# ENCIT-2018-0739
# HEURISTIC METHODOLOGY APPLIED TO WIND ENERGY: DETERMINATION OF WEIBULL CURVE PARAMETERS BY ANT COLONY 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.*** *To mitigate the use of fossil fuels and continue to attend to the demand for electricity, many countries are investing in wind power. As wind speed is a random variable, the wind speed variation over a period of time can be represented by probability density functions, with the Weibull distribution being the most used. This paper aims to estimate the Weibull distribution parameters by the Ant Colony Optimization Method, a heuristic method widely used in wind energy studies. The implementation was performed using the R language, and the results were compared by statistical tests with those obtained by deterministic methods usually used for this purpose. The statistical tests results, RMSE equal to 0.001692, MAE equal to 0.003701, $R^2$ equal to 0.996033 and the Wind Power Deviation equal to 1.3248%, showed that the proposed method is effective to estimate the Weibull distribution curve parameters for the analyzed region.*

***Keywords:*** *Wind energy, Weibull Distribution, Ant Colony Optimization, Heuristic, Deterministic*

## 1. INTRODUCTION

An increasing number of countries use wind power for electricity generation to meet their growing energy demand and as an alternative to mitigate the use of fossil fuel. In Brazil, the operation of this energy resource is gradually rising since 2000 according Energy Research Office (EPE). Renewable energy sources have gained great importance due to their inexhaustibility, sustainability, ecological awareness and providing energy security. Thus, it is expected that renewable energy sources will play an important role, especially in the generation of electricity.

For accurate assessment of wind power potential and characteristics, long-term meteorological observations are required. Wind speed is a random variable and the wind speed variation over a period of time is represented by probability density functions. Detailed knowledge of wind characteristics and distribution are crucial parameters for selecting the best wind power conversion system to optimize energy production and minimize electricity generation costs. The correctly estimation of the wind speed distribution parameters is important in terms of selection of the wind energy conversation system and obtains correct results on the wind power potential and economic feasibility of the project Pallabazzer (2003).

This paper aims to estimate $k$ e $c$ parameters by the application of Ant Colony 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 ($\chi^2$).

## 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 $\overline{v}$, $\sigma$, $\Gamma$ 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 $\overline{v}$ and $\sigma$ 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).

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

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

## 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):

$$E_{pf} = \frac{\overline{v^3}}{\overline{v}^3} \tag{11}$$

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

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

## 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:

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

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

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

## 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):

$$y_i = ax_i + b \tag{16}$$

$$\epsilon_i = y_i - (ax_i + b) \tag{17}$$

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

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

$$b = \overline{y} - a\overline{x} \tag{20}$$

$$k = a \tag{21}$$

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

## 2.8 Chi-Square Method ($\chi^2$)

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 Ant Colony Optimization (ACO)

In an ant colony the communication between individuals, or between the individuals and the environment, is based on the pheromone produced by them. The trail pheromone is a specific type of pheromone that some ant species use to mark paths on the ground. When detecting pheromone trails, forage ants may follow the path trodden by other ants to the food source. The first ants when sniffing the pheromone tend to choose, probabilistically, the trails marked with stronger concentrations of pheromone. The second group of ants will notice more intense the shortest path due to the shorter evaporation time. With the continuation of this procedure by all the ants, at one point in this process, one of the paths stands out for being the most frequented, being indicated by the intensity of ants pheromone and density superior to the others. At this point, the best path found by the ants is defined. This behavior inspired the optimization method by ant colonies Dorigo and stützle (2004).

In the ACO method, the parameters k and c of the Weibull curve form a Cartesian plane that is divided into N equal parts. The center point of each new area will be an ordered pair (k, c) where the curve fit will be evaluated Azevedo (2005).The probability of occurrence of each reticulum is defined by Equation 26.

$$P_r = \frac{\tau_r}{\sum_{r=1}^{R} \tau_r} \tag{26}$$

Where $\tau_r$ is the pheromone intensity for the reticle r.

