Original Research Paper

Statistical Modelling of Botswana's Monthly Maximum Wind Speed using a Four-Parameter Kappa Distribution

[1]Oliver Moses and [2]Bhagabat Prasad Parida

[1]Okavango Research Institute, University of Botswana, Maun, Botswana
[2]Department of Civil Engineering, University of Botswana, Gaborone, Botswana

Abstract: Wind speed modelling has been key to many environmental and engineering applications, particularly in environmentally friendly wind power generation to meet energy demands. Efficient assessment of wind speed at different recurrence intervals requires the choice of a suitable statistical distribution and an unbiased method of parameter estimation. This study suggests the use of a four parameter Kappa distribution, with its parameters estimated using the method of L-moments, to model Botswana's monthly maximum wind speed data at six meteorological weather stations. These synoptic weather stations are Gaborone, Sir Seretse Khama Airport, Tsabong, Tshane, Gantsi and Maun which are broadly spread across the country's economic activity centres. Reliable wind speed quantiles have been obtained for the selected stations and have been found to fall within the interval 13.80 to 21.69 m s^{-1}. Mean maximum wind speeds have been found to range between 12.65 and 14.97 m s^{-1}, with standard deviations ranging between 1.58 and 2.44 m s^{-1}. These results can reliably be used by environmentalists and technologists working in the energy sector in Botswana.

Keywords: Botswana, Four Parameter Kappa Distribution, L-Moments, Wind Power Production, Wind Speed Quantiles

Introduction

Wind speed is an important factor in many sectors. For instance, it is an important parameter in the estimation of evapotranspiration for agricultural purposes (Valipour and Eslamian, 2014; Valipour, 2014a; 2014b; 2015a; 2015b; 2015c; 2015d). It is also an important factor for engineers in the construction industry, who use it to calculate wind uplift loads (Liu, 1991; Dyrbye and Hansen, 1997; Rosowsky and Cheng, 1999; Kumar and Stathopoulos, 2000; Zhou *et al*., 2002). In this study, Botswana's wind speed is being studied mainly for renewable energy applications (Chiodo, 2013; Mukhopadhyay *et al*., 2014). With the growing rate of urbanisation, Botswana's urban clusters have recently been experiencing energy shortages. Therefore, it has become necessary to explore alternative, environmentally friendly energy sources such as wind mills to meet localised demands. This would reduce the use of energy from fossil fuel resources, the main sources of greenhouse gases. For efficient wind power evaluation, it is essential to model wind speed using a suitable statistical distribution that is able to represent

the observations accurately and the parameters of this distribution need to be estimated using an appropriate technique (Parida, 1999; Shabri and Jemain, 2010; Chiodo, 2013; Mukhopadhyay *et al*., 2014). Although researchers have suggested the use of Weibull distribution (Akpinar and Akpinar, 2004; Azad *et al*., 2014), this may not be applicable in semi-arid regions where wind speeds are highly variable. In view of this, a four parameter Kappa distribution with its parameters estimated using the L-moments method has been used to model Botswana's monthly maximum wind speeds. This distribution encompasses a family of distributions that can effectively describe the peculiarity in data variability (Hosking, 1986; 1990; 1994; Parida, 1999; Shabri and Jemain, 2010). When used for parameter estimation, the L-moments procedure gives unbiased parameter estimates, hence yields unbiased wind speed quantiles that can be used to develop growth curves at various recurrence intervals. Planners, designers and practicing engineers could easily make use of this information to plan and produce wind power efficiently (Chiodo, 2013; Mukhopadhyay *et al*., 2014).

Modelling Botswana's monthly maximum wind speed data for using a four parameter Kappa distribution, with its parameters estimated using the method of L-moments has never been done. This research is therefore the first of its kind in this country and uses one of the best techniques currently available in the world.

Materials and Methods

Data

Historical monthly maximum wind speed data at 10m height for the synoptic weather stations Maun, Gantsi, Tsabong, Sir Seretse Khama Airport (SSKA), Tshane and Gaborone (Fig. 1), were obtained from the Botswana Department of Meteorological Services. The stations were selected because of their data availability and because they are the centres of economic activities. The lengths of the data sets vary from station to station, but all fall within the period 1960 to 2005.

