Determination of Wind Potential by Two Components Mixture Probability Distribution Models in the Ankara, Turkey

Tayfun Servi*, Selim Gündüz**‡, Ülkü Erişoğlu***, Levent Yalçın****

*Department of Economics, Faculty of Economics and Administrative Sciences, Associate Professor, Adıyaman University, Adıyaman, 02040, Turkey

** Department of Business Administration, Faculty of Business, Assistant Professor, Adana Alparslan Türkeş Science and Technology University, Adana, 01250, Turkey

*** Department of Statistics, Faculty of Science, Associate Professor, Necmettin Erbakan University, Konya, 42090, Turkey

****Department of Forecasts Turkish State Meteorological Service, Ph.D., Department of Forecasts Turkish State Meteorological Service Ankara, 06120, Turkey

(tservi@adiyaman.edu.tr, sgunduz@atu.edu.tr, ugokal@erbakan.edu.tr, lyalcin@mgm.gov.tr)

‡ Corresponding Author; Selim Gündüz, Adana Alparslan Türkeş Science and Technology University, Adana, 01250, Turkey,

Tel: +90 322 455 0041-internal (2248), Fax: +90 322 455 0042, sgunduz@atu.edu.tr

Received: 05.10.2020 Accepted:23.11.2020

Abstract-In this study, hourly average wind speed data in the Ankara, Turkey are modeled with Weibull, Gamma and Rayleigh probability distribution and theirs two component mixture probability distributions. Expectation-Maximization (EM) algorithm is introduced for Maximum Likelihood Estimation (MLE) of mixture probability distributions used in modeling wind speed data. In comparing the modeling performances of probability distributions, the Akaike information criteria, the coefficient of determination, the root of the mean squares and chi-square criteria were used as comparison criteria. Also, in this study, the success in estimation of wind potential was evaluated with relative error. In the study results, it was observed that the mixture distribution models obtained from two different distributions were more successful in modeling wind speed data. The results obtained from the study revealed that the wind potential of Keçiören/Bağlum region is higher than Çankaya/Çaldağ region. According to the wind speed data observed in Keçiören/Bağlum and Çankaya/Çaldağ regions, wind power densities per unit area were calculated as $153.926\ \mathrm{W/m^2}$ and $62.785\ \mathrm{W/m^2}$, respectively.

Keywords Gamma-Rayleigh, relative error, Weibull-Gamma, Weibull-Rayleigh, wind speed

1. Introduction

Due to the population increase, growing industries and developing technologies, human beings need for electricity has been increasing [1]. Electrical energy, which is one of the most important consumption materials of today, is indispensable for the development and welfare of countries [2]. Most of the electricity we use now is produced from the fossil fuels [3].

The fact that the fossil resources are limited, decreasing day by day, being exhausted one day, caused the use of renewable energy sources in electrical energy production on account of their negative effects on natural vegetation, air and human health [4]. For this reasons, studies on renewable energy has been attracted more attention in the literature.

Many countries are invested in renewable energy sources in electricity generation to get rid of sustainability and foreign dependency. Wind energy is the fastest growing energy source in the world with its many advantages [5]. Wind energy is not external dependent, not causing to produce atmospheric events like acid rains or greenhouse gases, negative effects on nature and human life. In addition, wind energy technological development is a rapid source of energy.

Due to the fact that construction of wind turbine has a considerably high cost, exact estimation of wind potential is quite important [6]. The modeling with probability distributions of wind speed at a particular location and to determine the wind potential of the region depending on the model is a common approach [7, 8]. The most widely used probability distribution in determining wind energy potential is the Weibull distribution [9]. In the literature, Gamma and Rayleigh distributions are other well-known probability distributions in modelling of the wind speeds. Pishgar-Komleh et al. [10] modeled the wind speed data of Firouzkooh region in Iran with Weibull and Rayleigh distributions, while Jung and Schindler [11] used Weibull, Gamma and Rayleigh distributions in their study to evaluate the modeling performance of 21 different probability distributions in modeling wind speed data. Similarly, Mohammadi et al. [12] used Weibull and Rayleigh distributions to model the wind speeds in their study for examining the Birnbaum-Saunders distribution in modeling the wind speeds. Additionally, Aries et al. [13] determined the wind potential of four different regions in Algeria using Weibull and Gamma distributions. Finally, Bidaoui et al. [14] used Weibull and Rayleigh distributions to assess the wind energy potential of five major cities in Northern Morocco.

Mixture probability distributions are widely used models with ease of use and flexibility of mathematical structure in modeling of the wind speeds. Mixture probability distribution models have been suggested in a lot of studies for fitting wind speed data. Akpınar and Akpınar [15] used mixture distribution models to investigate an analysis of wind characteristics of four stations in Elazığ, Turkey. Shin et al. [16] used mixture distribution models for modelling heterogeneous wind speed data in the United Arab Emirates. Mazzeo et al. [17] used a mixture of two truncated normal distributions for modelling wind speed data measured at five Italian meteorological stations (Ancona, Cagliari, Naples, Reggio Calabria and Venice). In order to model wind speed in Quebec (Canada), Ouarda and Charron [18] examined the suitability of two-component mixture distribution models and emphasized that two-component mixture distribution models are successful in modeling wind speed data. Cook [19] showed that mixture distribution models are more successful than unimodal probability distributions in modeling wind speed data of four different regions.

The aim of this study is analyzed of hourly average wind speeds measured at Keçiören/Bağlum and Çankaya/Çaldağ meteorological stations by classical probability distributions and mixture probability distribution models. The classical probability distributions and mixture probability distribution models were used and constructed by mixing the three following probability distributions: Weibull (Wbl), Gamma (Gam) and Rayleigh (Rayl) distributions. The constructed mixture probability distribution models are included two component mixture Weibull (Wbl2), two component mixture Gamma (Gam2), two component mixture Rayleigh (Rayl2), Weibull-Gamma (WblGam), Weibull-Rayleigh (WblRayl) and Gamma-Rayleigh (GamRayl).

