Comparison of Some Estimation Methods of the two parameter Weibull Distribution for Unusual Wind Speed Data Cases

Yeliz Mert KANTAR, İbrahim ARIK, İlhan USTA, İsmail YENILMEZ

Department of Statistics, Faculty of Science, Anadolu University, Eskisehir, Turkey
ymert@anadolu.edu.tr, iarik@anadolu.edu.tr, iusta@anadolu.edu.tr, ismailyenilmez@anadolu.edu.tr
(Geliş/Received: 29.10.2015; Kabul/Accepted: 10.02.2016)
DOI: 10.17671/btd.61183

Abstract- Since the Weibull distribution has been accepted reference distribution in wind energy field, topics on its parameter estimation methods get much attention. In this context, the literature have generally focused on non-robust methods, which may yield questionable results in the cases of unusual and contaminated wind speed data. This paper discusses some robust estimation methods of the parameters of Weibull distribution for unusual wind speed data cases. The considered robust methods are evaluated by using artificially generated unusual wind speed data cases. It has been found that some of the considered robust methods provide reliable results compared the classical ones. The similar results are observed for the estimation of the mean power density error. As a result, the analyzes performed show that robust and efficient classical methods can be used together to check the results.

Keywords- Wind speed, wind power, Weibull distribution, robust estimation, unusual wind speed observation.

1. INTRODUCTION

It is clear that the importance of wind energy among renewable energy sources has considerable attention over the last half-century. Wind is one of clean and inexhaustible energy sources. Thus, public and private sector investments in wind energy are increasing day by day. It is well-known that it is necessary to observe long term wind speed data to understand whether there is the potential of wind energy of the specified region. Generally, observed empirical wind speed data is modeled by different statistical distributions such as exponential, Gamma, Beta, Weibull etc. [1-8]. It is well-known that the Weibull distribution is widely used reference distribution in the literature [1-3, 9]. Moreover, in some wind energy software, wind power is estimated based on the Weibull distribution [9]. Thus, since estimating wind energy potential is depend on the Weibull distribution, correctly estimation of its parameters is crucial. Several estimation methods have been used and proposed for the Weibull distribution such as maximum likelihood method, moment method, empirical method, modified maximum likelihood method and power method [9-12]. For example, Seguro and Lambert [10] claim that when the wind speed data is provided in time series format, the maximum likelihood (ML) estimation method can be used. [11-12] have compared some estimation methods for wind speed data.
As different previous studies, [9] have recently proposed power density method for estimating the parameters of Weibull distribution.

In literature, these estimation methods are handled from different angles. On the other hand, if there are unusual or outlier wind observations in data due to mistakes in measuring and recording data, the aforementioned estimation methods can produce non-reliable results. These unusual wind speed observations are sometimes called contaminated data. In such cases, robust estimation methods can be used to check reliability of the estimated parameters.

In this paper, we research the performance of robust estimation methods of the Weibull distribution for clear wind speed data and for contaminated (unusual) wind speed data. It should be expressed that preliminary study of this paper was presented in International Conference on Engineering & MIS (ICEMIS2015), Turkey, Istanbul, 24 - 26 September 2015.

The probability density function (pdf) and cumulative distribution function (cdf) of Weibull random variable is given as follows:

$$f(x)=\frac{\alpha}{\gamma}\left(\frac{x}{\gamma}\right)^{\alpha-1}\exp\left[-\left(\frac{x}{\gamma}\right)^{\alpha}\right],\quad x,\gamma,\alpha>0 \tag{1}$$

$$F(x)=1-\exp\left[-\left(\frac{x}{\gamma}\right)^{\alpha}\right] \tag{2}$$

where α is shape parameter and γ is scale parameter.

The objective of this study is to evaluate a number of robust estimation methods in estimating parameters of the Weibull distribution by using real wind speed data and also generated unusual wind speed observation. Considering all these issues, the remainder of this study is organized as follows: In Section 2, the commonly-used estimation methods are recalled. In section 3, the considered robust estimation methods are introduced briefly. The results of analysis based on real wind speed data are provided in section 4. Finally section 5 finishes the study by giving the conclusions and suggestions.

2. COMMONLY USED ESTIMATION METHODS

In this section, the most commonly used methods for parameter estimation of the Weibull distribution in wind energy field are recalled.

