

Evaluation of Weibull parameters by different methods for farms

Flavio Aristone¹, Amaury de Souza^{1*}, Ivana Pobocikova², Razika Ihaddadene³, Guilherme Henrique Cavazzana¹, Carlos José dos Reis⁴, Nabila Ihaddadene³, Hamilton Germano Pavão¹, Widinei Alves Fernandes¹ and Elias Silva de Medeiros⁵

¹Universidade Federal de Mato Grosso do Sul, Cx. Postal 549, Av. Costa e Silva, s/n., 79070-900, Campo Grande, Mato Grosso do Sul, Brazil. ²Departamento de Matemática Aplicada, Faculdade de Engenharia Mecânica, Universidade de Zilina, Zilina, Eslováquia. ³Departamento de Engenharia Mecânica, Universidade de M'Sila, M'Sila, Argélia. ⁴Departamento de Agricultura, Estatística e Experimentação, Universidade Federal de Alfenas, Alfenas, Minas Gerais, Brazil, ⁵Faculdade de Ciências Exatas e Tecnologia, Universidade Federal da Grande Dourados, Dourados 79825-070, MS, Brasil.*Author for correspondence. E-mail: amaury.de@uol.com.br

ABSTRACT. One of the most prevalent clean and sustainable forms of energy produced worldwide is the power created from wind runoff. Wind turbines ought to be installed in areas with favorable circumstances to transform mechanical wind energy into electricity. Finding appropriate ways to predict the energy produced by a wind farm using the Weibull distribution is the main goal of this work. Theoretical techniques have been applied to calculate Weibull selected characteristics using experimental data gathered at the campus of *Universidade Federal de Mato Grosso do Sul* (UFMS) in Brazil. These data were gathered 10 meters above the surface. The effectiveness of four statistical techniques that are frequently used in the energy industry are compared: the standard energy factor method; the least squares regression method; the moment method; and the mean standard deviation method in estimating Weibull parameters. The root mean square error, Chi-square error, Kolmogorov-Smirnov test, and coefficient of determination are used to contrast the statistical methodologies. The results demonstrated that the least squares regression approach performs less well than other methods. The standard energy factor approach, the moment method, and the mean standard deviation method are the most effective techniques when modifying Weibull distribution curves for the assessment of wind speed data. The data analysis confirms that these three strategies are fully applicable if the wind speed distribution closely matches the Weibull distribution.

Keywords: wind energy; Weibull distribution; probability distribution function (PDF); cumulative distribution function (CDF); shape parameter (k); scale parameter (c).

Received on September 12, 2022.

Accepted on April 12, 2023.

Introduction

The survival and progress of human civilization depend on energy. The production of energy problem extends to affect all of humanity. Together with population expansion and the quickening rise of the world economy, energy consumption is constantly growing. Concerns have been expressed regarding the rise of greenhouse gases (GHG) emissions and the growing consequences on the climate change due to the pervasive use of traditional fuels in almost all aspects of daily life (Liu et al., 2020). The finite nature of fossil fuel reserves must be noted. Humans may entirely exhaust fossil fuels within a few hundred years, after they gradually developed over millions of years in nature (Baz et al., 2021). Additionally, the carbon generated by fossil fuels is converted into carbon dioxide (CO₂) that raises the atmospheric CO₂ concentration and exacerbates the greenhouse gasses, changing the global temperature and disturbing the ecological balance (Arenas-López & Badaoui, 2020).

Future societies will mostly rely on renewable sources to meet their energy needs. Energy efficiency, environmental sustainability, carbon reduction and environmental preservation have all become progressively essential to political leaders. It appears that there is a substantial market for these cleaner energies. A viable renewable energy source that can be exploited is wind energy. Various nations have embraced wind energy growth to strengthen their energy grids and safeguard the environment. Furthermore, wind energy is essential for economic expansion since it expands opportunities for employment and promotes the progress of technology and research (Liu et al., 2019; Kandpal & Dhingra, 2021). In the energy production sector, wind energy already attracts many attention (Jansen et al., 2020). Currently, 10.6% of all power generated in Brazil comes from wind generation, according to data from the ONS (*Operador Nacional do*

Sistema Elétrico). This number has major implications, especially for Brazil's Northeast Region (NEB), which is already largely dependent on the wind matrix to supply its energy needs. In fact, the NEB is constantly lessening its societal vulnerability (Freire & Fontgalland, 2022).

According to the ONS, the most favorable periods for wind generation are winter and spring, mainly between August and September in the NEB, where records of production are routinely logged. Currently, most wind projects are concentrated in the NEB as such a region has multiscale meteorological systems which, therefore, favor the continuity and stability of the action of winds seasonally (Dantas, Rodrigues, Silva, Aquino, & Thomaz, 2021).

Shi, Dong, Xiao, and Huang (2021) reviewed the wind speed distribution models used for energy assessment, applicable to different wind regimes. Different potential models of wind speed distribution should be considered when modeling wind speed data in each location, therefore several methods of parameter estimation and evaluation criteria were selected, the quality of the adjustment was evaluated, and their advantages and disadvantages were discussed.

The characteristics of wind speed distribution vary continuously, both geographically and temporally. Wind energy assessment must consider local geographic elements such as climate, topography, differences in thermal properties of land and seas, and focus on long-term variations in wind characteristics.

Therefore, the purpose of this study is to evaluate the effectiveness of four statistical methods used to calculate the Weibull parameters applied to study the possibility of wind energy applications in Campo Grande.

Materials and methods

The *Centro de Monitoramento do Tempo e do Clima* of the Mato Grosso do Sul State (CEMTEC-MS) region provides the wind data used in this investigation. The location of the meteorological station is given in Table 1, where a Belford-three mug type anemometer was used to measure wind speed at 10 meter height. The uncertainty of the mean wind speeds is 2 within 95% of confidence level (Manwell, McGowan, & Rogers, 2010).

The monthly average wind speed has been derived from the measured wind speeds. There are obvious drawbacks from using monthly average wind speeds, such as losing exceptional high or low wind speeds throughout the month and being unable to monitor variations in wind speed during the day. Nevertheless, it is possible to evaluate seasonal fluctuations of the wind speed by using the available monthly average values, which makes it easier to analyze wind data.

Weibull distribution

Weibull probability distribution is used to model wind data in this study with two parameters. It is most easily introduced in terms of its probability density function (PDF), which has the form:

$$f(v) = \left(\frac{k}{c}\right) \left(\frac{v}{c}\right)^{k-1} \exp \left[-\left(\frac{v}{c}\right)^k \right] \quad (1)$$

The corresponding cumulative probability function (CDF) is given as:

$$F(v) = 1 - \exp \left[-\left(\frac{v}{c}\right)^k \right] \quad (2)$$

where:

c is the scale parameter in m s^{-1} , k is the dimensionless shape parameter, and v is the random variable (wind speed). Four techniques were used to determine those two parameters c and k . The Mean Standard Deviation Method (MSDM), the standard Energy Pattern Factor Method (EPFM), the Least Squares Regression Method (LSRM), and the Moment Method (MOM).

Table 1. Geographical coordinates of the area under study.

