ENCIT 2018

17th Brazilian Congress of Thermal Sciences and Engineering
November 25th-28th, 2018, Águas de Lindóia, SP, Brazil

ENCIT-2018-0735
DETERMINATION OF WEIBULL CURVE PARAMETERS APPLIED TO WIND ENERGY BY PARTICLE SWARM OPTIMIZATION

Marcus V. S. Macedo
Lindemberg F. dos Santos
Federal University of Ceara, 2853, Universidade Avenue, Benfica, Fortaleza
marcusmacedo87@gmail.com
lindemberg.li@hotmail.com

Felipe F. Gomes
Federal University of Ceara, 2853, Universidade Avenue, Benfica, Fortaleza
felipefagomes@outlook.com

Carla F. de Andrade
Federal University of Ceara, 2853, Universidade Avenue, Benfica, Fortaleza
carla@ufc.br

Paulo A. C. Rocha
Federal University of Ceara, 2853, Universidade Avenue, Benfica, Fortaleza
paulo.rocha@ufc.br

Abstract. *Wind power is one of the most widely used renewable energy sources in Brazil. To evaluate the wind potential of a region, a wind behavior model is required. Several probability distributions are used to represent the wind speed data, such as Weibull, Gamma and Normal, and Weibull is the most used for this purpose. This distribution is dependent of two parameters, shape factor k and scale factor c, usually estimated by deterministic methods. Studies have shown that some deterministic methods lead to an unsatisfactory fit. This paper aims to estimate the Weibul distribution curve parameters for the region of Triunfo-PE, by the heuristic method known as Particle Swarm Optimization Method. The method was implemented by R language and the results were compared using the statistical tests RMSE, MAE and RÂš, as well as the percentage value of the production deviation between the curve obtained and the histogram generated from the wind speed data - WPD. Particle Swarm Optimization method did not prove to be efficient to estimate the Weibull curve parameters for the analyzed region, since it obtained high values of RMSE, MAE and low value of R2 when compared to the deterministic methods. The WPD equal to 10.88% was higher than the acceptable maximum for the deviation of production.*

Keywords: *Wind energy, Weibull Distribution, Particle Swarm Optimization, Heuristic, Deterministic*

1. INTRODUCTION

As with hydropower, wind energy has been used for thousands of years for the same purposes, namely water pumping, grain grinding and other applications involving mechanical energy. For the generation of electricity, the first attempts appeared in the late nineteenth century, but only a century later, with the international oil crisis (1970), was there sufficient interest and investment to enable the development and application of equipment in commercial scale.

In order to evaluate wind power potential in a region, a stochastic wind behavior modeling is necessary, allowing the identification of seasonal patterns and prediction of the wind resources behavior Maceira *et al.*. There are several probability distributions used to represent wind speed data, such as Weibull, Gamma and Normal; The well-known Weibull distribution is the standard function used by the wind power community to model the frequency distribution of wind speeds around the world. However, studies have shown that some methods to determine the Weibull parameters lead to an unsatisfactory adjustment capacity for the used wind distribution histograms Silva (2003). In this sense, alternative methods to calculate the Weibull parameters should be investigated.

This paper aims to estimate k e c parameters by the application of Particle Swarm Optimization, and to compare them with those already obtained by the following deterministic methods: Least Squares Method (LSM), Moment Method (MM), Maximum Likelihood Method (MLM), Energy pattern factor method (EPFM), Modified Maximum Likelihood

Method (MMLM), Equivalent Energy Method (EEM), Empirical Method (EM) and Chi-Square Method (χ^2).

2. NUMERICAL METHODS FOR DETERMINING THE WEIBULL PARAMETERS

As the wind speed is a random variable, it is useful to use statistical analysis to determine the wind potential of a region Wais (2017), Celik (2003) and Akpinar and Akpinar (2004). Commonly, the two parameters Weibull distribution is the one that presents the best fit and is therefore the most used to estimate this potential Burton *et al.* (2001) and Manwell *et al.* (2009).

