Energy and Power Engineering, 2025, 17(10), 309-323
https://www.scirp.org/journal/epe
ISSN Online: 1947-3818
ISSN Print: 1949-243X

Monthly Reference Frequency Distributions of Hourly Relative Wind Speed Data from Burundian Stations

Mathias Bashahu*, **Janvier Ntirandekura**

Department of Physics and Technology, Institute of Applied Pedagogy, University of Burundi, Bujumbura, Burundi
Email: *bashahuma@yahoo.fr

How to cite this paper: Bashahu, M. and Ntirandekura, J. (2025) Monthly Reference Frequency Distributions of Hourly Relative Wind Speed Data from Burundian Stations. *Energy and Power Engineering*, **17**, 309-323.
https://doi.org/10.4236/epe.2025.1710017

Received: July 15, 2025
Accepted: October 14, 2025
Published: October 17, 2025

Abstract

In previous research works on one hand, terrain, meteorological and wind speed data from various Burundian sites have been processed in order to bring efficient tools for the planning and installation of wind energy conversion systems (WECSs) at those sites. On another hand, sunshine duration and global solar radiation data, respectively from different Burundian sites, have been used to implement monthly reference distributions of data at those sites for the next two dimensionless random variables: the daily relative sunshine duration and the daily clearness index. Those distributions were deemed to be important inputs in projects for setting solar energy conversion systems (SECSs) at the relevant sites. The present study comes out as a continuation of the afore-said works and it intends to further contribute to bringing useful help in projects for setting WECSs at selected Burundian localities. Using a 4-years period's hourly wind speed, v values from four Burundian stations as primary data, the specific objective of the study is three-fold. Firstly, to set up for each station frequency distributions, means and variances of the 12 monthly samples and the 48 sub-samples of hourly relative wind speed, v_r (another dimensionless random variable) data extracted from the primary wind speed values. Secondly, to use a suitable statistical test in order to ascertain whether or not, for each station, monthly samples of relative wind speed data for which mean values, $\bar{v}_r$ fall within a same interval with the width equal to 0.05, have also the same variance, originate from the same population and thus possess the same statistical distribution. Thirdly, to build such a distribution referred to as a monthly reference frequency distribution of relative wind speed data, for any of the identified mean relative wind speed ($\bar{v}_r$) intervals of interest and for each station. All the three objectives have been attained. Especially, the use of the one-way ANOVA conditions and the F-test has led to accept the null hypothesis for the following numbers of $\overline{v_r}$ intervals of interest: 3 out of 3 for Bujumbura, 2 out of 3 for Gitega, 2 out of 2 for Kirundo, 4 out of 4 for Musasa, and thus 11 out of 12 for the four stations. With the view to building the monthly reference frequency distribution for each identified $\overline{v}_r$ interval of interest, the v_r data of the relevant samples have been put together in a group. Then, the absolute, relative and cumulative relative frequency distributions of those data, respectively, have been inferred, and curves of the two last kinds of distribution have been plotted. Altogether, the obtained 11 monthly reference frequency distributions of hourly v_r data should be used as inputs into projects for setting WECSs at the relevant stations.

Keywords

ANOVA (One-Way), F-Test (One-Sided), Hourly Relative Wind Speed Data, Monthly Reference Frequency Distributions

1. Introduction

Only a low rate of the population in many developing countries has access to electricity, gas and even charcoal. On the contrary, a high percentage of households have recourse to firewood and wood scraps to satisfy their primary energy needs like food cooking, house heating and lighting. Electricity production in those countries is deficient in comparison with an increasing energy demand due to high rates of population growth and industrialization efforts. That is why attempt is made in those regions and even abroad to supplement the use of traditional energy sources with free, clean and inexhaustible ones such as solar and wind energies. Within this last energy source (wind energy), a huge number of research works have been (and are still being) implemented worldwide on the selection of suitable (onshore and offshore) sites to settle wind energy conversion systems (WECSs, such as lonely wind turbines (WTs) or wind farms) and on the decision about a system's parameters. All of those works require a good knowledge of different properties of a given site.

These properties are found out using firstly meteorological data (e.g. air density, temperature, humidity and pressure) [1] [2] and terrain (e.g. topography, obstacles and surface roughness) [3] [4]. The use of such data and boundary conditions through different GIS-based techniques [5] [6] allows one to capture the spatiotemporal distribution of wind resource in a given area [7] [8]. Secondly, the previous properties should be found out through a wind resource analysis. Experimental frequency distributions of wind speed (intensity, v and direction, θ) data may be drawn from *in situ* measurements over a given period, T (one or several years in general). Then, the use of a relevant probability density function (PDF) fitting very well the observed frequency distribution of wind speed (v) and/or direction data enables one to estimate different traditional wind energy potential characteristics of the site. These characteristics are notably [9]: average wind direction, wind speed distribution's mean and variance, wind speed shear exponent (when wind speed measurements have been performed at two or more heights above ground level (AGL)), most probable wind speed, optimum wind speed, mean wind power density, mean wind energy density over the period T, wind energy pattern factor.

The 2-parameter Weibull PDF is the most frequently used PDF in the above-mentioned fitting procedure [2] [9]-[13]. Its scale and modulus (or shape) parameters (c and k, respectively) should be estimated via one of the following techniques, notably [1] [11] [12] [14]-[17]: graphical method, maximum likelihood method, modified maximum likelihood method, standard deviation method, method of moments, generalized method of moments, power density method, equivalent energy method, Justus' empirical method, Lysen's empirical method, Bayes' method and non-linear Bernard's median rank regression method. Nevertheless, several other commonly tested PDFs include the following among others [18]-[21]: 3-parameter Weibull, Rayleigh, normal, exponential, gamma, lognormal, Burr, inverse Gaussian, generalized extreme value.