2.1 Maximum Likelihood Method

Maximum Likelihood (ML) estimates are obtained by maximizing the log-likelihood (LL) function of a sample from Weibull distribution. LL function for Weibull distribution is obtained as follows:

$$\log L(\gamma,\alpha|x_1,x_2,\ldots,x_n)=n\log\alpha-n\log\gamma+(\alpha-1)\sum_{i=1}^{n}\log\frac{x_i}{\gamma}-\sum_{i=1}^{n}\left(\frac{x_i}{\gamma}\right)^{\alpha} \tag{3}$$

ML estimates of parameters are calculated from the solution of the following equations obtained from derivatization of LL with respect to the parameters γ and α :

$$-\frac{n}{\gamma}+\frac{1}{\gamma^2}\sum_{i=1}^{n}x_i^{\alpha}=0 \tag{4}$$

$$\frac{n}{\alpha}+\sum_{i=1}^{n}\log x^{i}-\frac{1}{\gamma}\sum_{i=1}^{n}x_i^{\alpha}\log x_i=0 \tag{5}$$

It is reached from the equations (4) and (5) that we have to use numerical method to obtain the estimates [13].

2.2 Least Squares Method

Least squares (LS) estimators of the Weibull distribution are based on the linear form of $F(x)$ given in (6).

$$\begin{gathered}F(x)=1-\exp\left[-\left(\frac{x}{\gamma}\right)^{\alpha}\right]\\ \left(1-F(x)\right)=\exp\left[-\left(\frac{x}{\gamma}\right)^{\alpha}\right]\\ \log\left(-\log\left(1-F(x)\right)\right)=\alpha\log x-\alpha\log\gamma\end{gathered} \tag{6}$$

For sample size n, let $x_{(1)},x_{(2)},\ldots,x_{(n)}$ be order statistics and then, the linear regression model can be written as follows:

$$\log\left(-\log\left(1-F\left(x_{(i)}\right)\right)\right)=\alpha\log x_{(i)}-\alpha\log\gamma \tag{7}$$

To estimate $F(x)$, we use Bernard's median rank estimator given as following:

$$\hat{F}_i=\frac{i-0.3}{n+0.4} \tag{8}$$

The equation (7) can be converted to linear regression model as follows:

$$\begin{gathered}Y_i=\log\left(-\log\left(1-F\left(x_{(i)}\right)\right)\right),\quad X_i=\log x_{(i)}\\ \beta_0=-\alpha\log\gamma,\quad \beta_1=\alpha\end{gathered} \tag{9}$$

Finally the classical linear regression model can be written as following:

$$Y_i=\beta_0+\beta_1X_i+\varepsilon_i. \tag{10}$$

After estimating the parameters of linear regression model, the shape parameter α and the scale parameter γ can be obtained as follows:

$$\begin{gathered}\hat{\alpha}=\hat{\beta}_1\\ \hat{\gamma}=\exp\left(-\frac{\hat{\beta}_0}{\hat{\alpha}}\right)\end{gathered} \tag{11}$$

It is known that LS estimate can be calculated by minimizing the sum of squares of the errors in the equation (10). Thus, LS estimates are,

$$\hat{\beta}_1=\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})^2} \tag{12}$$

$$\hat{\beta}_0=\bar{Y}-\hat{\beta}_1\bar{X}. \tag{13}$$

3. ROBUST ESTIMATION METHODS

In this section, some robust estimation methods that could be used in wind energy field for the parameters of Weibull distribution, are introduced. In this estimation procedure, the linear form (7) of F(x) will be used.

3.1 M-Estimators

This estimator proposed by Huber [15] minimizes the sum of a less rapidly increasing function of residuals. The general form of M-estimator is,

$$\min \sum_{i=1}^{n} \rho\left(\frac{e_i}{\hat{\sigma}}\right) \tag{14}$$

where ρ is objective function of residuals ($e_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$) and $\hat{\sigma}$ is robust scale estimate to make $\hat{\beta}$ scale equivariant [16, 17]. One of the popular choice for scale estimate is mean absolute deviation (MAD) calculated as,

$$\frac{median\{|e_i - median(e_i)|\}}{0.6745} \tag{15}$$

While n is large and the error distribution is normal, the tuning constant 0.6745 provides nearly unbiased estimator for scale [17]. In the literature, S_n and Q_n are alternative for robust scale estimate, that are more efficient or higher breakdown point than MAD [18]. In this study, we use MAD for initial scale estimate due to the widely usage and ease of use.

