IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 27, 2021, accepted April 23, 2021, date of publication May 10, 2021, date of current version May 18, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3078511*

# Determination of Optimal Parametric Distribution and Technical Evaluation of Wind Resource Characteristics for Wind Power Potential at Jhimpir, Pakistan

**MUHAMMAD ARMOGHAN KHAN[1,2], YAO ZHANG[1], (Member, IEEE), JIANXUE WANG[1], (Senior Member, IEEE), JINGDONG WEI[1], MUHAMMAD ALI RAZA[3], AITIZAZ AHMAD[4], AND YIPING YUAN[1]**

[1]Shaanxi Key Laboratory of Smart Grid, School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[2]Department of Electronics and Power Engineering, National University of Sciences and Technology, Islamabad 24090, Pakistan
[3]Department Electrical and Computer Engineering, COMSATS University Islamabad at Lahore, Lahore 54000, Pakistan
[4]School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China

Corresponding author: Yao Zhang (yaozhang_ee@xjtu.edu.cn)

**ABSTRACT** The objective of this research work is to analyze wind characteristics and to assess wind power potential by selecting the best fit probability distribution function of Jhimpir Sindh Pakistan. This type of detailed investigation helps wind power generation companies in selecting suitable wind turbine and provides information of wind characteristics of potential site. Eight probability distribution functions are tested on the wind speed data from January 2015 to July 2018. Frequency bins of Weibull and Rayleigh distribution with maximum probabilities of 0.1210 and 0.1143 are most closest representation of our data. In order to, quantitatively analysis which distribution function is best fitting the local wind regime, we have applied the coefficient-of-determination, Kolmogorov-Smirnov, Chi square, Cramer-von Mises, Anderson-Darling tests along with Akaike information and Bayesian information criterion. These statistical test are used to rank the empirical distribution functions in order to identify two distribution function better fitting the actual wind speed data. After selecting two best fitted distribution functions, we analyze wind power potential and compare the error of wind power density based on these distribution functions (Weibull and Rayleigh). The power densities reported varied from 73.67 to 648.73W/m$^2$. Results indicate that power densities of Weibull and Rayleigh for the candidate site are 84.67–698.65W/m$^2$ and 83.67–1021.4W/m$^2$, respectively. The highest error for Weibull and Rayleigh are 0.1850 and 0.5745, respectively. Whereas lowest error are 0.0178 and 0.0180, respectively. Complete analysis suggested that Weibull distribution function is the most suitable for Jhimpir Sindh Pakistan and the studied site is suitable for wind power production. In addition, comprehensive analysis of wind direction at the candidate site suggested that Eastern and Southeastern wind directions are predominant with 38.52% and 33.24% of the total time.

**INDEX TERMS** Jhimpir Pakistan, parametric distribution, statistical analysis, Weibull distribution, wind power potential, wind speed and direction.

## I. INTRODUCTION

Due to an increase in the pollution, rising environmental issues and continuous decrease in the reservoirs of conventional energy sources worldwide, power generation companies are moving towards renewable energy resources. A lot of work is already being done for the reduction of environmental effects and produce low priced electricity. Among renewable energy resources, wind and solar power generation are considered as the most valuable and reliable sources. During the last decade, the total installed capacity of wind energy in countries like China, United States, Germany and India have shown a radical improvement and the share of wind energy has increased a lot. Beside countries leading the world in

The associate editor coordinating the review of this manuscript and approving it for publication was R. K. Saket.

term of installed capacity, a lot of other countries, like South Africa, Pakistan, Uruguay and Iran, have also shown a huge improvement in installed capacity of wind energy. Because of the environment-friendly, available abundantly in nature and most importantly affordability for weak economies of developing world, the total installed capacity of wind power all over the world has shown an increasing trend during the last decade [1]. In 2017, it has increased with an annual growth rate of 10.1% and reached 514798MW, as shown in Fig. 1 [1].

Pakistan is facing different problems in increasing the share of renewable energy resources from beginning. In the year 2003, energy sector has observed a major change when Federal Government decided to have a special division as Alternative Energy Development Board (AEDB). Since then, a lot of renewable energy projects have started working under this umbrella and the total installed capacity of wind power generation in Pakistan was 789MW by the end of 2017. In 2019, AEDB has published Alternative and Renewable Energy (ARE) Policy 2019 and decided to increase the share of renewable energy to at least 30% by the year 2030. In this regard, twelve new wind power projects with installed capacity of 610MW are in pipeline and are envisaged to come online by year 2021. Thus, wind power potential is very huge in Pakistan and its development would meet more and more electricity demand in Pakistan.

For any selected site, wind power potential can be calculated by analyzing wind speed characteristics (speed, direction, availability, and continuity) of that area. Researchers have analyzed these characteristics by using parametric and non-parametric methods. Parametric method requires estimation of distribution parameters and these parameters are then analyzed on different bases for the assessment of wind energy potential. Up to now, more than 100 distributions have been proposed to fit wind speed data. Among them, the most recommended distributions are Weibull [2], Rayleigh [3], [8], [9], Gamma [8]–[10] and Log-normal [10]–[13], depending on geographical locations. Ayodele *et al.* [2] have studied Weibull, Rayleigh and lognormal distributions function at the local site to estimate capacity factor of 20 commercially available wind turbines.

Keyhani *et al.* [14] have statistically analyzed eleven years of wind speed data with a resolution of three hours and wind power potential of Tehran, Iran, has been studied based on Weibull model. Dabbaghiyan *et al.* have employed Weibull probability distribution function to calculate the wind power density and energy for four locations in Bushehr province of Iran [15]. Multi-criteria decision making, using GIS software, is applied to study wind energy potential for Northwestern Iran by Bina *et al.* [16]. In their study, technical and environmental criteria have been applied discretely and results showed that sixty-four percent of the area is appropriate for wind farm installation. Saeed *et al.* [17] have assessed the wind power potential for Sanghar, Pakistan. Wind speed data of two years is analyzed by using three variations of Weibull parameters.

The proposed framework for estimating power output generation of wind turbines by Wacker *et al.* [18] depend upon wind speed distributions selection, wind speed data set and general power curve modeling. Which shows the importance of selection of wind speed distribution. Baser *et al.* [19] have assessed the wind energy resources and analyzed annual wind energy production from a wind turbine of 3 MW and estimated it to be 6285 MWh with plant capacity factor of 25%, for Jubail Saudi Arabia. Ramli *et al.* [20] have investigated the hybrid system configuration and concluded that both wind and solar have sufficient potential in the western part of Saudi Arabia. From the references, mention above it can be concluded that analysis of wind speed characteristics, selection of distribution function as assessment of wind power potential are the three most important attributes in analysis process at any particular site. These attributes are helpful for selection of suitable wind turbines, economical evaluation, auditing cost effectiveness and estimating future income of wind energy projects.

In Pakistan, although researchers have estimated the wind energy potential before the inauguration of AEDB, researchers have shown more interest in this area under the aegis of AEDB and have published some papers about the wind energy potential and wind characteristics for different locations of Pakistan. Ashfaq *et al.* have done wind power atlas modeling in their study and used the quasi exact method to calculate the power generation potential at a spatial resolution of $14 \times 14$ km$^2$ [21]. Khahro *et al.* have used the Weibull probability density function (PDF) and Rayleigh distribution function in their work [22]. They have also compared five different parameter selection methods to determine the shape and scale parameters of Weibull distribution for Babaurband, Sindh Pakistan. To meet the power shortage of Karachi city, the wind energy potential of Karachi city has been studied by Aman *et al.* [23]. For Jhimpir and Gharo, wind speed benchmark was developed by Mirza *et al.* [24]. Hulio *et al.* have assessed the wind characteristic and wind power potential for Hawksbay, Karachi and different parameter estimation methods are analyzed in their study at the candidate site [25]. An investigation of wind power potential from September 2003 to August 2005 for Jamshoro, Pakistan is performed in [26].

If we have a look at the data of installed capacity of wind energy published by [27]. Pakistan falls at 33th position out of 113 countries with a total increase of 783 MW in installed capacity from 2010 to 2017. However, the rate of increase in Pakistan was considerably high during this period. Pakistan falls at 2nd position in term of the rate of increase, among all those countries which were having installed capacity greater than or equal to that of Pakistan in 2010. Detail comparison of the rate of increase among different countries is discussed in Section 2.2 of this paper. Due to the increase in installed capacity and high rate of increase during the last decade, there is a need to study wind energy potential of different locations of the country. So far, no one has analyzed the wind power potential of Jhimpir, Pakistan that

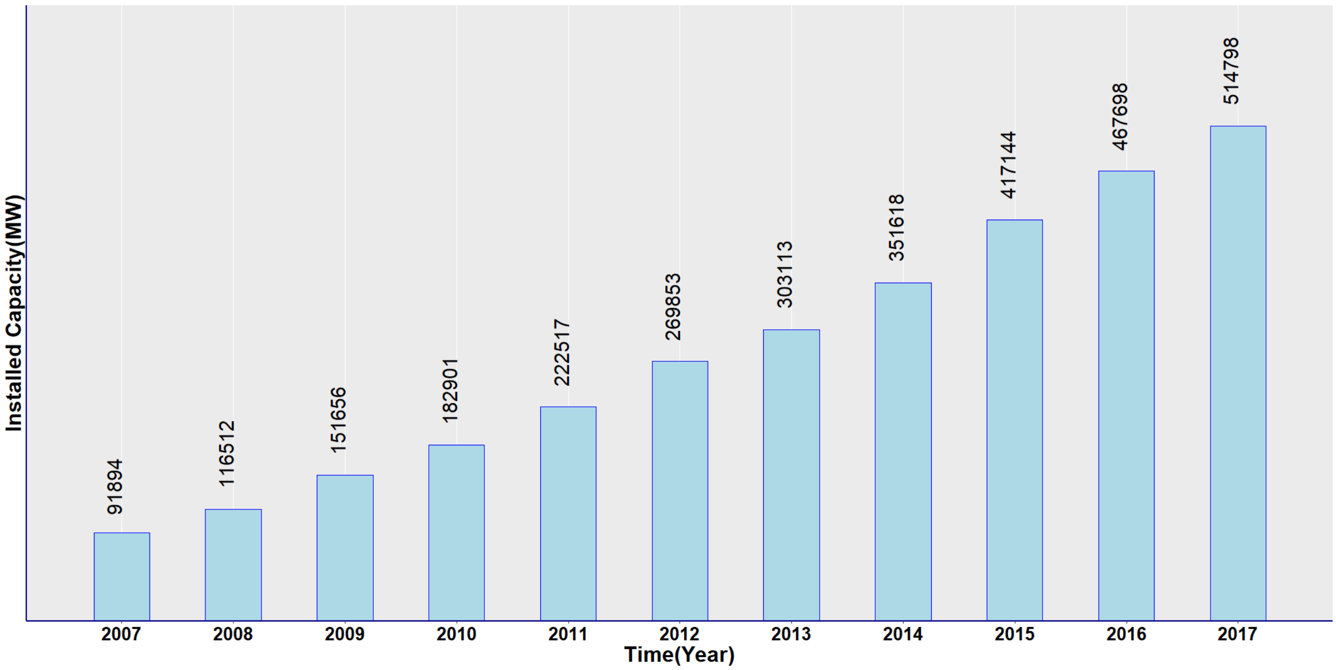



**FIGURE 1.** The total installed capacity of wind power generation all over the world.

lies in the wind energy corridor in the southern part of the country.

Unlike the conventional sources of electrical energy, power produced by renewable energy resources has a strong correlation with geographical location and regional climatic condition of the particular site. Understanding the impact of climatic parameters at particular site provides valuable input for wind energy projects at planning and financing stages. Better understanding of wind characteristics from historical data of wind makes business models more robust and budget tables more accurate. All-around-the-globe, researchers have been running similar studies in different countries and climates. The presence of such studies not only allows understanding wind characteristic in certain locations and climate conditions, but also helps compare wind power potential of studied site with existing studies at some other location having similar wind speed characteristics and distributional properties.

Up to now, there is lack of detailed study for the assessment of wind characteristics and wind power potential of Jhimpir, Pakistan. Consequently, in this study, we have analyzed the wind characteristics in Jhimpir, Pakistan and the competence of seven probability distribution functions (i.e., Gamma, Weibull, Rayleigh, Cauchy, Log-normal, Gumbel and Logistic) is assessed to find the best fit distribution function on the bases of seven tests applied (i.e., Coefficient of Determination, Kolmogorov-Smirnov test, Chi square test, Cramer-von Mises test, Anderson-Darling test, Akaike information criterion and Bayesian information criterion). After shortlisting the two best fitted distribution functions, we analyze the wind power potential and the comparison of wind power density error is performed for selecting the final best fit probability distribution function. In addition, comprehensive analysis of wind direction is also part of this research. The contribution of this paper is summarized as follows.

- To the best of our knowledge, this is the first effort to analyze wind data at Jhimpir, Pakistan. Seven probability density functions are employed to assess which one is the most suitable for the estimation of wind power density at Jhimpir, Pakistan.
- To test how good a probability function fits the observed wind data, seven statistical tests are applied to select the best fit distribution function. Wind power density error based on Weibull and Rayleigh distributions are compared. Final results show that Weibull is the most effective distribution for Jhimpir, Pakistan.
- Wind power potential at Jhimpir, Pakistan is also assessed, showing that wind power potential is very huge and Jhimpir is very suitable for the development of large-scale wind power projects.

Research findings of this study will help project developers understand the wind characteristics, distributional properties of wind, wind power potential and what to expect from the developed sites in the future. In addition to this, it will also help policy makers and financial supporter, in making decision about installation of wind farm at jhimpir, Sindh, Pakistan.

