https://doi.org/10.26552/com.C.2014.3A.136-141

Ivana Pobocikova - Zuzana Sedliackova - Jan Simon *

STATISTICAL ANALYSIS OF WIND SPEED DATA BASED ON WEIBULL AND RAYLEIGH DISTRIBUTION

In this paper the wind speed data from Meteorological observatory Bratislava-Mlynska dolina were statistically analyzed. The data were analyzed based on Weibull and Rayleigh distribution. The distribution parameters were estimated using maximum likelihood method. The both distributions were compared for their performance using the coefficient of determination and the root mean square error. The results indicate that the better performance can be obtained by the Weibull distribution.

Keywords: *Wind speed data, Weibull distribution, Rayleigh distribution, maximum likelihood method, coefficient of determination, root mean square error.*

1. Introduction

Global warming and environmental pollution become widely discussed issues within last decades. Current major energy sources have significant impact on ecosystems and should be replaced by alternative renewable sources of energy. Wind power plants are promising energy sources with minimal environmental impact and huge energetic potential. However, attention has to be paid to the optimal locality selection to maximize efficiency and reduce costs. Presented paper introduces the Weibull and Rayleigh distribution as the inventive tools in wind speed analyses as well as proposes a complex methodology recommended for evaluation of wind speed conditions in specific locality.

For the sake of this study the wind speed data from the Meteorological observatory Bratislava - Mlynska dolina were statistically analyzed. Processed data were collected during year 2009 in quasi-continuous regime by anemometer connected to electronic buffer. The main objective of presented paper is to propose better probability distribution functions for fitting the observed wind speed data and to establish methodology for wind conditions analyses. Based upon studies [1] - [14] we introduced the Weibull distribution and its special case Rayleigh distribution to approximate the measured wind speed data.

The maximum likelihood method was used to estimate the parameters of the distribution functions. The coefficient of determination (R^2) and the root mean square error ($RMSE$) were used to evaluate the fitting performance of the Weibull and Rayleigh distribution functions.

The Weibull distribution and its special case the Rayleigh distribution are commonly used and recommended probability distributions to describe the wind speed data. The probability density function of the Weibull distribution with parameters $k > 0$ and $c > 0$ is for $v > 0$ given by

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

where v is the wind speed, k is the dimensionless shape parameter and c is the scale parameter in units of the wind speed. The corresponding cumulative distribution function is given by

$$F(v) = 1 - \exp\left(-\left(\frac{v}{c}\right)^k\right), v > 0. \tag{2}$$

The Rayleigh distribution is a special case of the Weibull distribution where the shape parameter is set to $k = 2$. Consequently the probability density function of the Rayleigh distribution transforms as follows

$$f(v) = \frac{2v}{c^2} \exp\left(-\left(\frac{v}{c}\right)^2\right), v > 0. \tag{3}$$

2. Methods for estimating the parameters of the Weibull and Rayleigh distribution

There are several methods available in literature to estimate the Weibull and Rayleigh distribution parameters. In presented

* **Ivana Pobocikova, Zuzana Sedliackova, Jan Simon**
Department of Applied Mathematics, Faculty of Mechanical Engineering, University of Zilina, Slovakia
E-mail: ivana.pobocikova@fstroj.uniza.sk

paper we chose the maximum likelihood method (see [3], [5] and [15]) for estimation of the shape parameter k and the scale parameter c

$$\frac{1}{k}-\frac{\sum_{i=1}^{n} v_i^k \ln v_i}{\sum_{i=1}^{n} v_i^k}+\frac{1}{n}\sum_{i=1}^{n}\ln v_i=0, \tag{4}$$

$$c=\left(\frac{1}{n}\sum_{i=1}^{n} v_i^k\right)^{1/k}, \tag{5}$$

where v_i, $i = 1,2,...,n$, is the wind speed and n is the number of nonzero wind speeds. The shape parameter k was estimated by numerical solving of nonlinear equation (4). Newton method was employed to obtain numerical result. The scale parameter c was estimated by evaluating equation (5).