City	Latitude	Longitude	Altitude (m)	Area (km)	Period of measures
Campo Grande, Mato Grosso do Sul, Brazil	20 th 26' 34" S	54 th 38' 47" W	532	8118,4	Jan 1960 to Dec 2020

Least Squares Regression Method (LSRM)

The least squares method is commonly used in engineering and mathematical problems. From a double logarithmic transformation of CDF given by Equation (2) into a linear form, the parameters c and k as discussed in Justus, Hargraves, Mikhail, and Graber (1978) can be found thru the equations:

$$\ln[-\ln(1 - F(v))] = -k \ln(c) + k \ln(v) \quad (3)$$

which can be represented by the straight line:

$$Y = a + bX \quad (4)$$

where:

$$Y = \ln[-\ln(1 - F(v))], X = \ln(v), a = -k \ln(c), \text{ and } b = k.$$

The CDF of the Weibull distribution can be determined from Benard's approximation:

$$F(v) = \frac{i-0.3}{N+0.4} \quad (5)$$

where:

i is the measured wind speed order number and N is the number of observations (Benard & Bos-Levenbach, 1953). Therefore, the computation of the line coefficients a and b is a prerequisite for determining the parameters c and k . The coefficients a and b can be obtained applying the least squares method. The following Equation 6 and 7 can be used to estimate the Weibull shape and scale parameters:

$$k = \frac{N \sum_{i=1}^N X_i Y_i - \sum_{i=1}^N X_i \sum_{i=1}^N Y_i}{N \sum_{i=1}^N X_i^2 - (\sum_{i=1}^N X_i)^2} \quad (6)$$

$$c = \exp\left(\frac{k \sum_{i=1}^N X_i - \sum_{i=1}^N Y_i}{Nk}\right) \quad (7)$$

where:

$$Y_i = \ln\left[-\ln\left(1 - \frac{i-0.3}{N+0.4}\right)\right], \text{ and } X_i = \ln(v_i).$$

Standard Energy Pattern Factor Method (EPFM)

This method defines the energy pattern factor (EPF) as the ratio of the mean cube wind speed by the cube of the mean wind speed. The definition of the EPF is as follows:

$$EPF = \frac{\frac{1}{N} \sum_{i=1}^N v_i^3}{(\bar{v})^3} \quad (8)$$

where:

v_i is the observed wind speed (m s^{-1}), N is number of the wind speed observation, and \bar{v} is the mean wind speed. After estimating the EPF , the following Equation 9 and 10 can be used to estimate the Weibull shape and scale parameters:

$$k = 1 + \frac{3.69}{EPF^2} \quad (9)$$

$$c = \frac{\bar{v}}{\Gamma\left(1 + \frac{1}{k}\right)} \quad (10)$$

Moment Method (MOM)

The moment method is an additional technique for estimating the Weibull distribution parameters. It is calculated from the mean wind speed \bar{v} and the standard deviation σ of wind speed, which are expressed as:

$$\bar{v} = \frac{1}{N} \sum_{i=1}^N v_i \quad (11)$$

$$\sigma = c \left(\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right) \right)^{\frac{1}{2}} \quad (12)$$

$$\frac{\sigma}{\bar{v}} = \sqrt{\frac{\Gamma\left(1 + \frac{2}{k}\right)}{\Gamma\left(1 + \frac{1}{k}\right)^2}} - 1 \quad (13)$$

The estimates of the Weibull shape and scale parameters are calculated as follows:

$$k = \left(\frac{0.9874}{\frac{\sigma}{\bar{v}}} \right)^{1.0893} \quad (14)$$

$$v' = c\Gamma\left(1 + \frac{1}{k}\right) \quad (15)$$

$\Gamma(x)$ is the gamma function:

$$\Gamma(x) = \int_0^{\infty} y^{x-1} e^{-y} dy \quad (16)$$

where:

$y = \left(\frac{v}{c}\right)^k$. After transformation it becomes:

$$\frac{v}{c} = y^{x-1}; x = 1 + \frac{1}{k} \quad (17)$$

$$v = c + \Gamma\left(1 + \frac{1}{k}\right) = 0,8525 + 0,0135k + e^{-(2+3(k-1))} \quad (18)$$

Mean standard deviation Method (MSDM)

The mean standard deviation method is based on the mean wind speed v' and standard deviation from the wind speed σ . This method may be considered as a unique case of the MOM. The Weibull shape and scale parameters are estimated as follows:

$$k = \left(\frac{\sigma}{v'}\right)^{-1.086} \quad (19)$$

$$c = \frac{v'}{\Gamma\left(1 + \frac{1}{k}\right)} \quad (20)$$

Performance index

Four quantitative tests have been used to test the performance of the approaches mentioned above to determine the optimal method of analysis. These tests are used to validate the accuracy of the predicted wind speed distribution derived from the Weibull probability density function. They are described in the sequence.

Root Mean Square Error (RMSE):

$$RMSE = \left[\frac{1}{N} \sum_{i=1}^N (y_{i,m} - x_{i,w})^2 \right]^{\frac{1}{2}} \quad (21)$$

B) Chi-square error (χ^2):

$$\chi^2 = \frac{\sum_{i=1}^N (y_{i,m} - x_{i,w})^2}{x_{i,w}} \quad (22)$$

C) Kolmogorov - Smirnov test (KS):

$$Q_{95} = \frac{1.36}{\sqrt{N}} \quad (23)$$

D) Coefficient of determination (R^2):

$$R^2 = \frac{\sum_{i=1}^N (y_{i,m} - z_{i,v})^2 - \sum_{i=1}^N (y_{i,m} - x_{i,w})^2}{\sum_{i=1}^N (y_{i,m} - z_{i,v})^2} \quad (24)$$

For all these equations, N is the number of wind speed observations, $y_{i,m}$ is the observed data, $x_{i,w}$ is the estimated (predicted) data using the Weibull distribution, and $z_{i,v}$ is the average value calculated from observed data $y_{i,m}$.

The $RMSE$ has been used to assess the difference between the estimated and the observed values. The chi-square error (χ^2) has been used to determine if there exist any difference between the estimated and observed values. The method showing the best results is elected by the lowest values of $RMSE$ and χ^2 in each case, since both $RMSE$ and χ^2 should be as close to zero as possible. The coefficient of determination (R^2) determines the correlation between the estimated and observed values. The ideal value of R^2 is 1. In general, the best estimation method can be selected according to the lowest values of $RMSE$, χ^2 , and the highest values of R^2 .

Results

Figure 1 shows the monthly analysis of the wind speed at 10 m above the ground at Campo Grande from 1960-2020, namely monthly mean wind speed, standard deviation, coefficient of variation, median,

minimum, maximum, and skewness. It demonstrates clear fluctuations over time. The difference between the highest and lowest wind speeds can be used to depict the speed range averaging as 3.02 m s^{-1} . The mean wind speed varied from 2.37 to 3.91 m s^{-1} . Similarly, the standard deviations ranged from 1.45 to 0.80 , and the critical values at 95% of the reliable level in the Kolmogorov-Smirnov test (Q_{95}) were 0.0248 and 0.0246 for months with 31 days and 30 days, respectively. Finally, the maximum error in the CDF never exceeds the corresponding significant values, which implies that the proposed technique is applicable for generating the variables necessary in the selection of the viable site of potential wind power turbines.

