Fábio Sandro dos Santos

Modeling wind speed and solar radiation in Brazil using mathematical-computational techniques

Recife-PE

2023

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

Modeling wind speed and solar radiation in Brazil using mathematical-computational techniques

This thesis was considered suitable for obtaining the title Ph.D. in Biometrics and Applied Statistics defended and approved unanimously on 2023-02-24 by the examining committee

Concentration area: Biometrics and Applied Statistics

Orientador: Prof. Ph.D. Tiago Alessandro Espínola Ferreira

Recife-PE

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

S237m Santos, Fábio Sandro dos
Modeling wind speed and solar radiation in Brazil using mathematical-computational techniques / Fábio Sandro dos Santos. - 2023.
96 f. : il.

Orientador: Tiago Alessandro Espinola Ferreira.
Inclui referências.

Tese (Doutorado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Biometria e Estatística Aplicada, Recife, 2023.

1. Wind energy . 2. Solar energy. 3. Renewable sources. 4. probability distribution. 5. MFDFA. I. Ferreira, Tiago Alessandro Espinola, orient. II. Título

CDD 519.5

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

Modeling wind speed and solar radiation in Brazil using mathematical-computational techniques

Fábio Sandro dos Santos

This thesis was considered suitable for obtaining the title Ph.D. in Biometrics and Applied Statistics defended and approved unanimously on 2023-02-24 by the examining committee

Advisor:

Prof. Ph.D. Tiago Alessandro Espínola Ferreira
Advisor

Examining Board:

Prof. Ph.D. Jader da Silva Jale
Universidade Federal Rural de Pernambuco

Prof. Ph.D. Tatijana Stosic
Universidade Federal Rural de Pernambuco

Prof. Ph.D. Sílvio Fernando Alves Xavier Júnior
Universidade Estadual da Paraíba

Prof. Ph.D. Manoel Henrique da Nóbrega Marinho
Escola Politécnica da Universidade de Pernambuco

Acknowledgments

First of all, I thank God for allowing me to reach this very important moment in my life.

Here I want to thank my father João de Deus, and my mother Joana Darc de Souto Santos for always encouraging me and my brothers to study. Here I also want to thank my three brothers: the pedagogue Prof. Sandra Maria dos Santos, physicist Prof. Simone Mark Santos, and the historian and pedagogue Ph.D. Alexandro dos Santos. And of course, not forgetting my nieces Hiandra Mel dos Santos Alfredo and Hillary dos Santos Alfredo.

I would like to thank my girlfriend Kerolly Kedma Felix do Nascimento for always being by my side, helping in difficult times and happy times. In the most difficult moment, she was by my side, advising me not to give up on my goal.

Here are my sincere thanks to my supervisor Professor Tiago Alessandro Espínola Ferreira, for his advice, patience, dedication, and always help with my research. Professor Tiago, thank you very much for everything, I know I'm not the best student/researcher, but whenever you need it , I'll be available to help with whatever it takes.

To Professor Jader da Silva Jale, here I would like to thank him for all the patience he had with me during this doctorate, when I contacted him he was always willing to help me with the programming and ideas for the work to be done.

To Professor Tatijana Stosic for the example of a researcher always available, very helpful, and attentive to the students of the Department of Biometrics and Applied Statistics, here I would like to thank you for having believed that I would be able to achieve the expected results for the doctorate.

To Professor Silvio Fernando Alves Xavier Junior for his partnership in scientific research and willingness to share his knowledge and experiences, here is my thank you.

To Professor Manoel Marinho for his valuable contributions to the work and willingness to participate in my thesis defense.

I would like to thank my friends Diego Alves and Rosendo Chagas, who have always helped me with advice and guidance. To the girls who put up with my games: Girl 1 (Mickaelle), Girl 2 (Leika), Jackson, Luciano, Ademir, Pablo, Joelma, Henrique Santos, Henrique Antunes, Rayane, Allana, Denise, Jucarlos, Elielma, Laura, and Daniel. I would

also like to express my thanks to the professors of the Graduate Program in Biometrics and Applied Statistics, for all their dedication to the program, for being accessible to students, and for all the teachings transmitted. In particular, professors Moacyr, Samuel, Borko Stosic, Rômulo Simões, Lucian Bejan, Guilherme, and Paulo Duarte. I also thank Secretary Marco for his availability.

Finally, I would like to thank the research funding bodies which throughout my research financed me so that I could able to carry out my research, especially CAPES for the financial support and the Federal Rural University of Pernambuco for the support and necessary conditions for carrying out my research.

"Do your best, in the condition you have, while you don't have better conditions, to do even better"
(Mario Sergio Cortella)

Abstract

In the globalized world, there is always the need for new investments in energy sources to meet all the demands, not only industrial but also population ones. In a world where we already have more than 8 billion inhabitants, there is a very great demand for energy for the daily needs of the population, for example. In addition to the need for energy, one concern is rising temperatures on earth. For this reason, countries have been trying at all costs to reduce the global average temperature of the earth by 2°C. For this goal to be achieved, many countries are investing in renewable energy sources as one of the ways to contribute these reductions in greenhouse gases in the atmosphere which is one of the causes of global warming. For this reason, in November 2021, in Glasgow, Scotland, the Brazil it committed by the year 2030 to reduce its greenhouse gas emissions by around 50%, with investments in clean and renewable energy. In Brazil, the energy sources that can contribute to the country achieving this established goal are wind and solar power. From this perspective, one of our objectives in this work was to understand and analyze the persistence and mixtures of probability distributions, through statistical, numerical, and artificial intelligence methods to estimate the potential of wind and solar power generation. For this, mixtures of probability distributions, and the Multifractal Detrended Fluctuation Analysis-MFDFA Method are used in the modeling of the series. In addition, the geographic spatialization of the potential of wind velocity values was performed, and it was observed that for those velocities that are above 3.0m/s, the higher the height, the greater the occurrence of these observations of velocities above this threshold. Among the five Brazilian regions (North, Northeast, South, Southeast and Midwest), it is observed that the Northeast region has higher potential for wind power generation. The region also showed good results for the installation of solar panels. Wind and solar energy sources are important for generating clean and renewable energy across the country and can be considered complementary sources. It is expected that this research will be able to assist public agencies in decision-making about investments in renewable energies, in particular, in the wind and solar energy sources. It is important to highlight that investments in wind and solar energy are needed in Brazil and around the world due to the growing need to replace conventional and non-renewable energy sources with renewable and clean alternatives.

Keywords: Wind energy, Solar energy, Renewable sources, probability distribution, MFDFA.

Resumo

No mundo globalizado há sempre a necessidade de novos investimentos em fontes de energia para atender todas as demandas não só industriais, mas também populacionais. Em um mundo com uma população de mais de 8 bilhões de habitantes, há uma demanda muito grande de energia para as necessidades diárias da população, por exemplo. Além da necessidade de energia, uma preocupação é o aumento das temperaturas na Terra. Por esta razão, os países têm tentado a todo custo reduzir a temperatura média global da Terra em 2°C. Para que esse objetivo seja alcançado, muitos países estão investindo em fontes renováveis de energia como uma das formas de contribuir com essas reduções dos gases de efeito estufa na atmosfera que é uma das causas do aquecimento global. Por isso, em novembro de 2021, em Glasgow, na Escócia, o Brasil se comprometeu até o ano de 2030 a reduzir suas emissões de gases de efeito estufa em cerca de 50%, com investimentos em energia limpa e renovável. No Brasil, as fontes de energia que podem contribuir para que o país alcance essa meta estabelecida são a energia eólica e a solar. Nessa perspectiva, um de nossos objetivos neste trabalho foi entender e analisar a persistência e as misturas de distribuição de probabilidade por meio de métodos estatísticos, numéricos e de inteligência artificial para estimar o potencial de geração de energia eólica e solar. Para isso, misturas de distribuições de probabilidade e o Método Multifractal Detrended Fluctuation Analysis (MFDFA) são utilizados na modelagem das séries. Além disso, foi realizada a espacialização geográfica dos valores potenciais de velocidade do vento e observou-se que para aquelas velocidades acima de 3,0m/s, quanto maior a altura, maior a ocorrência dessas observações de velocidades acima desse limiar. Dentre as cinco regiões brasileiras (Norte, Nordeste, Sul, Sudeste e Centro-Oeste), observa-se que a região Nordeste apresenta maior potencial de geração eólica. A região também apresentou bons resultados para a instalação de paineis solares. As fontes de energia eólica e solar são importantes para a geração de energia limpa e renovável em todo o país e podem ser consideradas fontes complementares. Espera-se que esta pesquisa possa auxiliar os órgãos públicos na tomada de decisões sobre investimentos em energias renováveis, em especial, nas fontes de energia eólica e solar. É importante destacar que investimentos em energia eólica e solar são necessários no Brasil e no mundo devido à crescente necessidade de substituição de fontes de energia convencionais e não renováveis por alternativas renováveis e limpas.

Palavras-chaves: Energia eólica, Energia solar, Fontes renováveis, distribuição de probabilidade, MFDFA.

List of Figures

CHAPTER 2

Figure 1 - Geographic distributions of stations distributed throughout the nine states of the Brazilian Northeast region and altitudes in the locality. 11
Figure 2 - Average (a) and standard deviation (b) of the hourly average wind speed in Northeast Brazil, in the period from 2004-01-01 to 2018-09-29. 12
Figure 3 - Boxplots of the wind speed time series of all meteorological stations in the Northeast region of Brazil. 13
Figure 4 - Hourly boxplot for Petrolina-PE in the period from 2003-02-21 to 2018-09-30 and Recife-PE in the period from 2004-12-22 to 2018-09-30. 13
Figure 5 - Weibull-Weibull distribution adjustments using the expectation maximization algorithm to the series of hourly average wind speeds in Petrolina and Recife, Pernambuco. 14
Figure 6 - Weibull-Weibull estimated parameters using the iterative expectation maximization method for the 136 weather stations analyzed. 15
Figure 7 - Weibull-Weibull estimated parameters using the iterative expectation maximization method for the 136 weather stations analyzed. 16
Figure 8 - On the left original wind speed series (in m/s) and on the right time series of wind speed anomalies from Petrolina-PE station. 16
Figure 9 - Results obtained with MFDFA for the wind speed series from Petrolina-PE meteorological station: (A) Fluctuation functions for $q = -10$, $q = 0$ and $q = 10$; (B) Generalized Hurst exponent; (C) Rényi exponent; (D) Multifractal spectrum. 16
Figure 10 - Comparison of the values of the parameters α_{origi} (blue) and α_{rand} (red) for the wind speed series of the entire Brazilian Northeast region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.). 17
Figure 11 - Comparison of the values of the parameters w_{origi} (blue), w_{rand} (red) and $w = w_{orig} - w_{rand}$ (black) for the wind speed series of the entire Brazilian Northeast region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.). 17
Figure 12 - Comparison of the values of the parameters r_{origi} (blue) and r_{rand} (red) for the wind speed series of the Northeast of Brazil. (For interpretation of the references to color in

this figure legend, the reader is referred to the web version of this article.) 17
Figure 13 - On the left, parameter α_0 of the original series and on the right, α_0 of the randomized series in Northeast of Brazil 18
Figure 14 - On the left, parameter w of the original series and on the right, w of the randomized series in Northeast of Brazil. 18
Figure 15 - On the left, parameter r of the original series and on the right, r of the randomized series in Northeast of Brazil. 19
Figure 16 - Difference between w_{orig} and w_{rand} (Δw) 20

CHAPTER 3

Figure 1 - Geographical position of the municipality of Petrolina in the upper Sertão of Pernambuco. 29
Figure 2 - . Distribution models adjusted to the wind speed database in Petrolina-PE. 33
Figure 3 - PSO convergence process in the search for the optimal parameters of the LNW distribution mode. 35
Figure 4 - Movement of the best particle in the search space in relation to the parameters of the Lognormal and Weibull distribution, respectively.....36

CHAPTER 4

Figure 1 - The geographical location of meteorological stations in the Northeast region of Brazil. 43
Figure 2 - Hourly time series of solar radiation $\left(\frac{W^2}{m}\right)$ measured on August 29, 2022. 43
Figure 3 - On the left is the mean, and on the right is the standard deviation of solar radiation $\left(\frac{W^2}{m}\right)$ in the Brazilian Northeast.....47
Figure 4 - Behavior found with the MFDFA method for the series of solar radiation anomalies in Maceió-AL. (a) Fluctuation Functions, (b) Generalized Hurst Exponent, (c) Rényi Exponent, and (d) Multifractal Spectrum. 48
Figure 5 - Comparison between the multifractal process complexity parameters of the original series versus the randomized series: α_0 (blue) anomaly series and α_0 (red) randomized series for all solar radiation series. 50
Figure 6 - Comparison between the multifractal process complexity parameters of the original series versus the randomized series: w (blue) anomaly series and w (red) randomized series for all solar radiation series. 51
Figure 7 - Comparison between the multifractal process complexity parameters of the original series versus the randomized series: r (blue) anomaly series and r (red) randomized series for all solar radiation series. 51
Figure 8 - a) values of α_0, b) values of w, while c) contains the asymmetry values r. 53

CHAPTER 5

Figure 1 - Geographic location of meteorological stations in Brazil. 62
Figure 2 - On the left is the mean and on the right is the standard deviation of Brazil's average hourly wind speed from 2010-01-01 to 2022-03-31. 66
Figure 3 - On the left are the percentages for May, and on the right are the wind speed values above $3.0m/s$. 67
Figure 4 - (a) Scale 1, (b) shape 1, (c) scale 2, (d) shape 2, (e) w1 and (f) w2 for the hourly average wind speed series in Brazil. 69
Figure 5 - On the left is the Average Wind Power Density, and on the right is the Wind Power Density based on the Weibull-Weibull distribution for Northeast Brazil from 2010-01-01 to 2022-03-31. 71
Figure 6 - Average Wind Power Density was analyzed monthly for the five Brazilian regions from 2010-01-01 to 2022-31-03.....73
Figure 7 - Average Wind Power Density Estimate using the Weibull-Weibull distribution for the five regions of Brazil from 2010-01-01 to 2022-31-03. 74
Figure 8 - Predominant wind direction throughout Brazil. 76

List of Tables

CHAPTER 2

Table 1 - Estimated minimum and maximum parameters of the Weibull-Weibull distribution in each state of the Northeast. 19
Table 2 - Weibull-Weibull distribution parameters in each studied station 22

CHAPTER 3

Table 1 - Mathematical equations of the Weibull, Weibull-Weibull and Lognormal-Weibull distribution models.....30
Table 2 - Descriptive analysis of wind velocity observations in the municipality of Petrolina-PE.....33
Table 3 - Estimates of statistical tests.....34

CHAPTER 4

Table 1 - Multifractal complexity parameters for some meteorological stations.....49

Contents

CHAPTER 1

Introduction....01
Objectives....03
Scientific Questions....04
Thesis Structure....05
References....07

CHAPTER 2

Mixture distribution and multifractal analysis applied to wind speed in the Brazilian Northeast region 10
1. Introduction 10
2. Materials and methods 12
3. Results and discussion 15
4. Conclusion 20
Declaration of Competing Interest 21
Acknowledgments 21
Appendix A 21
References 23

CHAPTER 3

Comparison of methods and distribution models for the modeling of wind speed data in the municipality of Petrolina, Northeast Brazil 25
1. Introduction 27
2. Area of Study and Data Collection 28
3. Distribution Models 29
4. Numerical Methods for Estimating Parameters 30
5. Metaheuristic Optimization for Estimating Models Parameters 31
6. Statistical Estimates 32
7. Results and Discussion 32
8. Final Considerations 36
Acknowledgments 37

References 37

CHAPTER 4

Multifractal analysis of solar radiation in the northeastern region of Brazil 40

1. Introduction 41
2. Study area 42
3. Method 44
4. Inverse Distance Weighting (IDW) 46
5. Results and discussion 47

CONCLUSION 54
ACKNOWLEDGMENT 54
REFERENCES 54

CHAPTER 5

Prediction of wind energy generation potential in Brazil using mixtures of Weibull distributions. 58

1. Introduction 59
2. Description of geographic location and data source 61
3. Estimation of wind speed at different heights 62
4. Mixture distribution 63
5. Algoritmo Expectation Maximization EM 63
6. Wind Power Density - WPD 65
7. Results and Discussion 66
8. Conclusion 76

ACKNOWLEDGMENT 77
REFERENCES 77

GENERAL CONCLUSIONS

General Conclusions 82

Chapter 1

Introduction

Wind speed and solar radiation are complex and nonlinear variables, which require effort for their modeling. Nevertheless, in recent years, several studies have been carried out seeking a way to use winds and radiation in the generation of wind and solar energy, since these are energy sources considered clean and inexhaustible. Another reason for the generation of this type of energy refers to the growth of the world population, with the need for greater energy consumption and replacement of conventional energy matrices, such as those from the water forces that remains as the main energy source (ABEEOLICA, 2019a), by renewable sources. Moreover, unlike that achieved with fossil fuels, the energy produced by the winds does not pollute the environment (CAMELO et al., 2015) and can contribute to the decrease in global warming, one of the major global concerns.

In Brazil, only the Midwest and the North regions do not produce wind power. The Northeast region has stood out as the main producer of wind energy, followed by the southern region of the country. The winds of the Northeast have been favorable to this end, as this region is benefited by the trade winds of the South Atlantic, characterized as strong, stable, and coming from the same direction in much of the time (ABEEOLICA, 2019b). With the winds considered the strongest in the world, the Northeast broke three records in August 2019, managing to produce 89% of all energy consumed in the region in this period (ONS, 2019). The dry period in the Northeast region of Brazil occurs in the second half of the year, leaving water reservoirs scarcer to produce energy in hydroelectric plants. However, in this period the most intense winds of the year occur, offsetting consumers with wind energy in full production.

The intense use of wind for wind power generation is justified by the number of wind farms installed: more than 600 in Brazil, with a total of more than 7000 wind turbines, producing around 15 GW of wind power. By 2024 the country is expected to have a production of approximately 20 GW of power (ABEEOLICA, 2019a). However, it is recognized the need to check the average speed of winds before the installation of a wind power plant, because the irregularity or scarcity of wind over long periods can

cause damage in the maintenance of a wind turbine. It should be noted that the minimum speed required to generate wind power is around 3.1 m/s. Values below this average make it impossible to install the parks since wind turbines would produce only enough energy for their own operation (SANTOS et al., 2021).

Several mathematical models and computational tools help to understand wind behavior (NEDAEI; ASSAREH; WALSH, 2018), such as probabilistic models Weibull (DEEP et al., 2020; SUMAIR et al., 2020; GUARIENTI et al., 2020; ARSLAN et al., 2020)), Rayleigh (SERBAN; PARASCHIV; PARASCHIV, 2020; CHIODO; NOIA, 2020; GORLA; PALLIKONDA; WALUNJ, 2020), and Gamma (ÖZKAN; SEN; BALLI, 2020; NADIA et al., 2020) in addition to the mixtures of Weibull-Gamma and Weibull-Normal truncated probability distributions, for example (MIAO et al., 2019). Given the excellent results obtained with mixtures of distributions to the historical series of wind speed with bimodality or multimodality behavior, it is chosen to use these combinations of probability distributions in the adjustment of distributions to be performed with this thesis. Seeking to minimize the estimation error associated with the estimated data in relation to the real ones, it is proposed the use of optimization algorithms, such as Particle Swarm Optimization (CARNEIRO et al., 2016) and the Expectation-Maximization numerical method (BRACALE; CARPINELLI; FALCO, 2017). Being chosen the best distribution for modeling the database, the next step was to calculate the wind power density for each historical series and then interpolate the wind potential throughout Brazil. Based on interpolation it is possible to know the potential of wind generation in places where you do not have information from the actual information of each weather station, creating a wind map of the country at six different heights of wind speed. The wind heights studied were 10m, 25m, 50m, 75m, 100m, and 120m. In addition, although there are several models of probability distributions for wind speed estimation, in this work we used the Multifractal Detrended Fluctuation Analysis (MFDFA) model to analyze the persistence existing in wind speed series over the years.

The MFDFA method, in addition to the application in wind speed series, was also used in hourly time series of solar radiation in the Northeast region of Brazil. This study can be considered one of the pioneering studies in the country, taking into account, mainly, the large number of meteorological stations that were used in the research. Some studies already done with solar radiation in northeastern Brazil were: Silva et al. (2010), which analyzed reanalysis data at two stations in the Northeast region and used cluster analysis in the data. Andrade e Tiba (2016), made a detailed analysis of terrestrial measurements of global solar irradiation, in eight meteorological stations in northeastern Brazil, which are, Água Branca, Santana do Ipanema, Palmeira dos Índios, Laje, Pão de Açúcar, Arapiraca, Coruripe, and Maceió. Lima, Ferreira e Morais (2017) analyzed the performance of a

photovoltaic system connected to an existing power distribution network in northeastern Brazil. In our research, 137 meteorological stations distributed in the nine states of the Northeast region of Brazil were studied, using the Multifractal Detrended Fluctuation Analysis (MFDFA) methodology. After the estimates of the multifractality parameters, we performed the parameters estimation where no information on solar radiation was available using the Inverse Distance Weighting (IDW) interpolation method.

This method has been widely used in modeling different phenomena of science, such as in the financial market (CHOROWSKI; STRUZIK, 2021; XIAO; WANG, 2021), agricultural commodities market (NASCIMENTO et al., 2022), diseases such as COVID-19 (KHAN et al., 2022), energy market (ALI; ASLAM; FERREIRA, 2021), the efficiency of the electricity market during the COVID-19 pandemic (NAEEM et al., 2022) and climatology (ZHANG et al., 2021). In the application of solar radiation series, studies such as (MADANCHI et al., 2017) analyzed solar radiation in some locations distributed worldwide using the MFDFA method. In Brazil, until where it was researched, we did not find studies using the MFDFA method applied to hourly series of solar radiation.

Finally, through this research, it was possible to understand the behavior of the wind speed and solar radiation series to help in the predictive capacity of wind and solar power generation. For example, if the wind speed in a given location is below $3.1m/s$, an aerogenerator will not produce enough power to power the subsystems of power distributions. Based on the predictive results obtained with the time series of wind speed and solar radiation, it is possible to outline investment strategies for the implementation of wind and solar farms that are promising for investors. This contributes to the reduction of costs in the implementation, generation, and distribution of energy in all Brazilian regions, thus cooperating so that Brazil achieves self-sufficiency in the generation of clean and inexhaustible energy in the future.

Objectives

General Objective

The objective of this thesis was to understand and analyze the persistence and mixtures of probability distributions, through statistical, numerical, and artificial intelligence methods to estimate the potential of wind and solar power generation in Brazilian data.

Specific Objectives

For the overall objective to be achieved, the specific objectives were:

- To analyze the spatial trend of wind speed in northeastern Brazil and multifractal parameters using the Inverse Distance Weighting method, to provide useful information in the selection of regions with wind potential;
- Compare several mixtures of probability distributions using statistical, numerical, and artificial intelligence methods in wind speed series in Petrolina-PE, Brazil;
- To verify multifractality properties existing in solar radiation series in northeastern Brazil;
- Estimate the density of wind potential in the five Brazilian regions from the mixture of probability distributions.

Scientific Questions

Four main questions address this thesis:

1. Can understanding the behavior of hourly wind speed contribute to the identification of possible locations for the installation of wind farms?

   To answer this question, in the first article, we applied the Multifractal Detrended Fluctuation Analysis method and the Weibull-Weibull probability distribution mixture to analyze the behavior of wind speed series in the nine regions of northeastern Brazil. The results showed that the wind speed series indicated persistent behavior in all meteorological stations and the Weibull-Weibull mixture achieved good adjustments to the hourly series with bimodal behavior. The Expectation Maximization algorithm was used to estimate the Weibull parameters. From this, we used the Inverse Distance Weighting method to estimate wind speed behavior information where no a priori information was yet available. Based on this information, investors and public agencies can make their decisions to install wind farms in certain locations or not.

2. Through statistical methods and artificial intelligence, what better combination of mixtures of probability distributions to adjust wind speed at the Petrolina-PE weather station in the Northeast region of Brazil?

   To answer this question, in the second article we compared several methods of mixtures of probability distributions in the adjustment of bimodal series of wind

speed in Petrolina-PE, Northeastern Brazil. To estimate the parameters of the mixing models, the optimization methods were used: Moment Method, Maximum likelihood method, and the artificial intelligence method Particle Swarm Optimization algorithm-PSO. Based on the results, the best model that adjusted to the data was the LogNormal-Weibull distribution mixture via PSO.

3. Do the solar radiation series in the Northeast region of Brazil present persistent behavior?

   To answer this question, in the third article, we applied the Multifractal Detrended Fluctuation Analysis-MFDFA method in the modeling of solar radiation series throughout the Brazilian Northeast region. We observed that all weather stations presented persistent behavior in the solar radiation series. This fact indicates that radiation series over time do not present significant changes in their behavior.

4. Would it be possible to distribute the surplus of wind energy produced in a given region of Brazil to other regions with low energy production?

   To answer this question, in the fourth article we used the Weibull-Weibull distribution mixing model to adjust the wind speed series throughout Brazil. The Expectation-Maximization method was used to estimate the parameters. After estimating the parameters of the mixing model, wind power density was calculated. Based on the results, it was possible to notice that in regions such as the Northeast, wind power density is four times higher than in the North region. In this case, the Northeast could easily distribute its surplus of production to the North region if it needs to. The same thing was observed in the southern region of the country, where this region could also distribute its surplus to the Southeast and Midwest regions, in addition to selling to countries that border the region.

Thesis Structure

The structure of this thesis consists of five chapters described below:

Chapter 1 – Introduction: A brief general introduction of the modeling and analysis of wind speed and solar radiation series, presentation of the general and specific objective of the thesis, and the main questions that guided this research;

Chapter 2 – Mixture distribution and multifractal analysis applied to wind speed in the Brazilian Northeast region. Article published in Chaos, Solitons and Fractals;

Chapter 3 – Comparison of methods and distribution models for the modeling of wind speed data in the municipality of Petrolina, Northeast Brazil. Article published in Research, Society and Development;

Chapter 4 – Multifractal analysis of solar radiation in the northeastern region of Brazil. Article accepted for publication in Fractals;

Chapter 5 – Prediction of wind energy generation potential in Brazil using mixtures of Weibull distributions. Article to be submitted for evaluation in Energy.

General Conclusions - Main conclusions of this thesis and possible future work.

References

ABEEOLICA. **Eólica já é a segunda fonte da matriz elétrica brasileira com 15 GW de capacidade instalada**. 2019a. Disponível em: <http://abeeolica.org.br/noticias/eolica-ja-e-a-segunda-fonte-da-matriz-eletrica-brasileira-com-15-gw-de-capacidade-instalada/>. Acesso em: 25 jan. 2023.

ABEEOLICA. **Brasil avança em ranking de energia eólica com ventos entre os melhores do mundo (Reuters)**. 2019b. Disponível em: <http://abeeolica.org.br/noticias/brasil-avanca-em-ranking-de-energia-eolica-com-ventos-entre-os-melhores-do-mundo-reuters/>. Acesso em: 12 out. 2019.

ALI, H.; ASLAM, F.; FERREIRA, P. Modeling dynamic multifractal efficiency of US electricity market. **Energies**, MDPI, v. 14, n. 19, p. 6145, 2021.

ANDRADE, R. C. de; TIBA, C. Extreme global solar irradiance due to cloud enhancement in northeastern Brazil. **Renewable energy**, Elsevier, v. 86, p. 1433–1441, 2016.

ARSLAN, H. et al. Wind speed variability and wind power potential over Turkey: Case studies for çanakkale and istanbul. **Renewable Energy**, Elsevier, v. 145, p. 1020–1032, 2020.

BRACALE, A.; CARPINELLI, G.; FALCO, P. D. A new finite mixture distribution and its expectation-maximization procedure for extreme wind speed characterization. **Renewable Energy**, Elsevier, v. 113, p. 1366–1377, 2017.

CAMELO, H. do N. et al. Previsão de velocidade média do vento através da utilização de modelagem auto-regressiva de médias móveis (arma) em região serrana no estado do ceará-brasil. **Revista Brasileira de Energias Renováveis**, v. 4, n. 3, 2015.

CARNEIRO, T. C. et al. Particle swarm optimization method for estimation of Weibull parameters: a case study for the Brazilian northeast region. **Renewable energy**, Elsevier, v. 86, p. 751–759, 2016.

CHIODO, E.; NOIA, L. P. D. Stochastic extreme wind speed modeling and bayes estimation under the inverse Rayleigh distribution. **Applied Sciences**, MDPI, v. 10, n. 16, p. 5643, 2020.

CHOROWSKI, M.; STRUZIK, Z. Local Scaling Analysis with Continuous DFA in High Frequency Financial Time Series. **Acta Physica Polonica, A.**, v. 139, n. 4, 2021.

DEEP, S. et al. Estimation of the wind energy potential for coastal locations in india using the weibull model. **Renewable energy**, Elsevier, v. 161, p. 319–339, 2020.

GORLA, R.; PALLIKONDA, M.; WALUNJ, G. Use of Rayleigh distribution method for assessment of wind energy output in Cleveland–Ohio. **Renewable Energy Research and Applications**, Shahrood University of Technology, v. 1, n. 1, p. 11–18, 2020.

GUARIENTI, J. A. et al. Performance analysis of numerical methods for determining weibull distribution parameters applied to wind speed in Mato Grosso do Sul, Brazil. **Sustainable Energy Technologies and Assessments**, Elsevier, v. 42, p. 100854, 2020.

KHAN, K. et al. COVID-19 impact on multifractality of energy prices: Asymmetric multifractality analysis. **Energy**, Elsevier, v. 256, p. 124607, 2022.

LIMA, L. C. de; FERREIRA, L. de A.; MORAIS, F. H. B. de L. Performance analysis of a grid connected photovoltaic system in northeastern Brazil. **Energy for Sustainable Development**, Elsevier, v. 37, p. 79–85, 2017.

MADANCHI, A. et al. Strong short-term non-linearity of solar irradiance fluctuations. **Solar Energy**, Elsevier, v. 144, p. 1–9, 2017.

MIAO, S. et al. Determining suitable region wind speed probability distribution using optimal score-radar map. **Energy conversion and management**, Elsevier, v. 183, p. 590–603, 2019.

NADIA, A. et al. Interrelation Between TRM, Wind Power, and Wind Speed. In: IEEE. **2020 IEEE Region 10 Symposium (TENSYMP)**. [S.l.], 2020. p. 1743–1746.

NAEEM, M. A. et al. Comparing the asymmetric efficiency of dirty and clean energy markets pre and during COVID-19. **Economic Analysis and Policy**, Elsevier, v. 75, p. 548–562, 2022.

NASCIMENTO, K. K. F. D. et al. Covid-19 Influence Over Brazilian Agricultural Commodities and Dollarâ€“real Exchange. **FRACTALS (fractals)**, World Scientific Publishing Co. Pte. Ltd., v. 30, n. 06, p. 1–10, 2022.

NEDAEI, M.; ASSAREH, E.; WALSH, P. R. A comprehensive evaluation of the wind resource characteristics to investigate the short term penetration of regional wind power based on different probability statistical methods. **Renewable Energy**, Elsevier, v. 128, p. 362–374, 2018.

ONS. **Operador Nacional do Sistema Elétrico - ONS**. 2019. Disponível em: <http://www.ons.org.br/AcervoDigitalDocumentosEPublicacoes/Boletim_Eolica_ago2019.pdf>. Acesso em: 12 out. 2019.

ÖZKAN, R.; SEN, F.; BALLI, S. Evaluation of wind loads and the potential of turkey's south west region by using log-normal and gamma distributions. **Wind and Structures**, Techno-Press, v. 31, n. 4, p. 299–309, 2020.

SANTOS, F. S. dos et al. Mixture distribution and multifractal analysis applied to wind speed in the brazilian northeast region. **Chaos, Solitons & Fractals**, Elsevier, v. 144, p. 110651, 2021.

SERBAN, A.; PARASCHIV, L. S.; PARASCHIV, S. Assessment of wind energy potential based on Weibull and Rayleigh distribution models. **Energy Reports**, Elsevier, v. 6, p. 250–267, 2020.

SILVA, V. d. P. R. D. et al. Trends in solar radiation in NCEP/NCAR database and measurements in northeastern Brazil. **Solar Energy**, Elsevier, v. 84, n. 10, p. 1852–1862, 2010.

SUMAIR, M. et al. **RETRACTED: A newly proposed method for Weibull parameters estimation and assessment of wind potential in Southern Punjab**. [S.l.]: Elsevier, 2020.

XIAO, D.; WANG, J. Attitude interaction for financial price behaviours by contact system with small-world network topology. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 572, p. 125864, 2021.

ZHANG, L. et al. Application of an improved multifractal detrended fluctuation analysis approach for estimation of the complexity of daily precipitation. **International Journal of Climatology**, Wiley Online Library, v. 41, n. 9, p. 4653–4671, 2021.

Chapter 2

Chaos, Solitons and Fractals 144 (2021) 110651

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Mixture distribution and multifractal analysis applied to wind speed in the Brazilian Northeast region

Fábio Sandro dos Santos[a,*], Kerolly Kedma Felix do Nascimento[a], Jader da Silva Jale[a], Tatijana Stosic[a], Manoel H.N. Marinho[b], Tiago A.E. Ferreira[a]

[a] Department of Statistics and Informatics, Federal Rural University of Pernambuco, Recife, Pernambuco, Brazil
[b] University of Pernambuco, Recife PE, Brazil

ARTICLE INFO

Article history:
Received 28 August 2020
Revised 4 December 2020
Accepted 4 January 2021
Available online 17 January 2021

Keywords:
Seasonality
Persistence
Wind energy
Weibull distribution

ABSTRACT

The growing investments and installations of wind farms in the Brazilian Northeast have drawn attention to the region, leading investors and researchers to seek better ways of using the local wind regimen for energy production. In face of the complex behavior of wind speed time series, mixture distribution models have been applied to bimodal databases aiming at achieving the best modeling for series fitting. This paper used data from stations located in the nine states that make up the Brazilian Northeast region (Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, and Bahia) between January 1st, 2004 and August 29th, 2018. The two-component Weibull distribution model was employed to model the historical series using the Expectation Maximization (EM) algorithm to search for optimal parameters in data distribution. Multifractal detrended fluctuation analysis was applied to verify series persistence over time and, using spatialization obtained with inverse distance weighting, the results were estimated at the sites lacking meteorological wind information. The results obtained indicate that the highest mean wind speeds are found in the states of Rio Grande do Norte, Paraíba, and Pernambuco, whereas the lowest occur in parts of Bahia, Piauí, and Maranhão. The highest mean wind speeds were recorded between 10 a.m. and 8 p.m. of each day at every station. Multifractal analysis revealed that wind speed series exhibit persistent overall behavior for all stations, with multifractality dominated by small fluctuations. For most of the stations both long term correlations and broad probability density function of wind speed values are found to cause multifractality of the process. This study allows identifying favorable areas for the installation of wind farms in different locations of the Brazilian Northeast region.

