DOI: https://doi.org/10.54966/jreen.v1i1.1237

Journal of Renewable Energies

Revue des Energies Renouvelables

journal home page: https://revue.cder.dz/index.php/rer

Conference paper

The potential of mixing model of wind speed distribution in Algerian High Plateaus

Mohamed Amin Hellalbi [*], **Ahcene Bouabdallah**

Laboratory of Thermodynamics and Energy Systems, Faculty of Physics, University of Science and Technology Houari Boumediene, Bp 32 El Alia, 16111 Bab Ezouar, Algiers, Algeria

ARTICLE INFO

Article history:

Received July 29, 2024

Accepted September 4, 2024

Keywords:

Wind energy assessment

Mixed distribution,

Expectation Maximization,

Algerian High Plateaus.

ABSTRACT

The evaluation of wind energy relies primarily on the probability density function PDF, which corresponds well with the wind speed data. Single PDFs are widely used in the assessment of wind. In contrast, homogeneous or heterogeneous mixed models are rarely used, especially in Algeria, where the bimodal wind speed distribution is expected. This research aims to investigate the potential of heterogeneous PDFs Generalized Extreme Value-Weibul and Normal-Extreme Value PDFs in assessing wind energy at three meteorological stations in the high plateau against the single widespread PDFs Weibull and GEV by analyzing five years of archived wind speed data. The estimation of mixed model parameters is obtained by applying the Expectation-Maximization algorithm, and the identification of the appropriate PDF is obtained by four goodness-of-fit (GOF) criteria and compared with the widespread single distributions. The results show that the mixed model surpasses widely the single model for all the GOF criteria used at the three selected sites. The proposed mixed model fits all the wind speed distributions related to unimodal and bimodal regimes.

1. INTRODUCTION

Understanding the characteristics, patterns, and behaviors of wind power is vital to the selected sites (Alfawzan F & Alleman JE, 2020). Moreover, the random nature of the wind distribution must represented using the Power density function (PDF) (Jung C, & Schindler D, 2019). A wide range of PDFs are used to define the wind speed distribution all over the world (Jung C, & Schindler D, 2019). Besides, the wind distribution patterns change and the single PDF can not describe all the regimes that exist (Ouarda T M & Charron C, 2018). In Algeria, Weibull two parameters is the most utilized PDF (Hammouche R, 1990; Boudia S M et al., 2016). Meanwhile, (Aries et al., 2018) , (Guerri et al.,2020;

[*] *Corresponding author, E-mail address: Aminoo89@yahoo.fr*
Tel : + 213 664399087

ISSN: 1112-2242 / EISSN: 2716-8247

Ounis and Aries., 2021) proved that Generalized extreme value (GEV) is a more suitable PDF. On the other hand, (Hellalbi M A & Bouabdallah A, 2024) illustrates the bimodal nature of wind speed distribution in Algeria's Coastal regions using a mixed model and outperforming the single model. Algeria's high plateaus are a promising region for wind energy applications. Meanwhile, there is a lack of studies that apply the mixed PDFs representing wind speed distribution.
In this work, we used wind data at 10 m altitude archived at three meteorological stations located in Algerian's high plateaus for five years, from 2017 to 2021, and applied a mixed model PDF Weibull-GEV and Normal-extreme using the expectation-maximization method and compared with the widespread PDFs Weibull and GEV for the wind speed distribution at the selected sites.

2. DATASETS AND METHODS

2.1 Wind data

The datasets collected every three hours of wind data (wind speed) for five years from 2017 to 2021 at 10 m AGL at three meteorological stations in the high plateaus region of Algeria, whose names and geographic positions are listed in the table 1, are utilized in this study.

