Assessing the Best Fit Probability Distribution Model for Wind Speed Data for Different Sites of Burkina Faso

Drissa Boro[1*], Ky Thierry[1], Florent P. Kieno[1] and Joseph Bathiebo[1]

[1]*Laboratoire d'Energies Thermiques Renouvelables (L.E.T.R.E), UO1 10 BP: 13495 Ouaga 10, Burkina Faso.*

Authors' contributions

This work was carried out in collaboration among all authors. Author DB designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Authors KT and FPK managed the analyses of the study. Author JB managed the literature searches. All authors read and approved the final manuscript.

Original Research Article

ABSTRACT

In order to estimate the power output of a wind turbine, optimise its sizing and forecast the economic rate of return and risks of a wind energy project, wind speed distribution modelling is crucial. For which, Weibull distribution is considered as one of the most acceptable model. However, this distribution does not fit certain wind speed regimes. The objective of this study is to model the frequency distribution of the three-hourly wind speed at ten sites of Burkina Faso. In this context, we compared the accuracy of five distributions (Weibull, Hybrid Weibull, Rayleigh, Gamma and inverse Gaussian) which gave satisfactory results in this field. The maximum likelihood method was used to fit the distributions to the measured data. According to the statistical analysis tools (the coefficient of determination and the root mean square error), it was found that the Weibull distribution is most suited to the Bobo, Dédougou, Ouaga and Ouahigouya sites. On the other hand, for the sites of Bogandé, Fada and Po, the hybrid Weibull distribution is the most suitable one. As to the inverse Gaussian distribution, it is the most suitable for the Boromo, Dori and Gaoua sites. In addition, the

_____

*Corresponding author: E-mail: drissaboro5@gmail.com;

analysis focused on comparing the mean absolute error of the annual wind power density estimation using the distributions examined. The Hybrid Weibull distribution was found to have a minimal mean absolute error for most study sites.

Keywords: Wind potential; Weibull distribution; power density; vertical profile.

1. INTRODUCTION

Wind energy is one of the fastest growing sectors and one of the renewable energy sources widely used to meet energy demand in urban and rural areas in some African countries [1]. In fact, the pollution caused by the overuse of fossil fuels and their limited reserves make wind an alternative energy source for overcoming man-made environmental problems. The use of wind energy as green and sustainable energy can reduce dependence on fossil fuels which are the main sources in countries' energy supply chains. Wind energy has been used worldwide over the last decades, but its growth has been most significant in recent years. Assessing the wind energy characteristics and potential is a critical and major step in the economic development of wind energy. Actually, the wind speed frequency distribution represents wind speed data collected over a long period of time. Thus, this information is essential for assessing the wind energy potential of a particular location. It is also worth mentioning that wind turbines installed at two different sites with similar average wind speeds can generally produce completely different energy due to differences in wind speed characteristics. This further underlines the importance of knowing the wind speed distribution. As a rule, a frequency distribution can be calculated for a series of data. The wind speed frequency distribution can be determined using two approaches. The former is an approach based on wind speed time series and the latter on probability distribution functions. The time series approach seems to be more accurate owing to the direct use of the original wind speed data. However, since time series wind speed data are often huge, it would be preferable to have only a few main parameters so as to explain the behaviour and characteristics of a wide range of wind speed data [2]. In this respect, wind speed data can be fitted using probability distribution functions that simplify the characteristics of a wind regime into a limited set of parameters [2]. The probability density function of a random variable is a mathematical model that characterizes the probability that this variable will occur at a certain point in time in each observation interval. The cumulative distribution function, on the other hand, specifies the probability that a variable is less than or equal to a particular value.

In recent decades, the issue of modelling wind speed frequencies has motivated researchers to use different distribution functions in order to identify the most appropriate ones. The two-parameter Rayleigh and Weibull distribution functions are two popular functions that have been widely used in many studies [3,4,5]. There are many other functions that have been typically used for wind energy evaluations such as the Generalized Gamma [6], gamma [2], inverse gamma [7], inverse Gaussian [8], 2- and 3-parameter lognormal [9], Gumbel [10], Burr [8], Erlang [11], Kappa [12], Wakeby [11] and Generalized Extreme Value distribution [9]. In this regard, the identification of the most suitable functions that offer the best adjustments to the data are of vital importance. Indeed, the use of a distribution function that more precisely fits the wind speed dataset is useful to reduce uncertainties in wind energy estimation. To our knowledge, there is a lack of extensive research on the determination of appropriate frequency distribution models for estimating the wind speed distribution across the different regions of Burkina Faso. Therefore, in this study, the capacity of different distribution functions was evaluated to provide a better fit to wind speed data at ten sites located throughout the whole country.

The main objective is to identify the most appropriate distribution function for the wind speed dataset at the different study sites. With this end in mind, the performance of the Rayleigh, Weibull, hybrid Weibull, Gamma and inverse Gaussian distribution functions was compared. Their efficiency was also statistically evaluated based on statistical parameters. To estimate the wind energy available at these different sites, investors can directly use the distribution appropriate to each study site.

