IJSCIA

International Journal of Scientific Advances
ISSN: 2708-7972
Volume: 4 | Issue: 3 | May – Jun 2023 Available Online: www.ijscia.com
DOI: 10.51542/ijscia.v4i3.26

Geostatistical Modeling of Wind Speed Distribution in Uganda Using Ordinary Kriging Interpolation Technique

Noah Kisuule[1*], Mahmut Cetin[2], Nicholas Kiggundu[1], and Julia Kigozi[1]

[1]Department of Agricultural and Biosystems Engineering,
Makerere University, P.O. Box 7062, Kampala, Uganda

[2]Department of Agricultural Structures and Irrigation,
Cukurova University, 01330 Adana, Turkey

E-mail: nnkisuule@gmail.com; mcet64@cu.edu.tr; nskiggundu@gmail.com;
jbulyakigozi@yahoo.com

*Corresponding author details: Noah Kisuule; nnkisuule@gmail.com

ABSTRACT

Wind power is one of the thriving renewable energy technologies lately in the world. Therefore, the assessment of wind speed data is imperative for a specific site. This study focused on geostatistical modeling of wind speed distribution in Uganda. Wind speed data from 1981 to 2019, recorded at a height of 10 m above mean sea level was captured from NASA POWER Data Access Viewer and analyzed. The study area consisted of 35 stations evenly distributed over the country. Probabilistic assessment was performed using Minitab® statistical software to determine the best-fitting probability distribution function to the data sets. Experimental semi-variograms were calculated using exceedance probability data sets of 20%, 50%, 80%, and 95% obtained from probabilistic analysis. The theoretical models were then fitted to the experimental semi-variograms. Jack-knifing cross-validation approach was employed to assess the performance and validity of the theoretical semi-variogram models and their parameters. Kriging maps for the wind speed of pre-defined probabilities were then generated using JeoStat geostatistical software and Surfers 13®. Of the different theoretical semi-variogram models tested, the spherical model showed the best fit to all experimental semi-variograms. Cross-validation proved that the theoretical models obtained were in good agreement with the experimental data used. The kriging maps revealed that areas around Lake Victoria and Mt. Elgon experience higher wind speeds compared to other parts. Therefore, Kriging maps can be used to assess spatial distribution and magnitude of wind speed as well as the representativeness of geographical locations of observation stations.

Keywords: Geostatistics; wind speed; wind speed distribution; kriging; semi-variogram

INTRODUCTION

The wind is one of the notable atmospheric variables which directly influence physical and biological processes around the globe [1]. On the other hand, wind power has been one of the fastest-growing renewable energy technologies and receiving much attention lately. Wind power is a popular, sustainable, renewable energy source that has a much smaller impact on the environment than burning fossil fuels. As such, it is an important cornerstone of a non-polluting and sustainable energy supply [2,3]. Finding suitable wind speeds can be a basis for accurate estimation of the wind energy potential of a specific region [4]. Wind power is considered a viable alternative energy source of electricity in recent times. This is because it reduces the harmful effects of conventional electric power generation schemes and also ensures the security of the energy supply. But then, wind direction, as well as its velocity, has a significant influence on both pollination and crop growth in agriculture.

Uganda is considered a developing country with great wind energy potential. Wind energy studies in Uganda have concluded that possible applications for the technology exist. These include among others water pumping for irrigation in remote areas and small-scale power generation in mountainous and hilly areas [5,6]. Small-scale industries and factories in rural areas, where targets for a mill range from 2.5 kVA to 10 kVA, could also benefit from the wind resource [5]. Therefore, in concern of this development, it is necessary to have a strong spatial awareness of the characteristics and distribution of wind speed within the country. The wind speed distribution is of immense importance for the assessment of wind energy potential and the performance of wind energy conversion in a given area [7]. It is also essential for structural design and environmental analysis [7], in the pollination of plants, and the regulation of agricultural-related animal behaviors. Therefore, probabilistic analysis, i.e., determination of the magnitude of the frequency of wind speed, is needed in the engineering design of structures. For example, where wind loading on structures needs to be assessed or where long-term energy yields for wind farms need to be evaluated [8].

Several methods have been proposed for the prediction of wind speed distribution. These include Numerical Weather Prediction (NWP), Artificial Neural Networks (ANN), Statistical and Hybrid methods [9].

In engineering practices, Statistical methods i.e., probability distribution models have been widely used. These include Weibull distribution [10–12], Rayleigh distribution [13], Normal distribution, Log-normal distribution [14–16], and Gamma distribution [15]. Apart from the point values of wind speed, the spatial distribution of wind is of great importance to make region-based assessments. In this context, knowing that wind speed can be modeled as a regionalized variable, in turn, geostatistical methods can be used for modeling its spatial distribution in a given area. According to [17], geostatistics is a powerful tool for the estimation of the spatial distribution of geoscience variables. Previous studies have revealed that geostatistical methods are more accurate and objective estimation techniques. Accordingly, they outperform other methods in developing authentic models for any spatially distributed data [18,19]. Hence, this study focused on geostatistical modeling of wind speed distribution in Uganda using the ordinary kriging interpolation technique.

OVERVIEW OF GEOSTATISTICS

[19] defines geostatistics as a branch of statistics that focuses on the analysis and interpretation of any spatially or temporally referenced dataset. According to [20], the objectives of applying geostatistical techniques are twofold: 1) to characterize and interpret the behavior of the existing sample data set and, 2) to use the interpretation to predict likely values at the unsampled locations. In any geostatistical analysis, there are two major steps. Semi-variogram analysis is the first and more vital step. This involves the determination of the spatial dependence structure, called the semi-variogram function [20]. The most commonly used theoretical semi-variogram functions or models are Gaussian, Exponential, Spherical, and Linear [20]. The resulting semi-variogram is a measure of the spatial dependence of the semi-variogram model used to predict values of the modeled variables at unsampled locations. The second step is kriging estimation; then, mapping of kriged estimates and kriging estimation errors. Kriging Equation 1 is solved repeatedly to make the best linear unbiased estimation at the un-sampled locations over the sampling domain [21,22]. More generally, the two steps can be broken down into four steps for complete geostatistical analysis and mapping of regionalized variables [22] as the following: 1) determining an appropriate theoretical semi-variogram model used to fit experimental semi-variogram and possible anisotropy, 2) performing validation methods to the semi-variogram model, 3) generating kriging estimates and errors of estimates, i.e., kriging errors, for a point, zone or volume by kriging interpolation techniques and, 4) mapping the spatial distributions of the kriging estimates and kriging errors.

Ordinary Kriging (OK) Interpolation Method

Kriging is a geostatistical technique that has gained approval as a tool for spatial interpolation of different types of data, including meteorological data [22–25] such as precipitation, temperature, relative humidity, wind speed, etc. It is similar to the inverse distance weighting (IDW) method in that it uses a linear combination of weights at known points to estimate values at unknown points [20,26]. It is a technique of making optimal, unbiased estimates of regionalized variables at unsampled locations using the particular relationship between the spatial proximity among observational units and the numeric similarity among their values, i.e., spatial dependence, and the initial set of data values [27]. In this context, the kriging procedure takes into consideration the spatial structure of the given parameter, and thus it is better compared to other interpolation methods like the arithmetic method, nearest neighbor method, distance weighted method, and polynomial interpolation [27]. Apart from the other interpolation methods, kriging provides the estimation variance at every estimated point, which is an indicator of the estimation accuracy [28]. This is considered the major advantage of kriging over other estimation techniques.