The maximum likelihood method estimate for the parameter c of the Rayleigh distribution can be solved explicitly by equation (6)

$$c=\left(\frac{1}{n}\sum_{i=1}^{n} v_i^2\right)^{1/2}. \tag{6}$$

3. Descriptions of wind speed data

The wind speed data processed in presented paper were measured at the Meteorological observatory Bratislava - Mlynska dolina, situated in the campus of Faculty of mathematics, physics and informatics, Comenius University in Bratislava, within time frame January 2009 to December 2009. The wind speed and direction were measured continually by anemometer connected to the storage system. In order to remove accidental fluctuations continual data were hourly averaged and rounded to the nearest integer.

4. Statistical analysis of wind speed distributions

The wind speed data were generally divided into subsets with respect to the months and four seasons. Spring was considered to last from March to May. Summer lasts from June to August, autumn from September to November and winter from December to February. The monthly, yearly and seasonal average wind speeds $\bar{v}$ and the standard deviations s_v were calcutated by following equations

$$\bar{v}=\frac{1}{n}\sum_{i=1}^{n} v_i, \tag{7}$$

$$s_v=\sqrt{\frac{1}{n-1}\sum_{n-1}^{n}\left(v_i-\bar{v}\right)^2}, \tag{8}$$

where v_i, $i = 1,2,...,n$, is the averaged wind speed (month, year, season) and n is the number of records.

The estimates of the Weibull and Rayleigh distribution parameters were calculated using (4), (5) and (6) for each month, season and whole year, respectively.

The performance of the Weibull and Rayleigh distribution was evaluated by the coefficient of determination (R^2) and the root mean square error ($RMSE$). These parameters were calculated using equations (9) and (10)

$$R^2=1-\frac{\sum_{i=1}^{N}\left(y_i-x_i\right)^2}{\sum_{i=1}^{N}\left(y_i-\bar{y}\right)^2}, \tag{9}$$

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i-x_i\right)^2}, \tag{10}$$

where N is the number of wind speed data, y_i is the i^{th} ordered observed wind speed data ($y_1 \le y_2 \le ... \le y_N$), x_i is the i^{th} predicted data calculated using the Weibull or Rayleigh distribution, respectively, $i=1,2,...,N$, and $\bar{y}$ is average of values $y_1, y_2,..., y_N$.

The coefficient R^2 ranges from 0 to 1. The higher value of R^2 is better, R^2 approaches 1 in an ideal case. The coefficient $RMSE$ ranges from 0 to infinity. In this case lower value of $RMSE$ is better, in an ideal case it approaches 0. Therefore, the most suitable wind speed distribution is selected according to higher value of R^2 and lower value of $RMSE$. R^2 and $RMSE$ were calculated for each month, season and whole year.

5. Results and discussion

Table 1 shows the monthly and yearly descriptive statistics - average wind speeds, standard deviations, maximum, skewness, kurtosis and median. It has been shown that the yearly average wind speed is 10.485 *km/h* and the yearly standard deviation is 5.841 *km/h*. The monthly average wind speed varies between 8.272 and 13.617 *km/h* with maximum in March and minimum in September The same goes for monthly standard deviation which reaches the highest value in March (7.525 *km/h*) and the lowest one in September (4.327 *km/h*). The monthly average wind speeds are shown in Fig. 1.

Table 2 shows the seasonal wind speed descriptive statistics. One can see that the highest value of the average wind speed is observed in the winter season 11.657 *km/h* and the lowest value in the summer season 9.380 *km/h*. The highest value of the standard deviation was calculated for the spring season 6.260 *km/h* and the lowest one in the summer season 4.917 *km/h*.

Table 3 shows the monthly and yearly estimates of the Weibull and Rayleigh distribution parameters and statistical analysis for the monthly and yearly wind speeds distributions. One can see that the yearly shape parameter k of the Weibull distribution is

Monthly and yearly wind speed descriptive statistics Table 1

Months	Average wind speeds $\bar{v}\,(km/h)$	Standard deviation $s_v\,(km/h)$	Maximum $v_{max}\,(km/h)$	Skewness	Kurtosis	Median $\tilde{v}(km/h)$
Jan	11.013	5.545	29	0.233	-0.412	11
Feb	12.945	6.318	33	0.188	-0.324	13
Mar	13.617	7.525	38	0.533	-0.182	13
Apr	10.140	5.379	28	0.552	-0.219	9
May	10.298	4.910	25	0.336	-0.497	10
Jun	9.269	4.882	28	0.594	0.341	9
Jul	9.870	5.337	37	1.019	1.584	9
Aug	8.997	4.458	26	0.588	-0.037	9
Sep	8.272	4.327	24	0.556	0.063	8
Oct	9.784	6.260	30	0.799	-0.067	8
Nov	10.590	5.930	34	0.663	0.197	10
Dec	11.138	6.202	32	0.492	-0.108	11
Yearly	10.485	5.841	38	0.718	0.429	10