2. STUDY AREA AND DATA PRESENTATION

With a surface area of 274,200 km^2, Burkina Faso is a country located in the heart of West

Africa between latitudes 9°2'N and 15°05'N and longitudes 2°20'E and 5°30' W at an average altitude of 300 m above sea level. This landlocked country is surrounded by six other countries: Mali in the West and North, Niger in the East and Benin, Togo, Ghana and Ivory Coast in the South. In Burkina Faso, the climate classification highlights 3 major climate zones according to rainfall and temperature – namely the Sahelian, Sudano-Sahelian and Sudanese climate zone. The Sahelian climate zone lies to the North of the 14th parallel and characterized by an annual rainfall less than 650mm. The Sudano-Sahelian climate zone is located between latitudes 11°30N and 14° N characterized by an annual rainfall ranging from 650 to 1,000 mm. As to the Sudanese climate zone, it is situated in the South at a latitude of 11°30'N characterized by an annual rainfall above 1,000 mm [13]. Our study was conducted at the Ouahigouya site (Sahelian climatic zone). Data recorded and provided by the meteorological station of the Burkina Faso National Meteorological Agency (ANAM in French) over the period from January 2006 to December 2016 were used in this study. The series of data used consists of wind speeds measured every three hours using a cup anemometer positioned 10 m above ground level, mounted on a mast. Fig. 1 provides an overview of the study area and Table 1 shows the geographical coordinates of the sites selected for the study.
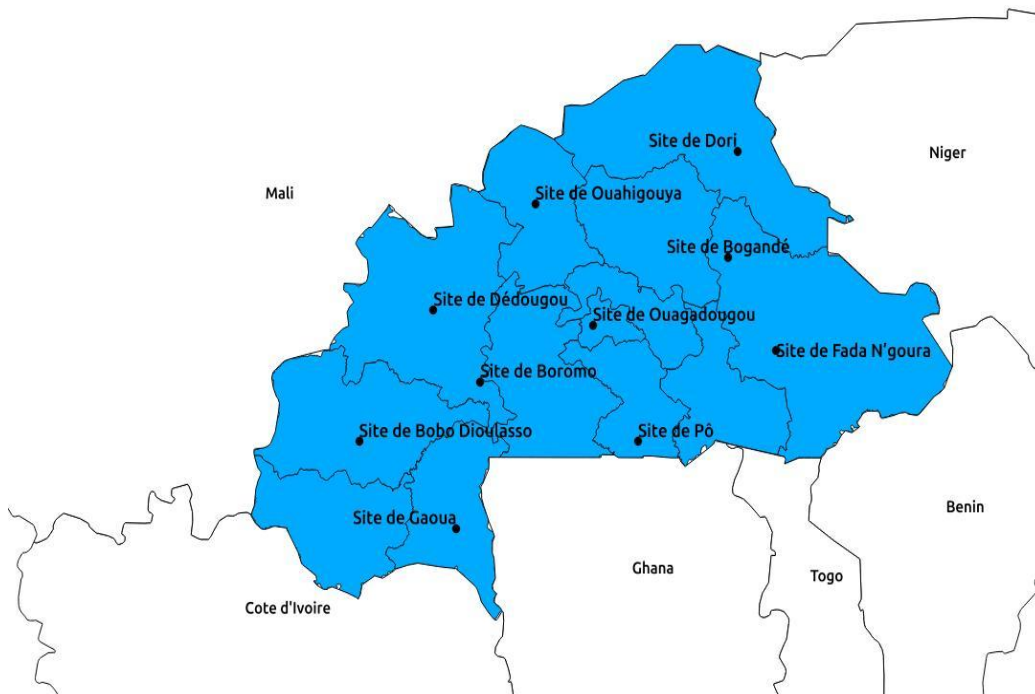
Fig. 1. Geographic location of the ten study sites

Table 1. Geographical coordinates of the study sites

Site	Longitude	Latitude	Altitude(m)
Dori	00°02' W	14° 02'N	282
Ouahigouya	02° 19' W	13° 31'N	328
Bogandé	00°08'W	12°59'N	295
Fada N'goura	00°25' E	12° 4' N	298
Po	01°09'W	11°10'N	305
Ouagadougou	01° 40'W	12° 19'N	299
Dédougou	03°28'W	12°28'N	302
Boromo	02°56'W	11°45'N	325
Bobo Dioulasso	04°18'W	11°10'N	423
Gaoua	03°12' W	10°18 N	329

3. METHODS

3.1 Modelling the Wind Speed Frequency Distribution

As a rule, the wind speed frequency distribution can be determined using two approaches – one based on wind speed time series and the other on probability distribution functions. The time series approach appears to be more accurate due to the direct use of the original wind speed data. However, since time series wind speed data are often huge, it would be preferable to have only a few main parameters so as to explain the behaviour and characteristics of a wide range of wind speed data [2]. For this purpose, the use of a continuous probability density function instead of a histogram makes it possible to calculate the statistical parameters analytically (such as mean velocity, median, Skewness, kurtosis, etc.). The empirical distribution of the measured mean wind speeds is then approximated by theoretical distribution functions. Knowledge of the statistical distribution law of wind speed from measured wind data is useful for wind energy applications, as the use of an analytical representation of the speed distribution has clear advantages. It is this approximate distribution that is used in the turbine formulas instead of the empirical histogram. It greatly simplifies the calculation of wind speed behaviour characterization as well as the potential and performance of wind energy systems [14]. Therefore, it is very important to determine the most appropriate functions that offer the best adjustments to wind speed data. Wind speed distribution modelling studies have been oriented towards models that combine power and exponential function. In most cases, the Weibull, Rayleigh and hybrid Weibull probability distributions are used in wind data analysis. In this study, five distribution functions were used to describe wind speed frequency distributions. These are the Weibull, Hybrid Weibull, Gamma, Rayleigh and inverse Gaussian distribution functions. Several methods are used to determine the distribution parameters [15,16, 17]. The one used in this study to calculate the parameters of the different distribution functions is the maximum likelihood method because of its accuracy at our study site [4].The theoretical desciptions of probability density functions are discussed in the following sections.