The rest of this study is organized as follows: the classical probability distributions and mixture probability distributions for modelling of wind speeds are defined in the section 2. Also, the parameter estimations of the defined wind distributions and performance criterions for comparison are presented in the section 2. The details about the data used for analysis and results derived from this study are discussed in section 3. Finally, some conclusions are noted in section 4.

2. Materials and Methods

The Weibull distribution is the most commonly used probability distribution in the studies which that using probability distribution for modeling wind speed in a region. Let V be continuous random variable for the hourly average wind speed. In this case, probability density function (pdf) and cumulative distribution function (cdf) of Weibull distribution is defined as [20]

$$f_{wbl}(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} e^{-\left(\frac{v}{c}\right)^k} \qquad v > 0 \; and \; k, c > 0 \tag{1}$$

$$F_{wbl}(v) = 1 - e^{-\left(\frac{v}{c}\right)^k} \tag{2}$$

where k and c are shown shape and scale parameters of Weibull distribution function respectively. The expected value (mean) and variance of Weibull distribution are defined $E_{wbl}(V) = c\Gamma\left(\frac{1}{k}+1\right)$ and $Var_{wbl}(V) = c^2\left\{\Gamma\left(\frac{2}{k}+1\right) - \Gamma\left(\frac{1}{k}+1\right)^2\right\}$ respectively. In these equations, $\Gamma(.)$ notation symbolizes gamma function.

The wind power density of the site per unit area considered based on any probability density function is defined as [15]

$$P(v) = \frac{1}{2}\rho A E(V^3) \tag{3}$$

where A is the wind turbine blade sweep area and ρ; is the air density and is calculated as 1.225 $\mathrm{kg/m^3}$. Wind power density by using Weibull distribution is calculated by [15]

$$P_{wbl}(v) = \frac{1}{2}\rho A c^3 \Gamma\left(\frac{3}{k}+1\right) \tag{4}$$

The pdf and cdf of the Gamma distribution, which is one of the distributions used in modeling the wind speed, are defined respectively as follows [20]

$$f_{gam}(v) = \frac{v^{a-1}}{b^a \Gamma(a)} e^{-\left(\frac{v}{b}\right)} \qquad v > 0 \; and \; a, b > 0 \tag{5}$$

$$F_{gam}(v) = \frac{\gamma\left(a, \frac{v}{b}\right)}{\Gamma(a)} \tag{6}$$

where a is a shape parameter and b is a scale parameter. In Eq. (6), $\gamma(.)$ notation is denoted an incomplete Gamma function and calculated by $\gamma\left(a, \frac{v}{b}\right) = \int_0^{v/b} v^{a-1} e^{-v} dv$.

The mean and variance for Gamma distribution are defined $E_{gam}(V) = ab$ and $Var_{gam}(V) = ab^2$ respectively.

Wind power density by using Gamma distribution is calculated as below [20].

$$P_{gam}(v) = \frac{1}{2}\rho A b^3 (a+2)(a+1)a \tag{7}$$

The pdf and cdf of the Rayleigh distribution which is a special case of Weibull distribution are defined by [14]

$$f_{rayl}s(v) = \frac{v}{\lambda^2} e^{-\left(\frac{v^2}{2\lambda^2}\right)} \qquad v > 0 \; and \; \lambda > 0 \tag{8}$$

$$F_{rayl}(v) = 1 - e^{-\left(\frac{v^2}{2\lambda^2}\right)} \tag{9}$$

where λ is a scale parameter. The mean and variance of the Rayleigh distribution are calculated by following equations $\mathrm{E_{rayl}(V)} = \lambda\sqrt{\frac{\pi}{2}}$ and $\mathrm{Var_{rayl}(v)} = \frac{(4-\pi)\lambda^2}{2}$. Wind power density by using Rayleigh distribution is calculated as below.

$$P_{rayl}(v) = \frac{1}{2}\rho A 3\lambda^3 \sqrt{\frac{\pi}{2}} \tag{10}$$

2.1. Modelling wind speed with two components mixture probability distributions

Finite mixture distribution models are using widely because of flexibility of the mathematical structure especially in the modelling of the heterogeneous datasets [21, 22]. In mixture distribution models, it is assumed that there are g subgroups with different characteristics in the population. Mixture distribution model is named two component mixture distribution models when it is obtained for two subgroups. The pdf and cdf for two component mixture distribution model are expressed respectively [21]

$$f_m(v) = p f_1(v, \theta_1) + (1-p) f_2(v, \theta_2) \tag{11}$$

$$F_m(v) = p F_1(v, \theta_1) + (1-p) F_2(v, \theta_2) \tag{12}$$

where $\mathrm{f_1(v, \theta_1)}$ and $\mathrm{f_2(v, \theta_2)}$ are pdf for each components with $\theta_1$and θ_2 parameter vectors. Similarly, $\mathrm{F_1(v, \theta_1)}$ and $\mathrm{F_2(v, \theta_2)}$ are cdf for each components. In the equations, p is mixture weight and $\mathrm{p} \in (0,1)$.