Methods

There are two aspects that make statistical modelling successful. The first aspect is the choice of an appropriate statistical distribution that can adequately describe the observations and the second one is the choice of an appropriate parameter estimation technique, which yields quantiles with the least bias and the least mean squared error. A good statistical model should be able to meet both the descriptive and predictive ability aspects (WMO, 1989), such that more meaningful answers can be deduced from it. Two parameter (2-P) distributions such as the Normal, Exponential, Gamma or Gumbel can explain the descriptive ability aspect well, but fail to account for the predictive ability aspect (Parida, 1999). Three parameter (3-P) distributions such as the Generalised Pareto, Generalised Extreme Value and Generalised Normal distributions are better in their predictive and descriptive ability aspects, but sometimes encounter problems with the conditions of separation.

Hosking (1986; 1990; 1994) showed that the method of Probability Weighted Moments (Greenwood *et al.*, 1979) or the method of L-moments for parameter estimation gives unbiased parameter estimates and hence unbiased quantiles, even when the choice of the parent distribution was inappropriate. This means that one should choose a 3-P distribution rather than a 2-P distribution, since it yields the least biased quantiles with the least mean squared error. Parida (1999) found that the use of a four parameter (4-P) distribution such as a 4-P Kappa can overcome the problem of inappropriate 2-P or 3-P distributions. This distribution can decrease to either of them, depending on the magnitude of the parameters. Therefore, a 4-P Kappa distribution is able to account for the descriptive ability aspects of the given data set very well.
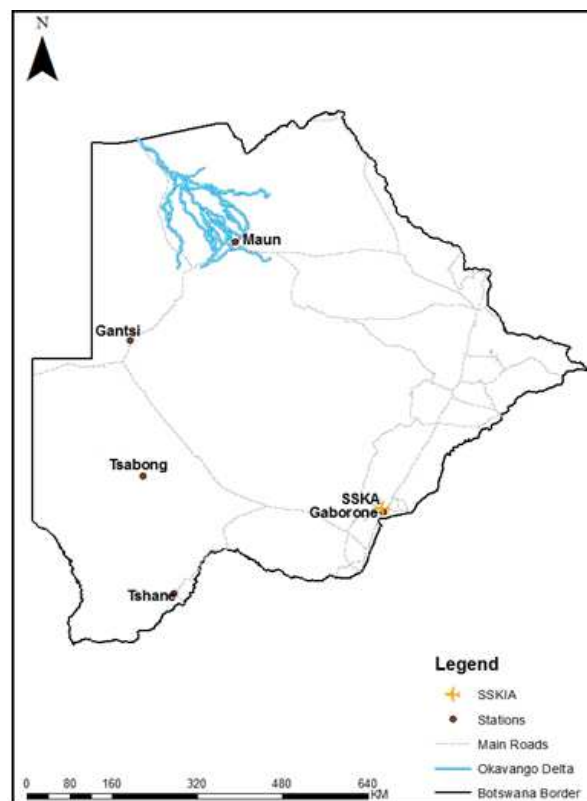

Fig. 1. Location of the six selected weather stations

Table 1. Family of distributions generated by the 4-P generalised Kappa distribution with different values of h and k (adapted from Parida (1999))

h	k	Distribution
1	$\neq 0$	3-P Generalised Pareto Distribution
0	$\neq 0$	3-P Generalised Extreme Value Distribution
-1	$\neq 0$	3-P Generalised Logistic Distribution
1	0	2-P Exponential Distribution
0	0	2-P Gumbel Distribution
-1	0	2-P Logistic D1istribution
1	1	2-P Uniform Distribution (one form of Normal Distribution)
0	1	2-P Reverse Exponential Distribution (i.e., 1 - F(x) is exponential)