3.1.1 The two-parameter Weibull distribution

Based on the literature review [18,19,20], it should be noted that the two-parameter Weibull

distribution is the most commonly used mathematical model to estimate the wind energy available at a given site. It was first used by Davenport in 1963 for the calculation of wind stress [21]. In 1974, Justus used it for wind power [22]. Weibull modelling is general in the sense that it includes exponential (k=1) and Rayleigh (k=2) distributions which are only special cases of this function. The Weibull distribution function and the cumulative distribution function are respectively given by the regular expressions (1) and (2) [23]:

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

$$F(v \le v_x) = \int_0^{v_x} f(v)dv = 1 - \exp\left(-\left(\frac{v_x}{c}\right)^k\right) \quad (2)$$

Where $f(v,c,k)$ is the frequency distribution of measured wind speeds or the frequency of occurrence (frequency of occurrence) of wind speeds, (k) is the dimensionless shape parameter ($k>0$) (it characterizes the shape of the frequency distribution), (c) is the scale parameter, expressed in metres per second and of positive value (it indicates the average wind speed characteristic of the site). The shape and scale parameters are estimated by the maximum likelihood method using equations (3) and (4) respectively.

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

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

Where v_i is the wind speed at time step i and n is the number of non-zero wind speed observations. Equation (3) can be solved using an iterative procedure (k=2 is the appropriate initial conjecture). Then, equation (4) is solved explicitly. Care should be taken to apply equation (3) only to non-zero wind speed data points.

3.1.2 The two-parameter hybrid Weibull distribution

At sites where the frequency of calm winds is relatively high (> 15% of the total number of wind observations) the Weibull distribution does not perfectly fit the situation. Indeed, this rather significant proportion of calm winds cannot be neglected. We therefore use the so-called hybrid Weibull distribution, defined by the probability density function $f_{hw}(v,k,c)$ described by:

$$f_{hw}(v,k,c) = (1-F_0)\left(\frac{k}{c}\right)\left(\frac{v}{c}\right)^{k-1}\exp\left(-\left(\frac{v}{c}\right)^k\right)$$

with $v > 0$ (5)

Where ff_0 is the frequency of calm winds, given by $ff_0 = f(v)$ for $v = 0$. The maximum likelihood method can be used to estimate the parameters of this distribution. As indicated above, the classical Weibull distribution is not appropriate to areas where calm wind frequencies are relatively high. In this case, it is advisable to process the data by removing the calm wind values and listing them separately [24]. The shape and scale parameters are given by the expressions obtained with the classical weibull distribution, i.e. equations (3) and (4).

3.1.3 Rayleigh distribution

The Rayleigh distribution is a special case of the Weibull distribution for the case where the shape factor k is equal to 2. The coefficient of skewness and the flattening coefficient are constants. Its probability density and cumulative distribution function are given by equations (6) and (7) respectively:

$$f(v) = \left(\frac{2v}{c^2}\right)\exp\left(-\left(\frac{v}{c}\right)^2\right)$$

       (6)

$$F(v,c) = 1 - \exp\left(-\left(\frac{v}{c}\right)^2\right)$$

       (7)

The determination of the scaling parameters by the maximum likelihood method does not require an iterative procedure and is given by equation (8) [24]:

$$c = \left(\frac{1}{2n}\sum_{i=1}^{n}v_i^2\right)^{1/2}$$

       (8)

3.1.4 The inverse Gaussian distribution function

The inverse Gaussian distribution is suggested as an alternative to the three-parameter Weibull distribution for the description of wind speed data with low wind frequencies [25]. Equations (9) and (10) respectively give the probability density function and cumulative probability function of the inverse Gaussian distribution in which k is the shape parameter and c the scale parameter [26].

$$f(v,c,k) = \left(\frac{k}{2\pi v^3}\right)^{\frac{1}{2}}.\exp\left[-\frac{k(v-c)^2}{2vc^2}\right]$$

       (9)

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

       (10)

With $$\Phi(z) = \frac{1}{2}\left[1 + erf\left(\frac{z}{\sqrt{2}}\right)\right]$$

Where Φ is the cumulative function of the standard normal distribution [27]. Following the maximum likelihood method, the expressions of the scale and shape parameters which do not require any iterative procedure are given by equations (11) and (12).

$$c = \frac{1}{n}\sum_{1=1}^{n}v_i$$

       (11)

$$k = n\left[\sum_{i=1}^{n}v_i^{-1} - n^2\left(\sum_{i=1}^{n}v_i\right)^{-1}\right]^{-1}$$

       (12)

3.1.5 The two-parameter gamma distribution function

Statistical studies have shown that the two-parameter gamma distribution is adequate to describe the distribution of surface wind speeds almost everywhere in Europe [28]. The probability density function and the cumulative function of the 2-parameter generalized Gamma distribution are given by equations (13) and (14), respectively.

$$f(v,k,c) = \frac{v^{k-1}}{c^k \Gamma(k)} \exp\left(-\frac{v}{c}\right), v > 0 \tag{13}$$

$$F(v,k,c) = \gamma\left(k, \frac{v}{c}\right) \Big/ \Gamma(k) \tag{14}$$

In the gamma distribution, k is the shape parameter, c is the scale parameter and γ is the lower incomplete gamma function given by equation (15) [29]:

$$\gamma(p,x) = \int_0^x t^{p-1} e^{-t} dt \tag{15}$$

The shape and scale parameters, evaluated using the maximum likelihood method [24,30, 31], are given by equations (16) and (17) and require no iterative procedure.

$$k = \frac{n \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i \ln x_i - \sum_{i=1}^n \ln x_i \cdot \sum_{i=1}^n x_i} \tag{16}$$

$$c = \frac{1}{n^2}\left(n \sum_{i=1}^n x_i \ln x_i - \sum_{i=1}^n \ln x_i \cdot \sum_{i=1}^n x_i\right) \tag{17}$$

3.2 Model Validation Test

For the purpose of testing the different methods, the statistical analysis parameters viz., Coefficient of Determination (R^2) and Root Mean Squared Error (RMSE) were computed by Eqs. (18) and (19), and are given as below [32]:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - x_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \tag{18}$$

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^N (y_i - x_i)^2\right]^{1/2} \tag{19}$$

Where n is the total number of intervals, y_i is the frequency of the observed values, x_i is the frequency of the values obtained from probability distribution models and $\bar{y}$ is the mean value of y_i. A model is said to be ideal if it is characterized by a value of zero for RMSE and 1 for the parameter R^2.