The monthly coefficient of variation (COV) ranged from 29.5 to 44.2%, with an average value of 39.1% (Figure 1), with its highest value in November and lowest in October, and when compared to other regions is lower than Cardiff – Wales, where the annual average is 59.0%, Ankara – Turkey with 64.6%, Athens – Greece with 72.8%, Davos – Switzerland with 79.4% and Canberra – Australia with 81.4% (Celik, 2003a; 2003b; 2004).

Only the month of November presented a negative asymmetry, which means that the tail has more left. The skewness in May is greater than 1, indicating a highly right skewed distribution. In March, August, and October the values of skewness indicate moderately right skewed distributions.

The monthly averages of the Weibull distribution parameters for all the period, estimated by four methods, are shown in Figure 1 and 2 from data extracted of an anemometer that was placed 10 m height. The variation of the Weibull distribution parameters is monthly dependent. The highest value of the monthly shape parameter k was observed in April for EPFM (4.13) and the lowest in February for MOM (2.34). For the scale parameter c , the highest value of 4.34 m s^{-1} was observed in October for EPFM and the lowest value of 2.43 m s^{-1} in September for LSRM.

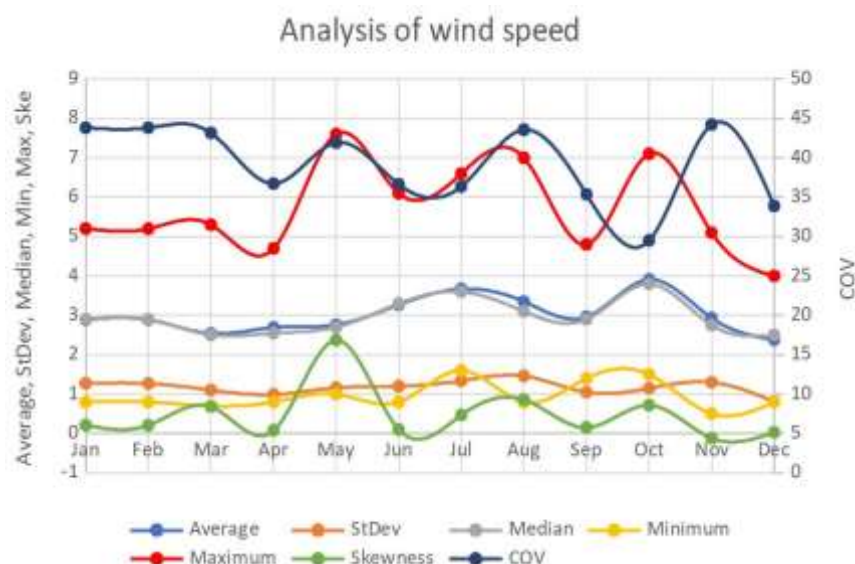


Figure 1. Monthly wind speed analysis at 10 m above ground level in Campo Grande (MS) – Brazil, from 1960-2020.

Figures 1, 2 and 3 show the descriptive statistics and monthly variations in the shape (k) and scale (c) parameters of the Weibull distribution using the four approaches (MSDM, MOM, EPFM and LSRM) considered in the study. The difference amongst the values of the shape parameters k obtained using different methods is more significant than for the scale parameters c , respectively. The dimensionless Weibull shape parameters k showed lesser significant values obtained by EPFM and LSRM. Consistency was achieved in Weibull scale parameters c . Both MOM and MSDM have a comparable range of Weibull parameters c and k .

Figure 4 presents the histograms of monthly wind speed variation in Campo Grande for all 61-year period. It has been found that the highest wind speeds have been measured in October as 3.91 m s^{-1} . Conversely, the lowest wind speeds have been recorded in December as 2.38 m s^{-1} . The variation in amplitude observed for the wind speed (1.6 m s^{-1}) can be attributed to the winter solstice. At this time, penetration of frontal systems (SF) occurred in the studied region (Souza et al., 2022).

Shaban, Resen, and Bassil (2020) have collected experimental data on the city of Al-Najaf, Iraq and determined the best theoretical ensemble to be used as Weibull's parameters. Data were collected at different heights (profiled) above the ground at 10, 30, and 50 m. Four criteria were used to evaluate different adjustment procedures of wind data; namely the root mean square error, chi-square, correlation coefficient,

and coefficient of determination. The valid distribution model has been determined according to the specific conditions that provide the best results. The results indicate that the Equivalent Energy Method is the finest when the value of the Weibull shape parameter and the Weibull scale parameter are calculated for the Weibull distribution curves at that location, based on the adjustment quality tests.

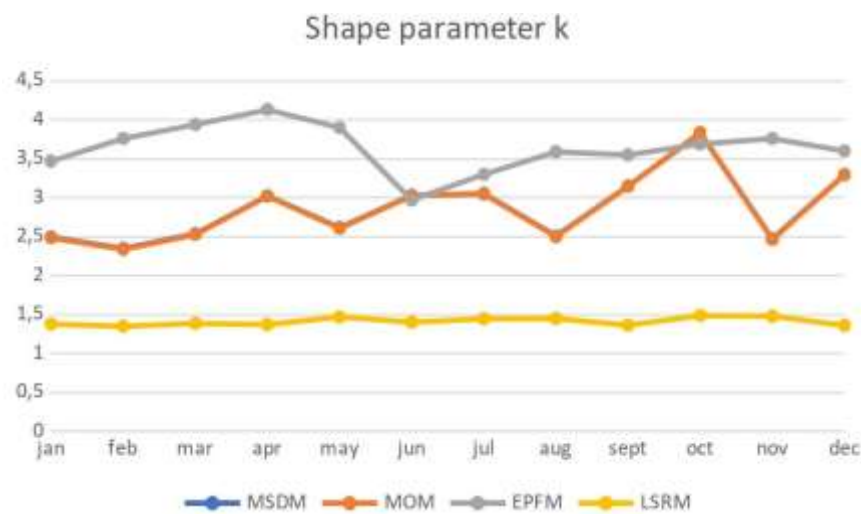


Figure 2. Monthly Weibull shape parameters k estimated by various methods at 10 m above ground level in Campo Grande in the period 1960-2020.

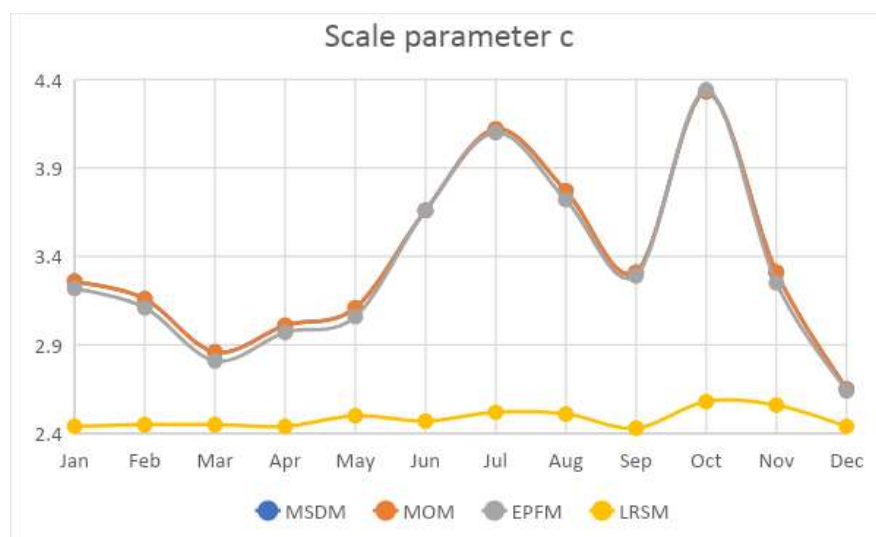


Figure 3. Monthly Weibull scale parameters c estimated by various methods at 10 m above ground level in Campo Grande in the period 1960-2020.

