity limitations, as noted by Oliveira et al. [7], who suggest using Bootstrap to address this. This study proposes superior modeling methods, keeping the current scenario generation technique to ensure improvements come from the proposed models.

2.3.4. Forecast of Wind Speed

Finally, it is possible to forecast from the created scenarios h steps ahead. Next, the process for the PAR methodology is presented. First, K scenarios are generated h steps ahead, which is already presented in Section 2.3.3:

$$z_{(t+h,m,k)} = \mu_m + \sigma_m \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t+h,m-i)} - \mu_{m-i}}{\sigma_{m-i}}\right) + \sigma_m a_{t+h,m}. \tag{27}$$

for $k = 1, ..., K$.

Then, the average of these scenarios is calculated to arrive at the forecast $\hat{y}$:

$$\hat{y}_{(t+h,m)} = \frac{\sum_{k=1}^{K} z_{(t+h,m,k)}}{K} \tag{28}$$

For the PARX methodology, the forecast process is also shown based on Section 2.3.3:

$$z_{(t+h,m,k)}^{PARX} = \mu_m + \sigma_m \sum_{i=1}^{p_m} \varphi_i^{(m)} \left(\frac{z_{(t+h,m-i)} - \mu_{m-i}}{\sigma_{m-i}}\right) + \sigma_m \sum_{j=0}^{v_m} \theta_j^{(m)} \left(\frac{x_{(t+h,m-j)} - \mu_{m-j}^{(x)}}{\sigma_{m-j}^{(x)}}\right) + \sigma_m a_{t+h,m}. \tag{29}$$

for $k = 1, ..., K$.

$$\hat{y}_{(t+h,m)}^{PARX} = \frac{\sum_{k=1}^{K} z_{(t+h,m,k)^{PARX}}}{K} \tag{30}$$

Moreover, the PAR-Cov and PARX-Cov forecasting methods resemble their respective models, with the addition of covariance integration.

3. Results

All the analyses and results were obtained using the R software version of December 2022 [50].

3.1. Descriptive Analysis of the Data

3.1.1. Wind Speed

Initially, after wind speed data were collected from the MERRA-2 database, as mentioned in Section 2.1.1, Figure 6 shows the monthly time series for the states of Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Rio Grande do Sul, and Santa Catarina from January 1980 to December 2023. Observing the graphs, it is noticeable that, overall, the states exhibit a well-defined seasonal behavior, with a particular emphasis on Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, and Sergipe, located in the northeast. It is important to emphasize that performing any data cleaning was unnecessary.

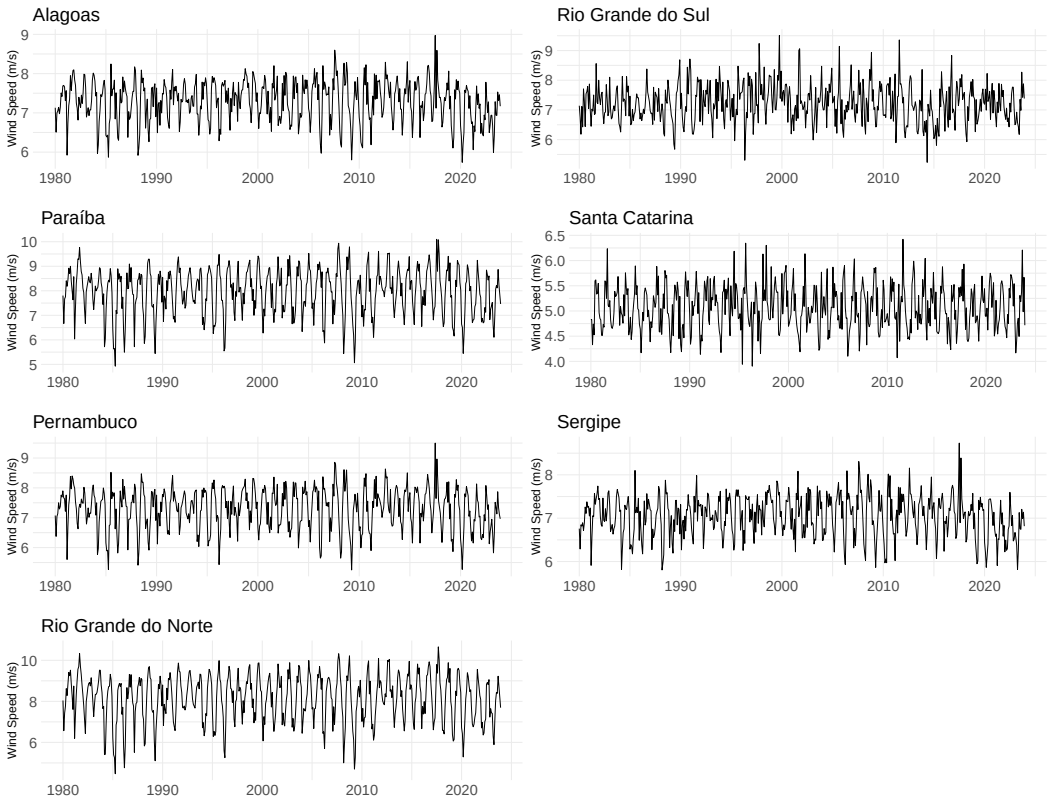


Figure 6. Wind speed time series by state.

It is also important to analyze the descriptive statistics of the time series shown in Table 2. Regarding measures of central tendency, there is a proximity between the mean and median values for the states. In this regard, the highest means are observed, as expected, in Rio Grande do Norte (8.15 m/s) and Paraíba (7.97 m/s) in the northeast, while the lowest mean is observed in Santa Catarina (5.08 m/s). Similarly, regarding standard deviation and coefficient of variation, Rio Grande do Norte and Paraíba exhibit the highest measures. As for skewness, all values fall within the interval $[-1, +1]$, typical of distributions with slight skewness. Regarding kurtosis, it is noted that the states exhibit kurtosis around three, indicating that these distributions have a frequency curve close to the normal distribution.

Table 2. Descriptive statistics of wind speed time series by state per m/s.