1. Introduction

Viable installation of a wind farm requires investigations on the conditions of the region considered promising for energy generation, such as its topography, temperature, and humidity. Since they can be strongly impacted by local terrain, wind characteristics such as direction and gust speed are essential for deciding the position of wind turbines and optimal use of wind power. Particularly, prior investigation on mean wind speed is key to verify whether investments at the site are suitable. The use of wind energy may reduce the dependence on hydroelectric generation, the main source of electricity in Brazil [1]. Such replacement means water can be used for population consumption, particularly during periods of drought, as was the case during the extended dry spell in 2013 in the Brazilian northeast. At the time, the volume of the biggest artificial lake of the world, the hydroelectric Sobradinho reservoir was greatly reduced and wind power was an alternative adopted to keep water from being used for energy generation. This initiative was responsible for supplying on average 30% of the energy consumed in the region by using a clean and limitless source [2]. In 2019, during some periods of August, the good winds of the Northeast were able to supply a record 89% of the regional energy generation [3].

In this sense, statistical tools have been adopted to help determine possible sites for the installation of wind farms, such as the probability distribution models. These models were employed to fit historical wind speed series in China [4], India [5], Iran [6], Iraq [7], Morocco [8], Nepal [9], Nigeria [10], Rwanda [11], Switzerland [12], and United States [13]. Among the several available models,

* Corresponding author.
E-mail address: fabio.sandropb@gmail.com (F.S.d. Santos).

https://doi.org/10.1016/j.chaos.2021.110651
0960-0779/

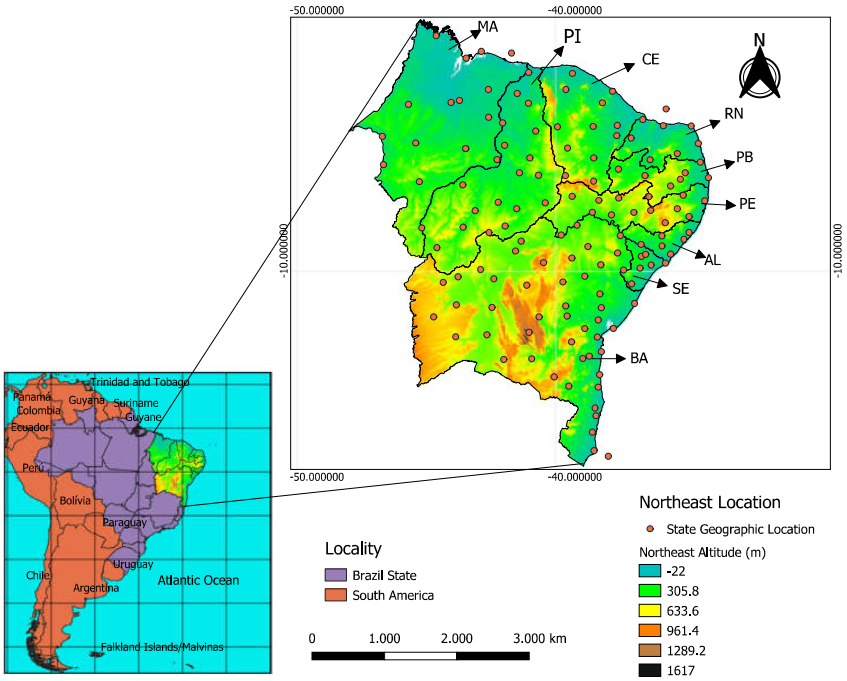


Fig. 1. Geographic distributions of stations distributed throughout the nine states of the Brazilian Northeast region and altitudes in the locality.

mixtures distribution have been standing out for their capacity of modeling bimodal databases [14], as in the case of wind data in Sri Lanka for which mixture Weibull distribution performed better than Lognormal, Gamma, Gamma and Weibull mixture, and Lognormal and Weibull mixture distributions, for example [15]. Shin et al. [16] applied twelve models of probability distributions, of which ten were mixtures of distributions, on wind speed data from the United Arab Emirates. Through the adjustments made with the Least Squares, Maximum Likelihood and Expectation Maximization methods, the results showed that the distribution mixtures provided the best adjustments to all analyzed data [16].

Another relevant piece of information on appropriate regions for the installation of wind farms is the persistence of wind regimens, i.e., whether the wind behavior repeats over time. Such knowledge can be acquired using multifractal analysis. The multifractal detrended fluctuation analysis methodology is employed in different scientific fields, such as to analyze the concentration of air pollutants [17], rainfall data [18], and wind speed series [12,19]. Laib et al. (2018) examined a wind speed monitoring system in Switzerland using correlation networks. In the study, stations are seen as nodes on the network and the application of MFDFA is made on the density of connectivity [20]. Telesca and Lovallo [21] analyzed wind speed series in Italy at seven different heights and observed through the MFDFA that there is a dependence between wind sensor height and the multifractal fluctuations present in wind speed. These results were consistent with those obtained through the Fisher-Shannon method [21]. Balkissoon et al. [22] investigated fractal and multifractal characteristics in wind speeds observed in the towers of three sites in Missouri. The authors studied Hurst's exponent in the monofractals series, using Rescaled Analysis and in the multifractal series using MFDFA. The results showed that the fractal dimensions of monofractals series are smaller to those found with multifractal analysis [22].

With that in mind, the present research analyzed wind speed databases in the Brazilian Northeast region aiming to verify the behavior of those series. The Weibull distribution model with five parameters was fitted to the data and after that, was checked whether they exhibit persistence over time. These investigations on the stochastic behavior of wind may contribute to estimating the wind power generation potential in the states of the Northeast region and assess whether said generation can be constant over time. Up to date, we are not aware that studies like this have been conducted in Brazil, considering the entire wind speed database of the Northeast region of Brazil (corresponding to more than 120,000 h of observations from 136 automatic stations). Previous studies have investigated only certain isolated regions and smaller time intervals or fewer stations. For example, Santos et al. [23] verify the existence of long-range correlations in some regions of Bahia, in a period of 60 days (January and February 2016). Rocha et al. [24] evaluated seven numerical methods in the adjustment of the Weibull distribution in the cities of Camocim and Paracuru in Ceará, from 2004 to 2006. Torres Silva dos Santos e Silva [25] analyzed linear trends, seasonality and interannual variability of wind speed recorded at 47 meteorological station for the period from 1986 to 2011.

Another important contribution of our work is the presentation of the spatial trend of wind speed in the Brazilian Northeast and the multifractal parameters through geostatistical techniques (via Inverse Distance Weighting), with the objective to provide useful information for the selection of regions with wind potential. We use Inverse Distance Weighting because it is a simple, fast and appropriate method when working with many observations [26]. It is a method commonly used in climate variable analysis [27], including wind speed time series. Considering that this study was carried out with methods already well consolidated in the literature and with data from Brazil, in particular from the Brazilian

Northeast region, other researchers can replicate the ideas adopted here to identify areas favorable to the installation of wind farms in other countries and thereby foster their economies, ensuring development and progress in different socioeconomic sectors.

The remainder of this paper is structured as follows. Section 2 presents the data used and the description of each methodology employed. Section 3 exhibits the results reached and discussions. Section 4 presents the main conclusions obtained.

2. Materials and methods

2.1. Data

The database used in the present study was obtained from the National Institute of Meteorology (INMET) with hourly measurements in m/s between January 1st, 2004 to September 29th, 2018, comprising over 120 thousand hours of wind speed observations. Such observations correspond to the information collected from 136 meteorological stations scattered across the Northeast region of Brazil, as shown in Fig. 1.

The area under study comprehends approximately 1,561,177.8 km^2, equivalent to 18.3% of the Brazilian territory, and comprises nine states (Maranhão (MA), Piauí (PI), Ceará (CE), Rio Grande do Norte (RN), Paraíba (PB), Pernambuco (PE), Alagoas (AL), Sergipe (SE) and Bahia (BA)) holding a total population estimated at 53,081,950 inhabitants. Mean annual wind speed in the region measured at a height of 10 m lies between 0.5 and 5.5 m/s [25].

2.2. Mixture of two Weibull distributions models

The complex non-linear behavior of wind speed series [28] makes it difficult to specify an optimal distribution that comprehensively models those series [29,30]. However, mixtures distributions have been a successful option as they tend to exhibit good fits to the frequencies observed.

A mixture of distributions can be obtained from the linear composition of two or more probability density functions (pdf). Eq. (1) expresses the mathematical formulation of a mixed distribution [14]

$$f(v;\rho_1,\ldots,\rho_d,\theta_1,\ldots,\theta_d)=\sum_{i=1}^{d}\rho_i f_i(v;\theta_i), \tag{1}$$

where d is the number of components and ρ_i are the weights of each mixture so that $\rho_1+\rho_2+\ldots+\rho_d=1$, θ_i are the parameters of the ith distribution and $f_i(v;\theta_i)$ are the independent distributions of the ith components.

Particularly when proposing the combination of two Weibull distribution models, one with two parameters and the other with three, the expression seen in Eq. (2) is obtained [31],

$$\begin{aligned} ff_{2,2}(v,\rho,\alpha_1,\beta_1,\alpha_2,\beta_2) &= \rho_1\frac{\alpha_1}{\beta_1}\left(\frac{v}{\beta_1}\right)^{\alpha_1-1} exp\left[-\left(\frac{v}{\beta_1}\right)^{\alpha_1}\right] \\ &+\rho_2\frac{\alpha_2}{\beta_2}\left(\frac{v}{\beta_2}\right)^{\alpha_2-1} exp\left[-\left(\frac{v}{\beta_2}\right)^{\alpha_2}\right], \end{aligned} \tag{2}$$

where v is the mean hourly wind speed (in m/s) above zero, ρ_1 and $\rho_2=1-\rho_1$ are the weights assigned to the distribution, α_1 and α_2 are shape parameters, and β_1 and β_2 are scale parameters. This model is known as five-parameter Weibull distribution (Weibull-Weibull).

2.3. Expectation maximization algorithm

Choosing the optimal method for parameter estimation may prevent convergence issues and ensures lower risk of false estimates. The Expectation Maximization (EM) algorithm is an iterative method used to estimate parameters based on an approach with maximum likelihood [32]. The EM algorithm has been successfully applied to models of mixtures distribution [33].

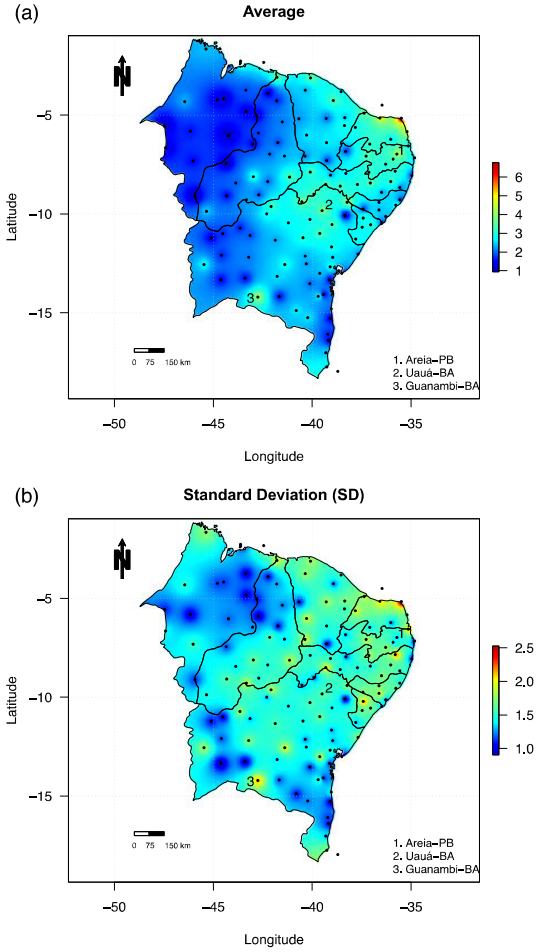


Fig. 2. Average (a) and standard deviation (b) of the hourly average wind speed in Northeast Brazil, in the period from 2004-01-01 to 2018-09-29.

In order to maximize a function given by

$$g(\theta)=\log\left(\sum\nolimits_{i=1}^{n}\theta_i\right), \tag{3}$$

where $\theta^\top=(\theta_1,\ldots,\theta_n)\in\mathbb{R}_+^n$, g can be minorized in $\psi\in\mathbb{R}_+^n$ by a Jensens inequality minorizer [34], such as

$$Q(\theta;\psi)=\sum_{i=1}^{n}\tau_i(\psi)\log(\theta_i)-\sum_{i=1}^{n}\tau_i(\psi)\log\tau_i(\psi), \tag{4}$$

where $\tau_i(\psi)=\psi_i/\sum_{j=1}^{n}\psi_j$.

Given $\mathbb{S}_n=\left\{s^\top=(s_1,\ldots,s_n):s_i\geq 0 \text{ for every } i\in[n].\sum_{i=1}^{n}s_i=1\right\}$, where $(.)^\top$ is the transposition operator and $[n]=1,\ldots,n$, the expression in Eq. (5) is obtained by applying Eq. (4) to the likelihood ratio at the rth iteration of the algorithm.

$$Q\left(\alpha;\alpha^{r-1}\right)=\sum_{k=1}^{N}\sum_{i=1}^{n}\tau_i\left(x_k;\alpha^{r-1}\right)log\left[\alpha_i\prod_{j=1}^{d}w_{ij}\phi(w_{ij}x_{kj}-w_{ij}y_{ij})\right]$$

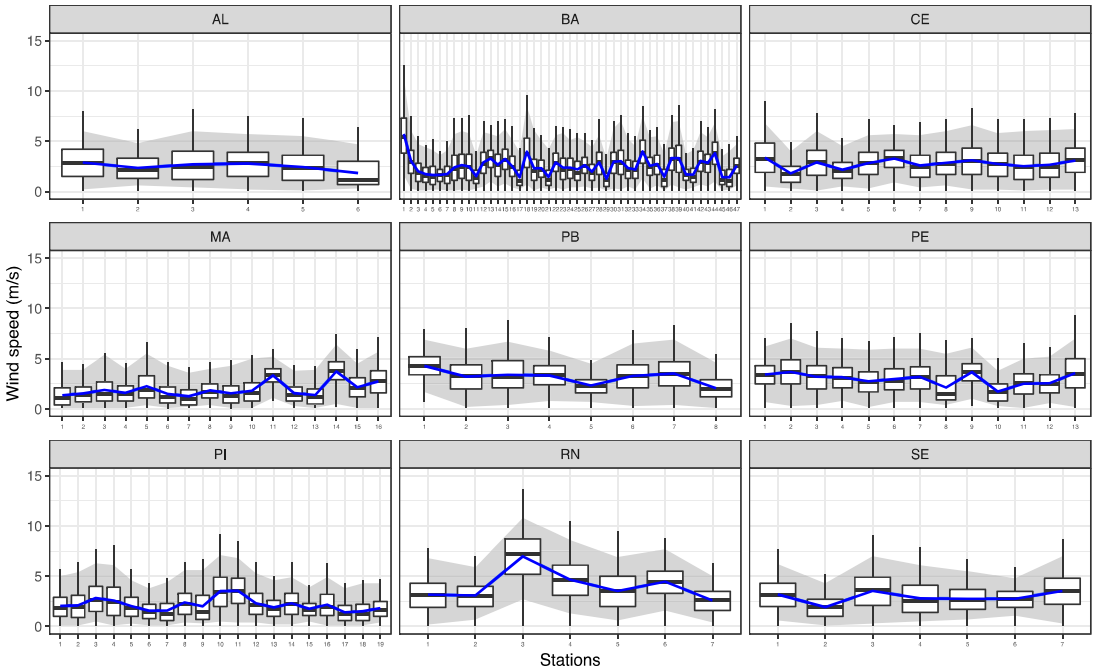


Fig. 3. Boxplots of the wind speed time series of all meteorological stations in the Northeast region of Brazil.

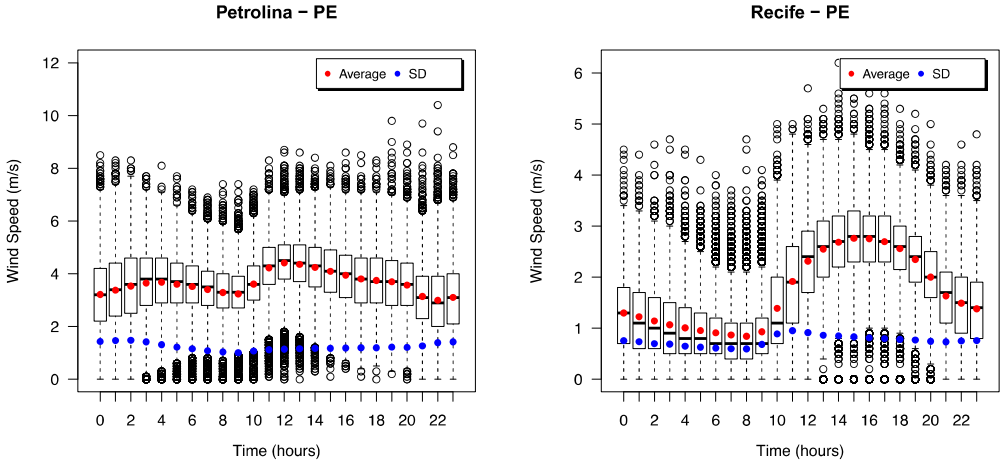


Fig. 4. Hourly boxplot for Petrolina-PE in the period from 2003-02-21 to 2018-09-30 and Recife-PE in the period from 2004-12-22 to 2018-09-30.

$$-\sum_{k=1}^{N}\sum_{i=1}^{n}\tau_i\left(x_k;\alpha^{r-1}\right)log\tau_i\left(x_k;\alpha^{r-1}\right) \tag{5}$$

where $x_k^{\top}=(x_{k1},\ldots,x_{kd})$ and $\tau_i(x;\alpha)=\frac{\alpha_i\prod_{j=1}^{d}w_{ij}\phi(w_{ij}x_j-w_{ij}y_{ij})}{\sum_{k=1}^{n}\left[\alpha_k\prod_{j=1}^{d}w_{kj}\phi(w_{kj}x_j-w_{kj}y_{kj})\right]}$.

With a solution vector $\alpha^*\in\mathbb{S}_n$, where $\nabla Q(\alpha^*;\alpha^{r-1})=0$, the rth iteration can be found. That yields solution $\alpha^{*\top}=(\alpha_1^*,\ldots,\alpha_n^*)$, where $\alpha_i^*=\frac{1}{N}\sum_{k=1}^{N}\tau_i\left(x_k;\alpha^{r-1}\right)$. $\alpha^r=\alpha^*$ defines the algorithm at each iteration $r\in\mathbb{N}$ until some stopping criterion is reached, such as the maximum number of iterations. The final iteration is then said to be the maximum likelihood estimator (MLE) $\widehat{\alpha}_N$.

2.4. Multifractal detrended fluctuation analysis

Multifractal Detrended Fluctuation Analysis (MFDFA) method was originally proposed in 2002 with the aim to detect multifractality in non-stationary time series [35]. In recent years MFDFA has been widely applied in different types of phenomena such as meteorological time series [36], air temperature [37], finances [38–41], oxygen in earths early atmosphere [42], air traffic flow [43], forest fires [44] and wind speed [45]. The implementation of the MFDFA method consists of five steps [35]:

Step 1: Integrate the time series x_i forming a new series $Y(k)=\sum_{i=1}^{N}[x_i-\bar{x}]$, where $\bar{x}$ is the average value of x_i and $k=1,\ldots,N$.

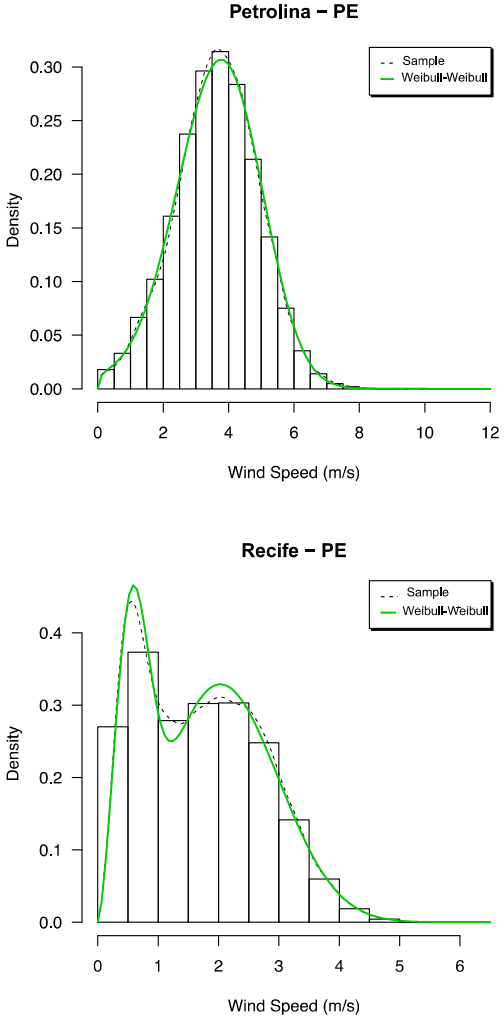


Fig. 5. Weibull-Weibull distribution adjustments using the expectation maximization algorithm to the series of hourly average wind speeds in Petrolina and Recife, Pernambuco.

Step 2: Divide the series Y_k into $N_s = int(N/s)$ non-overlapping segments of equal length s, starting the divisions on each of the two ends of the series and obtaining $2N_s$ segments.

Step 3: Determine local trend for each of the $2N_s$ segments using polynomial fit of degree m and calculate the corresponding variance using Eqs. (6) and (7).

$$F^2(v,s) \equiv \frac{1}{s}\sum_{i=1}^{s}\{Y[(v-1)s+i] - y_v(i)\}^2, \qquad v = 1, 2, \ldots, N_s \tag{6}$$

and

$$F^2(v,s) \equiv \frac{1}{s}\sum_{i=1}^{s}\{Y[N-(v-N_s)s+i] - y_v(i)\}^2, \text{ para } v = N_s+1, \ldots, 2N_s \tag{7}$$

where $y_v(i)$ is the fitting polynomial in segment v.

Step 4: Calculate the average of $F^2(v;s)$ over all segments and determine the fluctuation function according to the following expression.

$$F_q(s) \equiv \left\{\frac{1}{2N_s}\sum_{v=1}^{2N_s}\left[F^2(v,s)\right]^{\frac{q}{2}}\right\}^{\frac{1}{q}}, \tag{8}$$

where, in general, the index $q \in \mathbb{R}^*$.

Step 5: Repeat the calculation of fluctuation function $F_q(s)$ for all segment sizes. If long-term correlations are present in original series x_i, $F_q(s)$ will increase with s according to a power law $F_q(s) \sim s^{h(q)}$. For $q = 0$, $h(0) = lim_{h\to 0}h(q)$ is determined by means of the expression below.

$$F_0(s) \equiv exp\left\{\frac{1}{4N_s}\sum_{v=1}^{2N_s} ln\left[F^2(v,s)\right]\right\} \sim s^{h(0)} \tag{9}$$

For multifractal series $h(q)$ depends on q and decreases while q increases. On the other hand, when $h(q)$ does not depend on q the original series is a mono-fractal. For negative values of q, $h(q)$ characterizes the scaling behavior of subsets of the series with small fluctuations, while for positive values of q, $h(q)$ characterizes the scaling behavior of subsets with large fluctuations [35]. Multifractality of time series can also be characterized by Rényi's exponent [44]:

$$\tau(q) = \quad qh_q - 1 \tag{10}$$

which is linear function for monofractal series and nonlinear function for multifractal series. More information about multifractality can be obtained from multifractal spectrum which is defined as

$$f(\alpha) = q(\alpha - h(q)) + 1, \tag{11}$$

where

$$\alpha = h(q) + q\frac{dh(q)}{dq}, \tag{12}$$

is the Hölder exponent and $f(\alpha)$ is the dimension of the subset of the series that is characterized by the exponent α.

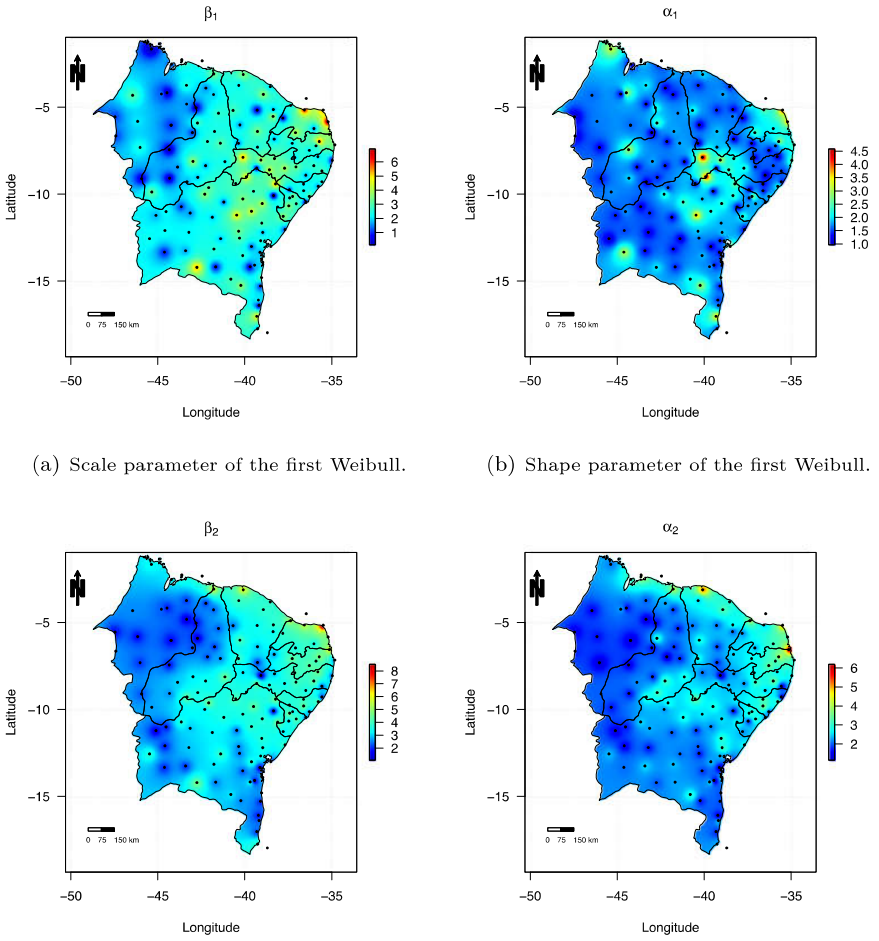
Graphically, a multifractal series is described by a concave down $f(\alpha)$ curve where the maximum point indicates the value of the overall Hurst exponent. For monofractals series the singularity spectrum is represented by only one point [46].

After a polynomial fitting of $f(\alpha)$, the complexity of the multifractal process can be described with the aid of the parameters presented below [47].

- Position of the maximum of the spectrum (α_0) - represents the overall Hurst exponent. If $\alpha_0 < 0.5$ the series is said to be antipersistent, if $\alpha_0 > 0.5$ the series is persistent and if $\alpha_0 = 0.5$ the series is said to be random;
- Asymmetry (r) - obtained by the expression $r = \frac{\alpha_{max} - \alpha_0}{\alpha_0 - \alpha_{\min}}$, where $\alpha_{\max}$ and $\alpha_{\min}$ are the two zero-crossings of the fitting polynomial. If $r = 1$ the $f(\alpha)$ spectrum is symmetric, indicating that both subsets with small and large fluctuations contribute equally to multifractality of process. If $r < 1$ there is asymmetry to the left and subsets with large fluctuations are dominant in multifractality of the process. The values $r > 1$ indicate the asymmetry to the right and the dominance of small fluctuations.
- Spectrum width (w) - given by $w = \alpha_{\max} - \alpha_{\min}$. It measures the degree of multifractality: the greater the width of the spectrum, the stronger the multifractality present in the process.

A time series can exhibit two types of multifractality:

i. Multfractality from a broad probability density function of the series.
ii. Multifractality from different long-range correlations for small and large fluctuations.



(a) Scale parameter of the first Weibull.

(b) Shape parameter of the first Weibull.

(c) Scale parameter of the second Weibull.

(d) Shape parameter of the second Weibull.

Fig. 6. Weibull-Weibull estimated parameters using the iterative expectation maximization method for the 136 weather stations analyzed.

Series randomization is recommended as an aid to identify the causes of multifractality in the process. After randomization the long term correlations are destroyed while probability density function is preserved. This leads to the loss of multifractality (f(α) spectrum is significantly reduced) for multifractals of type (ii) while for multifractals of type (i) $f(\alpha)$ spectrum remains the same. When both types of multifractality are present, the shuffled series shows weaker multifractality than the original series [35].

2.5. Inverse distance weighting

Among the interpolation methods available, Inverse Distance Weighting (IDW) has stood out since it does not require subjective conjectures or pre-modeling, besides being relatively simple to implement [48]. Originally, IDW was proposed aiming to solve bidimensional problems. However, it was later adapted to solve multidimensional problems [49].

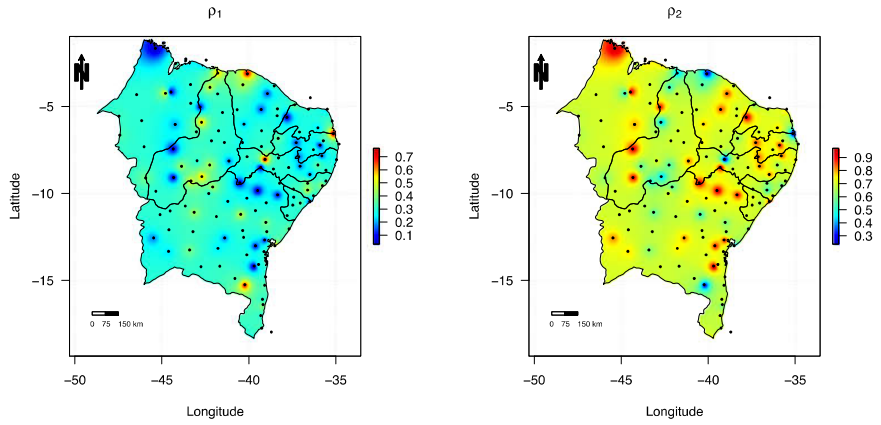
It is based on the principle that the closer the estimated value of the actual value is, the greater its influence will be on predicted values of more distant ones. The mathematical equation of IDW can be expressed as [50]

$$\widehat{Z}(u_0) = \frac{\sum_{i=1}^{N} Z(u_i) d_{0,i}^{-p}}{\sum_{i=1}^{N} d_{0,i}^{-p}}, \quad i = 1, \ldots, N, \tag{13}$$

where u_0 is an estimate, u_i are real points, $d_{0,i}^{-p}$ is the weight, d is the Euclidian distance between the values estimated and each station and p is the exponential power parameter. In general, $p = 2$.

3. Results and discussion

Based on the descriptive statistics of average and standard deviation of the 136 meteorological stations studied, the information was spatialized to the entire Brazilian Northeast region using the IDW method. Fig. 2 shows the plot obtained. The highest mean wind speeds were found on the coast of Rio Grande do Norte (RN) and at the stations of Areia-PB, Uauá-BA, and Guanambi-BA (in yellow). A comparison of the states revealed that the mean wind speeds in RN, PB, and PE were higher than in the other states. In PI, the fifth largest producer of wind power in Brazil by late 2018 [51], the highest mean wind speeds are found in the north and



(a) Weight of the first Weibull distribution. (b) Weight of the second Weibull distribution.

Fig. 7. Weibull-Weibull estimated parameters using the iterative expectation maximization method for the 136 weather stations analyzed.

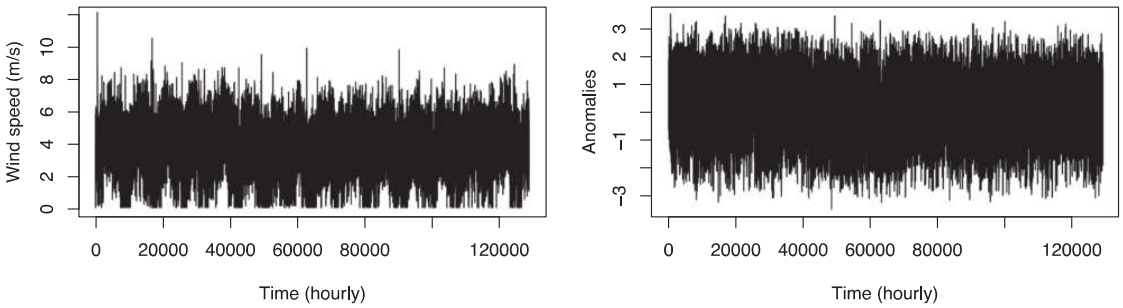


Fig. 8. On the left original wind speed series (in m/s) and on the right time series of wind speed anomalies from Petrolina-PE station.

