Time Series Analysis for Offshore Wind Energy Forecasting: Hybrid Model SARIMA and Genetic Algorithms.

Fabrício de A. Santos * **Nerval de J. S. Junior** *
Hellen. D. P. de Souza * **Shigeaki L. de Lima** *

* *Institute of Electrical Energy, Federal University of Maranhão, MA, (fabricio.as, nerval.junior, hellen.souza)@discente.ufma.br, shigeaki.lima@ufma.br*

Abstract: The conduct of studies on regional wind potential is essential to identify new projects and improve measurements, aiming for the efficient installation of turbines. The challenge of predicting wind behavior for energy extraction is of great interest, given the importance of accurate forecasts for the success of the sector. In this context, this article proposes a hybrid forecasting model, SARIMA and Genetic Algorithms, aiming to improve the accuracy of wind predictions using a wind measurement database, converting them into power generation scenarios. The SARIMA (5, 1, 5)x(6, 0, 6) model estimated by the algorithm was the best solution, demonstrating better fitness in tests. Its performance in MSE and MAE was 0.864 and 0.616, respectively.

Keywords: Time series; Forecast; SARIMA; Genetic Algorithms; Energy.

1. INTRODUCTION

The encouragement of wind energy production in Brazil began with government support through the creation of the Program for the Incentive of Alternative Sources of Electric Energy (Proinfa), established by Aneel Resolution No. 10438 of 2002. By 2009, a total of 14,324 MW of wind capacity had been contracted through auctions held under Proinfa. In 2024, the installed capacity of wind energy in Brazil reached 26.8 GW (ABEEólica, 2023; Moreira, 2017).

Brazil's wind generation potential is estimated at over 1500 GW, considering both onshore and offshore plants, according to surveys by Abeeólica. Despite the predominance of hydroelectric plants in Brazil's energy matrix, the share of renewable energies, especially wind, has been increasing, representing about 14% of the country's total generation (ABEEólica, 2023; EPE, 2023, 2020; Wang et al., 2021a).

Innovations in more efficient turbines, energy storage systems, and more accurate wind prediction methods are key areas that can further boost the Brazilian wind sector. At the same time, it is important to promote public policies that encourage investment in renewable energy and reduce barriers to entry for new players in the market (Moreira, 2017; Wang et al., 2021b; Liu et al., 2021).

In this context, with the continuous growth of the wind energy sector and the conduct of studies on the wind potential of different regions of Brazil becoming essential, aiming at the identification of new wind projects. Research focused on technological development in wind energy generation has been concentrated on improving measurements of regional wind potential. This allows for better selection of turbine installation sites, cost reduction, and increased efficiency (Moreira, 2017; EPE, 2020; Liu et al., 2020).

The challenge of predicting wind behavior for energy extraction has sparked the interest of researchers and companies, as accurate predictions are essential for the sector's success (Souza et al., 2022).

Recognizing the need to develop a forecasting model capable of meeting the sector's demands, this article presents a model developed based on regression from the SARIMA model and genetic algorithms. The main purpose is to make wind predictions using a database derived from atmospheric measurements, enabling the conversion of this data into energy generation scenarios.

2. TIME SERIES ANALYSIS

Regression is a crucial statistical technique for understanding and modeling the relationship between variables. When applied to time series data, regression plays a crucial role in forecasting patterns and trends over time. Time series analysis uses stochastic processes to model and predict events (Morettin, 2006; Mitsa, 2010).

This approach aims to understand the behavior of phenomena that evolve over time, such as trends, cycles, and seasonal variations (Morettin, 2006; Mitsa, 2010). Time series modeling seeks to predict future values of events by identifying significant patterns and behaviors.

The main objectives include understanding the underlying mechanism, predicting future values, and identifying components of the series (Morettin, 2006). The Auto-Regressive Integrated Moving Average (ARIMA) model, which combines autoregression, moving averages, and differencing to capture temporal patterns, is represented by:

$$Z_t = X_t + S_t + E_t \tag{1}$$

where Z_t is the observed time series, X_t is the trend, S_t is the seasonality, and E_t is the random component (Gujarati, 2011; Wang et al., 2021a; Yan et al., 2022).