Mdee (2020) used the Weibull distribution function with parameters of shape (k) and scale (c) to predict the wind speed distribution for different locations, and applied the correlation technique to evaluate the performance of analytical methods using various empirical methods. The analytical methods presented include Method of Moments (MOM), Standard Energy Density (EPD), Graphical Method (GM), and Maximum Likelihood Estimator (MLE). Five empirical methods were graphically evaluated using the GM to estimate the two Weibull parameters. Four empirical methods indicated a correlation coefficient of two Weibull parameters (r) greater than 0.9503. Again, five empirical methods for parameter k and four empirical methods for parameter c were used in the special case of MOM and EPD methods. The results indicated four empirical methods to estimate the parameter k and two empirical methods to estimate the parameter c with r greater than 0.7. When empirical methods of k and c parameters correlate, six matching patterns were identified from 20 possible pairs. The comparison of GM, MLE, and selected pairs of k and c values predicted by empirical methods resulting from MOM and EPD indicated the mean coefficient of best fit as 0.8851 and mean square error as 0.2083.

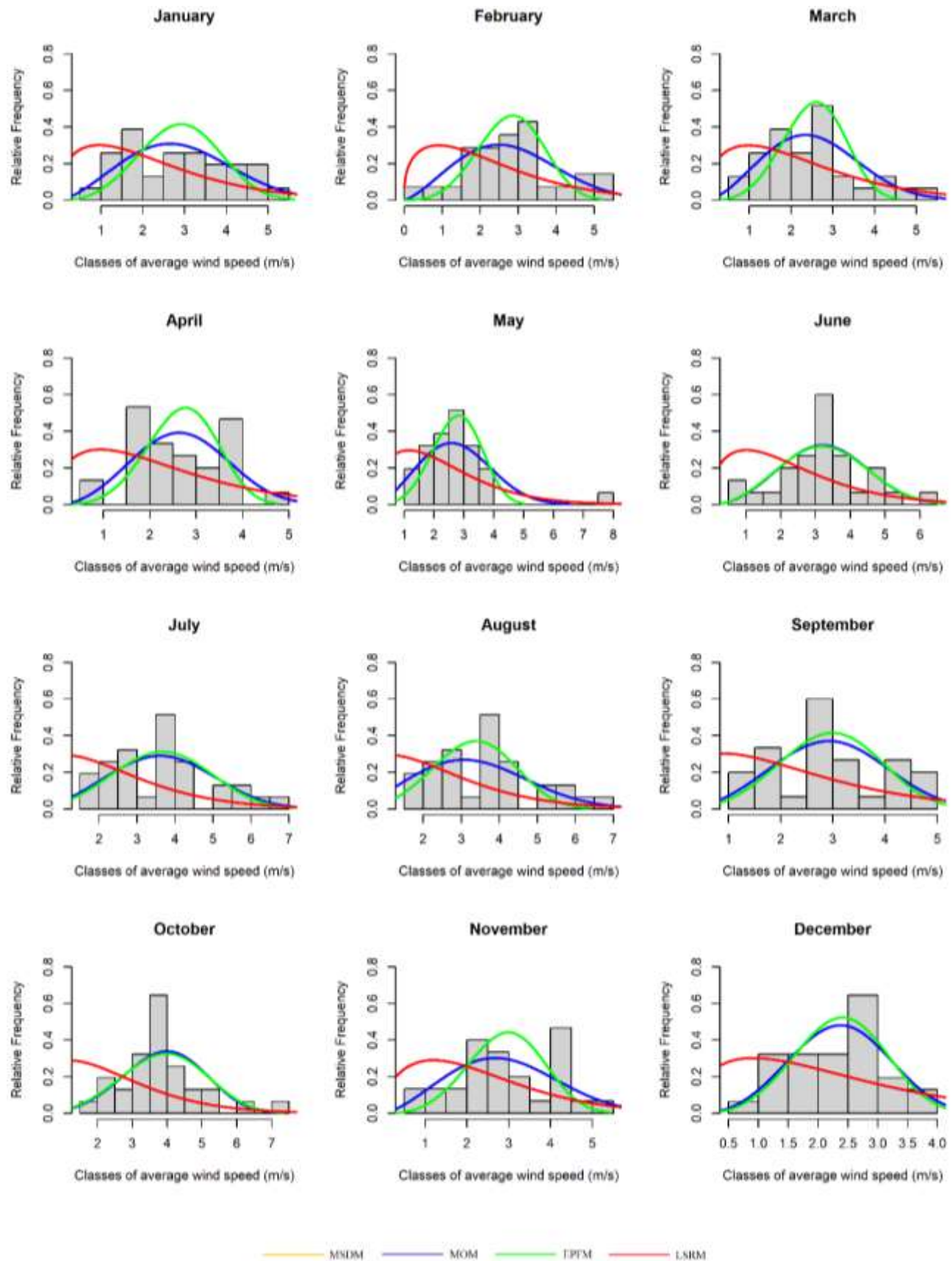


Figure 4. Histograms and comparison of Weibull probability functions for the observed monthly average wind speed.

Mdee (2020) and Kaplan (2017) argue that using more than one empirical method in an analytical series of MOM, EPD, and GM would improve the prediction of the Weibull distribution. In this case, the parameters k and c are estimated using various analytical methods, but provide different results for the same wind speed

distribution data. Kang, Khanjari, You, and Lee (2021) compared the 12 different numerical methods (MLM, PDM, MM, EMJ, MMLM, EML, STDM, MEPFM, GM, EPFM, EEM, and AMLM) to estimate the Weibull parameters for wind speed. They analyzed similarly collected wind data, at the same location and height, over 10 years, from 2010 to 2019, on Maldo Island, 1 km from a giant wind turbine. The EEM and EPFM form factors were the most accurate, while the GM and AMLM factors were the least accurate. The other methods were of intermediate precision. Scale factors were similar across all methods, with the exception of the GM, which was imprecise. Used the $RMSE$, R^2 and χ^2 methods to assess the accuracy of the method. GM, EPFM, EEM, and AMLM errors were the largest. The $RMSE$ standard deviations of GM, AMLM, EEM, and EPFM were higher than those of the other methods. The R^2 standard deviations of GM, AMLM, and EPFM were large. The GM, MMLM, and EEM standard deviations of χ^2 were also large. For all three methods, the GM had the highest standard deviations. The AMLM and EEM standard deviations were large for two statistical evaluation methods, while those for the MMLM and EPFM were large for one evaluation method. The MM, EMJ, and STDM standard deviations were small for all three statistical assessments. Based on the above, MM, EMJ, and STDM are higher. MLM, EML, MMLM, MEPF, and PDM can be accurate; however, if GM, EPFM, EEM or AMLM is used, statistical assessment of reliability is essential.