When its parameters are estimated using an L-moment procedure, the predictive ability aspects can also be adequately accounted for and the resulting quantiles would be quite reliable. We modelled the monthly maximum wind speed data from different sites in Botswana using a 4-P Kappa distribution, with its parameters estimated using L-moment methods, to obtain unbiased quantiles at each selected site and at the desired recurrence intervals (T). A 4-P Kappa distribution has a cumulative distribution function (Hosking and Wallis, 1997):

$$F(x) = [1 - h\{1 - k(x - u)/\alpha\}^{1/k}]^{1/h} \qquad (1)$$

where, u is the location parameter, α is the scale parameter, h and k are shape parameters, which implicitly include the continuous limits at $h = 0$ and $k = 0$. $F(x)$ is the probability of non exceedance and can also be expressed as $F(x) = 1-1/T$. Special cases of Equation 1 can take the form of different distribution functions with different values of h and k as shown in Table 1.

The quantile function $x(F)$ of the 4-P Kappa distribution can be expressed as:

$$x(F) = u + \frac{\alpha}{k}\left\{1 - \left(\frac{1 - F(x)^h}{h}\right)^k\right\} \tag{2}$$

This equation is the inverse of the cumulative distribution function (Equation 1). The first four L-moments for the data sample of size n arranged in ascending order are obtained using the expressions (Hosking and Wallis, 1997):

$$\lambda_1 = E\left[X_{(1:1)}\right] \tag{3a}$$

$$\lambda_2 = (1/2)E\left[X_{(1:2)} - X_{(2:2)}\right] \tag{3b}$$

$$\lambda_3 = (1/3)E\left[X_{(1:3)} - 2X_{(2:3)} + X_{(3:3)}\right] \tag{3c}$$

$$\lambda_4 = (1/4)E\left[X_{(1:4)} - 3X_{(2:4)} + 3X_{(3:4)} - X_{(4:4)}\right] \tag{3d}$$

where, λ_1, λ_2, λ_3 and λ_4 are related to location, scale, shape and peakedness respectively. In Equation 3b, the connotations (1:2) and (2:2) mean the first and second large values respectively, in a sample size of two drawn from the entire observations made at a station. Appropriate connotations are used in the other equations in a similar manner. It is convenient to present L-moments as L-moment ratios since their ratios measure the shape of a distribution independently of its scale of measurement. Dimensionless third and fourth L-moment ratios t_r, are defined as:

$$t_r = \lambda_r / \lambda_2, \text{ r } = 3, 4 \tag{4}$$

where, t_3 is the L-coefficient of skewness (L-Sk), t_4 is the L-coefficient of kurtosis (L-Ku). The ratio of L-coefficient of variation (L-Cv) is defined as:

$$t_2 = \lambda_2 / \lambda_1 \tag{5}$$

The L-moment ratios t_3 and t_4 are functions of only the shape parameters h and k (Hosking, 1986; 1990; 1994; Parida, 1999). The parameters h and k are restricted by conditions below (Equation 6a to 6d):

$$k > -1 \tag{6a}$$

$$if \ h < 0 \ then \ hk > -1 \tag{6b}$$

$$h > -1 \tag{6c}$$

$$k + h / 1.38 > -1 \tag{6d}$$

Existence of the L-moments is ensured by Equation 6a and 6b, while their uniqueness is ensured by Equation 6c and 6d. The first step of parameter estimation based on L-moment methods at a given station involves obtaining solutions for h and k that best describe t_3 and t_4 in the L-moment ratio diagram and also satisfying Equation 6d. For practical purposes, the interval $h \geq -1$ is the most useful one, while $k = (1-3 \ t_3)/(1+ t_3)$ is the equation that is used to calculate k (Hosking, 1986; 1990; 1994). These estimated values of h and k are used to estimate values of the parameters u and α in Equation 2.