Table 1. Geographic position of the meteorological stations

Station	Wilaya	Latitude	Longitude	Elevation [m]
BOUCHEKIF	TIARET	35.341136	1.463147	989.07
EL_BAYADH	EL_BAYADH	33.7166667	1.0833333	1347.0
CHEIKH_LARBI_TEBESSI	TEBESSA	35.431611	8.120717	811.07

2.2 Wind speed distribution models

2.2.1 Weibull PDF

The Weibull defined by the PDF and the CDF by the equations 1 and 2 respectively (Chen X et al., 2020) as**.**

$$f(v_i) = \left(\frac{\beta}{\alpha}\right)\left(\frac{v_i}{\alpha}\right)^{\beta-1} e^{\left[-\left(\frac{v_i}{\alpha}\right)^{\beta}\right]} \quad v_i > 0 \tag{1}$$

$$F(v_i) = 1 - e^{\left[-\left(\frac{v_i}{\alpha}\right)^{\beta}\right]} \tag{2}$$

where β and α denotes the shape and the scale parameters.

2.2.2 Generalized Extreme Value PDF

GEV is considered to include three functions (Gumbel, Fr´echer, Weibull) (Kotz S et al., 2000). The PDF and CDF are defined by the equations 3 and 4, respectively.

$$f(v_i) = \frac{1}{\alpha} e^{-\left(1+\frac{\beta(v_i-\mu)}{\alpha}\right)^{-1/\beta}} \left(1 + \frac{\beta(v_i - \mu)}{\alpha}\right)^{-\frac{1}{\beta}-1} \tag{3}$$

$$F(v_i) = e^{-\left(1+\frac{\beta(v_i-\mu)}{\alpha}\right)^{-1/\beta}} \tag{4}$$

where $\beta \neq 0$ and$\left(\frac{\beta(v_i-\mu)}{\alpha} + 1\right) > 0$, β, α and μ are the shape, the scale, and the location parameters, respectively.

2.2.3 Weibull_GEV mixed model PDF

The mixed model of Weibull and GEV is a linear combination of the two models given by the following PDF and CDF:

$$f(v_i) = \omega_1 \left(\frac{\beta_1}{\alpha_1}\right)\left(\frac{v_i}{\alpha_1}\right)^{\beta_1-1} \mathrm{e}^{\left[-\left(\frac{v_i}{\alpha_1}\right)^{\beta_1}\right]} + \omega_2 \frac{1}{\alpha_2} e^{-\left(1+\frac{\beta_2(v_i-\mu)}{\alpha_2}\right)^{-\frac{1}{\beta_2}}} \left(1 + \frac{\beta_2(v_i-\mu)}{\alpha_2}\right)^{-\frac{1}{\beta_2}-1} \tag{5}$$

$$F(v_i) = \omega_1 \left(1 - \mathrm{e}^{\left[-\left(\frac{v_i}{\alpha_1}\right)^{\beta_1}\right]}\right) + \omega_2 e^{-\left(1+\frac{\beta_2(v_i-\mu)}{\alpha_2}\right)^{-\frac{1}{\beta_2}}} \tag{6}$$

where $\sum_{j=1}^{2} \ \omega_j = 1 and \ \ 0 \leqslant \omega_j \leqslant 1, \beta_2 \neq 0$ and$\left(\frac{\beta_2(v_i-\mu_2)}{\alpha_2} + 1\right) > 0$. Furthermore, ω_1, ω_2 are the mixing weights related to the PDFs and CDFs.

2.2.4 Normal_Extreme mixed model PDF

The linear combination of Normal and Extreme is given by the following PDF and CDF:

$$f(v_i) = \omega_1 \frac{e^{\left(\frac{v_i-\beta_1}{2\alpha_1^2}\right)}}{\sqrt{2\pi}\alpha_1} + \omega_2 \frac{e^{-e^{\left(\frac{-v_i+\alpha_2}{\beta_2}\right)}+\left(\frac{-v_i+\alpha_2}{\beta_2}\right)}}{\beta_2} \tag{7}$$

$$F(v_i) = \omega_1 \left(\frac{1}{2} Erf\, c\left(\frac{-v_i+\beta_1}{\sqrt{2}\alpha_1}\right)\right) + \omega_2\, e^{-e^{\frac{-v_i+\alpha_2}{\beta_2}}} \tag{8}$$

where ω_1 is the weight of the Normal PDF and CDF with α_1 and β_1 being the shape and scale parameters (Guerri et al.,2020). The ω_2 is the weight of the Extreme Value PDF and CDF with α_2 and β_2 being the shape and the scale parameters (Akgül F G & Şenoğlu B, 2019).