State	Mean	Median	Standard Deviation	Coefficient of Variation	Skewness	Kurtosis
Alagoas	7.30	7.36	0.53	0.07	−0.40	2.87
Paraíba	7.97	8.12	0.93	0.12	−0.53	2.94
Pernambuco	7.31	7.43	0.69	0.09	−0.45	2.90
Rio Grande do Norte	8.15	8.34	1.13	0.14	−0.55	2.87
Rio Grande do Sul	7.23	7.21	0.64	0.09	0.24	3.36
Santa Catarina	5.08	5.06	0.44	0.09	0.13	2.71
Sergipe	7.05	7.12	0.48	0.07	−0.21	2.94

To verify the stationarity of the series, the Augmented Dickey–Fuller (ADF) [51] and Phillips–Perron tests [52] were applied. Since all *p*-values were below a significance level of 5%, there is sufficient statistical evidence to reject the null hypothesis of non-stationarity for all states.

3.1.2. ENSO

ENSO indicators are located in various regions, each with specific significance [53]. The Southern Oscillation Index (SOI) is one of the oldest, based on the sea-level atmospheric pressure difference between Tahiti and Darwin. However, SOI is sensitive to short-term fluctuations and is limited by its location south of the Equator, while ENSO centers closer to the Equator. The Equatorial SOI addresses this by measuring pressure differences directly along the Equator between Indonesia and the Eastern Pacific.

In 1969, Bjerknes identified Sea Surface Temperature (SST) in the equatorial Pacific as a primary ENSO indicator [54]. Initially, regions like Niño 1+2, Niño 3, and Niño 4 were used for measurements. Later, Niño 3.4 was deemed the most representative [55], and its temperature anomaly is measured by the Oceanic Niño Index (ONI), which removes regional warming trends. ENSO events are identified through anomaly time series of indices, with ONI employing a three-month moving average.

For SOI, La Niña occurs with five consecutive months of positive indices above 0.5 °C, while El Niño corresponds to five straight months of negative indices below −0.5 °C. For SST, the reverse applies: El Niño corresponds to positive anomalies and La Niña to negative ones [32].

Graphs for all ENSO indices are shown in Figure 7, presenting monthly historical data from 1931 to March 2024, with the start date varying according to each ENSO index. Regarding the SOI indices, sequences above the blue line indicate La Niña events, and sequences below the red line indicate El Niño events. For SST and ONI indices, sequences of points above the red line indicate El Niño events, while sequences below the blue line indicate La Niña events.

A cumulative series can be observed in Figure 8, showing that according to the SOI and Equatorial SOI indices, there is a trend of increasing sea-level atmospheric pressure both between the regions of Tahiti and Darwin, as well as between Indonesia and the Eastern Pacific, in recent years. Also, in Figure 8, cumulative indices for SST and ONI are presented. Note that after 1980, all these indices except ONI show a downward trend.

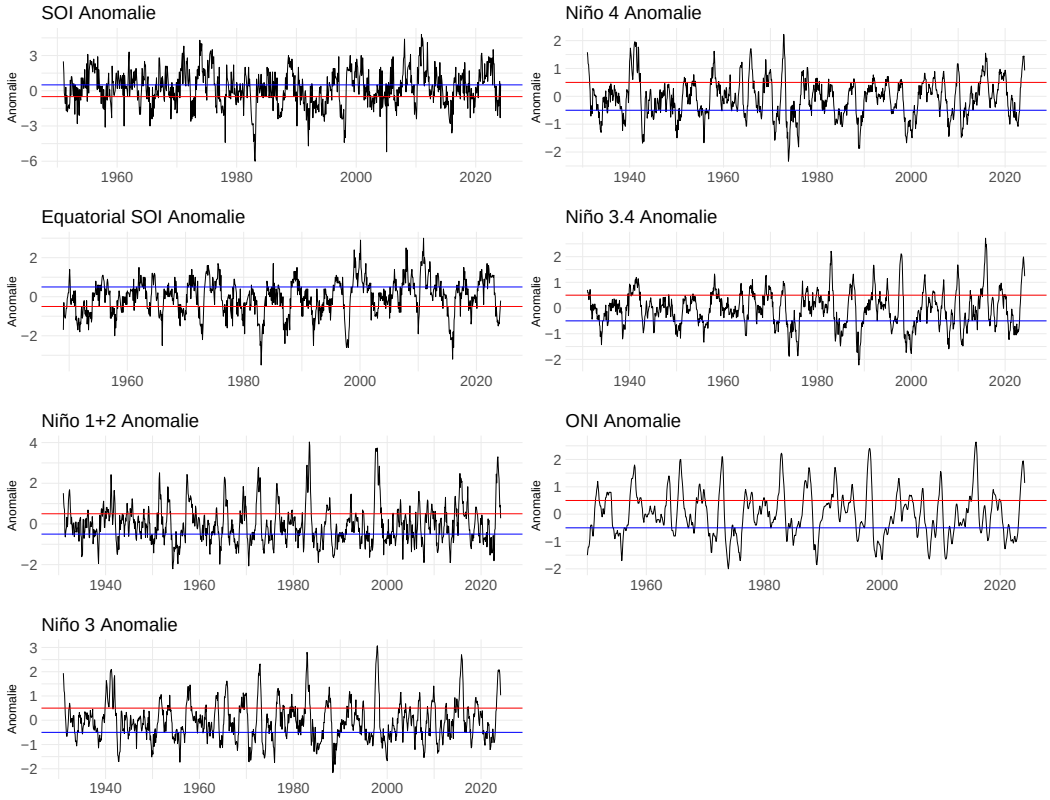


Figure 7. Time series of historical ENSO index anomalies from 1931 to March 2024. For SOI indices, sequences above the blue line indicate La Niña events, and sequences below the red line indicate El Niño events. For SST and ONI indices, the opposite is true.