The remaining paper is organized as follows. Wind energy development around the globe and Pakistan is summarized in Section 2. Parameter estimation of wind data is explained in Section 3. Parametric distributions for distributional

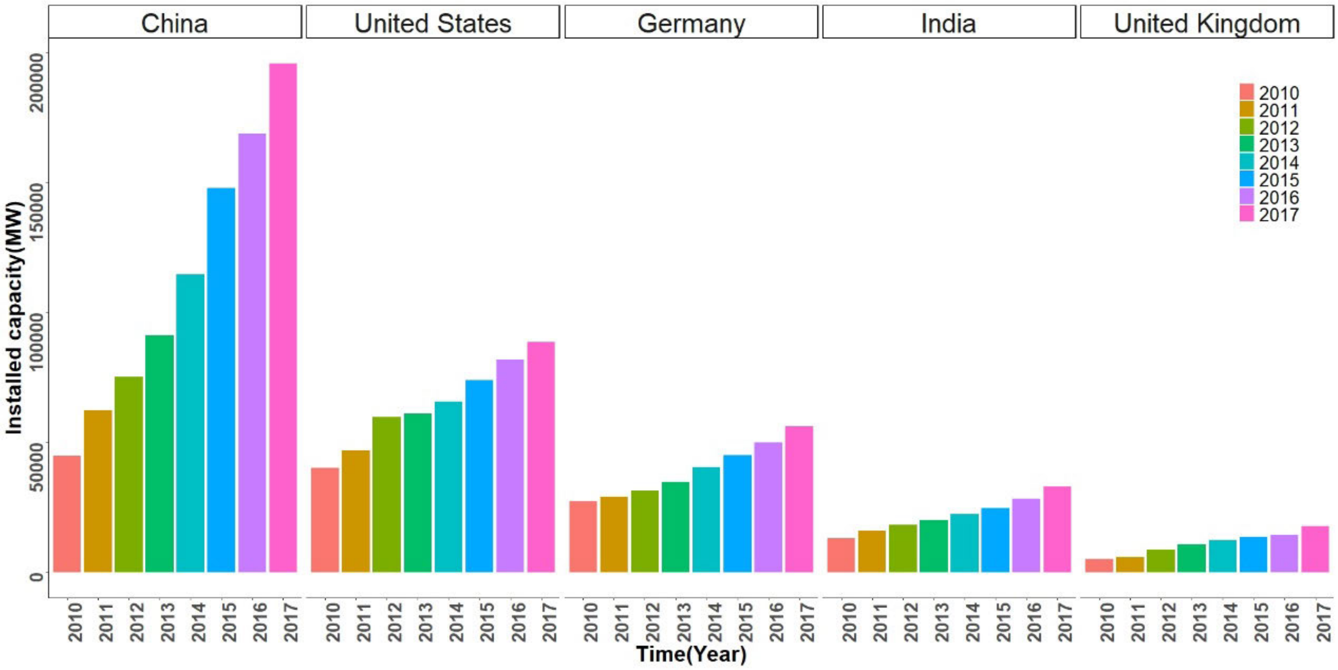



**FIGURE 2.** Top five countries with the largest increase in the installed capacity of wind power generation from 2010 to 2017.

assessment of wind speed along with the estimation method used and goodness to fit test applied for assessment of these distributions are presented in Section 4. Wind power density calculations are performed in Section 5. In Section 6, results of all the above mentioned sections are analyzed and discussed. Finally, conclusions are presented in Section 7.

## II. WIND ENERGY DEVELOPMENT

### A. GLOBAL WIND ENERGY DEVELOPMENT

Due to global warming and rapid weather changes across the globe, all international organizations have emphasized on the importance of renewable energy resources. During the last decade, many countries have shown great improvements in the usage of renewable energy resources like the wind on different scale to help the world be a healthier and safer place to live. In 2010, the installed capacity of China and United States were 44,733 MW and 40,180 MW, respectively. But over the years, China has shown a huge increase and in 2017, the installed capacity of wind energy in China was more than double of the total installed capacity of United States. If we have a look at the total increase in the installed capacity of wind energy from 2010 to 2017, with the significant addition of 150,977 MW, China is now leading the world in wind power development, followed by United States, Germany, India and United Kingdom. United States has shown the total increase of 48,595 MW across the years and its installed capacity was 88,775 MW in 2017. Likewise, Germany, India and United Kingdom have increased their installed capacity by 28,975 MW, 19,813 MW and 12,648 MW, respectively and the total installed capacity in these countries were 27,215 MW, 13,066 MW and 5,204 MW in 2017, respectively [27]. Bar graph of the installed capacity over the years for these five countries is shown in Fig. 2.

### B. WIND ENERGY DEVELOPMENT IN PAKISTAN

The installed capacity of Pakistan in 2010 was 6 MW and in 2017, it reached a value of 786 MW [27], showing a very rapid increase in recent years. Here, we use the rate of increase (ROI) to describe the growing rate of the installed capacity in different countries in recent years. ROI over a period of time ($n$ years) can be calculated by (1).

$$\mathrm{ROI} = \left(\frac{IC_t}{IC_{t-n}}\right)^{\frac{1}{n}} - 1 \tag{1}$$

where $IC_t$ is the installed capacity at year $t$, $IC_{t-n}$ is the installed capacity at year $t-n$ and $n$ is the total number of years across the period.

Top five countries with the biggest ROI of the installed wind capacity are shown in Table 1. South Africa is having the highest ROI of 114.42% with the total installed capacity of 2,085MW in 2017. Pakistan stands at the second place in Table 1 after South Africa, having the second highest ROI of 100.77%. Uruguay is at the third position with ROI of 74.13%. Brazil and Philippines have the ROI less than 50% and are at the fourth and fifth positions with the total installed capacity of 12,763MW and 427MW, respectively.

From Table 1, we can see that although the total installed capacity of Pakistan in 2017 is not very large, Pakistan now has the second highest ROI (100.77%) all over the world, showing that wind power development in Pakistan is growing very rapidly in recent years. Here, it is important to note that a zero value in the starting year will result in an infinite ROI.

**TABLE 1. Top five countries with the largest ROI of the installed capacity of wind energy generation from 2010 to 2017.**

| Rank | Country Name | Installed Capacity in 2017 (MW) | Installed Capacity in 2010 (MW) | ROI (%) |
|---|---|---|---|---|
| 1 | South Africa | 2085 | 10 | 114.43 |
| 2 | Pakistan | 789 | 6 | 100.77 |
| 3 | Uruguay | 1505 | 31 | 74.13 |
| 4 | Brazil | 12763 | 930 | 45.37 |
| 5 | Philippines | 427 | 33 | 44.16 |

Thus, to carry out a fair comparison with Pakistan, we have omitted the countries with the installed capacity less than the installed capacity of Pakistan in 2010 (i.e., 6MW).

## III. WIND DATA PARAMETER ESTIMATION

Wind energy potential of a particular site can be evaluated if we know wind features of that location. Analysis of different wind attributes like wind speed, the availability of wind at a particular speed, wind direction and the continuity of wind blowing is important for making initial estimates of wind energy potential. Distributional analysis of wind speed is considered to be the most important for wind energy potential of a site. By doing the distributional analysis, we can have information about the minimum and maximum speed of wind as well as wind speed frequency distribution that is the number of times a particular wind speed was available over the period of time. Beside wind features, other factors that can affect the wind production at a particular site are tower height and design of the wind turbine.

Average wind speed and standard deviation are considered to be two most important parameters for analysis of wind characteristics. The average wind speed and standard deviation can be calculated by using (2) and (3):

$$V_{avg} = \frac{1}{N}\left(\sum_{i=1}^{N} V_i\right) \tag{2}$$

$$\sigma = \left[\frac{1}{N-1}\sum_{i=1}^{N}\left(V_i - V_{avg}\right)^2\right]^{1/2} \tag{3}$$

where $V_i$ represents the wind speed (m/s) at time $i$ and $N$ is the total number of entries available.

Besides these two parameters, two more parameters are used to approximate the wind power potential. They are the most probable wind speed $V_{mp}$ and the wind speed having maximum energy $V_{maxE}$. Wind speed carrying maximum energy $V_{maxE}$ helps in estimating the design of wind turbine whereas the most probable wind speed $V_{mp}$ represents the peak of the PDF. They can be calculated by using (4) and (5).

$$V_{mp} = C\left(\frac{k-1}{k}\right)^{1/k} \tag{4}$$

$$V_{maxE} = C\left(\frac{k+2}{k}\right)^{1/k} \tag{5}$$

where $C$ and $k$ are the Weibull shape and scale parameters.

## IV. WIND SPEED DISTRIBUTION ASSESSMENT

It is important to note that wind power production has a strong correlation with the wind speed besides geographical statistics and meteorological data. So, some statistical analysis should be carried out to determine wind speed distribution. Literature review illustrates that several parametric distribution models have been proposed to guesstimate wind speed probability distributions and they have been used for wind farm planning, reliability evaluation of wind resources and long-term strategy of wind generators. Weibull is the most commonly used distribution and is recommended in many studies. In [2]–[7], researchers have performed the comparison of Weibull distribution with different distributions and recommended the two-parameter Weibull distribution. In [3], [8], [9], [28], researchers have used Rayleigh distribution and Drobinski *et al.* [28] recommended Rayleigh distribution after assessing four distributions for a location in France. Likewise, [3], [8]–[10], [25] compared Gamma distribution with other distributions. Lognormal has been used by [3], [10]–[13] and recommended for Iran and Algeria by Alavi *et al.* [3] and Aries *et al.* [11], respectively. Similarly, Gumbel [10], [11], Logistic [10], [13], Cauchy and generalized Lindley distributions are also used in different studies for the analysis of wind power potential.

Most studies in Pakistan region have used just Weibull and Rayleigh distributions and no one have tested other distributions so far. In our study, we have used Weibull, Rayleigh, Gumbel, Log-normal, Logistic, Gamma, generalized Lindley and Cauchy distributions for finding out the best fit distribution for the mentioned location in Pakistan. In this section, we first review eight probability distributions and then present how to estimate unknown parameters of these distributions. Finally, several goodness-to-fit measures are given to evaluate the quality of different distributions.

### A. PARAMETRIC DISTRIBUTIONS

#### 1) WEIBULL DISTRIBUTION

Weibull is the most widely used distribution for estimation of wind power potential [22], [25], [29], [30]. Weibull distribution depends upon three characteristic parameters known as shape, scale and location parameters. For a two-parameter Weibull distribution, we just have shape and scale parameters. In our study, we calculated Weibull parameters by using the maximum goodness-of-fit estimation approach explained in Section 4.2. After determining shape and scale parameters, the PDF and cumulative distribution function (CDF) for a two-parameter Weibull distribution can be calculated by using (6) and (7).

$$f(V) = \left(\frac{k}{C}\right)\left(\frac{V}{C}\right)^{k-1} e^{\left[-\left({}^{V}\!/_{C}\right)^{k}\right]} \tag{6}$$

$$F(V) = 1 - e^{\left[-\left({}^{V}\!/_{C}\right)^{k}\right]} \tag{7}$$

where $V$ is the wind speed, $k$ is the dimensionless Weibull shape parameter, and $C$ is the Weibull scale parameter having the same unit as $V$.

The PDF gives us the probability of availability of wind at a given speed $V$, and the corresponding CDF tells us about the probability of wind speed that is equal to or lower than a given speed $V$ or within a given interval of wind speed.

### 2) RAYLEIGH DISTRIBUTION

Rayleigh distribution can be considered as a special case of Weibull distribution in which the shape parameter has a fixed value of two. Rayleigh distribution is commonly applied to determine the wind power potential [29], [30]. In order to analyze the wind velocity in two dimensions (i.e., time and speed) with an assumption that each component is uncorrelated, normally distributed with equal variance and zero mean, Rayleigh distribution is an important tool and provides better results. The PDF and CDF for Rayleigh distribution can be calculated by using (8) and (9), respectively.

$$f(V) = \frac{\pi}{2}\left(\frac{V}{V_{avg}^2}\right) e^{\left[-\left(\frac{\pi}{4}\right)\left(\frac{V}{V_{avg}}\right)^2\right]} \tag{8}$$

$$F(V) = 1 - e^{\left[-\left(\frac{\pi}{4}\right)\left(\frac{V}{V_{avg}}\right)^2\right]} \tag{9}$$

### 3) LOGISTIC DISTRIBUTION

From the family of continuous distribution, Logistic distribution can be represented by location and scale parameters. As compared to normal distribution, logistic distribution has heavier tails but its shape resembles the normal distribution. After the parameter estimation, the PDF and CFD of Logistic distribution are given by (10) and (11).

$$f(V) = \frac{e^{-\left(\frac{V-a}{b}\right)}}{b\left(1 + e^{-\left(\frac{V-a}{b}\right)}\right)^2} \tag{10}$$

$$F(V) = \frac{1}{1 + e^{-\left(\frac{V-a}{b}\right)}} \tag{11}$$

where $a$ is the location parameter and $b$ is the scale parameter.

### 4) GAMMA DISTRIBUTION

Gamma distribution also belongs to the family of continuous probability distributions consisting of shape and rate parameters. The PDF and CDF can be expressed by using (12) and (13).

$$f(V) = \frac{b^a}{\Gamma(a)} V^{a-1} e^{-bV} \tag{12}$$

$$F(V) = \frac{1}{\Gamma(a)} \gamma(a, bV) \tag{13}$$

where $a$ is the shape parameter and $b$ is the rate parameter.