The Weibull distribution for the velocity v is expressed by the probability density function, wind velocity frequency curve, shown in Equation 1. Equation 2 expresses its cumulative probability function Ohunakin *et al.* (2011) and Chang (2011).

$$f(v) = (\frac{k}{c}) \cdot (\frac{v}{c})^{(k-1)} \cdot e^{-(\frac{v}{c})^k} \tag{1}$$

$$F(v) = \int_0^v f(v)\mathrm{d}v = 1 - e^{-(\frac{v}{c})^k} \quad v,\, k\, and\, c > 0 \tag{2}$$

Where c is the scaling factor with unit $m \cdot s^{-1}$, k is the shape factor (dimensionless) and $F(v)$ denotes the probability of velocities smaller than or equal to v.

2.1 Maximum Likelihood Method (MLM)

In the Maximum Likelihood Method, numerical iterations are required to determine the Weibull distribution parameters Fisher (1915). In this method Rocha *et al.* (2012), the parameters k and c are determined according to the Equations 3 and 4.

$$k = \left[\frac{\sum_{i=1}^n v_i^k ln(v_i)}{\sum_{i=1}^n v_i^k} - \frac{\sum_{i=1}^n ln(v_i)}{n}\right]^{-1} \tag{3}$$

$$c = \left(\frac{1}{n}\sum_{i=1}^n v_i^k\right)^{\frac{1}{k}} \tag{4}$$

Where n is the number of observed data and v_i is the wind speed measured in the interval i.

2.2 Moment Method (MM)

The Moment Method maybe used as an alternative to the Maximum Likelihood Method and it is recommended when the mean and standard deviation of the elements are known and are initially on an appropriate scale Justus *et al.* (1978). In this case Rocha *et al.* (2012), the k and c parameters are determined by the Equations 5 and 6.

$$\sigma = c \cdot \sqrt{\Gamma(1+\frac{2}{k}) - \Gamma^2(1+\frac{2}{k})} \tag{5}$$

$$\overline{v} = c \cdot \Gamma(1+\frac{1}{k}) \tag{6}$$

Where $\bar{v}$, σ, Γ are, respectively, the average wind speed, the standard deviation of the observed wind speed data, and the gamma function.

2.3 Empirical Method (EM)

The empirical method Rocha *et al.* (2012) and Chang (2011) is considered a simplified form of the Moment Method, in which the determination of the k parameter follows Equation 7 and the c parameter Equation 8.

$$k = (\frac{\sigma}{\overline{v}})^{-1{,}086} \tag{7}$$

$$\overline{v} = c \cdot \Gamma(1+\frac{1}{k}) \tag{8}$$

Where $\bar{v}$ and σ are respectively the mean wind speed and the standard deviation of the observed wind speed data.

2.4 Equivalent Energy Method (EEM)

The Equivalent Energy Method seeks the equivalence between the energy density of the observations and the theoretical Weibull curve. For this, the k parameter is estimated from the third moment of the velocity, by minimizing the square error related to the adjustment, represented by Equation 9 and the c parameter is adjusted by using Equation 10 Silva (2003) and Andrade *et al.* (2014).

$$\epsilon^2 = \sum_{i=1}^{n} \left\{ W_i - e^{-\left[\frac{(v_i - 1)(\Gamma(1+\frac{3}{k}))^{1/3}}{(\overline{v^3})^{1/3}}\right]^k} + e^{-\left[\frac{(v_i)(\Gamma(1+\frac{3}{k}))^{1/3}}{(\overline{v^3})^{1/3}}\right]^k} \right\}^2 \quad (9)$$

$$c = \left[\frac{\overline{v^3}}{\Gamma(1+\frac{3}{k})} \right]^{1/3} \quad (10)$$

2.5 Energy Pattern Factor Method (EPFM)

The energy pattern factor, E_{pf}, method is related to the averaged data of wind speed and is defined by the following equations 11 until 13 Akdag and Dinler (2009):

$$E_{pf} = \frac{\overline{v^3}}{\overline{v}^3} \quad (11)$$

$$k = 1 + \frac{3.69}{(E_{pf})^2} \quad (12)$$

$$\overline{v} = c \cdot \Gamma\left(1 + \frac{1}{k}\right) \quad (13)$$

2.6 Modified Maximum Likelihood Method (MMLM)

The modified maximum likelihood method can only be considered if the available data of wind speed are already in the shape of the Weibull distribution and, as in the maximum likelihood method, it requires numerical iterations for the solution of the equations:

$$k = \left[\frac{\sum_{i=1}^{n} v_i^k ln(v_i) f(v_i)}{\sum_{i=1}^{n} v_i^k f(v_i)} - \frac{\sum_{i=1}^{n} ln(v_i) f(v_i)}{f(v \geq 0)} \right]^{-1} \quad (14)$$

$$c = \left(\frac{1}{f(v \geq 0)} \sum_{i=1}^{n} v_i^k f(v_i) \right)^{\frac{1}{k}} \quad (15)$$

where $f(v_i)$ represents the Weibull frequency and $f(v \geq 0)$ is the probability of wind speed ≥ 0.