The pdf and wind power density of two component mixture Weibull distribution model consisting of two Weibull distributions with $(\mathrm{k_1, c_1})$ and $(\mathrm{k_2, c_2})$ parameter vectors are defined by [20]

$$f_{wbl2}(v) = \left\{ \begin{matrix} p \frac{k_1}{c_1}\left(\frac{v}{c_1}\right)^{k_1-1} e^{-\left(\frac{v}{c_1}\right)^{k_1}} \\ +(1-p)\frac{k_2}{c_2}\left(\frac{v}{c_2}\right)^{k_2-1} e^{-\left(\frac{v}{c_2}\right)^{k_2}} \end{matrix} \right\} \tag{13}$$

$$P_{wbl2}(v) = \frac{1}{2}\rho A \left\{ \begin{matrix} p c_1^3 \Gamma\left(\frac{3}{k_1}+1\right) \\ +(1-p) c_2^3 \Gamma\left(\frac{3}{k_2}+1\right) \end{matrix} \right\}. \tag{14}$$

The pdf and wind power density of two component mixture Gamma distribution model consisting of two Gamma distributions with $(\mathrm{a_1, b_1})$ and $(\mathrm{a_2, b_2})$ parameter vectors are defined by [20]

$$f_{gam2}(v) = \left\{ \begin{matrix} p \frac{v^{a_1-1}}{{b_1}^{a_1}\Gamma(a_1)} e^{-\left(\frac{v}{b_1}\right)} \\ +(1-p)\frac{v^{a_2-1}}{{b_2}^{a_2}\Gamma(a_2)} e^{-\left(\frac{v}{b_2}\right)} \end{matrix} \right\} \tag{15}$$

$$P_{gam2}(v) = \frac{1}{2}\rho A \left\{ \begin{matrix} p b_1^3 (a_1+2)(a_1+1)a_1 \\ +(1-p) b_2^3 (a_2+2)(a_2+1)a_2 \end{matrix} \right\} \tag{16}$$

The pdf and wind power density of the two component mixture Rayleigh distribution model consisting of two Rayleigh distributions with λ_1 and λ_2 parameters are defined as follows.

$$f_{rayl2}(v) = p\frac{v}{\lambda_1^2} e^{-\left(\frac{v^2}{2\lambda_1^2}\right)} + (1-p)\frac{v}{\lambda^2} e^{-\left(\frac{v^2}{2\lambda_2^2}\right)} \tag{17}$$

$$P_{rayl2}(v) = \frac{1}{2}\rho A 3\sqrt{\frac{\pi}{2}}\{p\lambda_1^3 + (1-p)\lambda_2^3\}. \tag{18}$$

The pdf and wind power density of two component mixture Weibull-Gamma (WblGam) distribution model consisting of the Weibull and Gamma distributions are expressed by

$$f_{wblgam}(v) = \left\{ \begin{matrix} p \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} e^{-\left(\frac{v}{c}\right)^{k}} \\ +(1-p)\frac{v^{a-1}}{b^a\Gamma(a)} e^{-\left(\frac{v}{b}\right)} \end{matrix} \right\} \tag{19}$$

$$P_{wblgam}(v) = \frac{1}{2}\rho A \left\{ \begin{matrix} p c^3 \Gamma\left(\frac{3}{k}+1\right) \\ +(1-p) b^3 (a+2)(a+1)a \end{matrix} \right\}. \tag{20}$$

The pdf, cdf and wind power density of two component mixture Weibull-Rayleigh (WblRayl) distribution model consisting of the Weibull and Rayleigh distributions are defined as follow

$$f_{wblrayl}(v) = \left\{ \begin{matrix} p \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} e^{-\left(\frac{v}{c}\right)^{k}} \\ +(1-p)\frac{v}{\lambda^2} e^{-\left(\frac{v^2}{2\lambda^2}\right)} \end{matrix} \right\} \tag{21}$$

$$P_{wblrayl}(v) = \frac{1}{2}\rho A \left\{ \begin{matrix} p c^3 \Gamma\left(\frac{3}{k}+1\right) \\ +(1-p) 3\lambda^3 \sqrt{\frac{\pi}{2}} \end{matrix} \right\}. \tag{22}$$

The pdf and wind power density of two component mixture Gamma-Rayleigh (GamRayl) distribution model consisting of the Gamma and Rayleigh distributions are defined by below.

$$f_{gamrayl}(v) = \left\{ \begin{matrix} p \frac{v^{a-1}}{b^a\Gamma(a)} e^{-\left(\frac{v}{b}\right)} \\ +(1-p)\frac{v}{\lambda^2} e^{-\left(\frac{v^2}{2\lambda^2}\right)} \end{matrix} \right\} \tag{23}$$

$$P_{gamrayl}(v) = \frac{1}{2}\rho A \left\{ \begin{matrix} p b^3 (a+2)(a+1)a \\ +(1-p) 3\lambda^3 \sqrt{\frac{\pi}{2}} \end{matrix} \right\} \tag{24}$$

2.2. Parameter estimation for two component mixture probability distributions

In finite mixture distribution model applications, the data is incomplete data because there is not component-label vector which that shown component memberships. Therefore, in the estimation of the parameters of the mixture distribution models, the estimation methods are applied with the Expectation Maximization (EM) algorithm [5]. In step E of the EM algorithm, the estimation of the component-label vector is performed. In step M, model parameters are estimated according to the component-label vector obtained in step E.

The elements of component-label vector z_i, which is the probability that the i-th unit belongs to the first component and z_i is updated according to parameter estimations in step M by below equation

$$\hat{z}_i = \frac{\hat{p} f_1(v_i, \hat{\theta}_1)}{\hat{p} f_1(v_i, \hat{\theta}_1) + (1-\hat{p}) f_2(v_i, \hat{\theta}_2)}. \tag{25}$$

The probability belongs to second component of the i-th unit is determined with $1 - z_i$. In this study, parameter estimations are defined respect to first component. The parameter estimations for second component are calculated substituting $1 - z_i$ instead of z_i in equations.

The parameter estimates are updated in step M according to the component-label vector renewed at the end of step E. The mixture weight p is estimated according to obtained component-label vector from step E.

$$\hat{p} = \frac{\sum_{i=1}^{n} \hat{z}_i}{n} \tag{26}$$

The algorithm is repeated until the determined convergence rule has been reached. In this study, the convergence rule is determined as $\left|\log L^{(t+1)} - \log^{(t)}\right| < 10^{-4}$.