The ARIMA(p, d, q) model is a procedure used in time series analysis for forecasting. Its characteristics include being a univariate model, working with only one variable across multiple equally spaced time units, and utilizing a transfer function with variables equally spaced for improved performance (Gujarati, 2011; BUENO, 2012).

Unlike the ARIMA model, the ARMA model does not involve differencing the series to make it stationary. Instead, it requires differencing the series d times until it becomes stationary, as described in (Gujarati, 2011; BUENO, 2012):

$$z_t = y_t - y_{t-1} = y'_t \tag{2}$$

Incorporating this characteristic into the ARIMA model yields:

$$y' = c + \phi_0 y'_{t-1} + \ldots + \phi_p y'_{t-p} + \theta_0 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{3}$$

The SARIMA(p, d, q)(P, D, Q) model, Seasonal Auto-Regressive Integrated Moving Average, contains all the characteristics of ARIMA, being an extension that takes into account the stochastic seasonality of the data (Araújo, 2021).

Hence, it requires a larger number of parameters, namely P representing the number of seasonal autoregressive coefficients depending on seasonality, D representing the number of seasonal differences needed to stationarize the series according to the number of periods, and Q representing the number of seasonal moving average coefficients (Araújo, 2021). Its structure is defined by: a non-seasonal part (ARIMA) and a seasonal part.

(1) Non-Seasonal Part (ARIMA)

- AR (AutoRegressive): This term represents the regression of the dependent variable against its own past observations, i.e., the time series is regressed against its own lags (previous values). The order of the AR term is denoted by p.
- I (Integrated): The integration part refers to the number of differentiations needed to make the time series stationary, i.e., to remove trends and seasonality. The order of differencing is denoted by d.
- MA (Moving Average): This term represents the moving average of the forecast errors of the time series, i.e., the regression of the time series against the residual errors of its lags. The order of the MA term is denoted by q.

(2) Seasonal Part (S)

- AR seasonal: Denoted by P, it refers to the order of the seasonal autoregressive component.
- I seasonal: Denoted by D, it refers to the order of seasonal differencing.
- MA seasonal: Denoted by Q, it refers to the order of the seasonal moving average component.
- Seasonal period: Denoted by s, it represents the periodicity of seasonal patterns in the time series.

Therefore, before applying the SARIMA model, it is essential to check the seasonality of the time series. Seasonality implies that the statistical properties of the series, such as mean and variance, remain constant over time. A common test used to check for seasonality is the Dickey-Fuller test, allowing to evaluate if the p-value of the test is less than the significance level (usually 0.05) (Shumway et al., 2000).

2.1 Augmented Dickey-Fuller Test

To determine if the time series is stationary, tests are performed using the Augmented Dickey-Fuller (ADF) test. With this test, it is possible to obtain the ADF variables, p-value, number of lags used, and observations, which aim to verify if there is a unit root, which is the factor that will determine if it is stationary or not (Bastos, 2024).

Finally, the performance metrics of the ARIMA model are presented, including *Root Mean Square Error (RMSE)*, *Mean Absolute Error (MAE)*, and *Mean Absolute Percentage Error (MAPE)*, as discussed earlier.

2.2 Performance Evaluation Metrics

Finally, the performance metrics of the ARIMA model are presented, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), as discussed earlier.

$$MAE = \frac{1}{n-m} \sum_{t=m+1}^{n} |y_t - \hat{y}_t| \tag{4}$$

$$MAPE = \frac{1}{n-m} \sum_{t=m+1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100 \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{6}$$

in which:

- n is the total number of observations.
- y_i are the observed values.
- $\hat{y}_i$ are the values predicted by the model.

The RMSE is a measure of the dispersion of residual values and provides an indication of the overall accuracy of the model, where the *mean_squared_error* library was used [1].

3. GENETIC ALGORITHMS

The Genetic Algorithms are methodologies aimed at heuristic search that can be applied to a wide range of problems, such as numerical function optimization, combinatorial optimization, and resource allocation. Moreover, they are stochastic algorithms, meaning they use randomness to minimize their functions. The fundamental principle of Genetic Algorithms is evolution, where the survival

[1] `https://scikit-learn.org/stable/`

of the fittest is the objective (Kramer, 2017; Lambora et al., 2019).