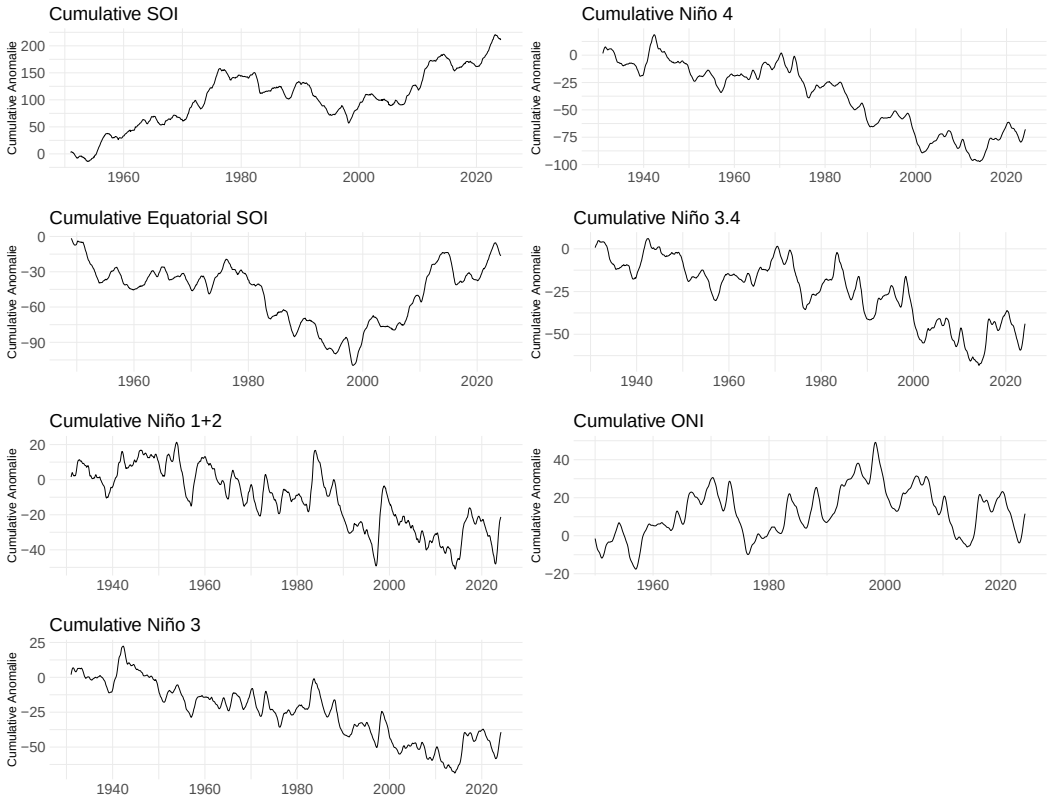


Figure 8. Time series of historical ENSO index cumulative anomalies.

3.1.3. Relationship Between Wind Speed and ENSO

One of the hypotheses of this study is that ENSO may impact wind speeds. A Kruskal–Wallis test assessed differences in wind speed distributions based on ENSO phases (El Niño, La Niña, or neutral periods). This test compares multiple independent groups using a quantitative response variable and can handle groups of different sizes. Importantly, it does not assume normality or equal variances. The tested hypotheses are as follows:

H_0: The k samples come from the same population.

H_1: At least one of the samples comes from a population different from the others.

When evaluating the test in Table 3, a significance level of 10% was chosen. Suppose the wind speed in the phases has the same distribution. In that case, the answer is "Yes", meaning that the null hypothesis is not rejected. If the phases do not have the same distribution, the answer is "No", indicating that the null hypothesis is rejected.

It is evident that all states, except Alagoas, exhibit at least one index that confirms distinct distributions between the ENSO phases. Therefore, it is statistically proven that the ENSO climatic phenomenon can influence wind speed patterns in the areas under study.

Table 3. The p-value and results of the Kruskal–Wallis test.

State	SOI	Equatorial SOI	Niño 1+2	Niño 3	Niño 4	Niño 3.4	ONI
Alagoas	0.426 Yes	0.93 Yes	0.123 Yes	0.801 Yes	0.662 Yes	0.801 Yes	0.645 Yes
Paraíba	0.262 Yes	0.918 Yes	0.013 No	0.312 Yes	0.097 No	0.197 Yes	0.788 Yes
Pernambuco	0.587 Yes	0.73 Yes	0.02 No	0.168 Yes	0.117 Yes	0.204 Yes	0.682 Yes
Rio Grande do Norte	0.065 No	0.852 Yes	0.023 No	0.787 Yes	0.193 Yes	0.26 Yes	0.609 Yes
Rio Grande do Sul	0.492 Yes	0.102 Yes	0.308 Yes	0.075 No	0.749 Yes	0.335 Yes	0.093 No
Santa Catarina	0.167 Yes	0.174 Yes	0.001 No	0.545 Yes	0.395 Yes	0.378 Yes	0.259 Yes
Sergipe	0.466 Yes	0.89 Yes	0.099 No	0.494 Yes	0.354 Yes	0.925 Yes	0.755 Yes

3.2. Forecast of ENSO Indices

The forecast of the ONI index from April 2024 to December 2024 is used to assist in predicting future wind speed scenarios, as mentioned in Section 2.1.1. This forecast is analyzed below, obtaining the estimates of anomalies for other indices of ONI, which the IRI does not provide. For this purpose, the linear regression is applied [35,36] to extrapolate the other climate variables.

Firstly, a fit was made between the observed data for each index and the ONI. The regression results can be seen in Table 4. The results show that the indices that obtained the best adjustments were Niño 3.4 and Niño 3 with R^2 values of 0.882 and 0.831, respectively. Furthermore, the sign of the coefficients reflects the relationship of each index with ONI, with the SOI and Equatorial SOI having a negative correlation. In contrast, the SST indices have a positive correlation.

Table 4. Fit of ENSO indices to ONI—April to December 2024.

Index	Coefficients	Estimated Value	Standard Deviation	*p*-Value	R^2
SOI	(Intercept) ONI	0.275 −1.349	0.036 0.043	≈0 ≈0	0.528
Equatorial SOI	(Intercept) ONI	0.001 −0.909	0.017 0.020	0.938 ≈0	0.701
Niño 1+2	(Intercept) ONI	−0.046 0.811	0.026 0.030	0.072 ≈0	0.444
Niño 3	(Intercept) ONI	−0.046 0.898	0.011 0.014	≈0 ≈0	0.831
Niño 4	(Intercept) ONI	−0.072 0.728	0.010 0.013	≈0 ≈0	0.792
Niño 3.4	(Intercept) ONI	−0.052 0.837	0.009 0.010	≈0 ≈0	0.882

From this, it is possible to construct the forecast of indices based on ONI. In Figure 9, these forecasts are shown in red, alongside their histories since 2010, shown in blue.