Guarienti et al. (2020) studied data from eight locations, collected by 27 automated systems for more than 6 years at heights varying from 79 to 727 m in Mato Grosso do Sul, Brazil, and estimated the monthly Weibull parameters for each location using the GM, MLM, MMLM, MM, EMJ, and PDM. Subsequent statistical analysis identified the optimal method for each location. Although accuracy varies from location and the monthly estimation method, the MLM and MMLM were generally highly accurate, whereas the GM was inaccurate.

The Atlantic Polar Mass originated in the sub-Antarctic waters of the south of South America, due to its continental displacement it penetrates the state by west and southwest predominantly in autumn and winter. It is dry and does not acquire moisture along the path. The Atlantic Polar Mass with maritime displacement also originates in the south of the South American continent, predominating during the winter and spring. It is dry in the origin but it absorbs moisture from the ocean, mainly from the hot current of Brazil (Wons, 1994).

The comparison of the monthly histograms and the Weibull PDF of the different methods tested are illustrated in Figure 3. After visual inspection, it becomes clear that the LSRM is not satisfying, compared to the other methods.

It has been verified that the shape parameter k oscillated significantly along the months, between 3.83 and 2.34 $m\ s^{-1}$ with the lowest value noted in February and the highest registered in October. Justus et al. (1978) observed that the k value is inversely related to the variance of wind speed around the average speed, which implies low variances if parameter k is high and vice versa. In this sense, the k values obtained for Campo Grande fully agreed with the previous statement, where the highest values of k were related to the lowest variances shown in Figure 1. The scale parameter c also varied, in the 4.33 to 2.65 range with minimum values, suggesting for this period a higher probability of occurrence of lower speed winds.

In this study, four methods have been used to estimate the Weibull shape and scale parameters. These average values have been calculated and are presented in Figures 3, 4 e 5. It has been determined that there is a conspicuous linear relationship between the monthly averages of the Weibull scale parameter (means of the four methods) and the measured monthly average wind speed.

$$c = 0.9031v - 0.0324 (R^2 = 0.9964) \quad (25)$$

The high value of R^2 coefficient demonstrate a counterpart to the linear model. Figure 5 shows the linear relationship connecting the Weibull scale parameter c and the monthly observed average values of wind speed in Campo Grande. The correlation involving the monthly scale parameter c and the measured monthly average wind speed has been determined with an inclination directly proportional to the average of the monthly scale parameters.

The degree of turbulence in relation to the coefficient of variation (COV) can be used to estimate the wind speed. It is characterized as the standard deviation divided by the wind speed. It is a highly helpful indicator for the operation and construction of the wind turbine since it is a turbulence indicator rather than an absolute value. Table 2 shows the average monthly COV values. As can be observed, the COV varies from 44.2 to 29.51%, with the months of January, February, March, and August showing the larger percentage of variance. The COV is typically lower when the wind speed is at higher or vice versa. However, further information on periodicity and, generally, wind speed time variability for a specific time scale is needed to fully analyze the potential for wind energy, the functioning of wind power conversion, or grid integration.

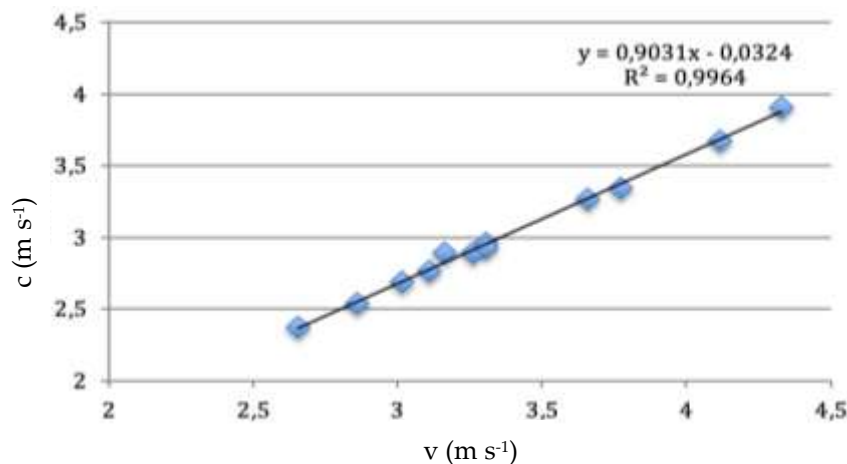


Figure 5. Relationship between the monthly Weibull scale parameters c and the wind speed means.

Table 2. Efficiency the tested methods.

Method	k	c (m s ⁻¹)	$RMSE$	χ^2	R^2
EPFM	3.64	3.35	0.34	3,71	0.99
LSRM	1.41	2.48	0.58	5,04	0.49
MOM	2.86	3.38	0.37	4,49	1.00
MSDM	2.86	3.38	0.37	4,49	1.00

Three different statistical tools, which are Root Mean Square Error (RMSE), Chi-square error (χ^2), and the coefficient of determination (R^2), are used in statistical analyzes to assess the efficiency of different methods. Since the approaches produce virtually identical results, classification of statistical methods typically only needs one column. These three statistical techniques were used to categorize the methods for a more accurate diagnosis.

Table 2 provides a summary of the test results for the four statistical approaches and categorizes them based on how well they work and how effectively they evaluate the wind data. EPFM, MOM, and MSDM produce better results than LSRM and are relatively similar to each other. The most important statistical tool, the chi-square error, has a value of $\chi^2 = 4.49$, and the effectiveness of the method has a value of $R^2 = 0.99$. The rankings were performed according to the respective minimum error and highest efficiency first places.

The results obtained in this study demonstrate the need to carry on further regionalized studies to test a larger number of probabilistic models for the adjustment of climatic variable, since peculiarities of the physical space and the interference of climatic phenomena in the region on a monthly, daily, and even hourly scale can change significantly the behavior of the wind climatic variable. As shown in Figure 1 and discussed earlier, the Weibull distribution adequately adjust the average wind speed data. This particularly important result could only be verified by investigating a probabilistic model less used in other studies and regions.

Discussion

Many studies have evaluated the accuracy of various estimation methods for determining the parameters of the Weibull distribution. Guarienti et al. (2020) analyzed the Weibull parameters over the months on municipalities of Mato Grosso do Sul. The higher values of the parameters were observed for Campo Grande stations, Chapadao do Sul, Costa Rica, and Sonora, reaching maximum values in August. The stations of Nhumirim, Aquidauana, Miranda, Sidrolandia, Porto Murtinho, Bela Vista, Jardim, Rio Brilhante, and Corumbá showed smaller variations in parameter values throughout the year.

The Maximum Likelihood Method (MLM) and the Modified Maximum Likelihood Method (MMLM) are more accurate to estimate the Weibull distribution parameters for the stations of Campo Grande, Chapadao do Sul, Costa Rica, Sonora, Nhumirim, Aquidauana, Sidrolandia, Porto Murtinho, Bela Vista, Jardim, Rio Brilhante, and Corumbá. The best-fit monthly estimator was generated by either one or the other of these two methods depending on the analyzed period.