Using this principle, the genetic algorithm is constituted, where the individual is the solution containing a chromosome in vector format. In this vector, we have N genes, which will be allocated values that, from the assignment, will be the alleles. There will be a population of individuals containing a number μ of parents and a number δ of offspring (Pacheco et al., 1999; MICHALEWICZ, 1996).

Next, genetic operators are used, such as crossover and mutation, to diversify the solutions. After this stage, a fitness testing process is initialized to select the best solution. This process is executed until one of its termination conditions is met (Pacheco et al., 1999; MICHALEWICZ, 1996). The implementation is as follows:

3.1 Initialization

The initial parent population is generated randomly, with their genes situated within a defined range according to the problem being solved. In the algorithm used, the random values have a lower limit of 0 and an upper limit of 3. These values can increase throughout the algorithm, always remaining integers and positive or equal to zero, that is, natural numbers including zero. (Pereira et al., 2001; Amirjanov and Sadikoglu, 2017).

3.2 Crossover

Crossover, or recombination, randomly selects a number of parents within the range of 2 to μ for each individual in the population. These are the parents that will pass on their genetic material, their characteristics, by combining their chromosomes. Each gene of the new individual is randomly chosen from the jth position of the parents to be copied into the offspring at position j (Kramer, 2017; Amirjanov and Sadikoglu, 2017).

3.3 Mutation

Mutation aims to generate new solutions based on the current ones, with a regular variation. When a population enters a new generation, it undergoes mutation processes to improve the diversity of solutions. There are a range of mutation operators, each with a specific purpose for each problem. The mutation used in individuals is to randomly choose one of the genes in the chromosome and add its value by a number x that can be -1 or 1, also randomly (Amirjanov and Sadikoglu, 2017).

3.4 Fitness Function

To evaluate each solution, a quality test is performed. For this, an evaluation function is needed to provide satisfaction parameters to determine how good the solution is. In the current problem, the Mean Squared Error (MSE) function will be used, applied after model fitting (MICHALEWICZ, 1996). The calculation of Mean Squared Error is performed by:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{7}$$

3.5 Selection

Just like in natural evolution methods, the genetic algorithm selects the best individuals based on the fitness function for the next generation. This process, called rank selection, returns the genetic material of the best individuals, with the best fitness value, passing them on to the subsequent generations (Gaspar-Cunha et al., 2012; Kramer, 2017; Amirjanov and Sadikoglu, 2017).

3.6 Stopping Criterion

For genetic algorithms, the stopping criterion is significant as it has the power to terminate the algorithm. The most commonly used criteria are:

- The number of generations, limiting the algorithm to a quantity of y generations to achieve the best solution within that limit.
- The fitness function value, where, upon reaching a certain value, the algorithm is stopped, aiming for the solution to be satisfactory.
- The absence of improvement after a certain number g of generations.

4. COMPUTATIONAL EXPERIMENTS

The region of the Mearim River in the state of Maranhão stands out for its geographical, economic, and environmental relevance. As one of the main rivers in the northeastern region of Brazil, it runs through a vast area within the state. Besides being crucial for the irrigation of agricultural lands, providing essential water resources for activities such as agriculture, fishing, and navigation, it can also be used by large vessels in the river sections closest to the Mearim Delta (Sousa Lima, 2013).

On another note, the river has shown significant potential for wind and hydrokinetic energy generation, prompting projects such as the National Institute of Oceanic and Fluvial Energies (INEOF) to conduct *in situ* research in the region, evaluating micrometeorological and hydrological effects on different time and space scales. Led by the Institute of Electrical Energy at the Federal University of Maranhão (IEE/UFMA), INEOF is co-funded by CNPq, CAPES, and FAPEMA.

The studies carried out by INEOF aim to contribute to the technical-scientific development in the use of oceanic and riverine energies, meeting the demands for renewable energy generation in the country. IEE has focused its studies on the regions of the Mearim River, São Marcos basin, and Arraial basin, using advanced technology to prospect potential offshore wind and hydrokinetic hotspots.

The research on energy generation forecasting from wind speed and height data, combined with genetic algorithms and SARIMAX, began with the development of the predictive model using Jupyter Colab [2], optimized for analyses, and Pandas [3], which has specific operations for time series. Additionally, the "statsmodels" library [4] was employed,

[2] https://colab.google/
[3] https://pandas.pydata.org/
[4] https://www.statsmodels.org/

allowing the use of modules that provide accuracy in modeling.