3.2.1 Relative percentage error in wind power density estimation

According to Celik [33], in the field of wind energy, speed distribution functions are ultimately used to model wind energy density. Therefore, the most important criterion for the suitability of a potential wind speed distribution function should be based on its effectiveness in predicting the observed wind power density. In addition, a distribution may be better for fitting the histogram of observations but not for estimating the power density. For example, the Weibull and Rayleigh models were compared by Ahmed and Mahammed in (2012) [34]. They found that, on an annual scale, the Weibull distribution matched the measured data better than the Rayleigh distribution, but the latter provides greater accuracy in estimating power density over 9 months. This is done by comparing the mean absolute errors in estimating wind energy using the five distributions and the observed data. For a given theoretical probability density function fitted to wind speed data, the wind power corresponding to a particular wind speed is given by equation (20).

$$P(v) = \frac{1}{2}\rho v^3 f(v) \tag{20}$$

Where ρ is the air density. Another popular approach is to compare the mean power density generated from the theoretical probability density function with the mean power density calculated from observed wind speed data. The former density is, then, obtained by integrating equation (20) and given by equation (21).

$$\bar{P}_{th} = \frac{1}{2}\int_0^\infty \rho v^3 f(v) dv \tag{21}$$

The latter density is given by equation (22).

$$\bar{P}_{ob} = \frac{1}{2}\rho \frac{1}{n}\sum_{i=1}^n v_i = \frac{1}{2}\rho \bar{v}^3 \tag{22}$$

The difference between the theoretical power density ($\bar{P}_{th}$) and the observed power density ($\bar{P}_{ob}$) is often represented by the mean absolute percentage error given by expression (23).

$$\varepsilon_d = \left| \frac{\overline{P}_{th} - \overline{P}_{ob}}{\overline{P}_{ob}} \right| \times 100 \tag{23}$$

4. RESULTS AND DISCUSSION

4.1 Adjustment of Measured Data

As mentioned earlier in this study, we tested the effectiveness of five distributions that are widely applied to adjust wind speed, namely Weibull, Hybrid Weibull, Rayleigh, Gamma and inverse Gaussian. The graphs in Fig. 2 illustrate the histograms of wind speed data fitted by the five distribution functions for each study site. Graphically, it can be noticed that the Rayleigh function is unsuitable for the data observed for all study sites while the other functions are acceptable. Table 2 presents the parameters of the considered distribution functions obtained using the maximum likelihood method in MATLAB software.

4.2 Adjustment Performance Indicators

In this study, the goodness of fit of the different distribution functions is evaluated to characterize wind speed distribution for ten sites throughout the whole Burkinabe territory. To better assess the fit quality, we applied the most commonly used statistical indicators, namely the coefficient of determination (R^2) and the root mean square error (RMSE). Table 3 presents the statistical indicator values obtained and the precision rank for the distributions studied. These indicators are presented graphically in Figs. 3 and 4. For the ten sites studied, the suitable distributions are those with the lowest root mean square error values (close to zero) and the highest coefficient of determination (closer to unity).

It should be noted that, on the basis of all the statistical indicators, the Weibull distribution is the most appropriate one for the sites of Bobo (k=2.58, c=3.76), Dédougou (1.83, c=3.27), Ouaga (k=1.97, c=3.15) and Ouahigouya (k=1.93, c=2.51). For the Bogandé, Fada and Po sites, the hybrid Weibull distribution is the most suitable. The shape and scale parameters of this distribution for these three sites are respectively (k=1.93, c=3.18), (k=1.18, c=2.48), (k=1.93, c=2.51). It is notable that the weibull distribution has a performance close to that of the hybrid distribution at these three sites. Therefore, it can be used to model wind speed because of its popularity and for comparison purposes. The inverse Gaussian distribution is most suited to the Boromo (λ=0.72, μ =1.09), Dori (λ=0.77, μ =1.06) and Gaoua (λ=1.70, μ =1.75) sites. The Rayleigh distribution is the least suitable one, especially at the sites of Bobo and Ouahigouya. Overall, the calculated statistical parameters show that the distributions used model the real distributions very well with a coefficient of determination ranging between 0.741 and 0.955. The major conclusion that can be inferred from the analysis of Figs. 3 and 4 is that the most efficient distribution function is not similar across sites.

4.3 Power Density Estimation Error

In addition to the goodness-of-fit indicators (R^2, RMSE), the reliability of each model is assessed in terms of estimating the wind energy density using the mean absolute error between the measured wind energy density and that expected from the models tested. Table 4 presents the annual wind power density calculated using the data measured and the annual wind power density estimated using the five

Table 2. Annual shape and scale parameters of five (05) distributions for the ten study sites using the maximum likelihood method

Station	Weibull		Hybrid Weibull		Gamma		Rayleigh		Inverse Gaussian	
	k	c	k	c	k	c	k	c	k	c
Bobo	2.58	3.76	2.58	3.76	6.34	0.52	2	2.42	10.60	3.02
Bogandé	1.93	3.18	1.93	3.18	3.42	0.81	2	2.06	4.07	2.32
Boromo*	2.25	2.55	2.25	2.55	4.87	0.46	2	1.23	0.72	1.09
Dédougou	1.83	3.27	1.83	3.27	3.43	0.84	2	2.21	5.03	2.57
Dori*	2.08	2.28	2.08	2.28	4.21	0.47	2	1.16	0.77	1.06
Fada	1.89	2.48	1.89	2.48	3.43	0.64	2	1.57	2.45	1.71
Gaoua	1.90	3.07	1.90	3.07	3.36	0.80	2	1.76	1.70	1.75
Ouaga	1.97	3.15	1.97	3.15	3.62	0.76	2	2.14	6.38	2.56
Ouahi.	1.79	2.97	1.79	2.97	3.09	0.84	2	1.93	2.99	2.10
Po	1.93	2.51	1.93	2.51	3.68	0.60	2	1.52	1.96	1.60