Results and Discussion

Estimated values of the mean maximum wind speeds for each station and their standard deviations, the period of the data, the location parameter u, the scale parameter α, the shape parameters h and k of the 4-P Kappa distribution are shown in Table 2. The mean maximum wind speeds varied between 12.65 and 14.97 m s^{-1}, with their standard deviations varying between 1.58 and 2.44 m s^{-1}. Scrutinizing the values of h and k in Table 2, it can be seen that only Tshane has approximately $h = -1$ and $k = 0$, suggesting that the underlying distribution is closer to the two parameter logistic distribution by Table 1. Three of the other stations have approximately $h = 0$ and $k \neq 0$, suggesting that the underlying distribution of their data is closer to the 3-P geneneralized extreme value distribution.

The estimated parameters of the 4-P Kappa distribution were substituted into Equation 2 to obtain estimates of wind speed quantiles (Park et al., 2001) for each station, corresponding to the recurrence intervals (T):10, 20, 50, 100, 200 and 500 years. These recurrence intervals were converted to the Gumbel reduced variate using Equation 7 (Hosking and Wallis, 1997; Fowler and Kilsby, 2003)

$$Y = -Ln(-Ln(F)) \tag{7}$$

where, Ln is the natural logarithm and F is the non-exceedance probability. The computed wind speed quantiles were used to draw the growth curves presented in Fig. 2, with the Gumbel reduced variate on the horizontal axis. The Gumbel reduced variates 2.3, 3.0, 3.9, 4.6, 5.3 and 6.2 (on the horizontal axis of Fig. 2), correspond to the recurrence intervals (T): 10, 20, 50, 100, 200 and 500 years respectively.
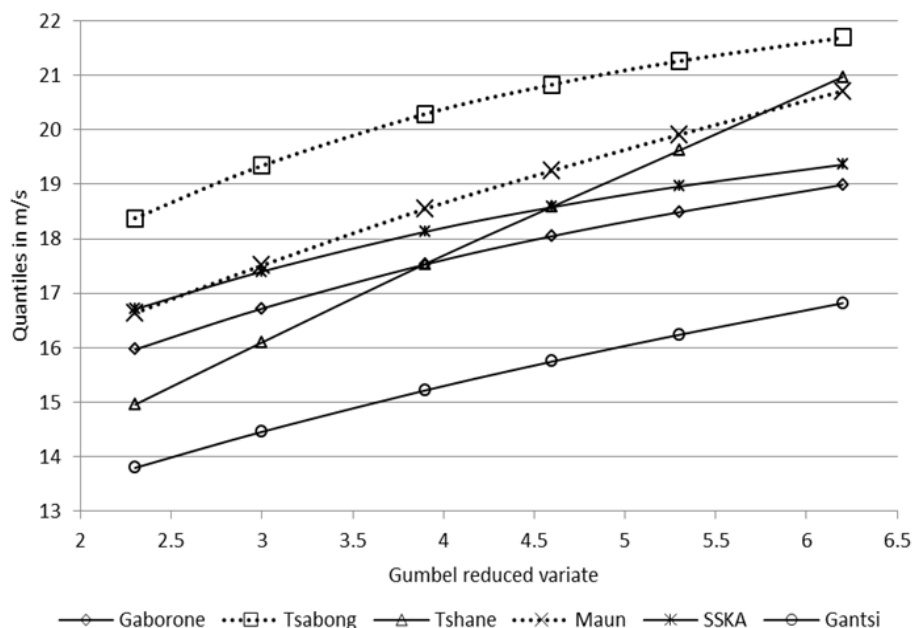
Fig. 2. Growth curves of estimated wind speed quantiles (m/s) for Gaborone, Tsabong, Tshane, Maun, SSKA and Gantsi

Table 2. Wind speed statistics and parameters of 4-P Kappa distribution at 6 stations in Botswana

Station name	Mean maximum (m/s)	Standard Deviation (m/s)	Data period	Parameters of 4-P Kappa distribution			
				Location (u)	Scale (α)	k	h
Gaborone	13.33	2.44	1962-1997	0.96	0.13	0.20	-0.07
Tsabong	14.97	1.71	1960-2002	0.89	0.21	0.34	0.12
Tshane	12.65	1.71	1960-2005	0.89	0.13	0.02	-0.80
Maun	13.34	1.58	1965-1990	1.01	0.12	0.10	-0.42
SSKA	13.21	1.58	1984-2004	1.02	0.14	0.25	0.00
Gantsi	13.38	1.72	1965-2003	0.86	0.09	0.11	-0.43