Once suitable sites for setting WECSs are selected on the basis of their properties and wind energy potential, a decision should be taken about the kinds of WTs to settle at one of those sites. This is made in terms of next WT's parameters and performance characteristics, notably [2]-[4] [9] [22]: hub height, blade shape and size, direction; cut-in, rated, and cut-out wind speeds; rated power, energy output (e.g. annual energy production, AEP), capacity factor (CF or C_f) and cost of electricity (COE) per unit kWh.

In Burundi, the use of renewable energy sources in order to supplement traditional ones is clearly stated in the current government's working agenda [23]. With respect to wind energy, there is currently no installed onshore or offshore wind park in the country. Only some single low-rated power WTs have been settled by private persons in different areas for useful purposes in dwellings and stock-farms. Nevertheless, one should mention, on one hand, the following among available research papers about suitable site selection and WECSs installation decision in Burundi. In a first paper, it has been evidenced that beta probability density functions (β-PDFs) fit very well related observed monthly relative frequency distributions of hourly relative wind speed (v_r) data. Furthermore, contrary to 2-parameter Weibull PDFs, those β-PDFs represent accurately the probabilities of observing zero (or very low) wind speed [24]. A second study has compared the effectiveness of different PDFs in fitting experimental frequency distributions of wind speed data [19]. In a third work, the wind energy potential has been assessed for two Burundian stations, and the economical feasibility, together with benefit of WECSs set at those stations has been evidenced for supplementing traditional electricity sources [9]. Results from a fourth study indicate among other things, that the western part of Burundi, including the Lake Tanganyika, is the optimal area to settle WECSs [25].

On another hand, as important inputs in projects of solar energy conversion systems (SECSs), monthly reference frequency distributions of data from different

Burundian stations have been established for two dimensionless random variables, *i.e.*: the daily relative sunshine duration (s_r) [26] and the daily clearness index (k_t) [27]. In the continuation of those last two works and in terms of an additional contribution to the knowledge of sites properties for WECSs setting, the present study makes use of a 4-year period hourly wind speed (v) data from four Burundian stations and its specific objective is three-fold. Firstly, to set up for each station, frequency distributions of monthly sub-samples and samples of the relative wind speed (v_r , another dimensionless random variable) data extracted from the above-mentioned hourly wind speed (v) ones, together with means and variances of those sub-samples' distributions. Secondly, to ascertain whether or not, for each station, monthly samples of v_r data having the same mean relative wind speed, $\overline{v}_r$ have also the same variance and thus the same statistical distribution. Thirdly, to build such a distribution, referred to as a monthly reference frequency distributions of v_r data, for any of the identified $\overline{v}_r$ intervals of interest and for any of the four stations.

2. Materials

2.1. Selected Stations and Primary Data

The primary data used in this work are hourly wind speed (v) values measured (in m/s) at the height of 2 m above ground level (AGL). They have been kindly provided by the Geographical Institute of Burundi (IGEBU) situated in Gitega and refer to the twelve-hourly intervals of the daily period from 6:00 to 18:00 local time, and to the 4-years period from 1991 to 1994. They also refer to the next four stations: Bujumbura ($L = 29.35^\circ\text{E}$; $\varphi = 3.38^\circ\text{S}$; $z = 800\text{ m}$), Gitega-Zege ($L = 29.92^\circ\text{E}$; $\varphi = 3.40^\circ\text{S}$; $z = 1663\text{ m}$), Kirundo ($L = 30.12^\circ\text{E}$; $\varphi = 3.58^\circ\text{S}$; $z = 1449\text{ m}$), and Musasa ($L = 30.10$; $\varphi = 4.00^\circ\text{S}$; $z = 1260\text{ m}$). L, φ and z are the station's longitude, latitude and elevation above sea level (ASL), respectively. Those data and stations have been selected owing to the next three major criteria: i) the consideration of different geographical regions of the country; ii) the matching of a reasonable long-term period records with a minimum of discontinuities for the relevant data; and iii) wind speed measurements at the same height AGL for all the stations.

In connection with this last criterion, it is worthy mentioning that wind speed records refer only to 2 m AGL for all Burundian meteorological stations, except for Bujumbura and Muyinga where short-period measurements have been performed at 2 m and 12 m AGL [9]. This statement does unfortunately not meet the WMO recommendation of using wind speed records at a reference height of 10 m AGL in wind resource analysis [14] [28].

2.2. Experimental Data and Monthly Frequency Distributions of the Random Variable v_r

For any of the four stations, the hourly wind speed (v) data of the 4-year period have been arranged into twelve monthly samples. A maximum v value ($v_{\max}$)

has been identified in each sample. Then, experimental hourly relative wind speed (v_r) data relating to that sample have been derived using the next dimensionless ratio:

$$v_r = \frac{v}{v_{\max}}; \quad 0 \le v_r \le 1 \tag{1}$$

At their turn, the v_r data of each sample have been arranged in four sub-samples relating to the four years of the above mentioned 4-year period. The next step has been to sort elements of each sub-sample in ten intervals of the same width equal to 0.10, and then to infer their absolute and relative frequency distributions, together with mean and variance. Results of that step have also allowed us to deduce the mean $\bar{v}_r$ of each sample. The total number of v_r data and mean $\bar{v}_r$ for any of the 48 samples are shown in **Table 1**. A discussion on the main features of the relative frequency distributions of those samples is out of the scope of this study, since it has been made in a previous work [24]. From results of **Table 1**, it should be noticed that, owing to discontinuities in wind speed records during the 4-year period, the total number of hourly v_r values actually used in this study is 66,489 instead of the expected total one which is 70,128.