The most commonly used kriging method in geostatistics is the ordinary kriging (OK) method [23,29,30]. The OK method aims at estimating a value of the random variable Z(x) at a point of a region x0 for which a semi-variogram model is known, using data in the neighborhood Z(xα) of the estimation location as shown by Equation 1.

$$Z^*_{OK}(x_0) = \sum_{\alpha=1}^{n(x_0)} \lambda_\alpha{}^{OK} z(x_\alpha) \quad (1)$$

Where λα is the OK weights and n(x0) is the number of data closest to the location x0 to be estimated. In particular, λα values must be evaluated to obtain an unbiased estimation and to minimize the variance.

The OK technique is applied on the basis of two assumptions: 1) the mean of the process is assumed constant over the sampling domain and is invariant within the spatial domain and, 2) the variance of the difference between two values is assumed to depend only on the distance vector h between the two points, and not on the location x [31].

METHODS AND MATERIAL

Characteristics of the Study Area

The study focused on Uganda, which lies in the East African region, across the equator between latitudes 4° 14′ N and 1° 29′ S and longitudes 29° 34′ E and 35° 29′ E. The country is divided into four major regions: Northern, Eastern, Western, and Central. The northernmost point is found at the latitude of 04° 14' N. In the south, Uganda extends to the latitude of 01° 29' S. The southernmost point of Uganda is in Kabale district in the Western region. This point is located along the border that Uganda shares with Tanzania. Uganda's most extreme point in the west is on the border between Uganda and the Democratic Republic of Congo. With a longitude of 29° 34' E, the westernmost point is situated in the Kisoro district. The country's easternmost point is on the dividing line between Uganda and Kenya. This point is found in Nakapiripirit district at a longitude of 35° 02' E. The central part of the country consists of a moderately flat plateau at an elevation of 1000 m to 1300 m above mean sea level (amsl), with high-altitude areas both to the east (Mt. Elgon) and the west (Mt. Rwenzori and Mt. Mufumbira). The western region of the country is also divided by the western rift valley running from north to south at an elevation of 600 m to 900 m amsl. Lake Victoria, which is the largest freshwater lake in Africa, occupies the southeastern corner of the country as seen in FIGURE 1.

FIGURE 1: Map of Uganda showing wind speed observation stations (Captured from Google Earth).

Data Sets

Thirty-five (35) meteorological stations were randomly selected to cover the four regions of Uganda (FIGURE 1). Average monthly wind speed (*WS*) data observed at a height of 10 m during the years 1981 to 2019 was used in this study. The data was downloaded from the POWER Data Access Viewer website *(DAV)* with geographical coordinates in decimal degrees (https://power.larc.nasa.gov/data-access-viewer/). The coordinates were converted to *UTM* coordinates of WGS 84 datum (EPSG:4326) with the help of Google Earth. The *GPS* coordinates of the stations both in geographical and *UTM* coordinate systems and elevation are summarized in TABLE 1.

Exploratory Data Analysis and Probabilistic Assessment

Exploratory data analysis (*EDA*) is a useful perspective to take any path in unraveling the mysteries in observational data. In this regard, *EDA* was carried out to a) check for the missing data, b) gain maximum insight into the data set and its underlying structure, and, c) check for the outliers and other anomalies in the data set.

Following the procedure given by [32] and [20], descriptive statistics such as mean, mode, median, minimum and maximum, variance, standard deviation; Coefficient of variation (*CV*), skewness (C_S), and kurtosis (C_K) for average monthly wind speed data series for each station were calculated.

On the other hand, frequency analysis and probabilistic assessment of average monthly wind speed data were done. Frequency analysis conveys, in a nutshell, the stochastic behavior pattern of the data.

In this regard, Normal, Lognormal, and 3-Parameter Lognormal distribution models have had wide application in engineering design works because of normal distribution's early connection with the "Theory of Errors" and for it has perfect mathematical properties [20]. Therefore, with the help of Minitab statistical analysis software, average monthly wind speed data for each station was empirically fitted to the distribution functions of Normal, Lognormal, and 3–Parameter Lognormal (*LN3*). The corresponding wind speed probability plots for each station were generated to find out which distribution best fits the data set as well as identify the parameters of the distribution functions. From the *LN3* probability plot, the wind speed (*WSx*) data sets with 20%, 50%, 80%, and 95% exceedance probabilities for each station were recorded. The threshold value (β) of the *LN3* distribution was also recorded.

TABLE 1: Geographical and *UTM* coordinates (Datum=WGS84) and elevation data for the selected stations.

No	Station	Geographical Coordinates		UTM Coordinates (Datum: WGS84)		
		Latitude (º)	Longitude (º)	Eastings (m)	Northings (m)	Elevation (amsl, m)
1	Arua	3.02291	30.90771	267 450.556	334 349.485	1 078.04
2	Bushenyi	-0.43629	30.11181	178 488.928	-48 284.908	1 350.92
3	Busia	0.46461	34.09021	621 316.485	51 362.736	1 249.44
4	Entebbe	0.06801	32.47601	441 692.255	7 517.465	1 171.83
5	Fort portal	0.67471	30.28761	198 086.170	74 659.962	1 225.60
6	Gulu	2.77881	32.29071	421 163.950	307 168.073	977.88
7	Hoima	1.45321	31.35271	316 731.040	160 690.542	1 080.10
8	Ibanda	-0.05489	30.47171	218 572.789	-6 072.946	1 350.92
9	Iganga	0.61431	33.47911	553 310.529	67 902.129	1 106.54
10	Jinja	0.43131	33.20731	523 067.814	47 673.037	1 145.36
11	Kabale	-1.20259	30.03431	169 916.867	-133 101.938	1 625.66

No	Station	Geographical Coordinates		UTM Coordinates (Datum: WGS84)		
		Latitude (º)	Longitude (º)	Eastings (m)	Northings (m)	Elevation (amsl, m)
12	Kalangala	-0.31849	32.16751	407 362.846	-35 206.459	1 157.64
13	Kampala	0.32171	32.59411	454 834.932	35 559.523	1 149.04
14	Kamuli	1.03481	33.14621	516 266.894	114 378.097	1 063.23
15	Kasese	0.19131	30.10081	177 255.845	21 172.786	1 384.35
16	Kiboga	0.92191	31.78171	364 441.978	101 922.034	1 180.24
17	Kisolo	-1.28419	29.68801	131 343.841	-142 180.818	1 768.53
18	Kitgum	3.29861	32.88581	487 314.415	364 600.285	975.38
19	Kotido	3.01201	34.11851	624 299.273	332 984.44	1 255.65
20	Kyenjejo	0.61061	30.64411	237 788.836	67 548.251	1 158.15
21	Lira	2.24101	32.88391	487 091.894	247 700.718	1 061.11
22	Masaka	-0.33599	31.73781	359 540.816	-37 146.068	1 194.88
23	Masindi	1.69071	31.72131	357 763.480	186 921.633	1 010.59
24	Mbale	1.07571	34.17671	630 924.699	118 923.659	1 568.84
25	Mbarara	-0.60659	30.66061	239 626.129	-67 102.744	1 398.92
26	Mityana	0.40141	32.04401	393 619.872	44 374.091	1 171.83
27	Moroto	2.52911	34.65961	684 520.236	279 662.416	1 111.35
28	Moyo	3.65531	31.72341	358 222.837	404 127.791	776.50
29	Mubende	0.54841	31.39721	321 634.300	60 639.685	1 222.07
30	Nakasongola	1.31041	32.50081	444 466.375	144 845.464	1 064.84
31	Palisa	1.21941	33.80051	589 059.597	134 794.889	1 077.76
32	Rukungiri	-0.67799	29.90171	155 092.650	-75 048.731	1 467.77
33	Soroti	1.71671	33.62141	569 117.908	189 759.834	1 075.05
34	Tororo	0.69251	34.18091	631 405.547	76 559.568	1 525.13
35	Yumbe	3.55491	31.34171	315 800.838	393 094.741	875.95

Geostatistical Analysis Procedure

The results of frequency analysis, i.e., wind speed (*WSx*) of 20%, 50%, 80%, and 95% exceedance probability data sets, were used for geostatistical analysis.