In Figure 1, the minimum, maximum, and average speed values of wind speed data at 100 meters height analyzed from the Lidar equipment are illustrated. In this context, the average wind speed at 100 meters height was used as an example.

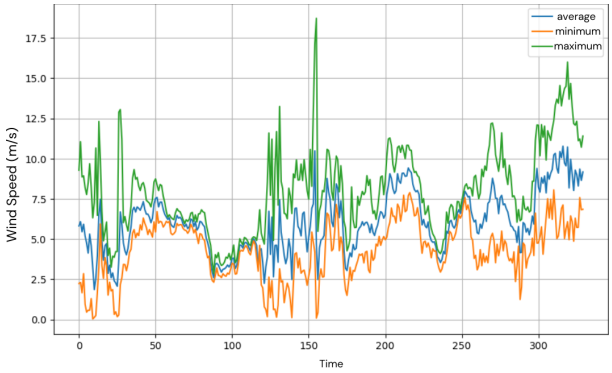

Figure 1. Minimum, maximum, and average speed at 100 m. Source: Developed by the Authors.

In Figure 2, the components Z_t, X_t, S_t, and E_t are illustrated, respectively. The graphs of the original series are presented, decomposed into trend, seasonality, and residual series using the SARIMA model. The residuals are the differences between the observed values and the values predicted by the model. This graph is used to check for patterns in the residuals, indicating whether the model adequately captured the data structure.

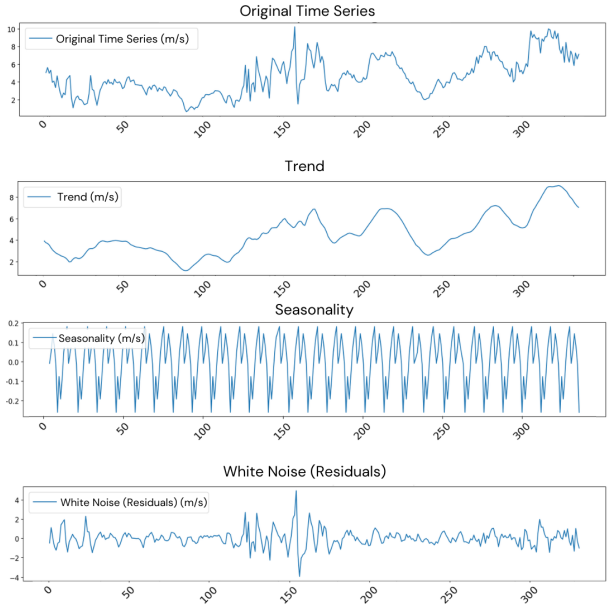

Figure 2. Time Series Decomposition. Source: Developed by the Authors.

Some relevant aspects are noted, such as the differentiation term in the genetic algorithm search being equal to 1, similar to the test performed (Figure 2), the Augmented Dickey-Fuller Test in Table 1, proving that to forecast this time series, differentiation is needed to make it stationary. Additionally, it is observed that differentiation is not necessary when it comes to seasonality.

The results of the Augmented Dickey-Fuller Test are presented in Table 1 below, obtained after differentiation:

Table 1. Augmented Dickey-Fuller Test

ADF Test Statistic	-11.786432097014963
Valor-p	1.0062434657410242e-21
Number of Lags Used	9
Number of Observations Used	318

These results indicate that the ADF test statistic is -11.7864 and the p-value is 1.006×10^{-21}. The p-value, which should be less than 0.05 to indicate stationarity of the time series, is significantly smaller, providing strong evidence against the null hypothesis. Therefore, the null hypothesis that the data have a unit root is rejected, concluding that the data is stationary.

The parameters of the genetic algorithm were as follows: maximum number of generations $= 20$, adopted as the stopping criterion; number of parents $\mu = 10$; number of children $\lambda = 10$; Lower and upper limit of gene initialization, respectively, $L_{\min} = 0$ and $L_{\max} = 3$.

In Table 2, the results of the generations that showed progress throughout the process are presented, including the generations that had progression, the parameter vector of the best individual, and their evaluation metrics.