Table 1. Total number (N) of hourly relative wind speed data and mean for any of the 48 monthly relative frequency distributions.

Stations and parameters→	Bujumbura		Gitega		Kirundo		Musasa	
	N	$\bar{v}_r$	N	$\bar{v}_r$	N	$\bar{v}_r$	N	$\bar{v}_r$
Months↓								
January	1440	0.2981	1473	0.2117	1336	0.1614	1476	0.1848
February	1231	0.1729	1341	0.1510	1356	0.2169	1350	0.2130
March	1433	0.2256	1473	0.1557	1115	0.1986	1486	0.1773
April	1437	0.2218	1375	0.2048	1240	0.2056	1440	0.2215
May	1481	0.2174	1488	0.2340	1368	0.1630	1488	0.2243
June	1428	0.2813	1440	0.2296	1403	0.1612	1435	0.2878
July	1472	0.3586	1487	0.2957	1488	0.1768	1486	0.3516
August	1456	0.3032	1486	0.3170	1485	0.2273	1482	0.3719
September	1187	0.2587	1435	0.2443	1434	0.2478	1438	0.4263
October	1342	0.2995	1356	0.2584	1336	0.1393	1476	0.2753
November	1077	0.2628	1140	0.2392	1080	0.1979	1428	0.2377
December	1239	0.1828	1468	0.2006	1195	0.1594	1483	0.2766
Total	**16,223**	---	**16,962**	---	**15,836**	---	**17,468**	---

3. Testing Procedure of the Null Hypothesis

For each station and the 4-years period of this study, the mean values $\bar{v}_r$ of the twelve samples have been firstly computed as described in Section 2.2. Then, they have been arranged into the relevant intervals of $\bar{v}_r$ among the twenty ones available in the range from 0.00 to 1.00, with the same width equal to 0.05, *i.e.*: [0.00;0.05[, [0.05; 0.10[, [0.10; 0.15[, [0.15; 0.20[, …, [0.95; 1.0]. The previous width (0.05) has been chosen too weak to allow the assertion according which samples with $\bar{v}_r$ values falling within one of the above-defined intervals, have the same mean relative wind speed. The next task has been to check whether or not those samples have also the same variance. The one-way analysis of variance (ANOVA) has been used for that purpose. As a matter of fact, the next current (or assumed) conditions take place in this study: i) one dependent continuous variable, v_r expressed in interval or ratio scale (and not in nominal or ordinal scales) of measurement; ii) one categorical independent variable or factor, I (with two or more categories, *i.e.*: samples here); iii) samples are randomly drawn from the population and observations within a group (sample) are independent of each other; iv) population values are normally distributed. Within those conditions, let us consider particularly the statement according which the populations from which the samples come are normally distributed. This allows us to ascertain that analyzing the differences between the means of two or more independent samples having the same variance (homogeneity of variances), is equivalent to test the following null hypothesis, H_0: those independent samples originate from the same population and have the same statistical distribution [29]-[31]. Moreover, two variances are estimated in ANOVA. Firstly, the within group variability (or error variance) which is based on random differences present in samples. Secondly, the between group variability (or effect variance) which is the result of our treatment. The test statistic is the quotient of those two variances, commonly referred to as the F-test, which is a parametric test.

For computations in the testing procedure, within the twelve monthly samples of v_r data for each station, we have called I ($I \geq 2$) the number of samples with the same mean $\bar{v}_r$ and i ($i = 1, 2, \cdots, I$) the code number of those samples. The i^{th} sample has been divided into J_i sub-samples, where each sub-sample is relating to one among the four years as stated in Section 2.2. We have also called k_{ij} the number of elements of the j^{th} sub-sample ($j = 1, \cdots, J_i$) in the i^{th} sample, and s_{ij}^2 the variance of that sub-sample. With values of the continuous dimensionless random variable v_r (defined by Equation (1)) ranging from 0 to 1, the null hypothesis H_0 was accepted if:

$$F_0 \leq F_c = F(\alpha, n_1, n_2) \tag{2}$$

where [32]:

$$n_1 = I - 1 \tag{3.1}$$

$$n_2 = \sum_{i=1}^{I} (J_i - 1) \tag{3.2}$$

$$\eta_i = \left[\sum_{j=1}^{J_i} (k_{ij} - 1) \ln s_{ij}^2\right] \Big/ \left[\sum_{j=1}^{J_i} (k_{ij} - 1)\right] \tag{4}$$

$$\eta = \left[\sum_{i,j}^{I,J_i} (k_{ij} - 1)\eta_i\right] \Big/ \sum_{i,j}^{I,J_i} (k_{ij} - 1) \tag{5}$$

$$F_0 = \left[n_2 \sum_{i,j}^{I,J_i} (k_{ij} - 1)(\eta_i - \eta)^2\right] \Big/ \left[n_1 \sum_{i,j}^{I,J_i} \eta_i (k_{ij} - 1)(\ln s_{ij}^2 - 1)^2\right] \tag{6}$$

and F_c is the critical value of the F-distribution, *i.e.*: the upper α point of that distribution with n_1 and n_2 degrees of freedom. The critical values of the Fisher-Snedecor distribution are given in tables [33] [34] for different values of

n_1 and n_2, together with different values of the significance level α or of the confidence range $\gamma = 1-\alpha$. In the present study, we have taken $\alpha = 0.005$ and therefore $\gamma = 0.995$.