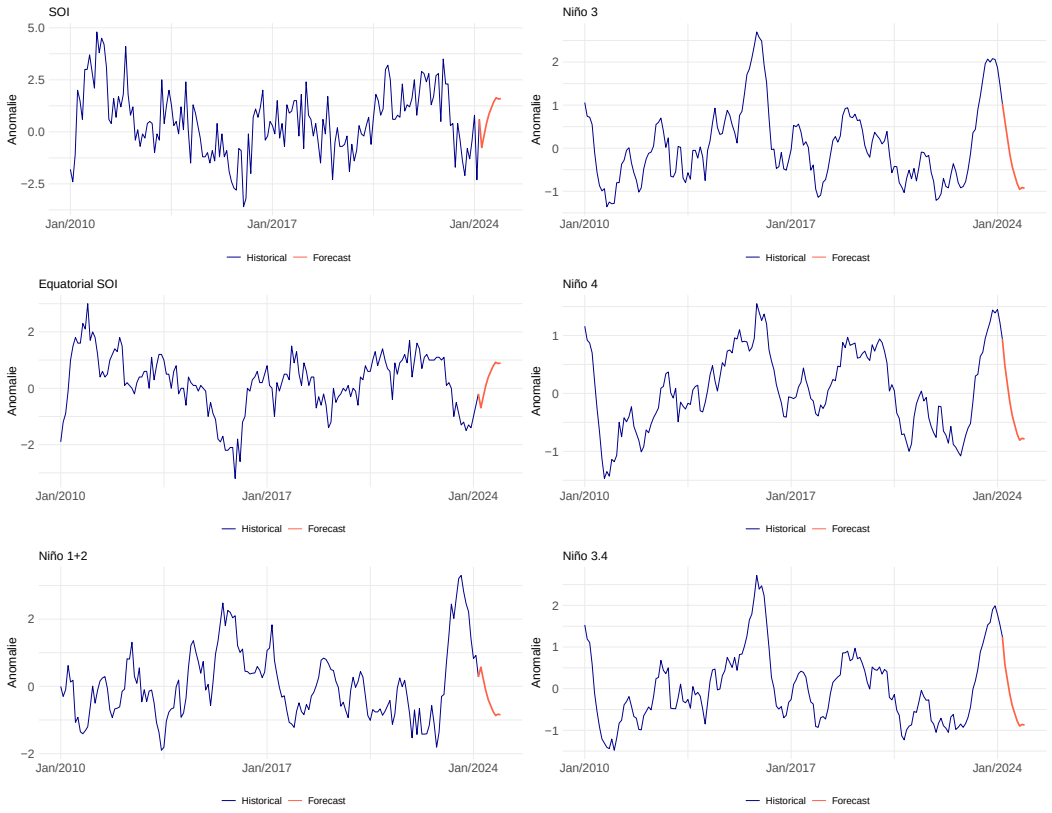


Figure 9. Forecast of ENSO indices—April to December 2024.

3.3. Wind Speed Simulation and Forecasting

The proposed models are now assessed in terms of their forecasting accuracy by generating wind speed scenarios using the selected model. The dataset was split into seven fitting and forecasting windows to evaluate each model's performance, as outlined in Table 1.

It is important to note that models with accumulated indices are referred to as "CUM" along with the name of the corresponding index to be abbreviated.

Table 5 highlights the models with the best performance according to the RMSE, MAE, and R^2 metrics, including improvements over the PAR model. It is important to note regarding the RMSE metric that the states with the largest improvements were Rio Grande do Sul (2.87%) and Santa Catarina (2.65%), using PARX-Cov models. On the other hand, the state of Pernambuco recorded the smallest improvement (0.87%), although it still showed an advantage over the PAR model. For the MAE metric, the largest improvement came, once again, from the South Region, with a 4.47% increase for Santa Catarina using a PARX-Cov model. Notably, Paraíba also showed a 2.19% improvement with a PARX model. Finally, the R^2 metric revealed a small gain for the state of Rio Grande do Norte (0.71%), while Rio Grande do Sul demonstrated a significant increase of 19.29% using a PARX-Cov model.

Table 5. Best Windows 2 to 5 models for the metrics RMSE, MAE, and R^2.

State	Metric	PAR	Best Model	Best Model	Improvement (%)
Alagoas	RMSE	0.4057	PARX + CUM ONI	0.401	1.15
	MAE	0.312	PARX + CUM ONI	0.3077	1.36
	R^2	0.476	PARX + CUM ONI	0.4878	2.48
Paraíba	RMSE	0.5013	PARX + CUM NINO4	0.4908	2.09
	MAE	0.3867	PARX + CUM ONI	0.3782	2.19
	R^2	0.7418	PARX + CUM NINO4	0.7522	1.4
Pernambuco	RMSE	0.4714	PARX + CUM ONI	0.4673	0.87
	MAE	0.3688	PARX-Cov + CUM ONI	0.3632	1.49
	R^2	0.6	PARX-Cov + CUM NINO3.4	0.6068	1.12
Rio Grande do Norte	RMSE	0.52	PARX-Cov + SOI	0.5102	1.88
	MAE	0.4064	PARX + CUM ONI	0.3996	1.69
	R^2	0.8045	PARX-Cov + SOI	0.8102	0.71
Rio Grande do Sul	RMSE	0.4888	PARX-Cov + CUM ONI	0.4748	2.87
	MAE	0.3867	PARX + CUM ONI	0.3798	1.8
	R^2	0.2263	PARX-Cov + CUM ONI	0.27	19.29
Santa Catarina	RMSE	0.3055	PARX-Cov + CUM ONI	0.2974	2.65
	MAE	0.2479	PARX-Cov + CUM ONI	0.2368	4.47
	R^2	0.429	PARX-Cov + CUM ONI	0.459	6.99
Sergipe	RMSE	0.3829	PARX + CUM ONI	0.3795	0.89
	MAE	0.2931	PARX + CUM ONI	0.2887	1.53
	R^2	0.422	PARX + CUM ONI	0.4321	2.4

Table 6 summarizes the models highlighted as the best, considering the three performance metrics for each of the seven states, resulting in 21 possible cases. All the selected models incorporate the exogenous variable ENSO in the modeling, indicating its significant contribution, with the ONI index being the most prevalent, with 16 occurrences. Additionally, it is noted that nine of the models are PARX-Cov, highlighting the importance of including covariance.

Table 6. Summary of the most selected best models.

Model	Frequency (Out of 21)
PARX + CUM ONI	10
PARX-Cov + CUM ONI	6
PARX + CUM NINO4	2
PARX-Cov + SOI	2
PARX-Cov + CUM NINO3.4	1

Considering the three performance metrics evaluated in Table 5, the best model for each state was selected based on most metrics indicating that model as the best. The exception was the state of Pernambuco, where each metric pointed to a different model as the best; in this case, the metric indicating the greatest improvement was chosen. The selected best models are presented in Table 7.