2.7 Least Squares Method (LSM)

The purpose of the method is to define a line where the values of a sample are contained by minimizing the square root of the discrepancy between the value of the sample and the value predicted by the line (objective function) according to Equations 16 until 22 Justus *et al.* (1978):

$$y_i = ax_i + b \quad (16)$$

$$\epsilon_i = y_i - (ax_i + b) \quad (17)$$

$$\epsilon^2 = \sum_{i=1}^{n} [y_i - (ax_i + b)]^2 \quad (18)$$

$$a = \frac{\sum_{i=1}^{n} x_i (y_i - \overline{y})}{\sum_{i-1}^{n} x_i (x_i - \overline{x})} \quad (19)$$

$$b = \overline{y} - a\overline{x} \quad (20)$$

$$k = a \quad (21)$$

$$c = e^{-\frac{b}{k}} = e^{[\overline{x} - (\frac{\overline{y}}{k})]} \quad (22)$$

2.8 Chi-Square Method (χ^2)

Similarly to the adjustment by the equivalent energy method, the Chi-Square method seeks to minimize the error of the Chi-Square test between measured data and the expected data, according to Equations 23 and 24 Dorvlo (2002).

$$\chi^2 = \sum_{i=1}^{n} \left\{ \frac{[F(v_i) - (1 + exp(\frac{v_i}{k})^k)]^2}{1 + exp(\frac{v_i}{k})^k} \right\} \tag{23}$$

$$\overline{v} = c \cdot \Gamma(1 + \frac{1}{k}) \tag{24}$$

It is worth mentioning that for all methods that use frequency distribution values (histogram), the value v_i represents the central value of the speed *(bin)*.

3. HEURISTIC METHODS

Heuristics encompasses a set of methods where, to solve a problem, the variables in question use the experience gained over the iterations. Heuristic methods combine different concepts intelligently to explore the search space, so that learning strategies are used to structure information and find efficient and almost optimal solutions Osman and Laporte (1996). Many of the heuristic approaches depend on probabilistic decisions made during the algorithm run. The main difference against pure random search is that in heuristic algorithms randomness is not used blindly but intelligently and biased Stutzle (1999). It is valid to emphasize that every optimization procedure searches for the best result of a function for the desired scenario. This function is called the Objective Function. In this paper, the objective function is the one presented in Equation 25, which represents the minimization of the square error sum applied to the frequency of occurrence values found by the curve adjusted by the method and the observed frequency of occurrence in the histogram of the data.

$$\epsilon^2 = \sum_{i=1}^{n} (f_{adjustment} - f_{observed})^2 \tag{25}$$

Where n is the number of histogram velocity intervals and $f_{adjustment}$ and $f_{observed}$ are the occurrence frequencies by the adjusted curve and observed in the histogram, respectively.

3.1 Particle Swarm Optimization (PSO)

In a PSO system, each particle "flies" through the multidimensional search space, adjusting its position in space according to its own experience, but also considering the experience of the neighboring particle. A particle uses the best position found by itself and the best position of its neighbors to position itself towards an ideal solution. The effect is that the particles "fly" towards a global optimum, while still investigating an area around the best current solution **?**. For each particle k positioned in a two-dimensional plane and for each iteration i, the positions are recorded and the best individual result (x_k^{best}, y_k^{best}) is recorded Then, the best result among the k particles is recorded $(x_{global}^{best}, y_{global}^{best})$. Each particles movement will be proportional to the distance between the current position of the particle and the resulting point of the weighted average between the best individual position of the particle and the best position of the swarm, according to Equations 26 and 27 Eberhart and kennedy (1995).