4. Results of the Test, Inferred Monthly Reference Frequency Distributions of v_r Data and Discussion

The results of the F-test applied to hourly relative wind speed data from the four Burundian stations and the 4-year period of this analysis, are reported in Table 2.

Table 2. Results of the F-test for the $\overline{v}_r$ intervals of interest and the four Burundian stations.

$\overline{v}_r$ intervals of interest	Stations → Parameters ↓	Bujumbura	Gitega	Kirundo	Musasa
[0.15;0.20[	n_1	1	1	6	1
	n_2	6	6	19	6
	F_0	1.2371	0.0357	1.5442	0.3632
	$F_c(\alpha = 0.005)$	18.63	18.63	4.5	18.63
	Order N° of months	2; 12	2; 3	1; 3; 5; 6; 7; 11; 12	1; 3
[0.20; 0.25[	n_1	2	6	3	3
	n_2	9	21	12	12
	F_0	0.3541	5.0999	0.2312	0.7631
	$F_c(\alpha = 0.005)$	10.11	4.39	7.23	7.23
	Order N° of months	3; 4; 5	1; 4; 5; 6; 9; 11; 12	2; 4; 8; 9	2; 4; 5; 11
[0.25; 0.30[	n_1	4	1		2
	n_2	14	6		9
	F_0	5.6463	2.1412		2.4966
	$F_c(\alpha = 0.005)$	6.00	18.63		10.11
	Order N° of months	1; 6; 9; 10; 11	7; 10		6; 10; 12
[0.35; 0.40[	n_1				1
	n_2				6
	F_0				0.6048
	$F_c(\alpha = 0.005)$				18.63
	Order N° of months				7; 8

According to those results, within statistical errors, the null hypothesis has been accepted (*i.e.*: $F_0 < F_c$) for the following numbers of $\overline{v}_r$ intervals of interest: 3 out of 3 for Bujumbura, 2 out of 3 for Gitega, 2 out of 2 for Kirundo, 4 out of 4 for Musasa and thus 11 out of 12 for the four stations. Therefore, it is ascertained that for any of those eleven intervals, the independent samples of relative wind speed data with the same mean, $\overline{v}_r$ have also the same variance. Consequently, those samples originate from the same mother population and have the same statistical distribution referred to as a monthly reference frequency distribution of v_r data. Moreover, that distribution depends only on the monthly mean $\overline{v}_r$ and not on the months of the year.

In order to construct such a distribution for any of the eleven $\overline{v}_r$ intervals of interest, the v_r data of the relevant independent monthly samples have been set together in a group, and then arranged in ten classes of the same width equal to 0.10. For any of the four stations and afore-said eleven $\overline{v}_r$ intervals of interest, the resulting relative and cumulative relative frequency distributions of v_r data are presented in Table 3.

Table 3. Results of the monthly reference frequency distributions of v_r data from the four Burundian stations, period 1991-1994. Ten classes of v_r data (with centers x_i) have been defined; f_{ijk} and F_{ijk} are the relative frequencies and cumulative relative frequencies, respectively; i, j and k are the code numbers of the classes of v_r data, the stations and the $\overline{v}_r$ intervals of interest, respectively.

1. Bujumbura

$\overline{v}_r$ intervals of interest	$0.15 \le \overline{v}_r < 0.20$		$0.20 \le \overline{v}_r < 0.25$		$0.25 \le \overline{v}_r < 0.30$	
Parameters → x_i ↓	fi11	Fi11	Fi12	Fi12	fi13	Fi13
0.05	0.3939	0.3939	0.3153	0.3153	0.2390	0.2390
0.15	0.2340	0.6279	0.2209	0.5362	0.1952	0.4342
0.25	0.1765	0.8044	0.1588	0.6950	0.1381	0.0.5723
0.35	0.1206	0.9250	0.1374	0.8324	0.1551	0.7274
0.45	0.0547	0.9797	0.1023	0.9347	0.1362	0.8636
0.55	0.0138	0.9935	0.0457	0.9804	0.0800	0.9436
0.65	0.0036	0.9971	0.0142	0.9946	0.0369	0.9805
0.75	0.0020	0.9991	0.0030	0.9976	0.0137	0.9943
0.85	0.0000	0.9991	0.0007	0.9983	0.0032	0.9974
0.95	0.0008	0.9999	0.0016	0.9999	0.0025	0.9999
Total	0.9999	---	0.9999	---	0.9999	---

2. Gitega

$\overline{v}_r$ intervals of interest →	$0.15 \le \overline{v}_r < 0.20$		$0.25 \le \overline{v}_r < 0.30$	
Parameters → x_i ↓	fi21	Fi21	fi22	Fi22
0.05	0.4971	0.4961	0.2406	0.2406
0.15	0.2050	0.7011	0.1235	0.3641
0.25	0.1535	0.8546	0.1794	0.5435
0.35	0.0917	0.9463	0.1917	0.7352
0.45	0.0327	0.9790	0.1600	0.8952
0.55	0.0142	0.9932	0.0661	0.9613
0.65	0.0036	0.9968	0.0257	0.9870
0.75	0.0018	0.9986	0.0088	0.9958
0.85	0.0007	0.9993	0.0025	0.9983
0.95	0.0007	1.0000	0.0018	1.0001
Total	1.0000	---	1.0001	---