Table 7. Best model with an ENSO index selected for each state.

State	Best Model
Alagoas	PARX + CUM ONI
Paraíba	PARX + CUM NINO4
Pernambuco	PARX-Cov + CUM ONI
Rio Grande do Norte	PARX-Cov + SOI
Rio Grande do Sul	PARX-Cov + CUM ONI
Santa Catarina	PARX-Cov + CUM ONI
Sergipe	PARX + CUM ONI

Finally, in Figure 10, the observed wind speeds during the validation period of Window 6 (Jan/2023–Dec/2023) are depicted in black. Forecasts obtained using the PAR model are shown in red, while forecasts from the best PARX or PARX-Cov models for each state are highlighted in dark blue.

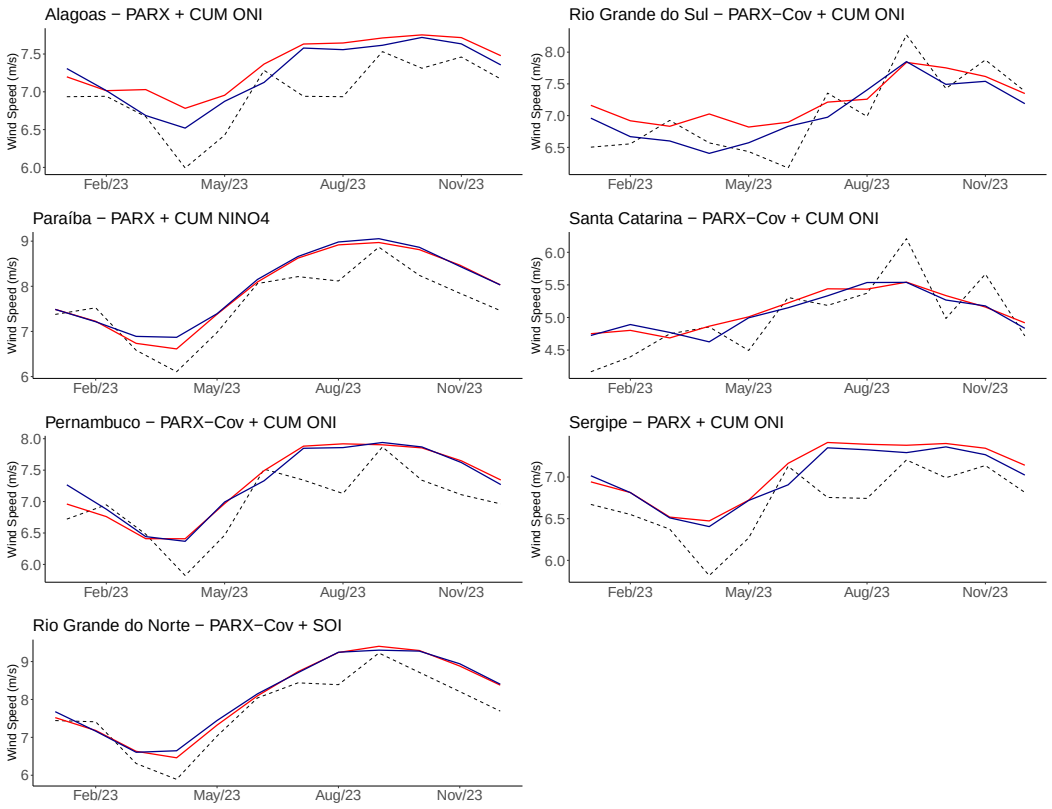


Figure 10. Observed wind speed (black) and forecasts obtained via the PAR model (red) and the best PARX or PARX-Cov model (dark blue) over Window 6.

As seen in Figure 4, it is expected that 2024 will see a transition from El Niño to La Niña, with a sharp decline in ONI index anomalies. Therefore, in Figure 11, scenarios (grey) with percentiles of 5% and 95% (dashed dark blue) and forecasts for Window 7 (Jan/2024–Dec/2024) are presented for each state. These forecasts were obtained using the

best PARX or PARX-Cov model (dark blue), which incorporates the out-of-sample ENSO climatic variable, and the PAR model (benchmark), in red.

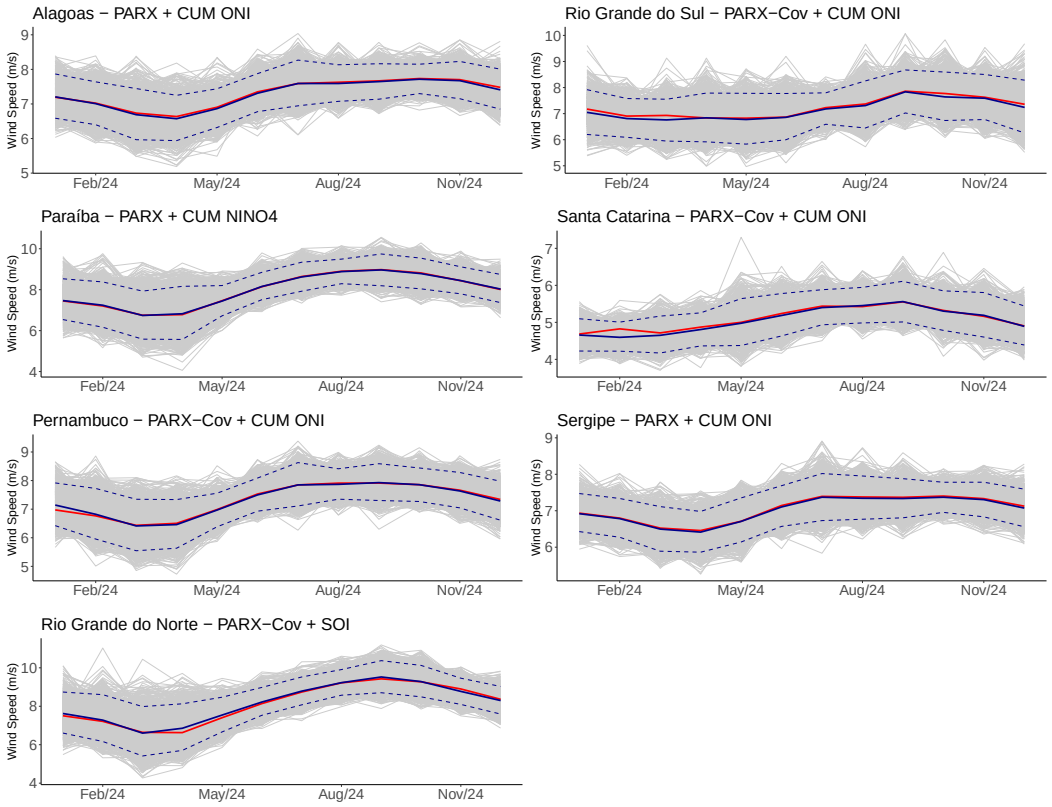


Figure 11. Scenarios (grey) with percentiles of 5% and 95% (dashed dark blue) and forecasts obtained by the best PARX or PARX-Cov model (dark blue) and the PAR model (red) over Window 7.

