



Wind speed distributions in the Italian coasts

T. Soukissian¹, F. Karathanasi¹, F. Falcieri²

¹ Hellenic Centre for Marine Research
Institute of Oceanography

PO BOX 712 – 19013 Anavyssos (Greece)

Phone/Fax number: +30 22910 76420/+30 22910 76323, e-mail: tsouki@hcmr.gr, flwrak@gmail.com

² CNR - ISMAR

Arsenale-Tesa 104, Castello 2737/F, 30122, Venezia, Italia

Phone/Fax number: +39 041 2407995/+39 041 2407940, e-mail: francesco.falcieri@ve.ismar.cnr.it

Abstract. For offshore wind energy assessment it is necessary to appropriately model and describe wind climate. In this connection, the Rayleigh and Weibull distributions are widely suggested for offshore wind speed modelling. Although the use of these distributions is theoretically consistent, in practice, they are often proved to be inadequate. In two recently published papers some, less known, multi-parameter distributions (Johnson S_B , Kappa and Wakeby) were introduced and proved to describe more accurately the stochastic behaviour of wind speed measurements obtained from buoys located in entirely different sea areas of the world. In order to evaluate their fitting performance with reference to coastal wind speed data, in this paper we assessed long-term time series obtained from ten meteorological land-based stations across the Italian coasts. The obtained results confirmed that the Johnson S_B , Kappa and Wakeby distributions are of general validity for any wind data set analysed, since they performed fairly well for the modelling of coastal wind speeds as well. These distributions adapted better than the Weibull and are suggested as reliable and prominent candidates for the modelling of offshore and coastal wind speed in any sea area.

Key words

Wind speed distribution, Johnson S_B distribution, coastal meteorological stations, Italian coasts.

1. Introduction

Over the last decades, renewable energy resources are becoming increasingly important due to the rise in fossil fuels prices, harmful emission and environmental problems. The development of more reliable and cost efficient technologies is therefore required, at both local and national level. Wind energy is considered one of the most advantageous comparing to the conventional forms of energy, which explains its steadily increasing growth. In Europe, the available onshore sites for the development of a wind park have been considerably reduced. In this connection, offshore wind power is an attractive alternative in combination with the higher and less variable sea wind intensities. In the Mediterranean Sea, no

offshore wind farms (OWFs) have been erected yet despite its priority amongst North European countries; however there is a huge anticipation for the investment of offshore wind energy projects, especially in the eastern (Greece) and western part (Spain, France, Italy) of the basin; see also [1].

Such investments require, first of all, a reliable and accurate knowledge of the prevailing long-term wind climate in the candidate area for OWF development. This is achieved by the accurate estimation of offshore wind speed and consequently, the available offshore wind power. In this way, the aforementioned prerequisites along with the local bathymetry comprise the basic technical criteria for selecting the potential zones for OWF installation. To this respect, the long-term probability density function (pdf) of offshore wind speed is a key component of wind climate description. Although the use of the Rayleigh and Weibull distributions (the latter generalizing Rayleigh with one or two additional parameters) in wind speed modeling is theoretically consistent, in practice, they are often proved to be inadequate; see for example [2, 3, 4]. Furthermore, many authors suggest that it is essential to select the most appropriate distribution for the wind climate of a specific area by implementing statistical fitting procedures and not use the Weibull distribution indiscriminately; see e.g. [5].

Two recent assessments of several potential offshore wind speed distributions, based on long-term wind speed time series obtained from oceanographic buoys, can be found in [6] and [7]. In [6], the five-parameter Wakeby and four-parameter Kappa were proposed for the first time for wind speed modeling and provided excellent fits to the analysed wind data. In [7], the four-parameter Johnson S_B distribution was also proposed for the first time and proved to be a very robust and flexible candidate distribution for wind energy assessment as well as wind speed modeling along with the Wakeby and Kappa distributions. The wind data analysed in [7] were obtained

from oceanographic buoys located in the Eastern (Spanish waters) and Western (Greek waters) Mediterranean Sea.