Experimental Semi-variogram Computation and Semi-variogram Modelling

The semi-variogram is a plot of semi-variances as a function of distances between the observations and is denoted by $\gamma(h)$ [22,25]. Assuming that wind speed data are regionalized variables, Omni-directional experimental semi-variograms of wind speed measurements were estimated using Equation 2. Anisotropic experimental semi-variograms with directional rotation of the data were also tried in this paper, but results were not discussed in detail. Experimental semi-variogram calculations were done with the help of JeoStat geostatistical software [22].

$$\gamma^*(h) = \frac{1}{2}S_h{}^2 = \frac{1}{2N_h}\sum_h (g_i - g_j)^2 \quad (2)$$

Where γ^* is the estimated value of the semi-variance, h is the distance between two stations, N_h is the number of pairs separated by distance vector h, g_i is the wind speed values included in the estimation, and $(g_i - g_j)$ is the difference in values between the two stations found in each pair.

The wind speed (*WSx*) data sets of 20% (*WS20*), 50% (*WS50*), 80% (*WS80*), and 95% (*WS95*) exceedance probability level in text file format was imported into JeoStat software.

To read the data in JeoStat, a two-dimensional data option of no transformation was adopted with the X and Y coordinates set as Eastings and Northings, respectively. The experimental semi-variograms were estimated using the lag spacing details provided in TABLE 2.

TABLE 2: Lag Spacing details for *WS20, WS50, WS80,* and *WS95* data sets.

Lag spacing Values	*WS20*	*WS50*	*WS80*	*WS95*
Minimum (km)	27.1	27.1	27.1	27.1
Maximum (km)	6956.7	6956.7	6956.7	6956.7
Increment (km)	34.1	34.1	34.1	34.1
Number of lags	20	20	20	20

Experimental semi-variograms were then fitted to the most frequently used theoretical models discussed by [20]. Of all the semi-variogram models tested, the spherical model showed the best fit to the experimental semi-variograms of wind speed data. This model represents volumes, and relies on two parameters: a) The range of influence (radius of the sphere) and, b) the sill (plateau) at which the graph reaches the range. In addition to these, there may be a positive *y*-intercept called the nugget effect variance, i.e., C_o. The spherical semi-variogram model is described below by Equation 3.

$$\gamma(h) = \begin{cases} C_0 & , \ h = 0 \\ C_0 + C\left\{\frac{3}{2}\left(\frac{h}{a}\right) - \frac{1}{2}\left(\frac{h}{a}\right)^3\right\} & , \ 0 < h \leq a \\ C_0 + C & , \ h > a \end{cases} \quad (3)$$

Where the parameter "*a*" represents the range of influence of the semi-variogram. This is interpreted as the distance beyond which pairs of sample values are unrelated [21]; C_0+C is the sill of the spherical component and C_0 is the nugget effect on the vertical axis. The final height of the semi-variogram, i.e., sill value, is the sum of C_0 and C [20].

Model Validation

Cross-validation was so used to evaluate the performance and validity of the fitted theoretical semi-variogram models. JeoStat employs a 'jack-knifing' cross-validation approach for validating the generated theoretical semi-variogram and adopted parameters [22,33]. Essentially, it checks the compatibility between the input data and the model outputs. The procedure considers each observation point in turn by removing it temporarily from the data set and using its neighboring information to predict the value of the variable at its location while using the fitted model. It also uses the information from surrounding stations to predict the value in different locations. In JeoStat, the estimate is compared with the measured value by calculating the reduced variable (RE_i), mean reduced error (*MRE*), and reduced variance (*RVAR*) as given by Equations 4 - 6.

$$\mathrm{RE_i} = \left[\frac{\mathrm{Z_i^* - Z_i}}{\sigma_{\mathrm{OK_i}}}\right] \quad (4)$$

$$\mathrm{RVAR} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{RE_i} \quad (5)$$

$$\mathrm{RVAR} = \frac{1}{\mathrm{N-1}}\sum_{\mathrm{i=1}}^{\mathrm{N}} (\mathrm{RE_i} - \mathrm{MRE})^2 \quad (6)$$

where Z^*_i is the estimated (predicted) value at station *I*, Z_i is the observed (measured) value at the station and σ_{OK_i} is the kriging standard deviation or kriging error. The positive and negative errors are defined as overestimates and underestimates, respectively.

The results of *MRE* and *RVAR* given by JeoStat software were used to evaluate whether the spherical semi-variogram models obtained depict the existing variance structure of wind speed. This was done using the procedure described by [22]

Kriging and Mapping of Kriged Estimates

JeoStat software was once again used to generate wind speed estimates and kriging estimation errors at the grid nodes. In this context, point values at grid nodes were estimated by following the criteria suggested by [22]. For all the trials, a total of 39,600 grids (200 x 198 in Eastings and Northings, respectively) with a grid size of 2766 m x 2766 m were established over the study area. Surfers 13® software package was used to generate the kriged (interpolated or prediction) maps for variables of *WS95*, *WS80*, *WS50*, and *WS20*.

To this end, the wind speed estimates generated by JeoStat were imported into Surfers 13®. A contour interval of 0.05 m/s and 0.01 m/s was used for the generation of contour maps (spatial distribution pattern) for the kriged wind speed and kriged errors, respectively.

RESULTS AND DISCUSSION

Descriptive Statistics

Descriptive statistics of wind speed data series by stations are summarized in TABLE 3. In the table, one can see that coefficients of variation (*CVs*) fluctuate between 2.58% (for Lira district) and 6.2% (for Kalangala district). This indicates that wind speed is more or less uniform and does not vary widely in most parts of the country. The mean wind speed also varies between 4.65 m/s and 2.88 m/s with the highest and lowest experienced in Busia (Eastern Uganda) and Fort Portal (Western Uganda), respectively. For most stations studied, the histograms of the data showed a uni-modal characteristic with the mode ranging from 3.14 m/s to 4.54 m/s. Unlike Iganga, Kiboga, Moroto, and Nakasongola, the data set indicated a moderately right (positive) skewed distribution, showing asymmetry (*Cs* > 0). The skewness is positive because the data is skewed to the upper tail of the histogram of wind speed values. A right-skewed distribution behavior might be attributed to the low variability of wind speed across all regions in Uganda. Data recorded from all the stations exhibited a coefficient of kurtosis less than three (<3) with the exception of wind speed data from Kigtum that had a kurtosis coefficient of slightly greater than 3. This means that, according to the normal distribution, wind speed distribution in Uganda is platykurtic. Due to the slight skewness in the sample data, the median is almost equal to the mode and mean wind speed across all stations, indicating that types of normal probability distribution might be the best candidate models to be fitted to the observed frequency of the data. Accordingly, as seen in Table 3, there seems no major discrepancy between the mode, median, and mean wind speed across the country.

TABLE 3: Descriptive statistics of average monthly wind speed (WS, m/s) were recorded from 1981-to 2019.