The MLM is the best adjustment estimator for the stations of Tres Lagoas, Agua Clara, Paranaíba, Cassilandia, and Sao Gabriel do Oeste. The MMLM is a better estimator of adjustment of the Weibull distribution parameters for the stations of Ponta Pora, Amambai, Juti, Sete Quedas, Itaquiraí, Ivinhema,

Dourados, and Maracaju. The Weibull distribution is not recommended for the stations of Coxim and Miranda, which are characterized by low monthly average wind speeds.

Kang, Ko, and Huh (2018) analyzed wind resources at nine sites on Jeju Island, South Korea (Chujado, Gapado, Udo, Gujwa, Hallim, Moseulpo, Aewol, Ohdeumg and Sunheul) using six Weibull methods, namely Justus empirical method (EMJ), MOM, GM, power density method (PDM), MLM, and MMLM. They concluded that the MOM was the most accurate, while the GM was the least accurate.

Saxena and Rao (2015) analyzed the data for Rajasthan, India, and used four methods the GM, EMJ, MMLM, and PDM. MMLM and GM were the best and less accurate performing methods, respectively. Accuracy varied by region and measurement period, but overall, GM performed poorly compared to the MOM, MLM, and MMLM.

Kumar and Gaddada (2015) analyzed data of four sites in northern Ethiopia. The accuracy was compared amongst the PDM, GM, MOM, and the standard deviation method (SDM). Rocha, Sousa, Andrade, and Silva (2012) analyzed the winds of Camocim and Paracuru (Ceara, Brazil) using the GM, MLM, PDM, MOM, EMJ, MMLM, and the equivalent energy method (EEM). Hove, Madiye, and Musademba (2014) analyzed data of Harare, Gweru, Bulawayo, and Masvingo (Zimbabwe) using the GM, STD, MOM, MLM, and energy pattern factor (EPFM) method and compared the data with actual wind energy densities. All studies established that GM is the best performing method. Hove et al. (2014) especially proved GM data to be 4% accurate, in contrast to results from previous studies (Saxena & Rao, 2015; Kang et al., 2018; Ouahabi, Elkhachine, Benabdelouahab, & Khamlichi, 2020; Guarienti et al., 2020).

Bertrand, Abraham, and Lucien (2020) analyzed data collected from 2012 to 2014 of Amban, Cameroon. EMJ was the best performing method among GM, EMJ, PDM, and MOM. Werapun, Tirawanichakul, and Waewsak (2015) used the EMJ, PDM, MLM, MMLM, and GM to assess wind data collected from 2012 to 2014 of Phangan Island, Thailand. MLM best represented the actual wind speed distribution, while PDM was the best method to represent wind energy density. Gugliani, Sarkar, Ley, and Mandal (2018) compared the modified EPFM (MEPFM) to six other methods; MEPFM was more accurate than PDM and MLM (Akgül, Şenoğlu, & Arslan, 2016). Chaurasiya, Ahmed, and Warudkar (2018a) used nine methods to estimate wind resources in Tamil, India. MLM and MMLM performed moderately, while GM and the alternative maximum likelihood method (AMLM) performed poorly. Although PDM generally performed the less accurate, it was the third best method in terms of wind energy density calculation. The performance of the method differed by height. Many similar studies have been carried out for several regions, at various times (Kidmo, Danwe, Doka, & Djongyang, 2015; Bilir, İmir, Devrim, & Albostan, 2015; Ihaddadene, Ihaddadene, & Mostefaoui, 2016; Chaurasiya, Ahmed, & Warudkar, 2018b). The results varied in terms of the best performing methods and the accuracy of wind speed distributions differed from those of wind energy density.

Mohammadi, Alavi, Mostafaeipour, Goudarzi, and Jalilvand (2016) used GM, EMJ, EML, PDM, MLM, and MMLM to calculate wind energy densities given on a daily or monthly scale of four stations in Alberta, Canada. The EMJ, EML, PDM, and MLM were developed together; EMJ and EML presented similar performance, with EMJ being the best performing method. These comparisons of the Weibull distribution methods were obtained using different statistical analyses.

The Weibull parameter estimation methods and statistical analysis methods have been used by Justus et al. (1978), Akgül et al. (2016), Carneiro, Melo, Carvalho, and Braga (2016), Mohammadi et al. (2016), Kang et al. (2018), Gugliani et al. (2018), Ali, Lee, and Jang (2018), Ouahabi et al. (2020), Guarienti et al. (2020), Bertrand et al. (2020), Kapen, Gouajio, and Yemélé (2020) and Shoaib et al. (2020). The GM, MLM, PDM, MOM, EMJ, and MMLM methods are the most used, while EML, AMLM, STD, EEM, EPFM, and MEPFM are less frequently used.

The best methods of statistical analysis used in studies were root mean square error (RMSE), regression coefficient (R^2), chi-square (χ^2), mean absolute percentage error (MAPE), Kolmogorov-Smirnov (KS) test, relative percentage error (RPE), mean percentage error (MPE), correlation coefficient (r), relative root mean square error (RRMSE), index of concordance (IA), mean absolute view error (MABE), relative view (RB), and best bin size methods (B). From all these models, the RMSE, χ^2 , and R^2 were the most widely used, e.g., by Justus et al. (1978), Bilir et al. (2015), Kidmo et al. (2015), Mohammadi et al. (2016), Saxena and Rao (2015), Akgül et al. (2016), Carneiro et al. (2016), Ihaddadene et al. (2016), Gugliani et al. (2018), Chaurasiya et al. (2018a), Chaurasiya et al. (2018b), Kang et al. (2018), Ouahabi et al. (2020), Guarienti et al. 2020, Bertrand et al. (2020), Kapen et al. (2020), Shoaib et al. (2020) and Kang et al. (2021).

Kang et al. (2021) showed different results in studies of the determination of estimation methods for Weibull parameters. They analyzed 10 years of data collected at the same location and height level at the island of Maldo, from 2010 to 2019, and at the seawall of Saemangeum, from 2011 to 2012, in the Republic of

South Korea. They compared the Weibull parameters using twelve methods and identified the most reliable and efficient methods to derive the Weibull probability distribution using the new approach comparing the variance of RMSE, χ^2 , and R^2 , which provide a comprehensive view of level errors and fluctuations. These twelve methods are AMLM, Equivalent Energy Method (EEM), EMJ, EML, EPFM, GM, MEPFM, MLM, MOM, MMLM, PDM, and STDM. The results showed that while the EMJ, EML, MOM, and STDM provided better accuracies in predicting the wind speed distribution, some methods such as the GM, AMLM, EEM, and PFM had the poorest prediction of the wind speed distribution based on all variances of the statistical methods for both regions.

Using MOM, EML, EMJ, and STDM methods with relatively small error levels and fluctuations can yield a highly acceptable prediction of the wind speed distribution. As the error level and error fluctuations of MLM, MMLM, MEPF, and PDM methods vary based on different values of RMSE, χ^2 and R^2 , they cannot be highly recommended methods. GM, EPFM, EEM, and AMLM are not the preferred methods for estimating the wind speed distribution as they have high levels of error and fluctuations for RMSE, χ^2 , and R^2 compared to the other methods, so it is necessary to proceed with statistical analysis. All methods have been used to predict the wind speed distribution of another region (Saemangeum Seawall), which showed the same occurrence in Maldo. As a result, for the relationship between wind speed and Weibull probability distribution, due to the importance of wind energy density, the reliability of the methods used to estimate the Weibull parameters requires further evaluation.