Table 3. Estimated wind speed quantiles (m/s) compared with the observed values at the six stations in Botswana, for recurrence intervals $T = 10$ years and $T = 20$ years

Station name	T = 10 years		T = 20 years	
	Estimated	Observed	Estimated	Observed
Gaborone	15.97	15.42	16.72	17.99
Tsabong	18.38	16.45	19.34	19.02
Tshane	14.96	15.42	16.10	15.42
Maun	16.64	15.42	17.51	16.96
SSKA	16.71	15.42	17.40	17.48
Gantsi	13.80	15.42	14.46	15.42

The estimated wind speed quantiles at all the stations fall within the range 13.80 to 21.69 m s^{-1}, but the growth curves vary from station to station. The growth curve for Tsabong has the highest wind speed quantiles, ranging between 18.38 and 21.69 m s^{-1}, while Gantsi has the growth curve with the lowest wind speed quantiles, ranging between13.80 and 16.82 m s^{-1}. Variations in wind speeds (hence variations in growth curves) can be attributed to variations in the local conditions or surroundings (e.g., buildings, vegetation and other topographical features) of each station (Wegley *et al.*,

1980; Larsson, 1986; Oke, 1987; Moses, 2007). Variations in wind speeds can also be attributed to weather systems that influence the country's weather. The influence of these weather systems depends on how near or far they are from the stations. It is worth noting that the weather in Botswana is influenced by synoptic scale weather systems such as the Indian ocean high pressure cell, the Atlantic ocean high pressure cell, surface lows, frontal systems, cut-off lows, Inter Tropical Convergence Zone (ITCZ), high pressure cells, easterly and westerly troughs. To check whether or not

the estimated wind speed quantiles corresponded to the observations at each station, the wind quantiles at the recurrence intervals $T = 10$ and 20 years were compared with the observed wind speeds at the respective stations (Table 3). Suppose that the observations were $X_1, X_2, ..., X_N$, where $X_1 < X_2 ... < X_N$, with N being the number of observations in the sample. The Weibull plotting formula (Shaw, 1983) was used to relate the observed wind speeds at $T = 10$ years and at $T = 20$ years, to the estimated wind speed quantiles:

$$i = (1 - 1/T) \times (N + 1) \tag{8}$$

where, i is the rank of the wind speed observation in the ordered data sample. At $T = 10$ years, the differences between the estimated wind speed quantiles and the observations range between 0.55 m s^{-1} (Gaborone) and 1.93 m s^{-1} (Tsabong), while these differences at $T = 20$ years range between 0.08 m s^{-1} (SSKA) and 1.27 m s^{-1} (Gaborone). The small magnitudes of these differences indicate that the estimated wind speed quantiles correspond well to the observations, which indicates that the growth curves presented in Fig. 2 are reliable. In view of this, the growth curves are useful information to the professionals in the energy sector who are concerned with environmentally friendly alternate energy sources such as wind power production.