4. Conclusions

Due to the significant growth of wind energy in Brazil in response to current climate changes [56], this study proposes a methodological approach capable of incorporating climate variables to enhance modeling and simulation in the electricity sector, contributing to reducing uncertainties. The proposal involves adding an explanatory variable to the Periodic Autoregressive (PAR) model for wind speed series, which is currently used in the Brazilian electricity sector [3], by employing the Periodic Autoregressive model with Exogenous Variables (PARX), including the exogenous variable ENSO. This model was applied to the reanalysis data of wind speed from coastal regions with high wind power generation, covering five states in the northeast (Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, and Sergipe) and two states in the south (Rio Grande do Sul and Santa Catarina).

The introduction of covariance between these states in each Brazilian region is proposed to improve the accuracy of modeling and simulation. The spatial correlation analysis is crucial for understanding interconnections, adding dimension to the proposed methodological approach, and creating the PAR-Cov and PARX-Cov models. Additionally, ENSO ONI forecasts for nine months were collected, along with a developed approach to account for the correlation between ENSO indicators and enable the out-of-sample forecasting of other climate variables, aiming for better adjustments and predictions than using observed values. Finally, out-of-sample climate variable forecasts were used to simulate future wind speed scenarios. These scenarios were generated while ensuring that no negative values occurred.

The descriptive analysis of wind speed series identified that the states exhibit a well-marked seasonal pattern, particularly in the northeast. The highest averages were observed

in Rio Grande do Norte and Paraíba, while the lowest was in Santa Catarina. Furthermore, the ACF, ADF, and Phillips–Perron tests verified stationarity and seasonality characteristics.

Regarding the descriptive analysis of climate variables related to ENSO, historical and forecast data were examined. Among its indices, the Southern Oscillation Index (SOI) and the Equatorial SOI address sea-level atmospheric pressure, while the Sea Surface Temperature (SST) indices, including the Oceanic Niño Index (ONI), analyze sea-surface temperature. All these indices were used to quantify the El Niño and La Niña phenomena, which refer to ocean warming and cooling, respectively. Additionally, new variables were created based on the dataset representing accumulated indices over time. This set of variables was relevant as it demonstrated whether the accumulated series exhibited any trends, indicating variations in sea pressure and temperature. The ONI index forecast from April 2024 to December 2024 was used to help predict future wind speed scenarios. The forecasts suggested a transition from El Niño to La Niña in 2024, with a sharp decline in ONI index anomalies. Initially, an adjustment was made between the observed data of each index and ONI. The regression results were satisfactory, emphasizing the Niño 3.4 and Niño 3 indices. From this, forecasts for the other indices were constructed based on ONI.

Finally, this analysis also showed that the ENSO climate phenomenon can influence wind speed behavior across the studied locations, with La Niña having the most notable impact.

To obtain the final results, seven training and validation windows were used in this study. The first window was employed to determine the best combination of parameters $p(v)$ e $m(v)$ for the PARX and PARX-Cov models, where the parameters selected for each state and model were those yielding the lowest error. These parameters were applied in Windows 2 to 5 to calculate performance metrics. Based on these results, the best models for each state were selected. Once the best models were chosen, the next step was to apply them in Window 6 using the predicted ENSO values. Finally, in Window 7, a forecast of future scenarios in 2024 was performed using out-of-sample ENSO data from the best models.

Compared to the current model (PAR), the proposed models demonstrated superior performance in modeling wind speed series. Suggesting that incorporating climate variables and covariance significantly influences wind speed in the analyzed Brazilian states, directly impacting the country's energy generation. Based on the model analysis, it is concluded that the PARX-Cov with the Accumulated ONI index is the most suitable for three states: Pernambuco, Rio Grande do Sul, and Santa Catarina. In contrast, the PARX-Cov with the SOI index is more appropriate for Rio Grande do Norte. Moreover, the PARX model with the Accumulated ONI index is the best choice for the states of Alagoas and Sergipe. In contrast, the PARX model with the Accumulated Niño 4 index is the most suitable for Paraíba.

As a continuation of this study, the application of advanced statistical models is proposed to address the nonlinear aspects of wind speed time series and model the non-Gaussian nature of the data, which is often evidenced by extreme events. Expanding the study to include other northeastern states, which are also major energy producers, is also relevant. The analysis can be conducted for the northeastern subsystem or by differentiating between inland and coastal regions. Finally, wind speed is suggested to be converted into energy generation estimates. Subsequently, integrating the proposed modeling into an optimization program for electro-energy operation, such as PDDE or NEWAVE, would be an interesting next step.

Author Contributions: Conceptualization, R.A.C., P.M.M.L. and F.L.C.O.; Data curation, R.A.C.; Formal analysis, R.A.C.; Investigation, R.A.C.; Methodology, R.A.C.; Resources, P.M.M.L. and F.L.C.O.; Software, R.A.C. and P.M.M.L.; Supervision, P.M.M.L. and F.L.C.O.; Validation, P.M.M.L. and F.L.C.O.;

Visualization, R.A.C.; Writing—original draft, R.A.C.; Writing—review and editing, P.M.M.L. and F.L.C.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by Rio Paraná Energia S.A. through the RD project ANEEL PD-10381-0322/2022, CAPES Finance Code 001, CNPq (422470/2021-0, 307084/2022-1, 311519/2022-9, 402971/2023-0), and FAPERJ (210.618/2019, 211.086/2019, 211.645/2021, 201.243/2022, 201.348/2022, 210.041/2023, 210.015/2024).