Let $\psi = \rho'$ be the derivative of ρ. System of two equations are produced after differentiating the objective function with respect to the coefficients and setting the partial derivatives to 0. Solution of following equations gives regression coefficients:

$$\sum_{i=1}^{n} x_i' \psi\left(\frac{y_i - x_i'\beta}{\hat{\sigma}}\right) = 0 \tag{16}$$

An iterative solution method, called iteratively reweighted least squares (IRWL), is used to solve equation (16) [17].
In robust literature, there are several objective functions which have been proposed for M-estimations. According to the behavior of ψ function, these estimators are classified as hard redescending, soft redescending and monotone. In this study we use Turkeys' Bisquare (soft redescending), Huber (monotone) and Cauchy (soft redescending) objective functions whose formulations are given respectively:

$$\rho(x) = \begin{cases} 1 - \left[1 - \left(\frac{x}{c}\right)^2\right]^3 & \text{if } |x| \le c \\ 1 & \text{if } |x| > c \end{cases} \tag{17}$$
$$c = 4.685$$

$$\rho(x) = \begin{cases} x^2 & if\ |x| \le c \\ 2c|x| - c^2 & if\ |x| > c \end{cases} \tag{18}$$
$$c = 1.345$$

$$\rho(x) = \frac{c^2}{2}\log\left[1 + (x/c)^2\right] \tag{19}$$
$$c = 2.385$$

where c is a tuning constant that plays a vital role to determine robustness and efficiency of an estimator [16].

3.2 Least Trimmed Squares

The estimates based on least trimmed squares (LTS) method proposed by Rousseeuw [21] are calculated by minimizing the sum of the trimmed squared residuals $(e_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 \mathrm{x}_i)$ The LTS estimator is,

$$\min \sum_{i=1}^{h} e_{(i)}^{\ 2} \tag{20}$$

where $e_{(i)}$ is ith ordered squared residual $h = [n(1-\theta)+1]$ is the number of observations that incorporated in the calculation of the estimator and proportion of trimming is provided as θ .

3.3 Least Median of Squares

Similar to LTS, the least median of squares (LMS) method proposed by Rousseeuw [20] is operated by minimizing the median of the squared residuals. The LMS estimator is,

$$\min\left(median\left(e_i^2\right)\right) \tag{21}$$

It is seen that while LS uses sum, LMS uses median. Thus, since median is more robust than sum, resulting estimates are resistant to outliers.

3.4 Least Absolute Deviation

The least absolute deviation (LAD) estimators are obtained via minimizing the absolute values of residuals, as follows:

$$\min \sum_{i=1}^{h} |e_i| \tag{22}$$

Although this estimator is more robust then LS against to the y-outliers, it is non-robust against to the x- outliers.

3.5 Quantile Estimator

The quantile function of Weibull distribution is,

$$x = \gamma\left[-\ln\left(1-F\left(x\right)\right)\right]^{1/\alpha} \quad (23)$$

For the percentiles P_1 and P_2, the quantile values are given as follows:

$$x_{P_1} = \gamma\left[-\ln\left(1-P_1\right)\right]^{1/\alpha} \quad (24)$$

$$x_{P_2} = \gamma\left[-\ln\left(1-P_2\right)\right]^{1/\alpha} \quad (25)$$

The estimators can be obtain via equations (24) and (25). It is clear that the number of equations must be equal to the number of unknown parameters [25].

3.6 Repeated Median (RM)

The repeated median (RM) estimate of β_1 is calculated as follows [19],

$$\hat{\beta}_1 = median_j median_{i\neq j} \frac{y_j - y_i}{x_j - x_i} \quad (26)$$

and the estimator of the coefficient β_0 is,

$$\hat{\beta}_0 = median_i\left(Y_i - \hat{\beta}_1 X_i\right). \quad (27)$$

4. ANALYSIS AND RESULTS

To evaluate the considered robust estimation methods, hourly mean wind speed data measured at 10 m above ground level in Izmir, Turkey is used. Evaluation criteria are taken as chi square (χ^2), the coefficient of determination (R^2) and root mean square error (RMSE). Their formulation are provided as follows:

$$\chi^2 = \frac{\sum_{j=1}^{n}\left(y_j - x_j\right)}{n-2} \quad (28)$$

$$R^2 = 1 - \frac{\sum_{j=1}^{n}\left(y_j - x_j\right)^2}{\sum_{j=1}^{n}\left(y_j - \bar{y}_j\right)^2} \quad (29)$$

$$RMSE = \left[\frac{1}{n}\sum_{j=1}^{n}\left(y_j - x_j\right)^2\right]^{\frac{1}{2}} \quad (30)$$

where n is the number of observations, y_i is the observed relative frequency of wind speed data, x_i is the relative frequency based on Weibull distribution and $\bar{y}$ is the average of y_i values [22].