Table 2. Fitness result

Iter	Parameters	MSE	MAE	RMSE	MAPE	AIC	BIC
1	[3. 1. 1. 1. 0. 1.]	0,912	0,623	0,955	0,126	788,162	813,995
3	[3. 1. 2. 1. 0. 3.]	0,909	0,618	0,954	0,125	793,217	830,121
4	[3. 1. 3. 1. 0. 1.]	0,906	0,620	0,952	0,125	790,148	823,361
5	[3. 1. 2. 2. 0. 2.]	0,906	0,624	0,952	0,126	791,892	828,795
6	[3. 1. 3. 2. 0. 1.]	0,905	0,623	0,951	0,126	791,735	828,638
7	[3. 1. 4. 2. 0. 2.]	0,889	0,618	0,943	0,126	789,484	833,769
11	[3. 1. 4. 3. 0. 3.]	0,887	0,625	0,942	0,127	792,381	844,046
12	[3. 1. 4. 3. 0. 4.]	0,886	0,622	0,941	0,125	794,244	849,599
13	[4. 1. 5. 3. 0. 3.]	0,884	0,617	0,940	0,125	795,676	854,722
15	[3. 1. 5. 5. 0. 3.]	0,882	0,616	0,939	0,125	796,809	859,545
18	[5. 1. 5. 6. 0. 6.]	0,864	0,616	0,930	0,125	801,670	886,548

The model performed a total of 20 iterations until establishing a consistency in data progression. Thus, Table 2 details the results that showed progress throughout the iterations and consequently presented the best results for the model. Regarding the associated metrics in the table, it can be observed that between the 1st and 18th iteration, the variation between the values is very close, such as in the MAPE index where the variation ranges from 0.125 to 0.127. Therefore, considering the best performance of the model, the 18th iteration exhibits the best and most balanced metrics within the model's proposal.

In Figure 3, a crucial diagnostic is illustrated, which is the residuals plot. This plot shows the difference between the actual and predicted values relative to the tests. It can be observed that the residuals have a random distribution around zero, indicating that the model captures the underlying patterns in the data.

Where the bars represent the difference between actual and predicted values, i.e., the noise. It is worth noting that when this error has a positive characteristic, it means the forecast is higher than the actual value, and when it is negative, it indicates the opposite. Therefore, the closer to zero, the better the forecast at that point. This error is caused by exogenous variables, which are external environmental factors not accounted for in the forecast.

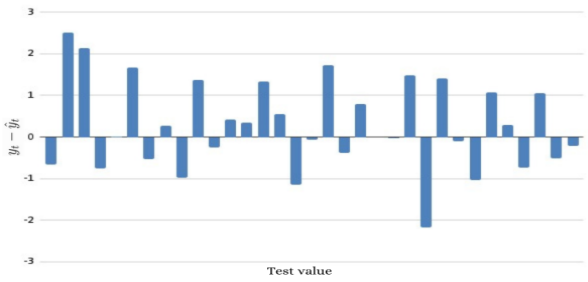

Figure 3. Residual plot. Source: Developed by the Authors.

Furthermore, most forecasts have an error variation of less than or equal to 20%, which is a favorable result for short-term forecasts.

During the search conducted by the algorithm, the values of the coefficients associated with the best individual found so far are revealed, as shown in Table 3. These coefficients represent the weights to be multiplied by the model variables, providing insights into how each variable influences the final outcome of the optimization process.

Table 3. Table of coefficients for the best solution.

Variable	Coefficient
ar.L1	-0.1848
ar.L2	-0.5416
ar.L3	-0.6189
ar.L4	0.2557
ar.L5	-0.3838
ma.L1	-0.0065
ma.L2	0.3829
ma.L3	0.2817
ma.L4	-0.5553
ma.L5	0.2803
ar.S.L6	0.1193
ar.S.L12	0.2277
ar.S.L18	2.2946
ar.S.L24	0.5747
ma.S.L30	-0.0421
ma.S.L36	-0.3130
ma.S.L6	-0.0772
ma.S.L12	-0.3081
ma.S.L18	-0.3651
ma.S.L24	-0.6503
ma.S.L30	0.1579
ma.S.L36	0.4534

For autoregressive variables (AR), the coefficients ($ar.L1$ to $ar.L5$) indicate the impact of past observations on the current observation, where negative coefficients suggest an inverse relationship, indicating that an increase in past observations leads to a decrease in the current observation, while positive coefficients suggest a direct relationship.