Seasonal wind speed descriptive statistics Table 2

Seasons	Average wind speeds $\bar{v}\,(km/h)$	Standard deviation $s_v\,(km/h)$	Maximum $v_{max}\,(km/h)$	Skewness	Kurtosis	Median $\tilde{v}(km/h)$
Winter	11.657	6.081	33	0.338	-0.250	12
Spring	11.366	6.260	38	0.770	0.586	11
Summer	9.380	4.917	37	0.806	1.007	9
Autumn	9.551	5.657	34	0.819	0.415	9

Monthly and yearly estimates of the Weibull and Rayleigh distribution parameters and statistical analysis for wind speed distributions Table 3

Months	Weibull distribution				Rayleigh distribution		
	k	*c (km/h)*	R^2	*RMSE*	c *(km/*h)	R2	*RMSE*
Jan	2.083	12.423	0.70181	0.01392	12.329	0.70568	0.01380
Feb	2.148	14.598	0.72251	0.01190	14.413	0.70296	0.01231
Mar	1.883	15.362	0.84599	0.00745	15.568	0.82786	0.00788
Apr	1.999	11.493	0.91894	0.00745	11.493	0.91894	0.00745
May	2.233	11.644	0.92244	0.00736	11.415	0.91642	0.00764
Jun	2.174	10.810	0.93649	0.00772	10.631	0.94539	0.00716
Jul	1.952	11.155	0.95882	0.00603	11.219	0.95942	0.00598
Aug	2.142	10.178	0.93706	0.00821	10.040	0.93843	0.00812
Sep	2.002	9.336	0.90796	0.01012	9.334	0.90804	0.01012
Oct	1.629	10.986	0.88958	0.00844	11.628	0.69942	0.01393
Nov	1.877	11.976	0.94372	0.00592	12.152	0.92346	0.00690
Dec	1.979	12.866	0.86210	0.00859	12.895	0.85901	0.00868
Yearly	1.902	11.897	0.99027	0.00251	12.034	0.97921	0.00367

Seasonal estimates of the Weibull and Rayleigh distribution parameters and statistical analysis for wind speed distributions Table 4

Seasons	Weibull distribution				Rayleigh distribution		
	k	*c (km/h)*	R^2	*RMSE*	*c (km/h)*	R^2	*RMSE*
Winter	2.042	13.257	0.85497	0.00868	13.203	0.85634	0.00864
Spring	1.911	12.853	0.97799	0.00355	12.987	0.97070	0.00410
Summer	2.064	10.713	0.98038	0.00430	10.641	0.98071	0.00427
Autumn	1.768	10.770	0.97509	0.00431	11.110	0.92833	0.00732

1.902, while the yearly scale parameter *c* is 11.897 *km/h*. The yearly parameter *c* of the Rayleigh distribution is 12.034 *km/h*. The comparison of the yearly Weibull and Rayleigh probability density distributions with the observed yearly probability density distribution of the wind speed is illustrated in Fig. 2. The top points of the curves are the most frequent wind speeds for the compared distributions (Weibull: 8.037 *km/h*, Rayleigh: 8.509 *km/h*). Both of the theoretical curves of the probability density distributions match the observed data satisfactorily at well acceptable levels of R^2 and *RMSE*.

Analyses show that the monthly shape parameter *k* of the Weibull distribution ranges from 1.629 to 2.233 and the monthly scale parameter *c* ranges from 9.336 to 15.362 *km/h*. The highest value of the parameter *c* was observed in March and the lowest one in September, which corresponds with observation well. The monthly parameter *c* of the Rayleigh distribution ranges from 9.334 to 15.568 *km/h* with maximum in March and minimum in September, thus corresponding to the monthly average wind speeds too.