The corresponding wind power density based on estimated parameters is calculated to evaluate the methods in terms of estimating wind energy potential. To make comparison, firstly reference mean wind power density is calculated based on the real wind speed values by using the formula given below:

$$P_{ref} = \frac{1}{2}\rho A v^3 f\left(v\right) \quad (31)$$

where ρ is air density (kg/m^3), v is wind speed and A is wind turbine blade sweep area (m^2). After calculating reference mean wind power density, the mean wind power density based on Weibull distribution can be calculated by the formula given following [23]:

$$P_W = \frac{1}{2}\rho\gamma^3\Gamma\left[1+\frac{1}{\alpha}\right] \quad (32)$$

where $\Gamma[\;]$ is the gamma function. The wind power density values calculated by the equations (31) and (32) are given in the Table 6.

ML, LS, LTS, LMS, LAD, Tukey, Huber, Cauc., Quan., RM estimates are given Table 1-5. All computations are performed in MATLAB and LIBRA library is used for LTS and LMS estimators.

Results of criteria are also provided in Table 1-5. The following conclusions can be derived from these tables. Most of M-estimators provide good performance for most of the considered months in terms of χ^2 criterion. According to R^2 criterion, all considered robust methods show comparable performance relative to ML and LS. M-estimation is generally better than ML and LS methods in terms of the considered criteria. Besides M-estimation methods, good performance of LAD is seen in almost all Tables. Thus, M-estimators and LAD can be alternatively used for modelling wind speed data. Additionally, among the robust methods with the property of high breakdown, LTS, LMS and RM provides comparable performance according to ML and LS which are non-robust methods.

Table 1. Estimates of parameters of Weibull distribution and the model selection criteria for monthly wind speed data (January-March).

Est.	γ	α	χ^2	R^2	*RMSE*
		January			
ML	2.5590	1.8044	0.0008	0.9546	0.0238
LS	2.5183	2.0284	0.0011	0.9471	0.0277
LTS	2.4793	1.6127	0.0013	0.9257	0.0306
LMS	2.4797	1.5972	0.0014	0.9205	0.0317
LAD	2.4910	1.7260	0.0008	0.9520	0.0244
Tukey	2.5125	1.7854	0.0007	0.9573	0.0230
Huber	2.5163	1.8567	0.0007	0.9587	0.0227
Cauc.	2.5179	1.8877	0.0007	0.9578	0.0231
Quan.	2.6824	1.8601	0.0012	0.9331	0.0291
RM	2.4966	1.6621	0.0011	0.9403	0.0275
		February			
ML	2.3740	1.8607	0.0006	0.9770	0.0204
LS	2.3381	2.1037	0.0003	0.9881	0.0150
LTS	2.3048	1.7652	0.0010	0.9576	0.0268
LMS	2.2954	1.7832	0.0009	0.9609	0.0256
LAD	2.3163	1.8940	0.0004	0.9804	0.0179
Tukey	2.3333	1.8911	0.0005	0.9803	0.0181
Huber	2.3311	1.9595	0.0003	0.9859	0.0151
Cauc.	2.3303	2.0006	0.0003	0.9879	0.0141
Quan.	2.4454	2.0140	0.0005	0.9765	0.0198
RM	2.3155	1.8283	0.0007	0.9712	0.0220
		March			
ML	2.6892	1.8441	0.0010	0.9464	0.0288
LS	2.6451	2.1554	0.0008	0.9560	0.0250
LTS	2.5975	1.8279	0.0010	0.9471	0.0279
LMS	2.5425	1.9464	0.0007	0.9611	0.0234
LAD	2.5771	1.9893	0.0007	0.9628	0.0229
Tukey	2.6192	1.9152	0.0008	0.9562	0.0252
Huber	2.6222	2.0132	0.0007	0.9607	0.0236
Cauc.	2.6184	2.0751	0.0007	0.9614	0.0233
Quan.	2.7824	2.0737	0.0011	0.9375	0.0299
RM	2.5860	1.9379	0.0007	0.9593	0.0241

Table 2. Estimates of parameters of Weibull distribution and the model selection criteria for monthly wind speed data (April-June)