Table 3. Root mean square error, coefficient of determination and rank between theoretical and observed distributions for the ten study sites

Site	Distributions	RMSE	R²	Rank
Bobo	Weibull	0.0418	0.8214	1
	Hybrid Weibull	0.0419	0.8204	2
	Gamma	0.0419	0.7753	3
	Rayleigh	0.1111	-0.2656	5
	Inverse Gaussian	0.0834	0.2883	4
Bogandé	Weibull	0.0281	0.8912	3
	Hybrid Weibull	0.0267	0.9012	1
	Gamma	0.0267	0.7483	2
	Rayleigh	0.0813	0.0880	5
	Inverse Gaussian	0.0648	0.4202	4
Boromo	Weibull	0.1363	0.1293	4
	Hybrid Weibull	0.1329	0.1716	3
	Gamma	0.1329	-0.0914	5
	Rayleigh	0.1130	0.4012	2
	Inverse Gaussian	0.0691	0.7760	1
Dédougou	Weibull	0.0236	0.9281	1
	Hybrid Weibull	0.0251	0.9189	2
	Gamma	0.0251	0.8531	3
	Rayleigh	0.0806	0.1606	5
	Inverse Gaussian	0.0667	0.4259	4
Dori	Weibull	0.1061	0.3997	5
	Hybrid Weibull	0.1035	0.4289	3
	Gamma	0.1035	0.2137	4
	Rayleigh	0.0878	0.5888	2
	Inverse Gaussian	0.0696	0.7416	1
Fada	Weibull	0.0229	0.9516	3
	Hybrid Weibull	0.0221	0.9550	1
	Gamma	0.0221	0.8650	2
	Rayleigh	0.0811	0.3943	5
	Inverse Gaussian	0.0633	0.6311	4
Gaoua	Weibull	0.0804	0.3996	4
	Hybrid Weibull	0.0785	0.4280	2
	Gamma	0.0785	0.2078	3
	Rayleigh	0.0992	0.0851	5
	Inverse Gaussian	0.0569	0.6993	1
Ouaga	Weibull	0.0232	0.9394	1
	Hybrid Weibull	0.0248	0.9309	2
	Gamma	0.0248	0.8799	3
	Rayleigh	0.0850	0.1892	5
	Inverse Gaussian	0.0633	0.5510	4
Ouahigouya	Weibull	0.0350	0.8629	1
	Hybrid Weibull	0.0352	0.8612	2
	Gamma	0.0352	0.7946	3
	Rayleigh	0.1287	-0.8587	5
	Inverse Gaussian	0.0951	-0.0150	4
Po	Weibull	0.0472	0.8101	3
	Hybrid Weibull	0.0460	0.8200	1
	Gamma	0.0460	0.6568	2
	Rayleigh	0.0908	0.2979	5
	Inverse Gaussian	0.0710	0.5702	4
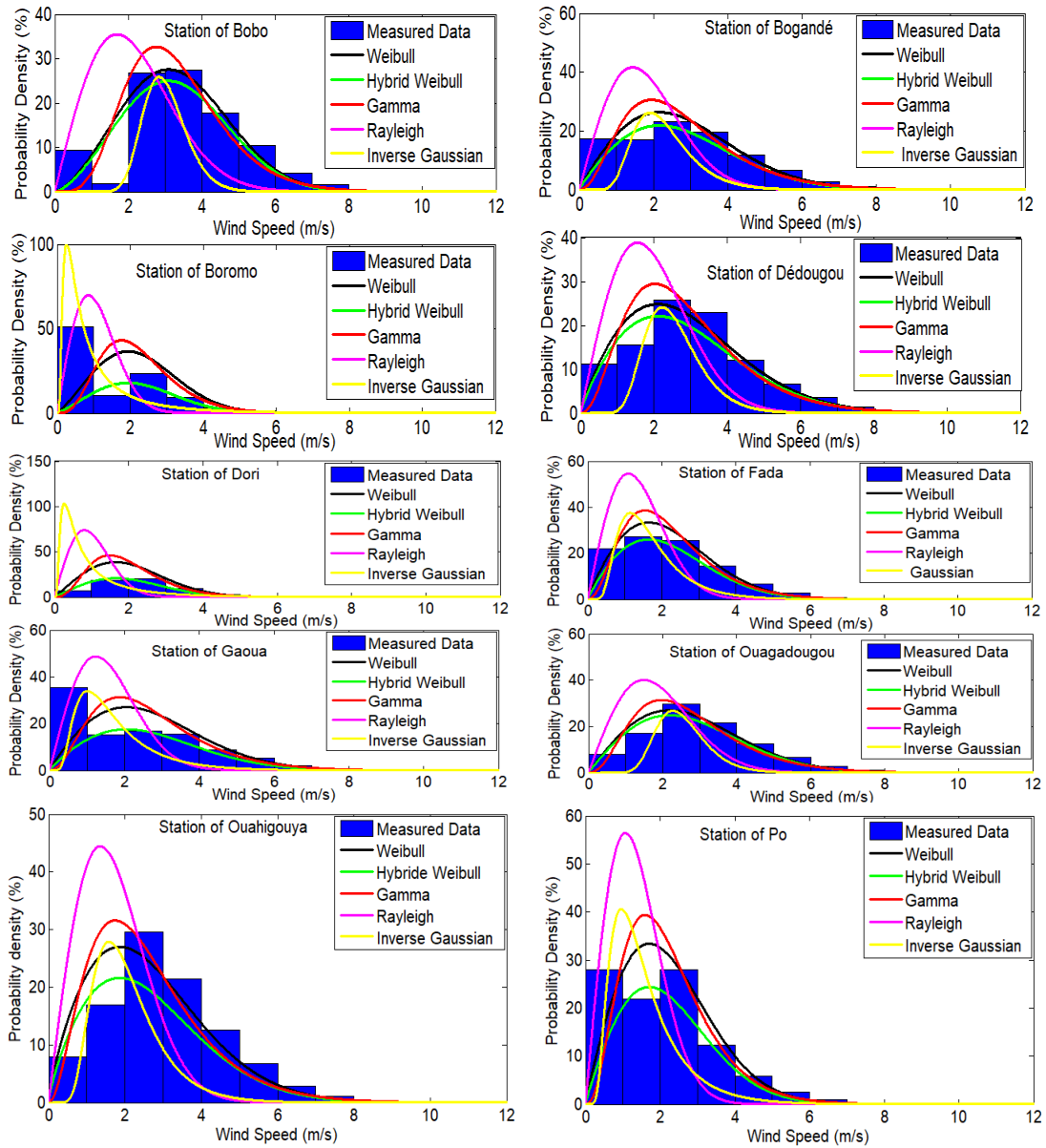
Fig. 2. Observed and theoretical distribution of the different models at the study site