Conclusion

In the recent past, Botswana has been highly dependent on energy imports. This has made it necessary to explore alternative energy sources such as wind mills to meet localised demands. For wind power applications, it is crucial to model wind speeds using an appropriate statistical distribution that can adequately describe the observations. In addition, the parameters of such a distribution need to be estimated using an appropriate technique. Monthly maximum wind speed data for Gaborone, Sir Seretse Khama Airport, Tsabong, Tshane, Gantsi and Maun have been modelled using a four parameter Kappa distribution based on L-moment procedure, which has made it possible to obtain reliable wind speed quantiles at recurrence intervals 10, 20, 50, 100, 200 and 500 years. Growth curves have been drawn to display the estimated wind speed quantiles. All the growth curves have wind speed quantiles falling within the range 13.80 to 21.69 m s^{-1}. A comparison between the estimated wind speed quantiles at the recurrence intervals $T = 10$ and 20 years and the observations corresponded well. Mean maximum wind speeds for each selected station have also been computed. They have been found to vary between 12.65 and 14.97 m s^{-1}, with their standard deviations varying between 1.58 and 2.44 m s^{-1}. The results of this study provide valuable information for many environmental and engineering sectors, which include environmentally friendly wind power production. The high correspondence between the estimated wind speed quantiles and the observations implies that the results of the study can be extended to other regions of the country with climates similar to those of the selected stations. To improve the results of the study or to reduce uncertainties, it is recommended that in similar future studies, reanalysis data should be used to extrapolate the wind data to longer-term data (Schwartz and George, 1999; Brower *et al.*, 2013; Jimenez *et al.*, "n.d."). Different reanalysis data sets will have to be compared to identify the most reliable one.

Author's Contributions

Oliver Moses: Conceptualised the research, organised ideas, requisition and analysis of the data, results interpretation and article write up.

Bhagabat Prasad Parida: Participated in the conceptualisation and organisation of ideas, data analysis, interpretation of the results, participated in the article write up.

Ethics

The two authors have read and approved the manuscript and give the assurance that no part of this original research article is being considered for publication in whole or in part elsewhere.

References

Akpinar, E.K. and S. Akpinar, 2004. Statistical analysis of wind energy potential on the basis of the Weibull and Rayleigh distributions for Agin-Elazig, Turkey. J Power Energy, 218: 557-565. DOI: 10.1243/0957650042584357

Azad, A.K., M.G. Rasul, M.M. Alam, S.M.A. Uddin and S.K. Mondal, 2014. Analysis of wind energy conversion system using Weibull distribution. Proc. Eng., 90: 725-732. DOI: 10.1016/j.proeng.2014.11.803

Chiodo, E. 2013. Wind speed extreme quantiles estimation. Proceedings of the International Conference on Clean Electrical Power, Jun. 11-13, IEEE Xplore Press, Alghero, pp: 760-765. DOI: 10.1109/ICCEP.2013.6586944

Dyrbye, C. and S.O. Hansen, 1997. Wind Loads on Structures. 1st Edn., John Wiley and Sons, Chichester, ISBN-10: 0471956511, pp: 13.

Fowler, H.J. and C.G. Kilsby, 2003. A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. Int. J. Climatol., 23: 1313-1334. DOI: 10.1002/joc.943

Greenwood, J.A., J.M. Landwehr, N.C. Matalas and J.R. Wallis, 1979. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. Water Resour. Res., 15: 1055-1064. DOI: 10.1029/WR015i005p01049

Hosking, J.R.M., 1994. The four-parameter Kappa distribution. IBM J. Res. Develop., 38: 251-258. DOI: 10.1147/rd.383.0251

Hosking, J.R.M., 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. J. Royal Stat. Soc., 52: 105-124.

Hosking, J.R.M. and J.R. Wallis, 1997. Regional Frequency Analysis: An Approach Based on L-Moments. 1st Edn., Cambridge University Press, United Kingdom, ISBN-10: 0521430453, pp: 242.

Hosking, J.R.M., 1986. The theory of probability weighted moments. IBM Res. Report RC 12210.

Jimenez, B., K. Moennich, J. Rey and F. Durante, "n.d.". Use of different globally available long-term data sets and its influence on expected wind farm energy yields.

Kumar, K.S. and T. Stathopoulos, 2000. Wind loads on low building roofs: A stochastic perspective. J. Struc. Eng., 126: 944-956. DOI: 10.1061/(ASCE)0733-9445(2000)126:8(944), 944-956

Larsson, F., 1986. Wind Resources in Botswana. 1st Edn., Ministry of Mineral Resources and Water Affairs, Gaborone, pp: 5.

Liu, H., 1991. Wind Engineering: A Handbook for Structural Engineers. 1st Edn., Prentice Hall, Englewood Cliffs, ISBN-10: 0139602798, pp: 209.