### 5) LOG-NORMAL DISTRIBUTION

Log-normal distribution best fits the theoretical distribution if the sample data is log-normally distributed, meaning that the logarithm of the sample data follows a normal distribution. Like a normal distribution, Log-normal distribution is also represented by two parameters, i.e., mean and standard deviation. However, we have to calculate these parameters after taking the logarithm of the sample data and then express them as mean-log and standard-deviation-log of the distribution. After the estimation of these two parameters, the PDF and CFD of Log-normal distribution can be expressed by (14) and (15).

$$f(V) = \frac{1}{V\sigma\sqrt{2\pi}} e^{\left(-\frac{(\ln(V)-\mu)^2}{2\sigma^2}\right)} \tag{14}$$

$$F(V) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left[\frac{\ln(V) - \mu}{\sigma\sqrt{2}}\right] \tag{15}$$

where erf(·) is the error function that is represented by (16).

$$\operatorname{erf}(v) = \frac{2}{\sqrt{\pi}} \int_0^v e^{-t^2} dt \tag{16}$$

### 6) CAUCHY DISTRIBUTION

The PDF and CDF of Cauchy distribution are determined by using the functions expressed by location and scale parameters in (17) and (18).

$$f(V) = \frac{1}{\pi b}\left[\frac{b^2}{(V-a)^2 + b^2}\right] \tag{17}$$

$$F(V) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{V-a}{b}\right) \tag{18}$$

where $a$ is the location parameter and $b$ is the scale parameter.

### 7) GUMBEL DISTRIBUTION

The Gumbel distribution is an extreme value distribution. After the estimation of location and scale parameters, the PDF and CFD of Gumbel distribution can be expressed by using (19) and (20).

$$f(V) = \frac{1}{\beta} e^{-(z+e^{-z})} \tag{19}$$

$$F(V) = e^{-(e^{z})} \tag{20}$$

where $z = \frac{V-a}{b}$, $a$ is the location parameter and $b$ is the scale parameter.

### 8) GENERALIZED LINDLEY DISTRIBUTION

Zakerzadeh and Dolati (2009) introduced generalized Lindley distribution having three parameters $\alpha$, $\beta$ and $\theta$. The corresponding PDF and CFD of the generalized Lindley distribution can be expressed as 21 and 22, respectively.

$$f(V) = \frac{\theta^{\alpha+1} x^{\alpha-1} (\alpha + \beta x)^{-\theta x}}{(\beta + \theta)\Gamma(\alpha + 1)} \tag{21}$$

$$F(V) = 1 - \frac{\alpha(\beta+\theta)\Gamma(\alpha,\theta x) + \beta(\theta x)^{\alpha} e^{-\theta x}}{(\beta+\theta)\Gamma(\alpha+1)} \tag{22}$$

where $\Gamma(\alpha, z)$ is the gamma function.

### *B. PARAMETER ESTIMATION FOR STATISTICAL DISTRIBUTION*

Many approaches have been proposed to accurately estimate unknown parameters of statistical distributions introduced before, such as shape and scale parameters of Weibull distribution. The study conducted by [31] has just considered the empirical method for parameter estimation of Weibull and Rayleigh distributions, while [25] has compared the empirical method, maximum likelihood method, modified maximum likelihood method, energy factor method and graphical method for parameter assessment of Weibull distribution for Hawky's Bay, Pakistan. [7] has compared the results of Empirical Method of Justus (EMJ), Energy Pattern Factor Method (EPFM), Maximum Likelihood Estimation Method (MLM) with the results of optimization techniques for parameter selection. This study concluded that the optimization method performs considerably well.

In our study, we applied the Nelder–Mead method for parameter estimation of probability distributions shown in Section 4.1. The Nelder–Mead method tries to find the parameter which can minimize the goodness-to-fit distance. The Nelder–Mead method is a numerical optimization technique used extensively for statistical estimation problems [32]. It uses the concept of a simplex in the factor space and repetitively forms new simplexes by replicating one point in the hyper plane of the remaining points. By doing so, it becomes accustomed to local landscape and converges to the final minimum. The stopping criteria to end up the optimization involves the maximization of log-likelihood function. For better understanding of the readers the complete procedure in form of a flow chart is given is Fig. 3. After that, standard errors of the estimates are calculated to minimize the Kolmogorov-Smirnov distance.

### *C. GOODNESS TO FIT STATISTICS*

A lot of statistical tests have been used by different researchers. Among them, Coefficient of Determination ($R^2$), Kolmogorov-Smirnov (K-S) and Chi square test ($\chi^2$) are the most widely used tests. Jung *et al.* [33] have presented a methodology that extends the scope of wind energy yield on high resolutions grid measured at 58 stations of german weather servive, goodness of fit of 67 therratical CDF was evaluated based on Coefficient of Determination ($R^2$). Reference [34] have evaluated the PDFs of wind speed data of five sites of north Dakota based on ($R^2$), Kolmogorov-Smirnov (K-S) and Chi square test ($\chi^2$) to find the best fit distribution function. In addition, Cramer-von Mises test (CvM) and Anderson-Darling (A-D) test have also been used. A-D test [35] gives more weight to the deviations at the tail of distribution, whereas, K-S test is more sensitive near the center of the distribution. Application of these two tests together will help us to find better fit distribution based on both center and tail regions of the distribution [36]. The goodness-to-fit assessment may also be done by using Akaike information criterion (AIC) and Bayesian information criterion (BIC). These methods are based on log-likelihood function of the distribution and a lower value of AIC and BIC indicates better fitting.

The measure of flexibility of AIC depends on the number of parameters, with more parameters resulting in harsh penalties. Whereas measure of flexibility of BIC not only depends on number of parameters but also on the number of observations (i:e sample length), with more parameters resulting in harsher penalties, and the relative penalty increases as the number of observations increase. It is important to mention here that unlike other statistical tests used for goodness-of-fit only AIC and BIC have used number of parameters and sample length.

For evaluation of how well the aforementioned seven distributions fit the original data, we have applied $R^2$, K-S, $\chi^2$, CvM and A-D tests along with AIC and BIC. Statistical results of these tests will tell us about the best fit and the worst fit distribution for our data. A brief description of all test methods is given below.

#### 1) COEFFICIENT OF DETERMINATION

Masseran *et al.* [37] and Shin*et al.* [38] have used $R^2$ along with two varients $R^2_{pp}$ and $R^2_{QQ}$ associated with P-P and Q-Q probability polt approaches. $R^2$ is the most commonly used test for measuring the goodness of fit for different distributions and tells us about the variance of the observed probabilities when fitted on some predicted probabilities. $R^2$ is given by

$$R^2 = 1 - \frac{SSE}{SST} \tag{23}$$

where SSE is the summed square of residuals and SST is the total sum of squares.

#### 2) KOLMOGOROV-SMIRNOV TEST

K-S test gives absolute deviations between the sample distribution function and the specified theoretical CDF [39]. The empirical CDF for wind speed samples $x_1, x_2, x_3, \ldots, x_n$ can be calculated by

$$F(x) = \frac{1}{N}[n(i)] \tag{24}$$

where $N$ is the sample size and for $x_i$ arranged from lowest to highest value, $n(i)$ are the number of points smaller than $x_i$. The K-S statistics can be calculated by using

$$D = \max_{1\le i\le N}\left(F(x_i) - \frac{i-1}{N}, \frac{i}{N} - F(x_i)\right) \tag{25}$$

Smaller value of K-S statistic represents better fitting of the theoretical distribution.

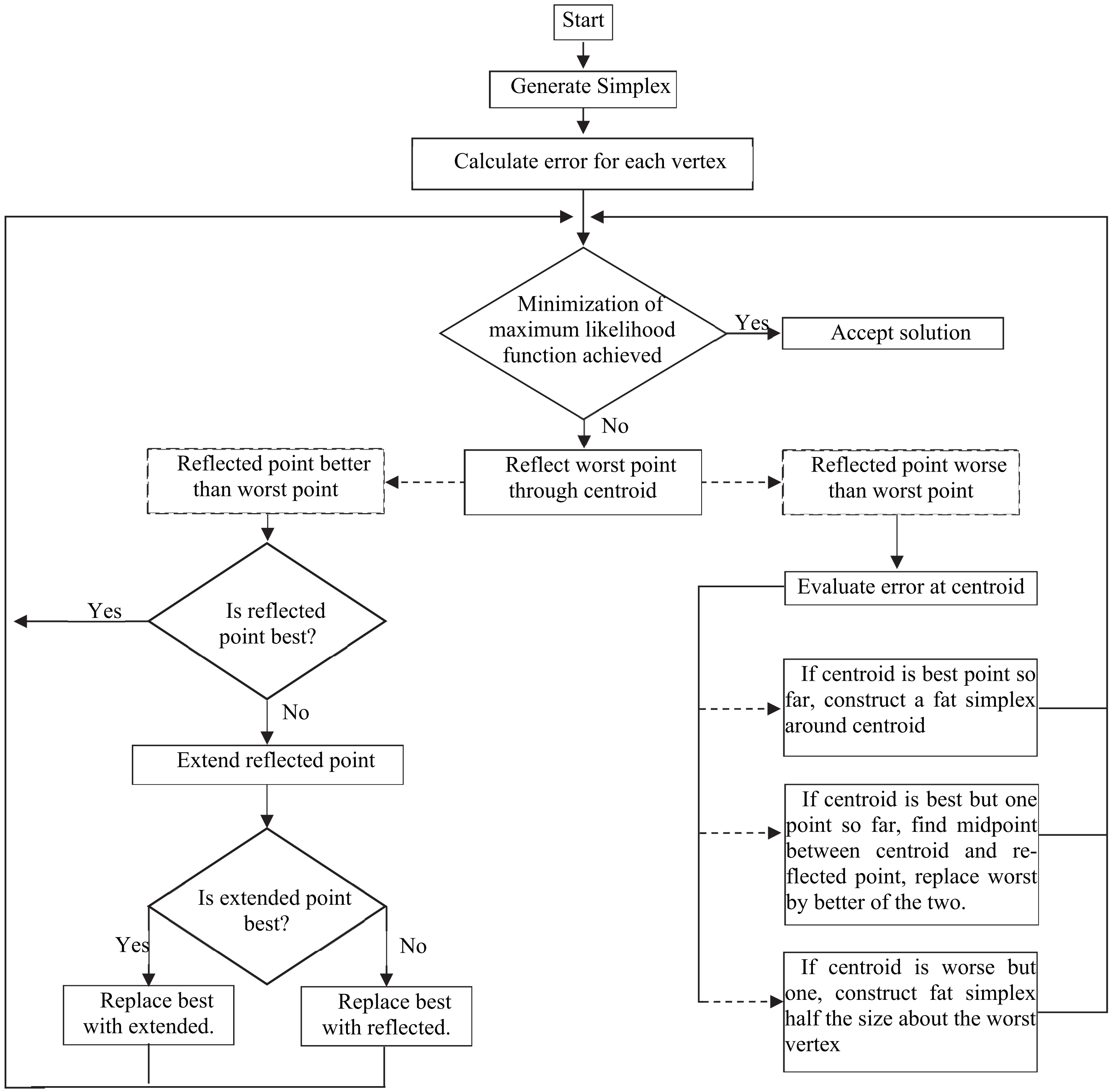



**FIGURE 3.** Flow chart for parameter selection of Nelder–Mead method.

### 3) CHI SQUARE TEST

$\chi^2$ test is used to statistically analyze how well the theoretical distribution is fitting the observed data and is used in other literatures [40]. The sample data is first divided into $k$ groups and then the difference between the observed frequency and the expected frequency of group $i$ is squared. Finally, the squared difference is divided by the expected frequency to give the error of that bin. These errors are then summed up for the overall error of probability distribution and can be represented mathematically by (26).

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{26}$$

where $O_i$ and $E_i$ are the observed and expected frequencies for group $i$. The expected frequency of each group $E_i$ can be calculated by using (27).

$$E_i = F(x_2) - F(x_1) \tag{27}$$

where $x_1$ and $x_2$ are the lower and upper limit of the group $i$, respectively. $F(\cdot)$ is the CDF of the theoretical distribution.

### 4) CRAMER-VON MISES TEST

CvM test is recommended for minimum distance estimation in statistic. For one sample application of empirical data arranged in the increasing order, CvM statistics can be calculated by (28).

$$CvM = n\omega^2 \tag{28}$$

where $\omega^2$ is defined by

$$\omega^2 = \int_{-}^{+} [F(v) - F^*(v)]^2 dF^*(v) \tag{29}$$

where $F(.)$ and $F^*(.)$ are CDFs of observed and theoretical distributions, respectively.

### 5) ANDERSON-DARLING TEST

A-D test is the modified version of K-S test and is also used to verify whether the sample data fits the specified distribution or not. The critical value of A-D test depends on the specified distribution and it gives more weight to the deviations at the tails of distributions. A smaller value of A-D statistic represents better fit of the theoretical curve and A-D statistic can be calculated by using (30).

$$AD = -N - \frac{1}{N}\sum_{i=1}^{N}(2i-1) \times \left\{\ln F(v_i) + \ln\left[1 - F\left(v_{n-(i+1)}\right)\right]\right\} \tag{30}$$

where $N$ is the total number of bins, $F(.)$ is the empirical CFD, and $v_i$ is the mean of wind speed in $i$-th bin.

### 6) AKAIKE INFORMATION CRITERION

AIC statistic is calculated by using model parameters based on maximum likelihood method and is used to determine the model accuracy. AIC is preferred due to its relation with maximum likelihood method for parameter estimation and is expressed by (31).

$$AIC = -2\ln(L) + 2k \tag{31}$$

where $k$ represents the number of parameters and $L$ is the maximum value of likelihood function.