distributions studied. The mean absolute error in estimating the power density using the distribution functions considered is also shown in the Table. According to the absolute mean error, the hybrid Weibull distribution is the most accurate and efficient one to model the wind speed frequency distribution for the sites of Bobo ($\varepsilon_d = 1.38\%$), Bogandé ($\varepsilon_d = 6.16\%$), Boromo ($\varepsilon_d = 4.70\%$), Dori ($\varepsilon_d = 6.74\%$), Fada ($\varepsilon_d = 9.85\%$), Gaoua ($\varepsilon_d = 11.29\%$), and Po ($\varepsilon_d = 11.29\%$). For the cities of Ouaga ($\varepsilon_d = 0.57\%$) and Ouahigouya ($\varepsilon_d = 7.50\%$), the most appropriate corresponding distribution is that of Rayleigh. Finally, for the town of Dédougou ($\varepsilon_d = 12.50\%$), the minimum error is observed with the Weibull distribution. We then conclude that the Hybrid Weibull distribution function has a minimum error in estimating the annual wind energy density at almost all study sites.
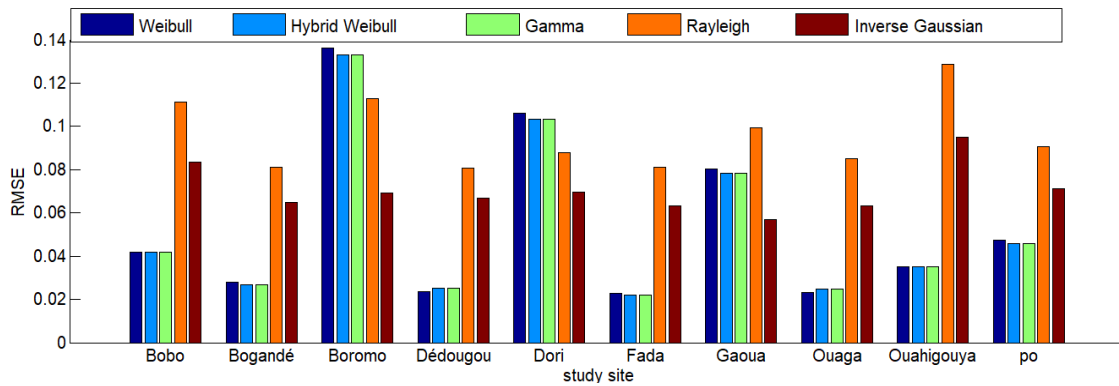
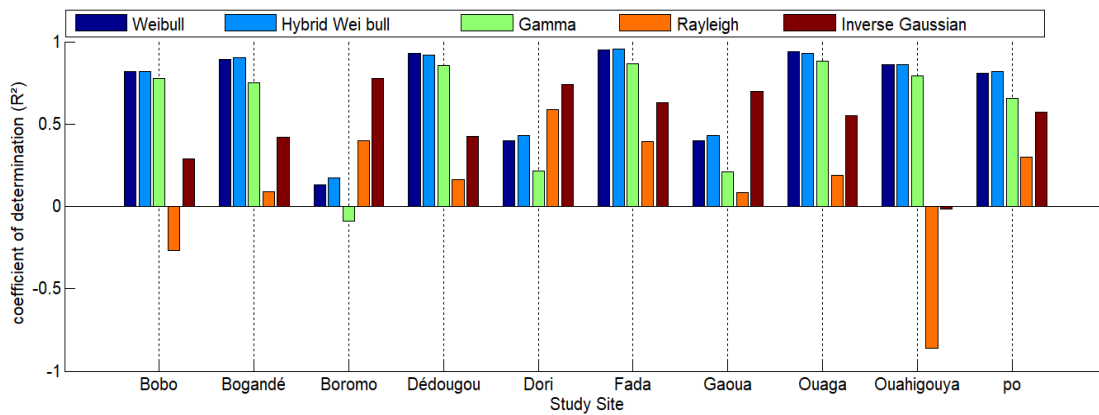**Fig. 3. Comparison of distribution models in terms of RMSE**

Fig. 4. Comparison of distribution models in terms of R^2

Table 4. Mean absolute error in power density estimation using the five distributions studied