The main objective for this paper is the assessment and fitting of various probability distributions in order to investigate their efficiency and capability in modeling wind speed data obtained from coastal meteorological stations. Another important aspect of this work is the evaluation of the Johnson S_B distribution using wind data from shore, which are characterized by very different wind climates with respect to the neighboring offshore areas. In particular, long-term wind speed measurements from ten coastal meteorological stations of the Italian coasts located in the Ionian/Adriatic and the Tyrrhenian Seas are analysed; see Section 2. All the examined locations exhibit diversity in the prevailing wind climates. In Section 3, an initial assessment of the performance of twenty analytic distributions is made and the quality of the statistical fits are firstly evaluated in terms of the Kolmogorov–Smirnov, the Anderson–Darling and χ^2 goodness-of-fit tests, and finally, in terms of the coefficient of determination R_a^2 . From this assessment, the fitting results of eleven selected probability distributions (all of which appearing consistently among the best fits) are presented in detail and compared with the Weibull distribution for each coastal station. Finally, in Section 4, some conclusions, remarks and suggestions are provided.

2. Wind measurements

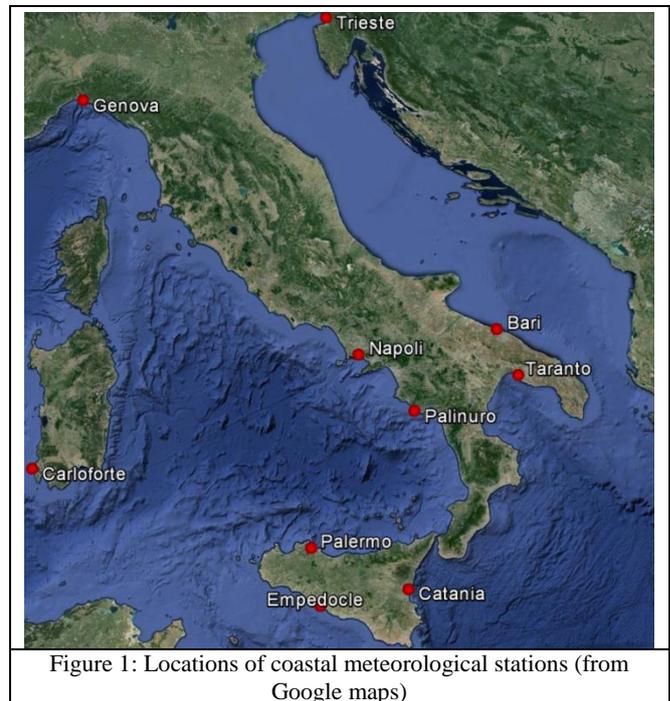
The analysed data in this work were obtained from ten meteorological stations across the Italian coasts covering Adriatic/Ionian as well as the Tyrrhenian Seas. The data were kindly provided by “Servizio Mareografico” of the “Istituto Superiore per la Protezione e Ricerca Ambientale” (APAT) in the context of the FP7-COCONET project; <http://www.coconet-fp7.eu/>.

The data consist of wind speed time series that cover a time period of 15 years. The wind measurements have recording period 600 s and recording interval 1 h. The locations of the stations are shown in Figure 1. The preliminary statistical analysis of these data (sample size, mean, minimum, maximum, standard deviation and coefficient of variation) with the corresponding measurement time periods are presented in Table 1. The highest mean wind speed corresponds to Bari location. The overall maximum value of wind speed (38.507ms^{-1}) is observed in Napoli and the highest values of standard deviation (3.635ms^{-1}) and coefficient of variation (106.357) correspond to Trieste location.

3. Fitting analytic probability distributions to wind speed data

Twenty analytic probability distributions were fitted to the wind speed data for the examined coastal locations. The analysis following in this section showed that the four-parameter Johnson S_B (called hereafter JSB), the four-parameter Kappa (KAP) and the five-parameter Wakeby (WAK) distributions were consistently the only

distributions providing fits in the top ten for all data sets. Thus, the results corresponding to these distributions will be analytically presented for the examined locations and will be compared with the results obtained from the Weibull (WEI) distribution.



The JSB, KAP and WAK distributions are very flexible, successfully mimicking several known distributions. The JSB distribution as well as the KAP and WAK distributions will be presented in the following subsection. For the other distributions examined here, see the relevant wind engineering literature as e.g.: for the lognormal, WEI, Gamma and Log-Pearson 3 see [6] and references cited therein, for the Generalized Gamma see [8], etc. The parameters of the WEI distribution have been estimated using the maximum likelihood method (MLM).