Data Availability Statement: Following the principles of open science, the transparency and accessibility of the methods used are promoted by making the entire methodology available on GitHub (https://github.com/Coutin22/DissertacaoMestradoRafaelCouto.git (accessed on 7 March 2025)).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. EPE. 2021. Available online: https://www.epe.gov.br/pt/publicacoes-dadosabertos/publicacoes/balanco-energetico-nacional-ben (accessed on 10 May 2023).
2. Rigotti, J.A.; Carvalho, J.M.; Soares, L.M.; Barbosa, C.C.; Pereira, A.R.; Duarte, B.P.; Mannich, M.; Koide, S.; Bleninger, T.; Martins, J.R. Effects of hydrological drought periods on thermal stability of Brazilian reservoirs. *Water* **2023**, *15*, 2877. [CrossRef]
3. Maceira, M.; Melo, A.; Pessanha, J.; Cruz, C.; Almeida, V.; Justino, T. Wind uncertainty modeling in long-term operation planning of hydro-dominated systems. In Proceedings of the 2022 17th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Manchester, UK, 12–15 June 2022; pp. 1–6.
4. Melo, G.; Barcellos, T.; Ribeiro, R.; Couto, R.; Gusmão, B.; Oliveira, F.L.C.; Maçaira, P.; Fanzeres, B.; Souza, R.C.; Bet, O. Renewable energy sources spatio-temporal scenarios simulation under influence of climatic phenomena. *Electr. Power Syst. Res.* **2024**, *235*, 110725. [CrossRef]
5. Pinson, P. Wind energy: Forecasting challenges for its operational management. *Stat. Sci.* **2013**, *28*, 564–585. [CrossRef]
6. Ferreira, P.G.C.; Oliveira, F.L.C.; Souza, R.C. The stochastic effects on the Brazilian Electrical Sector. *Energy Econ.* **2015**, *49*, 328–335. [CrossRef]
7. Souza, R.C.; Marcato, A.; Dias, B.H.; Oliveira, F.L.C. Optimal operation of hydrothermal systems with hydrological scenario generation through bootstrap and periodic autoregressive models. *Eur. J. Oper. Res.* **2012**, *222*, 606–615. [CrossRef]
8. EPE. Plano Nacional de Energia—PNE 2050. 2018. Available online: https://static.poder360.com.br/2020/12/PNE2050.pdf (accessed on 20 May 2022).
9. do Nascimento Camelo, H.; Lucio, P.S.; Junior, J.B.V.L.; de Carvalho, P.C.M. A hybrid model based on time series models and neural network for forecasting wind speed in the Brazilian northeast region. *Sustain. Energy Technol. Assess.* **2018**, *28*, 65–72. [CrossRef]
10. de Mattos Neto, P.S.; de Oliveira, J.F.; Júnior, D.S.d.O.S.; Siqueira, H.V.; Marinho, M.H.; Madeiro, F. An adaptive hybrid system using deep learning for wind speed forecasting. *Inf. Sci.* **2021**, *581*, 495–514. [CrossRef]
11. Corrêa, C.S.; Schuch, D.A.; Queiroz, A.P.d.; Fisch, G.; Corrêa, F.d.N.; Coutinho, M.M. The long-range memory and the fractal dimension: A case study for Alcântara. *J. Aerosp. Technol. Manag.* **2017**, *9*, 461–468. [CrossRef]
12. Lima, C.N.N.; Fernandes, C.A.C.; França, G.B.; de Matos, G.G. Estimation of the El Niño/La Niña impact in the intensity of Brazilian Northeastern winds. *Anuário Inst. Geociênci.* **2014**, *37*, 232–240. [CrossRef]
13. Arpe, K.; Molavi-Arabshahi, M.; Leroy, S.A.G. Wind variability over the Caspian Sea, its impact on Caspian seawater level and link with ENSO. *Int. J. Climatol.* **2020**, *40*, 6039–6054. [CrossRef]
14. Xu, Q.; Li, Y.; Cheng, Y.; Ye, X.; Zhang, Z. Impacts of climate oscillation on offshore wind resources in China seas. *Remote Sens.* **2022**, *14*, 1879. [CrossRef]
15. Coria-Monter, E.; Salas de León, D.A.; Monreal-Gómez, M.A.; Durán-Campos, E. Satellite observations of the effect of the "Godzilla El Niño" on the Tehuantepec upwelling system in the Mexican Pacific. *Helgol. Mar. Res.* **2019**, *73*, 3. [CrossRef]
16. Maçaira, P.; Thomé, A.; Cyrino Oliveira, F.; de Almeida, F. Time series analysis with explanatory variables: A systematic literature review. *Environ. Model. Softw.* **2018**, *107*, 199–209. [CrossRef]
17. de Mendonça, M.J.C.; Pessanha, J.F.M.; de Almeida, V.A.; Medrano, L.A.T.; Hunt, J.D.; Junior, A.O.P.; Nogueira, E.C. Synthetic wind speed time series generation by dynamic factor model. *Renew. Energy* **2024**, *228*, 120591. [CrossRef]
18. Ursu, E.; Pereau, J. Estimation and identification of periodic autoregressive models with one exogenous variable. *J. Korean Stat. Soc.* **2017**, *46*, 629–640. [CrossRef]