2.3 Log-likelihood parameter estimation

The log-likelihood method is expressed as

$$L(v_i) = l\,n\, l(v_i) = l\,n \prod_{i=1}^{n} f(v_i) = \sum_{i=1}^{n} l\,n\, f(v_i) \tag{9}$$

where $f(v_i)$ is the PDF and $i = 1, 2, 3, \ldots, n$ take the form 1 for Weibull, 3 for GEV, 5 for Weibull-GEV and 7 for Normal-Extreme. The log-likelihood related to Weibull, GEV, Weibull_GEV and Normal-Extremes can be written as the relations 10, 11, 12, and 13, respectively.

$$L_W(v_i|\beta,\alpha) = n\,ln(\beta) - n\,\beta ln(\alpha) + (\beta - 1)\sum_{j=1}^{n} ln(v_i) - \alpha^{-\beta}\sum_{j=1}^{n}(v_i)^{\beta} \tag{10}$$

$$L_{GEV}(v_i|\beta,\alpha,\mu) = n\,ln(\alpha) + \frac{\beta+1}{\beta}\sum_{i=1}^{n} ln\left(1+\beta\left(\frac{v_i-\mu}{\alpha}\right)\right) + \sum_{i=1}^{n}\left(1+\beta\left(\frac{v_i-\mu}{\alpha}\right)\right)^{-\frac{1}{\beta}} \tag{11}$$

$$L_{W\_GEV}(v_i|\omega_j,\alpha_j,\beta_j) = \sum_{i=1}^{n} \ln\left(\omega_1\left(\frac{\beta_1}{\alpha_1}\right)\left(\frac{v_i}{\alpha_1}\right)^{\beta_1-1} e^{\left(-\left(\frac{v_i}{\alpha_1}\right)^{\beta_1}\right)} + \omega_2\frac{1}{\alpha_2}e^{-\left(1+\frac{\beta_2(v_i-\mu)}{\alpha_2}\right)^{-\frac{1}{\beta_2}}}\left(1+\frac{\beta_2(v_i-\mu)}{\alpha_2}\right)^{-\frac{1}{\beta_2}-1}\right) \tag{12}$$

$$L_{N\_E}(v_i|\omega_j,\alpha_j,\beta_j) = \sum_{i=1}^{n} \ln(\omega_1\frac{e^{-\frac{(v_i-\alpha_1)^2}{2\beta_1^2}}}{\sqrt{2\pi}\beta_1} + \omega_2\frac{e^{-e^{\frac{-v_i+\alpha_2}{\beta_2}}+\frac{-v_i+\alpha_2}{\beta_2}}}{\beta_2}) \tag{13}$$

Estimating the parameters concerning each PDF is obtained by setting the partial derivative of the likelihood with respect to that parameter to zero, which makes maximum log-likelihood. The maximum log-likelihood regarding Weibull PDF is calculated in such a way $\frac{\partial \mathrm{L_W}}{\partial \beta} = 0, \frac{\partial \mathrm{L_W}}{\partial \alpha} = 0$

$$\begin{cases} \beta = [\frac{\sum_{i=1}^{n} v_i^{\beta}\ln(v_i)}{\sum_{i=1}^{n} v_i^{\beta}} - \frac{\sum_{i=1}^{n}\ln(v_i)}{n}]^{-1} \\ \alpha = \sqrt[\beta]{\frac{1}{n}\sum_{i=1}^{n}(v_i)^{\beta}} \end{cases} \tag{14}$$

The parameter β of Weibull PDF can be estimated by solving the equation 14 using fixed point method (Hoffman J D & Frankel S, 2018), whereas α is calculated directly if β is established. Similarly, for GEV, the maximum likelihood is reached by the following procedures

$$\frac{\partial \mathrm{L_{GEV}}}{\partial \alpha} = 0, \frac{\partial \mathrm{L_{GEV}}}{\partial \beta} = 0, \frac{\partial \mathrm{L_{GEV}}}{\partial \mu} = 0$$

respectively as.