The performace of the Weibull and Rayleigh distribution was evaluated by the coefficient of determination (R^2) and the root mean square error (*RMSE*). The value of R^2 is 0.99027 for the Weibull distribution and 0.97921 for the Rayleigh distribution when applied to yearly wind speed data. The value of *RMSE* is 0.00251 for the Weibull distribution and 0.00367 for the Rayleigh distribution when applied to the same set of data. The yearly comparison shows that the Weibull distribution returns higher value of R^2 and the smaller value of *RMSE*. This indicates that the Weibull distribution is slightly better choice for fitting the yearly wind speed data than the Rayleigh distribution.

For the monthly wind speed data the value of R^2 ranges from 0.70181 to 0.95882 for the Weibull distribution and from 0.69942 to 0.95942 for the Rayleigh distribution. The *RMSE* ranges from 0.00592 to 0.01392 for the Weibull distribution and from 0.00598 to 0.01393 for the Rayleigh distribution. The month to month comparison shows that, in general, the Weibull distribution leads to the higher values of R^2 and the smaller values of *RMSE* than the Rayleigh distribution. It holds true for 8 months of year 2009. It confirms that the Weibull distribution is slightly better for fitting the monthly wind speed data than the Rayleigh distribution. The values of R^2 and *RMSE* obtained by fitting the monthly probability density distributions derived from the observed data with the Weibull and Rayleigh probability density distributions are illustrated in Fig. 3.

Table 4 shows the seasonal estimates of the Weibull and Rayleigh distribution parameters and statistical analysis for seasonal wind speed. The comparison of seasonal Weibull and Rayleigh probability density distributions with the observed seasonal probability density distributions of the wind speed are illustrated in Fig. 4. In general, the value of the scale parameter *c* of the Weibull distribution is the highest in the winter season and the lowest in the summer season. Basically the same goes for the parameter *c* of the Rayleigh distribution. The seasonal value of the Weibull distribution parameter *k* ranges from 1.768 to 2.064. The value of the parameter *c* ranges from 10.713 to 13.257 *km/h*. The seasonal value of the Rayleigh distribution parameter *c* ranges from 10.641 to 13.203 *km/h*. The seasonal value of R^2 ranges from 0.85497 to 0.98038 for the Weibull distribution while for the Rayleigh distribution ranges from 0.85634 to 0.98071. The value of *RMSE* ranges from 0.00355 to 0.00868 for the Weibull distribution while for the Rayleigh distribution ranges from 0.00410 to 0.00864.

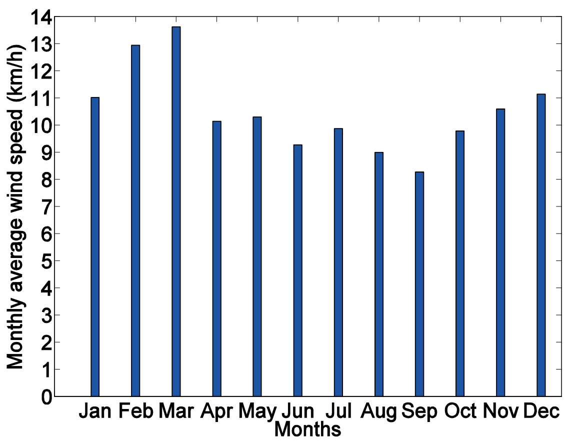

Fig. 1 Monthly average wind speeds

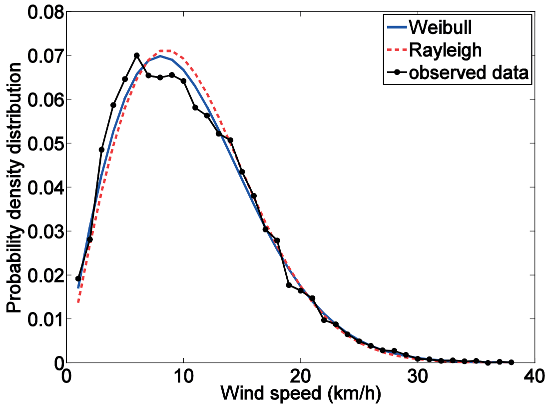

Fig. 2 Comparison of yearly Weibull and Rayleigh probability density distributions