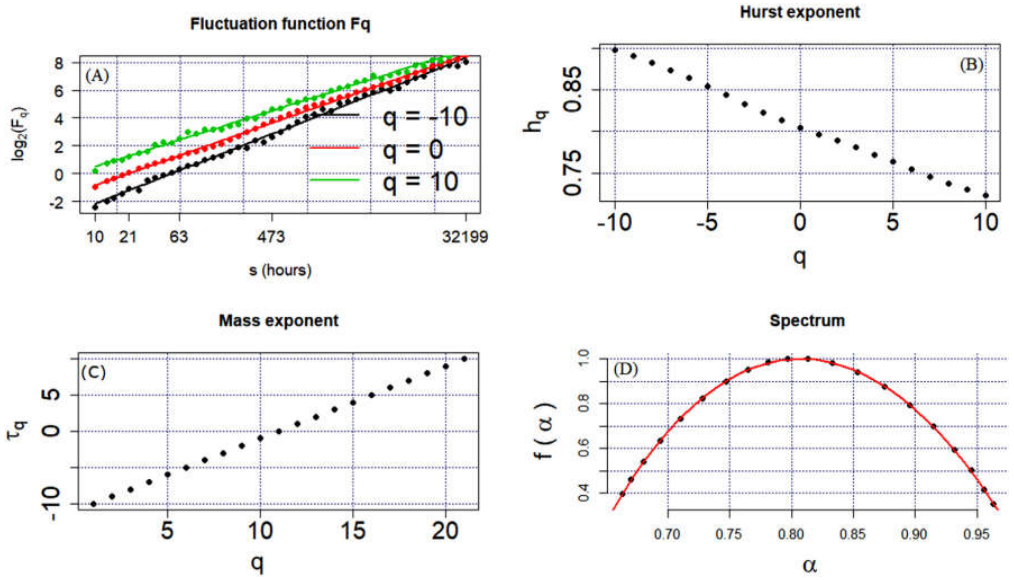


Fig. 9. Results obtained with MFDFA for the wind speed series from Petrolina-PE meteorological station: (A) Fluctuation functions for $q = -10$, $q = 0$ and $q = 10$; (B) Generalized Hurst exponent; (C) Rényi exponent; (D) Multifractal spectrum.

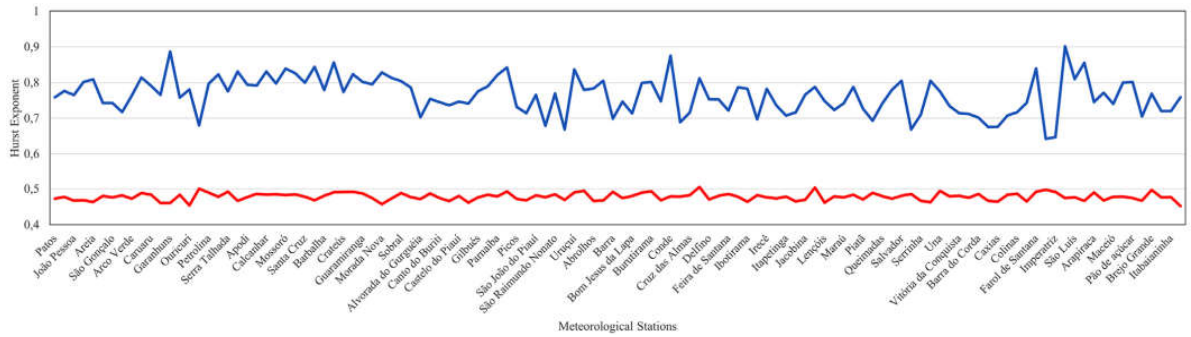


Fig. 10. Comparison of the values of the parameters α_{origi} (blue) and α_{rand} (red) for the wind speed series of the entire Brazilian Northeast region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

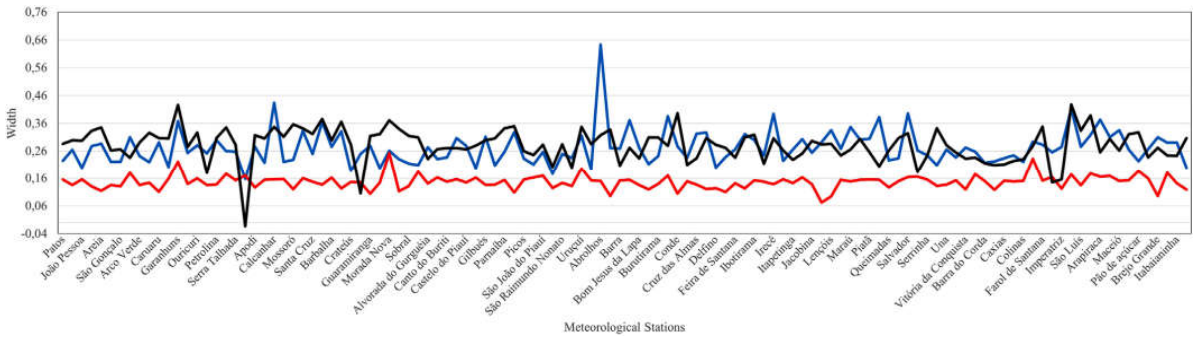


Fig. 11. Comparison of the values of the parameters w_{origi} (blue), w_{rand} (red) and $\Delta w = w_{orig} - w_{rand}$ (black) for the wind speed series of the entire Brazilian Northeast region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

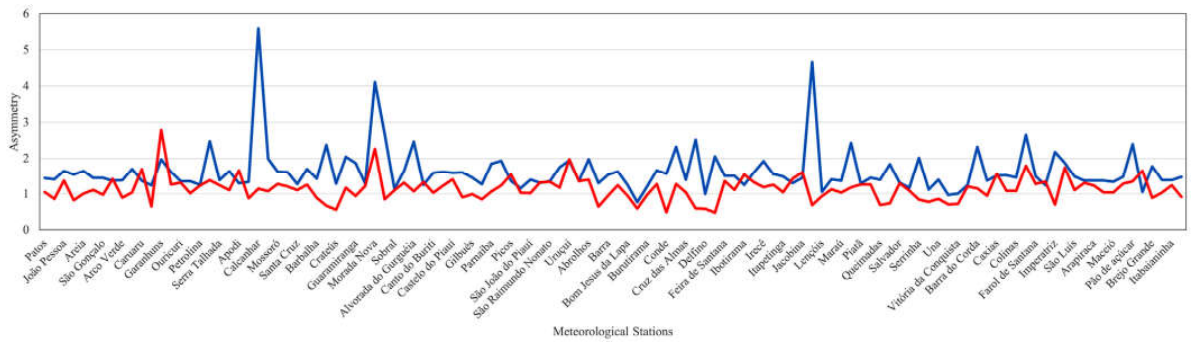


Fig. 12. Comparison of the values of the parameters r_{origi} (blue) and r_{rand} (red) for the wind speed series of the Northeast of Brazil. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

southeast portions of the state. In SE and AL, the averages largely ranged from 2.0 to 3.2 m/s.

On the other hand, in MA, the highest wind speeds were concentrated in the northeast coast of the state. Likely due to its proximity with the Amazon forest, the remainder of that state had mean speeds below 3.0 m/s, which is insufficient to interconnect the energy generation sub-systems. It was also found that the closer a site is to the equator, the higher the mean wind speed in the region, except when they are very close to the Amazon region. The standard deviation in the series indicate that the lowest data variability is found in MA, in southern BA, and center-north PI (see Fig. 2b).

In Fig. 3, we show the behaviors of all series of observations via boxplot, in which we visualize the variability of the data between the meteorological stations and the median value of wind speed (the numerical identification of each station can be found in Table 2 of the Appendix). Station 1 of Bahia (Abrolhos) presented a higher median than the other stations in the state. Like, station 3 of Rio Grande do Norte (Calcanhar) presented a higher median value among the other seasons. Overall, observing the entire Northeast region of Brazil, these two seasons were the only ones with medians above 5 m/s.

Fig. 4 presents the results of the boxplot for wind speed data over the 24 h of each day observed at the Petrolina and Recife stations, both located in PE. Overall, starting at 10 a.m. of each day,

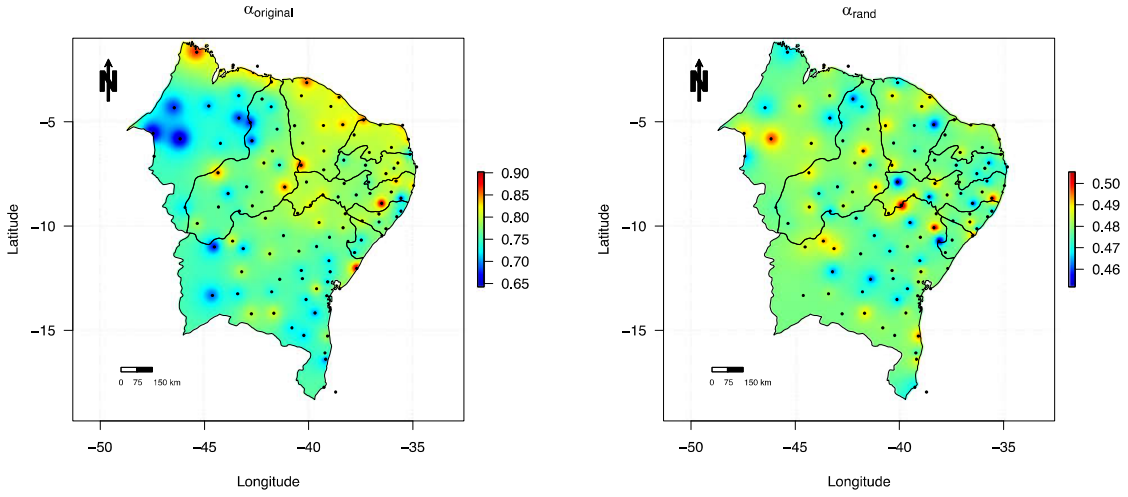


Fig. 13. On the left, parameter α_0 of the original series and on the right, α_0 of the randomized series in Northeast of Brazil.

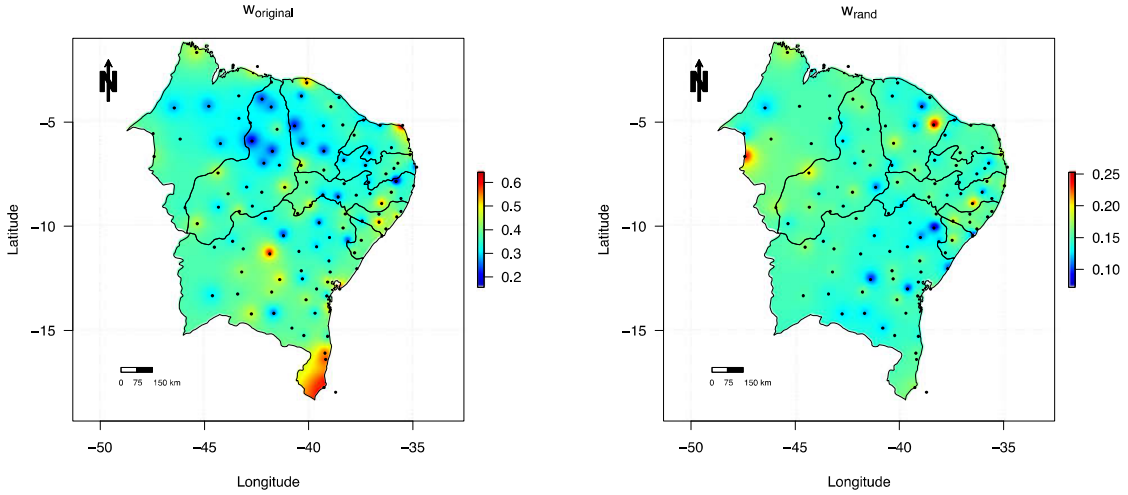


Fig. 14. On the left, parameter w of the original series and on the right, w of the randomized series in Northeast of Brazil.

the mean wind speed at the stations begins ramping up and, although values ramp down in the afternoon, the mean wind speed is higher until 8 p.m. than at other times. The variability in the series, assessed from their standard deviations (represented by blue dots in the plots), remained nearly constant over time and exhibited only slight oscillations. That shows electricity production from wind power may be best done between 10 a.m. and 8 p.m. of the days investigated. That was identified at all 136 meteorological stations, which featured different geographic conditions and climate characteristics. Likewise, the relationship between wind speed and temperature was explored in other studies, such as the peaks in temperature and cyclical components observed at different sites in Switzerland and China at different times of day [12,52,53].

Histograms were generated for all series studied and the Weibull-Weibull distribution, resulting from the mixture of a Weibull(α_1, β_1) distribution and a Weibull(α_2, β_2) distributions, with weight ρ fitted. An example of the results obtained is seen in Fig. 5 for the data of meteorological stations in Petrolina-PE and Recife-PE. It was found that the distribution adopted for the fit successfully modeled both unimodal and bimodal series. Similar results using the Weibull-Weibull distribution model were reported in other works [15,54]. The minimum and maximum values of shape parameters (α_1, α_2), scale parameters (β_1, β_2), and weights ρ_1 and ρ_2 obtained for each state of the Northeast region are shown in Table 1. The estimates of those values were obtained using the EM algorithm. Table 2 in the Appendix presents more details of the parameters estimates of the Weibull-Weibull distribution per meteorological station studied.

We collected new data to validate the study and adjusted the Weibull-Weibull distribution to the test series. We built a confidence interval at the significance level of 5% for the difference between the probability density functions of the training and test series. For example, for Petrolina-PE, the confidence interval [0.0015; 0.0107] obtained contains the value zero, that is, there is no statistically significant difference between the training and test distributions. Thus, statistically, the results found have analogous behaviors and reinforce the modeling performed in the research.

After estimating the parameters of the distribution in the fit to the databases from the stations studied, the information was spatialized for the entire Northeast region. That allows visualizing the

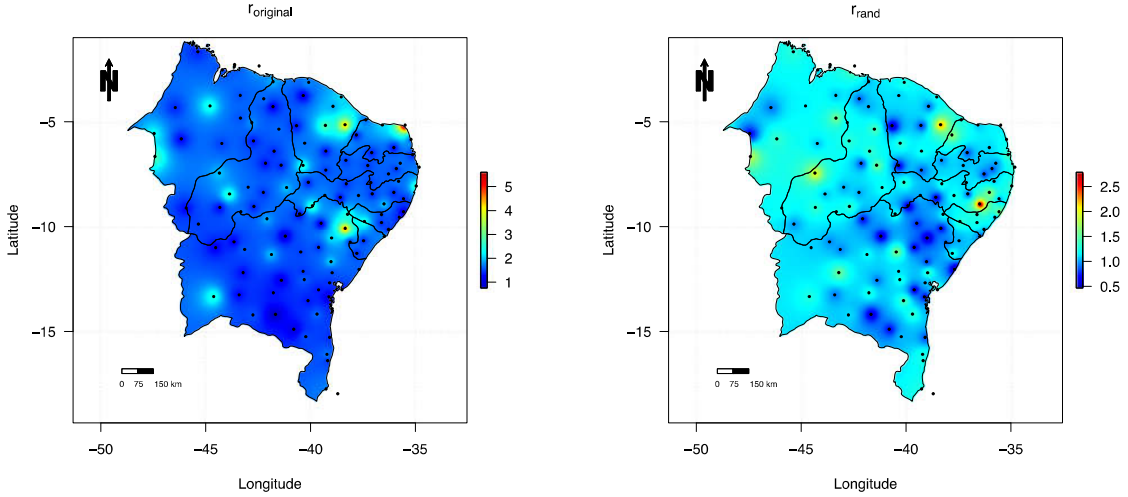


Fig. 15. On the left, parameter r of the original series and on the right, r of the randomized series in Northeast of Brazil.

Table 1
Estimated minimum and maximum parameters of the Weibull-Weibull distribution in each state of the Northeast.

States		β_1	β_2	α_1	α_2	ρ_1	ρ_2
Pernambuco	min	0.559974	1.060098	0.965443	1.151371	0.063402	0.249601
	max	4.530332	4.498235	4.571154	3.502933	0.750399	0.936598
Paraíba	min	1.504365	2.429077	1.207645	2.71571	0.100478	0.236674
	max	4.793033	4.717773	2.891606	6.182441	0.763326	0.899522
Alagoas	min	1.186569	1.360932	1.186569	1.360932	0.30000	0.515328
	max	2.212482	3.637992	2.212482	3.637992	0.484672	0.700000
Bahia	min	0.310714	1.522241	1.12348	1.259015	0.053154	0.320183
	max	6.898418	6.261342	4.274849	3.171728	0.679817	0.946846
Sergipe	min	0.322462	2.432078	1.219897	1.799703	0.123509	0.65512
	max	3.155649	4.497847	2.873872	3.052092	0.34488	0.876491
Maranhão	min	0.131801	1.375159	1.076328	1.272949	0.056881	0.619862
	max	2.851587	4.398315	3.07158	4.516304	0.502122	0.943119
Ceará	min	0.497569	2.031887	1.150285	1.806629	0.155091	0.297632
	max	3.228893	5.634987	2.103365	5.223665	0.702368	0.844909
Piauí	min	0.150473	1.376661	1.136813	1.323856	0.085073	0.425931
	max	3.185628	5.333431	2.411228	3.895098	0.574069	0.914927
Rio Grande do Norte	min	1.203313	2.231378	1.203313	2.231378	0.066031	0.631977
	max	4.047106	5.04857	4.047106	5.04857	0.368023	0.933969

sites exhibiting the highest and lowest estimated wind speeds with the scale parameters (β_1 and β_2), shape parameters (α_1 and α_2), and weight in the mixture Weibull-Weibull distribution model (ρ_1 and ρ_2). The plots resulting from the spatialization are presented in Figs. 6 and 7.

Only for the coastal region of Rio Grande do Norte we found a regard between altitude and the value of the parameter β_1: values higher than β_1 were found in regions with lower altitudes in RN. As for shape parameters, we observed that α_1 and α_2 may be related to the values of altitudes in this region. We found that the lowest values of α_1 and α_2 are concentrated in areas with lower altitudes, especially when we visualize the distribution of the α_1 parameter. For the weight parameters ρ_1 and ρ_2, we did not identify the relationship between them and the different altitudes in the studied region. Perhaps other features such as relief and air currents circulating the region may provide some direct relationship with these parameters.

Generally speaking, the wind speeds estimated using parameters β_1 and β_2 are higher in the coastal portion of RN, as expected from the real values observed presented in Fig. 2. The high wind speeds at that location were likely due to the influence of trade winds from the Atlantic Ocean blowing in a southeastward direction, according to the description of the main characteristics reported by Barros et al. [55] when studying the seasonality of the city of Natal, on the coast of RN. Such wind speeds in that region are a key reason for the state to be the current largest producer of wind energy in Brazil [51].

On the other hand, the state of MA, which is very close to the Amazon forest, exhibited the lowest estimated hourly wind speeds, a behavior expected given the proximity between the two regions. It is also seen that the maximum wind speed in the state of AL is 3.64 m/s, which confirms its low potential for wind energy generation.

Thus, based on the mean wind speed found for the Northeast region using parameter estimation of the Weibull-Weibull distribution, the states with the greatest potential for wind energy production are Rio Grande do Norte, Bahia, Pernambuco, Paraíba, Ceará, and Piauí. Of those, RN, BA, and CE are currently at the top of the ranking of Brazilian states in wind energy generation [56]. It is noteworthy that the northeastern portions of BA and PE, eastern PB, and RN coast feature wind behaviors that favor the installation of wind turbines and energy distribution.

Multifractal analysis was performed on wind speed records from 118 stations with more than 2 years of observations, using the MFDFA package [57] (with second degree polynomial, segment size from 10 to 32,199 h and the values of q ranging from −10 to 10) of the statistical software R [58]. In order to eliminate daily periodicity, we calculate wind speed anomalies using Eq. (14) [59].

$$\phi_{x_i} = \frac{(x_t - \bar{x}_t)}{\sigma}, \tag{14}$$

where x_i is the value observed at a given hour t, $\bar{x}_t$ is the average and σ is the standard deviation corresponding to the hour on which observation occurred. The results for the representative station Petrolina, located in Pernambuco state (with original and deseasonalized data shown on Fig. 8) are presented on Fig. 9.

It is seen from Fig. 9A that the fluctuation functions (Fq) versus segment size (s) exhibits linear behavior indicating the multifractality of wind speed series. The generalized Hurst exponent (hq) is a decreasing function (Fig. 9B), the Rényi exponent (τ_q) has a non-linear behavior (Fig. 9C) and the multifractal spectrum ($f(\alpha)$) shows concave shape downwards confirming that the hourly wind speed series from Petrolina station belongs to multifractal processes. It is also observed that $\alpha_0 > 0.5$, indicating overall persistent properties for the wind speed in the region during the studied period (see Fig. 9D).

To characterize the complexity of the multifractal process for all 118 available stations randomized series were generated according to the expression $rand = 20,000 * x_t$, where rand represents each randomized value and x_t corresponds to the value observed in the series at time t. For both, the original anomalies (origi) and randomized series (rand) the values of multifractal spectrum parameters, the position of maximum (α_0), width (w) and asymmetry (r) were calculated, as well as the difference in the spectrum width of the original and the randomized series ($\Delta w = w_{origi} - w_{rand}$). The results are shown in Figs. 10–12. Among all stations, Amargosa-BA (0.0959), Jeremoabo-BA (0.0721), Lençóis-BA (0.094), and Brejo Grande-SE (0.0954) stand out, for which w_{rand} values are very close to zero, indicating that long-range correlations are causes of multifractality in these wind speed series. On the other hand, for Surubim-PE station, $w_{rand} \approx w_{origi}$ ($\Delta w = -0.0129$), indicating the probability density function as a cause of multifractality in this series. For other stations the multifractality of randomized series is weaker than for original series, indicating that both, long-range correlations and probability distribution are causes of multifractality of the process.

With the values of the parameters α_0, w, r, and Δw for the original series and for the randomized ones, it was possible to perform the spatial interpolation for the entire Northeast of Brazil and generate the maps that can be viewed in Figs. 13–16.

It is observed from Fig. 13 that for the originals series the values of α_0 were greater than 0.5, indicating persistence in the wind speed series in the entire Northeast of Brazil. It is also possible to verify that the width of the spectrum (w) presented higher values in the south end of the map, with emphasis on Abrolhos-BA station for which $w \approx 0.64$. This indicates that in that region the multifractality is stronger in the process (see Fig. 14). High values of the width of the multifractal spectrum reflect greater stochastic complexity of the series and lower accuracy when making predictions in temporal observations. Thus, there is greater difficulty in making predictions of wind speed series for locations at the extreme south of Bahia state. Finally, it is seen from Fig. 15, that the multifractal spectrum shows an asymmetry to the right ($r > 1$) in a large part of the Northeast, indicating that small fluctuations contribute more to the multifractality of the process. The exceptions are Brumado-BA for which large fluctuations contribute more to multifractality of wind speed (left assymetry, $r \approx 0.76$), Delfino-BA ($r \approx 0.99$) and Valença-BA ($r \approx 0.97$) for which the multifractal spectrum is symmetric ($r \approx 1$), indicating equal contribution of both small and large fluctuations to multifractality of process (see Figs. 10–12 and 15). In general, the value of α_0 is varying from 0.64 to 0.91, the spectrum width from 0.15 to 0.65 and the asymmetry from 0.76 to 5.6. Similar results were obtained by Laib et al. [45] for 119 weather stations from Switzerland. Observing the different altitudes in the investigated region, we found stronger persistence in the coastal regions of the states of Ceará, Rio Grande do Norte and Maranhão, whose altitudes are among the lowest in northeastern Brazil. For parameters w and r, we did not find a relationship between them and the altitude values of northeastern Brazil.

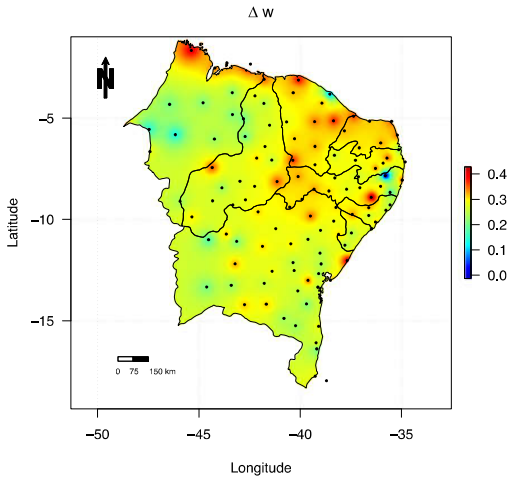


Fig. 16. Difference between w_{orig} and w_{rand} (Δw).

4. Conclusion

The behavior of wind regimens was investigated based on information of 14 years of observations from meteorological stations in the Northeast region of Brazil. The main conclusions of this study are:

- The coast of Rio Grande do Norte state (RN) exhibited the highest mean wind speed, followed by Paraíba (PB) and Pernambuco (PE) states. The lowest mean speed observed was found in Maranhão (MA) state. The lowest standard deviations were found in southern Bahia (BA), center-northern Piauí (PI), and much of Maranhão (MA);
- Between 10 a.m. and 8 p.m., the mean wind speed is the highest at all stations, with some well-defined cyclical components;
- By using the Expectation Maximization algorithm to optimize the parameters in the five-parameter Weibull distribution, it fit well to both unimodal data and bimodal data at all stations studied, showing it can be appropriately used to calculate estimated wind power;
- The spatialization of the parameters estimated in Weibull-Weibull allowed locating possible areas for the installation of wind turbines for electricity generation, characterized as having the highest mean wind speeds, such as in northeastern Bahia, eastern Paraíba, and northeastern Pernambuco, in addition to the Rio Grande do Norte coast;
- The wind speed series of the 118 stations examined with the Multifractal Detrended Fluctuation Analysis method showed persistent behavior over time. The complexity parameters of the multifractal spectrum showed that, the multifractality is

strongest in south region of Bahia state, and that for the entire Northeast of Brazil, small fluctuations contribute more to multifractality of the process;
- The results of multifractal analysis of randomized series showed that, except for the Amargosa-BA, Jeremoabo-BA, Lençóis-BA, Brejo Grande-SE, and Surubim-PE stations, both probability density function and long-range correlations are causes of multifractality in the wind speed series.

Such information may help identify seasonality patterns, besides predict wind potential for energy production in regions characterized as appropriate, convenient, and able to this type of use. The high wind energy potential makes its exploitation an important tool of social transformation, a vector of sustainable development, in addition to the possibility of insertion of wind energy together with the generation of energy by hydraulic/thermal source, significantly reducing the risks of unavailability of energy resources in the long term, especially in periods of longer droughts. Finally, the information generated serves as an important tool in the planning and decision-making of companies, management bodies, and the Government itself. Among the main areas benefited are agriculture, water resources, scientific studies, power generation, civil defense, transport, health, tourism, and leisure, among others.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work received funding and technical support from AES and Associated Companies (of the CPFL, Brookfield and Global group) as part of the ANEEL PD-0610-1004/2015 project, IRIS - Integration of Intermittent Renewables: A simulation model of the operation of the Brazilian electrical system to support planning, operation, commercialization, and regulation, which is part of an R&D program regulated by ANEEL, Brazil. The authors also thank IATI - Advanced Institute of Technology and Innovation for their support, the National Institute of Meteorology (INMET) for granting the data and the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financing Code 001, for support.

Appendix A

Table 2
Weibull-Weibull distribution parameters in each studied station.

Number	State	City	α_1	β_1	α_2	β_2	w_1	w_2
1	AL	Arapiraca	1.361002	1.860128	3.637992	4.410718	0.484672	0.515328
2	AL	Cururipe	1.997511	2.651583	1.997511	2.651583	0.3	0.7
3	AL	Maceió	2.212482	1.389615	3.306686	4.264418	0.43376	0.56624
4	AL	Palmeiras dos Índios	1.309499	1.779033	3.365652	3.997678	0.416994	0.583006
5	AL	Pão de Açúcar	1.186569	0.907616	2.729356	3.59838	0.331254	0.668746
6	AL	São Luís do Quitunde	1.360932	2.010956	1.360932	2.010956	0.3	0.7
1	BA	Abrolhos	2.375498	6.898418	2.469944	6.261342	0.301228	0.698772
2	BA	Amargosa	1.138864	1.16673	2.406377	3.652763	0.116959	0.883041
3	BA	Barra	1.419388	1.35684	2.339752	2.909727	0.427531	0.572469
4	BA	Barreiras	1.520375	1.908081	1.520375	1.908081	0.3	0.7
5	BA	Belmonte	1.483514	1.816563	1.483514	1.816563	0.3	0.7
6	BA	Bom Jesus da Lapa	1.318907	1.464866	2.522623	2.059686	0.382516	0.617484
7	BA	Brumado	1.426176	0.74462	2.320931	2.498592	0.31618	0.68382
8	BA	Buritirama	1.217464	0.972127	2.517363	3.552684	0.335387	0.664613
9	BA	Caravelas	1.49362	1.461031	2.34488	3.684235	0.330986	0.669014
10	BA	Conde	1.234157	1.064841	2.728441	3.779909	0.367956	0.632044
11	BA	Correntinha	3.027821	0.811623	2.090233	1.995246	0.236386	0.763614
12	BA	Cruz das Almas	1.360337	2.037314	2.392966	3.432033	0.1275	0.8725
13	BA	Curaçá	4.274849	4.009159	2.612726	3.508141	0.310044	0.689956
14	BA	Delfino	2.393844	1.262787	2.843464	3.571428	0.255689	0.744311
15	BA	Euclides da Cunha	2.281903	3.499871	2.564802	3.699729	0.281125	0.718875
16	BA	Feira de Santana	1.918934	1.521104	3.171728	3.505956	0.356036	0.643964
17	BA	Formoso do Rio Preto	1.377872	1.538132	1.377872	1.538132	0.3	0.7
18	BA	Guanambi	1.998138	4.458315	1.998138	4.458315	0.3	0.7
19	BA	Ibotirama	1.163259	1.808079	1.707159	2.897343	0.320815	0.679185
20	BA	Ilhéus	1.92145	2.612609	1.92145	2.612609	0.3	0.7
21	BA	Ipiaú	1.362948	1.588121	1.362919	1.596694	0.300001	0.699999
22	BA	Irecê	1.762847	0.310714	2.519272	3.34126	0.031339	0.968661
23	BA	Itaberaba	1.306211	1.638499	1.998534	3.141572	0.322552	0.677448
24	BA	Itamaraju	3.431148	3.814291	1.797364	2.152801	0.302785	0.697215
25	BA	Itapetinga	1.686045	2.472316	1.686045	2.472316	0.3	0.7
26	BA	Itiruçú	2.903056	3.103107	1.740374	2.563412	0.679817	0.320183
27	BA	Jacobina	1.85249	2.224183	1.85249	2.224183	0.3	0.7
28	BA	Jeremoabo	3.6458	4.109518	1.81983	2.405233	0.473761	0.526239
29	BA	Lençóis	2.307922	0.192138	1.636034	1.563554	0.106515	0.893485
30	BA	Luís Eduardo Magalhães	1.12348	1.952038	2.402721	3.61025	0.190699	0.809301
31	BA	Macajuba	1.129985	1.585205	2.10171	3.808834	0.190707	0.809293
32	BA	Macaúbas	1.919062	2.845817	1.979274	2.512856	0.30103	0.69897
33	BA	Maraú	1.464377	1.624356	2.696009	2.849761	0.326913	0.673087
34	BA	Paulo Afonso	2.627738	4.47758	2.627738	4.47758	0.3	0.7
35	BA	Piatã	1.319372	2.142884	2.421263	3.082899	0.243073	0.756927
36	BA	Pilão Arcado	2.250838	3.088288	2.250838	3.088288	0.3	0.7
37	BA	Porto Seguro	1.649587	0.588992	2.445269	2.345246	0.401503	0.598497
38	BA	Queimadas	2.172329	3.735228	2.172329	3.735228	0.3	0.7
39	BA	Remanso	1.486059	2.661509	2.78035	4.494104	0.428423	0.571577
40	BA	Salvador	2.782964	1.890932	1.518203	1.886781	0.573959	0.426041
41	BA	Santa Rita de Cássia	1.726166	1.890501	1.726122	1.890924	0.3	0.7
42	BA	Senhor do Bonfim	1.858513	2.994891	2.619747	3.5616	0.248703	0.751297
43	BA	Serrinha	1.818239	2.506236	3.086804	3.639633	0.402103	0.597897
44	BA	Uauá	1.294373	2.370793	2.912575	4.36742	0.053154	0.946846
45	BA	Una	1.355367	0.923997	1.939361	1.954935	0.346164	0.653836
46	BA	Valença	1.259015	1.522241	1.259015	1.522241	0.3	0.7
47	BA	Vitória da Conquista	2.13429	2.462333	2.957945	3.085342	0.333735	0.666265
1	CE	Acaraú	1.923043	2.930447	5.223665	5.634987	0.702368	0.297632
2	CE	Barbalha	1.806629	2.031887	1.806629	2.031887	0.3	0.7
3	CE	Campos Sales	1.150285	1.318324	2.757942	3.87106	0.253511	0.746489
4	CE	Crateús	2.103365	2.450825	2.103365	2.450825	0.3	0.7
5	CE	Fortaleza	1.443187	1.785075	3.020343	3.740161	0.301887	0.698113
6	CE	Guaramiranga	1.804959	3.063079	3.262508	3.763733	0.155091	0.844909
7	CE	Iguatu	1.880373	2.903189	1.881445	2.904571	0.3	0.7
8	CE	Jaguaribe	2.040087	3.228893	2.040087	3.228893	0.3	0.7
9	CE	Jaguaruana	1.200629	1.552842	2.613391	4.094945	0.249094	0.750906
10	CE	Morada Nova	1.275596	1.532838	2.760895	3.733637	0.311277	0.688723
11	CE	Quixeramobim	1.393703	0.497569	2.100248	3.274513	0.157483	0.842517
12	CE	Sobral	1.420848	1.885147	2.523791	3.856027	0.447294	0.552706
13	CE	Tauá	1.157571	1.443839	2.841617	3.99711	0.220544	0.779456
1	MA	Alto Parnaíba	1.395257	0.446617	2.01866	2.220045	0.380138	0.619862
2	MA	Bacabal	2.943047	0.152168	1.677186	1.862433	0.075647	0.924353
3	MA	Balsas	1.298371	2.046303	1.298371	2.046303	0.3	0.7
4	MA	Barra do Corda	1.212783	1.215002	2.289265	2.338432	0.502122	0.497878
5	MA	Buriticupu	1.625122	2.540452	1.625122	2.540452	0.3	0.7
5	MA	Carolina	1.272948	1.607364	1.272949	1.607364	0.3	0.7
6	MA	Caxias	1.318318	1.375159	1.318318	1.375159	0.3	0.7
7	MA	Chapadinha	2.010423	2.094673	2.010423	2.094673	0.3	0.7
8	MA	Colinas	1.97879	0.226093	1.567013	1.94253	0.123567	0.876433

(*continued on next page*)

Table 2 (*continued*)