Station	*Mean*	*Mode*	*Median*	*Min*	*Max*	*StDev**	*CV*	*Cs*	*Cκ*
Arua	3.93	3.79	3.91	3.76	4.29	0.1226	3.12	0.95	1.15
Bushenyi	3.68	3.62	3.69	3.35	4.18	0.1893	5.14	0.48	0.33
Busia	4.65	4.54	4.64	4.31	5.10	0.1799	3.87	0.33	-0.28
Entebbe	3.87	3.81	3.83	3.53	4.40	0.1808	4.67	0.77	0.92
Fort portal	2.88	2.88	2.88	2.68	3.20	0.1074	3.73	0.72	0.72
Gulu	3.98	3.93	3.99	3.73	4.36	0.1302	3.27	0.35	0.55
Hoima	3.90	4.00	3.91	3.62	4.22	0.1627	4.18	0.09	-0.97
Ibanda	3.68	3.62	3.69	3.35	4.18	0.1893	5.14	0.48	0.33
Iganga	3.90	3.92	3.92	3.58	4.14	0.1545	3.96	-0.51	-0.60

Station	Mean	Mode	Median	Min	Max	StDev*	CV	C_S	C_K
Jinja	3.74	3.69	3.70	3.48	4.05	0.1461	3.91	0.64	-0.36
Kabale	3.55	3.45	3.54	3.13	3.98	0.1898	5.35	0.30	-0.17
Kalangala	4.34	4.29	4.29	3.89	5.10	0.2689	6.20	0.82	0.64
Kampala	3.67	3.76	3.64	3.40	4.10	0.1506	4.10	0.56	0.34
Kamuli	4.03	3.93	4.03	3.79	4.30	0.1408	3.49	0.08	-0.87
Kasese	3.32	3.15	3.31	3.02	3.76	0.1576	4.75	0.71	1.02
Kiboga	3.81	3.72	3.82	3.50	4.14	0.1657	4.35	-0.09	-0.70
Kisolo	3.12	3.14	3.10	2.84	3.47	0.1621	5.19	0.42	-0.73
Kitgum	3.89	3.99	3.91	3.59	4.39	0.1409	3.62	0.91	3.11
Kotido	3.89	3.85	3.88	3.51	4.21	0.1965	5.05	0.02	-1.02
Kyenjejo	3.61	3.46	3.62	3.39	3.99	0.1397	3.87	0.49	0.15
Lira	3.93	3.93	3.93	3.68	4.23	0.1013	2.58	0.32	2.20
Masaka	4.40	4.42	4.39	3.98	5.11	0.2495	5.67	0.66	0.36
Masindi	3.82	3.87	3.83	3.57	4.08	0.1324	3.46	0.06	-0.64
Mbale	3.67	3.80	3.66	3.37	4.05	0.1750	4.77	0.30	-0.73
Mbarara	4.17	4.26	4.14	3.67	4.78	0.2358	5.65	0.54	0.64
Mityana	3.87	3.81	3.83	3.53	4.40	0.1808	4.67	0.77	0.92
Moroto	3.55	3.60	3.54	3.22	3.79	0.1475	4.16	-0.26	-0.55
Moyo	3.96	3.91	3.91	3.75	4.29	0.1361	3.44	0.85	0.05
Mubende	3.83	3.74	3.83	3.54	4.18	0.1654	4.32	0.12	-0.74
Nakasongola	3.94	3.99	3.95	3.66	4.19	0.1439	3.65	-0.31	-0.52
Palisa	4.14	4.27	4.14	3.83	4.52	0.1364	3.29	0.06	1.10
Rukungiri	3.04	2.97	3.02	2.79	3.41	0.1399	4.60	0.64	0.39
Soroti	4.03	4.00	4.00	3.84	4.27	0.1104	2.74	0.30	-0.57
Tororo	4.31	4.27	4.28	3.97	4.70	0.1846	4.28	0.21	-0.71
Yumbe	4.11	3.93	4.10	3.87	4.51	0.1573	3.83	0.56	-0.41

*StDev: Standard Deviation; CV: Coefficient of Variation; C_S: Coefficient of Skewness; C_K: Coefficient of Kurtosis.

Outcomes of Exploratory Data Analysis and Frequency Assessment

Exploratory data analysis of observation *WSx* data for each station showed that the data was continuous since there was no year with missing data.

Frequency analysis of wind speed data showed that the 3-parameter lognormal distribution function fitted well to the station's data. For example, the distribution characteristics of wind speed data for the Arua station are shown in Figure 2.

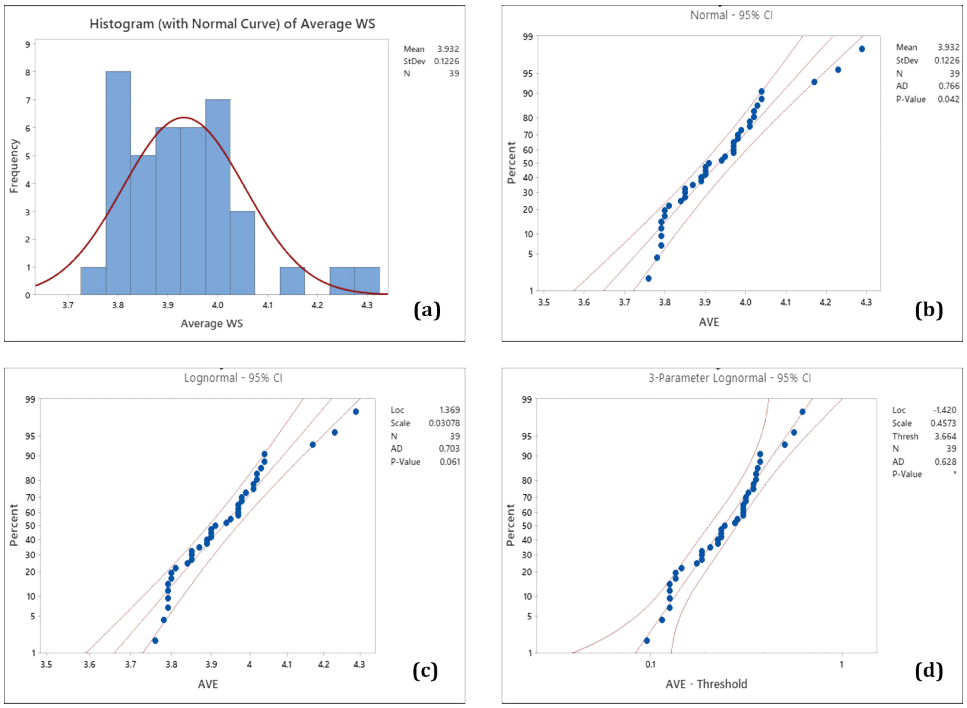

FIGURE 2: Wind speed probability plot of Arua station of 95% confidence intervals: (a) Histogram of average monthly wind speed, (b) Normal plot, (c) Lognormal plot, (d) 3-Parameter Lognormal (LN3) plot. CL stands for 95% confidence limits or intervals.

As seen in Figure 2b, the fit of the normal distribution to the data is not well enough. However, taking the logarithms of data forced a better fit of the normal distribution (Figure 2c). In the end, adding a threshold parameter (β=3.66 m/s) to the data provided a good fit for the LN3 distribution (Figure 2d), resulting in all the observed probabilities (dots in Figure 2d) remaining within the confidence limits. The wind speed data (WSx) of 95%, 80%, 50%, and 20% exceedance probabilities were estimated by using the distribution model of each station. The threshold parameter (β) of LN3 distribution models and estimated WS95, WS80, WS50, and WS20 values were given in Table 4.