Brower, M.C., M.S. Barton, L. Lledo and J. Dubois, 2013. A study of wind speed variability using global reanalysis data. AWS Truepower, Inc.

Moses, O., 2007. Study of stochastic and statistical behaviour of some climatic parameters in Botswana. MSc Thesis, University of Botswana, Botswana.

Mukhopadhyay, S., D. Dash, A. Mitra, P. Bhattacharya, 2014. A comparative study between seasonal wind speed by Fourier and Wavelet analysis.

Oke, T.R., 1987. Boundary Layer Climates. 1st Edn., Psychology Press, ISBN-10: 0415043190, pp: 435.

Parida, B.P., 1999. Modelling of Indian summer monsoon rainfall using a four-parameter Kappa distribution. Int. J. Climatol., 19: 1389-1398. DOI: 10.1002/(SICI)1097-0088(199910)19:12<1389::AID-JOC435>3.0.CO;2-T

Park, J.S., H.S. Jung, R.S. Kim and J.H. Oh, 2001. Modelling summer extreme rainfall over the Korean Peninsula using Wakeby Distribution. Int. J. Climatol., 21: 1371-1384. DOI: 10.1002/joc.701

Rosowsky, D.V. and N. Chen, 1999. Reliability of light-frame roofs in high-wind regions. II: Reliability analysis. J. Struc. Eng., 125: 725-739. DOI: 10.1061/(ASCE)0733-9445(1999)125:7(734), 734-739

Schwartz, M.N. and R.L. George, 1999. On the use of reanalysis data for wind resource assessment. National Renewable Energy Laboratory.

Shabri, A. and A.A. Jemain, 2010. LQ-moments: Parameter estimation for kappa distribution. Sains Malaysiana, 39: 845-850.

Shaw, E.M., 1983. Hydrology in Practice. 1st Edn., Van Nostrand Reinhold, Wokingham, ISBN-10: 0442305656, pp: 569.

Valipour, M., 2014a. Comparative evaluation of radiation-based methods for estimation of potential evapotranspiration. J. Hydrol. Eng. DOI: 10.1061/(ASCE)HE.1943-5584.0001066, 04014068

Valipour, M., 2014b. Use of average data of 181 synoptic stations for estimation of reference crop evapotranspiration by temperature-based methods. Water Resour Manage, 28: 4237-4255. DOI: 10.1007/s11269-014-0741-9

Valipour, M., 2015a. Evaluation of radiation methods to study potential evapotranspiration of 31 provinces. Meteorol. Atmos. Phys., 127: 289-303. DOI: 10.1007/s00703-014-0351-3

Valipour, M., 2015b. Investigation of Valiantzas' evapotranspiration equation in Iran. Theor. Applied Climatol., 121: 267-278. DOI: 10.1007/s00704-014-1240-x

Valipour, M., 2015c. Temperature analysis of reference evapotranspiration models. Meteorol. Applic., 22: 385-394. DOI: 10.1002/met.1465

Valipour, M., 2015d. Study of different climatic conditions to assess the role of solar radiation in reference crop evapotranspiration equations. Arch. Agronomy Soil Sci., 61: 679-694. DOI: 10.1080/03650340.2014.941823

Valipour, M. and S. Eslamian, 2014. Analysis of potential evapotranspiration using 11 modified temperature-based models. Int. J. Hydrol. Sci. Technol. DOI: 10.1504/IJHST.2014.067733

Wegley, H.L., J.V. Ramsdell, M.M. Orgill and R.L. Drake, 1980. A Siting Handbook for Small Wind Energy Conversion Systems. 1st Edn., WindBooks, Richland, ISBN-10: 0880160039, pp: 55.

WMO, 1989. Statistical distributions for flood frequency analysis: Secretariat of the World Meteorological Organization.

Zhou, Y., T. Kijewski and A. Kareem, 2002. Along-wind load effects on tall buildings: Comparative study of major international codes and standards. J. Struc. Eng., 128: 788-796. DOI: 10.1061/(ASCE)0733-9445(2002)128:6(788)