A. The Johnson S_B , Wakeby and Kappa distributions

The JSB distribution is defined as follows:

$$f(x) = \frac{\delta\lambda}{\sqrt{2\pi}(x-\xi)(\xi+\lambda-x)} \times \exp\left\{-\frac{1}{2}\left[\gamma + \delta \ln\left(\frac{x-\xi}{\xi+\lambda-x}\right)\right]^2\right\}, \quad (1)$$

and the corresponding cumulative distribution function (cdf) as follows:

$$F(x) = \Phi\left(\gamma + \delta \ln\left(\frac{x-\xi}{\xi+\lambda-x}\right)\right), \quad (2)$$

where $\Phi(\cdot)$ is the Gaussian cdf, with $\xi < x < \xi + \lambda$, $-\infty < \xi < \infty$, $\lambda > 0$, $-\infty < \gamma < \infty$, $\delta > 0$. The parameter

Name	Recording period	Valid N	Mean	Minimum	Maximum	Std. Var.	Coef. Var.
Bari	1998 – 2012	117558	4.318	0.000	32.700	2.890	66.924
Carloforte	1998 – 2012	120045	2.917	0.000	19.500	2.067	70.841
Catania	1998 – 2012	118062	3.412	0.000	24.200	2.466	72.284
Empedocle	1998 – 2012	103670	3.766	0.000	24.802	3.066	81.427
Genova	1998 – 2012	119247	2.803	0.000	21.700	2.236	79.766
Napoli	1998 – 2012	119349	3.033	0.000	38.507	2.442	80.492
Palermo	1998 – 2012	119610	2.274	0.000	17.800	1.726	75.923
Palinuro	1998 – 2012	119336	2.121	0.000	35.600	2.183	102.961
Taranto	1998 – 2012	120441	3.703	0.000	31.810	2.829	76.396
Trieste	1998 – 2012	115081	3.417	0.000	36.000	3.635	106.357

λ corresponds to the range, ξ is the location parameter (lower bound), δ and γ are shape parameters, while $\gamma = 0$ indicates a symmetrical distribution.

The JSB distribution can be symmetrical or asymmetrical (left-skewed or right-skewed) and uni-modal or bi-modal (for $\delta \leq 0.5$ JSB has a bimodal shape), while it can be partly shaped as uniform between the left and right tails of the variable. The JSB distribution covers an extended variety of distributional shapes (including, among others, the gamma and beta distributions) and an extended region in the skewness-kurtosis space; see [9]. For the estimation of the JSB distribution parameters the method of moments (MOM) has been used in this work; see [10, 11]. However, let us note that inference on JSB distribution is not an easy task, since it is, mathematically, a rather complex distribution.

The WAK distribution is defined, in an inverse manner, as follows:

$$x(F_x) = \xi + \frac{a}{\beta} \left[1 - (1 - F_x)^\beta \right] - \frac{\gamma}{\delta} \left[1 - (1 - F_x)^{-\delta} \right], \quad (3)$$

where $F_x \equiv F_x(x) = \Pr[X \leq x]$. The probability density is given implicitly as:

$$f(x) = \left[a(1 - F_x)^{\beta-1} + \gamma(1 - F_x)^{-\delta-1} \right]^{-1}.$$

The parameters ξ and a are location parameters and β , γ , δ are shape parameters. The domain of WAK is $\xi \leq x < \infty$, if $\delta \geq 0$ and $\gamma > 0$, and $\xi \leq x \leq \xi + a/\beta - \gamma/\delta$, if $\delta < 0$ or $\gamma = 0$. For relation (3) the following conditions hold: 1) $a \neq 0$ or $\gamma \neq 0$, 2) $\beta + \delta > 0$ or $\beta = \gamma = \delta = 0$, 3) if $a = 0$, then $\beta = 0$, 4) if $\gamma = 0$, then $\delta = 0$, 5) $\gamma \geq 0$, 6) $a + \gamma \geq 0$, see also [12], where these restrictions are provided. WAK distribution imitates several known and widely used distributions (as e.g. log-normal, extreme value, generalized Pareto and log-gamma distribution). For estimating the parameters of WAK the method of L-moments is used; see [12, 13].

The four-parameter KAP distribution includes as special cases the Generalized Logistic, Generalized Extreme

Value and Generalized Pareto distributions. The KAP probability density function is defined as follows:

$$f_x(x) = \frac{1}{\alpha} \left[1 - \frac{\kappa}{\alpha} (x - \xi) \right]^{\frac{1}{\kappa} - 1} [F_x(x)]^{1-h}, \quad (4)$$

where

$$F_x(x) = \left\{ 1 - h \left[1 - \frac{\kappa}{\alpha} (x - \xi) \right]^{\frac{1}{\kappa}} \right\}^{\frac{1}{h}}, \quad (5)$$

is the corresponding cumulative distribution function. The quantile function (inverse cdf) is given as

$$x(F) = \xi + \frac{a}{\kappa} \left\{ 1 - \left[\frac{1 - [F(x)]^h}{h} \right]^{\frac{1}{\kappa}} \right\}^{\kappa}. \quad (6)$$