$$x_k^{(i+1)} = x_k^{(i)} + V_{x,k}^{(i)} \tag{26}$$

$$y_k^{(i+1)} = y_k^{(i)} + V_{y,k}^{(i)} \tag{27}$$

Where,

$$V_{x,k}^{(i+1)} = \omega^{(i)} \cdot V_{x,k}^{(i)} + c_1 \cdot \lambda \cdot [x_k^{(best)} - x_k^{(i)}] + c_2 \cdot \mu \cdot [x_{(global)}^{(best)} - x_k^{(i)}] \tag{28}$$

$$V_{y,k}^{(i+1)} = \omega^{(i)} \cdot V_{y,k}^{(i)} + c_1 \cdot \eta \cdot [y_k^{(best)} - y_k^{(i)}] + c_2 \cdot \epsilon \cdot [x_{(global)}^{(best)} - y_k^{(i)}] \tag{29}$$

In these, λ, μ, η and ϵ are random numbers belonging to the set [0, 1] and ω is the particle inertia term, defined by Equation 30 Secchi and Biscaia (2012).

$$\omega^{(i)} = \omega_{initial} + (\omega_{final} - \omega_{initial}) \cdot \left(\frac{i}{m}\right) \tag{30}$$

Where m is the maximum number of iterations.

4. METHODOLOGY

The method application and data elaboration applied used the statistical tool called RStudio, integrated development environment of the R language. A wind data series was tested by combining a pair (*k*, *c*), composed by 52,560 speed values, number of values established according to the IEC 61400 PART 12-1(2005) standard that defines 1 (one) year of integrated data every 10 minutes, totaling 52,560 values.

The adjustment tests with real data were performed with public data from the Federal Government's SONDA project, referring to the TRI23 station, located in Triunfo, PE, at 50m at ground level and with one year of data for reasons of availability, the year 2006 was selected and the occurrence of inconsistent data was reviewed.

To analyze the efficiency of the aforementioned methods, the following tests are used: RMSE (Root Mean Square Error), Mean Absolute Error (MAE), R^2 (analysis of variance or efficiency of the method) and the percentage value of the production deviation between the obtained curve and the histogram was also evaluated. These tests are defined by Equation 31 until Equation 34 respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i^{calculated} - y_i^{measured})^2}{n}} \tag{31}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i^{calculated} - y_i^{measured}\right| \tag{32}$$

$$R^2 = \frac{\sum_{i=1}^{n}(y_i^{measured} - \bar{y}^{measured})^2 - \sum_{i=1}^{n}(y_i^{measured} - y_i^{calculated})^2}{\sum_{i=1}^{n}(y_i^{measured} - \bar{y}^{measured})^2} \tag{33}$$

$$WPD = \left(\frac{WPD_{estimated} - WPD_{measured}}{WPD_{measured}}\right) \cdot 100 \tag{34}$$

Where, according to (Jamil *et al.*, 1995), $\mathrm{WPD}_{measured}$ and $\mathrm{WPD}_{estimated}$ are calculated respectively by Equations 35 and 36

$$WPD_{medido} = \frac{1}{2} \cdot \rho \cdot c^3 \cdot \Gamma\left(1 + \frac{3}{k}\right) \tag{35}$$

$$WPD_{estimado} = \frac{1}{2} \cdot \rho \cdot v^3 \tag{36}$$

Where ρ is the specific mass of the air.

5. RESULTS AND DISCUSSION

Figures 1 present the Weilbull distribution curves, described by its probability function $f(v)$, versus wind speed. The PSO method was calculated based on the parameters, particles number equal to 27, cognitive term (c_1) equal to 1, social learning term (c_2), particle initial inertia term (w_i) equal to 1.98 and particle final inertia term (w_f) equal to 0.12. Figure 1 compares eight deterministic methods and the heuristic method, PSO.