In the M-step, an iterative method must be used to estimate the shape parameters of the Wbl2 distribution because there is not analytical solution. In this study, Newton-Raphson method, one of the iterative solution methods, was used for the estimation of the parameters. The maximum likelihood estimation of shape parameter k_1 of the Wbl2 distribution in the $(r+1)$-th iteration of Newton-Raphson method is defined by [20]

$$\hat{k}_{1,r} + \frac{A_{1,r} + (1/\hat{k}_{1,r}) - (C_{1,r}/D_{1,r})}{(1/\hat{k}_{1,r}^2) + (B_{1,r} D_{1,r} - C_{1,r}^2)/(B_{1,r}^2)} \tag{27}$$

where $A_{1,r} = \frac{\sum_{i=1}^{n} \hat{z}_i \log(v_i)}{\sum_{i=1}^{n} \hat{z}_i}$, $B_{1,r} = \sum_{i=1}^{n} \hat{z}_i v_i^{\hat{k}_{1,r}}$, $C_{1,r} = \sum_{i=1}^{n} \hat{z}_i v_i^{\hat{k}_{1,r}} \log(v_i)$ and $D_{1,r} = \sum_{i=1}^{n} \hat{z}_i v_i^{\hat{k}_{1,r}} \log(v_i)^2$. The initial value of k_1 parameter is obtained by equation $\hat{k}_{1,0} = \left(\frac{\frac{6}{\pi^2}\left\{\sum_{i=1}^{n} \hat{z}_i \log(v_i)^2 - \left(\sum_{i=1}^{n} \hat{z}_i \log(v_i)\right)^2 / \sum_{i=1}^{n} \hat{z}_i\right\}}{\sum_{i=1}^{n} \hat{z}_i - 1} \right)^{-1/2}$.

Then the scale parameter c_1 of the Wbl2 distribution is estimated by [20]

$$\hat{c}_1 = \left(\frac{\sum_{i=1}^{n} \hat{z}_i v_i^{\hat{k}_1}}{\sum_{i=1}^{n} \hat{z}_i} \right)^{\frac{1}{\hat{k}_1}}. \tag{28}$$

The Newton Raphson method is used to estimate the shape parameters of the Gam2 distribution. The maximum likelihood estimation of shape parameter b_1 of the Gam2 distribution in the $(r+1)$-th iteration of Newton-Raphson method is defined by [20]

$$\hat{b}_{1,r+1} = \hat{b}_{1,r} - \frac{\log(\hat{b}_{1,r}) - \log(S_1) - \psi(\hat{b}_{1,r}) + S_2}{\frac{1}{\hat{b}_{1,r}} - \psi'(\hat{b}_{1,r})} \tag{29}$$

where $S_1 = \frac{\sum_{i=1}^{n} \hat{z}_i v_i}{\sum_{i=1}^{n} \hat{z}_i}$ and $S_2 = \frac{\sum_{i=1}^{n} \hat{z}_i \log(v_i)}{\sum_{i=1}^{n} \hat{z}_i}$.

In Newton Raphson method, the initial value of b_1 parameter is obtained by equation $\hat{b}_{1,0} = \frac{1}{2\{\log(S_1) - S_2\}}$. Then the scale parameter a_1 of the Gam2 distribution is estimated by [20]

$$\hat{a}_{1,r+1} = \frac{\sum_{i=1}^{n} \hat{z}_i v_i}{\hat{b}_{1,r+1} \sum_{i=1}^{n} \hat{z}_i} \tag{30}$$

The estimation of the λ_1 parameter of the Rayl2 distribution in step M of the EM algorithm is obtained analytically directly.

$$\hat{\lambda}_1 = \frac{\sum_{i=1}^{n} \hat{z}_i v_i^2}{2 \sum_{i=1}^{n} \hat{z}_i} \tag{31}$$

2.3. Performance criterions

In this study, Akaike Information Criterion (AIC), root of mean square error (RMSE), the coefficient of determination (R^2), chi-square (χ^2) and the relative error (%) criterions are used in order to comparison the performances of the wind distributions [23-27].

AIC, defined based on log-likelihood value and penalty term, is used to determine the accuracy of a statistical model.

$$AIC = -2\log L + 2d \tag{32}$$

where 2d is penalty term and d is number of parameters. The good model in terms of AIC value is the one with the minimum AIC value.

RMSE is a measure of difference between empirical distribution function (EDF) probabilities and predicted cumulative distribution ($\hat{F}$) probabilities. A lower value of RMSE indicates a better statistical model.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(EDF(x_{(i)}) - \hat{F}(x_{(i)})\right)^2}{n}} \tag{33}$$

R^2 is represent linear relationship between EDF probabilities and $\hat{F}$ probabilities. A larger value of R^2 indicates a better statistical model. R^2 is calculated by

$$R^2 == \frac{\left\{\sum_{i=1}^{n} \left(EDF(x_{(i)}) - \overline{EDF}\right)\left(\hat{F}(x_{(i)}) - \bar{\hat{F}}\right)\right\}^2}{\sum_{i=1}^{n} \left(EDF(x_{(i)}) - \overline{EDF}\right)^2 \sum_{i=1}^{n} \left(\hat{F}(x_{(i)}) - \bar{\hat{F}}\right)^2} \tag{34}$$

where $\overline{EDF}$ is mean of the EDF probabilities, and $\bar{\hat{F}}$ is mean of the $\hat{F}$ probabilities.

χ^2 is used to assess whether the EDF probabilities differs from the $\hat{F}$ probabilities. A lower value of χ^2 indicates a better statistical model. χ^2 is given by below equation.

$$\chi^2 = \sum_{i=1}^{n} \frac{\left(EDF(x_{(i)}) - \hat{F}(x_{(i)})\right)^2}{\hat{F}(x_{(i)})} \tag{35}$$

The relative error between the wind power density calculated from actual wind data and that from probability distribution model is defined by

$$Relative\ Error(\%) = 100 \times \frac{|P_{actual} - P_{model}|}{P_{actual}}. \quad (36)$$

3. Results and Discussion

In this study, the hourly average wind speed measured at Keçiören/Bağlum and Çankaya/Çaldağ meteorological stations of the Turkish State Meteorological Service from 01 January 2019 to 31 December 2019 were used.