$$\begin{cases} -\frac{\beta+1}{\beta\,\alpha}\sum_{i=1}^{n}\frac{(v_i-\mu)}{(1+\beta(\frac{v_i-\mu}{\alpha}))}-\frac{1}{\beta^2}\sum_{j=1}^{n}\frac{(1+\beta(\frac{v_i-\mu}{\alpha}))}{(1+\beta(\frac{v_i-\mu}{\alpha}))^{\frac{1}{\beta}}}+\frac{1}{\beta\,\alpha}\sum_{i=1}^{n}\frac{(v_i-\mu)}{(1+\beta(\frac{v_i-\mu}{\alpha}))^{1+\frac{1}{\beta}}}= \\ \frac{\beta+1}{\alpha}\sum_{i=1}^{n}(1+\beta(\frac{v_i-\mu}{\alpha}))^{-1}-\frac{1}{\alpha}\sum_{i=1}^{n}(1+\beta(\frac{v_i-\mu}{\alpha}))^{-1-\frac{1}{\beta}}=0 \\ -\frac{n}{\alpha}+\frac{\beta+1}{\alpha^2}\sum_{i=1}^{n}(v_i-\mu)(1+\beta(\frac{v_i-\mu}{\alpha}))^{-1}-\frac{1}{\alpha^2}\sum_{i=1}^{n}(v_i-\mu)(1+\beta(\frac{v_i-\mu}{\alpha}))^{-1-\frac{1}{\beta}}= \end{cases} \quad (15)$$

The parameters α, β, and μ are calculated by solving the algebraic system of equations system 15 using Newton's method [13].

2.4 Expectation Maximization algorithm

Due to the complexity of the log-likelihood related to the mixed model method as written in equations 12 and 13, the parameters can be estimated numerically. Therefore, the Expectation Maximization (EM) iterative method is applied to find the maximum likelihood of a statistical parameters PDF model from selected datasets (McLachlan G J, Krishnan T, 2007). The iterative process of the EM method is achieved in two steps: the expectation (E-step) and the maximization (M-step). In the E-step, the expectation of log-likelihood is a function created to augment the observed data to estimate the unobserved or missing data and generate the parameters related to the log-likelihood function. In the M-step, the log-likelihood parameters created in the E-step will be maximized to obtain the new parameter values. The EM algorithm requires the initial values to start the iterative process between the E-step and M-step until the algorithm's convergence, as mentioned in fgure1.