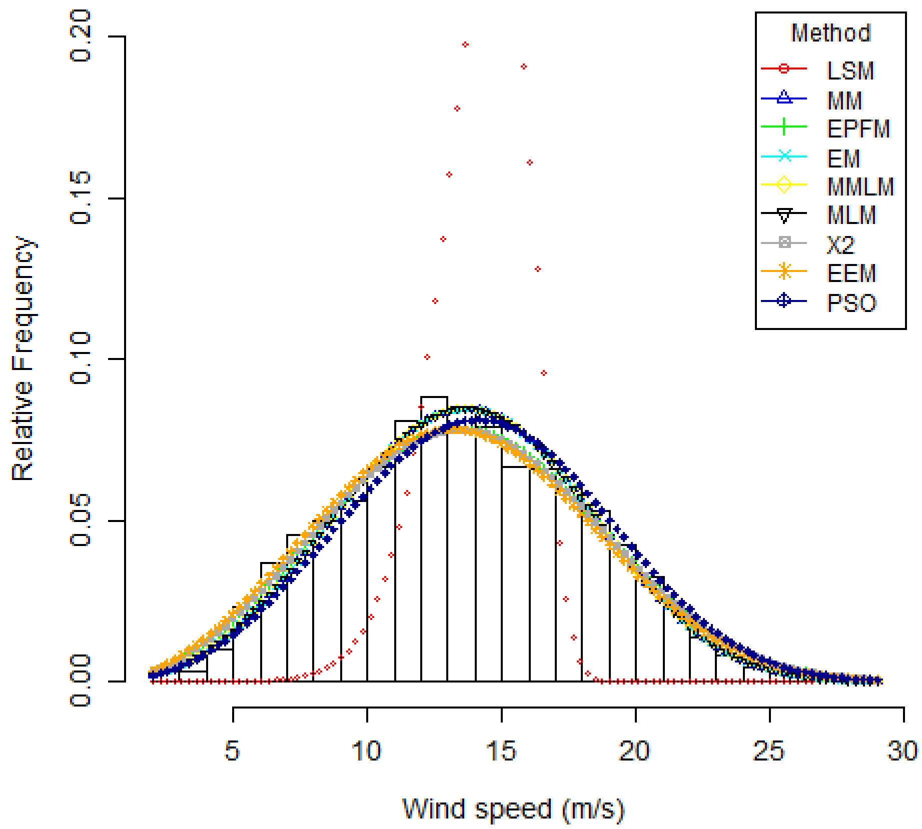

Figure 1. PSO and eight deterministic methods Comparison

The results of the statistical tests for the TRI23 station located in Triunfo are presented in Tab.1.

Table 1. Statistical Analysis of Triunfo, year 2006.

Method	k	c	RMSE	MAE	R^2	WPD (%)
LSM	9.9496	14.9003	0.010138	0.034751	-2.389724	-12.485331
MM	3.336	15.2103	0.000775	0.002973	0.980173	-0.241481
EPFM	3.0761	15.2702	0.000770	0.003117	0.980447	3.911729
EM	3.3290	15.2119	0.000768	0.002956	0.980510	-0.140927
MMLM	3.3337	15.2254	0.000775	0.002964	0.980183	0.080740
MLM	3.3283	15.2191	0.000769	0.002953	0.980493	0.007180
χ^2	3.0524	15.2756	0.000794	0.003234	0.979186	4.343732
EEM	3.0004	15.0249	0.000842	0.003614	0.976575	$-2.220446 \cdot 10^{-14}$
PSO	3.3012	15.7376	0.001034	0.004310	0.964724	10.875590

According to the Table 1, it can be observed that EM presented the lowest RMSE and R^2 tests values with 0.000768 and 0.980510, respectively. The MLM presented the best performance when it was analyzed the MAE test value, 0.002953. The WPD results showed a superiority of the EEM among all methods tested with value of $2.22 \cdot 10^{-14}$%.PSO did not perform well with statistical tests and the value obtained with WPD test presented value of 10.8755%, more than 2%, which one was below the acceptable limit for the deviation of Wind Power Density.

The EM obtained a better fit to the histogram when compared to the others deterministic methods and heuristic method. It is noticed that the curve suffers a slight shift to the right, in addition, the velocity peak becomes better represented. LSM obtained the worse fit to the histogram when compared to the others methods.

6. CONCLUSION

In this paper, eight deterministic and one heuristic optimization methods namely Particle Swarm Optimization were used to estimate the parameters, *k* and *c*, of the Weibull distribution for Triunfo, a city with good conditions, climate and geomorphology, for wind energy generation. The results were compared to each other. The deterministic methods were compared with the PSO method, using as a selection criteria the statistical tests. The following conclusions can be drawn based on the results presented in the previous sections:

1. For Triunfo, Equivalent Energy Method stood out, presenting the best performance among all methods tested for the cubic velocity energy production (WPD), obtained the best performance with value of $2.22{\cdot}10^{-14}\%$.