The box plots of the hourly average wind speeds obtained from Keçiören/Bağlum and Çankaya/Çaldağ meteorological stations are presented in Fig. 1.

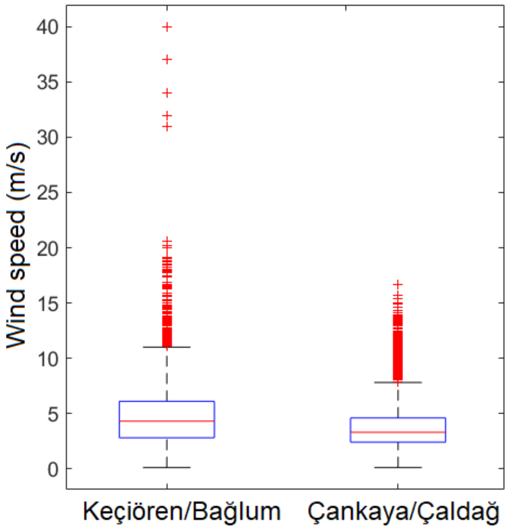


Fig. 1. The box plots of the wind speeds according to meteorological stations.

From Fig. 1 is seen that variability of the wind speeds obtained from Keçiören/Bağlum meteorological station more than variability of the wind speeds obtained from Çankaya/Çaldağ meteorological station. Furthermore, the observed wind speeds in the Keçiören/Bağlum meteorological station are higher according to the observed wind speeds in the Çankaya/Çaldağ meteorological station.

A wind rose diagram is an important graphical tool for evaluating wind speed measurements at a particular location. Wind rose diagram gives simultaneous information about wind speed and wind direction frequencies [28, 29]. The wind rose diagrams for the wind data measured at Keçiören/Bağlum and Çankaya/Çaldağ meteorological stations are given Fig. 2.

In this section of the study, hourly average wind speed data is modeled with Wbl, Gam, Rayl, Wbl2, Gam2, Rayl2, WblRay and GamRayl distributions. The parameter estimations and log-likelihood values for probability distribution used in wind speeds modeling are given in Table 1.

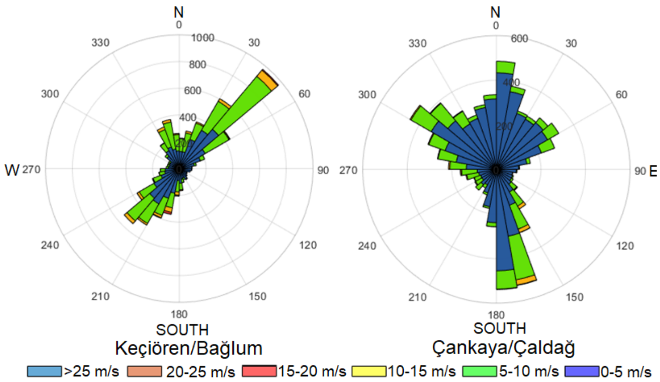


Fig. 2. Wind rose diagrams according to meteorological stations.

Table 1. The parameter estimates and log-likelihood values.

	Models	Parameter Estimates					***logL***
Keçiören/Bağlum	Wbl	$\hat{c}$ 5.35	$\hat{k}$ 1.83				-20214.4
	Gam	$\hat{a}$ 3.08	$\hat{b}$ 1.54				-20048.4
	Rayl	$\hat{\lambda}$ 3.87					-20286.3
	Wbl2	$\hat{p}$ 0.16	$\hat{c}_1$ 7.31	$\hat{k}_1$ 1.60	$\hat{c}_2$ 4.96	$\hat{k}_2$ 2.14	-20023.6
	Gam2	$\hat{p}$ 0.013	$\hat{a}_1$ 0.65	$\hat{b}_1$ 9.63	$\hat{a}_2$ 3.11	$\hat{b}_2$ 1.52	-20006.2
	Rayl2	$\hat{p}$ 0.96	$\hat{\lambda}_1$ 3.60	$\hat{\lambda}_2$ 7.81			-20033.0
	WblGam	$\hat{p}$ 0.06	$\hat{c}$ 5.33	$\hat{k}$ 1.03	$\hat{a}$ 3.51	$\hat{b}$ 1.34	-19992.8
	WblRayl	$\hat{p}$ 0.06	$\hat{c}$ 9.07	$\hat{k}$ 1.64	$\hat{\lambda}$ 3.59		-20029.4
	GamRayl	$\hat{p}$ 0.48	$\hat{a}$ 3.34	$\hat{b}$ 1.65	$\hat{\lambda}$ 3.21		-20014.8
Çankaya/Çaldağ	Wbl	$\hat{c}$ 4.20	$\hat{k}$ 2.06				-17181.7
	Gam	$\hat{a}$ 4.28	$\hat{b}$ 0.87				-16729.8
	Rayl	$\hat{\lambda}$ 2.95					-17189.5
	Wbl2	$\hat{p}$ 0.29	$\hat{c}_1$ 5.86	$\hat{k}_1$ 2.19	$\hat{c}_2$ 3.49	$\hat{k}_2$ 2.93	-16686.0
	Gam2	$\hat{p}$ 0.12	$\hat{a}_1$ 2.50	$\hat{b}_1$ 2.10	$\hat{a}_2$ 5.03	$\hat{b}_2$ 0.70	-16636.1
	Rayl2	$\hat{p}$ 0.94	$\hat{\lambda}_1$ 2.74	$\hat{\lambda}_2$ 5.23			-17041.6
	WblGam	$\hat{p}$ 0.15	$\hat{c}$ 6.37	$\hat{k}$ 2.02	$\hat{a}$ 5.55	$\hat{b}$ 0.61	-16616.0
	WblRayl	$\hat{p}$ 0.69	$\hat{c}$ 3.52	$\hat{k}$ 2.95	$\hat{\lambda}$ 3.96		-16693.0
	GamRayl	$\hat{p}$ 0.85	$\hat{a}$ 5.55	$\hat{b}$ 0.61	$\hat{\lambda}$ 4.48		-16616.0

The calculated performance criterions for probability distribution models used in wind speeds modeling are given in Table 2.