Number	State	City	α_1	β_1	α_2	β_2	w_1	w_2
9	MA	Estreito	1.189088	0.973999	1.781245	2.558716	0.336706	0.663294
10	MA	Farol de Santana	2.085395	2.851587	4.516304	3.857066	0.189804	0.810196
11	MA	Grajaú	1.685672	1.768069	1.685672	1.768069	0.3	0.7
12	MA	Imperatriz	1.076328	0.99303	1.49608	1.761861	0.319573	0.680427
13	MA	Preguiças	1.153051	1.854552	3.575603	4.398315	0.108609	0.891391
14	MA	São Luís	1.29837	1.106898	3.097371	3.076234	0.336366	0.663634
15	MA	Turiaçú	3.07158	0.131801	2.162783	3.271396	0.056881	0.943119
1	PB	Areia	2.891606	4.793033	4.047802	4.717773	0.336706	0.663294
2	PB	Cabaceiras	1.301109	1.815377	3.486614	4.267541	0.28611	0.71389
3	PB	Camaratuba	2.010676	3.183325	6.182441	5.650085	0.763326	0.236674
4	PB	Campina Grande	1.49055	1.95856	3.187892	3.923041	0.100478	0.899522
5	PB	João Pessoa	2.511171	3.034667	2.794838	2.429077	0.332101	0.667899
6	PB	Monteiro	1.26609	1.675091	2.722817	4.090184	0.170055	0.829945
7	PB	Patos	1.2434	1.786132	2.71571	4.329456	0.150585	0.849415
8	PB	São Gonçalo	1.207645	1.504365	2.540796	2.737598	0.358182	0.641818
1	PE	Arco Verde	1.447448	2.867465	3.181265	4.013963	0.175061	0.824939
2	PE	Cabrobó	1.334876	0.559974	2.686999	4.377489	0.063402	0.936598
3	PE	Caruaru	1.541998	2.289491	3.502933	4.282608	0.361672	0.638328
4	PE	Floresta	2.391384	3.56342	2.391384	3.56342	0.3	0.7
5	PE	Garanhuns	0.965443	1.482901	2.489687	3.416215	0.183256	0.816744
6	PE	Ibimirim	2.219628	3.361993	2.219628	3.361993	0.3	0.7
7	PE	Ouricuri	4.571154	4.530332	2.082035	3.109644	0.314188	0.685812
8	PE	Palmares	1.477135	2.375182	1.477135	2.375182	0.3	0.7
9	PE	Petrolina	1.364484	2.487726	3.490777	4.179945	0.069061	0.930939
10	PE	Recife	2.530471	0.675868	2.697016	2.41101	0.26175	0.73825
11	PE	Salgueiro	3.029233	3.369175	1.151371	1.060098	0.750399	0.249601
12	PE	Serra Talhada	2.236253	2.905932	2.236253	2.905932	0.3	0.7
13	PE	Surubim	1.132969	1.647545	2.622215	4.498235	0.174066	0.825934
1	PI	Alvorada do Gurguéia	1.330104	1.375011	2.200349	2.948802	0.426036	0.573964
2	PI	Bom Jesus do Piauí	1.951208	0.225103	1.662523	2.587719	0.100305	0.899695
3	PI	Canto do Buriti	1.799856	3.185628	1.799856	3.185628	0.3	0.7
3	PI	Caracol	1.230657	1.523925	2.917669	4.187517	0.500044	0.499956
5	PI	Castelo do Piauí	1.736184	2.290221	1.736184	2.290221	0.3	0.7
6	PI	Esperantina	1.236877	1.339308	2.299955	2.086901	0.474529	0.525471
7	PI	Floriano	1.136813	0.930553	1.6361	2.185466	0.342166	0.657834
8	PI	Gilbués	1.728864	2.689923	1.728864	2.689923	0.3	0.7
9	PI	Oeiras	1.323856	2.196636	1.323856	2.196636	0.3	0.7
10	PI	Parnaíba	1.802202	2.8133	3.895098	5.333431	0.563427	0.436573
11	PI	Paulistana	1.281114	1.809216	2.920173	4.451634	0.182392	0.817608
12	PI	Picos	1.794231	2.604822	1.794231	2.604822	0.3	0.7
13	PI	Piripiri	1.310893	1.645811	1.854913	2.332707	0.313746	0.686254
14	PI	São João do Piauí	1.366669	1.537622	3.148349	3.38832	0.447285	0.552715
15	PI	São Pedro do Piauí	2.411228	2.440907	2.978614	1.376661	0.573582	0.426418
16	PI	São Raimundo Nonato	1.392952	1.596405	3.186823	3.509533	0.574069	0.425931
17	PI	Teresina	2.166081	0.207841	1.608906	1.750046	0.11113	0.88887
18	PI	Uruçuí	2.95669	0.150473	1.502236	1.815776	0.085073	0.914927
19	PI	Valença do Piauí	1.804591	2.074498	1.804591	2.074498	0.3	0.7
1	RN	Apodi	1.203313	0.605652	2.231378	3.782606	0.066031	0.933969
2	RN	Caicó	2.382226	3.478441	2.382957	3.478966	0.3	0.7
3	RN	Calcanhar	4.047106	4.572036	5.04857	8.503106	0.228696	0.771304
4	RN	Macau	2.511187	5.253387	2.51119	5.253392	0.3	0.7
4	RN	Mossoró	2.057301	1.994142	3.784013	5.049751	0.368023	0.631977
5	RN	Natal	3.587064	5.827875	3.217895	4.548699	0.322727	0.677273
6	RN	Santa Cruz	2.051828	1.302669	3.151695	3.405675	0.237819	0.762181
1	SE	Aracajú	1.805514	2.477546	2.996538	4.142791	0.34488	0.65512
2	SE	Brejo Grande	1.623168	0.322462	2.150305	2.432078	0.123509	0.876491
3	SE	Carira	1.292074	2.016623	3.052092	4.6831	0.276301	0.723699
4	SE	Itabaiana	1.799705	3.155649	1.799703	3.155644	0.3	0.7
5	SE	Itabaianinha	2.231012	3.11913	2.231012	3.11913	0.3	0.7
6	SE	Nossa Senhora da Glória	2.873872	3.100695	2.873872	3.100695	0.3	0.7
7	SE	Poço Verde	1.219897	1.826628	2.791558	4.497847	0.201771	0.798229

References

[1] Silva AM, Vieira RMF. Energia eólica: conceitos e características basilares para uma possível suplementação da matriz energética brasileira. Rev Direito Ambient Soc 2016;6(2):53–76.

[2] Energia eólica salva abastecimento do Nordeste JC Online. 2017. Acessed on 2019-11-12; URL http://abeeolica.org.br/noticias/3696/.

[3] O. N. do Sistema Elétrico ONS. Boletim Mensal de Geração Eólica - Agosto/2019. 2019.

[4] Wang J, Huang X, Li Q, Ma X. Comparison of seven methods for determining the optimal statistical distribution parameters: A case study of wind energy assessment in the large-scale wind farms of China. Energy 2018;164:432–48. https://doi.org/10.1016/j.energy.2018.08.201.

[5] Sumair M, Aized T, Gardezi SAR, ur Rehman SU, Rehman SMS. A newly proposed method for weibull parameters estimation and assessment of wind potential in southern punjab. Energy Rep 2020;6:1250–61. https://doi.org/10.1016/j.egyr.2020.05.004.

[6] Mohammadi K, Mostafaeipour A. Using different methods for comprehensive study of wind turbine utilization in Zarrineh, Iran. Energy Convers Manag 2013;65:463–70. https://doi.org/10.1016/j.enconman.2012.09.004.

[7] Mahmood FH, Resen AK, Khamees AB. Wind characteristic analysis based on weibull distribution of al-salman site, iraq. Energy Rep 2020;6:79–87. https://doi.org/10.1016/j.egyr.2019.10.021.

[8] Allouhi A, Zamzoum O, Islam M, Saidur R, Kousksou T, Jamil A, et al. Evaluation of wind energy potential in Morocco's coastal regions. Renew Sustain Energy Rev 2017;72:311–24. https://doi.org/10.1016/j.rser.2017.01.047.

[9] Parajuli A. A statistical analysis of wind speed and power density based on Weibull and Rayleigh models of Jumla, Nepal. Energy Power Eng 2016;8(7):271. https://doi.org/10.4236/epe.2016.87026.
[10] Olaofe ZO. Assessment of the offshore wind speed distributions at selected stations in the South-West Coast, Nigeria. Int J Renew Energy Res 2017;7(2):565–77.
[11] Safari B. Modeling wind speed and wind power distributions in Rwanda. Renew Sustain Energy Rev 2011;15(2):925–35. https://doi.org/10.1016/j.rser.2010.11.001.
[12] Telesca L, Lovallo M, Kanevski M. Power spectrum and multifractal detrended fluctuation analysis of high-frequency wind measurements in mountainous regions. Appl Energy 2016;162:1052–61. https://doi.org/10.1016/j.apenergy.2015.10.187.
[13] Zhou J, Erdem E, Li G, Shi J. Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites. Energy Convers Manag 2010;51(7):1449–58. https://doi.org/10.1016/j.enconman.2010.01.020.
[14] Ouarda TB, Charron C. On the mixture of wind speed distribution in a Nordic region. Energy Convers Manag 2018;174:33–44. https://doi.org/10.1016/j.enconman.2018.08.007.
[15] Rajapaksha K, Perera K. Wind speed analysis and energy calculation based on mixture distributions in Narakkalliya, Sri Lanka. J Natl Sci Found Sri Lanka 2016;44(4). http://doi.org/10.4038/jnsfsr.v44i4.8023.
[16] Shin J-Y, Ouarda TB, Lee T. Heterogeneous mixture distributions for modeling wind speed, application to the UAE. Renew Energy 2016;91:40–52. https://doi.org/10.1016/j.renene.2016.01.041.
[17] Wang Q. Multifractal characterization of air polluted time series in China. Phys A: Statistical Mechanics and its Applications 2019;514:167–80. https://doi.org/10.1016/j.physa.2018.09.065.
[18] Júnior SFAX, Stosic T, Stosic B, Jale JdS, Xavier ÉFM. A Brief multifractal analysis of rainfall dynamics in Piracicaba, São Paulo, Brazil. Acta Sci Technol 2018;40:e35116. https://doi.org/10.4025/actascitechnol.v40i1.35116.
[19] Laib M, Telesca L, Kanevski M. Periodic fluctuations in correlation-based connectivity density time series: Application to wind speed-monitoring network in Switzerland. Phys A: Statistical Mechanics and its Applications 2018;492:1555–69. https://doi.org/10.1016/j.physa.2017.11.081.
[20] Laib M, Telesca L, Kanevski M. Long-range fluctuations and multifractality in connectivity density time series of a wind speed monitoring network. Chaos: An Interdisciplinary Journal of Nonlinear Science 2018;28(3):033108. https://doi.org/10.1063/1.5022737.
[21] Telesca L, Lovallo M. Analysis of the time dynamics in wind records by means of multifractal detrended fluctuation analysis and the Fisher–Shannon information plane. J Stat Mech: Theory and Experiment 2011;2011(07):P07001. https://doi.org/10.1088/1742-5468/2011/07/P07001.
[22] Balkissoon S, Fox N, Lupo A. Fractal characteristics of tall tower wind speeds in missouri. Renew Energy 2020. https://doi.org/10.1016/j.renene.2020.03.021.
[23] Santos J, Moreira D, Moret M, Nascimento E. Analysis of long-range correlations of wind speed in different regions of Bahia and the Abrolhos Archipelago, Brazil. Energy 2019;167:680–7. https://doi.org/10.1016/j.energy.2018.11.015.
[24] Rocha PAC, de Sousa RC, de Andrade CF, da Silva MEV. Comparison of seven numerical methods for determining Weibull parameters for wind energy generation in the northeast region of Brazil. Appl Energy 2012;89(1):395–400. https://doi.org/10.1016/j.apenergy.2011.08.003.
[25] Torres Silva dos Santos A, e Silva S, Moisés C. Seasonality, interannual variability, and linear tendency of wind speeds in the Northeast Brazil from 1986 to 2011. Sci World J 2013;2013. https://doi.org/10.1155/2013/490857.
[26] 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.
[27] Bilal M, Araya G, Birkelund Y. Preliminary assessment of remote wind sites. Energy Procedia 2015;75:658–63.
[28] Rind D. Complexity and climate. Science 1999;284(5411):105–7. https://doi.org/10.1126/science.284.5411.105.
[29] Li H-N, Zheng X-W, Li C. Copula-based joint distribution analysis of wind speed and direction. J Eng Mech 2019;145(5):04019024. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001600.
[30] Hu Q, Wang Y, Xie Z, Zhu P, Yu D. On estimating uncertainty of wind energy with mixture of distributions. Energy 2016;112:935–62. https://doi.org/10.1016/j.energy.2016.06.112.
[31] Qin X, Zhang J-s, Yan X-d. Two improved mixture Weibull models for the analysis of wind speed data. J Appl Meteorol Climatol 2012;51(7):1321–32. https://doi.org/10.1175/JAMC-D-11-0231.1.
[32] Arellano-Valle RB, Ferreira CS, Genton MG. Scale and shape mixtures of multivariate skew-normal distributions. J Multivar Anal 2018;166:98–110. https://doi.org/10.1016/j.jmva.2018.02.007.
[33] Bracale A, Carpinelli G, De Falco P. A new finite mixture distribution and its expectation-maximization procedure for extreme wind speed characterization. Renew Energy 2017;113:1366–77. https://doi.org/10.1016/j.renene.2017.07.012.
[34] Nguyen HD, Wang D, McLachlan GJ. Randomized mixture models for probability density approximation and estimation. Inf Sci 2018;467:135–48. https://doi.org/10.1016/j.ins.2018.07.056.
[35] Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE. Multifractal detrended fluctuation analysis of nonstationary time series. Phys A: Statistical Mechanics and its Applications 2002;316(1-4):87–114. https://doi.org/10.1016/S0378-4371(02)01383-3.
[36] Baranowski P, Gos M, Krzyszczak J, Siwek K, Kieliszek A, Tkaczyk P. Multifractality of meteorological time series for Poland on the base of MERRA-2 data. Chaos Solitons Fractals 2019;127:318–33. https://doi.org/10.1016/j.chaos.2019.07.008.
[37] Kalamaras N, Philippopoulos K, Deligiorgi D, Tzanis C, Karvounis G. Multifractal scaling properties of daily air temperature time series. Chaos Solitons Fractals 2017;98:38–43. https://doi.org/10.1016/j.chaos.2017.03.003.
[38] Grech D. Alternative measure of multifractal content and its application in finance. Chaos Solitons Fractals 2016;88:183–95. https://doi.org/10.1016/j.chaos.2016.02.017.
[39] Gajardo G, Kristjanpoller W. Asymmetric multifractal cross-correlations and time varying features between Latin-American stock market indices and crude oil market. Chaos Solitons Fractals 2017;104:121–8. https://doi.org/10.1016/j.chaos.2017.08.007.
[40] Zhang X, Yang L, Zhu Y. Analysis of multifractal characterization of Bitcoin market based on multifractal detrended fluctuation analysis. Phys A: Statistical Mechanics and its Applications 2019;523:973–83. https://doi.org/10.1016/j.physa.2019.04.149.
[41] Stosic D, Stosic D, de Mattos Neto PS, Stosic T. Multifractal characterization of Brazilian market sectors. Phys A: Statistical Mechanics and its Applications 2019;525:956–64. https://doi.org/10.1016/j.physa.2019.03.092.
[42] Kumar S, Cuntz M, Musielak ZE. Fractal and multifractal analysis of the rise of oxygen in Earths early atmosphere. Chaos Solitons Fractals 2015;77:296–303. https://doi.org/10.1016/j.chaos.2015.06.007.
[43] Zhang X, Liu H, Zhao Y, Zhang X. Multifractal detrended fluctuation analysis on air traffic flow time series: A single airport case. Phys A: Statistical Mechanics and its Applications 2019:121790. https://doi.org/10.1016/j.physa.2019.121790.
[44] de Benicio RB, Stošić T, de Figueirêdo P, Stošić BD. Multifractal behavior of wild-land and forest fire time series in Brazil. Phys A: Statistical Mechanics and its Applications 2013;392(24):6367–74. https://doi.org/10.1016/j.physa.2013.08.012.
[45] Laib M, Golay J, Telesca L, Kanevski M. Multifractal analysis of the time series of daily means of wind speed in complex regions. Chaos Solitons Fractals 2018;109:118–27. https://doi.org/10.1016/j.chaos.2018.02.024.
[46] Delbianco F, Tohmé F, Stosic T, Stosic B. Multifractal behavior of commodity markets: Fuel versus non-fuel products. Phys A: Statistical Mechanics and its Applications 2016;457:573–80. https://doi.org/10.1016/j.physa.2016.03.096.
[47] Shimizu Y, Thurner S, Ehrenberger K. Multifractal spectra as a measure of complexity in human posture. Fractals 2002;10(01):103–16. https://doi.org/10.1142/S0218348X02001130.
[48] Zaki MFM, Ismail MAM, Govindasamy D, Zainalabidin MH. Interpretation and development of top-surface grid in subsurface ground profile using inverse distance weighting (IDW) method for twin tunnel project in Kenny Hill Formation. Bull Geol Soc Malays 2019;67:91–7. https://doi.org/10.7186/bgsm67201911.
[49] Ballarin F, D'Amario A, Perotto S, Rozza G. A POD-selective inverse distance weighting method for fast parametrized shape morphing. Int J Numer Methods Eng 2019;117(8):860–84. https://doi.org/10.1002/nme.5982.
[50] Ozelkan E, Chen G, Ustundag BB. Spatial estimation of wind speed: a new integrative model using inverse distance weighting and power law. Int J Digit Earth 2016;9(8):733–47. https://doi.org/10.1080/17538947.2015.1127437.
[51] Eólica já é a segunda fonte da matriz elétrica brasileira com 15 GW de capacidade instalada. 2019a. Acessed on 2019-10-12; URL http://abeeolica.org.br/noticias/eolica-ja-e-a-segunda-fonte-da-matriz-eletrica-brasileira-com-15-gw-de-capacidade-instalada/.
[52] Kruyt B, Lehning M, Kahl A. Potential contributions of wind power to a stable and highly renewable Swiss power supply. Appl Energy 2017;192:1–11. https://doi.org/10.1016/j.apenergy.2017.01.085.
[53] He Y, Chan P, Li Q. Wind characteristics over different terrains. J Wind Eng Ind Aerodyn 2013;120:51–69. https://doi.org/10.1016/j.jweia.2013.06.016.
[54] Carta J, Ramirez P. Analysis of two-component mixture Weibull statistics for estimation of wind speed distributions. Renew Energy 2007;32(3):518–31. https://doi.org/10.1016/j.renene.2006.05.005.
[55] Barros JD, Furtado MLS, Costa AMdB, Marinho GS, da Silva FM. Sazonalidade do vento na cidade de Natal/RN pela distribuição de Weibull. Soc Territ 2013;25(2):78–92.
[56] Gerao de energia elica cresce 15% em 2018. 2019b. Acessed on 2019-02-22; URL http://abeeolica.org.br/noticias/geracao-de-energia-eolica-cresce-15-em-2018-2/.
[57] Laib M., Telesca L., Kanevski M.. Mfdfa: multifractal detrended fluctuation analysis for time series(2017) R package version 01 0
[58] R CORE TEAM. R: A language and environment for statistical computing. 2012. URL https://www.r-project.org/.
[59] Kantelhardt JW, Koscielny-Bunde E, Rybski D, Braun P, Bunde A, Havlin S. Long-term persistence and multifractality of precipitation and river runoff records. J Geophys Res: Atmospheres 2006;111(D1). https://doi.org/10.1029/2005JD005881.

Chapter 3

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

Nascimento, KKF, Santos, FS, Jale, JS & Ferreira, TAE. (2020). Comparison of methods and distribution models for the modeling of wind speed data in the municipality of Petrolina, Northeast Brazil. ***Research, Society and Development***, 9(X): 1-00, **eXX**

Comparison of methods and distribution models for the modeling of wind speed data in the municipality of Petrolina, Northeast Brazil

Comparação de métodos e modelos de distribuição para a modelagem de dados de velocidade do vento no município de Petrolina, Nordeste brasileiro

Comparación de métodos y modelos de distribución para el modelado de datos de velocidad eólica en el municipio de Petrolina, noreste de Brasil

Received: 04/05/2020 | Revised: 04/05/2020 | Accepted: 06/05/2020 | Published: 12/05/2020

Kerolly Kedma Felix do Nascimento
ORCID: https://orcid.org/0000-0001-7360-0961
Universidade Federal Rural de Pernambuco, Brasil
E-mail: kerollyfn@gmail.com
Fábio Sandro dos Santos
ORCID: https://orcid.org/0000-0002-0135-4981
Universidade Federal Rural de Pernambuco, Brasil
E-mail: fabio.sandropb@gmail.com
Jader da Silva Jale
ORCID: https://orcid.org/0000-0001-7414-1154
Universidade Federal Rural de Pernambuco, Brasil
E-mail: jsj_ce@yahoo.com.br
Tiago Alessandro Espínola Ferreira
ORCID: https://orcid.org/0000-0002-2131-9825
Universidade Federal Rural de Pernambuco, Brasil
E-mail: taef.first@gmail.com

Abstract

The identification of the probability distribution model that provides the best fit to the wind speed databases is necessary for defining investment and developing projects about the wind potential of different locations. For this, the estimation of the parameters of the models is essential in this process. The aim of this study is to investigate among the distribution models

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

and methods for estimating their respective parameters with better modeling in the literature which of them provides better fit to the wind speed data of Petrolina-PE. Through the case study, of quali-quanti nature, the adjustment of the Moment Method, the Estimation of Maximum Likelihood and the Particle Swarm Optimization (PSO) algorithm with Weibull were evaluated in this work, as well as the PSO with the Lognormal-Weibull and Weibull-Weibull distributions to the historical series of information. The results, investigated with the $RMSE$, R^2 and χ^2 error measures and by verifying the percentage of correctness between the theoretical and sample quantiles, demonstrated a better modeling of the Lognormal-Weibull distribution model with the PSO algorithm to the historical speed series of the wind. Thus, from the determination of the best distribution model that fits the data in the region, it may be possible to generate estimated wind speed series for areas where these historical series do not exist.

Keywords: Weibull; Lognormal; MM; EMV; PSO; Adjustment.

Resumo

A identificação do modelo de distribuição de probabilidade que forneça o melhor ajuste às bases de dados de velocidade do vento é necessária para definição de investimento e desenvolvimento de projetos acerca do potencial eólico de diversas localidades. Para isso, a estimativa dos parâmetros dos modelos é essencial nesse processo. O objetivo deste estudo é investigar dentre os modelos de distribuição e métodos para estimativa de seus respectivos parâmetros com melhor modelagem na literatura qual deles fornece melhor ajuste aos dados de velocidade do vento de Petrolina-PE. Através do estudo de caso, de natureza quali-quanti, foram avaliados neste trabalho o ajuste do Método dos Momentos, da Estimação de Máxima Verossimilhança e do algoritmo *Particle Swarm Optimization* (PSO) com a Weibull, bem como o PSO com as distribuições Lognormal-Weibull e Weibull-Weibull à série histórica de informações. Os resultados, investigados com as medidas de erro $RMSE$, R^2 e χ^2 e pela verificação da porcentagem de acerto entre os quantis teóricos e amostrais, demonstraram melhor modelagem do modelo de distribuição Lognormal-Weibull com o algoritmo PSO à série histórica de velocidade do vento. Dessa maneira, através da determinação do melhor modelo de distribuição que se ajuste aos dados na região, pode ser possível gerar séries de velocidade do vento estimadas para áreas onde não existem essas séries históricas.

Palavras-chave: Weibull; Lognormal; MM; EMV; PSO; Ajuste.

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

Resumen

La identificación del modelo de distribución de probabilidad que proporciona el mejor ajuste a las bases de datos de velocidad del viento es necesario para definir la inversión y desarrollar proyectos sobre el potencial eólico de diferentes ubicaciones. Para esto, la estimación de los parámetros de los modelos es esencial en este proceso. El objetivo de este estudio es investigar entre los modelos y métodos de distribución para estimar sus respectivos parámetros con un mejor modelado en la literatura que de ellos proporciona un mejor ajuste a los datos de velocidad del viento de Petrolina-PE. A través del estudio de caso, de naturaleza quali-quanti, el ajuste del Método momento, la Estimación de Máxima Probabilidad y el algoritmo de Optimización de Enjambre de Partículas (PSO) con Weibull fueron evaluados en este trabajo, así como el PSO con las distribuciones Lognormal-Weibull y Weibull-Weibull a la serie histórica de información. Los resultados, investigados con las medidas de error $RMSE$, R^2 y χ^2 y al verificar el porcentaje de corrección entre los cuantiles teóricos y de muestra, demostraron un mejor modelado del modelo de distribución Lognormal-Weibull con el algoritmo PSO a la serie de velocidad histórica del viento. Por lo tanto, a partir de la determinación del mejor modelo de distribución que se ajuste a los datos de la región, puede ser posible generar series estimadas de velocidad del viento para áreas donde estas series históricas no existen.

Palabras clave: Weibull; Lognormal; MM; EMV; PSO; Ajuste.

1. Introduction

Wind energy is an important energy source in the replacement of sources obtained by conventional and exhausting resources. According to ABEEolica, it is expected that by 2023 the production of wind energy on Brazilian soil will reach the mark of almost 20 GW of installed capacity (ABEEolica, 2019), reducing the consumption of fossil fuels to generate this energy source.

An advantage in the case of the substitution of water by winds in strengthening the energy matrix is the greater use of the water reserve of reservoirs for human consumption, animal and irrigation, among others, especially in regions affected by scarcity of rains and large droughts.

However, before the installation of wind farms, investigations are needed on the potential to obtain energy from winds in localities that have conditions considered favorable to this end (dos Santos et al., 2019). For this, the modeling of wind speed is being performed

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

through the Probability Density Functions in the adjustment of distribution models to the databases.

In the literature, several probabilistic distributions have been used (Gamma, Raileigh, Log-Normal, Logística, Burr), especially the Weibull distribution (Rocha et al., 2018) and the mixtures of distributions, such as Lognormal-Weibull e Weibull-Weibull (Rajapaksha & Perera, 2016).

To obtain the parameters of the model that fits the data it is necessary to choose the methods that perform the best estimates, such as the Moment Method (MM) and the Maximum Likelihood Estimation (EMV) or optimization algorithms such as Particle Swarm Optimization (PSO), in order to minimize the estimation errors obtained with traditional methods. In this choice, the performance of the adjustment of the parameters can be evaluated according to the values of the Coefficient of Determination (R^2) and the Mean Quadratic Error of The Residue (RMSE), as well as the Chi-square test statistic (χ^2).

Thus, the objective of this work is to compare the adjustments made with MM, EMV, and PSO and to seek the values for Weibull parameters that allow lower errors in the estimation of the parameters of this distribution.

The mixtures of Lognormal-Weibull and Weibull-Weibull distributions were observed with optimization with the PSO algorithm. For the comparison, we used hourly historical series of wind speed of Petrolina-PE, due to the need to reduce water consumption in the vicinity of this region for energy production purposes, since the locality requires aquifer reserves for irrigation of different fruit trees predominant in the surroundings (Melo et al., 2014).

The rest of this article is structured as follows. Section 2 describes the region investigated and the database used. Section 3 is explored the description of the distribution models Weibull, Weibull-Weibull, and Lognormal-Weibull. Section 4 presents the Moment Method and the Maximum Likelihood Estimation Method. Section 5 sets out a description of the Particle Swarm Optimization algorithm. Section 6 shows statistical estimates adopted in this article. In section 7 the results and discussions are covered. Finally, section 8 finds the final considerations of the article.

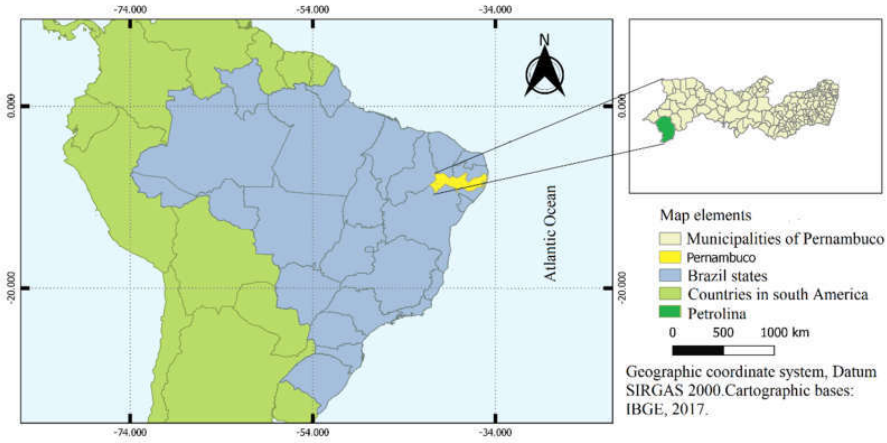
2. Area of Study and Data Collection

The development of this study was carried out in the municipality of Petrolina, located in the Brazilian Semiarid, to the extreme west of the state of Pernambuco, between the states of Bahia and Piauí (Jatobá et al., 2017), under the geographic coordinates of latitude -9.38832 and

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

-40.5233 (see Figure 1). Rainfall is concentrated in three to four months of the year and temperatures range from 18.7 º C to 33.6 º C. (da Silva et al., 2017). According to the last census conducted, the municipality has 293,962 inhabitants and a population density of 64.44 $inhabitants/km^2$ (IBGE, 2019).

Figure 1. The geographical position of the municipality of Petrolina in the upper Sertão of Pernambuco.



Source: Prepared by the authors.

The hourly wind speed observations used in this research were obtained from the National Institute of Meteorology - INMET, from 21/02/2003 to 09/30/2018. Such observations were collected from the automatic meteorological station located at the geographic coordinates of longitude -40.367 and latitude -9.150, at a height of 10 *m* and altitude of 366 *m*, in the city of Petrolina-PE. For the analyses, the daily means were evaluated, and all calculations were made in the R software.

According to Pereira et al. (2018), this study is a case study, in which a quali-quanti method was applied, in which the qualitative results reinforce the numerical ones, complementing them.

3. Distribution Models

Knowledge about the distribution model that best characterizes the behavior of the wind

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

regime is fundamental in the evaluation of wind potential in a locality. In this sense, Weibull (W) is one of the most used distributions, especially due to the simplicity necessary to estimate the parameters of its model and the good adherence of the model to the different wind speed databases. Other distributions also applied to this end are the mixtures of Weibull-Weibull (WW) and Lognormal-Weibull (LNW) distributions, which present good adjustments to bimodal series. The mathematical expressions of the probability density functions of the Weibull, Weibull-Weibull and Lognormal-Weibull distribution model, with their respective parameters are found in Table 1 (Rajapaksha & Perera, 2016).

Table 1. Mathematical equations of the Weibull, Weibull-Weibull and Lognormal-Weibull distribution models.

Distribution	Equation	Parameters
Weibull	$f_W(v;\alpha,\beta)=\frac{\alpha}{\beta}\left(\frac{v}{\beta}\right)^{\alpha-1}e^{-\left(\frac{v}{\beta}\right)^{k}}$	$(\alpha\ ,\beta)$
Weibull-Weibull	$f_{WW}(v;w,\alpha_1,\beta_1,\alpha_2,\beta_2)=wf(v;\alpha_1,\beta_1)+(1-w)f(v;\alpha_2,\beta_2)$	$(w,\alpha_1,\beta_1,\alpha_2,\beta_2)$
Lognormal-Weibull	$f_{LNW}(v;w,\lambda,\phi,\alpha,\beta)=wl(v;\lambda,\phi)+(1-w)f(v;\alpha,\beta)$	$(w,\lambda,\phi,\alpha,\beta)$

Source: Prepared by the authors.

in which v corresponds to the observation of wind speed, α is the shape parameter, β is the scale, w is the weight of the mixture of distributions, λ is the average and ϕ is the standard deviation, being $l(v;\lambda,\phi)=\frac{1}{v\,\phi\sqrt{2\pi}}exp\left[\frac{-(ln(v)-\lambda^2)}{2\phi^2}\right]$.

4. Numerical Methods for Estimating Parameters

In the Moment Method, the estimation of population parameters occurs through an iterative process based on the sample and theoretical (population) moments of the random variables, equaling them (de Souza et al., 2019), that is,

$$m_n = \mu_n, n = 1,2,\ldots,r, \quad (1)$$

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

in which m_n are the sampling moments, and μ_n are the theoretical moments.

While Maximum Likelihood Estimation is performed from the product of the probability density functions (continuous random variables) or from the probability functions (discrete random variables) of the series observations, provided that each random variable event is independent (Seckin et al., 2010). It is a methodology widely used in the literature for the estimation of parameters. Particularly, to estimate the parameters $\hat{\theta}$ the Weibull distribution used in the modeling of wind speed data, can be used the log-likelihood function given by (Ouarda et al., 2016).

$$lnL = ln\left(\prod_{i=1}^{n} f_{\hat{\theta}}(v_i)\right) \tag{2}$$

where n is the sample size, and v_i is every wind speed observation in the instant i.

5. Metaheuristic Optimization for Estimating Models Parameters

The Particle Swarm Optimization algorithm was originally developed by Kennedy and Eberhart and is a population-based stochastic investigation procedure. Every possible solution in search space is called a particle. All particles move Iteratively throughout optimization according to information from the best swarm experiences (particle set) and their own experience (Zhou et al., 2018). The velocity and position equations that guide the movement of each particle can be seen below.

$$V_i(t+1) = w * V_i(t) + c_1 * r_1(p_i - X_i) + c_2 * r_2(p_g - X_i) \tag{3}$$

$$X_i(t+1) = X_i(t) + V_i(t) \tag{4}$$

in which V_i represents particle velocity i, t the iteration, w the inertial weight, c_1 the local cognitive, c_2 the social cognitive, r_1 and r_2 the vectors of random numbers, p_i the best position of particle i, p_g the best position among all the particles in the swarm and X_i the position of particle i.

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

A function, called an adaptation function or fitness, is used to evaluate the performance of particles, and keep them in the direction of the best solution to the problem investigated. The implementation of the traditional PSO algorithm can be described as:

(i) To generate p random particles and initialize them in the search space;
(ii) To evaluate each particle and to calculate its respective fitness functions;
(iii) To calculate the best local (p_i) and global (p_g);
(iv) To update particle velocities and positions using Equations 3 and 4;
(v) Update the inertial weight w according to the current iteration information.

The process must be repeated until some stop criterion is reached, such as reaching the maximum expected error or a maximum number of iterations of the algorithm.

6. Statistical Estimates

With the methods of estimation of parameters used, the evaluation of the adjustment of the models to wind speed observations can be performed through the statistics Mean Quadratic Error of Residue (RMSE), Coefficient of Determination (R^2), and Chi-square (χ^2). Their respective expressions can be viewed in Equations 5, 6, and 7 (Pishgar-Komleh & Keyhani, 2015; Kumar et al., 2019).