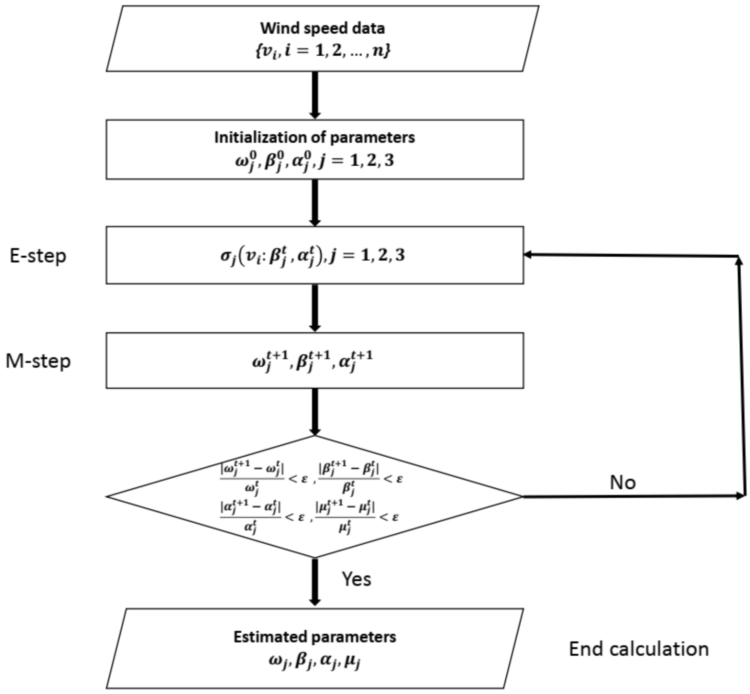

Fig 1. Flowchart of Expectation Maximization algorithm

2.5 Validation Models

To validate and select the best PDF model that reveals the wind speed distribution, four common goodness of fit criteria (Ouarda T, et al., 2016; Jung C, & Schindler D, 2019) are used in the present study. R^2, RMSE, SSE, and KS, are described in the table2. The wind speed datasets v_i in n category intervals and the relative probability $p(v_i)$ are calculated for each category. For the three goodness of fitR^2, RMSE, SSE formulas, the $f(v_i)$is the predicted PDF related to the theoretical model for i^{th} category interval, except for R^2 used $\overline{p}(v_i) = \frac{1}{n}\sum_{i=1}^{n}(p(v_i))$. The Kolmogorov-Smirnov KS statistic test corresponds to the maximum difference between F predicted and P observed CDF

Table 2.Goodness of fit metrics formulas

Criteria	Symbols	Formulas
Coefficient of determination	R^2	$1 - \frac{\sum_{i=1}^{n}(p(v_i) - f(v_i))^2}{\sum_{i=1}^{n}(p(v_i) - \overline{p}(v_i))^2}$
Root Mean Square Error	RMSE	$\sqrt{\frac{\sum_{i=1}^{n}(p(v_i) - f(v_i))^2}{n}}$
Sum of Squared Error	SSE	$\sum_{i=1}^{n}(p(v_i) - f(v_i))^2$
Kolmogorov-Smirnov	KS	$max\,\lvert P(v_i) - F(v_i)\rvert$

3. RESULTS

The wind speed of three meteorological stations in Algeria's high plateaus was used as datasets to estimate the parameters related to Weibull and GEV through the maximum likelihood method and Expectation Maximization algorithm for Weibull-GEV and Normal-Extreme mixing models. Furthermore, four goodness-of-fit criteria were applied to evaluate the PDF models for each site. Table 3 illustrates each PDF model's estimated parameters and the PDF models goodness of fit criterion was calculated. For the EL_BAYADH station, the Normal_Extreme PDF model ranked highest, achieving the best fit with an R^2 value of 0.755 and the lowest RMSE, KS, and SSE values at 4.49×10^{-4}, 0.0559, and 0.0073, respectively. Conversely, the GEV model performed the worst, with a minimum R^2 of 0.644 and the highest RMSE, KS, and SSE values of 5.41×10^{-4}, 0.1307, and 0.0111, respectively. At the CHEIKH_LARBI_TEBESSI station, the Weibull_GEV PDF was chosen for its higher R^2 of 0.6293 and lower RMSE of 8.39×10^{-4}. However, according to KS and SSE, the Normal_Extreme PDF was superior, with values of 0.0929 and 0.0216. The Weibull model performed the worst, mainly due to a negative R^2 of -0.0615, indicating its inadequacy in representing wind speed distribution at this station. Finally, at the Buchakif station, the Weibull_GEV PDF was the best model based on an R^2 of 0.718 and RMSE of 5.9×10^{-4}. The Normal_Extreme PDF was superior according to KS and SSE values of 0.0618 and 0.0106, while the Weibull model ranked last across all metrics.

Table 3. PDFs and four different goodness of fit models applied for the three meteorological

		Weibull	GEV	Weibull_GEV	Normal_Extreme
EL_BAYADH	R^2	0.696	0.644	0.704	0.755
	RMSE	$5 * 10^{-4}$	$5.41 * 10^{-4}$	$4.93 * 10^{-4}$	$4.49 * 10^{-4}$
	KS	0.0786	0.1307	0.0699	0.0559
	SSE	0.00817	0.0111	0.0080	0.0073
CHEIKH_LARBI_TEBESSI	R^2	-0.0615	0.4613	0.6293	0.6138
	RMSE	$1.42 * 10^{-3}$	$1.01 * 10^{-3}$	$8.39 * 10^{-4}$	$8.5 * 10^{-4}$
	KS	0.485	0.199	0.108	0.0929
	SSE	0.029	0.041	0.0227	0.0216
BOUCHEKIF	R^2	0.598	0.625	0.718	0.694
	RMSE	$7.04 * 10^{-4}$	$6.80 * 10^{-4}$	$5.90 * 10^{-4}$	$6.10 * 10^{-4}$
	KS	0.145	0.128	0.0624	0.0618
	SSE	0.0178	0.0167	0.0110	0.0106