Table 2. The performance criterions for probability distribution models.

	Models	*AIC*	R^2	*RMSE*	χ^2
Keçiören / Bağlum	Wbl	40432.85	0.99794	0.14705	0.01602
	Gam	40100.85	0.99977	0.10429	0.00614
	Rayl	40574.67	0.99674	0.19315	0.02291
	Wbl2	40057.16	0.99965	0.02640	0.00721
	Gam2	40022.34	0.99980	0.01856	0.00580
	Rayl2	40071.99	0.99954	0.03643	0.00766
	WblGam	**39995.62**	**0.99996**	0.01799	**0.00393**
	WblRayl	40066.76	0.99955	0.03817	0.00744
	GamRayl	40037.59	0.99984	**0.01586**	0.00519
Çankaya / Çaldağ	Wbl	34367.40	0.99308	0.43086	0.02863
	Gam	33463.60	0.99829	0.14685	0.01500
	Rayl	34381.03	0.99341	0.48616	0.02826
	Wbl2	33382.00	0.99941	0.05040	0.00918
	Gam2	33282.17	0.99965	**0.02736**	0.00729
	Rayl2	34089.09	0.99675	0.53385	0.02405
	WblGam	33241.94	**0.99983**	0.02838	0.00612
	WblRayl	33394.06	0.99940	0.05680	0.00885
	GamRayl	**33239.99**	**0.99983**	0.02751	**0.00610**

The results in Table 2 show clearly that two component mixture probability distribution models are more successful than classical probability distributions in modelling of the hourly average wind speed data. WblGam distribution is most successful model according to performance criterions AIC, R2, and χ^2 in modelling of the obtained wind speed data from Keçiören/Bağlum station. However, in terms of RMSE criteria for Keçiören station, the most successful model is GamRayl distribution. For Çankaya/Çaldağ station, GamRayl distribution is most successful model according to performance criterions AIC, R2, and χ^2 in modelling of the wind speed data. Wherein further, the WblGam distribution showed the same performance to the GamRayl distribution in term of R^2 criteria. In terms of RMSE, Gam2 distribution is the most successful model in modelling of the obtained wind speed data from Çankaya/Çaldağ station. The pdf curves of the best fit model and histograms of wind speed data for Keçiören/Bağlum and Çankaya/Çaldağ stations are illustrated in Fig. 3.

According to Fig. 3, it is clearly seen that goodness-of-fit of the WblGam and GamRayl distributions, which are selected as the most successful model in modeling wind speeds, are very good for wind speed data.

The estimations of the mean wind speed, variance and wind power density according to the probability distributions used in the modeling of wind speeds, and the obtained actual values from the observed data are given in Table 3. In addition, relative error values (%) for wind power density estimations are given in the last column of Table 3.

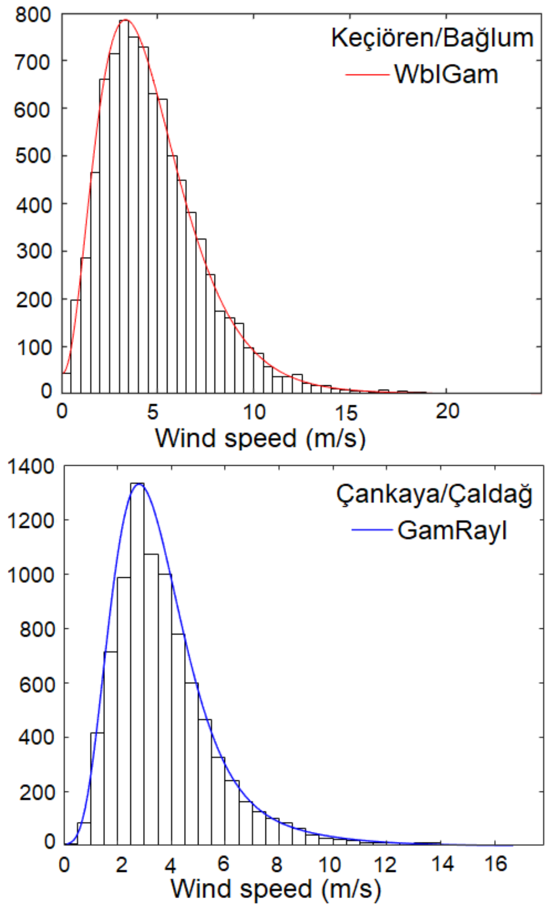


Fig. 3. The pdf curves of the best fit model and histograms of wind speed data

Table 3. Wind characteristics and relative error values (%) for all stations

	Models	v_{mean}	P_{Model}	Relative Error (%)
Keçiören / Bağlum	Data(actual)	4.7397	153.926	
	Wbl	4.7495	137.974	10.363
	Gam	4.7397	142.432	7.467
	Ray	4.8521	133.624	13.189
	Wbl2	4.7437	147.491	4.180
	Gam2	4.7397	158.284	2.831
	Ray2	4.7347	149.211	3.063
	WblGam	4.7399	149.808	**2.675**
	WblRayl	4.7282	149.306	3.001
	GamRayl	4.7312	141.199	8.268
Çankaya / Çaldağ	Data(actual)	3.7108	62.785	
	Wbl	3.7225	58.567	6.718
	Gam	3.7108	56.690	9.708
	Ray	3.6975	59.134	5.815
	Wbl2	3.7135	62.263	0.831
	Gam2	3.7108	65.822	4.837
	Ray2	3.6212	64.437	2.632
	WblGam	3.7105	62.723	0.098
	WblRayl	3.7041	62.711	0.118
	GamRayl	3.7102	62.747	**0.061**

According to the relative error values (%) in Table 3, WblGam and GamRayl distribution respectively are the most successful models for Keçiören/Bağlum and Çankaya/Çaldağ stations in modeling of the hourly average wind speeds. It is observed that the wind potential of Keçiören/Bağlum region more than twice the wind potential of Çankaya/Çaldağ region in terms of wind power density.