### 7) BAYESIAN INFORMATION CRITERION

AIC and BIC are closely related to each other and the only difference is in the penalty term (the second term) of the BIC. Penalty term of BIC is larger than AIC, as the number of parameters is multiplied with the logarithm of the sample length rather than a fix number "2". BIC statistic depends on the length of the empirical data and can be calculated by using (32).

$$BIC = -2\ln(L) + k\ln(n) \tag{32}$$

## V. WIND POWER DENSITY CALCULATION

Wind speed can be converted into wind power. Given wind speed, Ideal wind power $P_{ideal}$ can be calculated by using (33).

$$P_{Ideal} = \frac{1}{2}\rho A_T V^3 \tag{33}$$

where $\rho$ is the air density in kg/m$^3$, $A_T$ is the rotor swept area in m$^2$ and $V$ is wind speed in m/sec.

Wind power cannot be entirely extracted by wind turbines because the air particles would stop in the intercepting area of the rotor, thereby blocking the cross-sectional area for the subsequent air particles. According to Betz, if the turbine swirl and transmission losses were ignored, 59% of the wind power available can theoretically be extracted from the wind. Therefore, for any wind turbine, the wind turbine efficiency should not exceed 0.59 percent of the ideal power that can be extracted from the wind. Theoretical maximum power that can be extracted from a given wind speed is $C_p$ time the ideal power known as $P_{Th}$ can be calculated by (34).

$$P_{Th} = \frac{1}{2}\rho C_P A_T V^3 \tag{34}$$

where $C_p$ is the coefficient of performance of the turbine, it is a function of tip speed ratio and the pitch angle. Theoretically, $C_p$ has a maximum value of 0.59 known as Betz limit. The variable wind speed turbine can track the maximum $C_p$ by adjusting the turbine speed according to the wind speed. Practically, we have to consider the turbine swirl and transmission losses, which results in further reduction of power extracted from wind. For practical purposes, we have to consider the thrust of wind and power coefficient $C_p$ is replaced by the Thrust coefficient $C_T$. The developed thrust T on the turbine blades is given by equation (35), see [R1.] for more details, as follows:

$$T = \frac{1}{2}\rho C_T A_T V^2 \tag{35}$$

By combining equation (34) and (35), we can produce equation (36).

$$P_{actual} = TV \tag{36}$$

In our study, we have considered the $P_{Th}$ for assessment of wind power as it is the theoretical maximum that can be achieved and fulfill the aim of this study that's is the assessment of wind power potential for Jhimpir, Pakistan. After selection of wind turbine and getting the value of thrust coefficient, Actual power extracted from the wind can be calculated by using (36).

Wind power density can be calculated by dividing the wind power $P_{Th}$ with the area of the turbine $A_T$. Inserting the value of $P_{Th}$ from (34) into (37), the wind power density can be calculated by using (38).

$$P_D = \frac{P}{A_T} \tag{37}$$

$$P_D = \frac{1}{2}\rho C_P V^3 \tag{38}$$

Similarly, wind power density using Weibull distribution function can be calculated by using (39).

$$P_{D_W} = \frac{1}{2}\rho C^3 \Gamma\left(1 + \frac{3}{k}\right) \tag{39}$$

And, wind power density using Rayleigh distribution function can be calculated by using (40).

$$P_{D_R} = \frac{3}{\pi}\rho V^3 \tag{40}$$

**TABLE 2.** Yearly average values of wind characteristics and Weibull parameters from 2015 to 2018.

| Year | $V_{avg}$ | $\sigma$ | $k$ | $C$ | $V_{mp}$ | $V_{maxE}$ | $P_D$ (W/m$^2$) | $E_D$ (kWh/m$^2$) |
|---|---|---|---|---|---|---|---|---|
| 2015 | 6.62 | 3.09 | 2.24 | 7.66 | 5.88 | 10.18 | 290.05 | 2988.014 |
| 2016 | 6.40 | 3.16 | 2.09 | 7.35 | 5.38 | 10.14 | 275.86 | 2849.628 |
| 2017 | 6.47 | 2.97 | 2.28 | 7.41 | 5.75 | 9.763 | 268.25 | 2763.403 |
| 2018 | 6.51 | 3.19 | 2.06 | 7.53 | 5.45 | 10.47 | 289.11 | 1729.886 |
| Complete | 6.50 | 3.09 | 2.17 | 7.48 | 5.64 | 10.09 | 279.84 | 10330.93 |

The difference between the wind power density calculated using the wind data recorded at site and the wind power density calculated through Weibull or Rayleigh distribution is considered as an error and this error can be calculated by using (41).

$$Error\ (\%) = \frac{P_{D_{W,R}} - P_{D_M}}{P_{D_M}} \tag{41}$$

where $\boldsymbol{P_{D_{W,R}}}$ is the wind power density calculated using Weibull or Rayleigh distribution and $\boldsymbol{P_{D_M}}$ is the wind power density calculated using wind data.

For better understanding of complete methodology id presented in form of a flow chart in Fig. 4.

## VI. RESULT AND DISCUSSION

Pakistan Meteorological Department (PMD) has identified the wind corridor in the southern region of Pakistan. Jhimpir is located in the wind corridor area in the southern part of Sindh province of Pakistan. In this study, we have analyzed wind characteristics and wind power potential of Jhimpir. Real-time wind data of 43 months from January 2015 to July 2018 of the investigated site is considered in this paper. For assessment of wind power potential of Jhimpir, detail analysis of wind characteristics is first carried out. For wind characteristics, we have analyzed the average wind speed on monthly and yearly bases, the availability of duration of wind at a particular speed and wind direction. Analysis of wind power density and wind energy density are also part of this study. Monthly, yearly and seasonally variations of wind characteristics are also analyzed for observation of seasonality pattern of the site.

Wind speed data was collected from two wind masts located at Jhimpir, Sindh Pakistan. The land surface is barren with very less population and no agricultural activities, making it suitable for installation of wind farms. Data set used for this study was given in tabular form and is stretched from January 01, 2015 to July 31, 2018 with total 31392 entries of wind speed and wind direction entries corresponding to 1308 hours. Range of average wind speed varies from 0 m/sec to 18.02 m/sec and wind direction ranges from zero degree to 356 degree. The wind speed data was recorded in meter per seconds and wind direction was recorded in degrees, at a height of 80 meters above ground level. The data was available from two wind masts. Thus, missing values from mast A are either imported from the wind mast B or forecasted by linear regression for further analysis.

### A. GOODNESS TO FIT STATISTICS

Time series plot of hourly wind speed is shown in Fig. 5. Mean wind speed for the complete data set is calculated to be 6.5 m/s and is represented by a black line on the graph. It is important to note that most of the time the average wind speed is high than 3 m/s (represented by a dotted line on the graph). Wind speed of 3 m/s is usually the cut-in speed for most of the commercially available wind turbines. Table 2 shows the average wind speed of each year. The average wind speed of each year is above 6 m/s, ranging from 6.40 m/s to 6.62 m/s.

Bar graph of monthly average wind speed of each year is shown in Fig. 6(a) and the overall average wind speed of each month is shown in Fig. 6(b). It is important to mention that in the month of May, June and July each year, the temperature goes high and these months are considered to be hot weather months. There is a rise in the demand of electricity in hot months, which can be supplied by wind power generation. However, for the months of November, December and January, the electricity demand is usually low because of the cold weather in this region. It can be seen from Fig. 6(a) that the average wind speed from the months of May, June, July and August are relatively high and the maximum mean monthly wind speeds for 2015, 2016, 2017 and 2018 are calculated to be 9.46m/s, 9.26m/s, 8.52m/s and 9.22m/s, respectively. Whereas, the minimum is in December 2015, November and December 2016, October 2017 and January 2018 (they are 4.56m/s, 4.07m/s, 4.46m/s and 4.28m/s, respectively). Fig. 6(b) shows that the overall monthly wind speed is high for hot weather months and is low for cold weather months. The average wind speed for each month along with the standard deviation are given in Table 3.

Monthly mean values of the most probable wind speed $V_{mp}$ and wind speed carrying maximum energy $V_{maxE}$ calculated by using equation 4 and 5 are given in Table 4. The highest values of $V_{mp}$ for each year are 10.1m/s, 9.68m/s, 8.75m/s and 9.41m/s in July 2015, May 2016, June 2017 and July 2018, respectively. Likewise, $V_{maxE}$ has high values of 11.8m/s for July 2015 and June 2016, 11.5m/s for July 2017 and 12.1m/s for June 2018. Yearly values of these two speeds are also shown in Table 2. The range of $V_{mp}$ varies from 5.38m/s to 5.88m/s and $V_{maxE}$ varies from 9.76m/s to 10.47m/s over the years.

### B. WIND SPEED DISTRIBUTION

#### 1) RESULTS OF PARAMETER ESTIMATION

Parameters of statistical distributions explained in Section 4.2 are estimated by using the numerical optimization technique

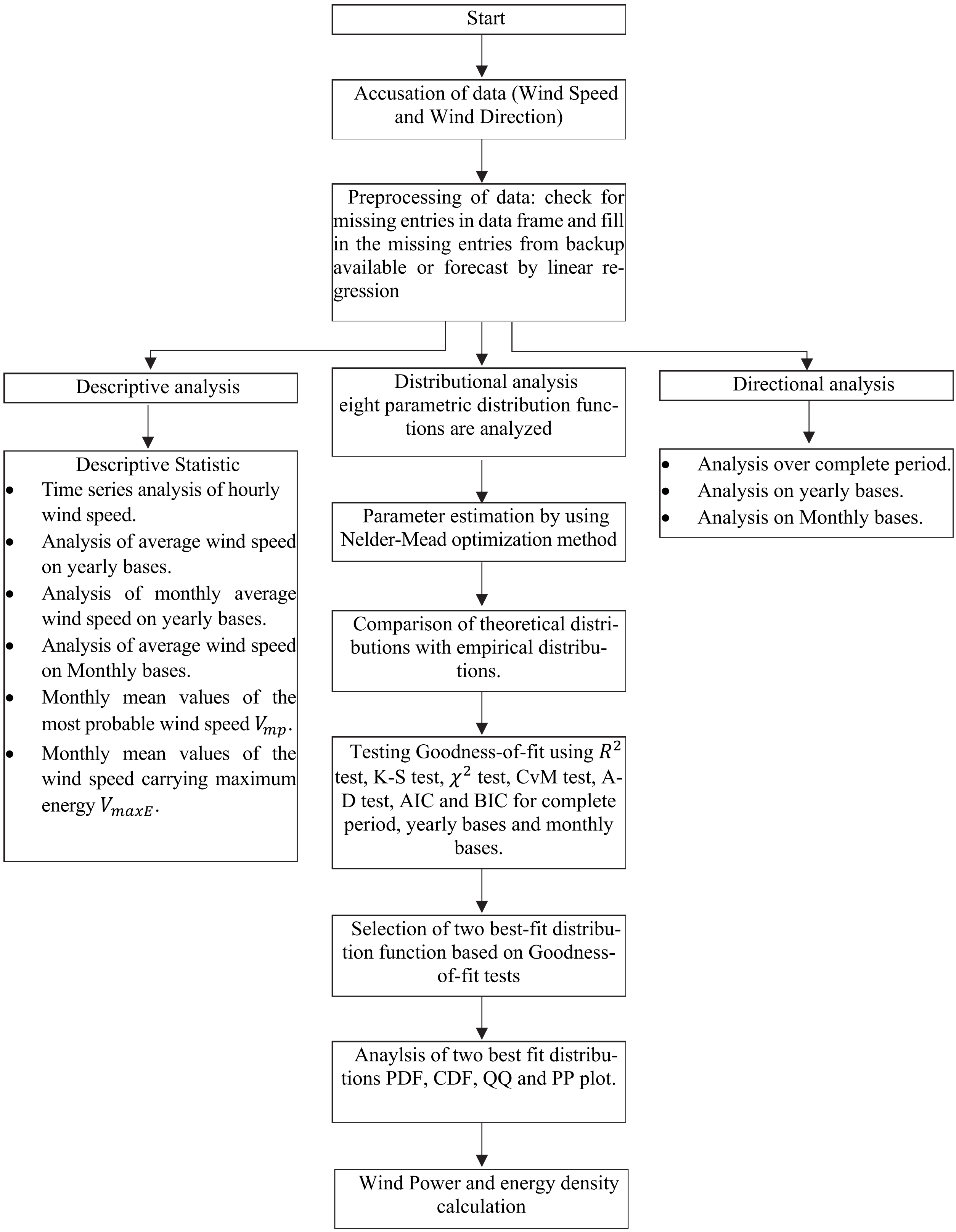



**FIGURE 4. Flowchart of proposed methodology.**

to minimize the goodness to fit distance (known as the Nelder-Mead optimization method). For any numerical optimization method we have to provide initial values before starting the process, initial values for the optimization process