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}\left(Y_{obs} - Y_{esp}\right)^2\right]^{\frac{1}{2}} \quad (5)$$

$$R^2 = \frac{\left(\sum_{i=1}^{n}(Y_{obs} - \overline{Y_{obs}}) * \left(Y_{esp} - \overline{Y_{esp}}\right)\right)^2}{\sum_{i=1}^{n}(Y_{obs} - \overline{Y_{obs}})^2 * \sum_{i=1}^{n}\left(Y_{esp} - \overline{Y_{esp}}\right)^2} \quad (6)$$

$$\chi^2 = \sum_{i=1}^{n}\left[\frac{\left(Y_{obs} - Y_{esp}\right)^2}{Y_{esp}}\right] \quad (7)$$

in which, Y_{obs} indicates the observed values, Y_{esp} indicates the expected values and, $\overline{Y_{obs}}$ and $\overline{Y_{esp}}$ indicate the means of the observed and expected values, respectively.

7. Results and Discussion

Table 2 presents the descriptive statistics of the analyzed wind velocity database. It is

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

verified that there is low variability in the observations, with coefficient of variation (CV) of 35.12% and an average velocity of 3.66 *m/s*.

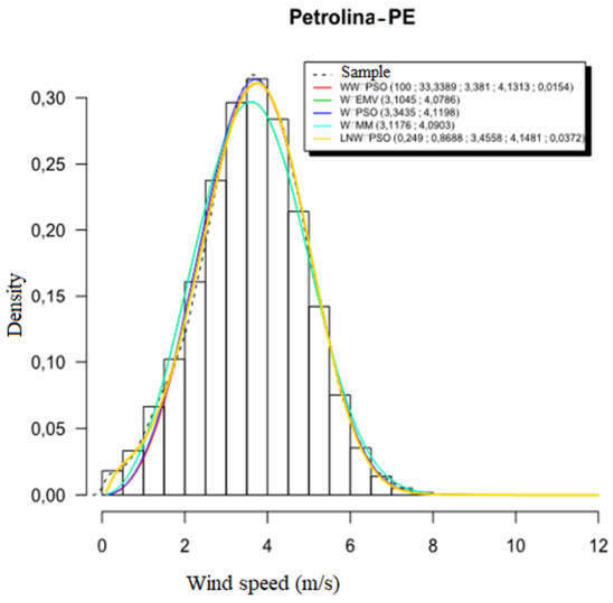
Table 2. Descriptive analysis of wind velocity observations in the municipality of Petrolina-PE.

Min.	1º Quartile	Median	Mean	3º Quartile	Max.	Skewness	Kurtosis	CV%
0.10	2.80	3.70	3.66	4.50	12.00	-0.06	2.98	35.12

Source: Prepared by the authors.

Figure 2 illustrates the curves of the probability density functions resulting from the adjustments of the parameters of the Weibull distribution model using EMV, PSO and MM, as well as the results obtained with the Weibull-Weibull and Lognormal-Weibull mixtures applying PSO.

Figure 2. Distribution models adjusted to the wind speed database in Petrolina-PE.



Source: Prepared by the authors.

According to Figure 2, although the adjustment performed by the Lognormal-Weibull mixture with PSO (in yellow) is standing out to the detriment of the others, visually it is not possible to define whether this is the optimal distribution for the modeling of the Petrolina-PE database in the studied period.

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

From this, the values of statistics R^2, $RMSE$, and χ^2 were examined, as shown in Table 3.

Table 3. Estimates of statistical tests.

Statistics	W-EMV	W-MM	W-PSO	WW-PSO	LNW-PSO
R^2	0.99473	0.994817	0.997848	0.997992	**0.998507**
$RMSE$	0.008134	0.008057	0.005949	0.005528	**0.004266**
χ^2	6.64e-05	6.52e-05	3.55e-05	3.09e-05	**1.84e-05**

Source: Prepared by the authors.

Table 3 was observed that although the Coefficient of Determination (R^2) of all analyzed methods had values very close to one, the adjustment of the Lognormal-Weibull distribution model with the optimization of parameters by the PSO algorithm (LNW-PSO) resulted in the best approximation. Regarding the Mean Quadratic Residual Error (RMSE), all the values obtained were very small and close to each other. However, the LNW-PSO's RMSE stands out as the smallest among them, revealing a better fit. Table 3 also shows that all adjusted distributions obtained lower statistics than the p-value of 5% for the statistic χ^2. However, the lowest of the values were obtained by the PSO algorithm applied to the Lognormal-Weibull distribution mixture. This fact corroborates the results obtained by statistics R^2 e $RMSE$, demonstrating that the optimization of Lognormal-Weibull parameters with the Particle Swarm Optimization algorithm expresses a better fit to the investigated database.

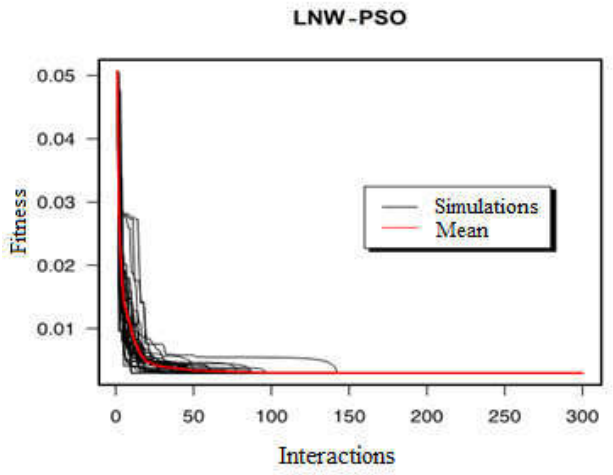
The percentages of success of the quantiles of the data set of the adjusted distribution in relation to the quantiles of the empirical set of data were calculated for each adjusted model. The Lognormal distribution mixture with Weibull, with parameters optimized with PSO, showed a higher percentage of correct answers than the other models analyzed (71.14%). Next, weibull via EMV (70.91%) and Weibull with MM (70.8%). These results reveal that, for the investigated database, the numerical methods EMV and MM provide good adjustments if applied to the Weibull distribution. On the other hand, for mixtures of distributions the PSO optimization algorithm provides better approximation of model parameters to the sample data if the Weibull distribution combined with the Lognormal distribution is applied, to the detriment of the Weibull-Weibull distribution for the data studied. This result is interesting because, in general, Weibull and Weibull-Weibull distributions are better adjusted to wind speed series, there is no optimal model that can perform modeling on all wind speed series.

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

(Qin et al., 2012).

Figure 3 illustrates the convergence graph for 30 PSO simulations during the optimization of the parameters of the Lognormal and Weibull distribution mixture (in black) and for the average behavior of these simulations (in red).

Figure 3. PSO convergence process in the search for the optimal parameters of the LNW distribution model.



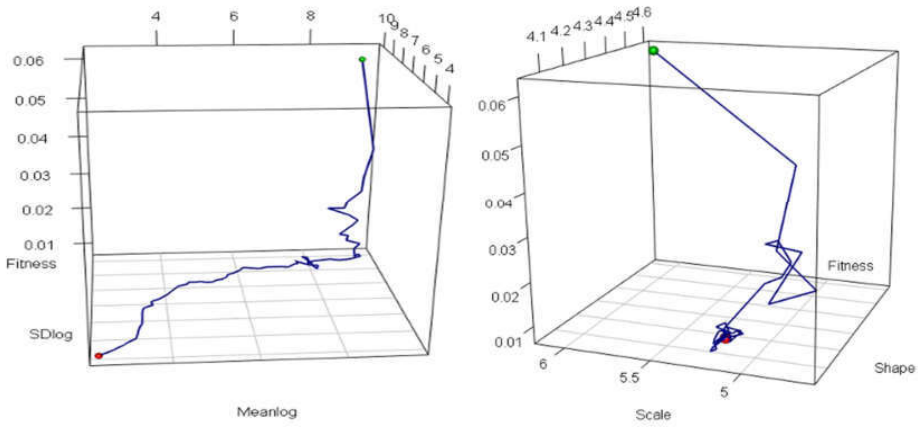
Source: Prepared by the authors.

The Ox axis in Figure 3 indicates the number of iterations performed and the Oy axis the fitness value obtained at each iteration. The fitness function was used to minimize the distance between the sample density and the theoretical of the LNW. A decline in fitness values is observed throughout the iterations, in addition to the good and rapid convergence of the PSO in the search for the ideal parameters of the Lognormal-Weibull distribution model.

The swarm in the search for optimal parameters was composed of 30 particles. Figure 4 illustrates the movement of the best particle.

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

Figures 4. Movement of the best particle in the search space in relation to the parameters of the Lognormal and Weibull distribution, respectively.



Source: Prepared by the authors.

It is possible to verify that Figure 4 displays the movement of the best particle in the search space at each iteration of the PSO, in relation to the parameters of average and standard deviation of the Lognormal (Meanlog and SDlog, respectively) and shape (Shape) and Scale (Scale) of Weibull. The green sphere indicates the starting position, while the red sphere indicates the final position on the movement.

8. Final Considerations

From the comparison of the adjustments made between the Weibull distribution with the Moment Method, with the Maximum Likelihood Estimation and with the Particle Swarm Optimization algorithm, as well as the Lognormal-Weibull and Weibull-Weibull adjustments both with the PSO, it was verified that the lowest estimation errors of the parameters of these distributions with the adopted methods were achieved with Lognormal-Weibull via PSO, with a percentage of hit of 71.14% of the adjusted data compared to the empirical data set of Petrolina-PE in the analyzed period.

This result is important in the sense that it is of paramount importance to determine the distribution model that offers better quality in the adjustment to wind speed data in order to assist in making decisions about the wind potential of the region, being able to minimize operational costs of wind power management, generation and distribution.

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

As future studies, the Lognormal-Weibull distribution combined with the Particle Swarm Optimization algorithm to obtain the parameters of the model can be used to calculate the potential of wind generation in the municipality of Petrolina, Pernambuco. In addition, studies on the gust of winds in this region can be conducted and assist public and private managers for better wind use, aiming at reducing losses arising from the by damage caused by high gusts in wind turbines.

Acknowledgments

This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financing Code 001.

References

ABEEolica. (2019). *Eólica já é a segunda fonte da matriz elétrica brasileira com 15 gw de capacidade instalada*. Acesso em 06 maio 2020. Em: http://abeeolica.org.br/noticias/eolica-ja-e-a-segunda-fonte-da-matriz-eletrica-brasileira-com-15-gw-de-capacidade-instalada/.

Da Silva, K. A., Rodrigues, M. S., Cunha, J. C., Alves, D. C., Freitas, H. R., & Lima, A. M. N. (2017). Levantamento de solos utilizando geoestatística em uma área de experimentação agrícola em Petrolina-PE. *Comunicata Scientiae*, 8(1): 175-180. https://doi.org/10.14295/cs.v8i1.2646.

De Souza, A., De Oliveira, S. S. & Ozonur, D. (2019). Análise estatística de parâmetros de Weibull para avaliação de potencial de energia eólica em Campo Grande. *Journal of Environmental Analysis and Progress*, 4.3: 168-179. https://doi.org/ 10.24221/JEAP.4.3.2019.2468.168-179.

Dos Santos, F. S., Nascimento, K. K. F., Jesus, E. S., Jale, J. S., Stosic, T. & Ferreira, T. A. E. (2019). Análise estatística da velocidade do vento em Petrolina-PE utilizando as distribuições Weibull e a Burr. *Journal of Environmental Analysis and Progress*, 4(1): 057-064. https://doi.org/10.24221/JEAP.4.1.2019.2057.057-064.

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

IBGE. (2019). *Instituto Brasileiro de Geografia e Estatística*. Acesso em 06 maio 2020 Em: https://cidades.ibge.gov.br/brasil/pe/petrolina/panorama.

Jatobá, L., Silva, A. F. & Galvíncio, J. D. (2017). A dinâmica climática do Semiárido em Petrolina-PE. *Embrapa Semiárido-Artigo em periódico indexado (ALICE)*.

Kumar, M. B. H., Balasubramaniyan, S., Padmanaban, S., & Holm-Nielsen, J. B. (2019). Wind Energy Potential Assessment by Weibull Parameter Estimation Using Multiverse Optimization Method: A Case Study of Tirumala Region in India. *Energies*, *12*(11), 2158. https://doi.org/10.3390/en12112158.

Melo, E. C. D. S., Aragão, M. R. D. S., & Correia, M. D. F. (2014). Regimes do vento à superfície na área de Petrolina, Submédio São Francisco. *Revista Brasileira de Meteorologia*, *29*(2): 229-241. https://doi.org/10.1590/S0102-77862014000200007.

Ouarda, T. B., Charron, C. & Chebana, F. (2016). Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and L-moment ratio diagram methods, with a case study. *Energy Conversion and Management*, 124: 247-265. http://dx.doi.org/10.1016/j.enconman.2016.07.012 0196-8904/.

Pereira, A.S. et al. (2018). *Metodologia da pesquisa científica*. [*e-book*]. Santa Maria. Ed. UAB/NTE/UFSM. Acesso em: 5 maio 2020. Disponível em: https://repositorio.ufsm.br/bitstream/handle/1/15824/Lic_Computacao_Metodologia-Pesquisa-Cientifica.pdf?sequence=1.

Pishgar-Komleh, S. H., Keyhani, A., & Sefeedpari, P. (2015). Wind speed and power density analysis based on Weibull and Rayleigh distributions (a case study: Firouzkooh county of Iran). *Renewable and Sustainable Energy Reviews*, *42*, 313-322. http://dx.doi.org/10.1016/j.rser.2014.10.028.

Qin, X., Zhang, J. & Yan, X. (2012). Two improved mixture Weibull models for the analysis of wind speed data. *Journal of applied meteorology and climatology, 51.7: 1321-1332.* https://doi.org/10.1175/JAMC-D-11-0231.1.

Research, Society and Development, v. 9, n. 7, eXX, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i7.XX

Rajapaksha, K. W. G. D. H., & Perera, K. (2016). Wind speed analysis and energy calculation based on mixture distributions in Narakkalliya, Sri Lanka. *Journal of the National Science Foundation of Sri Lanka*, *44*(4). http://dx.doi.org/10.4038/jnsfsr.v44i4.8023.

Rocha, L. C. S., Aquila, G., Junior, P. R., de Paiva, A. P., de Oliveira Pamplona, E., & Balestrassi, P. P. (2018). A stochastic economic viability analysis of residential wind power generation in Brazil. *Renewable and Sustainable Energy Reviews*, 90(1): 412-419. https://doi.org/10.1016/j.rser.2018.03.078.

Seckin, N., Yurtal, R., Haktanir, T., & Dogan, A. (2010). Comparison of probability weighted moments and maximum likelihood methods used in flood frequency analysis for Ceyhan River Basin. *Arabian Journal for Science and Engineering*, *35*(1), 49.

Zhou, J., Yang, J., Lin, L., Zhu, Z., & Ji, Z. (2018). Local best particle swarm optimization using crown jewel defense strategy. In *Critical developments and applications of swarm intelligence* (pp. 27-52). IGI Global. https://doi.org/10.4018/978-1-5225-5134-8.ch002.

Percentage of contribution of each author in the manuscript

Kerolly Kedma Felix do Nascimento – 25%

Fábio Sandro dos Santos – 25%

Jader da Silva Jale – 25%

Tiago Alessandro Espínola Ferreira – 25%

Chapter 4

Multifractal analysis of solar radiation in the northeastern region of Brazil

Fábio Sandro dos Santos[1*, a], Kerolly Kedma Felix do Nascimento[b], Jader Silva Jale[a], Sílvio Fernando Alves Xavier Júnior[c] and Tiago A. E. Ferreira[a]

[a] Department of Statistics and Informatics, Federal Rural University of Pernambuco, Rua Dom Manoel de Medeiros s/n, Dois Irmãos, 52171-900 Recife, Pernambuco, Brazil

[b] Teacher Training Center, Federal University of Pernambuco, Av. Marielle Franco, s/n - Km 59 - Nova, PE, 55014-900 Caruaru, Pernambuco, Brazil

[c] Department of Statistics, State University of Paraíba, Rua Baraúnas, nº 351, Bodocongó, Campina Grande/PB, 58109-753 Campina Grande, Paraíba, Brazil

---

ABSTRACT: In this work, we used hourly data of high frequency of solar radiation from the entire Northeast region of Brazil. We used the Multifractal Detrended Fluctuation Analysis (MFDFA) method to analyze the characteristics of the solar radiation series in 137 meteorological stations from 2010 to 2022. For all analyzed series, the parameter $\alpha_0 > 0.5$ characterizes persistent series. The values of $r > 1$ reveal asymmetry to the right, indicating that large fluctuations contributed to the multifractality process. The states of Maranhão and Bahia presented the highest values of spectrum width W, indicating greater complexity. We found that long-range correlations are the leading cause of multifractality observed in the dynamics of the series of solar radiation anomalies.

Keywords: Multifractal, solar energy, solar radiation, persistence, complexity.

---

[1] * Corresponding author.

Email address: fabio.sandropb@gmail.com (Fábio Sandro dos Santos)

1. INTRODUCTION

Photovoltaic (PV) solar energy can act as a complementary energy source to existing renewable sources worldwide. There is uncertainty about each country's compliance with the Paris Agreement on reducing CO2 in the atmosphere. Regarding countries' actions at COP26, in Glasgow, Scotland, in 2021[1], there is a concern for the agreements to be served up to 2030, with a temperature reduction of around 1.5°C [2]. At COP26, Brazil committed to reducing about 50% of greenhouse gas emissions in the atmosphere, showing that by 2019, around 53.8% of all Brazilian energy was derived from fossil sources [3]. On June 09, 2022, more than 83% of all energy in Brazil was from renewable energies, exceeding 184 GW of installed power [4]. Based on these observations, photovoltaic energy can contribute to the reduction of global warming, as this energy source is clean and renewable. In the globalized world, there is an expanding need for investment in new renewable sources to supply all the energy demand.

The photovoltaic energy market has been extensively studied. In recent years, several researchers have studied the photovoltaic energy potential [5], [6], [7], [8], [9], [10]. Other studies used the MFDFA Multifractal Detrended Fluctuation Analysis method in the solar radiation series [11], [12], [13]. The studies verified the complexity properties of multifractality, which are pretty relevant in solar radiation series since it is possible to understand the characteristics of time series, such as persistence, anti-persistence, and dynamics. Plocoste and Pavón-Domínguez [14] found evidence of long-range correlations and small and significant fluctuations in the solar radiation series.

In Brazil, in recent years, investments in photovoltaic energy generated from solar panels have been gaining more and more space in its energy matrix [15]. Some research is being done with the objective of understanding this clean and renewable energy source in the country [16], [17], [18]. In this scenario, Brazil has great potential for photovoltaic energy; the country contributes to reducing the greenhouse effect in the atmosphere. Water is the main source of Brazilian energy generation used for energy generation by hydroelectric plants, which account for about 60% of the entire energy matrix in Brazil [19].

Photovoltaic energy is expected to be responsible for about 3.9% of all Brazilian energy by 2025, reaching 36% by the year 2050 [20]. Marchetti and Rego (2022) [21] observed that in Brazil's Northeast and Southeast regions, solar park performance capacity peaks occur at noon, with large productions between 10 am and 3 pm. In the state of Paraíba, located in the Northeast region of Brazil, water use for hydroelectric energy can be replaced by energy generated from solar radiation [22]. In the researched literature, no work was found using the Multifractal Detrended Fluctuation Analysis – MFDFA method to analyze the time series of solar radiation in Brazil. To fill this gap in the literature, we used the MFDFA method in the daily time series of solar radiation from 2010 to 2022. In this article, we investigated the multifractality properties of the time series of solar radiation in the entire Northeast region of Brazil, intending to verify suitable areas for installing solar panels for the possible generation of solar energy.

The rest of the article is organized as follows: In section 2, study areas, the study stations are located. In section 3, we present the MFDFA methodology. In subsection 4, we present the Inverse Distance Weighting (IDW) interpolation method for estimating values in which we do not have any a priori information. Section 5 shows the results found and compares them with the literature. In section 6, we conclude our findings in this article.

2. STUDY AREA

The Northeast region of Brazil has a geographic area of approximately 19,427 km² [23]. The Northeast is characterized as a dry and semi-arid region. Most of the time, it has low precipitation rates in much of the region. Detailed solar radiation $\left(\frac{W^2}{m}\right)$ databases containing hourly measurements from 2010 to 2022 in the Northeast region of Brazil were obtained from the National Institute of Meteorology (INMET[2]). Figure 1 shows the geographical locations of the meteorological stations investigated in the states Alagoas-AL, Bahia-BA, Ceará-CE, Maranhão-MA, Rio Grande do Norte-RN, Paraíba-PB, Pernambuco-PE, Piauí-PI and Sergipe -SE in the Northeast region of Brazil.

[2] https://tempo.inmet.gov.br/TabelaEstacoes/

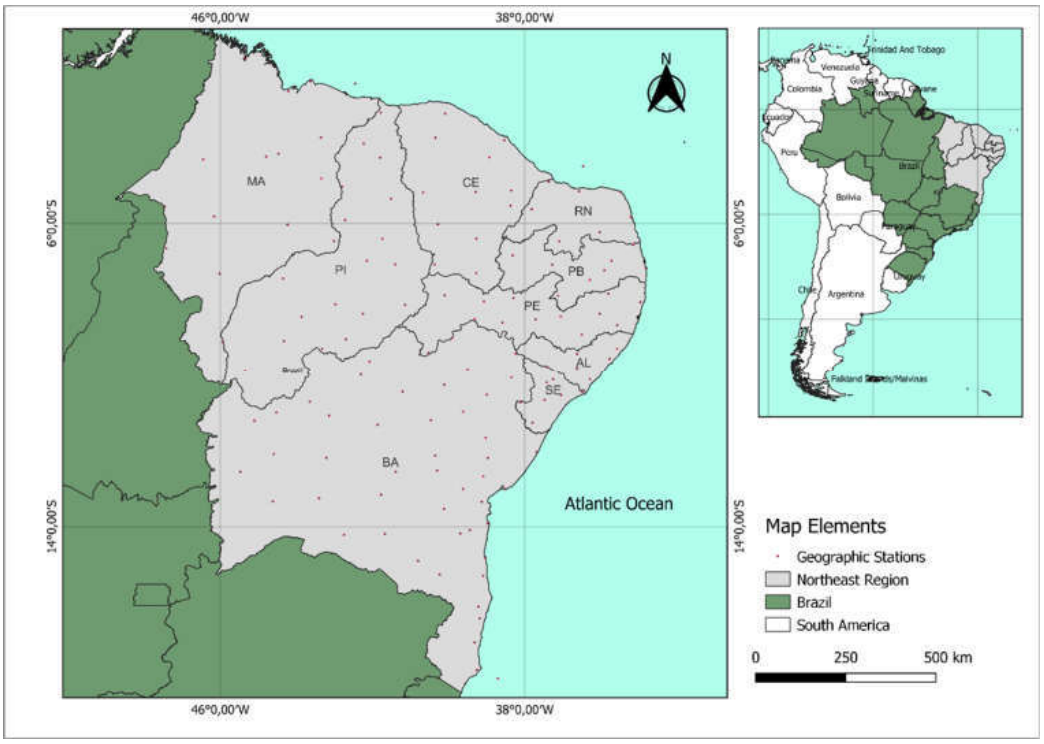


Figure 1. The geographical location of meteorological stations in the Northeast region of Brazil.

In Figure 2, to exemplify the times with the highest incidence of solar radiation, we present the historical series of the behavior of a measurement day for the nine capitals of the Northeast Region of Brazil. For all capitals, the peak of solar radiation occurs between 10:00 am and 3:00 pm on August 29, 2022 (Time zone in Brasilia, Federal District (GMT-3)).

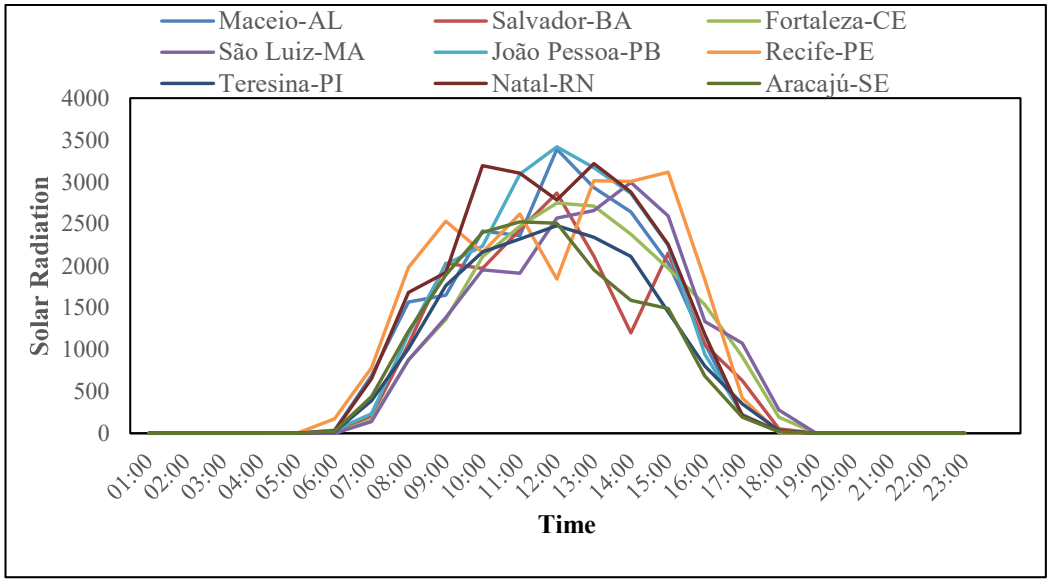


Figure 2. Hourly time series of solar radiation $\left(\frac{W^2}{m}\right)$ measured on August 29, 2022.

3. Method

The Multifractal Detrended Fluctuation Analysis (MFDFA) method is a generalization of the Detrended Fluctuation Analysis (DFA) method. MFDFA is a powerful tool for detecting multifractal in a non-stationary time series with small and large fluctuations.

The MFDFA method has been widely used in financial series applications [24], energy market [25], the impact of COVID-19 on energy prices [26], rainfall characteristics in India [27], closing prices of commodities [28], in addition to the time series of temperature and solar radiation [29]. The implementation of the MFDFA method considers any set of a non-stationary time series x_i $(i = 1,2, \dots, N)$ with a specific length N. We determine the following steps of the MFDFA:

(i) The first step is to construct a new integration of the original series $x(i)$ after subtracting the mean of the time series $\bar{x}$, in which $\bar{x} = \frac{1}{N}\sum_{i=1}^{k} x_i$,

$$X(k) = \sum_{i=1}^{k} \left[x_{(i)} - \bar{x}\right], k = 1,2, \dots, N. \tag{1}$$

(ii) The second step is to generate an integrated series, subdividing $X(k)$ into $N_n = int(\frac{N}{n})$ non-overlapping segments of sizes equal to n(scale). In the segments $v = 1,2, \dots, \frac{N}{n}$, the local trend $X_{n,v}(k)$ is estimated by the least squares fit of the series.

(iii) In the third step, the unbiased variance is estimated by the following equation:

$$F^2(n, v) = \frac{1}{N} \sum_{k=(v-1)n+1}^{vn} \left[X(k) - X_{n,v}(k)\right]^2 \tag{2}$$

(iv) In the fourth step, we calculate the average of all segments and obtain the fluctuation function of order q, given by the following equation:

$$F_q(n) = \left\{ \frac{1}{N_n} \sum_{v=1}^{N_n} [F^2(n, v)]^{\frac{q}{2}} \right\}^{\frac{1}{q}}, \tag{3}$$

where q can take on any real value other than zero.

(v) In the fifth step, the scales of the fluctuation functions are analyzed using $log - log\ F_q(n)$ graphs at each scale n, in which they follow a power law $F_q(n) \propto n^{h(q)}$ for cases where long-range correlations are present in the time series.

The scale exponent $h(q)$ is called the generalized Hurst exponent. In the case where a time series is stationary, $h(q = 2)$ is identical to the well-known Hurst exponent (2). For positive values of q, $h(q)$ presents a scaling behavior of large fluctuations; in contrast to values of q negative, $h(q)$ shows small fluctuations. If $h(q)$ is independent of q, the time series has a monofractal process, and if $h(q)$ decays with respect to q, the time series has a multifractal process.

The relationship between the generalized Hurst exponent and the Rényi exponent $\tau(q)$ can be defined by the equation,

$$\tau(q) = qh(q) - 1, \tag{4}$$

for the monofractal process, $\tau(q)$ is a linear function of q, $h(q)$ is a constant, and for time series with a multifractal process, $\tau(q)$ is a nonlinear function. The latter can also be characterized by the singularity spectrum or multifractal spectrum $f(\alpha)$, where it is obtained through the Legendre transform,

$$\alpha(q) = \frac{\partial\tau(q)}{\partial q}, \tag{5}$$

$$f\big(\alpha(q)\big) = q\alpha(q) - \tau(q), \tag{6}$$

where α is the Holder exponent, $f(\alpha)$ indicates the multifractal dimension of the singularity measure of the series characterized by α. For the monofractal process, the singularity spectrum is represented by a single point, while a downward concave function gives the multifractal spectrum.

To measure the degree of complexity of the multifractal spectrum of a time series, we fit a fourth-degree polynomial, and from there, we can obtain the parameters of the MFDFA method. The three parameters of the singularity spectrum are estimated from the following equation:

$$f(\alpha) = A + B(\alpha - \alpha_0) + C(\alpha - \alpha_0)^2 + D(\alpha - \alpha_0)^3 + E(\alpha - \alpha_0)^4 \tag{7}$$

When fitted, the fourth-degree polynomial reveals the value of α_0, which is the maximum point of the singularity spectrum. For the value of $\alpha_0 = 0.5$, the times series are said to be random. For values of $\alpha_0 < 0.5$, the series are anti-persistent, showing that the time series show behavior of sudden changes over time. For values of $\alpha_0 > 0.5$, the series exhibit persistent behavior, indicating that the time series do not show sudden changes over time [28].

The asymmetry value, obtained by $r = \frac{\alpha_{max} - \alpha_0}{\alpha_0 - \alpha_{min}}$, is the asymmetry value of the multifractal spectrum, where the value of $r = 1$ reveals multifractal asymmetry and indicates that the time series process is governed by small and large fluctuations. Values of $r > 1$ reveal asymmetry to the right, indicating large fluctuations in the multifractal process. Values of $r < 1$ reveal asymmetry to the left, indicating small fluctuations in the multifractal process [30]. The spectrum width is calculated by $W = \alpha_{max} - \alpha_{min}$. Values of $W = 0$ indicate that the series is uniformly distributed. Higher W values indicate a greater degree of multifractality.

4. Inverse Distance Weighting (IDW)

The Inverse Distance Weighting (IDW) method was used to estimate values of multifractal complexity parameters. Through this interpolation method, it is possible to predict values of observations based on the information closest to the point we want to estimate. The mathematical equation of the IDW method is defined as follows [31]:

$$x = \frac{\sum_{i=1}^{n} Z_i W_i}{\sum_{i=1}^{n} W_i}, \tag{8}$$

where x is the unobserved point to be predicted, Z_i is the control value of the i-th point of the time series, and $W_i = d_{x,y,i}^{-\beta}$ is the weight defined in the interpolation for estimating the new point, where $d_{x,y,i}$ is the distance between $z, x\ y$ and z_i and β is an exponent defined by the searcher, in our work was using the value of $\beta = 3$.

5. Results and discussion

In Figure 3, the highest averages of solar radiation in the Northeast region are concentrated in the West and Northeast parts of the region, with emphasis on the state of Ceará, Southwest of Bahia, in the southern part of the states of Piauí and Maranhão. In addition to a large part of the coast of Rio Grande do Norte and west of the border of the states of Sergipe and Alagoas. On the other hand, the coastal part between Bahia and Alagoas is the area with the lowest incidence of solar radiation in the Northeast region of Brazil. Also, according to Figure 3, the Southwest and Northeast parts of the map present the areas with the highest values of the standard deviation of solar radiation. To the north of the region, the lowest values of the standard deviation of solar radiation were verified. The state of Maranhão stands out because, among all the states, throughout its area, the values of the standard deviations had a uniform behavior. This behavior of the state of Maranhão may have been influenced by its proximity to the Amazon rainforest. This somehow reflects a lower standard deviation in the region. In this sense, deforestation can somehow influence the increase in UV indices in the atmosphere [32]. In the coastal part of the Northeast, the lowest values of the mean and standard deviation of radiation may be related to the presence of Atlantic Forest areas that still resist in this region.

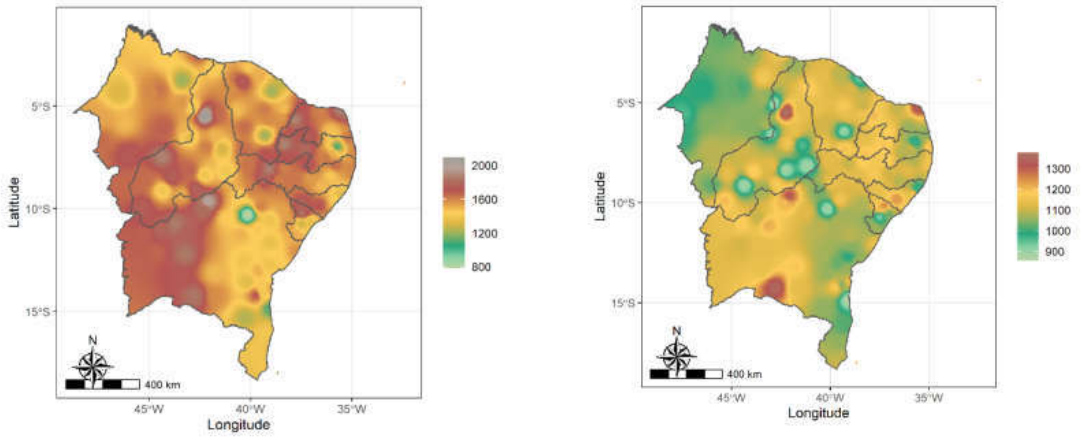


Figure 3. On the left is the mean, and on the right is the standard deviation of solar radiation $\left(\frac{W^2}{m}\right)$ in the Brazilian Northeast.