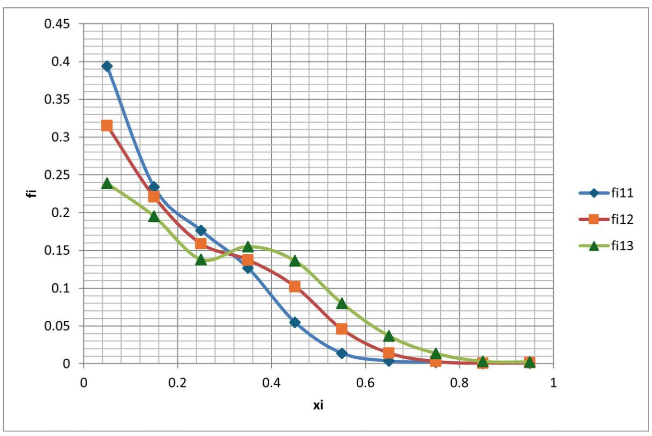
3. Kirundo

$\overline{v}_r$ intervals of interest →	$0.15 \le \overline{v}_r < 0.20$		$0.20 \le \overline{v}_r < 0.25$	
Parameters → x_i ↓	fi31	Fi31	fi32	Fi32
0.05	0.3934	0.3934	0.2986	0.2986
0.15	0.2732	0.6666	0.2286	0.5272
0.25	0.1765	0.8431	0.2025	0.7297
0.35	0.0838	0.9269	0.1396	0.8693
0.45	0.0420	0.9689	0.0687	0.9380
0.55	0.0158	0.9847	0.0359	0.9739
0.65	0.0073	0.9920	0.0145	0.9884
0.75	0.0042	0.9962	0.0056	0.9940
0.85	0.0021	0.9983	0.0038	0.9978
0.95	0.0016	0.9999	0.0020	0.9998
Total	0.9999	---	0.9998	---

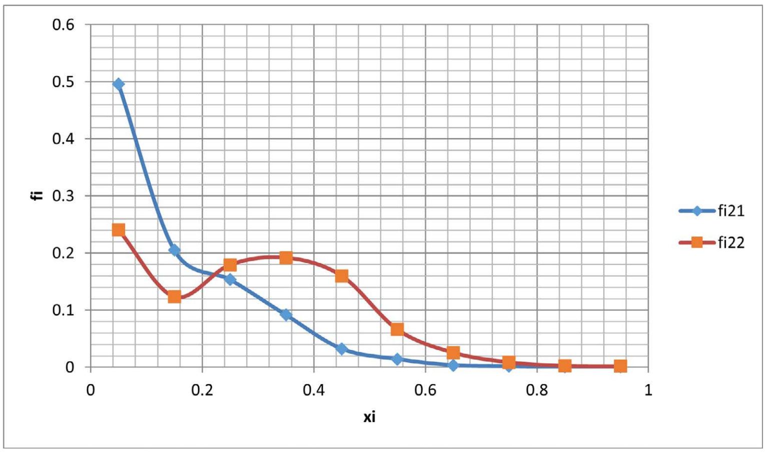
4. Musasa

$\overline{v}_r$ intervals of interest →	$0.15 \le \overline{v}_r < 0.20$		$0.20 \le \overline{v}_r < 0.25$		$0.25 \le \overline{v}_r < 0.30$		$0.35 \le \overline{v}_r < 0.40$	
Parameters → x_i ↓	fi41	Fi41	fi42	Fi42	fi43	Fi43	fi44	Fi44
0.05	0.2633	0.2633	0.2266	0.2266	0.1934	0.1934	0.1004	0.1004
0.15	0.3423	0.6056	0.2352	0.4618	0.1698	0.3629	0.1203	0.2207
0.25	0.2677	0.8733	0.2627	0.7245	0.1839	0.5468	0.1341	0.3548
0.35	0.0922	0.9655	0.1739	0.8984	0.1982	0.7450	0.1769	0.5317
0.45	0.0233	0.9888	0.0701	0.9685	0.1559	0.9009	0.2453	0.7770
0.55	0.0084	0.9972	0.0189	0.9874	0.0683	0.9692	0.1546	0.9316
0.65	0.0003	0.9975	0.0072	0.9946	0.0200	0.9892	0.0451	0.9767
0.75	0.0010	0.9985	0.0026	0.9972	0.059	0.9951	0.0155	0.9922
0.85	0.0003	0.9988	0.0014	0.9986	0.0034	0.9985	0.0047	0.9969
0.95	0.0010	0.9998	0.0014	1.0000	0.0014	0.9999	0.0030	0.9999
Total	0.9998	---	1.0000	---	0.9999	---	0.9999	---

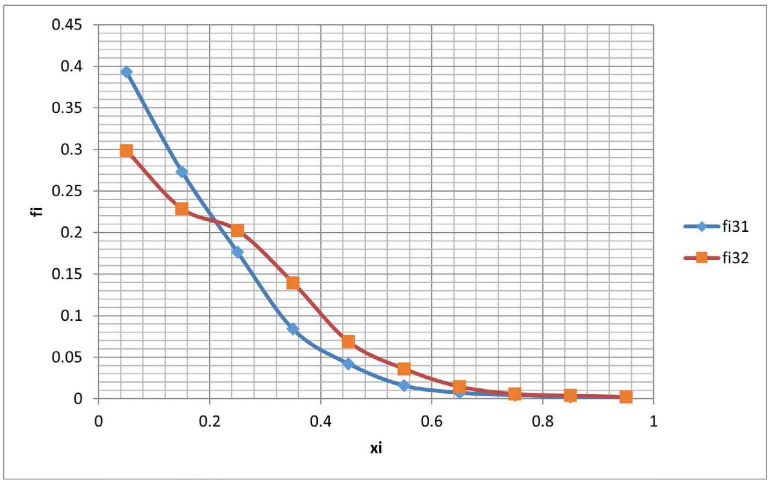
At their turn, curves of the obtained monthly reference relative frequency distributions are presented in **Figure 1**, while integrals of those distributions are exhibited in **Figure 2**.