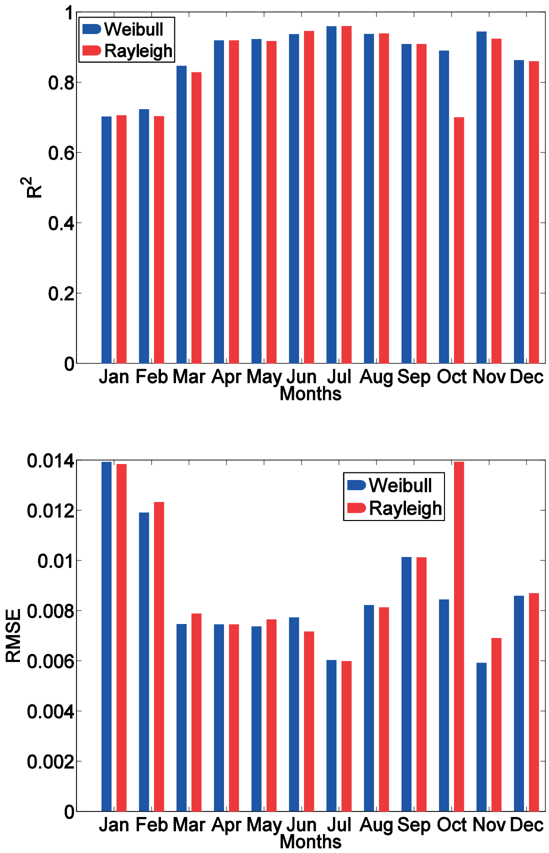

Fig. 3 R^2 and RMSE values obtained by fitting observed monthly probability density distributions with Weibull and Rayleigh probability density distributions

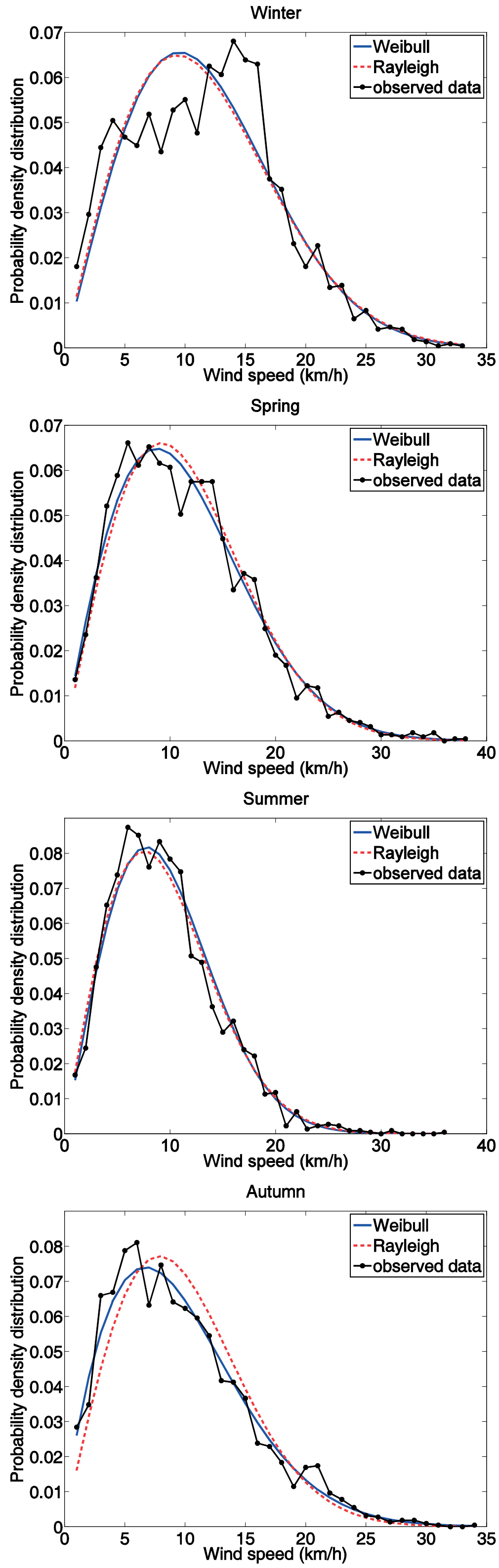

Fig. 4 Comparison of seasonal Weibull and Rayleigh probability density distributions

6. Concluding remarks

In this paper the wind speed data from Bratislava have been statistically analyzed using the Weibull and Rayleigh probability distributions. The probability distributions have been derived from the measured wind speed data for the year 2009. The monthly, yearly and seasonal Weibull and Rayleigh distribution parameters have been calculated. To evaluate the performance of the considered probability distributions the coefficient of determination and the root mean square error have been used. The following conclusions can be made.