TABLE 4: Wind speed values with exceedance probability.

Station	Threshold (β)	WS20	WS50	WS80	WS95
Arua	3.66	4.02	3.91	3.83	3.78
Bushenyi	2.67	3.83	3.67	3.52	3.41
Busia	3.58	4.79	4.63	4.49	4.38
Entebbe	3.14	4.01	3.85	3.72	3.62
Fort portal	2.47	2.96	2.87	2.79	2.73
Gulu	2.81	4.09	3.98	3.88	3.79
Hoima	1.73	4.03	3.89	3.76	3.64
Ibanda	2.67	3.83	3.67	3.52	3.41
Iganga	3.13	4.03	3.90	3.77	3.64
Jinja	3.26	3.85	3.72	3.62	3.54
Kabale	1.95	3.70	3.54	3.39	3.26
Kalangala	3.54	4.54	4.29	4.11	3.98
Kampala	3.01	3.79	3.66	3.55	3.46
Kamuli	1.77	4.15	4.03	3.91	3.80
Kasese	2.61	3.44	3.30	3.19	3.09
Kiboga	3.04	3.95	3.81	3.67	3.54
Kisolo	2.54	3.25	3.10	2.98	2.90
Kitgum	3.06	4.00	3.88	3.77	3.69
Kotido	-5.11	4.05	3.89	3.73	3.58
Kyenjejo	3.01	3.72	3.60	3.49	3.41
Lira	1.97	4.01	3.93	3.85	3.77
Masaka	3.56	4.59	4.36	4.19	4.05
Masindi	-0.09	3.93	3.82	3.71	3.61
Mbale	2.80	3.81	3.65	3.52	3.41
Mbarara	2.70	4.36	4.15	3.97	3.82
Mityana	3.14	4.01	3.85	3.72	3.62
Moroto	3.05	3.67	3.55	3.42	3.30
Moyo	3.65	4.05	3.93	3.84	3.79
Mubende	1.80	3.96	3.82	3.69	3.57
Nakasongola	3.12	4.06	3.94	3.82	3.70
Palisa	-5.73	4.26	4.24	4.03	3.92
Rukungiri	2.45	3.15	3.03	2.92	2.84
Soroti	3.39	4.12	4.02	3.94	3.87
Tororo	2.94	4.46	4.30	4.16	4.03
Yumbe	3.74	4.23	4.08	3.97	3.90

Semi-variogram Analysis

The resulting Omni-directional experimental and theoretical semi-variograms are seen in FIGURE . The experimental semi-variograms were in good agreement with the *CV* values presented in Table 3. All the plotted experimental semi-variograms increased gradually with the distance between pairs up to a distance equal to 200 km. At this point, the semi-variograms displayed random oscillations around a sill value slightly beyond the range of influence. This implied that the experimental semi-variograms would best be fitted by the spherical model described by Equation 3. Apart from the nugget effect component, there exist no nested structures in the experimental semi-variograms. Therefore, nested theoretical models have not been tried in the phase of semi-variogram modeling. The parameters for the developed fitted spherical models for *WS95, WS80, WS50,* and *WS20* are given in TABLE 5.

TABLE 5: Parameters for the spherical semi-variogram models fitted to data.

Parameters	Variable Analysed			
	WS95	WS80	WS50	WS20
Nugget (C_0, (m/s)²)	0.025	0.017	0.020	0.024
Sill (C_0+C, (m/s)²)	0.114	0.107	0.118	0.130
Range (a, km)	295.8	212.4	185.5	184.1
$\frac{Co}{Co+C}$ (%)	22.20	16.30	16.60	18.20
Modified Cressie goodness of fit statistics	0.279	0.314	0.297	0.217

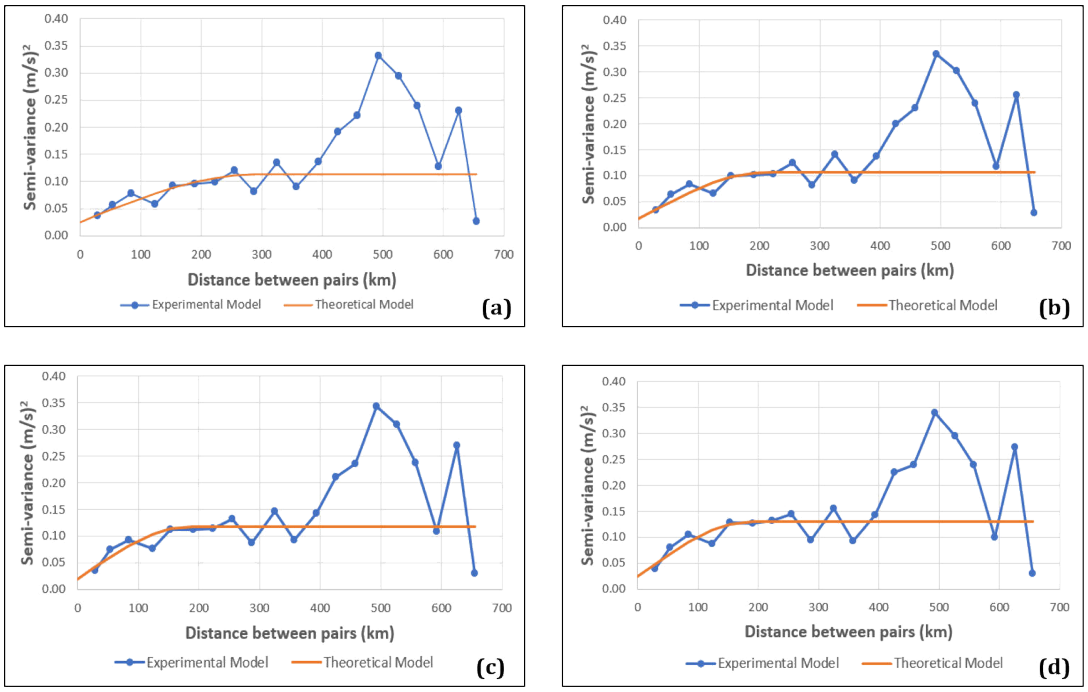

FIGURE 3: Experimental and theoretical semi-variogram models of wind speed for: (a) *WS95*, (b) *WS80*, (c) *WS50*, (d) *WS20*.

The nugget-to-sill ratio, i.e., $\frac{Co}{Co+C} * 100$, is used to classify the spatial dependency in variables [34]. If the nugget-to-sill ratio is less than 25%, the variable is said to have strong spatial dependence. If it is between 25% and 75%, then the variable exhibits moderate spatial dependence; otherwise, the variable has a weak spatial dependence. Results provided in TABLE 5 show that the nugget-to-sill ratio for all four fitted semi-variogram models is less than 25%; therefore, the wind speed distribution structure in Uganda exhibits a strong spatial dependence that almost remained stable over the study years. For all four theoretical semi-variograms generated, the nugget values were found to be relatively small and close to zero. This indicated that there was very small random variability, i.e., inherent spatial variability in the wind speed data, associated with the wind speed sample values. The modified Cressie goodness-of-fit statistic for the fitted theoretical semi-variogram models also indicated good agreement with the experimental ones. The range of influence is the distance within which the wind speed values are spatially dependent. The minimum spatial dependence (184.1 km) was encountered with *WS20* semi-variogram and the maximum (295.8 km) with *WS95*.

Cross-validation Results

According to [20], if the theoretical semi-variogram model and kriging technique are appropriate, the *MRE* and *RVAR* (standard deviation) should be approximately equal to zero and one, respectively.