19. Silveira, C.; Alexandre, A.; de Souza Filho, F.; Vasconcelos Junior, F.; Cabral, S. Monthly streamflow forecast for National Interconnected System (NIS) using Periodic Auto-regressive Endogenous Models (PAR) and Exogenous (PARX) with climate information. *Braz. J. Water Resour.* **2017**, *22*. [CrossRef]
20. Maçaira, P.M.; Oliveira, F.L.C.; Ferreira, P.G.C.; Almeida, F.V.N.d.; Souza, R.C. Introducing a causal PAR (p) model to evaluate the influence of climate variables in reservoir inflows: A Brazilian case. *Pesqui. Oper.* **2017**, *37*, 107–128. [CrossRef]
21. Huang, X.; Maçaira, P.M.; Hassani, H.; Oliveira, F.L.C.; Dhesi, G. Hydrological natural inflow and climate variables: Time and frequency causality analysis. *Phys. A Stat. Mech. Its Appl.* **2019**, *516*, 480–495. [CrossRef]
22. Duran, M.J.; Cros, D.; Riquelme, J. Short-term wind power forecast based on ARX models. *J. Energy Eng.* **2007**, *133*, 172–180. [CrossRef]
23. Golia, S.; Grossi, L.; Pelagatti, M. Machine learning models and intra-daily market information for the prediction of Italian electricity prices. *Forecasting* **2022**, *5*, 81–101. [CrossRef]
24. Iung, A.M.; Cyrino Oliveira, F.L.; Marcato, A.L.M. A review on modeling variable renewable energy: Complementarity and spatial—Temporal dependence. *Energies* **2023**, *16*, 1013. [CrossRef]
25. Global Wind Atlas. Available online: https://globalwindatlas.info/en (accessed on 27 November 2023).
26. de Aquino Ferreira, S.C.; Oliveira, F.L.C.; Maçaira, P.M. Validation of the representativeness of wind speed time series obtained from reanalysis data for Brazilian territory. *Energy* **2022**, *258*, 124746. [CrossRef]
27. Gruber, K.; Klöckl, C.; Regner, P.; Baumgartner, J.; Schmidt, J. Assessing the Global Wind Atlas and local measurements for bias correction of wind power generation simulated from MERRA-2 in Brazil. *Energy* **2019**, *189*, 116212. [CrossRef]
28. Olauson, J.; Bergkvist, M. Modelling the Swedish wind power production using MERRA reanalysis data. *Renew. Energy* **2015**, *76*, 717–725. [CrossRef]
29. Renewables.ninja. Available online: https://www.renewables.ninja/ (accessed on 27 November 2023).
30. Pfenninger, S.; Staffell, I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy* **2016**, *114*, 1251–1265. [CrossRef]
31. Staffell, I.; Pfenninger, S. Using bias-corrected reanalysis to simulate current and future wind power output. *Energy* **2016**, *114*, 1224–1239. [CrossRef]
32. NOAA. Available online: https://www.ncei.noaa.gov/access/monitoring/enso/ (accessed on 31 March 2024).
33. National Weather Service, Climate Prediction Center. Available online: https://www.cpc.ncep.noaa.gov/data/indices/ (accessed on 31 March 2024).
34. International Research Institute. Available online: https://iri.columbia.edu/our-expertise/climate/enso/ (accessed on 31 March 2024).
35. Barhmi, S.; Elfatni, O.; Belhaj, I. Forecasting of wind speed using multiple linear regression and artificial neural networks. *Energy Syst.* **2020**, *11*, 935–946. [CrossRef]
36. James, G.; Witten, D.; Hastie, T.; Tibshirani, R.; Taylor, J. Linear regression. In *An Introduction to Statistical Learning: With Applications in Python*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 69–134.
37. Hipel, K.; McLeod, A. *Time Series Modelling of Water Resources and Environmental Systems*; Elsevier: Amsterdam, The Netherlands, 1994.
38. McLeod, A. Diagnostic Checking Periodic Autoregression Models with Application. *J. Time Ser. Anal.* **1995**, *15*, 221–233. [CrossRef]
39. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
40. Vrieze, S. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol. Methods* **2012**, *17*, 228–243. [CrossRef]
41. Makubyane, K.; Maposa, D. Forecasting short-and long-term wind speed in Limpopo province using machine learning and Extreme Value Theory. *Forecasting* **2024**, *6*, 885–907. [CrossRef]
42. Dismuke, C.; Lindrooth, R. Ordinary least squares. *Methods Des. Outcomes Res.* **2006**, *93*, 93–104.
43. Suryawanshi, A.; Ghosh, D. Wind speed prediction using spatio-temporal covariance. *Nat. Hazards* **2015**, *75*, 1435–1449. [CrossRef]
44. Ezzat, A.A.; Jun, M.; Ding, Y. Spatio-temporal short-term wind forecast: A calibrated regime-switching method. *Ann. Appl. Stat.* **2019**, *13*, 1484. [CrossRef]
45. Jacondino, W.D.; da Silva Nascimento, A.L.; Calvetti, L.; Fisch, G.; Beneti, C.A.A.; da Paz, S.R. Hourly day-ahead wind power forecasting at two wind farms in northeast Brazil using WRF model. *Energy* **2021**, *230*, 120841. [CrossRef]
46. de Souza, N.B.P.; Nascimento, E.G.S.; Santos, A.A.B.; Moreira, D.M. Wind mapping using the mesoscale WRF model in a tropical region of Brazil. *Energy* **2022**, *240*, 122491. [CrossRef]
47. Mugware, F.W.; Sigauke, C.; Ravele, T. Evaluating wind speed forecasting models: A comparative study of CNN, DAN2, Random Forest and XGBOOST in diverse South African weather conditions. *Forecasting* **2024**, *6*, 672–699. [CrossRef]

48. Charbeneau, R. Comparison of the two- and three-parameter log normal distributions used in streamflow synthesis. *Water Resour. Res.* **1978**, *14*, 149–150. [CrossRef]
49. Pereira, M.; Oliveira, G.; Costa, C.; Kelman, J. Stochastic streamflow models for hydroeletric systems. *Water Resour. Res.* **1984**, *20*, 379–390. [CrossRef]
50. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022.
51. Paparoditis, E.; Politis, D.N. The asymptotic size and power of the augmented Dickey–Fuller test for a unit root. *Econom. Rev.* **2018**, *37*, 955–973. [CrossRef]
52. Escobari, D.; Garcia, S.; Mellado, C. Identifying bubbles in Latin American equity markets: Phillips-Perron-based tests and linkages. *Emerg. Mark. Rev.* **2017**, *33*, 90–101. [CrossRef]
53. Barnston, A. Why Are There So Many ENSO Indexes, Instead of Just One? Available online: https://www.climate.gov/news-features/blogs/enso/why-are-there-so-many-enso-indexes-instead-just-one (accessed on 27 November 2023).
54. Bjerknes, J. Atmospheric teleconnections from the equatorial Pacific. *J. Phys. Oceanogr.* **1969**, *97*, 163–172.
55. Barnston, A.; Chelliah, M.; Goldenberg, S. Documentation of a highly ENSO-related SST region in the equatorial Pacific. *Atmosphere-Ocean* **1997**, *35*, 367–383.
56. IRENA. 2021. Available online: https://www.irena.org/newsroom/pressreleases/2021/Apr/World-Adds-Record-New-Renewable-Energy-Capacity-in-2020 (accessed on 10 May 2023).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.