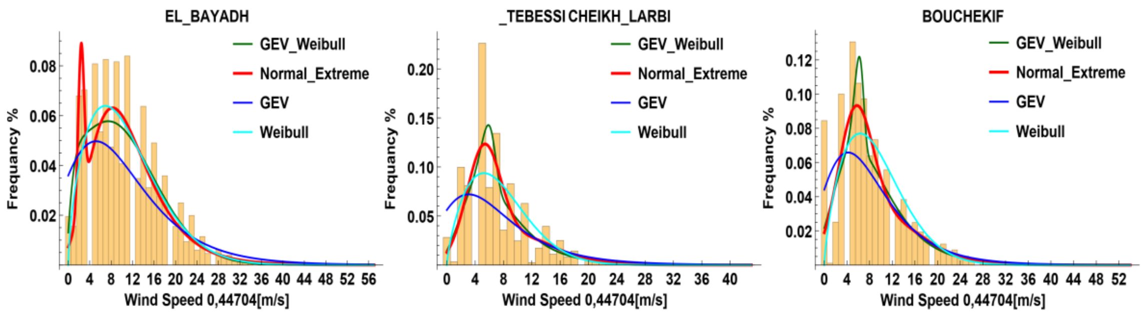

Fig 2. Wind speed histograms and the four probability density function predictions applied at three meteorological stations.

Figure 2. shows the wind speed histograms and the three fit PDFs plotted for the three sites selected in this study. Moreover, the proposed models can represent the wind speed distribution with unimodal, bimodal wind speed regimes. Figure 3. illustrates the CDFs of the three models and the discrete wind speed cumulative distributions. The figure shows that the mixed CDF model adapts well to the other CDF models with cumulative wind speed distribution at all the selected sites.

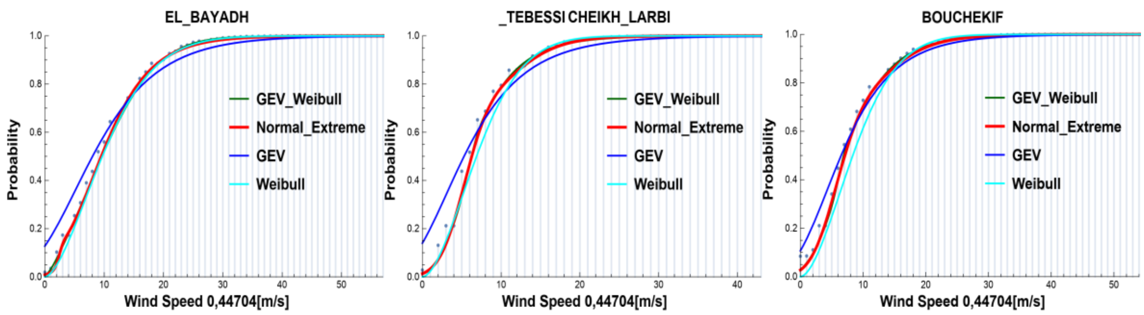

Fig 3. Wind speed cumulative probability and the four cumulative density functions estimated at the three meteorological stations

4 CONCLUSION

In this study, an investigation was conducted to evaluate the Probability Density Function PDF suitable for wind speed distribution in Algeria's high plateaus. The investigation used a five-year wind dataset from 2017 to 2021 from three distinct meteorological stations. Two mixed model PDFs, Weibull-GEV and Normal_Extreme, were applied to fit the wind speed distribution at the selected high plateau stations. The parameters of the mixed model PDFs were estimated using the expectation-maximization method. To evaluate the effectiveness of the PDFs, we used four goodness-of-fit criteria R^2, RMSE, SSE, KS to compare it with two widely used PDFs, Weibull two parameters and GEV. The newly used models outperform the two alternative PDFs, Weibull and GEV, for all sites concerning all the GOFs utilized. Furthermore, the mixed model's PDF can fit all wind speed distributions related to different unimodal and bimodal regimes at all meteorological stations in Algeria's high plateau regions, which is considered an appropriate model.