Nomenclature

Symbols:

v_3^- : mean of wind speed cubes, $3 \text{ m}^3 \text{ s}^{-1}$

\bar{v} : mean wind speed, m s^{-1}

c : scale parameter of Weibull distribution, m s^{-1}

COV: coefficient of variation

EPF: energy pattern factor, dimensionless

E_w : wind energy per unit area by Weibull function, kWh m^{-2}

$f(v)$: Weibull PDF

$F(v)$: Weibull CDF

k : shape parameter of Weibull distribution, dimensionless

N : Total number of wind speed observations

R^2 : coefficient of determination

T : time, hour

v : wind speed, m s^{-1}

$x_{i,w}$: the frequency of Weibull or i th calculated value from the Weibull distribution

$y_{i,m}$: the frequency of observation or i th calculated value from measured data

$z_{i,v}$: the mean of i th calculated value from measured data

χ^2 : Chi-square error

Greek letters

σ : standard deviation of wind speed, m s^{-1}

$\Gamma()$: Gamma function

ρ : air density, kg m^{-3}

List of Abbreviations

AMLM: Alternative maximum likelihood method

EEM: Equivalent energy method

EMJ: Empirical method of Justus

EML: Empirical method of Lysen

EPFM: Energy pattern factor method

GM: Graphical method

LMOM: L-moment estimation method

MEPFM: Modified energy pattern factor method

MLM: Maximum likelihood method

MOM: Moment method

MMab: Mabchour method

MMLM: Modified maximum likelihood method
PDM: Power density method
PE: Percentile estimator
RMSE: Root mean square error
STDM: Standard deviation method
WM: WASP method
WPD: Wind power density

Conclusion

Four statistical methods to derive Weibull parameters for wind energy application are compared in Campo Grande - MS, Brazil. The statistical performance of the best Weibull distribution methods for wind data analysis are discussed and presented. From the analysis of the test results, it has been shown that the LSRM presents lesser accurate performance than other methods. The EPFM, MOM, and MSDM methods are the most efficient methods to fit the Weibull distribution curves for the evaluation of wind speed data in the study region. This study provides a new way to assess viable wind power locations that is applicable in any location with high winds in any country in the world.

Acknowledgments

This research was supported by the Universities and the UFMS Air Quality Laboratory, data was available at <https://crudata.uea.ac.uk/cru>.

References

- Akgül, F. G., Şenoğlu, B., & Arslan, T. (2016). An alternative distribution to Weibull for modeling the wind speed data: inverse Weibull distribution. *Energy Conversion and Management*, 114, 234-240. DOI: <https://doi.org/10.1016/j.enconman.2016.02.026>
- Ali, S., Lee, S.-M., & Jang, C.-M. (2018). Forecasting the long-term wind data via measure-correlate-predict (MCP) methods. *Energies*, 11(6), 1541. DOI: <https://doi.org/10.3390/en11061541>
- Arenas-López, J. P., & Badaoui, M. (2020). A Fokker–Planck equation based approach for modelling wind speed and its power output. *Energy Conversion and Management*, 222, 113152. DOI: <https://doi.org/10.1016/j.enconman.2020.113152>
- Baz, K., Cheng, J., Xu, D., Abbas, K., Ali, I., Ali, H., & Fang, C. (2021). Asymmetric impact of fossil fuel and renewable energy consumption on economic growth: a nonlinear technique. *Energy*, 226, 120357. DOI: <https://doi.org/10.1016/j.energy.2021.120357>
- Benard, A., & Bos-Levenbach, E. C. (1953). Het uitzetten van waarnemingen op waarschijnlijkheidspapier1. *Statistica Neerlandica*, 7(3), 163-173.
- Bertrand, K. S. E., Abraham, K., & Lucien, M. (2020). Sustainable energy through wind speed and power density analysis in Ambam, south region of Cameroon. *Frontiers in Energy Research*, 8, 176. DOI: <https://doi.org/10.3389/fenrg.2020.00176>
- Bilir, L., İmir, M., Devrim, Y., & Albostan, A. (2015). Seasonal and yearly wind speed distribution and wind power density analysis based on Weibull distribution function. *International Journal of Hydrogen Energy*, 40(44), 15301-15310. DOI: <https://doi.org/10.1016/j.ijhydene.2015.04.140>
- Carneiro, T. C., Melo, S. P., Carvalho, P. C. M., & Braga, A. P. S. (2016). Particle Swarm Optimization method for estimation of Weibull parameters: a case study for the Brazilian northeast region. *Renewable Energy*, 86, 751-759. DOI: <https://doi.org/10.1016/j.renene.2015.08.060>
- Celik, A. N. (2003a). Energy output estimation for small-scale wind power generators using Weibull-representative wind data. *Journal of Wind Engineering and Industrial Aerodynamics*, 91(5), 693-707. DOI: [https://doi.org/10.1016/S0167-6105\(02\)00471-3](https://doi.org/10.1016/S0167-6105(02)00471-3)
- Celik, A. N. (2003b). Weibull representative compressed wind speed data for energy and performance calculations of wind energy systems. *Energy Conversion and Management*, 44(19), 3057-3072. DOI: [https://doi.org/10.1016/S0196-8904\(03\)00075-X](https://doi.org/10.1016/S0196-8904(03)00075-X)