Each ant is then randomly positioned in the plane through a roulette draw, where each slice of the roulette represents a reticle and is defined by the probability of occurrence. The visited quadrants are indicated by the pheromone deposit according to Equation 27. At each iteration the amount of the hormone is also reduced at a constant rate to simulate the hormone volatility, according to Equation 28.

$$\tau_{i,r} = \tau_{i-1,r} + \left( \frac{\mu}{err_f} \right) \tag{27}$$

$$\tau_{i,r} = \rho \cdot \tau_{i,r} \tag{28}$$

Where $\tau_{i,r}$ is the pheromone intensity for the reticle r, at iteration i, $err_f$ is the error evaluated by the ant f, $\mu$ the deposition constant and $\rho$ is the evaporation constant Socha (2009).

While the iterations follow up, some reticles will be more attractive to ants because they have a large amount of pheromone, this attraction being symbolized by the larger slices of the roulette, until most of the ants follow the same path.

## 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 SMS08 station, located in São Martinho da Serra, RS, at 50m at ground level and with one year of data for reasons of availability, the year 2010 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 29 until Equation 32 respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i^{calculated} - y_i^{measured})^2}{n}} \tag{29}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i^{calculated} - y_i^{measured}\right| \tag{30}$$

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

$$WPD = \left(\frac{WPD_{estimated} - WPD_{measured}}{WPD_{measured}}\right) \cdot 100 \tag{32}$$

Where, according to (Jamil *et al.*, 1995), WPD$_{measured}$ and WPD$_{estimated}$ are calculated respectively by Equations 33 and 34

$$WPD_{medido} = \frac{1}{2} \cdot \rho \cdot c^3 \cdot \Gamma\left(1 + \frac{3}{k}\right) \tag{33}$$

$$WPD_{estimado} = \frac{1}{2} \cdot \rho \cdot v^3 \tag{34}$$

Where $\rho$ 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 ACO method was calculated based on the parameters, Ants number equal to 100, evaporation constant equal to 0.1 and deposition constant equal to 0.2. Figure 1 compares eight deterministic methods and the heuristic method, ACO.
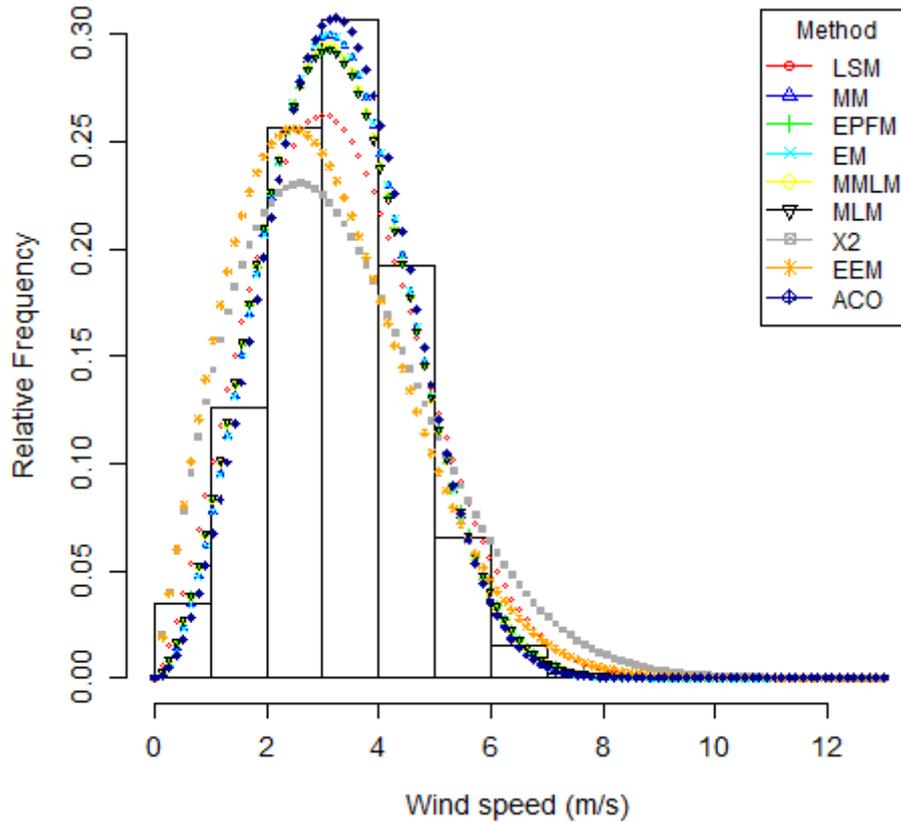