For moving average variables (MA) ($ma.L1$ to $ma.L5$), the coefficients represent the impact of past forecast errors on the current forecast, where negative coefficients indicate a downward correction and positive coefficients indicate an upward correction. The coefficients for SARIMA ($ar.S.L6$ to $ma.S.L36$) capture seasonal patterns in the data, where positive coefficients indicate an increase in associated seasonal observations and negative coefficients indicate a decrease.

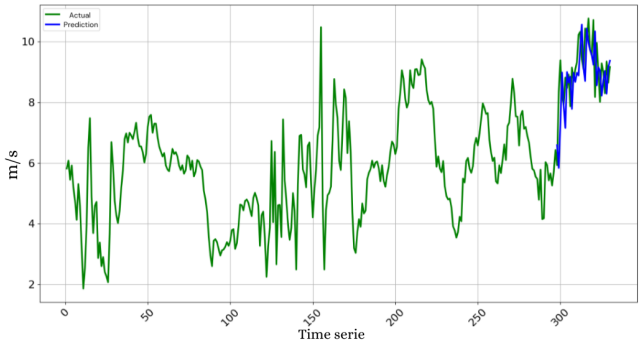

Figure 4. Forecasting with Genetic Algorithm and SARIMA. Source: Developed by the Authors.

In Figure 4, it can be observed how the time series compared to the best prediction obtained by the algorithm. There is a noticeable similarity between them and their behavior, which proves that such prediction is relevant to the objective.

The prediction was made using 90% of the data for training, resulting in 50.09 hours, while the remaining 10% for testing showed 5.56 hours. In total, the time series totaled approximately 55.66 hours, with a testing duration of about 5 hours and 30 minutes. This will be the interval for the predicted data. Additionally, the adequacy of the coefficients of the series equation will be done using the maximum likelihood method.

5. CONCLUSION

The study highlighted the feasibility of using algorithms for parameter optimization, particularly in the context of the SARIMA model. The genetic algorithm played a fundamental role in selecting the parameters to achieve the best performance in the regression model used. In this way, the study demonstrated the feasibility of the proposal for application in forecasting, emphasizing the adaptive capability of the genetic algorithm in improving and optimizing the model's input parameters.

Developed with a limited amount of data for training predictive models, the proposed model was able to demonstrate its feasibility, showing results with minimal errors, indicating the possibility of creating other combinations to solve complex problems. It is worth noting that the approach used involves significant computational expense, while alternative methods tend to reduce this aspect. However, its implementation becomes necessary when the goal is to achieve more accurate forecasting.

Additionally, it was possible to observe the statistical data resulting from the model and develop an automated approach to identify and classify the time series in order to adjust it in the best possible way. Thus, executing the genetic algorithm for enhanced adaptation and better results becomes feasible.

In this work, a wind speed forecast was made using 10% of the data for testing and 90% as predictor data for the model. The result was 50.09 hours, predicting an interval of 5 hours and 30 minutes, with an error variation of 20% or less. The model performed a total of 20 iterations to determine the optimal parameters. The iteration with the

chosen parameters was the 18th, with a MAPE index of 0.126.

Future work may involve the diversification of energy sources, the use of a more significant amount of data, as well as the application of other models in a hybrid manner.

ACKNOWLEDGMENT

The authors thank the Federal University of Maranhão (UFMA), the National Council for Scientific and Technological Development (CNPq) - Brazil, and the Foundation for Support to Research and Scientific and Technological Development of Maranhão (Fapema) for supporting our research. Special thanks also to CAPES and INEOF for the support provided.

REFERENCES

ABEEólica (2023). *INfovento*. ABEEólica, São Paulo.

Amirjanov, A. and Sadikoglu, F. (2017). Linear adjustment of a search space in genetic algorithm. *Procedia computer science*, 120, 953–960.

Araújo, E.G.d. (2021). *Modelagem e previsão de modelos de séries temporais do consumo de energia elétrica na região Nordeste do Brasil*. Master's thesis.