NOMENCLATURE

$f(v_i)$ Probability Density Function PDF	$RMSE$ Root Mean Square Error
$p(v_i)$ Probability relative to wind speed	SSE Sum of Square Error
$\overline{p}(v_i)$ Mean of relative probability	v_i Wind speed
$F(v_i)$ Cumulative Density Function CDF	β Shape parameter
KS Kolmogorov Smirnov criteria	α Scale parameter
$L(v_i)$ Log likelihood	μ Location parameter
R^2 Coefficient of determination	ω Mixing weight parameter

REFERENCES

Akgül F G, Şenoğlu B, (2019) Comparison of wind speed distributions: a case study for Aegean coast of Turkey, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects. doi.org/10.1080/15567036.2019.1663309 .

Alfawzan F, Alleman JE, (2020) Wind energy assessment for NEOM city, Saudi Arabia, Energy Science & Engineering 8 755–67. doi.org/10.1002/ese3.548.

Aries N, Boudia S M, Ounis H, (2018) Deep assessment of wind speed distribution models: A case study of four sites in Algeria, Energy Conversion and Management 155. doi.org/10.1016/j.enconman.2017.10.082.

Boudia S M, Benmansour A, Tabet Hellal M A, (2016) Wind resource assessment in Algeria,Sustainable Cities and Society 22. doi.org/10.1016/j.scs.2016.02.010.

Chen X, Foley A, Zhang Z, Wang K, O'Driscoll K, (2020) An assessment of wind energy potential in the Beibu Gulf considering the energy demands of the Beibu Gulf Economic Rim, Renewable and Sustainable Energy Reviews 119. doi.org/10.1016/j.rser.2019.109605.

Guerri O, Dali A, Boudia S M, Yassaa N, (2020) Performance evaluation of a wind farm using different power density distributions, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects.

Hammouche R, (1990) Atlas vent de l'Algérie, Office National de la Météorologie. https://backend.orbit.dtu.dk/ws/portalfiles/portal/116709838/Atlas_Vent_de_l_Algerie.pdf.

Hellalbi M A, Bouabdallah A, (2024) Elaboration of a generalized mixed model for the wind speed distribution and an assessment of wind energy in Algerian coastal regions and at the Capes, Energy Conversion and Management 305. doi.org/10.1016/j.enconman.2024.118265.

Hoffman J D, Frankel S,(2018) Numerical methods for engineers and scientists, CRC press.

Jung C, Schindler D, (2019) Wind speed distribution selection – a review of recent development and progress, Renewable and Sustainable Energy Reviews 114 109-290. doi.org/10.1016/j.rser.2019.109290.

Kotz S, Nadarajah S,(2000) Extreme Value Distributions, PUBLISHED BY IMPERIAL COLLEGE PRESS AND DISTRIBUTED BY WORLD SCIENTIFIC PUBLISHING CO, URL: https://www.worldscientific.com/doi/abs/10.1142/p191. doi:10.1142/p191.

McLachlan G J, Krishnan T,(2007) The EM algorithm and extensions, John Wiley & Sons.

Ouarda T M, Charron C, (2018) On the mixture of wind speed distribution in a Nordic region, Energy Conversion and Management 174 33–44. doi.org/10.1016/j.enconman.2018.08.007.

Ouarda T, Charron C, Chebana F, (2016) Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and l-moment ratio diagram methods, with a case study, Energy Conversion and Management 124. doi.org/10.1016/j.enconman.2016.07.012.

Ounis H, Aries N, (2021) On the wind resource in Algeria: Probability distributions evaluation, Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy235. doi.org/10.1080/15567036.2020.1775727.