Est.	γ	α	χ^2	R^2	*RMSE*
		April			
ML	2.5033	1.9043	0.0014	0.9364	0.0325
LS	2.4647	2.1562	0.0013	0.9391	0.0318
LTS	2.4223	1.7257	0.0021	0.9042	0.0399
LMS	2.4048	1.6528	0.0027	0.8783	0.0448
LAD	2.4499	1.8908	0.0014	0.9361	0.0323
Tukey	2.4658	1.9590	0.0013	0.9408	0.0309
Huber	2.4658	2.0266	0.0012	0.9424	0.0304
Cauchy	2.4647	2.0663	0.0012	0.9423	0.0305
Quantile	2.6821	2.0609	0.0020	0.9104	0.0384
RM	2.4442	1.7852	0.0018	0.9201	0.0365
		May			
ML	2.4611	2.1262	0.0004	0.9842	0.0153
LS	2.4334	2.2982	0.0006	0.9808	0.0204
LTS	2.4294	1.9724	0.0006	0.9762	0.0197
LMS	2.4216	1.9520	0.0007	0.9730	0.0210
LAD	2.4291	2.0755	0.0004	0.9836	0.0157
Tukey	2.4368	2.0963	0.0004	0.9843	0.0153
Huber	2.4369	2.1558	0.0004	0.9848	0.0154
Cauchy	2.4367	2.1886	0.0004	0.9844	0.0161
Quantile	2.5840	2.1336	0.0007	0.9709	0.0217
RM	2.4350	2.0240	0.0004	0.9811	0.0171
		June			
ML	2.3766	1.9104	0.0003	0.9870	0.0138
LS	2.3549	2.0290	0.0004	0.9838	0.0177
LTS	2.3383	1.7133	0.0007	0.9682	0.0216
LMS	2.3253	1.6934	0.0008	0.9624	0.0234
LAD	2.3336	1.8375	0.0003	0.9840	0.0152
Tukey	2.3483	1.8676	0.0003	0.9862	0.0141
Huber	2.3505	1.9257	0.0003	0.9873	0.0141
Cauchy	2.3523	1.9550	0.0003	0.9869	0.0148
Quantile	2.5412	1.9471	0.0006	0.9702	0.0211
RM	2.3410	1.7756	0.0004	0.9785	0.0176

Table 3. Estimates of parameters of Weibull distribution and the model selection criteria for monthly wind speed data (July-September)

Est.	γ	α	χ^2	R^2	*RMSE*
July					
ML	2.7455	1.9914	0.0007	0.9586	0.0225
LS	2.7312	2.0440	0.0007	0.9552	0.0238
LTS	2.7412	1.8677	0.0006	0.9590	0.0223
LMS	2.7675	1.8293	0.0007	0.9588	0.0227
LAD	2.7543	1.9522	0.0006	0.9603	0.0219
Tukey	2.7357	1.9522	0.0006	0.9595	0.0222
Huber	2.7335	1.9844	0.0007	0.9585	0.0226
Cauchy	2.7374	1.9998	0.0007	0.9580	0.0228
Quantile	2.7692	2.0023	0.0006	0.9587	0.0224
RM	2.7568	1.9092	0.0006	0.9608	0.0218
August					
ML	2.6174	2.0879	0.0005	0.9745	0.0191
LS	2.5881	2.2602	0.0007	0.9679	0.0230
LTS	2.5785	1.9400	0.0007	0.9690	0.0218
LMS	2.5692	1.8634	0.0009	0.9580	0.0258
LAD	2.5889	2.0255	0.0005	0.9747	0.0192
Tukey	2.5898	2.0579	0.0005	0.9753	0.0188
Huber	2.5900	2.1223	0.0005	0.9749	0.0190
Cauchy	2.5911	2.1499	0.0005	0.9740	0.0195
Quantile	2.6723	2.0678	0.0006	0.9719	0.0206
RM	2.5950	1.9685	0.0006	0.9722	0.0207
September					
ML	2.5357	2.0855	0.0003	0.9860	0.0145
LS	2.5124	2.2055	0.0004	0.9840	0.0166
LTS	2.5123	1.9780	0.0004	0.9817	0.0172
LMS	2.5145	1.9741	0.0004	0.9815	0.0174
LAD	2.5242	2.0428	0.0003	0.9854	0.0150
Tukey	2.5189	2.0547	0.0003	0.9856	0.0148
Huber	2.5177	2.0926	0.0003	0.9861	0.0144
Cauchy	2.5220	2.1081	0.0003	0.9861	0.0145
Quantile	2.6234	2.1035	0.0004	0.9807	0.0176
RM	2.5220	1.9945	0.0004	0.9833	0.0165

Table 4. Estimates of parameters of Weibull distribution and the model selection criteria for monthly wind speed data (October-December).