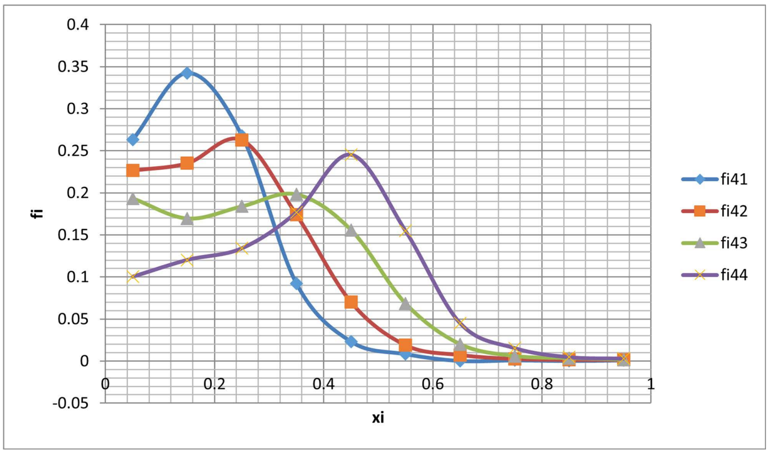
1. Bujumbura



2. Gitega

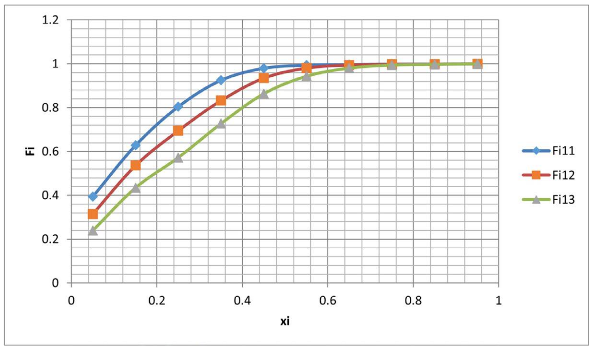


3. Kirundo

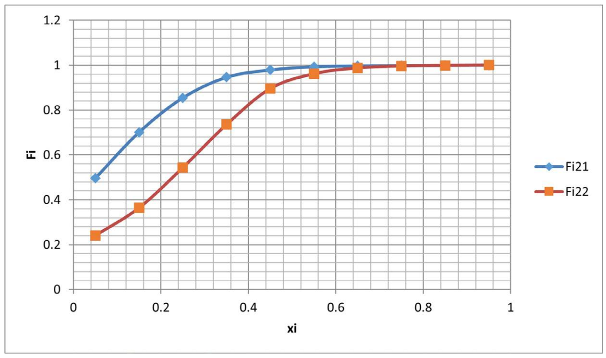


4. Musasa

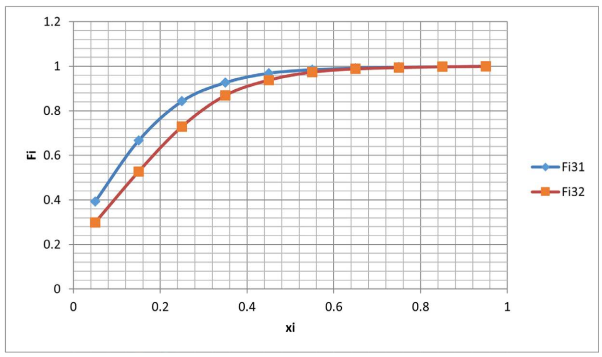
Figure 1. Monthly reference relative frequency (f_i) distributions of v_r data for the eleven $\overline{v}_r$ intervals of interest and the four stations, period 1991-1994. In the symbol f_{ijk}, the subscripts i, j and k are the code numbers of classes of v_r data, stations (1. Bujumbura, 2. Gitega, 3. Kirundo, 4. Musasa) and $\overline{v}_r$ intervals of interest per station, respectively.



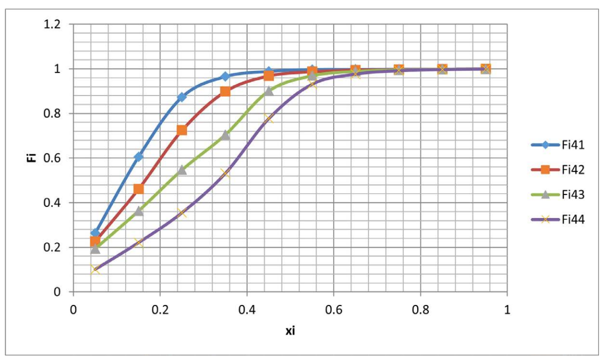
1. Bujumbura



2. Gitega



3. Kirundo



4. Musasa

Figure 2. Integrals of monthly reference relative frequency (F_i) distributions of v_r data for the eleven $\overline{v}_r$ intervals of interest and the four stations, period 1991-1994. In the symbol F_{ijk} , each of the subscripts i, j, k keeps the same meaning as in Figure 1.