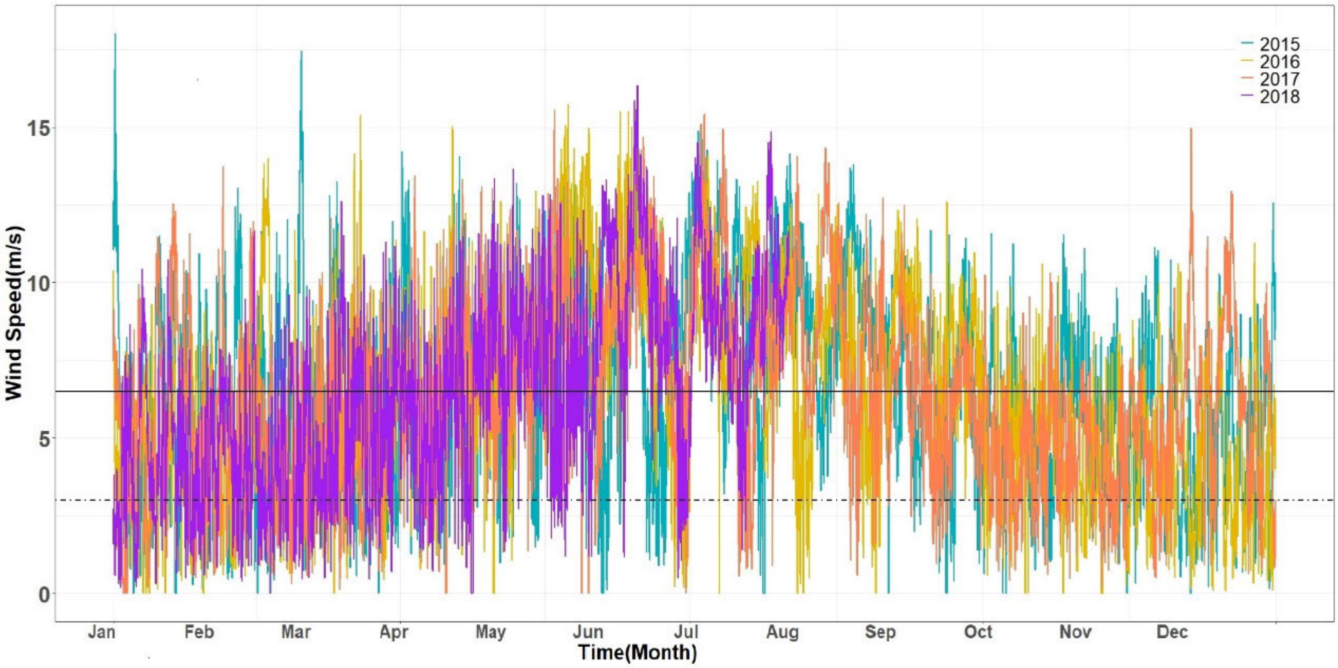


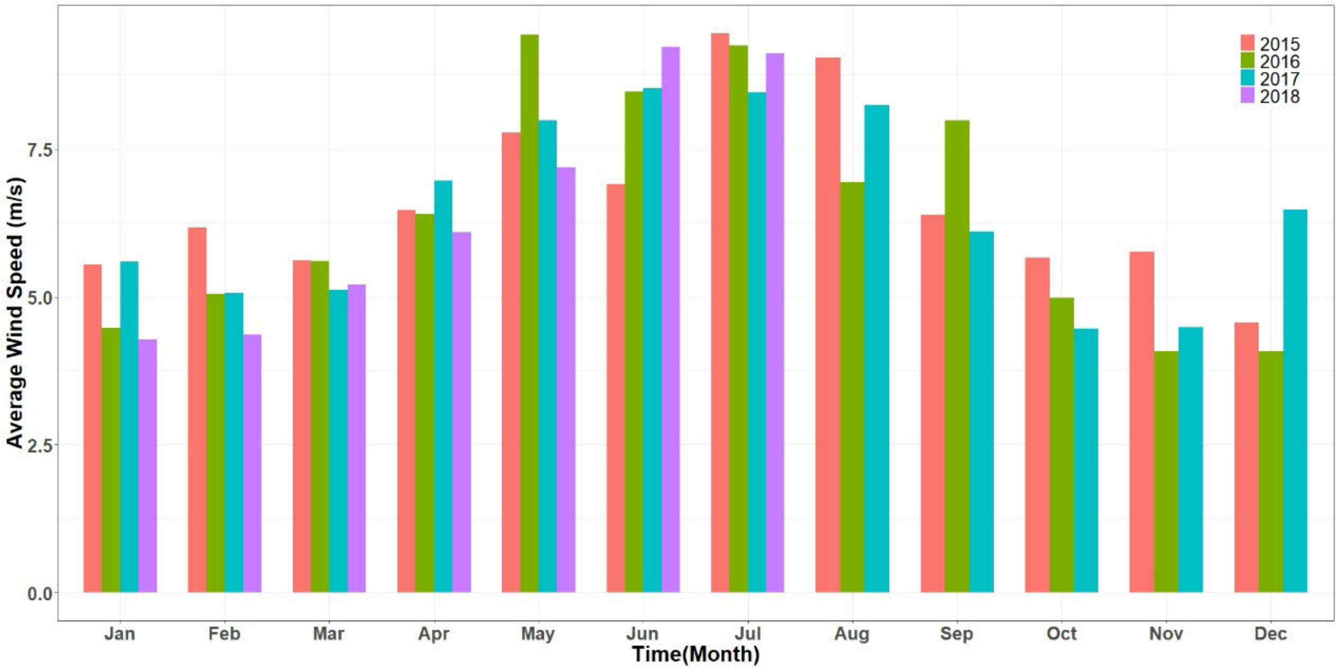

**FIGURE 5.** Hourly wind speed for Jhimpir, Pakistan.

**TABLE 3.** Monthly average values of wind speed, standard deviation and Weibull parameters from 2015 to 2018.

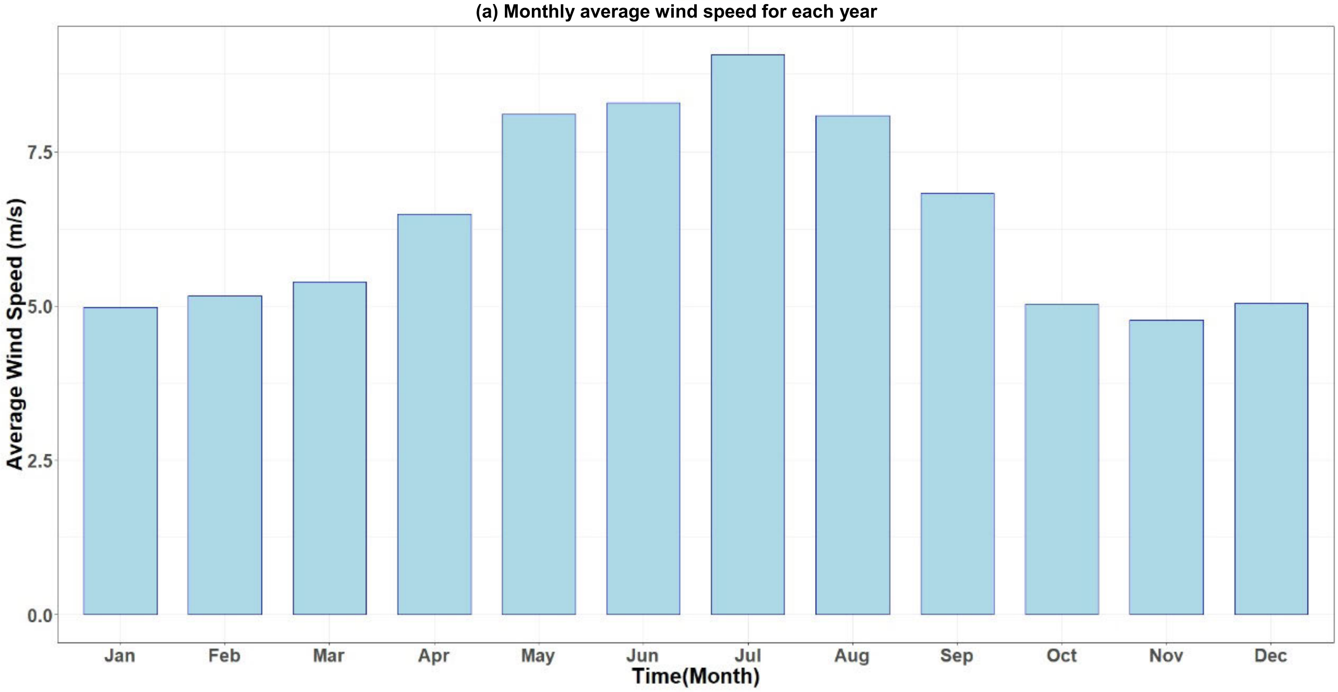
| | | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2015 | $V_{avg}$ | 5.55 | 6.18 | 5.62 | 6.47 | 7.78 | 6.91 | 9.46 | 9.05 | 6.38 | 5.66 | 5.77 | 4.56 |
| | $\sigma$ | 3.32 | 3.13 | 3.02 | 2.86 | 2.63 | 3.02 | 3.01 | 2.06 | 2.51 | 2.28 | 2.39 | 2.51 |
| | $k$ | 1.84 | 2.05 | 2.02 | 2.35 | 3.19 | 2.39 | 4.26 | 5.03 | 2.98 | 2.73 | 2.54 | 1.82 |
| | $C$ | 6.54 | 7.24 | 6.34 | 7.31 | 8.75 | 8.02 | 10.8 | 9.84 | 7.37 | 6.49 | 6.6 | 5.24 |
| 2016 | $V_{avg}$ | 4.47 | 5.11 | 5.61 | 6.4 | 9.44 | 8.48 | 9.26 | 6.94 | 7.99 | 4.98 | 4.07 | 4.07 |
| | $\sigma$ | 2.18 | 3.06 | 2.59 | 2.66 | 2.64 | 3.30 | 2.15 | 2.79 | 2.01 | 1.93 | 2.07 | 2.48 |
| | $k$ | 2.28 | 1.69 | 2.36 | 2.65 | 4.11 | 2.91 | 4.98 | 2.84 | 4.28 | 2.97 | 1.98 | 1.54 |
| | $C$ | 5.28 | 5.70 | 6.37 | 7.31 | 10.4 | 9.84 | 10.1 | 8.06 | 8.84 | 5.70 | 4.68 | 4.68 |
| 2017 | $V_{avg}$ | 5.59 | 5.07 | 5.12 | 6.97 | 7.98 | 8.52 | 8.46 | 8.24 | 6.10 | 4.46 | 4.48 | 6.47 |
| | $\sigma$ | 3.04 | 2.58 | 2.37 | 2.71 | 2.50 | 2.83 | 3.08 | 2.78 | 1.98 | 1.94 | 1.88 | 2.79 |
| | $k$ | 1.87 | 2.09 | 2.19 | 2.88 | 3.81 | 3.48 | 3.14 | 3.27 | 3.63 | 2.4 | 2.59 | 2.45 |
| | $C$ | 6.63 | 5.67 | 5.89 | 7.96 | 8.91 | 9.65 | 9.81 | 9.28 | 6.89 | 5.01 | 5.16 | 7.41 |
| 2018 | $V_{avg}$ | 4.28 | 4.36 | 5.21 | 6.09 | 7.19 | 9.22 | 9.12 | - | - | - | - | - |
| | $\sigma$ | 2.28 | 2.25 | 2.48 | 2.55 | 2.61 | 3.17 | 2.31 | - | - | - | - | - |
| | $k$ | 1.87 | 2.05 | 2.13 | 2.72 | 2.92 | 3.23 | 4.74 | - | - | - | - | - |
| | $C$ | 4.81 | 5.00 | 5.90 | 7.09 | 8.20 | 10.4 | 9.89 | - | - | - | - | - |

**TABLE 4.** Most probable wind speed and wind speed carrying maximum energy of each month from 2015 to 2018.

| Month | 2015 | | 2016 | | 2017 | | 2018 | |
|---|---|---|---|---|---|---|---|---|
| | $V_{mp}$ | $V_{maxE}$ | $V_{mp}$ | $V_{maxE}$ | $V_{mp}$ | $V_{maxE}$ | $V_{mp}$ | $V_{maxE}$ |
| Jan | 4.27 | 9.76 | 4.10 | 6.95 | 4.4 | 9.79 | 3.19 | 7.10 |
| Feb | 5.24 | 10.1 | 3.34 | 9.06 | 4.15 | 7.81 | 3.62 | 6.97 |
| Mar | 4.51 | 8.92 | 5.04 | 8.26 | 4.46 | 7.92 | 4.38 | 8.06 |
| Apr | 5.77 | 9.5 | 6.11 | 9.04 | 6.87 | 9.56 | 5.99 | 8.69 |
| May | 7.78 | 10.2 | 9.68 | 11.4 | 8.23 | 9.95 | 7.11 | 9.80 |
| Jun | 6.40 | 10.3 | 8.51 | 11.8 | 8.75 | 11.0 | 9.30 | 12.1 |
| Jul | 10.1 | 11.8 | 9.66 | 10.8 | 8.68 | 11.5 | 9.41 | 10.7 |
| Aug | 9.41 | 10.5 | 6.91 | 9.72 | 8.30 | 10.7 | - | - |
| Sep | 6.42 | 8.76 | 8.31 | 9.67 | 6.30 | 7.77 | - | - |
| Oct | 5.49 | 7.95 | 4.96 | 6.78 | 4.00 | 6.45 | - | - |
| Nov | 5.42 | 8.3 | 3.28 | 6.67 | 4.27 | 6.44 | - | - |
| Dec | 3.39 | 7.86 | 2.36 | 8.06 | 5.98 | 9.46 | - | - |




(a) Monthly average wind speed for each year




(b) Monthly average wind speed for all four years

**FIGURE 6.** Monthly average wind speed for Jhimpir, Pakistan.

can be fixed up with a constant vector of all ''1's'' or it can be a vector of random variables. In this paper, we have programmed the algorithm by adding the stochastic variable that randomly choose the value and start the iterations that leads to the convergence of program. Moreover, It is worth mentioning here that the complete data set is required for parameter estimation of statistical distributions. Table 5 shows the estimated parameters of each distribution calculated by using the data of every year separately and the estimated parameters calculated by using the complete data are also given in the last two columns of Table 5.