In order to analyze the multifractality properties of the time series of solar radiation in the Brazilian Northeast Region, the anomalies of 137 meteorological stations distributed throughout the nine states

were calculated. In applying the MFDFA method, we used the MFDFA Package [33]. The anomalies were calculated using the equation below:

$$\phi_{x_i} = \frac{x_i - \bar{x}_i}{\sigma_i}, \quad (9)$$

where x_i stands for the hourly values observed in the solar radiation time series, $\bar{x}_i$ is the average of the time series observations, and σ_i is the value of the standard deviation of the solar radiation series. To exemplify the process of implementation and output of the results, we chose a station randomly to represent the graphs built with the MFDFA method.

According to Figure 4A, the fluctuation function $F_q(n)$ presents a linear behavior in the logarithmic scale q of $(-10$ a $10)$; this means that the solar radiation series indicates a multifractality process over time. In Figure 4B, the generalized Hurst exponent h_q has a decreasing behavior. The Rényi exponent $\tau_{(q)}$ (Figure 4C) points out the nonlinear form of the process. The multifractal spectrum $f_{(\alpha)}$(Figure 4D) reveals a concave downward curve, confirming the presence of multifractal in the hourly time series of solar radiation for this randomly selected meteorological station.

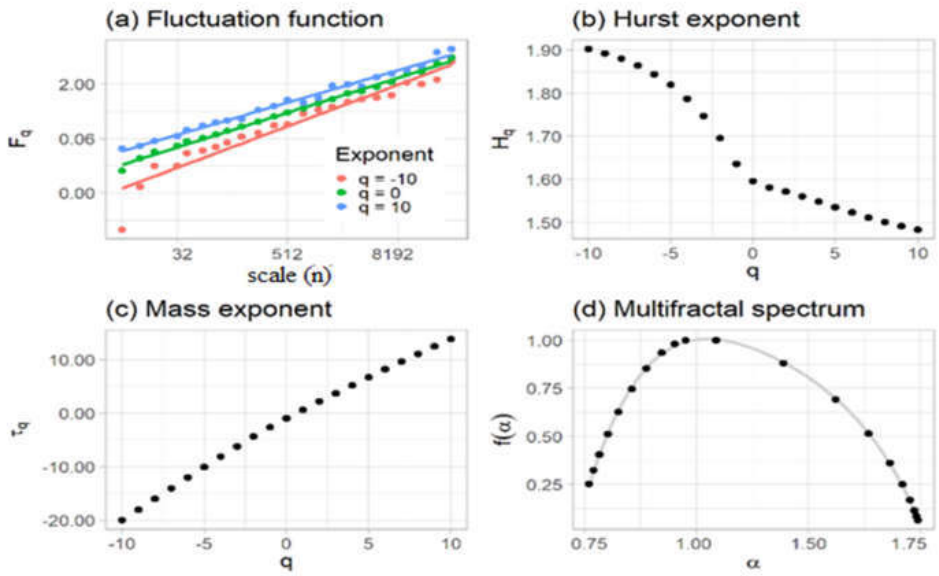


Figure 4. Behavior found with the MFDFA method for the series of solar radiation anomalies in Maceió-AL. (a) Fluctuation Functions, (b) Generalized Hurst Exponent, (c) Rényi Exponent, and (d) Multifractal Spectrum.

We shuffled the series of solar radiation anomalies, then applied the MFDFA method again to determine the multifractal complexity characteristics in all stations geographically distributed in all nine states of the Brazilian Northeast region. The shuffling was performed with the mathematical expression $random = 10.000 \times x_i$, repeating 1000 times with different random seeds.

From the series of anomalies and randomized series, it was possible to analyze the causes of multifractality present in the processes: (i) Long-range correlations; and/or (ii) The probability density function [34]. We observed in Figure 6 that all randomized series showed weaker multifractality in relation to anomaly series. This fact indicates that the long-range correlations and the probability density are the main causes of the multifractality observed in the radiation series. Figures 5, 6 and 7 represent the variation of the multifractal complexity parameters, α_0, W, and r, for the original time series (anomalies) versus the scrambled series.

Table 1 shows some stations where the cause of multifractality was the probability density function. This occurs when the value of W_r randomized is approximately equal to the value of W_a of the series of anomalies, that is when the difference (ΔW) between them is equal to or close to zero.

Table 1: Multifractal complexity parameters for some meteorological stations.

Station	**Latitude**	**Longitude**	$\boldsymbol{W}$ **anomalies** $(\boldsymbol{W_a})$	$\boldsymbol{W}$ **randomized** $(\boldsymbol{W_r})$	$\boldsymbol{\Delta W = W_a - W_r}$
Iguatu-CE	-6.396	-39.268	0.509	0.493	0.016
Barreiras-BA	-12.091	-44.592	0.491	0.461	0.029
Itaporanga-PB	-7.516	-38.233	0.638	0.591	0.048
Alvorada do Gurgueia-PI	-8.441	-43.865	0.568	0.512	0.056
Apodi-RN	-5.626	-37.815	0.504	0.439	0.065

Bom Jesus da Lapa-BA	-13.251	-43.405	0.543	0.473	0.069
Recife-PE	-8.059	-34.959	0.513	0.439	0.074
Imperatriz-MA	-5.555	-47.459	0.508	0.433	0.075
Nossa Senhora da Glória-SE	-10.130	-37.251	0.716	0.639	0.077
Picos-PI	-7.071	-41.404	0.551	0.472	0.078
Colinas-MA	-8.150	-48.783	0.538	0.457	0.081
Correntinha-BA	-13.332	-44.617	0.533	0.448	0.085
Santa Rita de Cassia-BA	-11.002	-44.525	0.511	0.426	0.085
Ribeira do Amparo-BA	-11,046	-38,432	0.661	0.570	0.091

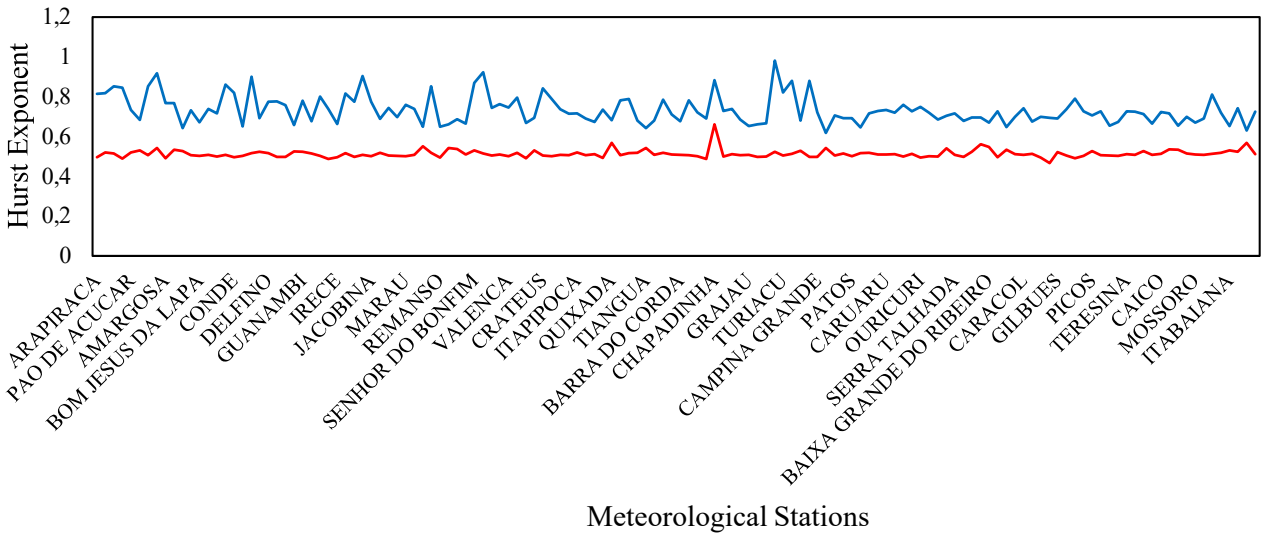


Figure 5. Comparison between the multifractal process complexity parameters of the original series versus the randomized series: α_0 (blue) anomaly series and α_0 (red) randomized series for all solar radiation series.

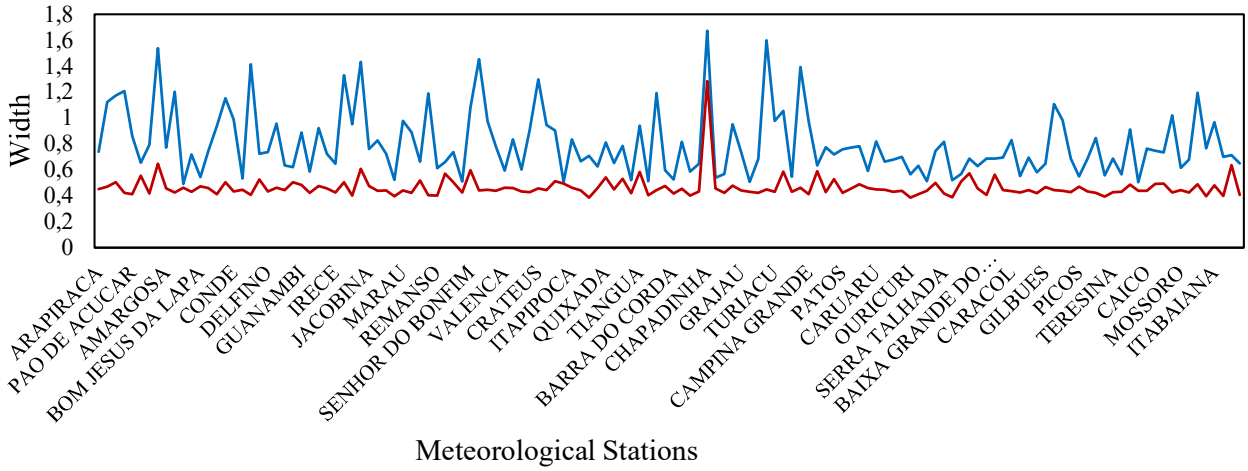


Figure 6. Comparison between the multifractal process complexity parameters of the original series versus the randomized series: W (blue) anomaly series and W (red) randomized series for all solar radiation series.

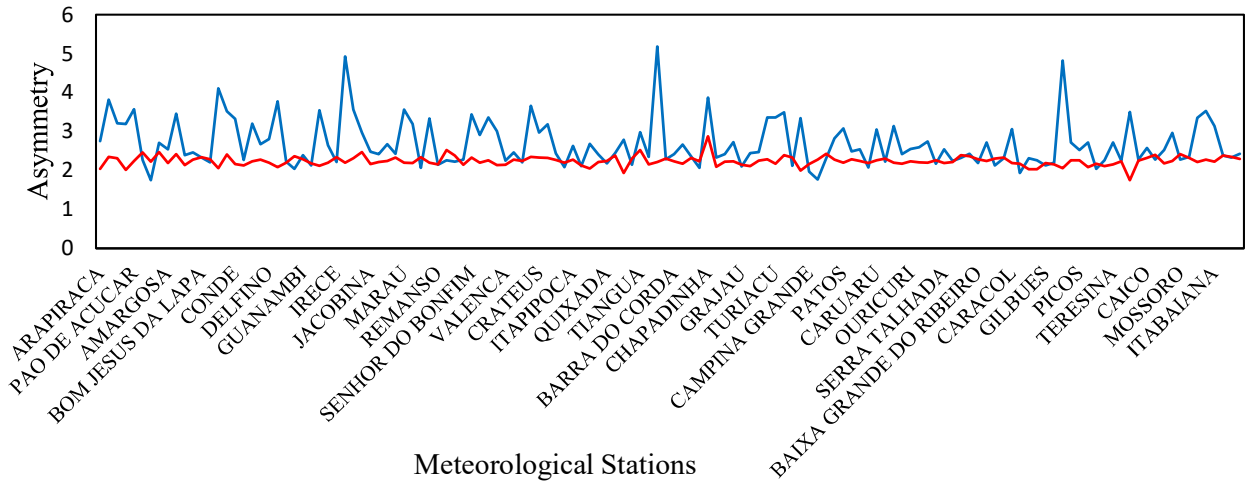


Figure 7. Comparison between the multifractal process complexity parameters of the original series versus the randomized series: r (blue) anomaly series and r (red) randomized series for all solar radiation series.

Figure 8 exhibits the results obtained after applying the IDW method to estimate the MFDFA parameters in places where information on these parameters was not available. We notice in Figure 8A that all the values estimated for the parameter α_0 were greater than 0.5, indicating that the solar radiation series present persistent behavior. This situation justifies that the observed values did not decrease over

these 12 years of study. We also note that long-range correlations are present in the analyzed series because the closer $\alpha_0 \to 1$, the longer the long-range correlation period is present in the solar radiation series.

In the state of Maranhão, specifically on the coast, a greater presence of long-range correlation was observed among all the states of the Northeast region. In the southern and northeastern parts of the state of Bahia, the highest values of α_0 were observed, revealing a more extanding period of long-range correlation. In the states of Paraíba and Sergipe, the presence of long-range correlation was observed in the coastal part. In the other states, the values of α_0 behaved uniformly.

In Figure 8B, the spectrum width (W) values ranged between (0.490 and 1.673). The highest values of the width of the multifractal spectrum were observed in Maranhão. This state is under the influence of the Amazon rainforest and is close to the equator. What may be influencing such expressive values of the spectrum width W. The states of Bahia, Piauí, and Paraíba presented relatively high values of spectrum width (W) in the Atlantic Forest and Caatinga biomes.

One can see that these values decrease in the East/West direction. The lowest values were observed in the Cerrado portion, in the west of Bahia and Piauí and southeast of Maranhão. The greater the width of the multifractal spectrum, the greater the complexity of the stochastic process that generates the analyzed series and, consequently, the higher the difficulty in making predictions. In this case, the western regions of the states of Bahia, Piauí, and Maranhão are the places where the best regions to perform solar radiation prediction were observed, as well as the coastal part of Rio Grande do Norte, East/West of Pernambuco and West of Paraíba.

Figure 8C shows the map of Asymmetry values (r), which ranged between (1.751 and 5.188). It was possible to observe that all values of r were greater than 1. Based on this spatial information, we can say that the multifractality of hourly solar radiation in the entire Northeast region is more influenced by small fluctuations.

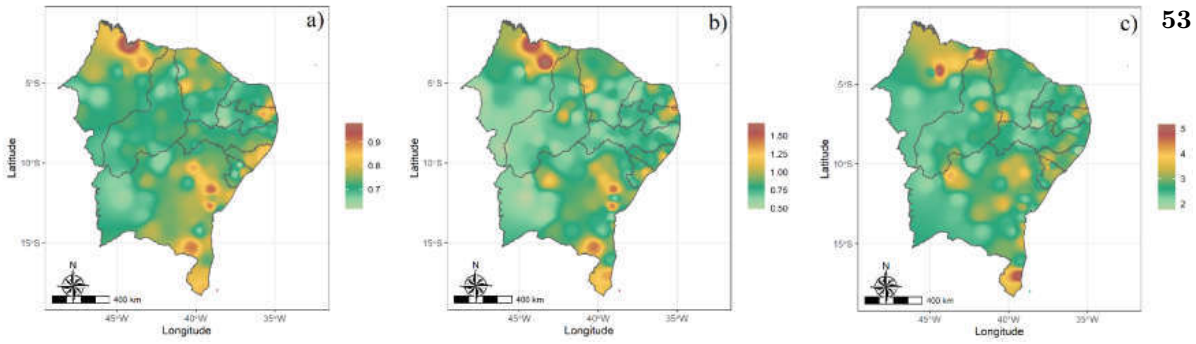


Figure 8. a) values of α_0, b) values of w, while c) contains the asymmetry values r.

Similar work to our approach was carried out in the Guadeloupe archipelago located in the Caribbean. The authors observed that time series of solar radiation contribute to long-range correlations and small and large fluctuations [35, 11]. This type of characteristic can be caused by the influences of air masses. It is possible to observe that the locations with the highest asymmetry values are geographically distributed in the North and Northeast parts of the state of Maranhão, as well as in the southern part of Bahia and in some areas distributed in the central part of this state (Figure 8C). In these areas, the influence of air masses associated with the southeast trade winds [36] often hinders the arrival of rains during the winter period. This episode leads to an increase in solar radiation rates. Unlike other regions, during the winter period, the air masses can move forward with more force, and there is a decrease in the effect of solar radiation.

In specific years, the main phenomena that influence the climate in the Northeast region of Brazil are El Niño and La Niña. In El Niño years, there is a gradual decrease in the amount of rainfall in the region [37]. On the other hand, in years of the presence of the La Niña phenomenon, the Northeast region is more susceptible to the more constant presence of rain in the region [38]. Such phenomena act directly on the increase or decrease of solar radiation in the Brazilian Northeast. Our findings contribute to a better understanding of the multifractal process of solar radiation series. The MFDFA method proved to be efficient in characterizing the regimes and behavior of the solar radiation series in a large study area such as the Brazilian Northeast region.

With the increasing need for investments in new energy sources, the energy generated from solar radiation becomes an efficient source to meet the energy demand across the planet and as a complementary source to existing energy sources. In Brazil, solar energy already has a production of

approximately 16 GW (Gigawatts) of installed photovoltaic solar energy. Also, according to Associação Brasileira de Energia Solar Fotovoltaica (ABSOLAR), Brazil failed to send about 23.6 million tons of CO2 into the atmosphere with the generation of electricity. Another significant result of the investment in photovoltaic energy is in the economic part since approximately 500 thousand jobs have already been generated, and more than R$ 86.2 billion have already been applied in new investments for Brazil in recent years [39].

CONCLUSION

In this article, the multifractal process of the hourly anomalies of the time series of solar radiation in the Northeast region of Brazil was investigated. For that, we applied the MFDFA method in the series of hourly anomalies over 12 years studied and obtained the parameters of multifractal complexity α_0, W and r, as well as the multifractal spectrum in each of the 137 stations. We observed that of the 137 meteorological stations studied, in 14 of them, the probability density function was the leading cause of the multifractality process present in the time series of solar radiation. In the remaining 123 stations, the leading causes of multifractality present in the series were long-range correlations and the probability density function. In future works, a monthly application over the years using the MFDFA method should be made in order to compare the characteristics of multifractality in the series, as well as a future study for the entire Brazilian territory.

ACKNOWLEDGMENT

This work received financial support from the Coordination for the Improvement of Higher Education Personnel — Brazil (CAPES) — Financing Code 001.

REFERENCES

1. Smith, P., *et al*., (2022). Essential outcomes for COP26. *Global change biology*, 28(1), 1-3.
2. Gasser, T., Ciais, P., & Lewis, S. L. (2022). How the Glasgow Declaration on Forests can help keep alive the 1.5° C target. *Proceedings of the National Academy of Sciences*, 119(23), e2200519119.
3. Maia, R. G. T., & Bozelli, H. (2022). The importance of GHG emissions from land use change for biofuels in Brazil: An assessment for current and 2030 scenarios. *Resources, Conservation and Recycling*, 179, 106131.

4. ANEEL, Agência Nacional de Energia Elétrica. 2022. https://www.gov.br/aneel/pt-br. Accessed in 2022-09-15.
5. Kabir, E., Kumar, P., Kumar, S., Adelodun, A. A., & Kim, K. H. (2018). Solar energy: Potential and future prospects. *Renewable and Sustainable Energy Reviews*, 82, 894-900.
6. Zhao, X., Huang, G., Lu, C., Zhou, X., & Li, Y. (2020). Impacts of climate change on photovoltaic energy potential: A case study of China. *Applied Energy*, 280, 115888.
7. Silalahi, D. F., Blakers, A., Stocks, M., Lu, B., Cheng, C., & Hayes, L. (2021). Indonesia's Vast Solar Energy Potential. *Energies*, 14(17), 5424.
8. Kan, A., Zeng, Y., Meng, X., Wang, D., Xina, J., Yang, X., & Tesren, L. (2021). The linkage between renewable energy potential and sustainable development: Understanding solar energy variability and photovoltaic power potential in Tibet, China. *Sustainable Energy Technologies and Assessments*, 48, 101551.
9. Muhammad, A., Muhammad, U., & Abid, Z. (2021). Potential of floating photovoltaic technology in Pakistan. *Sustainable Energy Technologies and Assessments*, 43, 100976.
10. lrwashdeh, S. S. (2021). Investigation of the energy output from PV panels based on using different orientation systems in Amman-Jordan. *Case Studies in Thermal Engineering*, 28, 101580.
11. Plocoste, T., & Pavón-Domínguez, P. (2020). Temporal scaling study of particulate matter (PM10) and solar radiation influences on air temperature in the Caribbean basin using a 3D joint multifractal analysis. *Atmospheric Environment*, 222, 117115.
12. Akinsusi, J., Ogunjo, S., & Fuwape, I. (2022). Nonlinear dynamics and multifractal analysis of minimum–maximum temperature and solar radiation over Lagos State, Nigeria. *Acta Geophysica*, 1-8.
13. Li, Q., & Li, P. (2021). Intermittency study of global solar radiation under a tropical climate: case study on Reunion Island. Scientific Reports, 11(1), 1-16.
14. Li, Q., & Li, P. (2021). Intermittency study of global solar radiation under a tropical climate: case study on Reunion Island. Scientific Reports, 11(1), 1-16.
15. dos Santos Carstens, D. D., & da Cunha, S. K. (2019). Challenges and opportunities for the growth of solar photovoltaic energy in Brazil. *Energy policy*, 125, 396-404.
16. Garlet, T. B., Ribeiro, J. L. D., de Souza Savian, F., & Siluk, J. C. M. (2019). Paths and barriers to the diffusion of distributed generation of photovoltaic energy in southern Brazil. Renewable and Sustainable Energy Reviews, 111, 157-169.
17. Santos, A. J. L., & Lucena, A. F. (2021). Climate change impact on the technical-economic potential for solar photovoltaic energy in the residential sector: A case study for Brazil. Energy and Climate Change, 2, 100062.
18. de Martino Jannuzzi, G., & de Melo, C. A. (2013). Grid-connected photovoltaic in Brazil: Policies and potential impacts for 2030. *Energy for Sustainable Development*, 17(1), 40-46.

19. Zuluaga, C. F., Avila-Diaz, A., Justino, F. B., Martins, F. R., & Ceron, W. L. (2022). The climate change perspective of photovoltaic power potential in Brazil. Renewable Energy.
20. Marchetti, I., & Rego, E. E. (2022). The impact of hourly pricing for renewable generation projects in Brazil. Renewable Energy, 189, 601-617.
21. Mendes, H. A. (2021, October). On AutoMLs for Short-Term Solar Radiation Forecasting in Brazilian Northeast. In 2021 International Conference on Engineering and Emerging Technologies (ICEET) (pp. 1-6). IEEE.
22. Medeiros, S. E. L., Nilo, P. F., Silva, L. P., Santos, C. A. C., Carvalho, M., & Abrahão, R. (2021). Influence of climatic variability on the electricity generation potential by renewable sources in the Brazilian semi-arid region. *Journal of Arid Environments*, 184, 104331.
23. Tabarelli, M., Melo, M. D. V. C., & Lira, O. C. (2006). A Mata Atlântica do nordeste. Rio de Janeiro: *MMA*.
24. Choi, Sun-Yong. Analysis of stock market efficiency during crisis periods in the US stock market: Differences between the global financial crisis and COVID-19 pandemic. *Physica A: Statistical Mechanics and Its Applications*, v. 574, p. 125988, 2021.
25. Wang, Q., Yang, X., and L, R. The impact of the COVID-19 pandemic on the energy market–A comparative relationship between oil and coal. *Energy Strategy Reviews*, v. 39, p. 100761, 2022.
26. Khan, K. *et al*. COVID-19 impact on multifractality of energy prices: Asymmetric multifractality analysis. *Energy*, v. 256, p. 124607, 2022.
27. Sarker, A., and Mali, P. Detrended multifractal characterization of Indian rainfall records. *Chaos, Solitons & Fractals*, v. 151, p. 111297, 2021.
28. Nascimento, K. K. F. do et al. COVID-19 influence over Brazilian agricultural commodities and dollar-real exchange. *Fractals*, v. 30, p.1-10, 2022.
29. Akinsusi, J., Ogunjo, S., and Fuwape, I. Nonlinear dynamics and multifractal analysis of minimum–maximum temperature and solar radiation over Lagos State, Nigeria. *Acta Geophysica*, p. 1-8, 2022.
30. Kantelhardt, J. W. *et al*. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, v. 316, n. 1-4, p. 87-114, 2002.
31. Masoudi, M. (2021). Estimation of the spatial climate comfort distribution using tourism climate index (TCI) and inverse distance weighting (IDW) (case study: Fars Province, Iran). *Arabian Journal of Geosciences*, 14(5), 1-13.
32. Reis, G., *et al*., (2022). Solar Ultraviolet Radiation Temporal Variability Analysis from 2-Year of Continuous Observation in an Amazonian City of Brazil. Atmosphere, 13(7), 1054.
33. Laib M., Telesca L., Kanevski M. Mfdfa: multifractal detrended fluctuation analysis for time series (2017) R package version 01 0.
34. dos Santos, F. S., do Nascimento, K. K. F., da Silva Jale, J., Stosic, T., Marinho, M. H., & Ferreira, T. A. (2021). Mixture distribution and multifractal analysis applied to wind speed in the Brazilian Northeast region. Chaos, Solitons & Fractals, 144, 110651.

35. Plocoste, T., & Pavón-Domínguez, P. (2020). Multifractal detrended cross-correlation analysis of wind speed and solar radiation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30(11), 113109.
36. de Albuquerque Wanderley, L. S., & Nóbrega, R. S. (2022). Desenvolvimento de um sistema de classificação climática com base na metodologia dos tipos sinóticos de tempo para a Região Nordeste do Brasil. *GEOUSP Espaço e Tempo (Online)*, 26(1).
37. Jardim, A. M. D. R. F., *et al.* (2021). Spatiotemporal climatic analysis in Pernambuco state, Northeast Brazil. *Journal of Atmospheric and Solar-Terrestrial Physics*, 223, 105733.
38. dos Santos Fernandes, É., & de Sousa Lopes, J. L. (2021). Natural disasters in the state of Alagoas northeast region of BRAZIL-arising from the climate events in La NIÑA. *International Journal Semiarid*, 4(4).
39. ABSOLAR (Associação Brasileira de Energia Solar Fotovoltaica). Accessed on 2022-09-12; URL https://www.absolar.org.br/noticia/energia-solar-brasil-ultrapassa-marca-historica-de-16-gw/.

Chapter 5

Prediction of wind energy generation potential in Brazil using mixtures of Weibull distributions.

Fábio Sandro dos Santos[a,*], Kerolly Kedma Felix do Nascimento[a], Jader da Silva Jale[a], Sílvio Fernando Alves Xavier Júnior[b], Tiago A. E. Ferreira[a]

[a]*Department of Statistics and Informatics, Federal Rural University of Pernambuco, Recife, Pernambuco, Brazil*

[b]*Department of Statistics, State University of Paraíba, Campina Grande, Paraíba, Brazil*

Abstract

Due to the need to reduce the emission of carbon dioxide in the atmosphere as a result of the greenhouse effect, there is a growing demand to intensely reduce the use of fossil fuels in energy generation to meet the energy demand of the population, which is only growing. Among the options for renewable sources, we highlight wind energy, which is clean and inexhaustible. Wind energy can contribute to the new era of energy sources that do not pollute the environment. The contributions of wind energy are countless so that nations around the planet can reach goals established in the Paris Agreement, in which, by 2030, greenhouse gases must be reduced by half and zeroed by 2050. Brazil is one of the countries that possess great potential for generating energy from renewable sources. In particular, the energy from the winds. The knowledge and understanding of wind speed behavior at different heights and the use of mixtures of probability distributions for estimating the potential of wind generation in large territorial areas are of paramount importance for several reasons. Among them, to evaluate the types of wind turbines that best suit a given height and the potential for energy generation in each region, reducing uncertainties in the development and distribution of wind energy. In this work, we apply the Weibull-Weibull distribution mixture in adjusting the wind speed series of each of the 575 meteorological stations at six different heights. Next, we estimate the wind power density. In estimating the Weibull-Weibuill parameters, the Expectation-Maximization (EM) algorithm was used. With the results obtained at each meteorological station and using the Inverse Distance Weighting (IDW) method, we predicted wind energy generation at points where there was no information about wind speed. We observed in this study that over large areas, the Weibull-Weibull distribution mixture proved to be excellent in estimating the wind energy density distribution. It was able to provide a suitable fit for the different regions of Brazil. The results obtained here demonstrate that the model used can be an excellent candidate to be applied in other countries as a form of integration with other energy sources in a complementary way. In Brazil, we indicate this result in the integration of Solar, Biomass, and Hydro energy sources, making the country not suffer from rationing due to lack of energy, given that an auxiliary energy source could be distributed quickly and efficiently using the subsystems existing electrical. This action can help avoid using energy sources that pollute the environment,

such as the electrical energy obtained by nuclear power plants, which are used in critical situations.

Keywords: Weibull-Weibull; Wind Power Density; Carbon dioxide; Forecast; Expectation-Maximization.

1. Introduction

The world will have a projected 8.5 billion inhabitants by 2030 [1]. The planet is in the process of changing its economic structure, and electricity generation act of one of the ways to supply all this population growth [2].In this sense, there is a demand to invest in new energy sources to meet the population and industrial growth demands. However, one of the biggest concerns in today's globalized world is global warming [3].In 2015, in Paris, about 197 countries signed some proposals in the so-called Paris agreement, committing themselves to reduce the earth's warming by $1,5°C$ [4, 5]. Countries like China, the United States, India, and Russia are the most significant greenhouse gas polluters [6]. In South America, Brazil plays a crucial role in reducing greenhouse gases, as it holds 60% of the largest forest on the planet, the Amazon rainforest [7]. At the 26th (COP26) Conference of the Parties of the (UNFCCC) United Nations Framework Convention on Climate Change, in Glasgow, Scotland, Brazil committed to reducing greenhouse gases by about 50% [8]. For this, investments in renewable energy sources can contribute to this reduction of gases since they are clean and inexhaustible sources. Germany, for example, has committed that by 2030 100% of all energy generated will come from sources that do not pollute the environment [9, 10]. The European Union has adopted the European Green Deal, which envisions zero greenhouse gas emissions in all bloc countries by 2050 [11].

Researchers generally seek ways to estimate the energy consumption needed to meet the world's demand [12]. In order for the climate to improve in certain regions, such as those that suffer from heavy rains or hurricanes due to deforestation, for example, and for the reduction of the greenhouse effect in the atmosphere to be achieved, it is necessary to invest in global policies that adopt low-cost renewable energy sources emission of pollutants [13]. From this perspective, renewable energy research has increased in recent years. As an efficient and ecological alternative, the wind is one of the primary sources of clean energy, as it is an inexhaustible, renewable, accessible, and sustainable resource [14]. Wind energy emerges as one of the leading clean and renewable energy sources, with predictions for the year 2050 to supply about 1/4 to 1/3 of all global electricity demand [15]. Thus, wind speed has been proposed as a raw material for exploration without pollution. The process of converting kinetic energy into wind energy takes place through the movement of air masses; that is, from the wind speed, wind turbines are moved, and energy generation takes place [16]. However, some locations have characteristics that are more favorable to wind energy development than others.

*I am corresponding author: Fábio Sandro dos Santos
Email address: fabio.sandropb@gmail.com (Fábio Sandro dos Santos)

Countries such as China, the United States, Germany, India, Spain, the United Kingdom, and France have frequently stood out as significant wind energy producers [17]. In Brazil, the Northeast region has a high potential for wind generation, and several regional investments are being made [18, 19, 20]. In particular, the state of Bahia is a major producer of energy from wind [21]. On the other hand, the state of Rio Grande do Norte is the biggest producer of wind energy in Brazil [22]. Still, in Brazil, wind energy is the country's second primary energy source [23]. Wind energy contributes to the country being among the world's largest wind energy producers.

Due to the complexity present in wind speed time series, several efforts are constantly employed to characterize the dynamics of these series to develop forms of planning for the generation and distribution of wind energy to society. Statistical models such as the Weibull [24], Gamma [25], Lognormal [26], and Rayleigh [27] probability distributions have broad applications in modeling wind speed series. However, it is only possible to determine a probabilistic model that applies to some historical series. Probability distribution mixture models are essential for predictive models to be applied to the wind potential of regions considered favorable for wind energy generation.

In hourly series with wind speed bimodal behavior, two-parameter probability distribution models fail to fit well [28, 29, 30]. Therefore, mixtures of probability distributions are the most used in these situations, presenting a better fit when compared to distributions with only two parameters [17]. Khamees et al. (2022) [31] used mixtures of probability distributions of two and three components generated from the combination of Weibull, Gamma, and Inverse Gaussian distributions to verify the adequacy of mixture models to bimodal series. Ouarda and Charron (2018) [28] investigated the potential of mixtures of homogeneous and heterogeneous distributions to model bimodal wind speed series, and ten models of the combination of components were used, such as Gamma, Weibull, Gumbel, and Truncated Normal. Indhumathy et al. (2021) [32] used the mixture of the Weibull distribution from the linear combination of two and three components to estimate wind speed series. Wang and Liu (2021) [33] employed the Expectation-Maximization algorithm to estimate the mixing parameters of the two- and three-component Weibull distribution in the multimodal wind speed series adjustment.

In estimating wind generation potential based on probability distribution, estimating parameters of distribution functions is of paramount importance in wind speed applications. There is a range of numerical and artificial intelligence methods that have been developed to improve the estimation of distribution parameters, such as Moment Methods [34], Maximum Likelihood Method [35], Empirical Method [36], and artificial intelligence methods, such as Particle Swarm Optimization (PSO) [37] and Genetic Algorithms (GA) [38]. In our work, we used the Expectation-Maximization (EM) numerical method to estimate the mixing parameters of the Weibull-Weibull distribution. This method was successfully used in fitting bimodal wind speed series in the Northeast region of Brazil at an altitude of $10m$ from the surface [17].

In view of this, this work aims to estimate the potential of wind energy at different heights throughout the Brazilian territory, using models of mixtures of probability distribution and predicting the estimate of wind energy where we do not have prior information based on the

Inverse Distance Weighting (IDW), method predicts a value for some unmeasured locations using values sampled from surrounding weather stations. As far as we searched, we did not find a study similar to the work presented here: with many meteorological stations and using the Weibull-Weibull mixing distribution at different heights. The methodology presented here can be helpful in countries that produce wind energy above their consumption capacity and wish to sell their surplus to neighboring countries. In this way, those countries that do not produce enough wind energy to meet their domestic demand could be supplied with energy generated by another country. Consequently, some countries may decrease the need for energy generated from sources that pollute the environment. For the global temperature to stay below $2°C$, transformations in the electrical system and investments in low-carbon technologies would help countries meet the Paris agreement's goals [39].