5. Conclusions

Up to 48 monthly samples of hourly relative wind speed (v_r) data from four Burundian stations and a 4-year period (1991-1994) have been firstly settled in this work. Each sample of v_r data has been divided into four sub-samples relating to the four years. The monthly frequency distributions of all the sub-samples have been inferred, together with their means and variances. Using those statistical properties in a further task, the F-test has been implemented to ascertain whether or not, for each station, the independent monthly samples of v_r data with the same mean, $\bar{v}_r$, have also the same variance. The null hypothesis has been accepted (*i.e.*: $F_0 < F_C$) for eleven out of twelve identified $\bar{v}_r$ intervals of interest for the four stations. At the opposite, the null hypothesis has been rejected (*i.e.*: $F_0 > F_C$) for the $\bar{v}_r$ interval [0.20; 0.25[at Gitega station. This indicates that among the seven independent monthly samples (related to January, April, May, June, September, November and December, respectively) with mean $\bar{v}_r$ falling in the previous interval at that station, there is at least one pair with different variances. The one-way ANOVA cannot tell us what pairs of samples violate the homogeneity of variance [29]. One may need to run an ad hoc test (e.g. the Games-Howell test [35]) to identify exactly which samples present differences in variance.

For any of the above-mentioned eleven $\bar{v}_r$ intervals of interest, it is therefore ascertained that the independent samples originate from the same mother population and have the same frequency distribution, referred to as a monthly reference frequency distribution of v_r data. Such reference distributions have been built for the four stations and the relevant eleven $\bar{v}_r$ intervals of interest. Those distributions should be used as input data in projects of setting wind energy conversion systems (WECSs) at the selected stations, especially in helping to predict a wind turbine long-term power output or to assess site viability.

The primary wind speed (v) data sets provided us by the IGEBU weather service and from which samples and sub-samples of hourly relative wind speed data (v_r) have been shaped, are from a quite remote period. This is a limitation of our results, indicating that one needs to use more recent v_r data to account for potential long-term climate variability. Moreover, wind speed measurements at the selected stations were taken only at the height of 2 m AGL, which is below the WMO standard height in wind speed analysis ($z_0 = 10\ \text{m}$). If those *in situ* measurements were performed at several heights AGL, then we could use available vertical wind speed profile models in order to interpolate our results to typical wind turbine hub heights.

Acknowledgements

The authors are grateful to the IGEBU weather service for providing them with the primary data used in this study.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Teetz, H.W., Harms, T.M. and von Backström, T.W. (2003) Assessment of the Wind Power Potential at SANAE IV Base, Antarctica: A Technical and Economic Feasibility Study. *Renewable Energy*, **28**, 2037-2061. https://doi.org/10.1016/s0960-1481(03)00076-4

[2] Velázquez, M.T., Rodríguez, J.H., Carmen, M.V.D., Murrieta, F.E.F. and Eslava, G.T. (2016) Application of the Weibull Distribution to Estimate the Volume of Water Pumping by a Windmill. *Journal of Power and Energy Engineering*, **4**, 36-51. https://doi.org/10.4236/jpee.2016.49004

[3] Kazet, M.Y., Mouangue, R., Kuitche, A. and Ndjaka, J.M. (2016) Wind Energy Resource Assessment in Ngaoundere Locality. *Energy Procedia*, **93**, 74-81. https://doi.org/10.1016/j.egypro.2016.07.152

[4] Boudia, S.M., Berrached, S. and Bouri, S. (2016) On the Use of Wind Energy at Tlemcen, North-Western Region of Algeria. *Energy Procedia*, **93**, 141-145. https://doi.org/10.1016/j.egypro.2016.07.162

[5] Messaoudi, D., Settou, N., Negrou, B., Rahmouni, S., Settou, B. and Mayou, I. (2019) Site Selection Methodology for the Wind-Powered Hydrogen Refueling Station Based on AHP-GIS in Adrar, Algeria. *Energy Procedia*, **162**, 67-76. https://doi.org/10.1016/j.egypro.2019.04.008

[6] Ummhani, I.A. and Manahil, E.E.M. (2024) Site Selection for Wind Energy Farm Using GIS-Based Multicriteria Decision Analysis (MCDA) in Qassim Area. *Migration Letters*, **21**, 57-68.

[7] Qiu, X., Li, Y., Li, J., Wang, B. and Liu, Y. (2024) WindFormer: Learning Generic Representations for Short-Term Wind Speed Prediction. *Applied Sciences*, **14**, Article 6741. https://doi.org/10.3390/app14156741

[8] Xu, D., Xue, F., Wu, Y., Li, Y., Liu, W., Xu, C., *et al.* (2024) Analysis of Wind Resource Characteristics in the Ulanqab Wind Power Base (Wind Farm): Mesoscale Modeling Approach. *Energies*, **17**, Article 3540. https://doi.org/10.3390/en17143540

[9] Bashahu, M., Nsabimana, P., Barakamfitiye, J. and Niyukuri, F. (2022) Assessment of the Wind Energy Potential of Two Burundian Sites. *Energy and Power Engineering*, **14**, 181-200. https://doi.org/10.4236/epe.2022.145010

[10] Baseer, M.A., Meyer, J.P., Rehman, S. and Alam, M.M. (2017) Wind Power Characteristics of Seven Data Collection Sites in Jubail, Saudi Arabia Using Weibull Parameters. *Renewable Energy*, **102**, 35-49. https://doi.org/10.1016/j.renene.2016.10.040

[11] Azad, K., Rasul, M., Halder, P. and Sutariya, J. (2019) Assessment of Wind Energy Prospect by Weibull Distribution for Prospective Wind Sites in Australia. *Energy Procedia*, **160**, 348-355. https://doi.org/10.1016/j.egypro.2019.02.167

[12] Galarza, J. (2021) Assessment of Wind Energy Potential and the Application for Micro-Turbines. *International Journal of Electrical Engineering and Technology*, **12**, 20-31. https://doi.org/10.34218/ijeet.12.9.2021.003