In Gamma distributions, P1 and P2 represent the rate and scale parameters, respectively. P1 value for each year varies from 3.5 to 4.19 and it has a value of 3.86 for the complete data. Likewise, P2 has a minimum value of 0.51 in 2018 and

**TABLE 5. Parameters of each distribution estimated by the Nelder-Mead optimization method.**

| Distribution | 2015 | | | 2016 | | | 2017 | | | 2018 | | | All Years | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P1 | P2 | P3 | P1 | P2 | P3 | P1 | P2 | P3 | P1 | P2 | P3 |
| Gamma | 4.03 | 0.58 | -- | 3.60 | 0.54 | -- | 4.19 | 0.63 | -- | 3.50 | 0.51 | -- | 3.86 | 0.57 | -- |
| Weibull | 2.24 | 7.66 | -- | 2.09 | 7.35 | -- | 2.28 | 7.41 | -- | 2.06 | 7.53 | -- | 2.18 | 7.47 | -- |
| Log-normal | 1.82 | 0.52 | -- | 1.77 | 0.55 | -- | 1.79 | 0.51 | -- | 1.79 | 0.56 | -- | 1.79 | 0.53 | -- |
| Cauchy | 6.65 | 1.52 | -- | 6.34 | 1.57 | -- | 6.40 | 1.42 | -- | 6.42 | 1.57 | -- | 6.46 | 1.51 | -- |
| Logistic | 6.61 | 1.78 | -- | 6.34 | 1.83 | -- | 6.39 | 1.70 | -- | 6.38 | 1.86 | -- | 6.42 | 1.77 | -- |
| Gumbel | 5.34 | 2.73 | -- | 5.02 | 2.78 | -- | 5.18 | 2.59 | -- | 5.12 | 2.88 | -- | 5.17 | 2.72 | -- |
| Rayleigh | 2.00 | 7.65 | -- | 2.00 | 7.34 | -- | 2.00 | 7.45 | -- | 2.00 | 7.50 | -- | 2.00 | 7.47 | -- |
| Lindley | 3.22 | 0.59 | 1.58 | 2.83 | 0.55 | 1.49 | 3.43 | 0.64 | 1.60 | 2.93 | 0.56 | 1.33 | 3.10 | 0.59 | 1.52 |

a maximum value of 0.63 in 2017. For the complete data set, it has the value of 0.57.

Weibull and Rayleigh distributions are represented by the shape parameter (P1) and the scale parameter (P2). The only difference between them is that the shape parameter (P2) of Rayleigh distribution has a fix value of 2. It can be observed that there is not too much difference between P1 and P2 of these two distributions. P1 of Weibull ranges from 2.06 to 2.24 for yearly analysis and has a value of 2.18 for the complete data set, whereas for Rayleigh, its P1 has a fix value of 2. Similarly, P2 of Weibull ranges from 7.35 to 7.66 and for Rayleigh, its P2 lies between 7.34 and 7.65 for yearly analysis. On the other hand, P1 (i.e., the parameter $k$) and P2 (i.e., the parameter $C$) estimated on monthly bases for Weibull distribution are also given in Table 3.

P1 and P2 of Log-normal distributions represent the mean and standard deviation of the data. For yearly analysis, the value of mean and standard deviation lies in the intervals [1.77, 1.82] and [0.51, 0.56], respectively. For the complete data, P1 and P2 are 1.79 and 0.53, respectively.

For Cauchy, Logistics and Gumbel, distributions P1 and P2 represent the location and scale parameters, respectively. For yearly analysis, the location parameter (P1) for Cauchy, Logistics and Gumbel ranges from [6.34, 6.65], [6.34, 6.36] and [5.02, 5.34]. The scale parameter (P2) ranges from [6.34, 6.65], [1.70, 1.86] and [5.02, 5.34], respectively. Likewise, for the complete data, P1 values are 6.46, 6.42 and 5.17 and P2 values are 1.51, 1.77 and 2.72 for these three distributions, respectively.

P1, P2 and P3 represents three parameters of generalized Lindley distribution. P1 value for each year varies from 2.83 to 3.43 and it has a value of 3.10 for the complete data. P2 ranges from 0.55 to 0.64 and for complete data it has a value of 0.59. Similarly for yearly analysis P3 ranges from 1.33 to 1.66. For the complete data set, it has the value of 1.52.

### 2) COMPARISON OF FREQUENCY DISTRIBUTION

It is better to arrange the data in frequency distribution for statistical analysis and compare theoretical distributions with empirical distributions. For this reason, we have organized the complete time series data of wind speed ranging from January 2015 to July 2018 in the frequency distribution format and results are shown in Table 6. Depending on the wind speed characteristic and for better comparisons of the original data with parametric distributions applied, our data is divided into 20 bins with each bin having a range of 1m/s mentioned in Column 1. Column 2 shows the number of times that the particular wind speed of time series data lies between the mentioned range in the corresponding row. From Column 2, it can be observed that the maximum number of times is 3788 and wind speed lies between 6-7m/s. The second, third, fourth and fifth most occurring speed lies between 7-8m/s, 5-6m/s, 8-9m/s and 4-5m/s and they have occurred 3561, 3491, 3128 and 3018 times, respectively.

Probability density distribution of the actual wind speed data is shown in Column 3. Probability density tells us about how the wind speed of a particular range is distributed. Wind speed bins of 6-7m/s, 7-8m/s, 5-6m/s, 8-9m/s and 4-5m/s are top five probabilities of actual distribution in the ascending order. The sum of probabilities of wind speed from 4m/s to 9m/s is 0.5442 showing that wind speed was within this range for about 54.42% of the total time. For the actual wind speed data, the highest probability has a value of 0.1214 for the bin of 6-7m/s, followed by 0.1141 and 0.1118 for bins of 7-8m/s and 5-6m/s, respectively.

Likewise, results extracted from PDFs of different theoretical distributions are given in Column 4-11. From Table 6, it can be seen that top five probabilities of Gamma, Weibull, Log-normal, Gumbel, Rayleigh and generalized Lindley distributions have occurred at bins of 3-4m/s, 4-5m/s, 5-6m/s, 6-7m/s and 7-8m/s, which are different from top five bins of the actual wind speed distribution. Among these six distributions, Gamma, Weibull, Gumbel and Rayleigh have their maximum value of probabilities as 0.1289, 0.1210, 0.1332 and 0.1143 for wind speed bin of 5-6m/s. whereas, Log-normal and generalized Lindley distributions have their maximum value of probability as 0.1429 and 0.1295 for wind speed bin of 4-5m/s, which is also different from the peak probability of the actual data.

For Cauchy and Logistic distributions, top five probabilities occurred at the same wind speed bins with those of the actual wind distribution, i.e., five bins between 4 m/s to 9 m/s. Moreover, the highest probabilities of these two distributions

**TABLE 6.** Frequency distributions of actual wind speed data and tested distributions using the complete data.

| Bins | $f_i$ | Actual | Gamma | Weibull | Log-normal | Cauchy | Logistic | Gumbel | Rayleigh | Lindley |
|---|---|---|---|---|---|---|---|---|---|---|
| 0-1 | 451 | 0.0145 | 0.0038 | 0.0125 | 0.0004 | 0.0860 | 0.0450 | 0.0100 | 0.0178 | 0.0068 |
| 1-2 | 1601 | 0.0513 | 0.0318 | 0.0427 | 0.0185 | 0.0181 | 0.0315 | 0.0311 | 0.0514 | 0.0408 |
| 2-3 | 2573 | 0.0824 | 0.0753 | 0.0729 | 0.0757 | 0.0271 | 0.0505 | 0.0683 | 0.0797 | 0.0841 |
| 3-4 | 2722 | 0.0872 | 0.1109 | 0.0980 | 0.1260 | 0.0443 | 0.0766 | 0.1063 | 0.1003 | 0.1159 |
| 4-5 | 3018 | 0.0967 | 0.1285 | 0.1145 | 0.1429 | 0.0803 | 0.1063 | 0.1296 | 0.1118 | 0.1295 |
| 5-6 | 3491 | 0.1118 | 0.1289 | 0.1210 | 0.1345 | 0.1507 | 0.1311 | 0.1332 | 0.1143 | 0.1270 |
| 6-7 | 3788 | 0.1214 | 0.1174 | 0.1177 | 0.1145 | 0.2034 | 0.1399 | 0.1214 | 0.1090 | 0.1141 |
| 7-8 | 3561 | 0.1141 | 0.0999 | 0.1065 | 0.0922 | 0.1436 | 0.1280 | 0.1018 | 0.0980 | 0.0960 |
| 8-9 | 3128 | 0.1002 | 0.0807 | 0.0901 | 0.0718 | 0.0760 | 0.1016 | 0.0805 | 0.0834 | 0.0770 |
| 9-10 | 2438 | 0.0781 | 0.0626 | 0.0716 | 0.0548 | 0.0423 | 0.0720 | 0.0612 | 0.0676 | 0.0594 |
| 10-11 | 1770 | 0.0567 | 0.0471 | 0.0536 | 0.0414 | 0.0261 | 0.0470 | 0.0452 | 0.0523 | 0.0444 |
| 11-12 | 1374 | 0.0440 | 0.0345 | 0.0378 | 0.0311 | 0.0175 | 0.0291 | 0.0327 | 0.0387 | 0.0324 |
| 12-13 | 753 | 0.0241 | 0.0247 | 0.0252 | 0.0233 | 0.0125 | 0.0174 | 0.0234 | 0.0274 | 0.0231 |
| 13-14 | 372 | 0.0119 | 0.0174 | 0.0158 | 0.0175 | 0.0093 | 0.0102 | 0.0166 | 0.0186 | 0.0162 |
| 14-15 | 126 | 0.0040 | 0.0120 | 0.0094 | 0.0131 | 0.0072 | 0.0059 | 0.0117 | 0.0121 | 0.0111 |
| 15-16 | 34 | 0.0011 | 0.0082 | 0.0053 | 0.0099 | 0.0057 | 0.0034 | 0.0082 | 0.0076 | 0.0076 |
| 16-17 | 5 | 0.0002 | 0.0056 | 0.0028 | 0.0075 | 0.0047 | 0.0019 | 0.0057 | 0.0046 | 0.0051 |
| 17-18 | 6 | 0.0002 | 0.0037 | 0.0014 | 0.0057 | 0.0039 | 0.0011 | 0.0040 | 0.0026 | 0.0034 |
| 18-19 | 1 | 0.0000 | 0.0025 | 0.0007 | 0.0043 | 0.0033 | 0.0006 | 0.0028 | 0.0015 | 0.0022 |
| 19-20 | 0 | 0.0000 | 0.0045 | 0.0005 | 0.0150 | 0.0382 | 0.0008 | 0.0063 | 0.0016 | 0.0040 |

are calculated to be 0.2034 and 0.1399 falling in the wind speed bin of 6-7m/s.

Besides doing the analysis of the complete data set, we have also done these calculations on the yearly base. Fig. 7 is the graphical representation of PDFs of actual and theoretical distributions for the year 2015. The X-axis represents the wind speed and the Y-axis represents the probability density. PDFs of all applied distributions are shown by different line types over the observed distribution represented in form of histogram. Although peak densities of Gamma, Weibull, Log-normal, Gumbel, Rayleigh and generalized Lindley distributions are not in line with the peak density of the actual distribution, it can be observed from naked eyes that these distributions can fit the actual distribution very well. For Cauchy and Logistic distributions, they have better fitting across the tail region. Whereas in the central region, their peak densities are far away from the actual distribution.

### 3) GOODNESS TO FIT STATISTICS

In this study, the comparison of goodness-to-fit statistics among eight statistical distributions for wind speed at Jhimpir, Pakistan is performed using $R^2$ test, K-S test, $\chi^2$ test, CvM test, A-D test, AIC and BIC given in Section 4.3. The statistical value of each distribution for the complete data is given in Table 7. Each distribution is then ranked according to these statistics in the ascending order mentioned by a number given in small brackets. It is important to mention that $R^2$ is a positively oriented test, which means that the value of $R^2$ close to 1 represents better fitting of the distribution. Other tests are negatively oriented and smaller values represent better goodness-to-fit.

Based on all tests performed, Weibull is always ranked as number one, meaning that Weibull is best fitting the actual distribution. For $R^2$ statistics in Table 7, Weibull has the highest value of 0.974, followed by Rayleigh that is having the second highest value of 0.965. This shows that based on $R^2$ test, Weibull and Rayleigh are two best distributions for our data. Based on K-S tests, Weibull and Logistic are found to be two best fit distributions with statistical values of 0.0203 and 0.0322, respectively. Whereas, Cauchy and generalized Lindley distributions are having statistical values of 0.0763 and 0.06260, respectively, showing that these distributions are not suitable for our data.

Likewise, statistical values of $\chi^2$, CvM, A-D, AIC and BIC tests for Weibull are calculated to be 733.5796, 5.128, 48.4, 157333.1 and 157349.8, respectively. Weibull is ranked in the first position representing the best fit distribution over the actual data. Rayleigh is ranked as the second best distribution with statistical values of 1181.229, 13.82, 108.92, 158019.3 and 158027.6 for the above mentioned tests. Based on $\chi^2$ test, Log-normal is ranked eighth and Cauchy is at number seven on the list. For CvM, A-D, AIC and BIC tests, Cauchy and Log-normal are at number eight and seven, respectively.

The comparison of statistical values of seven goodness-to-fit tests applied on yearly bases is shown in Fig. 8. As $R^2$ is a positive orientated test, it can be observed that for all years, Weibull distribution has the maximum $R^2$ value followed by Rayleigh distribution. In 2016 and 2018, values of Weibull distribution are calculated to be 0.971 and 0.941. Whereas for Rayleigh distribution function these values are observed to be 0.970 and 0.940. The difference of statistical value between these distribution is calculated to be very small i:e 0.001, but it cannot be ignored. These statistical values showing that Weibull is best fit distribution for our data. Cauchy is least suitable for our data as it has the lowest $R^2$ value for all years.