Figure 1. ACO and eight deterministic methods Comparison

The results of the statistical tests for the SMS08 station located in São Martinho da Serra are presented in Tab.1.

Table 1. Statistical Analysis of São Martinho da Serra, year 2010.

| Method | k | c | RMSE | MAE | $R^2$ | WPD (%) |
|--------|------|------|------|------|------|---------|
| LSM | 2.4119 | 3.7395 | 0.005247 | 0.011540 | 0.961880 | 14.946297 |
| MM | 2.7760 | 3.6717 | 0.002251 | 0.005308 | 0.992982 | -0.080201 |
| EPFM | 2.7050 | 3.6751 | 0.002684 | 0.006282 | 0.990020 | 1.578068 |
| EM | 2.7833 | 3.6713 | 0.002214 | 0.005208 | 0.993208 | -0.244181 |
| MMLM | 2.7188 | 3.6600 | 0.002665 | 0.006216 | 0.990160 | 0.059538 |
| MLM | 2.6946 | 3.6662 | 0.002799 | 0.006508 | 0.989148 | 1.056167 |
| $\chi^2$ | 1.9734 | 3.6870 | 0.010051 | 0.024919 | 0.860125 | 31.587672 |
| EEM | 2.0573 | 3.4135 | 0.010163 | 0.021280 | 0.856987 | $-1.110223 \cdot 10^{-14}$ |
| ACO | 2.9100 | 3.7170 | 0.001692 | 0.003701 | 0.996033 | 1.324812 |

According to the Table 1, it can be observed that ACO method presented the lowest RMSE and MAE tests with values of 0.001692 and 0.003701, respectively. The heuristic method also presented the best performance when it was analyzed the and $R^2$ test with value of 0.996033. The WPD results showed a superiority of the EEM among all methods tested with value of $1.11 \cdot 10^{-14}\%$. ACO performed well, since the value obtained, 1.324812%, less than 2%, which was below the acceptable limit for the deviation of Wind Power Density.

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

## 6. CONCLUSION

In this paper, eight deterministic and one heuristic optimization methods namely Ant Colony Optimization were used to estimate the parameters, $k$ and $c$, of the Weibull distribution for São Martinho da Serra, 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 ACO 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 São Martinho da Serra, 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 $1.11 \cdot 10^{-14}\%$.

2. Ant Colony Optimization was an efficient method, for determining the Weibull distribution, $k$ and $c$ parameters, for São Martinho da Serra, RS, Brazil.

## 7. ACKNOWLEDGEMENTS

## 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.

Azevedo, D.C.R., 2005. *Métodos Heurísticos Aplicados no Ajuste de Curvas de Weibull em Energia Eólica*. Ph.D. thesis, Mechanical Engineering Course, Mechanical Engineering Department, Federal University of Ceará, Fortaleza.

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.

Dorigo, M. and stützle, T., 2004. *Ant Colony Optimization*. MIT Press, England. ISBN 0-262-04219-3.

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

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.

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.

Pallabazzer, R., 2003. "Parametric analysis of wind siting efficiency". *J Wind Eng Indus Aerod*, Vol. 91, pp. 1329 – 1352.

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.

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.

Socha, K., 2009. *Ant Colony Optimisation for Continuous and Mixed-variable Domains*. Ph.D. thesis, UniversitÂťe Libre de Bruxelles, Av. Franklin D. Roosevelt 50, 1050 Brussels, Belgium.

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.