Kriging search and estimation parameters are of great importance in the '*jack-knifing*' cross-validation procedure in the kriging process.

In this context, kriging search parameters obtained as a result of many trials are given in TABLE 6. By using search parameters, the cross-validation exercise produced the reduced mean statistic of -0.0023, 0.0081, 0.0412, and 0.0275 for the *WS95*, *WS80*, *WS50,* and *WS20* theoretical semi-variograms respectively. The corresponding reduced variances were 0.9937, 1.0071, 1.0457, and 0.9911, respectively (Table 7). This indicates that the theoretical models obtained are in good agreement with the experimental data used and, that the estimation process is unbiased over the sample area. However, theoretical semi-variogram models for the *WS95* and *WS80* data sets were able to predict more accurate wind speed values than the other models because their *MRE* and *RVAR* statistics were found to be so close to zero and one, respectively. The errors are defined so that a positive error is an over-estimate and a negative error is an under-estimate. To this end, as seen clearly from the scatter plots of cross-validation results given in Table 4, the wind speed prediction at 95% exceedance probability is under-estimated whereas others are over-estimated. On the whole, cross-validation results given in
TABLE **7** lead us to conclude that, the accuracy of kriging estimates of wind speed data is acceptable enough over Uganda.

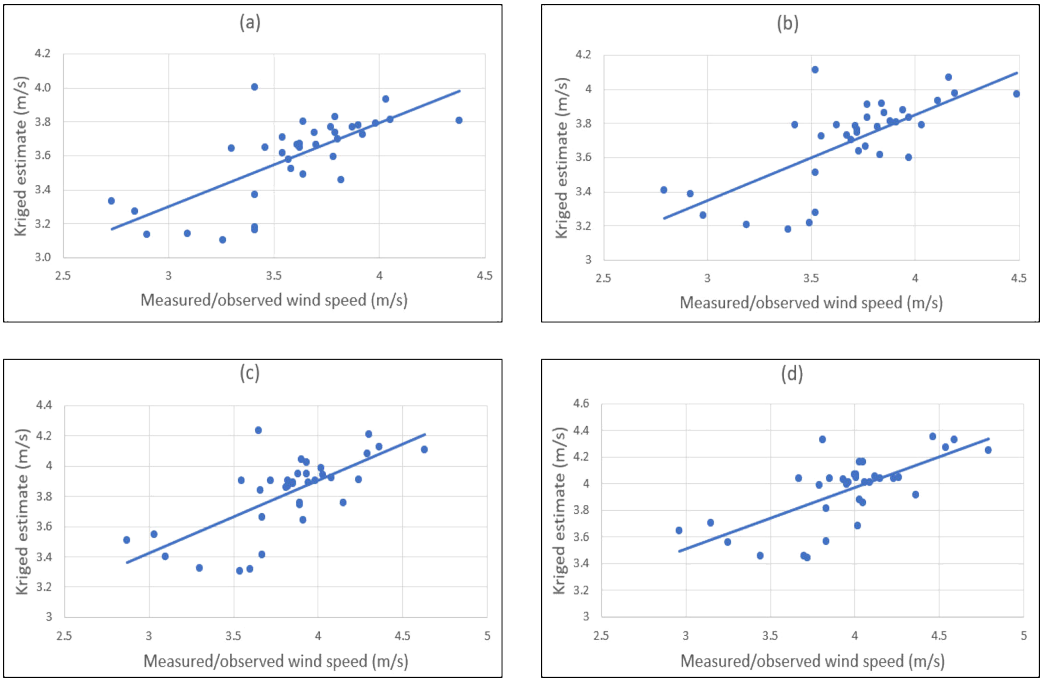

FIGURE 4: Cross-validation scatter plots for: (a) WS95, (b) WS80, (c) WS50, and (d) WS20.

TABLE 6: Kriging search and estimation parameters.

Model Parameters	***Variables***			
	WS95	***WS80***	***WS50***	***WS20***
Major range (*km*)	29.8	212.4	185.5	184.1
Minor range (*km*)	29.8	212.4	185.5	184.1
Angle (^{0})	0	0	0	0
Dip angle (^{0})	0	0	0	0
Plunge angle (^{0})	0	0	0	0
Ratio (*3D*)	1	1	1	1
Max. Pts/sector	7	6	8	8
Min. Pts/sector	2	2	2	2

TABLE 7: Cross-validation results for the fitted theoretical semi-variogram models for Wind Speed.

Wind speed data sets	***Means of measured WS (m/s)***	***Means of estimated WS (m/s)***	***MRE***	***RVAR***
WS95	4.562	4.516	-0.0023	0.9937
WS80	4.612	4.648	0.0081	1.0071
WS50	4.751	4.814	0.0412	1.0457
WS20	4.960	4.971	0.0275	0.9911

Mapping of Kriged Estimates

In the wind speed prediction map (Figure 5), the areas shown in red are the areas where a high amount of wind speed is predicted, and the areas shown in dark blue are the areas with low wind speed. The wind speed prediction (kriged or kriging estimate) maps reveal that prospects for wind energy utilization in Uganda are low for large-scale wind resource applications. However, they indicate that potential exists in areas around the shores of Lake Victoria (Kalangala, Entebbe, Masaka, and Mityana), in the Eastern region, especially around Mt. Elgon (Mbale, Tororo, Busia, and Soroti), North (Yumbe, Kitgum, Gulu, and Arua) and North-Eastern/Karamoja Region (Kitgum, Kapchorwa, Palisa, and Lira). On the other hand, as shown in the kriged error maps (Figure 6), areas with high wind speed are predicted with high errors as compared to low wind speed areas.

The extreme western part of the country is predicted with lower wind speed values, which implies that the probability of operating any wind energy application or development in this area is very low as compared to other parts of the country. The low levels of wind speed in the western region of the country could be attributed to the abundance of dense forests that act as wind buffers. Nonetheless, experimental evidence shows that these amounts of wind speeds are possibly fit for specific wind power applications such as water pumping in villages, small-scale irrigation for agricultural production, and small-scale power generation [35–37]. As might be figured out, it is important to note that the 95% exceedance probability wind speed is predicted with fewer errors compared to others.

The corresponding uncertainties/errors were less than 5% in almost all stations. This confirms that the observation station density is sufficient enough. However, kriging estimation errors tend to increase in Lake Victoria and its environs. It should be addressed that the decrease in station density over the area caused a conspicuous increase in ordinary kriging estimation errors.