For K-S test, Weibull is found to be the best fit distribution in each year. However, there is strong competition for the second, third and fourth best fit distributions among all years. Gamma, Gumbel and Rayleigh are having very close

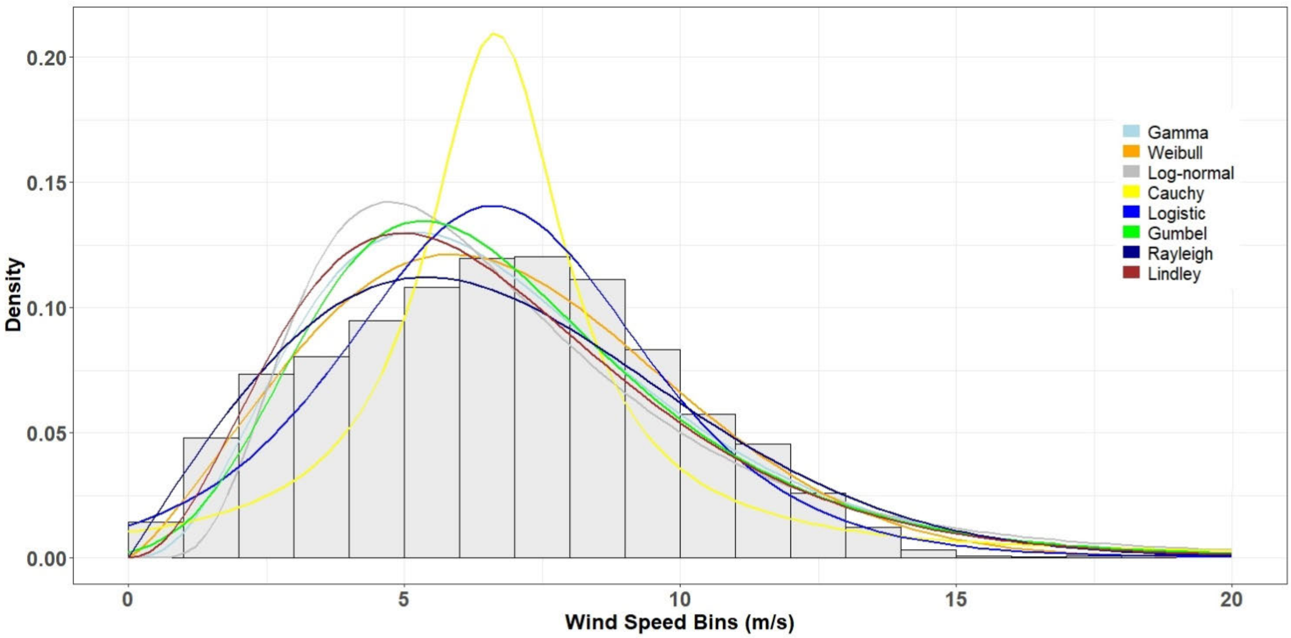



**FIGURE 7.** PDF comparison of all distributions for year 2015.

**TABLE 7.** Goodness-to-fit statistical indicators of different distribution functions for the whole data.

| | $R^2$ | K-S | $\chi^2$ | CvM | A-D | AIC | BIC |
|---|---|---|---|---|---|---|---|
| Gamma | 0.911(4) | 0.0385(4) | 3164.779(6) | 21.30(4) | 217.83(5) | 159971.0(5) | 159987.7(5) |
| Weibull | 0.974(1) | 0.0203(1) | 733.5796(1) | 5.128(1) | 48.400(1) | 157333.1(1) | 157349.8(1) |
| LogN | 0.815(7) | 0.0571(6) | 23022.26(8) | 48.14(7) | 564.61(7) | 166822.8(7) | 166839.5(7) |
| Cauchy | 0.586(8) | 0.0763(8) | 15844.04(7) | 81.70(8) | 601.50(8) | 174099.5(8) | 174116.2(8) |
| Logistic | 0.914(3) | 0.0322(2) | 2546.555(5) | 15.12(3) | 125.43(3) | 160049.6(6) | 160066.3(6) |
| Gembel | 0.907(5) | 0.0391(5) | 2541.110(4) | 22.02(5) | 193.92(4) | 159790.1(4) | 159806.8(4) |
| Rayleigh | 0.965(2) | 0.0345(3) | 1181.229(2) | 13.82(2) | 108.92(2) | 158019.3(2) | 158027.6(2) |
| Lindley | 0.902(6) | 0.0626(7) | 2443.300(3) | 40.49(6) | 233.65(6) | 159423.4(3) | 159448.4(3) |

statistical values of 0.0417, 0.0428 and 0.0432 in 2015. Whereas in 2016, Rayleigh is found to be the second best fit distribution with the statistical value of 0.0268, followed by Gamma and Gumbel (both having statistical values of 0.0374). Logistic is found to be the second best fit distribution in 2017 (having the statistical value of 0.0306), followed by Gamma, Gumbel and Rayleigh (having statistical values of 0.0358, 0.0377 and 0.0404). In 2018, once again Rayleigh came back to the second position with the statistical value of 0.0318, followed by Logistic, Gumbel and Gamma distributions. Overall, Rayleigh is found to be the second best fit distribution in 2016 and 2018. However, in 2015 and 2017, Logistic is in the second place.

Based on statistical values of $\chi^2$ test, Weibull distribution is the best fit distribution with statistical values of 260.476, 223.453, 157.726 and 201.736 in each year. Rayleigh has the second position on the list with statistical values of 467.829, 265.449, 395.117 and 221.976 from 2015 to 2018, respectively. Lognormal is the worst fit distribution for 2015, 2016 and 2017. But for the year 2018, the worst fit distribution is Cauchy distribution.

For CvM test, Weibull is ranked first with statistical values of 2.017, 1.462, 0.776 and 1.390 from 2015 to 2018, while Rayleigh is ranked second in 2016 and 2018 with statistical values of 2.116 and 1.636, respectively. Whereas, Logistic is ranked second in 2015 and 2017 with statistical values of 3.159 and 3.213, respectively. Cauchy and Log-normal fall at number six and seven in term of goodness-to-fit for each year. Statistical values of A-D test are analogous to results of CvM test. Weibull is ranked first for each year. Rayleigh is ranked second for 2016 and 2017. Logistic distribution is ranked second in 2015 and 2017. Bar graphs of AIC and BIC are shown in Fig. 8, respectively. It can be seen that Weibull is the best fit distribution based on statistical values of AIC and BIC for each year, followed by Rayleigh distribution which is the second best fit distribution for both cases every year.

Based of AIC, Weibull is the best fit distribution with statistical value of 43781.68, 44292.74, 43326.60 and 25864.04 in each year. Rayleigh stands at second position with statistical values of 44084.55, 44361.78, 43686.68 and 25895.38 from 2015 to 2018, respectively. Likewise, Minimum value of BIC is observed to be 43795.82, 44306.89, 43340.74 and

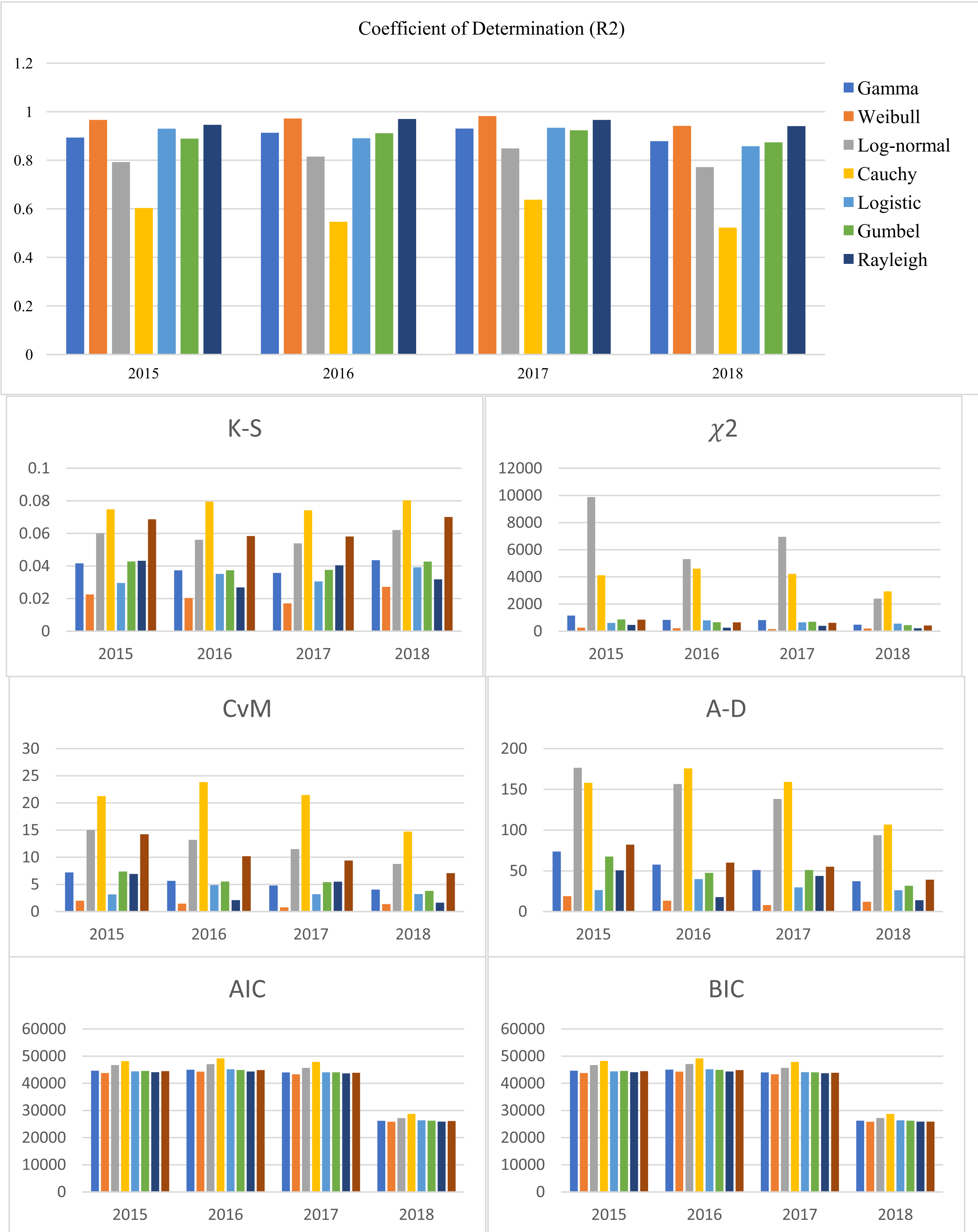
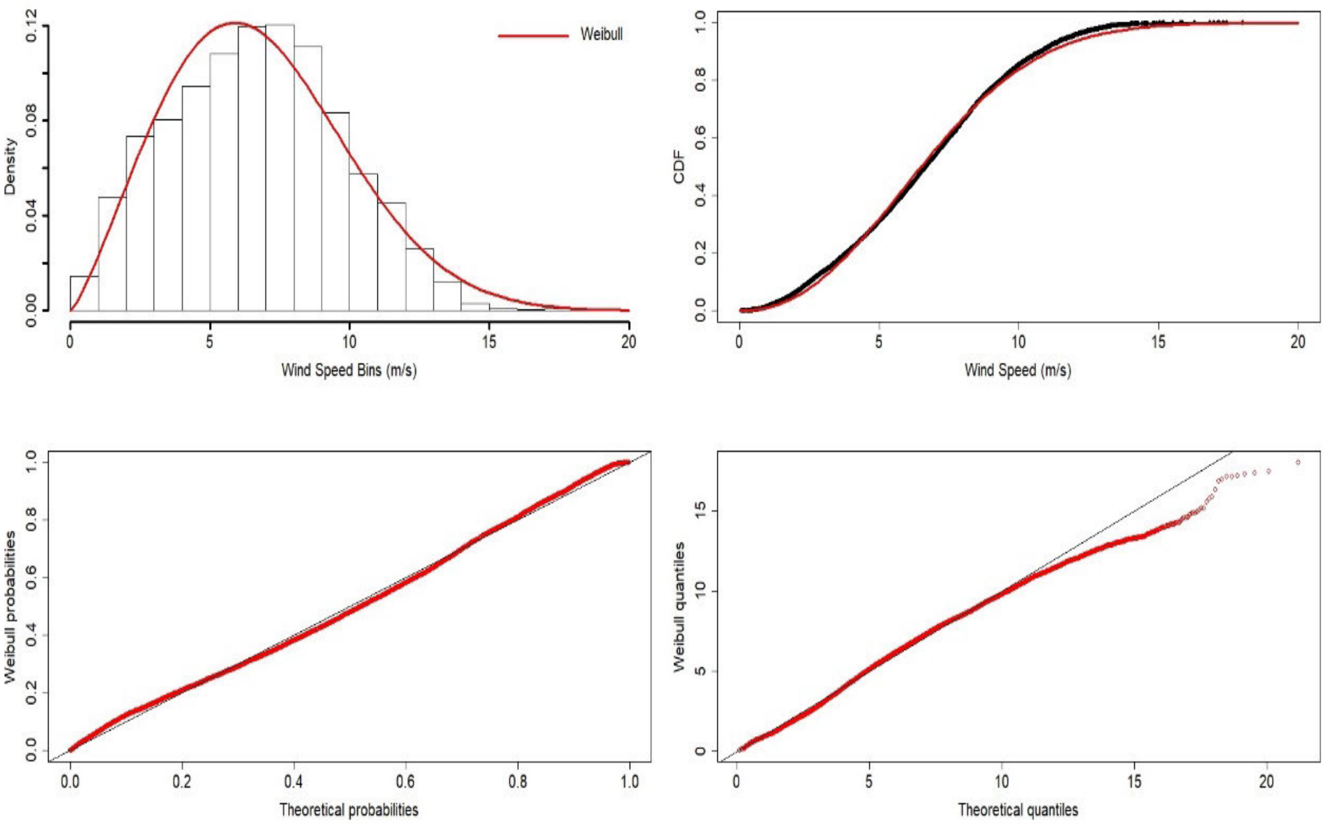
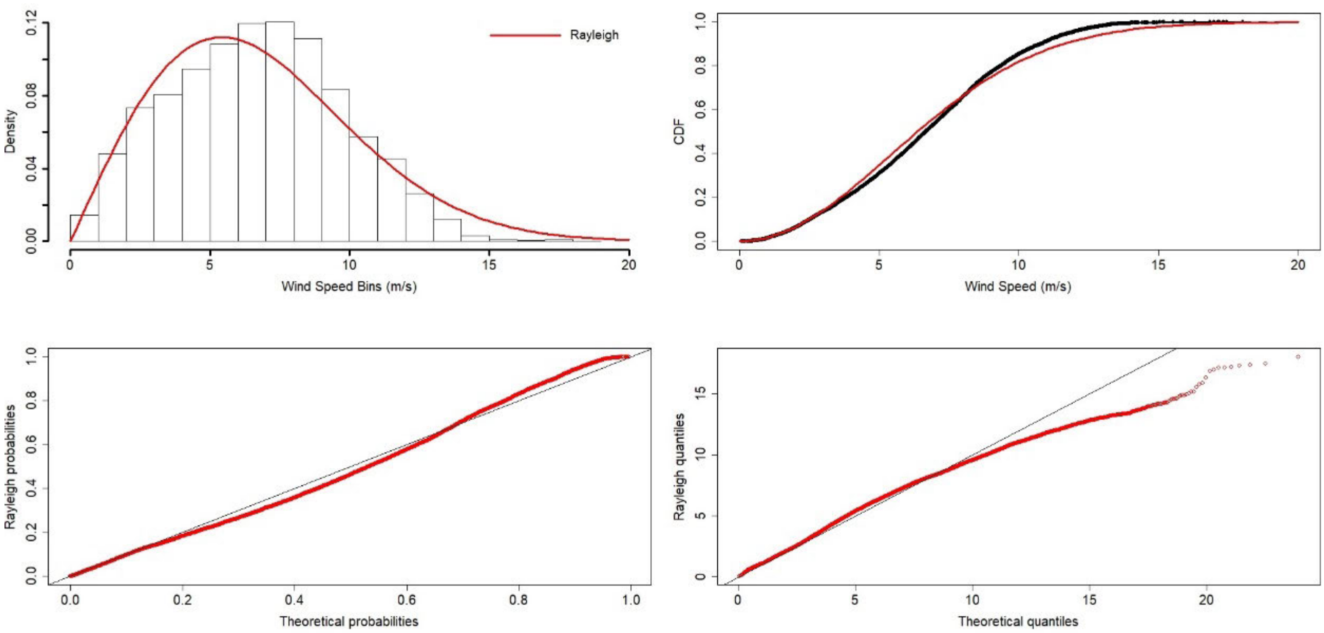