Est.	γ	α	χ^2	R^2	*RMSE*
October					
ML	2.1432	1.7642	0.0007	0.9742	0.0222
LS	2.1069	2.0371	0.0007	0.9757	0.0225
LTS	2.0534	1.8102	0.0004	0.9819	0.0182
LMS	2.0230	1.7459	0.0006	0.9754	0.0212
LAD	2.0518	1.8767	0.0004	0.9853	0.0168
Tukey	2.0916	1.7900	0.0005	0.9790	0.0196
Huber	2.0833	1.9098	0.0004	0.9833	0.0178
Cauchy	2.0809	1.9569	0.0005	0.9830	0.0185
Quantile	2.2650	1.9164	0.0012	0.9522	0.0297
RM	2.0616	1.8273	0.0004	0.9827	0.0178
November					
ML	2.5396	1.7202	0.0003	0.9804	0.0155
LS	2.5013	1.9104	0.0004	0.9801	0.0173
LTS	2.4614	1.5603	0.0008	0.9507	0.0246
LMS	2.4526	1.5398	0.0009	0.9432	0.0263
LAD	2.4839	1.6508	0.0004	0.9730	0.0182
Tukey	2.4974	1.7184	0.0003	0.9812	0.0151
Huber	2.5002	1.7939	0.0003	0.9843	0.0141
Cauchy	2.5024	1.8287	0.0003	0.9839	0.0146
Quantile	2.6824	1.8601	0.0007	0.9588	0.0224
RM	2.4820	1.5984	0.0006	0.9625	0.0215
December					
ML	2.1589	1.8639	0.0015	0.9424	0.0320
LS	2.1263	2.1064	0.0019	0.9406	0.0351
LTS	2.0639	1.5758	0.0036	0.8667	0.0487
LMS	2.0474	1.4432	0.0058	0.7838	0.0619
LAD	2.1367	1.7482	0.0019	0.9309	0.0354
Tukey	2.1225	1.8932	0.0014	0.9459	0.0310
Huber	2.1230	1.9628	0.0015	0.9471	0.0311
Cauchy	2.1290	1.9827	0.0015	0.9461	0.0315
Quantile	2.3970	2.0507	0.0036	0.8647	0.0490
RM	2.1115	1.6120	0.0030	0.8909	0.0445

Table 5. Estimates of parameters of Weibull distribution and the model selection criteria for yearly wind speed data.

Est.	γ	α	χ^2	R^2	*RMSE*
Whole Year					
ML	2.4734	1.8811	0.0004	0.9803	0.0169
LS	2.4367	2.1016	0.0005	0.9774	0.0191
LTS	2.4097	1.7321	0.0006	0.9658	0.0224
LMS	2.4075	1.7274	0.0006	0.9649	0.0227
LAD	2.4288	1.8306	0.0004	0.9781	0.0178
Tukey	2.4389	1.8983	0.0003	0.9816	0.0162
Huber	2.4391	1.9659	0.0003	0.9823	0.0160
Cauchy	2.4402	1.9955	0.0003	0.9818	0.0163
Quantile	2.5885	1.9164	0.0005	0.9723	0.0204
RM	2.4293	1.7637	0.0005	0.9713	0.0206

Table 6 provides the wind power density based on the estimated parameters of Weibull distribution. Comparisons of ML and robust methods according to the reference mean wind power density value show that ML provides closer wind power to reference power value than the other methods. Next to ML, the considered M-estimators calculate mean power density correctly.

To evaluate robustness of the estimators for wind speed data, artificially large wind speed data is created. In other words, we modified 5 % of the right side of the data by adding $\mu+3\sigma$ for generating unusual wind speed data. μ and σ are respectively the mean and standard deviation of the original wind speed data. The obtained wind power density results for the modified data sets are given in Table 7.

Table 6. The estimated mean wind power density values and reference wind power.