- Celik, A. N. (2004). A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey. *Renewable Energy*, 29(4), 593-604.
DOI: <https://doi.org/10.1016/j.renene.2003.07.002>
- Chaurasiya, P. K., Ahmed, S., & Warudkar, V. (2018a). Comparative analysis of Weibull parameters for wind data measured from met-mast and remote sensing techniques. *Renewable Energy*, 115, 1153-1165.
DOI: <https://doi.org/10.1016/j.renene.2017.08.014>
- Chaurasiya, P. K., Ahmed, S., & Warudkar, V. (2018b). 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), 2299-2311. DOI: <https://doi.org/10.1016/j.aej.2017.08.008>
- Dantas, G. C. B., Rodrigues, M. V. S., Silva, L. M. X., Aquino, M. D., & Thomaz, A. C. F. (2021). Panorama do setor eólico no Estado do Rio Grande do Norte no período 2004-2017. *Estudos Avançados*, 35(102), 79-94.
DOI: <https://doi.org/10.1590/s0103-4014.2021.35102.005>
- Freire, A. Í., & Fontgalland, I. L. (2022). Perspectives and economic challenges of wind energy generation in the Northeastern region of Brazil. *Research, Society and Development*, 11(1), e58911125429.
DOI: <https://doi.org/http://dx.doi.org/10.33448/rsd-v11i1.25429>
- Guarienti, J. A., Almeida, A. K., Menegati Neto, A., Ferreira, A. R. O., Ottonelli, J. P., & Almeida, I. K. (2020). 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*, 42, 100854.
DOI: <https://doi.org/10.1016/j.seta.2020.100854>
- Gugliani, G. K., Sarkar, A., Ley, C., & Mandal, S. (2018). New methods to assess wind resources in terms of wind speed, load, power and direction. *Renewable Energy*, 129(Part A), 168-182.
DOI: <https://doi.org/10.1016/j.renene.2018.05.088>
- Hove, T., Madiye, L., & Musademba, D. (2014). Mapping wind power density for Zimbabwe: a suitable Weibull-parameter calculation method. *Journal of Energy in Southern Africa*, 25(4), 37-47.
- Ihaddadene, R., Ihaddadene, N., & Mostefaoui, M. (2016). Estimation of monthly wind speed distribution basing on hybrid Weibull distribution. *World Journal of Engineering*, 13(6), 509-515.
DOI: <https://doi.org/10.1108/wje-09-2016-0084>
- Jansen, M., Staffell, I., Kitzing, L., Quoilin, S., Wiggelinkhuizen, E., Bulder, B., ... Müsgens, F. (2020). Offshore wind competitiveness in mature markets without subsidy. *Nature Energy*, 5, 614-622.
DOI: <https://doi.org/10.1038/s41560-020-0661-2>
- Justus, C. G., Hargraves, W. R., Mikhail, A., & Graber, D. (1978). Methods for estimating wind speed frequency distributions. *Journal of Applied Meteorology and Climatology*, 17(3), 350-353.
DOI: [https://doi.org/10.1175/1520-0450\(1978\)017<0350:MFEWSF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1978)017<0350:MFEWSF>2.0.CO;2)
- Kandpal, D., & Dhingra, T. (2021). Migrating to reverse auction mechanisms in wind energy sector: Status and challenges. *Energy Policy*, 156, 112352. DOI: <https://doi.org/10.1016/j.enpol.2021.112352>
- Kang, D., Ko, K., & Huh, J. (2018). Comparative study of different methods for estimating Weibull parameters: a case study on Jeju Island, South Korea. *Energies*, 11(2), 356. DOI: <https://doi.org/10.3390/en11020356>
- Kang, S., Khanjari, A., You, S., & Lee, J.-H. (2021). Comparison of different statistical methods used to estimate Weibull parameters for wind speed contribution in nearby an offshore site, Republic of Korea. *Energy Reports*, 7, 7358-7373. DOI: <https://doi.org/10.1016/j.egy.2021.10.078>
- Kapen, P. T., Gouajio, M. J., & Yemélé, B. (2020). Analysis and efficient comparison of ten numerical methods in estimating Weibull parameters for wind energy potential: application to the city of Bafoussam, Cameroon. *Renewable Energy*, 159, 1188-1198. DOI: <https://doi.org/10.1016/j.renene.2020.05.185>
- Kaplan, Y. A. (2017). Determination of the best Weibull methods for wind power assessment in the southern region of Turkey. *IET Renewable Power Generation*, 11(1), 175-182. DOI: <https://doi.org/10.1049/iet-rpg.2016.0206>
- Kidmo, D. K., Danwe, R., Doka, S. Y., & Djongyang, N. (2015). Statistical analysis of wind speed distribution based on six Weibull methods for wind power evaluation in Garoua, Cameroon. *Revue des Energies Renouvelables*, 18(1), 105-125.
- Kumar, K. S. P., & Gaddada, S. (2015). Statistical scrutiny of Weibull parameters for wind energy potential appraisal in the area of Northern Ethiopia. *Renewables: Wind, Water, and Solar*, 2, 14.
DOI: <https://doi.org/10.1186/s40807-015-0014-0>

- Liu, G., Zhou, J., Jia, B., He, F., Yang, Y., & Sun, N. (2019). Advance short-term wind energy quality assessment based on instantaneous standard deviation and variogram of wind speed by a hybrid method. *Applied Energy*, 238, 643–667. DOI: <https://doi.org/10.1016/j.apenergy.2019.01.105>
- Liu, S., Zheng, Y., Ma, R., Yu, K., Han, Z., Xiao, S., ... Zou, J. (2020). Increased soil release of greenhouse gases shrinks terrestrial carbon uptake enhancement under warming. *Global Change Biology*, 26(8), 4601–4613. DOI: <https://doi.org/10.1111/gcb.15156>
- Manwell, J. F., McGowan, J. G., & Rogers, A. L. (2010). *Wind energy explained: theory, design and application* (2nd ed.). Wiltshire, GB: Wiley.
- Mdee, O. J. (2020). Performance evaluation of Weibull analytical methods using several empirical methods for predicting wind speed distribution. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 1(1), 24. DOI: <https://doi.org/10.1080/15567036.2020.1832161>
- Mohammadi, K., Alavi, O., Mostafaeipour, A., Goudarzi, N., & Jalilvand, M. (2016). Assessing different parameters estimation methods of Weibull distribution to compute wind power density. *Energy Conversion and Management*, 108, 322–335. DOI: <https://doi.org/10.1016/j.enconman.2015.11.015>
- Ouahabi, M. H., Elkhachine, H., Benabdelouahab, F., & Khamlichi, A. (2020). Comparative study of five different methods of adjustment by the Weibull model to determine the most accurate method of analyzing annual variations of wind energy in Tetouan - Morocco. *Procedia Manufacturing*, 46, 698–707. DOI: <https://doi.org/10.1016/j.promfg.2020.03.099>
- Rocha, P. A. C., Sousa, R. C., Andrade, C. F., & Silva, M. E. V. (2012). Comparison of seven numerical methods for determining Weibull parameters for wind energy generation in the northeast region of Brazil. *Applied Energy*, 89(1), 395–400. DOI: <https://doi.org/10.1016/j.apenergy.2011.08.003>
- Saxena, B. K., & Rao, K. V. S. (2015). Comparison of Weibull parameters computation methods and analytical estimation of wind turbine capacity factor using polynomial power curve model: case study of a wind farm. *Renewables: Wind, Water, and Solar*, 2(3), 1–11. DOI: <https://doi.org/10.1186/s40807-014-0003-8>
- Shaban, A. H., Resen, A. K., & Bassil, N. (2020). Weibull parameters evaluation by different methods for windmills farms. *Energy Reports*, 6(suppl. 3), 188–199. DOI: <https://doi.org/10.1016/j.egy.2019.10.037>
- Shi, H., Dong, Z., Xiao, N., & Huang, Q. (2021). Wind speed distributions used in wind energy assessment: a review. *Frontiers in Energy Research*, 9, 769920. DOI: <https://doi.org/10.3389/fenrg.2021.769920>
- Shoaib, M., Rehman, S., Siddiqui, I., Rehman, S., Khan, S., & Ibrahim, Z. (2020). Comparison of Weibull and Gaussian mixture models for wind speed data analysis. *International Journal of Economic and Environmental Geology*, 11(1), 10–16.
- Souza, A., Abreu, M. C., Oliveira-Júnior, J. F., Fernandes, W. A., Aviv-Sharon, E., & Graf, R. (2022). Climate regionalization in Mato Grosso do Sul: a combination of hierarchical and non-hierarchical clustering analyses based on precipitation and temperature. *Brazilian Archives of Biology and Technology*, 65, e22210331. DOI: <https://doi.org/10.1590/1678-4324-2022210331>
- Werapun, W., Tirawanichakul, Y., & Waewsak, J. (2015). Comparative study of five methods to estimate Weibull parameters for wind speed on Phangan Island, Thailand. *Energy Procedia*, 79, 976–981. DOI: <https://doi.org/10.1016/j.egypro.2015.11.596>
- Wons, I. (1994). *Geografia do Paraná – física, humana e econômica* (6. ed.). Curitiba, PR: Ensino Renovado.