[13] Idriss, A.I., Ahmed, R.A., Atteyeh, H.A., Omar, A.I. and Akinci, T.C. (2024) Techno-economic Assessment and Wind Energy Potential of Nagad in Djibouti. *International Journal of Applied Power Engineering* (*IJAPE*), **13**, 91-101. https://doi.org/10.11591/ijape.v13.i1.pp91-101

[14] Mathew, S., Pandey, K.P. and Kumar.V, A. (2002) Analysis of Wind Regimes for Energy Estimation. *Renewable Energy*, **25**, 381-399. https://doi.org/10.1016/s0960-1481(01)00063-5

[15] Karsli, V.M. and Geçit, C. (2003) An Investigation on Wind Power Potential of

Nurdağı-Gaziantep, Turkey. *Renewable Energy*, **28**, 823-830. https://doi.org/10.1016/s0960-1481(02)00059-9

[16] Parajuli, A. (2016) A Statistical Analysis of Wind Speed and Power Density Based on Weibull and Rayleigh Models of Jumla, Nepal. *Energy and Power Engineering*, **08**, 271-282. https://doi.org/10.4236/epe.2016.87026

[17] Nielsen, M.A. (2011) Parameter Estimation for the Two-Parameter Weibull Distribution. Ph.D. Thesis, Brigham Young University.

[18] Ayik, A., Ijumba, N., Kabiri, C. and Goffin, P. (2021) Preliminary Wind Resource Assessment in South Sudan Using Reanalysis Data and Statistical Methods. *Renewable and Sustainable Energy Reviews*, **138**, Article ID: 110621. https://doi.org/10.1016/j.rser.2020.110621

[19] Gatoto, P., Lollchund, M.R. and Dalson, G.A. (2021) Wind Energy Potential Assessment of Some Sites in Burundi Using Statistical Modeling. *Proceedings of* 2021 *IEEE PES/IAS Power Africa Virtual Conference*, Nairobi, 23-27 August 2021, 219-223.

[20] Alzubaidi, M., Hasan, K.N. and Meegahapola, L. (2020) Identification of Suitable Probability Density Function for Wind Speed Profiles in Power System Studies. *Proceedings of the* 2020 *Australian Universities Power Engineering Conference* (*AUPEC*), Hobart, 29 November-2 December 2020, 1-6.

[21] Wais, P. (2017) Two and Three-Parameter Weibull Distribution in Available Wind Power Analysis. *Renewable Energy*, **103**, 15-29. https://doi.org/10.1016/j.renene.2016.10.041

[22] Kidmo, D.K., Deli, K., Raidandi, D. and Yamigno, S.D. (2016) Wind Energy for Electricity Generation in the Far North Region of Cameroon. *Energy Procedia*, **93**, 66-73. https://doi.org/10.1016/j.egypro.2016.07.151

[23] Plan National de Développement (PND Burundi 2018-2027) (2018) Annexe2: 73-87.

[24] Bashahu, M. and Buseke, M. (2016) Statistical Analysis of Hourly Wind Speed Data from Some Burundian Stations Using Beta Probability Density Functions. *Modern Environmental Science and Engineering*, **2**, 740-746. https://doi.org/10.15341/mese(2333-2581)/11.02.2016/005

[25] Placide, G. and Lollchund, M.R. (2024) Wind Farm Site Selection Using GIS-Based Mathematical Modeling and Fuzzy Logic Tools: A Case Study of Burundi. *Frontiers in Energy Research*, **12**, Article 1353388. https://doi.org/10.3389/fenrg.2024.1353388

[26] Bashahu, M., Nsabimana, J.C. and Manirakiza, F. (2006) β Probability Density Functions and Reference Distributions of Relative Sunshine Duration Data from Burundian Stations. *ISESCO Science and Technology Vision*, **2**, 30-34.

[27] Bashahu, M. and Manirakiza, E. (2013) Monthly Reference Frequency Distributions of Daily Clearness Index Data from Burundian Stations. *International Review of Physics* (*IREPHY*), **7**, 352-357.

[28] Ilinca, A., McCarthy, E., Chaumel, J. and Rétiveau, J. (2003) Wind Potential Assessment of Quebec Province. *Renewable Energy*, **28**, 1881-1897. https://doi.org/10.1016/s0960-1481(03)00072-7

[29] Naskar, S. and Das, P. (2018) Application of Different Statistical Tests in Educational Research: An Overview. *Journal of Emerging Technologies and Innovative Research* (*JETIR*), **5**, 129-137.

[30] Kumar, N.K. (2024) F-Test and Analysis of Variance (ANOVA) in Economics. *Mikailalsys Journal of Mathematics and Statistics*, **2**, 102-113. https://doi.org/10.58578/mjms.v2i3.3449

[31] Kebalepile, M.M. and Chakane, P.M. (2022) Commonly Used Statistical Tests and

their Application. *Southern African Journal of Anaesthesia and Analgesia*, **28**, 580-584.

[32] Barbaro, S., Cannata, G. and Coppolino, S. (1983) Monthly Reference Distribution of Daily Relative Sunshine Values. *Solar Energy*, **31**, 63-67. https://doi.org/10.1016/0038-092x(83)90034-8

[33] Sarty, G. (2020) F-Distribution Table. University of Saskatchewan. https://www.saskoer.ca_sites_2020/05

[34] Dougherty, C. (2002) Introduction to Econometrics. 2nd Edition, Oxford University Press.

[35] Marshall, E., Boggis, E., Patel, C., Emmett, M., Owen, A., Russell, J. and Fieller, N. The Statistics Tutor's Quick Guide to Commonly Used Statistical Tests. Statstutor Community Project, University of Sheffield. http://www.statstutor.ac.uk