Bastos, S.M.D. (2024). *Aplicação de modelos preditivos de aprendizado de máquina para geração de energia solar*. B.S. thesis.

BUENO, R.D.L.d.S. (2012). *Econometria de Séries Temporais*. CENGAGE Learning, 2 edition.

EPE (2020). *Plano Nacional de Energia 2050,Brasil, Ministério de Minas e Energia, Empresa de Pesquisa Energética*. MME/EPE, Brasília.

EPE (2023). Balanço energético nacional, empresa de pesquisa energética (epe). Technical report, Brasil.

Gaspar-Cunha, A., Takahashi, R., and Antunes, C.H. (2012). *Manual de computação evolutiva e metaheurística*. Imprensa da Universidade de Coimbra/Coimbra University Press.

Gujarati, Damodar N. e PORTER, D.C. (2011). *Econometria Básica*. McGraw Hill Bookman, 5 edition.

Kramer, O. (2017). *Genetic Algorithm Essentials*, volume 1. Springer.

Lambora, A., Gupta, K., and Chopra, K. (2019). Genetic algorithm-a literature review. In *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)*, 380–384. IEEE.

Liu, X., Lin, Z., and Feng, Z. (2021). Short-term offshore wind speed forecast by seasonal arima - a comparison against gru and lstm. *Energy*, 227, 120492. doi:https://doi.org/10.1016/j.energy.2021.120492. URL https://www.sciencedirect.com/science/article/pii/S0360544221007416.

Liu, Z., Jiang, P., Zhang, L., and Niu, X. (2020). A combined forecasting model for time series: Application to short-term wind speed forecasting. *Applied Energy*, 259, 114137. doi:https://doi.org/10.1016/j.apenergy.2019.114137. URL https://www.sciencedirect.com/science/article/pii/S0306261919318240.

MICHALEWICZ, Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programs*. Springer.

Mitsa, T. (2010). *Temporal Data Mining*. Data Mining and Knowledge Discovery Series. Chapman & Hall/CRC.

Moreira, J.R.S. (2017). *Energias renováveis, geração distribuída e eficiência energética*. LTC, Rio de Janeiro.

Morettin, Pedro A. e Toloi, C.M. (2006). *Análise de Séries Temporais*. Editora Blucher, 2ª edition.

Pacheco, M.A.C. et al. (1999). Algoritmos genéticos: princípios e aplicações. *ICA: Laboratório de Inteligência Computacional Aplicada. Departamento de Engenharia Elétrica. Pontifícia Universidade Católica do Rio de Janeiro. Fonte desconhecida*, 28.

Pereira, R.S., Mendez, O.R.S., and Cortes, O.A.C. (2001). Parallel distributed evolution strategies. In *V Simpósio Brasileiro de Automação Inteligente*. Trabalho Completo.

Shumway, R.H., Stoffer, D.S., and Stoffer, D.S. (2000). *Time series analysis and its applications*, volume 3. Springer.

Sousa Lima, A.d. (2013). Análise geomorfológica da bacia hidrográfica do rio mearim-ma a partir do quadro geológico regional.

Souza, H.D.P., Pereira, M.M., Costa, A.M.S.C., de Lima, S.L., and Oliveira, D.Q. (2022). Nanorrede cc como solução de geração de energia e telecomunicação para microestações e perfiladores micrometeorológicos. In *Congresso Brasileiro de Automática-CBA*.

Wang, Y., Zou, R., Liu, F., Zhang, L., and Liu, Q. (2021a). A review of wind speed and wind power forecasting with deep neural networks. *Applied Energy*, 304, 117766.

Wang, Y., Zou, R., Liu, F., Zhang, L., and Liu, Q. (2021b). A review of wind speed and wind power forecasting with deep neural networks. *Applied Energy*, 304, 117766. doi:https://doi.org/10.1016/j.apenergy.2021.117766. URL https://www.sciencedirect.com/science/article/pii/S0306261921011053.

Yan, J., Möhrlen, C., Göçmen, T., Kelly, M., Wessel, A., and Giebel, G. (2022). Uncovering wind power forecasting uncertainty sources and their propagation through the whole modelling chain. *Renewable and Sustainable Energy Reviews*, 165, 112519.