	P_ref	*P_ml*	*P_ls*	*P_lts*	*P_lms*	*P_lad*	*P_tukey*	*P_huber*	*P_cauchy*	*P_quant*	*P_rm*
Jan.	15.6915	15.3940	12.8146	16.4757	16.7434	15.0804	14.7739	14.1180	13.8653	17.0641	16.0460
Feb.	12.1107	11.8248	9.8937	11.5795	11.2836	10.7520	11.0111	10.5441	10.2996	11.8203	11.2144
March	18.5097	17.3800	14.0063	15.8340	13.7824	14.0152	15.3466	14.5798	14.0800	16.9057	14.5730
April	13.7299	13.4870	11.3273	13.8685	14.4631	12.7469	12.4826	12.0402	11.7913	15.2344	13.6038
May	11.2232	11.4234	10.3298	11.8497	11.8706	11.2394	11.2379	10.9501	10.8022	13.1796	11.6097
June	11.3090	11.4979	10.4752	12.6043	12.6078	11.4065	11.3943	11.0225	10.8611	13.7553	12.0389
July	16.6190	16.9273	16.2195	18.1208	19.1325	17.4629	17.1125	16.7692	16.7037	17.2689	17.9577
Aug.	13.7939	13.9792	12.5949	14.4293	14.9618	13.9426	13.7360	13.3372	13.1954	15.0194	14.4726
Sep.	12.4976	12.7248	11.7611	13.0648	13.1266	12.8128	12.6579	12.4153	12.3918	13.9763	13.0986
Oct.	9.7701	9.3169	7.4721	7.9210	7.9491	7.5554	8.4944	7.7471	7.5113	9.9165	7.9202
Nov.	16.4699	16.0562	13.4047	17.0298	17.2383	15.9691	15.2927	14.4667	14.1503	17.0641	16.7698
Dec.	8.8958	8.8739	7.4314	9.8732	11.3189	9.3478	8.2774	7.9500	7.9307	10.9290	10.1838
Year	13.3472	13.1979	11.2095	13.5836	13.5971	12.9207	12.5177	12.0356	11.8590	14.8024	13.5739

Table 7. The mean wind power density values for 5% contaminated data.

	P_ref	*P_ml*	*P_ls*	*P_lts*	*P_lms*	*P_lad*	*P_tukey*	*P_huber*	*P_cauchy*	*P_quant*	*P_rm*
Jan.	15.6915	37.5687	20.4670	16.2435	16.8609	15.6192	15.6522	15.6090	15.2608	17.0641	16.1792
Feb.	12.1107	29.2530	15.8517	11.2843	11.2836	10.9852	11.1276	10.9200	10.6815	11.8203	11.2204
March	18.5097	42.6058	22.5282	15.3851	13.7824	14.3711	14.8950	15.0943	15.0647	16.9057	14.5730
April	13.7299	32.7640	18.0985	13.5207	14.4142	13.1517	13.1240	13.1402	13.1255	15.2344	13.8333
May	11.2232	26.9169	16.1134	11.9064	11.9926	11.6355	11.7085	11.8096	11.7357	13.1796	11.8188
June	11.3090	27.7735	16.3906	12.8241	12.6872	12.3744	12.5545	12.4473	12.3746	13.7553	12.5288
July	16.6190	40.3481	24.8295	18.2529	19.1098	18.2457	17.7401	18.1179	18.3327	17.2689	18.6169
Aug.	13.7939	33.2406	19.6847	14.4443	14.9124	14.5102	14.3065	14.3137	14.0824	15.0194	14.7594
Sep.	12.4976	30.4075	18.2854	13.2763	13.2801	13.1875	12.9494	13.1660	13.0705	13.9763	13.4229
Oct.	9.7701	22.8985	11.8919	7.8753	7.9491	7.8004	7.8095	8.0121	8.3918	9.9165	7.9202
Nov.	16.4699	39.4908	21.1761	16.9734	17.1392	16.4657	16.5930	16.0629	15.6694	17.0641	16.8914
Dec.	8.8958	21.5738	12.0651	9.9257	11.3189	10.1502	8.7943	9.5889	9.5450	10.9290	10.7828
Year	13.3472	31.7370	17.5509	13.5420	13.8234	13.2692	13.3412	13.1097	12.8449	14.8024	13.7637

5. CONCLUSIONS

Some robust estimation methods of the Weibull distribution are introduced for wind energy applications. The considered robust methods are evaluated in terms of estimating parameters of Weibull distribution and wind power. It is based on the analyses, the main findings of this study are listed as follows:

1. Some of robust methods provide good performance for both clear wind speed data and contaminated wind speed data.
2. Particularly M-estimation performs better than classical estimation methods.
3. Also, LAD estimator provides good performance for clear and contaminated wind speed data cases.
4. It is also observed that the calculated wind power density values based on classical estimation methods change with contaminated wind data.
5. As a result, we suggest that robust methods not only can be used to estimate wind power for real wind speed, but also can simultaneously be used with classical efficient estimators to check the reliability of the estimated wind power results.