For the KAP distribution the following conditions hold: $x < \xi + \alpha/\kappa$, if $\kappa > 0$, $x < \infty$ if $\kappa \leq 0$, $x > \xi + \alpha(1 - h^{-\kappa})/\kappa$, if $h > 0$, $x > \xi + \alpha/\kappa$, if $h \leq 0$, $\kappa < 0$ and $x > -\infty$, if $h \leq 0$, $\kappa \geq 0$. For estimating the parameters of KAP distribution the method of L-moments is used.

B. Fitting criteria and goodness-of-fit tests

Before the application of any fitting and parameter estimation procedure, the wind data were corrected and filtered and erroneous values were discarded. The Kolmogorov-Smirnov (K-S) goodness-of-fit test was the first criterion for the selection procedure of the most efficient distributions that were examined, being the most widely used; see [14]. Anderson-Darling (A-D) and χ^2 goodness-of-fit tests were additionally used. These three tests are not equivalent. The K-S test aims to examine whether the empirical cdf of the data closely follows the theoretical cdf. For the K-S test, the critical values are independent from the examined cumulative distribution function, which should be fully pre-specified. The A-D test is more sensitive to deviations in the tails of the distribution than the K-S test. The χ^2 goodness-of-fit test examines whether the observed frequencies of some possible events are consistent with the theoretical

frequencies obtained from the examined analytic distribution. Let us note that different goodness-of-fit statistics may generate inconsistent ranking orders of fit performance among the candidate pdfs; see [15, 16].

The JSB and WAK distributions are the only distributions appearing consistently among the three best fits for 8 out of 10 examined coastal locations according to all three goodness-of-fit tests. In addition to this, according to the K-S test, the JSB distribution provided the best fit for 1 out of 10 examined locations and the second best fit for 5 out of 10 examined locations. According to the A-D test, the JSB provided also the best fit for 7 out of 10 examined locations and the second best fit for 3 out of 10 examined locations and, as for the χ^2 -test, the best and second best fit for 1 and 3 out of 10 examined locations respectively. These results clearly support the generality, robustness and efficiency of the JSB and WAK distributions to model wind speed for coastal locations of diverse wind climates. In addition to the JSB, the following distributions provided fair fits for all examined locations and were further assessed: KAP, Gamma (GAM), Generalized Gamma (GNG), Pearson 6 (PEAR6), Generalized Extreme Value (GEV), Burr (BUR), Generalized Pareto (GPAR), Rayleigh (RAY) and WEI distributions were also included for comparison purposes.

For the final evaluation of the obtained statistical fits the coefficient of determination R_a^2 , has been estimated. Since in the estimation of R_a^2 the respective plotting positions (see below relation (7)) for the examined distributions are unbiased, this criterion is used by many authors for fitting evaluation purposes; see e.g. [2, 6, 17, 18].

The specific plotting position that is used for representing the empirical distribution in the PP plots is

$$F(x_i) = i/(n+1), \quad i = 1, 2, \dots, n. \quad (7)$$

$F(x_i)$ provides an unbiased estimate of the observed cumulative probabilities independently of the distribution considered. The coefficient of determination R_a^2 is given by

$$R_a^2 = \frac{\sum_{i=1}^n (\tilde{F}_i - \bar{F})^2}{\sum_{i=1}^n (\tilde{F}_i - \bar{F})^2 + \sum_{i=1}^n (F_i - \tilde{F}_i)^2}, \quad (8)$$

where

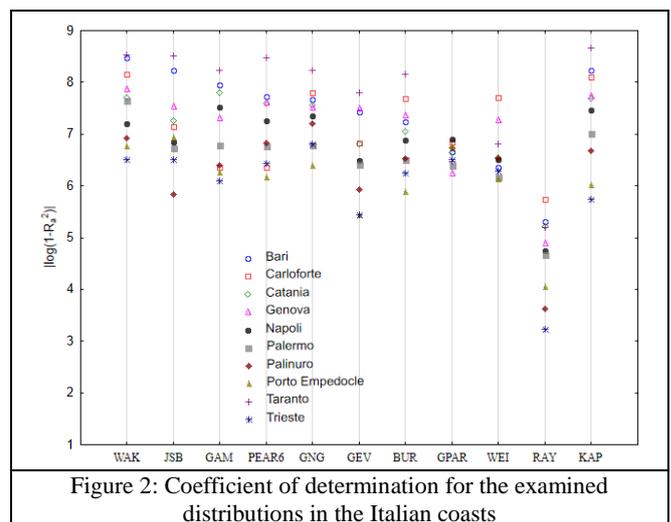
$$\bar{F} = \frac{1}{n} \sum_{i=1}^n \tilde{F}_i. \quad (9)$$

The estimates \tilde{F} are obtained from the examined theoretical probability models, while F_i , $i = 1, 2, \dots, n$, is provided by relation (7) above.