Station	Distributions	$\overline{P_{th}}(W/m^2)$	$\overline{P_{ob}}(W/m^2)$	$\varepsilon_d(\%)$	Rank
Bobo	Weibull	35.88	32.99	8.76	3
	Hybrid Weibull	32.53		1.38	1
	Gamma	11.69		64.56	5
	Rayleigh	34.10		3.37	2
	Inverse Gaussian	17.21		47.82	4
Bogandé	Weibull	27.77	24.51	13.31	3
	Hybrid Weibull	23.00		6.16	1
	Gamma	7.21		70.58	5
	Rayleigh	27.21		11.00	2
	Inverse Gaussian	7.80		68.16	4
Boromo	Weibull	12.33	6.29	96.00	5
	Hybrid Weibull	5.99		4.70	1
	Gamma	1.53		75.61	2
	Rayleigh	11.94		89.78	4
	Inverse Gaussian	0.80		87.14	3
Dédougou	Weibull	32.19	36.79	12.50	1
	Hybrid Weibull	28.57		22.32	3
	Gamma	8.90		75.79	5
	Rayleigh	30.56		16.92	2

Station	Distributions	$\overline{P_{th}}(W/m^2)$	$\overline{P_{ob}}(W/m^2)$	$\varepsilon_d(\%)$	Rank
	Inverse Gaussian	10.60		71.16	4
Dori	Weibull	9.46	5.38	75.82	3
	Hybrid Weibull	5.01		6.74	1
	Gamma	1.28		76.07	4
	Rayleigh	8.83		64.20	2
	Inverse Gaussian	0.74		86.17	5
Fada	Weibull	13.49	11.67	15.59	3
	Hybrid Weibull	10.52		9.85	1
	Gamma	3.19		72.66	4
	Rayleigh	13.51		15.74	2
	Inverse Gaussian	3.12		73.24	5
Gaoua	Weibull	25.45	18.52	37.37	3
	Hybrid Weibull	16.43		11.29	1
	Gamma	4.49		75.72	4
	Rayleigh	25.12		35.63	2
	Inverse Gaussian	3.34		81.91	5
Ouaga	Weibull	26.39	25.93	1.75	2
	Hybrid Weibull	24.30		6.28	3
	Gamma	8.08		68.82	5
	Rayleigh	25.78		0.57	1
	Inverse Gaussian	10.48		59.57	4
Ouahigouya	Weibull	24.81	22.16	11.95	3
	Hybrid Weibull	19.81		10.59	2
	Gamma	5.93		73.24	4
	Rayleigh	23.82		7.50	1
	Inverse Gaussian	5.78		73.88	5
Po	Weibull	13.66	11.09	23.09	3
	Hybrid Weibull	9.84		11.29	1
	Gamma	2.89		73.89	4
	Rayleigh	13.20		19.00	2
	Inverse Gaussian	2.56		76.93	5

5. CONCLUSION

Under- or overestimation of the wind energy potential of a given site negatively influences the economic rate of return of wind energy projects. This is caused by the uncertainty associated with the mathematical modelling of the wind speed frequency distribution. Indeed, it is common practice to model the wind speed distribution systematically using the Weibull function. However, this function is not always suitable for all wind regimes. Therefore, in this study, wind speed data in three-hourly time series format over the period from January 2006 to December 2016 at ten sites in Burkina Faso are statistically analyzed and fitted by five candidate distribution functions (Weibull, Hybrid Weibull, Rayleigh, Gamma and inverse Gaussian).

- Graphically, the Rayleigh distribution curve indicates a poor fit with the measurements at most sites. In order to make a scientific decision, two suitability tests (coefficient of determination and RMSE) were used to select the effective distribution that best fits the histogram of observations.

- Use of the maximum likelihood method to compare the five distributions shows that the Weibull distribution is the most suitable one for the sites of Bobo (R^2 =0.82), Dedougou (R^2 =0.92), Ouaga (R^2 =0.93) and Ouahigouya (R^2 =0.86). On the other hand, for the sites of Bogandé (R^2 =0.90), Fada (R^2 =0.95), and Po (R^2 =0.82), the hybrid weibull distribution is the most appropriate. As for the inverse Gaussian distribution, it is most suited to the sites of Boromo (R^2 =0.77), Dori (R^2 =0.74) and Gaoua (R^2 =0.69).

- In addition, the calculation of the mean absolute error between the annual wind power density estimated using theoretical

distributions and the wind power density calculated using measurements was performed. It was found that the hybrid Weibull distribution has a minimum absolute error in the estimation of power density at the sites of Bobo ($\varepsilon_d = 1.38\%$), Bogandé ($\varepsilon_d = 6.16\%$), Boromo ($\varepsilon_d = 4.70\%$), Dori ($\varepsilon_d = 6.74\%$), Fada ($\varepsilon_d = 9.85\%$), Gaoua ($\varepsilon_d = 11.29\%$), Po ($\varepsilon_d = 11.29\%$). For the cities of Ouaga ($\varepsilon_d = 0.57\%$) and Ouahigouya ($\varepsilon_d = 7.50\%$), the most appropriate corresponding distribution is that of Rayleigh. Finally, for the town of Dédougou ($\varepsilon_d = 12.50\%$), the minimum error is observed with the Weibull distribution.

We then conclude that the Hybrid Weibull distribution function has a minimum error in estimating the annual wind power density at almost all study sites. Thus Hybrid Weibull distribution function is the mostsuitable distribution that models wind speed at almost all study sites.

Due to the shortage of wind turbines in the regions of Burkina, the results of this study can provide useful information for the development of wind energy.