The remaining parts of this article are organized as follows: in Section 2, we provide an overview of how the hourly series of wind speeds across Brazil were obtained and treated using the Python programming language and the R software for estimating wind speed series for higher heights from wind series $10m$ from the surface. In section 3 through Hellmann's Exponential Law Equation we determine wind speed to upper heights. In Section 4, we present the Weibull-Weibull distribution mixing model. Section 5 shows the Expectation-Maximization Algorithm's development process for estimating the Weibull-Weibull distribution parameters. In section 6, we show the Wind Power Density estimates. Section 7 presents Inverse Distance Weighting to predict estimates in places where no information was available. Section 8 presents and discusses the results obtained with the mixture model. Finally, in Section 9, we present the conclusions of the article.

2. Description of geographic location and data source

The data used in this work were obtained from the Instituto Nacional de Meteorology (INMET - refers to the National Institute of Meteorology) and correspond to hourly measurements of wind speed in m/s, from 2000/01/01 to 2022/03/01, in the five Brazilian regions (Northeast, North, South, Southeast, and Midwest). In the study, 575 automatic stations were used (as shown in Figure 1) at $10m$ above the ground height. Data were downloaded in real-time from the INMET website and processed using Python. Analyzes were performed using the R software, version 4.1.3.

The Brazilian territory is comprised of around 8,516,000 km^2 [40]. Brazil is divided into five regions: Northeast, North, South, Southeast, and Midwest, according to estimates by IBGE (refers to the Brazilian Institute of Geography and Statistics). The Northeast has an estimated population of 57,667,842 inhabitants. In the North, there are 18,906,962 inhabitants; in the South, 30,402,587 inhabitants; in the Southeast, 89,632,912 inhabitants; in the Midwest, accounting for fewer inhabitants, there are 16,707,336 [41]. The total population of Brazil exceeds 213 million inhabitants, with a Gross Domestic Product of R$ 8.7 trillion in 2021 [42]. Brazil occupies the fifth place among the most populous countries in the world, behind China, with the largest population (about 1.3 billion inhabitants), India with 1.1 billion inhabitants, the United States with 314 million inhabitants, and Indonesia with 229 million inhabitants [43]. In terms of territory size, Brazil is smaller than the Europe

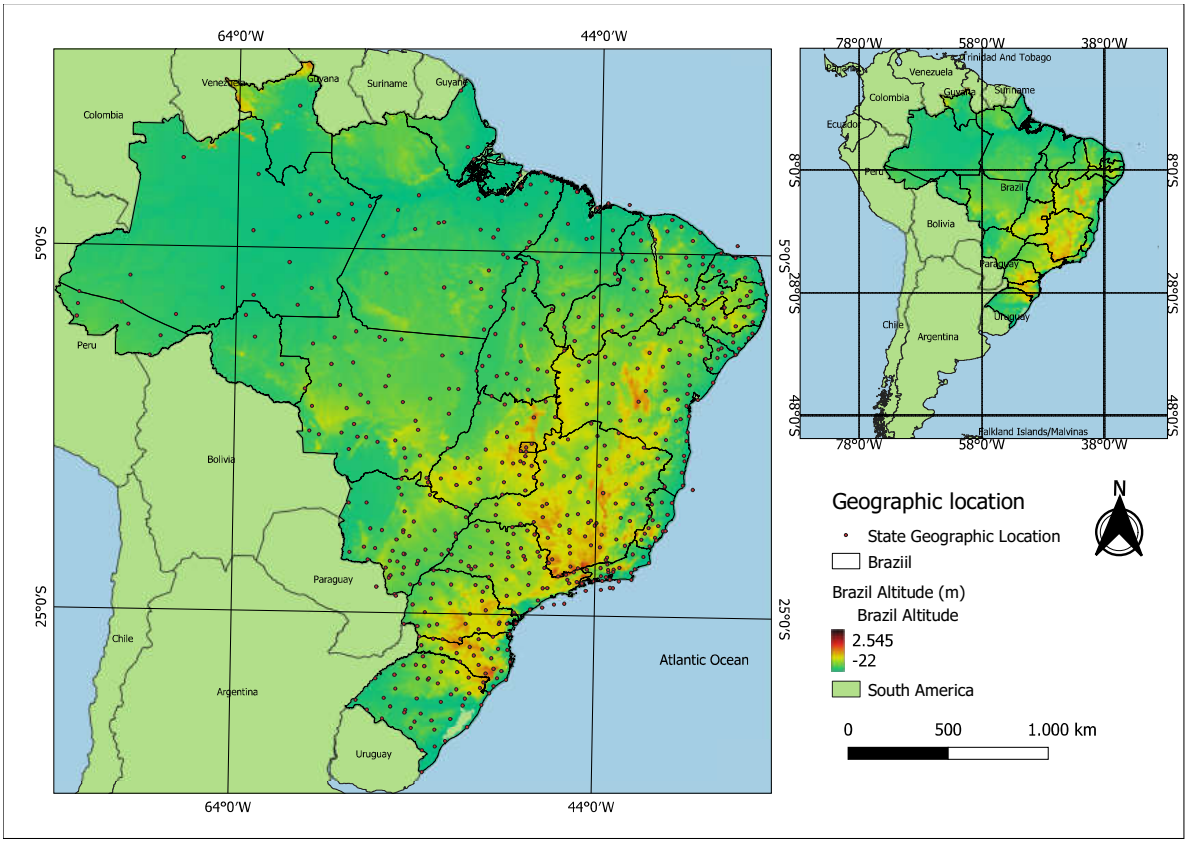


Figure 1: Geographic location of meteorological stations in Brazil.

continent, whose territory is around 10,180,000 km^2. However, Europe is made up of 48 countries.

3. Estimation of wind speed at different heights

To estimate the monthly wind power density throughout Brazil, we collected the wind speed series (at $10m$ from the surface) collected on the INMET website. Subsequently, we estimate the wind speed series for the heights ($25m$, $50m$, $75m$, $100m$, $120m$) based on Hellmann's exponential law, as seen in Eq. 1 below:

$$\frac{v}{v_0} = \left(\frac{h}{h_0}\right)^{\alpha} \tag{1}$$

where v is the wind speed at height h, v_0 is the wind speed at height h_0, and α is the roughness coefficient relative to the nature of the terrain [44].

4. Mixture distribution

In the literature, there is a range of models of unimodal probability distributions with only a single component. Often these models may need to be more suitable for applying frequency to wind speed data. To solve this modeling problem, component mixture models are the most suitable when the series presents a bimodal behavior. In some works, many authors have already achieved satisfactory results using models of mixtures and probability distributions, known as multicomponent distributions [45].

Formed by the linear combination of two or more components of distributions, the probability density function (pdf) of a mixed distribution is given by Eq. 2 [28]:

$$f(v; w, \theta) = \sum_{i=1}^{d} w_i f_i(v; \theta_i), \tag{2}$$

where $\theta_i s$ are the parameters of the i-th distribution, $f_i(v; \theta_i)$ are the independent distributions of the i-th components, d is the number of components, and w_i are the weights of each mixture, so that $\sum_{i=1}^{d} w_i = 1$. The mixture of two distributions is given by the equation below:

$$f(v; w, \theta_1, \theta_2) = w f_1(v; \theta_1) + (1 - w) f_2(v; \theta_2), \tag{3}$$

where the value of w must be in the range $0 < w < 1$, the parameter vectors θ_1 and θ_2 correspond to the first and second components of the distribution.

From this, different combinations of distribution mixtures can be generated, such as Gamma-Gamma [46], Wiebull-Weibull, Weibull-Gamma, Gamma-Weibull [47], Weibul-Burr, Weibull-Lognormal, Weibull-Gumbel, among others existing in the literature [26]. In this article, we work with the mixture of two Weibull distribution models, given in Eq. 4, known as Weibull-Weibull [17].

$$ff_{2,2}(v) = w_1 \frac{k_1}{c_1} \left(\frac{v}{c_1}\right)^{k_1 - 1} exp\left(-\frac{v}{c_1}\right)^{k_1} + w_2 \frac{k_2}{c_2} \left(\frac{v}{c_2}\right)^{k_2 - 1} exp\left(-\frac{v}{c_2}\right)^{k_2}, \tag{4}$$

where $v > 0$, $w_1 = w$ and $w_2 = (1 - w)$, that is, $w_1 + w_2 = 1$ and $k_1, c_1, k_2, c_2 > 0$.

5. Algorithm Expectation Maximization EM

The iterative Expectation-Maximization (EM) algorithm was developed with the aim of fitting models of mixtures of probability distributions. The EM iteration method is based on the maximum likelihood method. EM presents reasonable estimates of parameters in Weibull-Weibull mixture models applied to wind speed series [17, 48].

Given a time series of hourly observations of wind speed $v = (v_1, v_2, v_3, \cdots, v_n)$, we will define the likelihood function of vector v based on the Eq. 5 in the defined sequence:

$$f(v; \Theta) = \prod_{i=1}^{n} f(v_i; \Theta) = \prod_{j=1}^{d} w_j \frac{k_j}{c_i} \left(\frac{v_i}{c_i}\right)^{k_j - 1} exp^{-\left(\frac{v_i}{c_j}\right)^{k_j}} \tag{5}$$

where d is the number of mixture components, w_j are the weights of each mixture, k_j and c_j are the shape and scale parameters of the Weibull probability density function and $\Theta = (w_1, w_2, \cdots, w_d, k_1, k_2, \cdots, k_d, c_1, c_2, \cdots, c_d)$ is a set of parameters.

The log-likelihood function can be defined using Eq. 6 [49]:

$$L\left(v;\Theta\right)=\sum_{i=1}^{N}log\sum_{j=1}^{d}w_i\frac{k_j}{c_j}\left(\frac{v_i}{c_j}\right)^{k_j-1}exp^{-}\left(\frac{v_i}{c_j}\right)^{k_j} \tag{6}$$

As the log-likelihood function is complex in obtaining the set of parameters Θ, they cannot be obtained directly through the partial derivative. In this case, we use the EM iterative algorithm for estimating the parameters of the five-parameter Weibull distribution function.

The Expectation-Maximization algorithm consists of steps E and M [49]:

1. Step E: Define the function Q according to the equation below

$$Q=\sum_{i=1}^{N}\sum_{j=1}^{d}\gamma_{ij}\,log\,w_j+\sum_{i=1}^{N}\sum_{j=1}^{d}\gamma_{ij}\times\left[log\;k_j-log\;c_j+(k_j-1)(log\;v_i-log\;c_j)-\left(\frac{v_i}{c_j}\right)^{h_j}\right]$$

   where γ_{ij} is the prior probability of the j-th observation coming from the v_i mixture component [50], the d-th components of the Weibull-Weibull mixture are obtained by Eq. 7:

$$\gamma_{ij}=E(z_{ij})=\frac{w_jf(vi;k_j,c_j)}{\sum_{j=1}^{d}w_jf(vi;k_j,c_j)}=\frac{w_j\frac{v_i}{c_j}\left(\frac{v_i}{c_j}\right)^{k_j-1}exp^{-(\frac{v_i}{c_j})^{k_j}}}{\sum_{j=1}^{d}w_j\frac{v_i}{c_j}\left(\frac{v_i}{c_j}\right)^{k_j-1}exp^{-(\frac{v_i}{c_j})^{k_j}}} \tag{7}$$

2. Step M: Find the log-likelihood maximization for the estimated parameter of the Weibull-Weibull probability distribution by maximizing the function Q.

5.1. Algorithm execution process EM

The Algorithm 1 shows the basic structure of the EM algorithm.

Algorithm 1: Expectation-Maximization algorithm.

Input: $v=\{v_1,\cdots,v_N\}$ /* v is the wind speed */
Parameter Initialization: $w^{(0)}$, $\boldsymbol{\theta}^{(0)}$ /* Initial guess */
/* Iterations */
for t = *1:T* **do**
  /* E-step */
  $Q^{(t)}=Q\left(\boldsymbol{\theta}^{(t-1)}\right)$ /* Update w and Q */
  /* M-step */
  $\boldsymbol{\theta}^{(t)}=arg_{\,\boldsymbol{\theta}}\,maxL(\boldsymbol{\theta}^{(t-1)};v)$ /* Update $\widehat{\boldsymbol{\theta}}$ */
end
return $\widehat{\boldsymbol{\theta}}_{MLE}$ /* Maximum Likelihood Estimator */

6. Wind Power Density - WPD

In the wind energy sector, the frequency of the wind speed series is used to estimate wind energy density [51]. The wind power density for the Weibull-Weibull distribution mix $ff_{2,2}(v)$ is defined as follows:

$$WPD = \frac{1}{2}\rho\bar{v}^3 W/m^2 \tag{8}$$

$$WPD(v) = \int_0^{+\infty} \frac{1}{2}\rho v^3 ff_{2,2}(v)\, dv W/m^2 \tag{9}$$

where ρ is the air density, v^3 is the wind speed (in m/s) cubed, $ff_{2,2}(v)$ is the Weibull-Weibull probability density function, $\bar{v}^3$ is the average wind speed cubed. For this article, WPD estimates were calculated across Brazil at different heights. The different heights of wind speed were estimated by Eq.1. The difference between Eq. 8 and Eq. 9 is that the first considers the average wind speed cubed, and the second considers the probability distribution that was fitted to the wind speed series. This process is essential for the generation of wind energy in the country. Such energy can be integrated with other types, such as hydro power, and those from renewable sources, such as solar energy and biomass energy, both important sources for Brazil. Investments in these types of energy are of great importance for Brazil to comply with the Paris agreement to reduce greenhouse gases and reduce the risks of lack of electricity due to climate change [52]. By 2029, Brazil aims to double its current wind energy capacity [53], further reinforcing the importance of a study like this. The perspective for Brazil until the year 2027 is that investments in technologies can replace the sources of thermoelectric plants powered by diesel and fuel oil [54].

6.1. Inverse Distance Weighting (IDW)

In this work, we used the deterministic spatial interpolation method Inverse Distance Weighting [55]. In the literature, this method has been applied in different studies, obtaining good results with spatial interpolation in the mapping of annual precipitation in Bosnia and Herzegovina [56], wind speed estimation for Iraq [57], the study of precipitation and trends over the basin Mahaweli, Sri Lanka [58] and monthly rainfall in Thailand [59], for example. This method was initially proposed by Shepard [60]. The IDW is based on the idea that the closer the estimated value is to the real one, the greater the influence on the predicted value of the more distant ones. The IDW mathematical equation can be defined from the following formula [61]:

$$\hat{Z}_{(s_0)} = \sum_{i=1}^{n} \frac{1/d(s_0, s_i)^p}{\sum_{i=1}^{n}(1/d(s_0, s_i)^p)} z(s_i), \qquad p > 1 \tag{10}$$

where $\hat{Z}_{(s_0)}$ is the estimated value for each point is the observed value in s_0, $z(s_i)$ is the distance between ordered pairs s_0 and s_i and p is a parameter where the value is defined. In this article, the value of $p = 3$ was used. A mesh of $(0.1, 0.1)$ was used to obtain good results, equivalent to an area of $11km^2$.

7. Results and Discussion

Based on information from 575 automatic stations distributed throughout the Brazilian territory, Figure 2 shows the maps of the mean statistics and standard deviation of the analyzed hourly historical series. Wind speed information was spatialized for all of Brazil using the IDW method. Although the analyses were carried out for all months studied, we illustrate the results for the month with the lowest average wind speed at the height of $120m$ and for the month with the highest average: May and September, respectively.

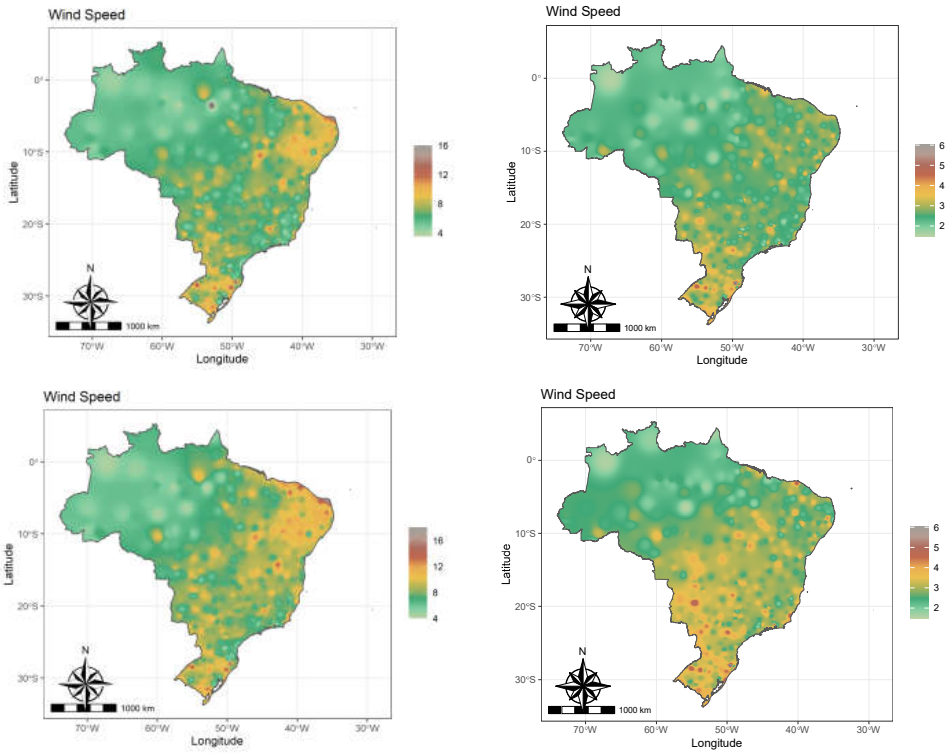


Figure 2: On the left is the mean and on the right is the standard deviation of Brazil's average hourly wind speed from 2010-01-01 to 2022-03-31. Upper figures are for May and lower figures are for September

.

The highest average speed in Brazil was observed in the Northeast region, followed by the country's South, Southeast and Midwest regions. Specifically, the coastal part of the Northeast region, with emphasis on Rio Grande do Norte and Ceará, presented the highest average wind in all of Brazil. On the other hand, the region with the lowest average wind incidence was the Brazilian North region. One reason for this is that this region contains the Amazon rainforest, which directly and indirectly influences the climate in Brazil. In the North, the average wind for May ranged from 2.0 to $7.5m/s$. In the Northeast region, the average varied between 5.0 and $12.0m/s$. In the South Region, the average wind varied between 3.0 and $8.0m/s$. While in the Southeast and Midwest regions, the wind speed varied between 3.0 and $8.0m/s$.

For September, the country's North region had the lowest average wind speed (around $4.0m/s$). On the other hand, the South and Northeast regions presented an average above $8.0m/s$, which is excellent for installing wind farms. Furthermore, the Southeast and Midwest regions presented relatively low averages for wind farm installations, except for the coastal areas of Rio de Janeiro and Espírito Santo. In Brazil, it was found that the farther from the equator, the greater the variability in the wind speed series. The South, Southeast, and Midwest regions have the greatest dispersion in the series. The North region had the lowest standard deviation.

Fig. 3 illustrates the wind speed (in m/s) at $120m$ for May and September in all Brazilian regions. It is crucial to observe regions with wind speeds exceeding $3.0m/s$, as a wind turbine only generates enough energy to supply a subsystem if it reaches this minimum speed [17]. The Northeast and South regions have more wind speeds for both May and September. The North region had the lowest percentage of wind. For more than 90% of the time, the Northeast region indicated that the wind speed is more significant than $3.0m/s$. However, in the North region, only a few meteorological stations showed a percentage of wind speed values above $3.0m/s$. These stations do not even represent 50% of the total, thus making a possible installation of a wind farm in this locality unfeasible. The Southeast and Midwest regions, on average, have percentages of wind speed values above $3.0m/s$ that vary from 75% to up to 100% of the total observations in some stations. In the case of the installation of a wind farm in the Northeast, the wind turbine can spend more than 90% of the time generating energy in the region, as well as in large part of the South of the country.

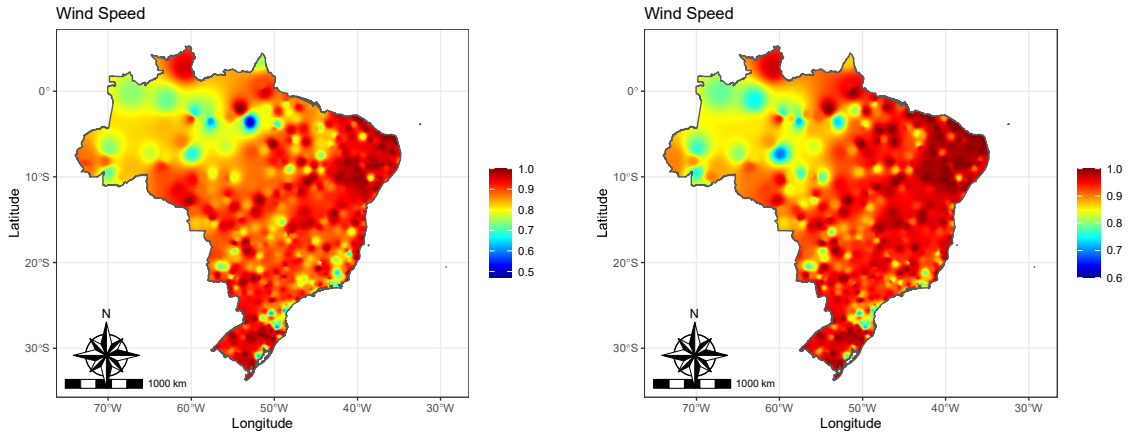


Figure 3: On the left, the percentage of values above 3.0m/s for the month of May. On the right, the percentage of values above 3.0m/s for the month of September at a height of 120m.

Figure 4 shows the estimates of the six parameters for the mixture of two Weibull-Weibull, resulting from the mixture of a Weibull(k_1, c_1) and another Weibull(k_2, c_2), with a weight parameter w_1 for the first Weibull and w_2 for second Weibull. The Weibull-Weibull mixture has shown promising results regarding series with bimodal behavior. A similar result was

seen in the work of Carta and Ramirez [62]. The parameters of the Weibull-Weibull mixture were obtained with the EM algorithm, as shown in Section 5.

Through the spatial mapping of the parameters of the Weibull-Weibull mixture, it was possible to identify the geographic location of the Brazilian regions according to their climatic and topological characteristics in each of the regions of Brazil. The five Brazilian regions have different characteristics. The Northeast is dry and arid [63], and the average annual rainfall is $500mm$ in the central part of the region and $1500mm$ in the coastal part [64]. The North region is influenced by the Amazon rainforest and has an average of $2000mm$ rainfall [65]. The Southeast has reasonable annual precipitation rates (around $1314mm$) [66]. The Midwest is characterized by different precipitation conditions, such as the South Atlantic Convergence Zone (SACZ) [40]. The southern region is characterized by having a cold and humid climate for most of the year, with the Pampa Biome and the Atlantic Forest Biome [63]. The results of this article can contribute to a better understanding of wind speed behavior in regions close to the equator, extensive forests such as the Amazon, and areas of Caatinga, Cerrado, Atlantic Forest, Pampa, and Pantanal. In addition to areas close to the coastal marine, Brazil has approximately 7.300 kilometers of the coastal area [67]. This fact can directly affect the behavior of the wind speed series in different Brazilian regions, as there are differences in energy production between regions.

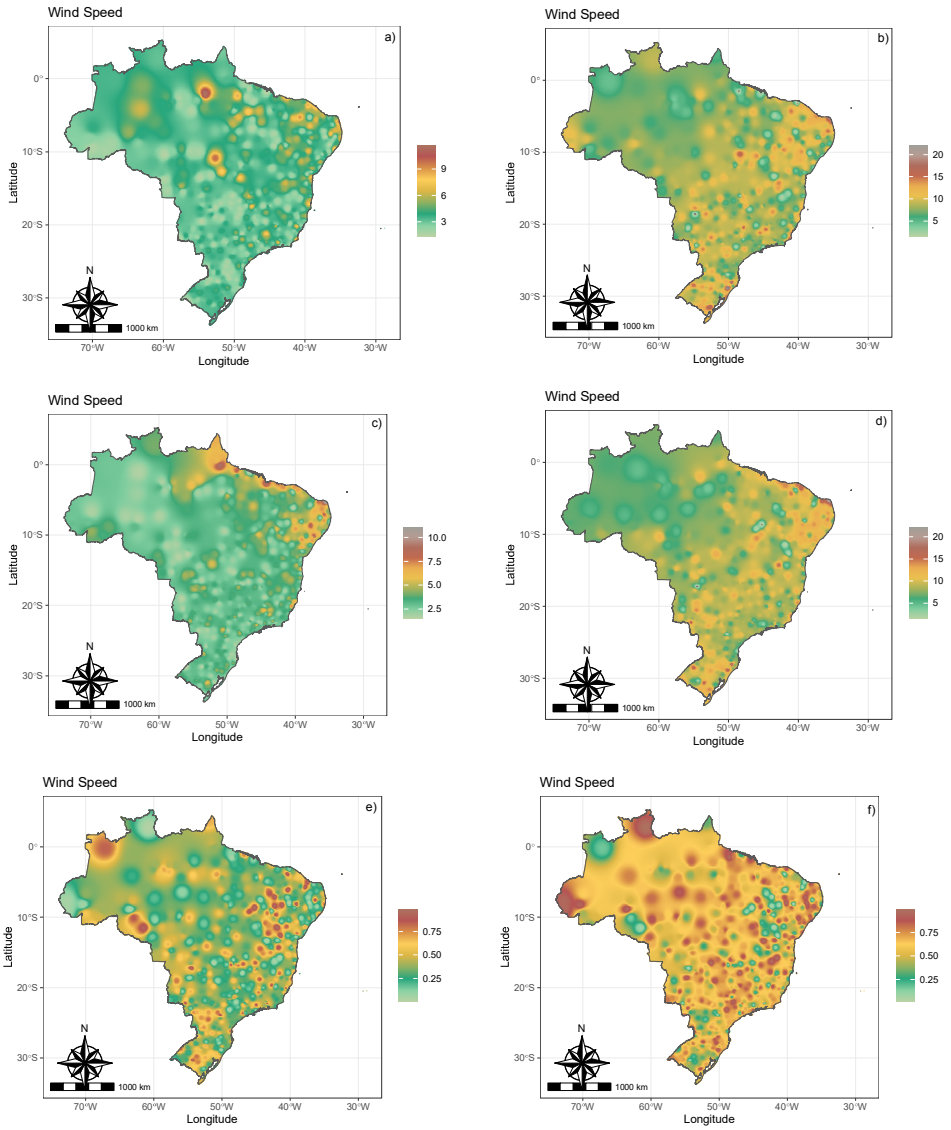


Figure 4: (a) Scale 1, (b) shape 1, (c) scale 2, (d) shape 2, (e) w1 and (f) w2 for the hourly average wind speed series in Brazil.

Figure 4 presents the estimates of the Weibull-Weibull parameters for a height of 120 m. In the coastal area of the states of Pernambuco, Paraíba, Rio Grande do Norte, Ceará, Piauí, Maranhão, Rio de Janeiro, and in parts of the state of Bahia, these coastal areas, according to the parameter c_1, showed good potential for the installation of wind farms. The highest values found for c_1 were in the coastal region of Rio Grande do Norte, in the coastal and central parts of Ceará, in the east and west of the state of Pará, in the Northeast of Mato Grosso, and a significant part of the state of Bahia. Similar characteristics were observed in the coastal part of Rio Grande do Norte [68, 69]. The lowest values of c_1 are concentrated in the states of Santa Catarina, Mato Grosso do Sul, Paraná, Goiás, Acre, Rondônia, Roraima,

and Amapá, the latter two being above the equator. As for the parameter c_2, the highest values are concentrated among the states of Bahia, Pará, and Amapá, the latter two partially above the equator. For the parameters of k_1 and k_2, the North region presented the lowest values concerning the other Brazilian regions. For parameters w_1 and w_2, we did not observe a pattern of behavior in the investigated regions.

According to our analyses, the ordering of regions for better use of the wind resource (considering its greater availability) is Northeast, South, Southeast, Midwest, and North. This result is coherent with the literature since Brazil's Northeast region has the largest wind power generation, followed by the South and Southeast regions [67]. From the parameters of the Weibull-Weibull distribution mixture, it was possible to estimate the Wind Power Density of the five Brazilian regions, as we can see in Fig. 5.

Aquila et al. [70] used the two-parameter Weibull to estimate wind energy in the 63 main cities of São Paulo (Brazil). The author noted that some cities had higher costs in implementing a wind farm. Juárez et al. [71] also used the two-parameter Weibull in their work and showed remarkable growth in the wind energy sector for the year 2021. Duca, Fonseca, and Oliveira [72] used the Gamma and Weibull probability distributions with two parameters for forecasting wind speed in Bahia. In this study, it was observed that Gamma had a better fit in June, while in September, Weibull had a better fit in the series. The authors also suggest in future works the use of probability distribution mixtures in a more significant time window for a better performance of the models in the adjustments of the wind speed series.

We use probability distribution mixture models via the EM algorithm to improve wind power estimates in the five Brazilian regions. The series used in these works present a bimodal behavior. Therefore, the Weibull-Weibull mixtures present good results in long-term hourly series [17]. May be proved to be one of the months with the lowest potential for wind energy generation in Brazil. The Northeast and South regions were the two regions that showed the most significant potential for wind generation. September was the month with the most significant potential for wind energy generation. The Northeast and South regions were the ones with the highest energy potential. The North of Brazil had the lowest wind potential, followed by the Southeast and Midwest, respectively. The results showed that the Northeast and the South are the two central regions for installing wind farms in Brazil.

In Figure 7, we present the results for WPD in each of the five Brazilian regions for all 12 months of the year. In the Northeast region, we note that between January and May, there is a decrease in the wind potential, where in May, at the height of 100 m, the WPD value was 295.1259 W/m^2, and at the height of 120m the WPD was 347.7526 W/m^2. From May, there is a gradual increase in wind generation potential. The projection peak for the average wind energy generation occurs in September, in which the WPD at the height of 100m was 553.386 W/m^2, and for a height of 120 m, the WPD estimate was 652 .0656 W/m^2. In the Northeast region, September has almost double the potential of May. This fact may be related to the dry period in September, contributing to an increase in the average wind speed and, consequently, a more significant potential for wind power generation. In Figure 7, one can see that the other regions of Brazil have the same behaviors of Wind

Power Density, except in the South Region, where August has the highest WPD values. From the information obtained in each meteorological station, new information was forecast where there still needed more information on wind generation potential based on the IDW method.

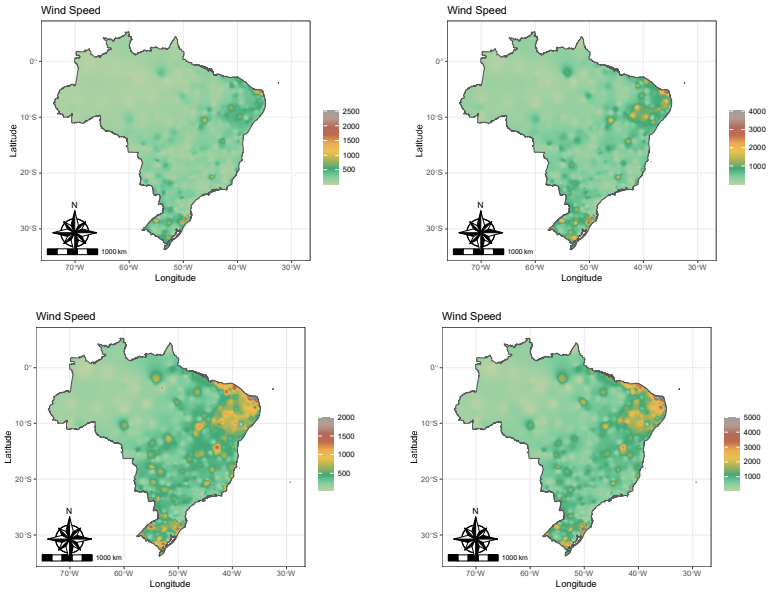


Figure 5: On the left is the Average Wind Power Density, and on the right is the Wind Power Density based on the Weibull-Weibull distribution for Northeast Brazil from 2010-01-01 to 2022-03-31. Upper figures are for May and lower figures are for September.

Thus, we can identify the possible places where wind farms can be installed. The Northeast region had the most significant potential for wind power generation in the country. In second place is the South region, third is the Southeast region, fourth is the Midwest region, and fifth is the North region. The wind energy projection for the month with the greatest potential for energy generation in 120m of the Northeast is 1465.167 W/m^2 in September. In the South region, August's estimate was $1098.041 W/m^2$. In the Southeast, it was $758.8887 W/m^2$ for August. In the Midwest, it was $744.854 W/m^2$ for September, and in the North, it was $569.3334 W/m^2$ for September. The month with the lowest wind power generation potential for the Northeast region was April, with an estimated value of $690.935 W/m^2$. The South region has a smaller production in January, April, and March with amounts of 762.0361 W/m^2, 758.974 W/m^2 and 755.036 W/m^2, respectively, with slight variation between these observed months.

The lowest production occurred in the Southeast and North regions in April and May. For the Southeast region, the month of May presented an estimate of $427.2629 W/m^2$, while for the North, in April, the estimate was 281.5192. In the Midwest region, April was the month with the lowest density of wind energy generation, with an estimate of $449.9184 W/m^2$. We can then define the wind potential density in all five Brazilian regions at different heights

based on the hourly wind speed series. Mixing the Weibull-Weibull distributions made it possible to adjust the different series spread across the country. The model proved efficient in adjusting wind speed series in large regions and different types of geography. The Weibull-Weibull model can also be used in large areas such as Brazil and, for example, continents such as South America and Europe. A similar result using the Weibull-Weibull mixture was observed by Santos et al. [17], where the authors showed that for the Northeast region of Brazil, at the height of $10m$ from the surface, mixing this model using the EM algorithm achieved excellent results.