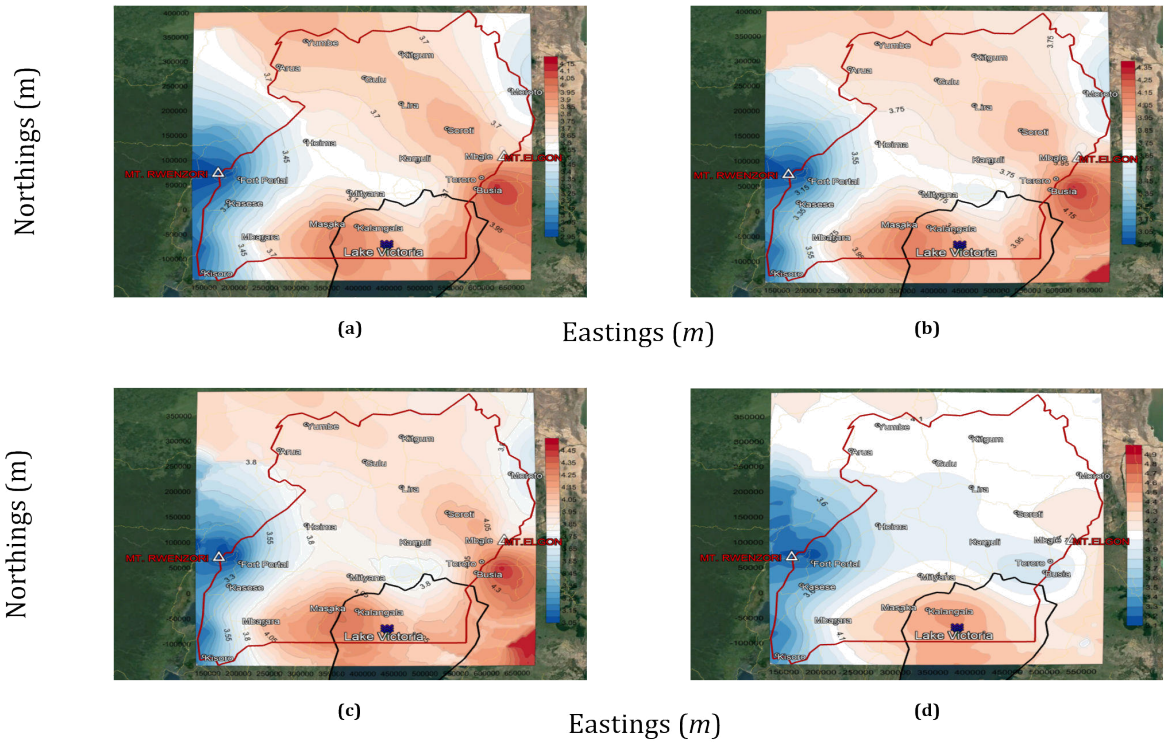

FIGURE 5: Wind speed prediction map for: (a) WS95, (b) WS80, (c) WS50, (d) WS20.

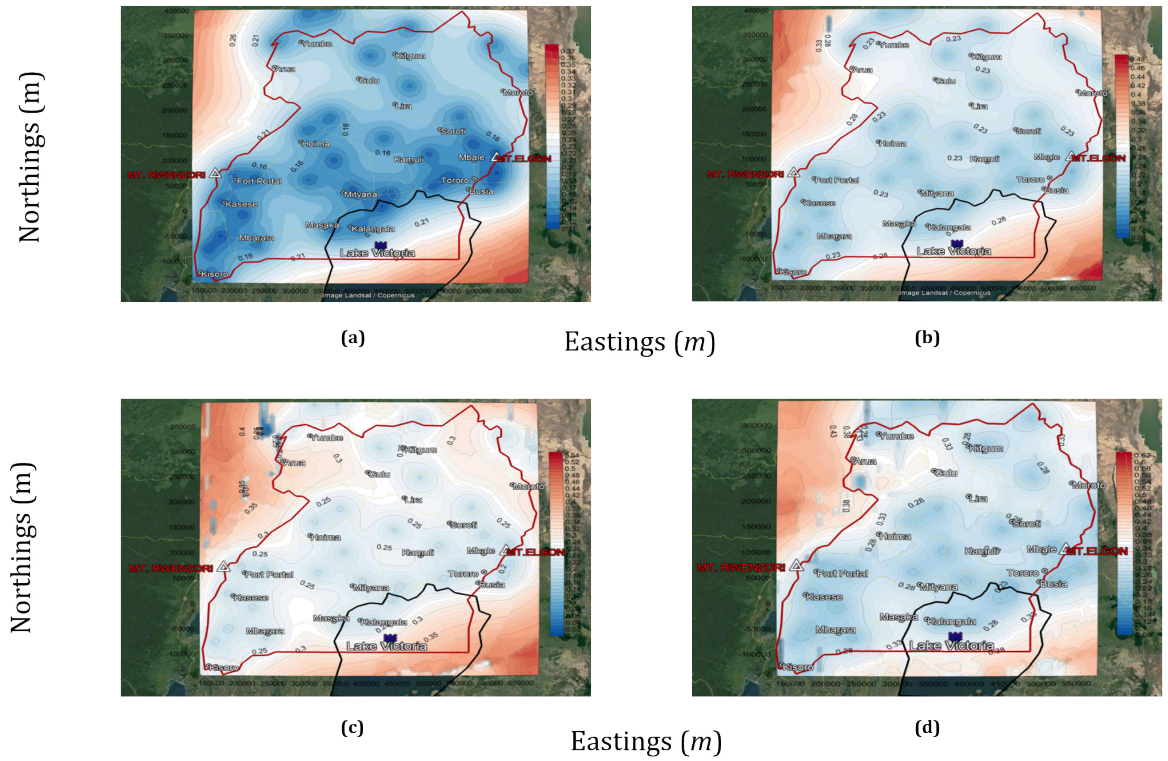

FIGURE 6: Wind speed prediction error map for: (a) *WS95*, (b) *WS80*, (c) *WS50*, (d) *WS20*.

CONCLUSION AND RECOMMENDATIONS

Based on the research results obtained, the following conclusion and recommendations were drawn; Following a systematic statistical analysis procedure plays an important role in obtaining tangible results to check inconsistencies in the data. In this context, exploratory data analysis (*EDA*) showed no anomalies in the observation wind speed data. Lognormal transformation helped to make wind speed data more normal or at least roughly symmetric. However, the additive parameter caused the data to conform well enough to the three-parameter log-normal distribution to make probabilistic assessment of wind speed data. Probability distribution models are the only tools for estimating a value of exceedance probability to assess uncertainty in the data. Therefore, wind speed data of 95% (*WS95),* 80% (*WS80)*, 50% (*WS50)*, and 20% (*WS20*) exceedance probabilities were estimated from probability distribution models. Geostatistical analysis methods were used to reveal spatial variability of wind speed data of pre-defined probabilities, i.e., exceedance probabilities, over large areas.

Experimental semi-variograms indicated no anisotropic behavior in the wind speed data by direction. The spherical model fitted well to the Omni-directional experimental semi-variograms. Kriging estimation and error maps may be used to assess spatial distribution and magnitude of wind as well as the representativeness of geographical locations of observation stations. The decrease in station density over the area caused a conspicuous increase in ordinary kriging estimation errors. In turn, priority should be given to establishing stations at points where ordinary kriging estimation errors are high.

ACKNOWLEDGMENT

This research was carried out within the scope of Erasmus+ study mobility program in Adana, Turkey under the financial support of the European Union (*EU*). The corresponding author, therefore, wishes to thank the Department of Agricultural and Biosystems Engineering of Makerere University for giving him an opportunity to go to Cukurova University (Department of Agricultural Structures and Irrigation), Adana, Turkey where he studied a course in "*Applied Geostatistics*" that enabled him to undertake this study. He is also grateful to Cukurova University for accepting him as a guest student and also availing all the necessary study resources during the study period. Finally, the authors would like to thank Mrs. Semra Sadik Krupka, the Erasmus+ Institutional Coordinator (*Erasmus Mummy*) of Cukurova University for making sure that the funds are received in time plus all the mentorships and guidance throughout the study period.

REFERENCES

[1] Li Y. Spatial interpolation of temperature in the United States using residual Kriging 2017. https://doi.org/10.1016/j.apgeog.2013.07.012.

[2] Zhou J, Erdem E, Li G, Shi J. Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites. Energy Convers Manag 2010; 51:1449–58. https://doi.org/10.1016/j.enconman.2010.01.020.