2. Particle Swarm Optimization was not an efficient method for determining the Weibull distribution, *k* and *c* parameters, for Triunfo, PE, Brazil.

3. Least Squares Method obtained results very distant compared to the other methods, becaming the method unfeasible for determining the Weibull distribution, *k* and *c* parameters, for Triunfo, PE, Brazil.

4. Energy Method was an efficient method, for determining the Weibull distribution, *k* and *c* parameters, for Triunfo, PE, Brazil.

7. ACKNOWLEDGEMENTS

This research was jointly supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazilian governmental agencies.

8. REFERENCES

Akdag, S.A. and Dinler, A., 2009. “A new method to estimate weibull parameters for wind energy applications”. *Energy Conversion and Management*, Vol. 50, pp. 1761–1766.

Akpinar, E.K. and Akpinar, S., 2004. “Determination of the wind energy potential for maden-elazig”. *Energy Convers Manage*, Vol. 45, pp. 2901–2914.

Andrade, C.F., Neto, M.F.H., Rocha, P.A.C. and Silva, M.E.V., 2014. “An efficiency comparison of numerical methods for determining weibull parameters for wind energy applications:a new approach applied to the northeast region of brazil”. *Energy Conversion and Management*, Vol. 86, pp. 801–808.

Burton, T., Sharpe, D., Jenkins, N. and Bossanyi, E., 2001. *Wind Energy Handbook*. Jonh Wiley Sons Ltd, England. ISBN 0 471 48997 2.

Celik, A.N., 2003. “A statistical analysis of wind power density based on the weibull and rayleigh models at the southern region of turkey”. *Renew Energy*, Vol. 29, pp. 593–604.

Chang, T.P., 2011. “Estimation of wind energy potential using different probability density functions”. *Applied Energy*, Vol. 88, pp. 1848–1856.

Dorvlo, A.S., 2002. “Estimating wind speed distribution”. *Energy Conversion and Management*, Vol. 43, pp. 2311–2318.

Eberhart, R.C. and kennedy, J., eds., 1995. *A new optimizer using particle swarm theory*, Vol. 1. Sixth International Symposium on, IEEE.

Fisher, R.A., 1915. “Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population”. *Biometrika*, Vol. 10, pp. 507–521.

Jamil, M., Parsa, S. and Majidi, M., 1995. “Wind power statistics and an evaluation of wind energy density”. *Renewable Energy*, Vol. 6, pp. 623–628.

Justus, C.G., Hargraves, W.R., Mikhail, A. and Graber, D., 1978. “Methods forestimating wind speed frequency distributions”. *Journal Of Applied Meteorology*, Vol. 17, pp. 350–353.

Maceira, M., Penna, D., Pessanha, J. and Melo, A., ???? “Modelagem estatística de ventos para utilização nos modelos de planejamento e operação”.

Manwell, J.F., McGowan, J.G. and Rogers, A.L., 2009. *Wind energy explained : theory, design, and application*. Jonh Wiley Sons Ltd, United Kingdom, 2nd edition. ISBN 978-0-470-01500-1.

Ohunakin, O.S., Adaramola, M.S. and Oyewola, O.M., 2011. “Wind energy evaluation for electricity generation using wecs in seven selected locations in nigeria”. *Applied Energy*, Vol. 88, pp. 3197–3206.

Osman, I.H. and Laporte, G., 1996. “Metaheuristics: A bibliography”. *Annals of Operational Research*, Vol. 63, pp. 513–628.

Rocha, P.A.C., Souza, R.C., Andrade, C.F. and 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*, Vol. 89, pp. 395–400.

Secchi, A.R. and Biscaia, E.C., 2012. “Otimização de processos”.

Silva, G.R., 2003. *Características de vento da Região Nordeste, análise, modelagem e aplicações para projetos de centrais eólicas*. Master's thesis, Federal University of Pernambuco, Recife.

Stutzle, T., 1999. *Local Search Algorithms for Combinatorial Problems-Analysis, Algorithms and New Applications*. Ph.D. thesis, DISKI - Dissertation en zur Kunstliken Intelligenz, Augustin, Germany.

Wais, P., 2017. "Two and three-parameter weibull distribution in available windpower analysis". *Renewable Energy*, Vol. 103, pp. 15–29.

9. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.