The bar plots of wind power density and relative error values (%) are given in Fig. 4-5.

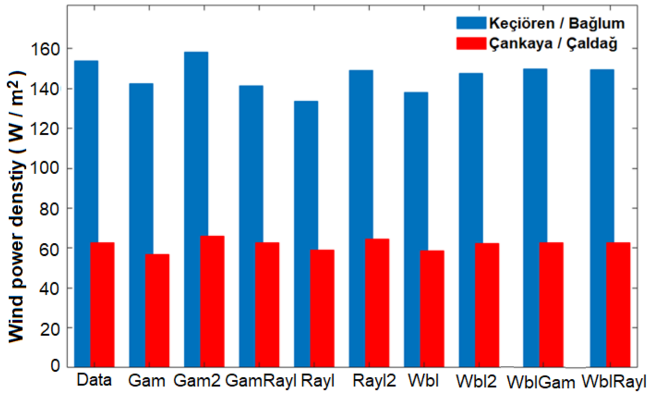
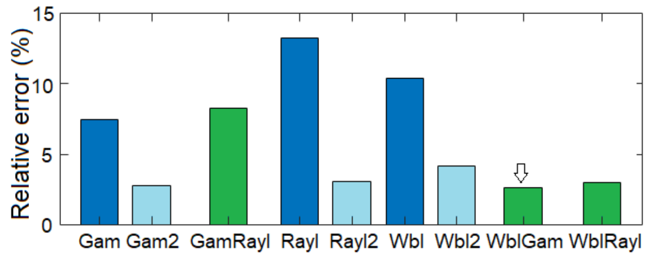


Fig. 4. The bar plot of the wind power densities for all the stations.



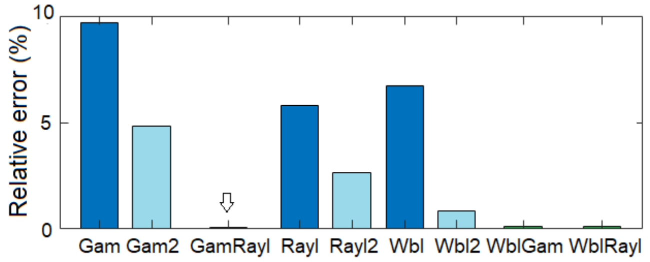


Fig. 5. The bar plot of the relative error values (%) for all the stations.

The wind power density per unit area was calculated as 153.926 W/m^2 based on the observed wind speed data in Keçiören/Bağlum region. According to the WblGam mixture distribution model, the most successful model in modeling the wind speeds in Keçiören/Bağlum region, wind power density per unit area was calculated as 149.211 W/m^2. The relative error of the WblGam model in estimating the wind power density was 2.67%. The wind power density per unit area was calculated as 62.785 W/m^2 based on the observed wind speed data in Çankaya/Çaldağ region. According to the GamRayl mixture distribution model, the most successful model in modeling the wind speeds in Çankaya/Çaldağ region, wind power density per unit area was calculated as 62.747 W/m^2. The relative error of the GamRayl model in estimating the wind power density was 0.06%.

4. Conclusion

In this study, hourly average wind speed data measured at Keçiören/Bağlum and Çankaya/Çaldağ meteorological stations are modeled with Weibull, Gamma and Rayleigh probability distribution and theirs two component mixture probability distributions. The performances of investigated wind distributions were measured with AIC, RMSE, R^2, χ^2 and the relative error (%) criterions. In this study, it has been shown that the models created by the mixture of two different distributions are more successful in modeling the wind speeds. The conclusions of the study are as follows:

- The mixture distributions have lead to better fit than classical distributions in modeling wind speed data at Keçiören/Bağlum and Çankaya/Çaldağ meteorological stations in the Ankara.
- Weibull-Gamma mixture distribution was best fit model for modeling wind speed data at Keçiören/Bağlum meteorological station.
- Gamma-Rayleigh mixture distribution was best fit model for modeling wind speed data at Çankaya/Çaldağ meteorological station.
- The wind potential of Keçiören/Bağlum region was found more than twice the wind potential of Çankaya/Çaldağ region in terms of wind power density.

Future studies in modeling wind speeds, may employ the effect of different parameter estimation methods in the estimation of model parameters.

Acknowledgements

We would like to thank the Editor and anonymous referees for their comments which significantly enhanced the quality of this manuscript.

References

[1] P. A. Owusu and S. Asumadu-Sarkodie, "A review of renewable energy sources, sustainability issues and climate change mitigation", *Cogent Engineering,* vol. 3, no. 1, pp. 1-14, 2016.

[2] S. M. O. Thierry, D. A. Anyouzo, and A. N. Ndachigam, "Numerical Study of the Power Range of Wind Devices Adapted to the Wind Potential of the Coastal Region of Cameroon", *International Journal of Renewable Energy Research (IJRER)*, vol. 10(3), pp. 1439-1450, 2020.

[3] S. E. Hosseini and M. A. Wahid, "Hydrogen production from renewable and sustainable energy resources: promising green energy carrier for clean development", *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 850-866, 2016.