5. ACKNOWLEDGMENT

This study was supported by Anadolu University Scientific Research Projects Commission under the grant no: 1506F532

6. REFERENCES

[1] S. Akdag and O. Guler, "Calculation of Wind Energy Potential and Economic Analysis by Using Weibull Distribution. A Case Study from Turkey. Part 1: Determination of Weibull Parameters." *Energy sources part B-economics planning and policy,* 4: 1–8, 2009.

[2] D. Weisser, "A wind energy analysis of Grenada: an estimation using the 'Weibull' density function." *Renewable Energy*, 28:1803–1812, 2003.

[3] K. Philippopoulos, D. Deligiorgi and G. Karvounis, "Wind speed distribution modeling in the Greater Area of Chania." *Greece. Int J of Green Energy*, 9(2):174–93, 2012.

[4] J. A. Carta, P. Ramirez and S. Velazquez, "A review of wind speed probability distributions used in wind energy analysis – case studies in the Canary Islands." *Renew. Sustain Energy Rev.,* 13:933–55, 2009.

[5] V. T. Morgan, "Statistical distributions of wind parameters at Sydney, Australia." *Renew. Energy,* 6(1); 39–47, 1995.

[6] I. Usta and Y. M. Kantar, "Analysis of some flexible families of distributions for estimation of wind speed distributions." *Appl. Energy*, 89(1): 355-367, 2012.

[7] J. Zhou, E. Erdem, G. Li, and J. Shi, "Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites." *Energy Convers. Manage.,* 51(7):1449–1458, 2010.

[8] T. P. Chang, "Estimation of wind energy potential using different probability density functions." *Appl. Energy*, 88(5):1848-1856, 2011.

[9] S. Akdag and A. Dinler, "A new method to estimate Weibull parameters for wind energy applications." *Energy Convers. Manage.* 2009: 50: 1761–1766, 2009.

[10] J. V. Seguro and T. W. Lambert, "Modern estimation of the parameters of the Weibull probability density distribution." *J. Wind Eng. Ind. Aerodyn.,* 85:75–84, 2000.

[11] T. P. Chang, "Performance comparison of six numerical methods in estimating Weibull parameters for wind energy application." *Appl. Energy,* 88:272–82, 2011.

[12] A. Genc, M. Esiroglu, A. Pekgor, G. Oturanc, A. Hepasli and K. Ulgen "Estimation of the wind power potential using Weibull distribution." *Energy Sources,* 27: 809–22, 2005.

[13] Y. M. Kantar and B. Senoglu, "A comparative study for the location and scale parameters of the Weibull distribution with given shape parameter." *Computers and Geosciences,* 34:1900–1909, 2008.

[14] M. Tiryakioglu and D. Hudak, On estimating "Weibull modulus by the linear regression method." *J. Mater. Sci.* 42:10173–10179, 2007.

[15] P. J. Huber, Robust estimation of a location parameter. *Annals of Mathematical Statistics,* 35, 73-101, 1964.

[16] R. A. Maronna, D. R. Martin and V. J. Yohai, **Robust Statistics Theory & Methods**, Wiley: England, 2006.

[17] D. C. Montgomery, E. A. Peck and G. G. Vining, **Introduction to Linear Regression Analysis***,* John Wiley & Sons: New York, 2001.

[18] P. J. Rousseuw and C. Croux, "Alternatives to the Median Absolute Deviation." *Journal of the American Statistical Association*, 88, 424, 1273-1283, 1993.

[19] A. F. Siegel, "Robust regression using repeated medians." *Biometrika*, 69, 242-244, 1982.

[20] P. J. Rousseeuw, "Least Median of Squares Regression." *Journal of American Statistical Association*, 79, pp. 871-880, 1984.

[21] P. J. Rousseeuw, A. M. Leroy and B. Daniels, "Resistant line fitting in actuarial science." *Advanced Science Institutes Series C*, 88, pp. 315-332, 1984.

[22] P. A. C . Rocha, R. C. de Sousa, C. F. de Andrade and M. E. V. da Silva, "Comparison of seven numerical methods for determining Weibull parameters for wind energy generation in the northeast region of Brazil." *Applied Energy*, 89(1), 395-400, 2012.

[23] A. N. Celik, "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, 2004.

[24] L. F. Zhang, M. Xie and L. C. Tang, Robust regression using probability plots for estimating the Weibull shape parameter. *Qual. and Rel. Eng. Int.*, 22:905-917, 2006.

[25] N. B. Marks, "Estimation of Weibull Parameters from Common Percentiles", Journal of Applied Statistics, 32, 1, 17-24, 2005.

[26] Internet: LIBRA, http://wis.kuleuven.be/stat/robust.html.