The Weibull distribution has been found to be more suitable for fitting the wind speed data obtained in year 2009 in Bratislava than the Rayleigh distribution at the yearly base.

The Weibull distribution has been found to be more suitable for fitting the wind speed data in eight months than the Rayleigh distribution.

The Weibull distribution can be recommended for fitting the wind speed data at the seasonal base.

Acknowledgement

This paper was supported by the Slovak Grant Agency VEGA through the projects No. 1/1245/12. The authors gratefully acknowledge the Meteorological observatory Mlynska dolina and Faculty of mathematics, physics and informatics for providing us with measured wind speed data.

References

[1] ALBUHAIRI, M. H.: A Statistical Analysis of Wind Speed Data and an Assesment of Wind Energy Potential in Taiz-Yemen. *Assiut University Bulletin for Environmental Research* 9(2), 21-33, 2006

[2] AKDAG, S. A., DINLER, A.: A New Method to Estimate Weibull Parameters for Wind Energy Applications. *Energy Conversion and Management* 50, 1761-1766, 2009

[3] CARTA, J. A., RAMIREZ, P., VELASQUEZ, S.: A Review of Wind Speed Probability Distributions Used in Wind Energy Analysis: Case Studies in the Canary Islands. *Renewable and Sustainable Energy Reviews* 13, 933-955, 2009

[4] CELIK, A. N.: A Statistical Analysis of Wind Power Density Based on the Weibull and Rayleigh Models at the Southern Region of Turkey. *Renewable Energy* 29, 593-604, 2003

[5] CONRADSEN, K., NIELSEN, L. B. PRAHM, L. P.: Review of Weibull Statistics for Estimation of Wind Speed Distributions. *J. of Climate and Applied Meteorology* 23, 1173-1183, 1994

[6] DORVLO, A. S. S.: Estimating Wind Speed Distribution. *Energy Conversion and Management* 43, 2311-2318, 2002

[7] FADARE, D. A.: A Statistical Analysis of Wind Energy Potential in Ibadan, Nigeria, Based on Weibull Distribution Function. *The Pacific J. of Science and Technology* 9 (1), 110-119, 2008

[8] GUPTA, B. K.: Weibull Parameters for Annual and Monthly Wind Speed Distribution for Five Locations in India. *Solar Energy* 37 (6), 673-678, 1986

[9] GOKCEK, M., BAYULKEN, A., BEKDEMIR, S.: Investigation of Wind Characteristics and Wind Energy Potential in Kirklareli, Turkey. *Renewable Energy* 32, 1739-1752, 2007

[10] LUN, I. Y. F., LAM, J. C.: A Study of Weibull Parameters Using Long-term Wind Observation. *Renewable Energy* 20, 145-153, 2000

[11] RAMIREZ, P., CARTA, J. A.: Influence of the Data Sampling Interval in the Estimation of Parameters of the Weibull Wind Speed Probability Density Distribution: a Case Study. *Energy Conversion and Management* 45, 2419-2438, 2005

[12] SEGURO, J. V., LAMBERT, T. W.: Modern Estimation of the Parameters of the Weibull Speed Distribution for Wind Energy Analysiss. *J. of Wind Engineering and Industrial Aerodynamics* 85, 75-84, 2000

[13] ULGEN, K., HEPBASLI, A.: Determination of Weibull Parameters for Wind Energy Analysis of Izmir, Turkey. *Intern. J. of Energy Research* 26, 495-506, 2002

[14] FERANEC, V.: Wind Engineering in the 21th Century. Communications - Scientific Letters of the University of Zilina, vol. 2, No. 4, 74-83, 2000

[15] CHU, Y. K., KE, CH. J.: Computation Approaches for Parameter Estimation of Weibull Distribution. *Mathematical and Computational Applications*, vol. 17, No. 1, 39-47, 2012.