[4] J. Huenteler, C. Niebuhr, and T. S. Schmidt, "The effect of local and global learning on the cost of renewable energy in developing countries", *Journal of Cleaner Production*, vol. 128, pp. 6-21, 2016.

[5] S. Ozdemir, U. S. Selamogullari and O. Elma, “Analyzing the effect of inverter efficiency improvement in wind turbine systems”, International Conference on Renewable Energy Research and Application (ICRERA), pp. 572-575, IEEE, 19-22 October 2014.

[6] A. Pigazo, Z. Qin, M. Liserre, and F. Blaabjerg, “Generation of random wind speed profiles for evaluation of stress in WT power converters”, International Conference on Renewable Energy Research and Applications (ICRERA), IEEE, pp. 436-441, 20-23 October 2013.

[7] T. P. Chang, “Estimation of wind energy potential using different probability density functions”, *Applied Energy*, vol. 88, no. 5, pp. 1848-1856, 2011.

[8] S. Basile, R. Burlon, and F. Morales, F, “Joint probability distributions for wind speed and direction: A case study in Sicily”, International Conference on Renewable Energy Research and Applications (ICRERA), IEEE, pp. 1591-1596, 22-25 November 2015.

[9] 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”, *Renewable and Sustainable Energy Reviews*, vol. 13, no. 5, pp. 933-955, 2009.

[10] S. H. Pishgar-Komleh, A. Keyhani, and P. Sefeedpari, “Wind speed and power density analysis based on Weibull and Rayleigh distributions (a case study: Firouzkooh county of Iran)”, *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 313-322, 2015.

[11] C. Jung and D. Schindler, “Global comparison of the goodness-of-fit of wind speed distributions”, *Energy* Conversion *and* Management, vol. 133, pp. 216-234, 2017.

[12] K. Mohammadi, O. Alavi, and J. G. McGowan, “Use of Birnbaum-Saunders distribution for estimating wind speed and wind power probability distributions: A review”, *Energy Conversion and Management,* vol. 143, pp. 109-122, 2017.

[13] N. Aries, S. M. Boudia, and H. Ounis, “Deep assessment of wind speed distribution models: A case study of four sites in Algeria”, *Energy Conversion and Management,* vol. 155, pp. 78-90, 2018.

[14] H. Bidaoui, I. El Abbassi, A. El Bouardi, and A. Darcherif, “Wind speed data analysis using Weibull and Rayleigh distribution functions, case study: five cities northern Morocco”, *Procedia Manufacturing*, vol. 32, pp. 786-793, 2019.

[15] S. Akpinar and E. K. Akpinar, “Estimation of wind energy potential using finite mixture distribution models”, *Energy Conversion and Management*, vol. 50, no. 4, pp. 877-884, 2009

[16] J. Y. Shin, T. B. M. J. Ouarda, and T. Lee, “Heterogeneous mixture distributions for modeling wind speed, application to the UAE”, *Renewable Energy*, vol. 91, pp. 40-52, 2016.

[17] D. Mazzeo, G. Oliveti, and E. Labonia, “Estimation of wind speed probability density function using a mixture of two truncated normal distributions”, *Renewable Energy*, vol. 115, pp. 1260-1280, 2018.

[18] T. B. M. J. Ouarda, and C. Charron, “On the mixture of wind speed distribution in a Nordic region”, *Energy Conversion and Management*, vol. 174, pp. 33-44, 2018.

[19] N. J. Cook, “The OEN mixture model for the joint distribution of wind speed and direction: A globally applicable model with physical justification”, *Energy Conversion and Management,* vol. 191, pp. 141-158, 2019.

[20] M. Erisoglu, T. Servi, U. Erisoglu, and N. Calis, “Mixture Gamma Distribution for Estimation of Wind Power Potential”, *International Journal of Applied Mathematics and Statistics*, vol. 40, pp. 232-241, 2013.

[21] U. Erisoglu and M. Erisoglu, “Percentile Estimators for two-component mixture distribution models”, *Iranian Journal of Science and Technology, Transactions A: Science,* vol. 43, no. 2, pp. 601-619, 2019.

[22] S. Miao, Y. Gu, D. Li, and H. Li, “Determining suitable region wind speed probability distribution using optimal score-radar map”, *Energy conversion and management*, vol. 183, pp. 590-603, 2019.

[23] J. Zhou, E. Erdem, G. Li, and J. Shi, “Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites”, *Energy Conversion and Management*, vol. 51, no. 7, pp. 1449-1458, 2010.

[24] I. Usta and Y. M. Kantar, “Analysis of some flexible families of distributions for estimation of wind speed distributions”, *Applied Energy*, vol. 89, no. 1, pp. 355-367, 2012.

[25] V. Katinas, G. Gecevicius, and M. Marciukaitis, “An investigation of wind power density distribution at location with low and high wind speeds using statistical model”, *Applied Energy,* vol. 218, pp. 442-451, 2018.

[26] M. Sumair, T. Aized, S. A. R. Gardezi, S. U. ur Rehman, and S. M. S. Rehman, “A newly proposed method for Weibull parameters estimation and assessment of wind potential in Southern Punjab”, *Energy Reports*, vol. 6, pp. 1250-1261, 2020.

[27] O. Kaplan, and M. Temiz, “The analysis of wind speed potential and energy density in Ankara”, IEEE International Conference on Renewable Energy Research and Applications (ICRERA), IEEE, pp. 919-923, 20-23 November 2016.

[28] A. Allouhi, O. Zamzoum, M. R. Islam, R. Saidur, T. Kousksou, A. Jamil, and A. Derouich, “Evaluation of wind energy potential in Morocco’s coastal regions”, *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 311-324, 2017.

[29] I. Colak, M. S. Ayaz, and K. Boran, “CFD based wind assesment in west of Turkey”, International Conference on Renewable Energy Research and Applications (ICRERA), IEEE, pp. 727-731, 22- 25 November 2015.