A significant result is that the Northeast region has the potential to generate wind energy about four times more when compared to the North region. With this result, the Northeast region will likely be able to produce wind energy to supply its entire population and sell or distribute its surplus to the North region. During the so-called wind harvest, which runs from June to December, the wind energy generated is sufficient to meet the entire demand of the Northeast region and about 20% of the demand of all Brazilian people [73]. In Brazil, energy is distributed in four subsystems: South (S), Southeast/Midwest (SE/CW), Northeast (NE), and North (N) [74]. To distribute the surplus to energy companies in the regions, existing subsystem networks could be used and reduce energy generation costs in the country.

In the same way, the North region can send the excess energy produced in its hydroelectric plants to the Northeast since the Northeast is a region that suffers a lot from a lack of rain. In the same sense, the South Region of Brazil can distribute surplus energy to neighboring regions such as the Southeast and Midwest, and it can sell surplus energy to neighboring countries such as Paraguay, Argentina, and Uruguay. The projection of Brazil's energy integration when comparing the years from 2015 to 2050, going from $130.1GW$ to $283.1GW$, with a growth of 117.6% in the year 2050 [75].

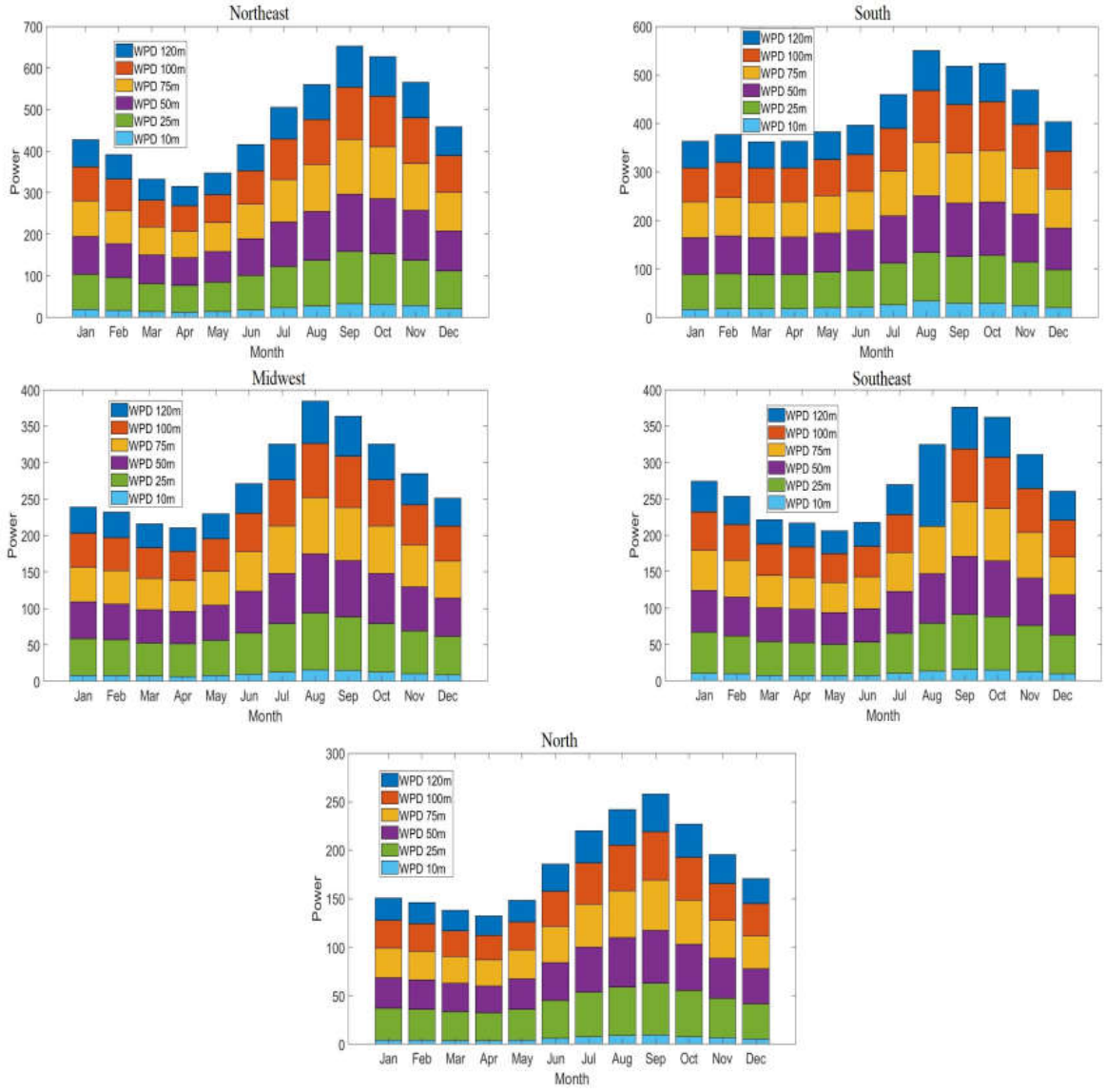


Figure 6: Average Wind Power Density was analyzed monthly for the five Brazilian regions from 2010-01-01 to 2022-31-03.

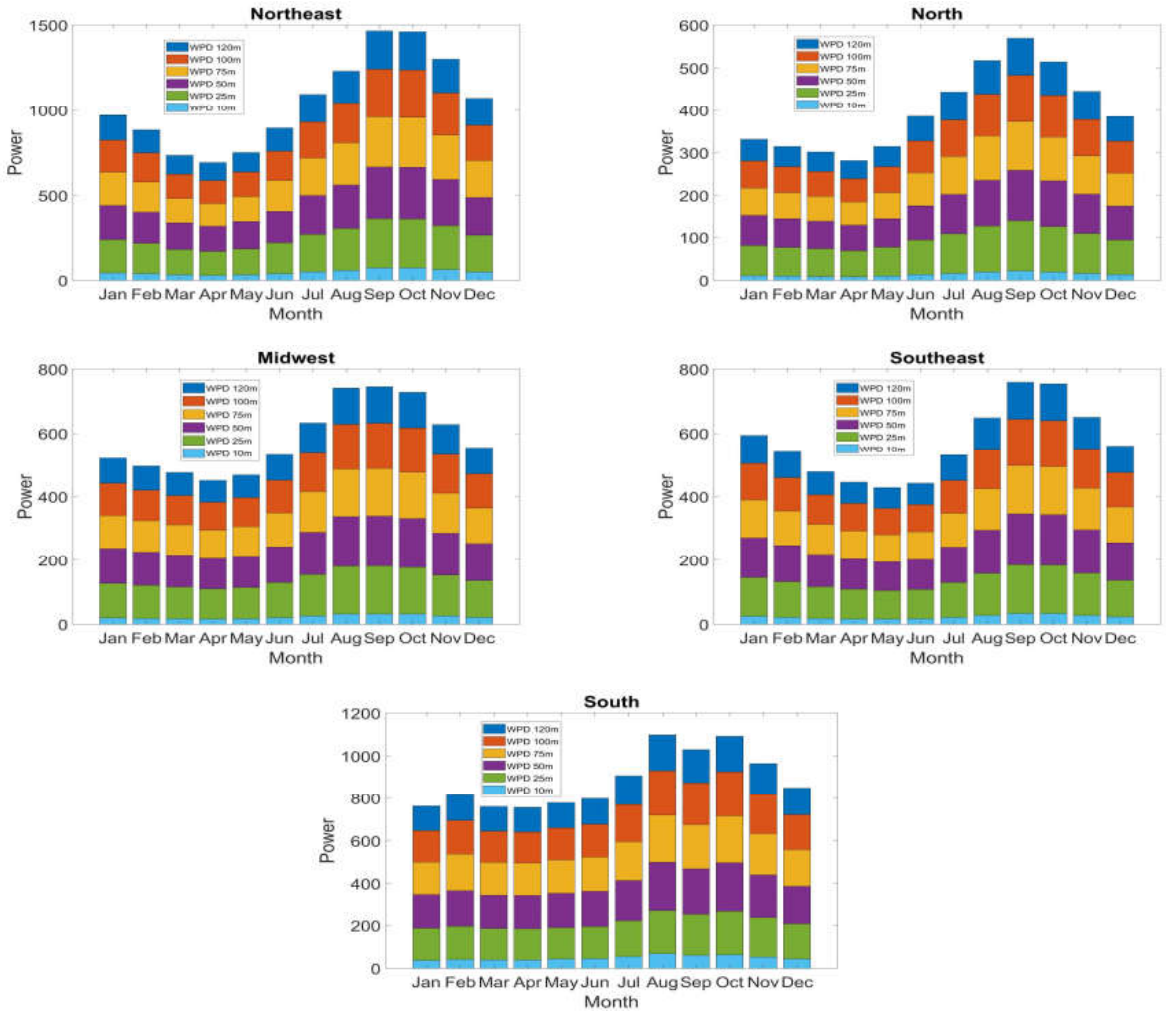


Figure 7: : Average Wind Power Density Estimate using the Weibull-Weibull distribution for the five regions of Brazil from 2010-01-01 to 2022-31-03.

The authors emphasize the importance of this work in investments in technologies to help integrate wind energy with other energy sources that already exist in Brazil, mainly the Wind-Hydric integration. In this integration, it would be possible to produce hydrogen as a new energy source, as discussed by Nadaleti et al.[76], using the surplus energy generated. In addition, other integrations are possible, such as Wind-Solar and Wind-Biomass. This could greatly help the country to reduce greenhouse gases in the atmosphere. Thus, Brazil would meet the goals agreed upon at the United Nations Conference on Climate Change at COP26 in Glasgow, Scotland, in 2021 [77], where Brazil committed to reducing greenhouse gases by 50% by 2030 [8]. At COP26, around 29 countries, including the United Kingdom, Canada, Germany, and Italy, committed to investing in clean and renewable sources by 2022. Brazil had already committed to achieving the Paris target of limiting global temperature to 1.5 ∘C [66]. In this way, wind energy can contribute to Brazil achieving this critical

reduction of the greenhouse effect, being an alternative form of new sources of income and technologies in land management through government incentives [77]. Faced with such an adverse scenario caused by the COVID-19 pandemic, investment in renewable energy can collaborate with GDP growth in the country [78].

In Figure 8, we analyze the predominant wind direction for Brazil, where the arrows indicate the wind's direction. We note that the wind speed is predominant in the Northeast from East to West. In the North, the direction oscillates between the East/West and North/South, with some moments in the Southeast/North direction. In the Midwest and Southeast regions of the country, the direction that most predominates is the East/West direction and sometimes the North/South direction. This indicates that in these two regions of Brazil, the wind direction oscillates a lot compared to the Northeast region. In the South region, wind direction predominance is from Brazil's East/West direction.

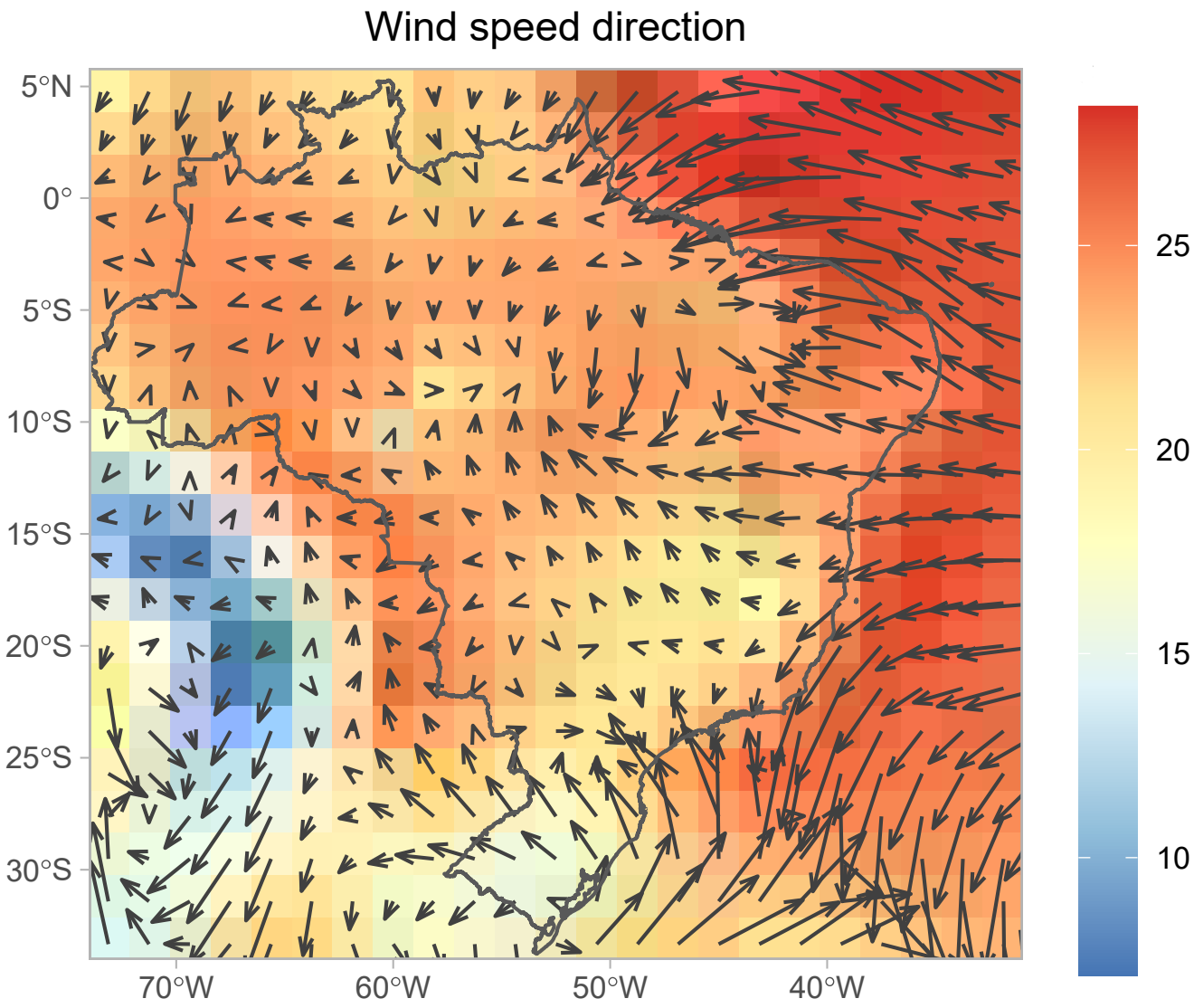


Figure 8: Predominant wind direction throughout Brazil.

8. Conclusion

In the results obtained in this article, the authors show a new way of treating and modeling the wind speed series for installing wind farms and wind energy generation in the five regions of Brazil. In addition, the authors also emphasize the importance of using probability distribution mixture models, especially the Weibull-Weibull model, to solve series adjustment problems with bimodal behavior in large areas such as Brazil, in addition to research in the energy field renewable. Other results that we considered important were the predictability of wind speed and the density of potential wind energy generation in large

areas, making it possible to obtain information regarding wind energy in places where we needed prior information. With this, the investment sectors of the public and private sectors can make their decisions about the implementation of wind farms in certain regions.

This work identified the months with each region's highest and lowest wind power density. It is possible to help public and private initiatives with this information to adapt and protect themselves from possible energy rationing due to the lack of water in the reservoirs since more than 60% of Brazil's energy is generated from its hydroelectric plants. Our results can cooperate with the possible integration of wind energy with other energy sources, thus reducing all risks of a collapse in the country's energy distribution, which would be a disaster. In future work, we suggest using artificial intelligence algorithms such as Particle Swarm Optimization (PSO) in the Weibull-Weibull mixing model.

ACKNOWLEDGMENT

This work received financial support from the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financing Code 001.

References

[1] U. N. Organization, https://www.un.org/en/global-issues/population.

[2] G. Cui, S. Pei, Z. Rui, B. Dou, F. Ning, J. Wang, Whole process analysis of geothermal exploitation and power generation from a depleted high-temperature gas reservoir by recycling CO2, Energy 217 (2021) 119340.

[3] M. Usman, D. Balsalobre-Lorente, A. Jahanger, P. Ahmad, Pollution concern during globalization mode in financially resource-rich countries: Do financial development, natural resources, and renewable energy consumption matter?, Renewable Energy 183 (2022) 90–102.

[4] R. T. Gotama, A. T. H. Seng, M. Tasaloti, R. T. Fernandez, A critical study on the state of globalization today, focusing on selected current global issues (2021).

[5] U. N. F. C. on Climate Change. Key aspects of the Paris Agreement, https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement/key-aspects-of-the-paris-agreement.

[6] D. Nong, P. Simshauser, D. B. Nguyen, Greenhouse gas emissions vs CO2 emissions: Comparative analysis of a global carbon tax, Applied Energy 298 (2021) 117223.

[7] T. M. Rosan, K. K. Goldewijk, R. Ganzenmüller, M. O'Sullivan, J. Pongratz, L. M. Mercado, L. E. Aragao, V. Heinrich, C. Von Randow, A. Wiltshire, et al., A multi-data assessment of land use and land cover emissions from Brazil during 2000–2019, Environmental Research Letters 16 (7) (2021) 074004.

[8] A. Fünfgeld, " Brazil must be back!"-but real climate action is possible only after Bolsonaro (2021).

[9] E. Winquist, M. Van Galen, S. Zielonka, P. Rikkonen, D. Oudendag, L. Zhou, A. Greijdanus, Expert views on the future development of biogas business branch in Germany, the Netherlands, and Finland until 2030, Sustainability 13 (3) (2021) 1148.

[10] M. Z. Jacobson, The cost of grid stability with 100% clean, renewable energy for all purposes when countries are isolated versus interconnected, Renewable Energy 179 (2021) 1065–1075.

[11] B. Hrnčić, A. Pfeifer, F. Jurić, N. Duić, V. Ivanović, I. Vušanović, Different investment dynamics in energy transition towards a 100% renewable energy system, Energy 237 (2021) 121526.

[12] N. D. Rao, J. Min, A. Mastrucci, Energy requirements for decent living in India, Brazil and South Africa, Nature Energy 4 (12) (2019) 1025–1032.

[13] D. Gielen, F. Boshell, D. Saygin, M. D. Bazilian, N. Wagner, R. Gorini, The role of renewable energy in the global energy transformation, Energy Strategy Reviews 24 (2019) 38–50.

[14] A. Khosravi, R. Koury, L. Machado, J. Pabon, Energy, exergy and economic analysis of a hybrid renewable energy with hydrogen storage system, Energy 148 (2018) 1087–1102.

[15] V. Kati, C. Kassara, Z. Vrontisi, A. Moustakas, The biodiversity-wind energy-land use nexus in a global biodiversity hotspot, Science of The Total Environment 768 (2021) 144471.
[16] A. Kalmikov, Wind power fundamentals in wind energy engineering: A handbook for onshore and offshore wind turbines ed letcher tm (2017).
[17] F. S. dos Santos, K. K. F. do Nascimento, J. da Silva Jale, T. Stosic, M. H. Marinho, T. A. Ferreira, Mixture distribution and multifractal analysis applied to wind speed in the Brazilian Northeast region, Chaos, Solitons & Fractals 144 (2021) 110651.
[18] H. Vasconcellos, L. C. Couto, Estimation of socioeconomic impacts of wind power projects in Brazil's Northeast region using Interregional Input-Output Analysis, Renewable and Sustainable Energy Reviews 149 (2021) 111376.
[19] W. D. Jacondino, A. L. da Silva Nascimento, L. Calvetti, G. Fisch, C. A. A. Beneti, S. R. da Paz, Hourly day-ahead wind power forecasting at two wind farms in northeast brazil using wrf model, Energy 230 (2021) 120841.
[20] V. P. da Silva, M. L. d. M. Galvão, Onshore Wind Power Generation and Sustainability Challenges in Northeast Brazil: A Quick Scoping Review, Wind 2 (2) (2022) 192–209.
[21] N. B. P. de Souza, J. V. C. dos Santos, E. G. S. Nascimento, A. A. B. Santos, D. M. Moreira, Long-range correlations of the wind speed in a northeast region of Brazil, Energy 243 (2022) 122742.
[22] M. F. Sobrinho Junior, M. C. Ramirez Hernandez, S. S. Albano Amora, E. R. Costa de Morais, Perception of Environmental Impacts of Wind Farms in Agricultural Areas of Northeast Brazil, Energies 15 (1) (2021) 101.
[23] M. T. Tolmasquim, T. de Barros Correia, N. A. Porto, W. Kruger, Electricity market design and renewable energy auctions: The case of Brazil, Energy Policy 158 (2021) 112558.
[24] S. Miao, Y. Gu, D. Li, H. Li, Determining suitable region wind speed probability distribution using optimal score-radar map, Energy conversion and management 183 (2019) 590–603.
[25] R. K. Samal, M. Tripathy, Estimating wind speed probability distribution based on measured data at Burla in Odisha, India, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 41 (8) (2019) 918–930.
[26] K. Rajapaksha, K. Perera, Wind speed analysis and energy calculation based on mixture distributions in Narakkalliya, Sri Lanka, Journal of the National Science Foundation of Sri Lanka 44 (4) (2016).
[27] S. Suwarno, R. Rohana, Wind speed modeling based on measurement data to predict future wind speed with modified Rayleigh model, International Journal of Power Electronics and Drive Systems 12 (3) (2021) 1823.
[28] T. B. Ouarda, C. Charron, On the mixture of wind speed distribution in a Nordic region, Energy Conversion and Management 174 (2018) 33–44.
[29] W. Xie, P. Huang, Extreme estimation of wind pressure with unimodal and bimodal probability density function characteristics: A maximum entropy model based on fractional moments, Journal of Wind Engineering and Industrial Aerodynamics 214 (2021) 104663.
[30] W. Wang, T. Okaze, Statistical analysis of low-occurrence strong wind speeds at the pedestrian level around a simplified building based on the Weibull distribution, Building and Environment 209 (2022) 108644.
[31] A. K. Khamees, A. Y. Abdelaziz, Z. M. Ali, M. M. Alharthi, S. S. Ghoneim, M. R. Eskaros, M. A. Attia, Mixture probability distribution functions using novel metaheuristic method in wind speed modeling, Ain Shams Engineering Journal 13 (3) (2022) 101613.
[32] D. Indhumathy, D. Narmatha, C. Indirani, K. Meenambika, Wind resource assessment using improved mixture Weibull distribution, in: 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), IEEE, 2021, pp. 1–4.
[33] Z. Wang, W. Liu, Wind energy potential assessment based on wind speed, its direction and power data, Scientific reports 11 (1) (2021) 1–15.
[34] U. C. Ben, A. E. Akpan, C. C. Mbonu, C. H. Ufuafuonye, Integrated technical analysis of wind speed data for wind energy potential assessment in parts of Southern and Central Nigeria, Cleaner Engineering and Technology 2 (2021) 100049.

[35] P. K. Chaurasiya, S. Ahmed, V. Warudkar, Study of different parameters estimation methods of Weibull distribution to determine wind power density using ground based Doppler SODAR instrument, Alexandria Engineering Journal 57 (4) (2018) 2299–2311.
[36] S. Deep, A. Sarkar, M. Ghawat, M. K. Rajak, Estimation of the wind energy potential for coastal locations in India using the Weibull model, Renewable Energy 161 (2020) 319–339.
[37] T. C. Carneiro, S. P. Melo, P. C. Carvalho, A. P. d. S. Braga, Particle swarm optimization method for estimation of Weibull parameters: a case study for the Brazilian northeast region, Renewable energy 86 (2016) 751–759.
[38] M. B. Koca, M. B. Kilic, Y. Şahin, Using genetic algorithms for estimating Weibull parameters with application to wind speed, An International Journal of Optimization and Control: Theories & Applications (IJOCTA) 10 (1) (2020) 137–146.
[39] P. Fragkos, H. L. van Soest, R. Schaeffer, L. Reedman, A. C. Köberle, N. Macaluso, S. Evangelopoulou, A. De Vita, F. Sha, C. Qimin, et al., Energy system transitions and low-carbon pathways in Australia, Brazil, Canada, China, EU-28, India, Indonesia, Japan, Republic of Korea, Russia and the United States, Energy 216 (2021) 119385.
[40] K. F. Silveira Marinho, L. d. M. Barbosa Andrade, M. H. Constantino Spyrides, C. M. Santos e Silva, C. P. de Oliveira, B. Guedes Bezerra, P. Rodrigues Mutti, Climate profiles in Brazilian microregions, Atmosphere 11 (11) (2020) 1217.
[41] I. B. de Geografia e Estatística, https://www.gov.br/pt-br/noticias/financas-impostos-e-gestao-publica/2021/08/populacao-brasileira-chega-a-213-3-milhoes-de-habitantes-estima-ibge.
[42] I. B. de Geografia e Estatística, https://www.ibge.gov.br/explica/pib.php.
[43] A. F. Nelwan, R. Dalimi, C. Hudaya, A new formula to quantify the national energy security of the world's top ten most populous nations, International Journal of Energy Economics and Policy 11 (1) (2021) 394.
[44] L. C. S. Rocha, G. Aquila, P. R. Junior, A. P. de Paiva, E. de Oliveira Pamplona, P. P. Balestrassi, A stochastic economic viability analysis of residential wind power generation in Brazil, Renewable and Sustainable Energy Reviews 90 (2018) 412–419.
[45] R. K. Samal, Probabilistic modelling of 80 m mast measured wind resource: A case study., Energy Sources, Part A: Recovery, Utilization, and Environmental Effects (2021) 1–15.
[46] S. Atapattu, C. Tellambura, H. Jiang, A mixture gamma distribution to model the snr of wireless channels, IEEE transactions on wireless communications 10 (12) (2011) 4193–4203.
[47] R. Kollu, S. R. Rayapudi, S. Narasimham, K. M. Pakkurthi, Mixture probability distribution functions to model wind speed distributions, International Journal of energy and environmental engineering 3 (1) (2012) 1–10.
[48] J.-Y. Shin, T. B. Ouarda, T. Lee, Heterogeneous mixture distributions for modeling wind speed, application to the uae, Renewable Energy 91 (2016) 40–52.
[49] Q. Hu, Y. Wang, Z. Xie, P. Zhu, D. Yu, On estimating uncertainty of wind energy with mixture of distributions, Energy 112 (2016) 935–962.
[50] J.-Y. Shin, J.-H. Heo, C. Jeong, T. Lee, Meta-heuristic maximum likelihood parameter estimation of the mixture normal distribution for hydro-meteorological variables, Stochastic Environmental Research and Risk Assessment 28 (2) (2014) 347–358.
[51] K. Bagci, T. Arslan, H. E. Celik, Inverted Kumarswamy distribution for modeling the wind speed data: Lake Van, Turkey, Renewable and Sustainable Energy Reviews 135 (2021) 110110.
[52] M. Lima, L. Mendes, G. Mothé, F. Linhares, M. de Castro, M. Da Silva, M. Sthel, Renewable energy in reducing greenhouse gas emissions: Reaching the goals of the paris agreement in brazil, Environmental Development 33 (2020) 100504.
[53] O. Turkovska, G. Castro, M. Klingler, F. Nitsch, P. Regner, A. C. Soterroni, J. Schmidt, Land-use impacts of brazilian wind power expansion, Environmental Research Letters 16 (2) (2021) 024010.
[54] N. Carvalho, D. B. Viana, M. M. de Araújo, J. Lampreia, M. Gomes, M. Freitas, How likely is Brazil to achieve its NDC commitments in the energy sector? A review on Brazilian low-carbon energy perspectives, Renewable and Sustainable Energy Reviews 133 (2020) 110343.

[55] K. Shukla, P. Kumar, G. S. Mann, M. Khare, Mapping spatial distribution of particulate matter using Kriging and Inverse Distance Weighting at supersites of megacity Delhi, Sustainable Cities and Society 54 (2020) 101997.
[56] D. Nusret, S. Dug, Applying the inverse distance weighting and kriging methods of the spatial interpolation on the mapping the annual precipitation in Bosnia and Herzegovina (2012).
[57] S. Ali, A. Mahdi, A. H. Shaban, Wind speed estimation for iraq using several spatial interpolation methods, Environ. Prot 1 (2) (2012).
[58] U. Pawar, U. Rathnayake, Spatiotemporal rainfall variability and trend analysis over Mahaweli Basin, Sri Lanka, Arabian Journal of Geosciences 15 (4) (2022) 1–16.
[59] N. Chutsagulprom, K. Chaisee, B. Wongsaijai, P. Inkeaw, C. Oonariya, Spatial interpolation methods for estimating monthly rainfall distribution in Thailand, Theoretical and Applied Climatology 148 (1) (2022) 317–328.
[60] D. Shepard, A two-dimensional interpolation function for irregularly-spaced data, in: Proceedings of the 1968 23rd ACM national conference, 1968, pp. 517–524.
[61] A. Bărbulescu, C. Șerban, M.-L. Indrecan, Computing the beta parameter in idw interpolation by using a genetic algorithm, Water 13 (6) (2021) 863.
[62] J. Carta, P. Ramirez, Analysis of two-component mixture weibull statistics for estimation of wind speed distributions, Renewable energy 32 (3) (2007) 518–531.
[63] P. Regoto, C. Dereczynski, S. C. Chou, A. C. Bazzanela, Observed changes in air temperature and precipitation extremes over Brazil, International Journal of Climatology 41 (11) (2021) 5125–5142.
[64] D. T. Rodrigues, W. A. Gonçalves, M. H. C. Spyrides, C. M. Santos e Silva, Spatial and temporal assessment of the extreme and daily precipitation of the Tropical Rainfall Measuring Mission satellite in Northeast Brazil, International Journal of Remote Sensing 41 (2) (2020) 549–572.
[65] J. Villamayor, T. Ambrizzi, E. Mohino, Influence of decadal sea surface temperature variability on Northern Brazil rainfall in CMIP5 simulations, Climate dynamics 51 (1) (2018) 563–579.
[66] V. T. M. de Morais Junior, L. A. G. Jacovine, C. M. M. E. Torres, E. B. B. M. Alves, H. N. de Paiva, R. Alcantara-de la Cruz, J. C. Zanuncio, Early assessment of tree species with potential for carbon offset plantations in degraded area from the southeastern Brazil, Ecological Indicators 98 (2019) 854–860.
[67] A. Vinhoza, R. Schaeffer, Brazil's offshore wind energy potential assessment based on a spatial multi-criteria decision analysis, Renewable and Sustainable Energy Reviews 146 (2021) 111185.
[68] E. J. d. A. Dantas, L. P. Rosa, N. F. d. Silva, M. G. Pereira, Wind power on the Brazilian Northeast Coast, from the whiff of hope to turbulent convergence: the case of the Galinhos Wind Farms, Sustainability 11 (14) (2019) 3802.
[69] A. Gorayeb, C. Brannstrom, Toward participatory management of renewable energy resources (wind-farm) in northeastern Brazil, Mercator (Fortaleza) 15 (2016) 101–115.
[70] G. Aquila, W. T. Nakamura, P. R. Junior, L. C. S. Rocha, E. de Oliveira Pamplona, Perspectives under uncertainties and risk in wind farms investments based on Omega-LCOE approach: An analysis in São Paulo state, Brazil, Renewable and Sustainable Energy Reviews 141 (2021) 110805.
[71] A. A. Juárez, A. M. Araújo, J. S. Rohatgi, O. D. Q. de Oliveira Filho, Development of the wind power in brazil: Political, social and technical issues, Renewable and sustainable energy reviews 39 (2014) 828–834.
[72] V. E. Duca, T. C. Fonseca, F. L. C. Oliveira, A generalized dynamical model for wind speed forecasting, Renewable and Sustainable Energy Reviews 136 (2021) 110421.
[73] Desenvolvimento da eólica no brasil (2022).
URL https://abeeolica.org.br/energia-eolica/o-setor/
[74] G. R. Moraes, B. A. Ambrósio, J. L. Pereira, D. Issicaba, A. F. Aquino, I. C. Decker, Impact analysis of covid-19 pandemic on the electricity demand, frequency control and electromechanical oscillation modes of the Brazilian Interconnected Power System using low voltage WAMS data, International Journal of Electrical Power & Energy Systems 142 (2022) 108266.
[75] T. Santos, Regional energy security goes South: Examining energy integration in South America, Energy Research & Social Science 76 (2021) 102050.

[76] W. C. Nadaleti, G. B. dos Santos, V. A. Lourenço, The potential and economic viability of hydrogen production from the use of hydroelectric and wind farms surplus energy in brazil: A national and pioneering analysis, International Journal of Hydrogen Energy 45 (3) (2020) 1373–1384.

[77] C. H. Silva Junior, N. S. Carvalho, A. Pessôa, J. B. Reis, A. Pontes-Lopes, J. Doblas, V. Heinrich, W. Campanharo, A. Alencar, C. Silva, et al., Amazonian forest degradation must be incorporated into the COP26 agenda, Nature Geoscience 14 (9) (2021) 634–635.

[78] C. Magazzino, M. Mele, G. Morelli, The relationship between renewable energy and economic growth in a time of covid-19: a machine learning experiment on the brazilian economy, Sustainability 13 (3) (2021) 1285.

General Conclusions

Through the analysis performed, it was possible to identify the areas favorable to installing wind farms and installing solar panels in Brazil. The results were obtained using the Multifractal Detrended Fluctuation Analysis methods and the mixtures of probability distributions, such as the Weibull-Weibull mixture. In the adjustment of the mixture models, the methods of parameter optimization were used: Moments' Method, Maximum Likelihood Method, Expectation-Maximization Algorithm, and the Particle Swarm Optimization artificial intelligence algorithm. Based on these methods, the best model fitting the study series of observations was defined. In addition, the parameters of the models were estimated to identify areas conducive to installing wind and solar farms in Brazilian regions. Based on the Inverse Distance Weighting method, it was possible to predict the most accurate behavior of wind speed and solar radiation where there was no a priori information.

With the inclusion of the methods adopted, it was found that in Brazil, in regions such as the Northeast of the country, the production of energy from renewable and clean sources can be about four times higher than in the North region, for example. Therefore, it is possible to distribute all the surplus energy generated to the North, if this region at some point will need it. The same situation was observed in the Southern region, where this region can distribute its surplus energy to the Southeast and Midwest regions and sell its surpluses to countries bordering the region, such as Argentina, Uruguay, and Paraguay.

Finally, through this research, new proposals for future work are using the Random Forest algorithm and Artificial Neural Networks to estimate wind speed at different heights based on the existing series and taking into account the geography of each locality. This proposal is necessary for the continuity of this research and may generate an even more accurate range of results compared to those in this research. After using the proposed methods, it's significant to use neural networks to predict the potential of wind power generation and distribution throughout Brazil. In this way, it creates the distribution map of surplus renewable energies in the country.