C. Best fit results

As mentioned in the previous subsection, the evaluation of R_a^2 was made for WAK, JSB, KAP, GAM, PEAR6, GEV, GNG, BUR, GPAR, RAY and WEI distributions. The corresponding results were plotted appropriately scaled, i.e., using $|\log(1 - R_a^2)|$ instead of R_a^2 , since the latter often results in values very close to each other and thus, is hardly distinguishable in a graph.

In Figure 2, the WAK distribution provided the best fit in 4 out of 10 locations (for 5 locations provided the second best fit), KAP provided the best fit in 1 location (the second and third best fits for 4 and 2 locations respectively) and JSB distribution provided the best and second fit once and the third best fit also for 2 locations. In general, the JSB, GAM, GNG, KAP and WAK distributions appeared within the ten best fits (among all the examined distributions) for all the examined locations. It is worth noticing that, especially for the JSB, KAP and WAK distributions, the coefficient of determination R_a^2 took constantly high values.



In conclusion, the JSB and WAK distributions were systematically among the five best fits for both goodness-of-fit tests and coefficient of determination for all locations in the Italian coastline. Consequently, these distributions were the most reliable and prominent candidates for the probabilistic modeling of coastal wind speed.

In Figure 3 and Figure 4, the histograms of wind speed together with the fitted WEI, WAK, JSB, GNG, GAM and KAP distributions were plotted for the coastal locations Catania and Taranto, respectively. It is clear that the WEI distribution underperforms compared to the other examined distributions.

Let us also note that the WAK distribution is double bounded for all coastal locations, except for Palinuro and Trieste, while the JSB distribution is, by construction, always double bounded. Potential ramifications of the

bounded-ness of the examined distributions with extreme value analysis are discussed in [7].

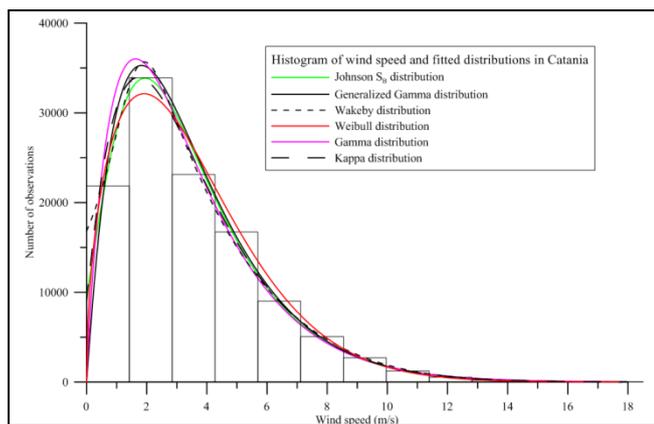


Figure 3: Wind speed histograms and fitted distributions for Catania

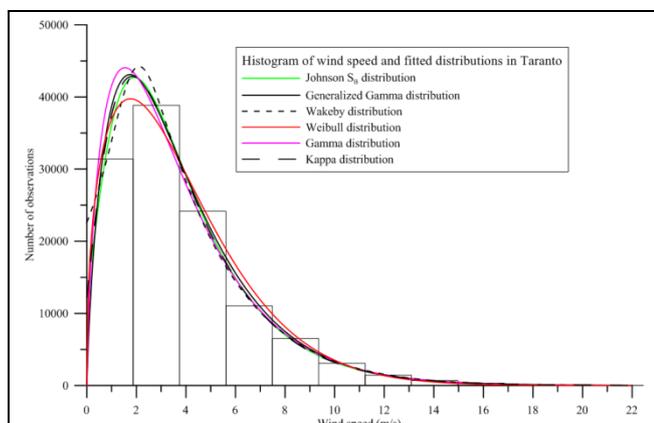


Figure 4: Wind speed histograms and fitted distributions for Taranto

4. Conclusions

In this paper, we model long-term wind speed time series obtained from ten meteorological stations, located at the Italian coasts, examining twenty analytic probability models. An initial evaluation of the obtained fits was based on three statistical tests and, at the final stage, on the coefficient of determination. Among the examined probability distributions, the Weibull and Rayleigh distributions have been also considered for comparison purposes.