ACKNOWLEDGEMENTS

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Gabra S, Miles J, Scott SA. Techno-economic analysis of stand-alone wind micro-grids, compared with PV and diesel in Africa. Renewable Energy. 2019;43: 1928-1938.
   DOI: 10.1016/j.renene.2019.05.119
2. Alavi O, Mohammadi K, Mostafaeipour A. Evaluating the suitability of wind speed probability distribution models: A case of study of east and southeast parts of Iran. Energy Conversion and Management. 2016;119:101-108.
   DOI: 10.1016/j.enconman.2016.04.039
3. Mohammadi K, Alavi O, Mostafaeipour A, Goudarzi N, Jalilvand M. Assessing different parameters estimation methods of Weibull distribution to compute wind power density. Energy Conversion and Management. 2016;108:322-335.
   DOI: 10.1016/j.enconman.2015.11.015
4. Statistical analysis of wind data and assessment of wind potential in Burkina Faso. IJESRT. 2019;8:2.
5. Olatayo KI, Wichers JH, Stoker PW. Energy and economic performance of small wind energy systems under different climatic conditions of South Africa. Renewable and Sustainable Energy Reviews. 2018;98:376-392.
   DOI: 10.1016/j.rser.2018.09.037
6. Mert İ, Karakuş C. A statistical analysis of wind speed data using Burr, generalized gamma, and Weibull distributions in Antakya, Turkey. Turkish Journal of Electrical Engineering & Computer Sciences. 2015;23:1571-1586.
   DOI: 10.3906/elk-1402-66
7. Llera A, Beckmann CF. Estimating an Inverse Gamma distribution; 17.
8. Lo Brano V, Orioli A, Ciulla G, Culotta S. Quality of wind speed fitting distributions for the urban area of Palermo, Italy. Renewable Energy. 2011;36(3):1026-1039.
   DOI: 10.1016/j.renene.2010.09.009
9. Alayat M, Kassem Y, Çamur H. Assessment of wind energy potential as a power generation source: A case study of eight selected locations in Northern Cyprus. Energies. 2018;11(10):2697.
   DOI: 10.3390/en11102697
10. Kang D, Ko K, Huh J. Determination of extreme wind values using the Gumbel distribution. Energy. 2015;86:51-58.
    DOI: 10.1016/j.energy.2015.03.126
11. Jung C, Schindler D. Wind speed distribution selection – A review of recent development and progress. Renewable and Sustainable Energy Reviews. 2019; 114:109290.
    DOI: 10.1016/j.rser.2019.109290
12. Ouarda TBMJ, et al. Probability distributions of wind speed in the UAE. Energy Conversion and Management. 2015;93:414-434.
    DOI: 10.1016/j.enconman.2015.01.036

13. *Septieme Reunion du Sous-Comite Ouest and Centre Africain de Correlation des Sols pour la Mise en Valeur des Terres, Ouagadougou*, 1985. Food & Agriculture Org.; 1986.

14. Tizgui I, El Guezar F, Bouzahir H, Benaid B. Wind speed distribution modeling for wind power estimation: Case of Agadir in Morocco. Wind Engineering. 2019;43(2): 190-200.
    DOI: 10.1177/0309524X18780391.

15. Usta I. An innovative estimation method regarding Weibull parameters for wind energy applications. Energy. 2016;106: 301-314.
    DOI: 10.1016/j.energy.2016.03.068

16. Kidmo DK, Danwe R, Doka SY, Djongyang N. Statistical analysis of wind speed distribution based on six Weibull; 21.

17. Tizgui I, El Guezar F, Bouzahir H, Benaid B. Comparison of methods in estimating Weibull parameters for wind energy applications. International Journal of Energy Sector Management. 2017;11(4): 650-663.
    DOI: 10.1108/IJESM-06-2017-0002

18. Katinas V, Gecevicius G, Marciukaitis M. An investigation of wind power density distribution at location with low and high wind speeds using statistical model. Applied Energy. 2018;218:442-451.
    DOI: 10.1016/j.apenergy.2018.02.163

19. Akdağ SA, Dinler A. A new method to estimate Weibull parameters for wind energy applications. Energy Conversion and Management. .2009;50(7):1761-1766.
    DOI: 10.1016/j.enconman.2009.03.020

20. Weibull Parameters_download.pdf

21. Justus CG, Mikhail A. Height variation of wind speed and wind distributions statistics. Geophysical Research Letters. 1976;3(5):261-264.
    DOI: 10.1029/GL003i005p00261

22. Sarkar A, Singh S, Mitra D. Wind climate modeling using Weibull and extreme value distribution. International Journal of Engineering. 2011;3(5):7.

23. Ihaddadene R, Ihaddadene N, Mostefaoui M. Estimation of monthly wind speed distribution basing on hybrid Weibull distribution. World Journal of Engineering. 2016;13(6):509-515.
    DOI: 10.1108/WJE-09-2016-0084

24. Lo Brano, et al. Quality of wind speed fitting distributions for th.pdf; 2011.

25. Available:bardsley1980.pdf

26. Dookie I, Rocke S, Singh A, Ramlal CJ. Evaluating wind speed probability distribution models with a novel goodness of fit metric: A Trinidad and Tobago case study. International Journal of Energy and Environmental Engineering. 2018;9(3): 323-339.
    DOI: 10.1007/s40095-018-0271-y

27. Samal RK, Tripathy M. Estimating wind speed probability distribution based on measured data at Burla in Odisha, India. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects. 2019;41(8):918-930.
    DOI: 10.1080/15567036.2018.1521888

28. Abdulkarim A, Abdelkader SM, Morrow DJ, Falade AJ, Adediran YA. Statistical analysis of wind speed for electrical power generation in some selected sites in northern Nigeria. Nigerian Journal of Technology. 2018;36(4):1249.
    DOI: 10.4314/njt.v36i4.35

29. Available:2016-14.pdf

30. Available:philippopoulos2012.pdf

31. Forbes CS, Éd. Statistical distributions: Catherine Forbes [et al.].4th Ed. Hoboken, N.J: Wiley; 2011.

32. Zohbi GA, Hendrick P. Evaluation du potentiel d'énergie éolienne au Liban; 14.

33. Celik AN. A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey. Renewable Energy. 2004;29(4): 593-604.
    DOI: 10.1016/j.renene.2003.07.002

34. Ahmed SA, Mahammed HO. A statistical analysis of wind power density based on the Weibull and Ralyeigh models of "Penjwen Region" Sulaimani/ Iraq. 2012; 6(2):7.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/59487