[3] Marcos R. Characterization of the near-surface wind speed distribution at global scale: ERA-Interim reanalysis and ECMWF seasonal forecasting system. Clim Dyn 2019; 52:3307–19. https://doi.org/10.1007/s00382-018-4338-5.

[4] Bilal BO, Kebe CMF, Sambou V, Ndiaye A, Ndiaye PA, Sava A. Modeling of wind speed distribution for wind power analysis in the northwestern coast of Senegal. 6th International Metrology Conference – CAFMET 2016.

[5] Ssenyimba S, Kiggundu N, Banadda N. Designing a solar and wind hybrid system for small-scale irrigation: a case study for Kalangala district in Uganda 2020;1:1–18.

[6] Wabukala BM, Otim J, Mubiinzi G. Assessing wind energy development in Uganda: Opportunities and challenges. Wind Engineering 2021:1–19. https://doi.org/10.1177/0309524X20985768.

[7] Zaharim A, Najid SK, Razali AM, Sopian K. Analyzing Malaysian Wind Speed Data Using Statistical Distribution. Proceedings of the 4th IASME / WSEAS International Conference on ENERGY & ENVIRONMENT (EE'09) Analyzing, 2010, p. 363–70.

[8] Watson S, Kritharas P. Wind Speed Variability across the UK between 1957 and 2011 2013. https://doi.org/10.1002/we.1679.

[9] Wang Y, Wang D, Zhao J, Zhu C. Wind speed spatial estimation using geostatistical kriging. IOP Conf Ser Earth Environ Sci, 2020. https://doi.org/10.1088/1755-1315/619/1/012049.

[10] Buhairi MH Al. A Statistical Analysis of Wind Speed Data and an Assessment of Wind Energy Potential in Taiz-Yemen. Ass Univ Bull Environ Res 2006; 9:21–33.

[11] Waduge P, Sonnadara U. Modelling Wind Speed Distributions in Selected Weather Stations Modelling Wind Speed Distributions in Selected Weather Stations 2014. https://doi.org/10.13140/2.1.3923.6806.

[12] Kantar YM, Usta İ, Yenİlmez İ, Arİk İ. A Study on Estimation of Wind Speed Distribution by Using the Modified Weibull Distribution 2016:63–70. https://doi.org/10.17671/btd.49478.

[13] Quan P, Leephakpreeda T. Assessment of wind energy potential for selecting wind turbines: An application to Thailand. Sustainable Energy Technologies and Assessments 2015; 11:17–26. https://doi.org/10.1016/j.seta.2015.05.002.

[14] Chang T p. Estimation of wind energy potential using different probability density functions. Appl Energy 2011; 88:1848–56. https://doi.org/10.1016/j.apenergy.2010.11.010.

[15] Sohoni V, Gupta S, Nema R. A comparative analysis of wind speed probability distributions for wind power assessment of four sites. Turkish Journal of Electrical Engineering & Computer Sciences 2016; 24:4724–35. https://doi.org/10.3906/elk-1412-207.

[16] Jung C, Schindler D. Wind speed distribution selection – A review of recent development and progress. Renewable and Sustainable Energy Reviews 2019; 114:109290. https://doi.org/10.1016/j.rser.2019.109290.

[17] Shoji T. Statistical and geostatistical analysis of wind: A case study of direction statistics 2006; 32:1025–39. https://doi.org/10.1016/j.cageo.2005.01.021.

[18] Asa E. Nonlinear Spatial Characterization and Interpolation of Wind Data. Wind Engineering 2012; 36:251–72.

[19] Hengl T, Minasny B, Gould M. A geostatistical analysis of geostatistics. Scientometrics 2009; 80:491–514. https://doi.org/10.1007/s11192-009-0073-3.

[20] Clark I, Harper W V. Practical Geostatistics 2000. Columbus - Ohio: Ecosse North America Llc; 2000.

[21] Cetin M, Kirda C. Spatial and temporal changes of soil salinity in a cotton field irrigated with low-quality water Spatial and temporal changes of soil salinity in a cotton field irrigated with low-quality water 2003;1694. https://doi.org/10.1016/S0022-1694(02)00268-8.

[22] Mert BA, Dag A. A Computer Program for Practical Semivariogram Modeling and Ordinary Kriging: A Case Study of Porosity Distribution in an Oil Field. Journal of Open Geoscience 2017; 9:663–74. https://doi.org/https://doi.org/10.1515/geo-2017-0050.

[23] Apaydin H, Kemal Sonmez F, Yildirim YE. Spatial interpolation techniques for climate data in the GAP region in Turkey. Clim Res 2004; 28:31–40. https://doi.org/10.3354/cr028031.

[24] Ozturk D, Kilic F. Geostatistical approach for spatial interpolation of meteorological data. An Acad Bras Cienc 2016; 88:2121–36. https://doi.org/10.1590/0001-3765201620150103.

[25] Pellicone G. Application of several spatial interpolation techniques to monthly rainfall data in the Calabria region (southern Italy) 2018:3651–66. https://doi.org/10.1002/joc.5525.

[26] Luo W, Taylor MC, Parker SR. A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales 2008; 959:947–59. https://doi.org/10.1002/joc.

[27] Táany RA, Tahboub AB, Saffarini GA. Geostatistical analysis of spatiotemporal variability of groundwater level fluctuations in Amman-Zarqa basin, Jordan: A case study. Environmental Geology 2009; 57:525–35. https://doi.org/10.1007/s00254-008-1322-0.

[28] Kumar V, Remadevi. Kriging of Groundwater Levels – A Case Study. Journal of Spatial Hydrology 2006; 6:12.

[29] Ryu J, Kim M. Kriging interpolation methods in geostatistics and DACE model Kriging Interpolation Methods in Geostatistics and DACE Model. KSME International Journal 2002; 16:619–32. https://doi.org/10.1007/BF03184811.

[30] Oliver MA, Webster R. Catena A tutorial guide to geostatistics: Computing and modelling variograms and kriging. Catena (Amst) 2014; 113:56–69. https://doi.org/10.1016/j.catena.2013.09.006.

[31] Cellura M, Cirrincione G, Marvuglia A, Miraoui A. Wind speed spatial estimation for energy planning in Sicily: Introduction and statistical analysis 2010; 33:1237–50. https://doi.org/10.1016/j.renene.2007.08.012.

[32] David M. Geostatistical Ore Reserve Estimation. Amsterdam: Elsevier Inc; 1977.

[33] Cetin M, Kirda C. Spatial and temporal changes of soil salinity in a cotton field irrigated with low-quality water 2003;272:238–49.

[34] Cambardella CA, Moorman TB, Novak JM, Parkin TB, Karlen DL, Turco RF, et al. Field-Scale Variability of Soil Properties in Central Iowa Soils. Soil Science Society of America Journal 1994; 58:1501–11. https://doi.org/10.2136/sssaj1994.03615995005800050033x.

[35] Adaramola MS, Oyewola OM. On wind speed pattern and energy potential in Nigeria. Energy Policy 2011; 39:2501–6. https://doi.org/10.1016/j.enpol.2011.02.016.

[36] Paul SS, Oyedepo SO, Adaramola MS. Economic assessment of water pumping systems using wind energy conversion systems in the southern part of Nigeria. vol. 30. 2012.

[37] Ohunakin OS, Oyewola OM, Adaramola MS. Economic analysis of wind energy conversion systems using levelized cost of electricity and present value cost methods in Nigeria. International Journal of Energy and Environmental Engineering 2013;4. https://doi.org/10.1186/2251-6832-4-2.