The Johnson S_B , Kappa and Wakeby distributions describe quite accurately the empirical distribution of wind speed at the coasts of Italy and have much better adaptability than the Weibull distribution. Taking into consideration the overall behavior of these distributions, one may conclude that they are very promising candidate distributions for the probabilistic modeling of both offshore and coastal wind speed. In this respect, they shall be taken into consideration in any attempt for wind speed modeling along with any other widely used distributions.

Acknowledgement

- 1) The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 287844 for the project "Towards COast to Coast NETworks of marine protected areas (from the shore to the high and deep sea), coupled with sea-based wind energy potential" (COCONET).
- 2) This research has been also funded from the Greek General Secretariat for Research and Technology for the project "National programme for the utilization of offshore wind potential in the Aegean Sea: preparatory actions" (AVRA).
- 3) The authors wish to thank the "Istituto Superiore per la Protezione e Ricerca Ambientale (APAT)" and its "Servizio Mareografico" for the wind data (data are available at www.idromare.it).

References

- [1] Esteban, M. D., Diez, J. J., López, J. S., Negro, V. (2011), Why offshore wind energy?. *Renewable Energy*; 36(2):444-50.
- [2] Garcia, A., Torres, J. L., Prieto, E., de Francisco, A. (1998), Fitting wind speed distributions: a case study, *Solar Energy*; 62(2): 139-44, 10.1016/S0038-092X(97)00116-3.
- [3] Manwell, J. F., McGowan, J.G., Rogers, A. L. (2010), *Wind energy explained: theory, design and application*. Wiley.
- [4] Safari, B., Gasore, J. (2010), A statistical investigation of wind characteristics and wind energy potential based on the Weibull and Rayleigh models in Rwanda, *Renewable Energy*; 35(12):2874-80.
- [5] Carta, J. A., Ramirez, P., Velazquez, S. (2009), A review of wind speed probability distributions used in wind energy analysis Case studies in the Canary Islands. *Renew Sustain Energy Rev*; 13(5):933-55.
- [6] Morgan, E. C., Lackner, M., Vogel, R. M., Baise, L. G. (2011), Probability distributions for offshore wind speeds", *Energy Conversion and Management*; 52:15-26.
- [7] Soukissian, T. (2013), Use of multi-parameter distributions for offshore wind speed modeling: The Johnson S_B distribution, *Applied Energy*, 982-1000.
- [8] Kiss, P., Jánosi, I. M. (2008), Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe, *Energy Conversion and Management*; 49(8): 2142-51, 10.1016/j.enconman.2008.02.003.
- [9] Johnson, N. L., Kotz, S., Balakrishnan, N. (1994), *Continuous Univariate Distributions*. New-York: John Wiley and Sons.
- [10] Draper, J. (1952), *Properties of distributions resulting from certain simple transformations of the normal distribution*. *Biometrika*; 39(3-4):290-301.
- [11] Hill, I. D., Hill, R., Holder R. (1976), Algorithm AS99: fitting Johnson curves by moments. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*; 25:180-89.
- [12] Hosking, J. R. M., Wallis, J. R. (1997), *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press, Appendix A.11.
- [13] Hosking, J.R.M. (1990), L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B* 52: 105-124.
- [14] Lo Brano, V., Orioli, A., Ciulla, G., Culotta, S. (2011), Quality of wind speed fitting distributions for the urban area of Palermo, Italy, *Renewable Energy*; 36(3): 1026-39, 10.1016/j.renene.2010.09.009.

- [15] Chang, TP. (2011), Estimation of wind energy potential using different probability density functions, *Applied Energy*; 88(5):1848-56.
- [16] Zhou, J., Erdem, E., Li, G., Shi, J. (2010), Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites, *Energy Conversion and Management*; 51(7):1449-58.
- [17] Kollu, R., Rayapudi, S. R., Narasimham, S. V. L. (2012), Pakkurthi, K. M., Mixture probability distribution functions to model wind speed distributions, *International Journal of Energy and Environmental Engineering*, 3:27 doi:10.1186/2251-6832-3-27.
- [18] Zhang, J., Chowdhury, S., Messac, A., Castillo, L. A. (2013), Multivariate and Multimodal Wind Distribution model, *Renewable Energy*; 51: 436-47, 10.1016/j.renene.2012.09.026.