**FIGURE 8.** Yearly comparison of empirical distributions on goodness-to-fit indices using the actual wind data.




**(a) Weibull distribution of the year 2015**




**(b) Rayleigh distribution of the year 2015**

**FIGURE 9.** **PDF, CFD, PP plot and QQ plot for the comparison between Weibull and Rayleigh distributions.**

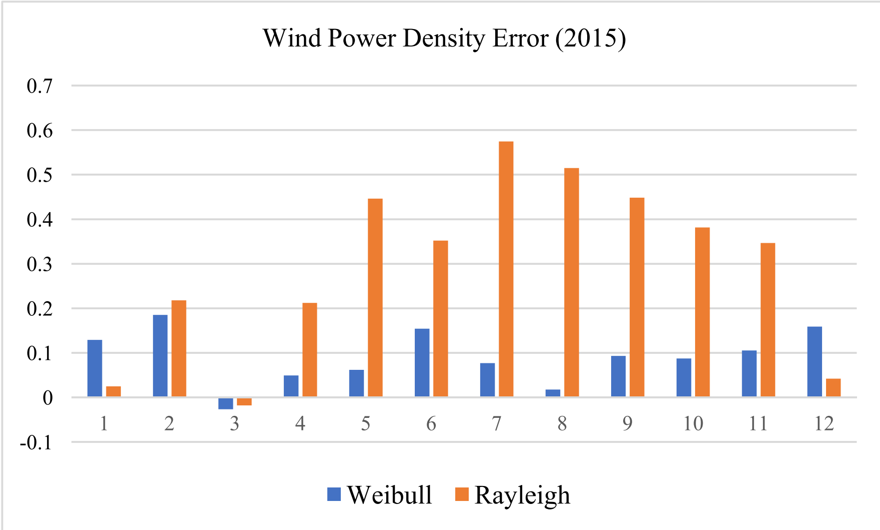



**FIGURE 10.** Wind power density error of Weibull and Rayleigh distributions in 2015.

**TABLE 8.** Monthly wind power densities calculated from wind speed data and two parametric distributions.

| | | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2015 | Actual | 222.13 | 254.07 | 210.94 | 262.10 | 377.39 | 310.70 | 648.73 | 511.97 | 224.96 | 161.47 | 174.15 | 112.41 |
| | Weibull | 250.78 | 301.09 | 205.36 | 275.02 | 400.81 | 358.70 | 698.65 | 521.07 | 245.89 | 175.61 | 192.49 | 130.30 |
| | Rayleigh | 227.65 | 309.48 | 207.15 | 317.74 | 545.83 | 420.07 | 1021.4 | 775.54 | 325.85 | 223.07 | 234.54 | 117.15 |
| 2016 | Actual | 91.459 | 177.50 | 179.00 | 240.75 | 620.40 | 519.86 | 550.05 | 294.47 | 362.58 | 106.80 | 73.672 | 90.438 |
| | Weibull | 105.96 | 186.91 | 181.40 | 254.40 | 622.80 | 590.92 | 564.30 | 328.34 | 385.10 | 113.87 | 84.677 | 133.34 |
| | Rayleigh | 119.74 | 150.78 | 210.30 | 317.79 | 905.08 | 775.16 | 839.26 | 425.79 | 563.30 | 150.76 | 83.681 | 110.73 |
| 2017 | Actual | 200.44 | 144.94 | 133.98 | 293.93 | 393.66 | 489.12 | 500.50 | 448.14 | 178.25 | 85.081 | 82.350 | 255.66 |
| | Weibull | 256.26 | 141.70 | 152.46 | 314.70 | 402.09 | 522.35 | 567.33 | 473.58 | 187.76 | 87.165 | 90.835 | 262.18 |
| | Rayleigh | 237.26 | 148.08 | 166.26 | 410.96 | 576.04 | 731.18 | 768.02 | 650.51 | 265.84 | 102.37 | 111.98 | 316.32 |
| 2018 | Actual | 90.120 | 92.207 | 146.2 | 205.77 | 309.59 | 627.69 | 541.36 | - | - | - | - | - |
| | Weibull | 92.588 | 94.491 | 149.25 | 213.83 | 319.07 | 653.50 | 553.89 | - | - | - | - | - |
| | Rayleigh | 90.417 | 101.99 | 167.24 | 290.46 | 449.23 | 924.55 | 788.71 | - | - | - | - | - |

25877.11 for Weibull distribution followed by Rayleigh distribution with statistical values of 44489.67, 44870.02, 43901.64 and 25901.91 for each year.

After implementing seven goodness-to-fit tests on yearly bases, we have a total of twenty eight scenarios to rank these distributions based on their statistical values. One important observation is that for all cases, Weibull distribution is ranked first and no other distributions has a better statistical value than Weibull distribution. Rayleigh is ranked second twenty two times, followed by Logistic distribution which is ranked second six times.

After evaluating the results of yearly analysis extracted from the complete data, it can be concluded that Weibull and Rayleigh are best suitable distributions for our data. For more in depth analysis and verification of our results, we have plotted PDF, CFD, Q-Q and P-P diagrams of these distributions for 2015, as shown in Fig. 9.

Fig. 9(a) and (b) show the detailed graphical comparison of the actual data with Weibull and Rayleigh distributions for 2015. PDFs of these theoretical distributions are represented by the solid line on bar graphs of the actual data distribution. It can be noted that the peak and tail of Weibull has better fitting with the actual data, while the peak and tail of Rayleigh are a little misaligned compared to Weibull distribution. Likewise, CFD, Q-Q and P-P plots also indicate that Weibull distribution has better fitting compared with Rayleigh distribution.

### C. WIND POWER AND ENERGY DENSITY

Wind power density discussed in Section 5 plays the pivot role for assessment of wind power potential for a candidate site. After quantitative and qualitative analysis of parametric distributions, we have concluded that Weibull and Rayleigh are two most suitable distributions for assessment of wind power potential at Jhimpir, Pakistan. A comparison of monthly wind power densities calculated from the actual wind data with wind power densities of Weibull and Rayleigh distributions is given in Table 8.

Wind power density calculated in the months of May, June, July and August is relatively higher than other months of the year. During these months, the temperature is usually

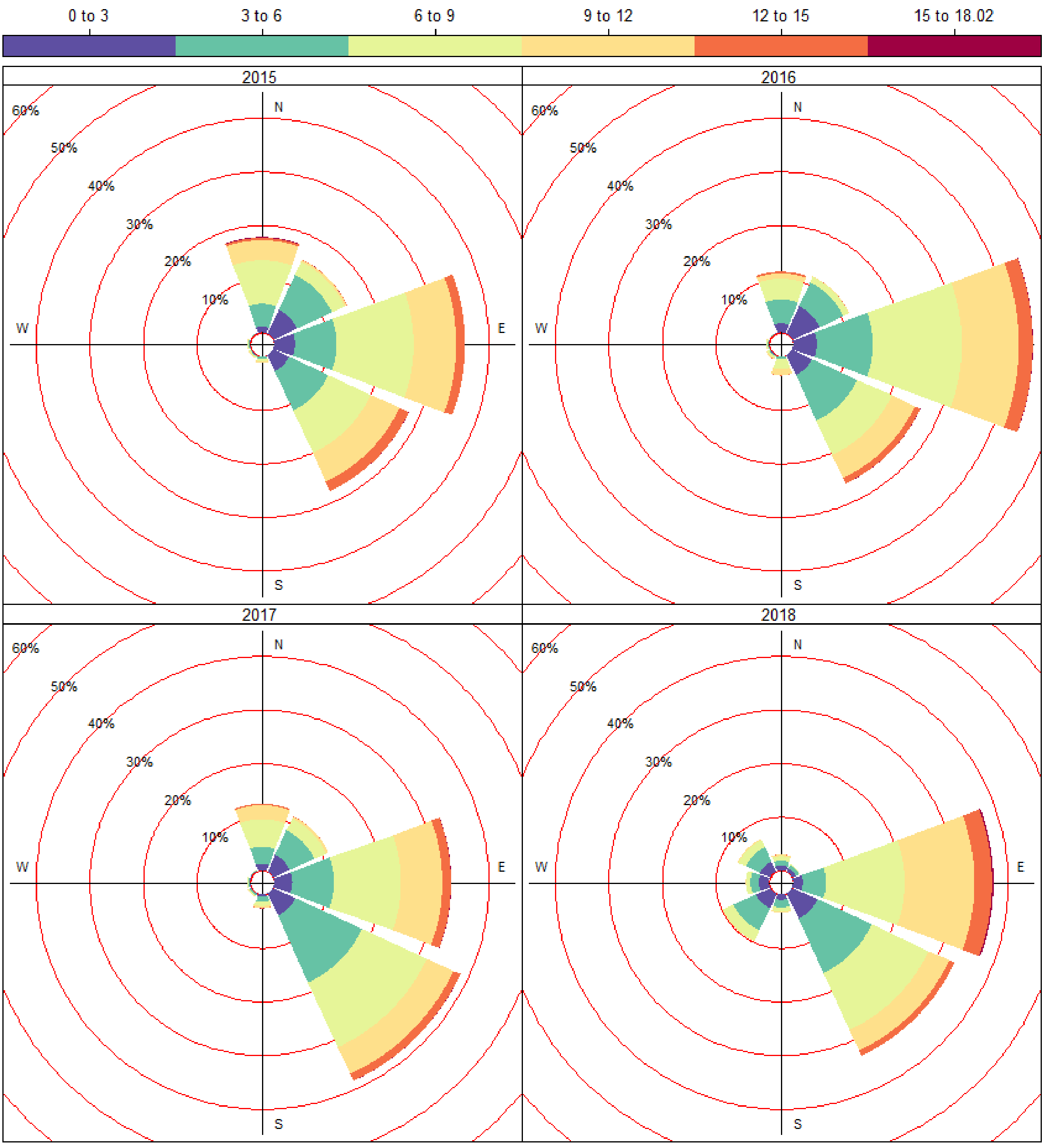



**FIGURE 11.** **Yearly wind rose graph of wind data collected at Jhimpir, Pakistan.**

high and these months are considered to be the summer in Pakistan. The demand of electricity is usually high in summer. The low value of wind power density is observed for the months ranging from October to January. These months are usually cold and the demand of electricity is usually low during this period. Actual wind power density has its maximum value of 648.73W/$m^2$, 620.40W/$m^2$, 500.50W/$m^2$, and 627.69W/$m^2$ for the months of July 2015, May 2016, July 2017 and June 2018. Whereas, wind power density has its lowest value of 112.41W/$m^2$ for December 2015, 73.672W/$m^2$ and 82.35W/$m^2$ for November 2016 and 2017, 90.12W/$m^2$ for January 2018, respectively.

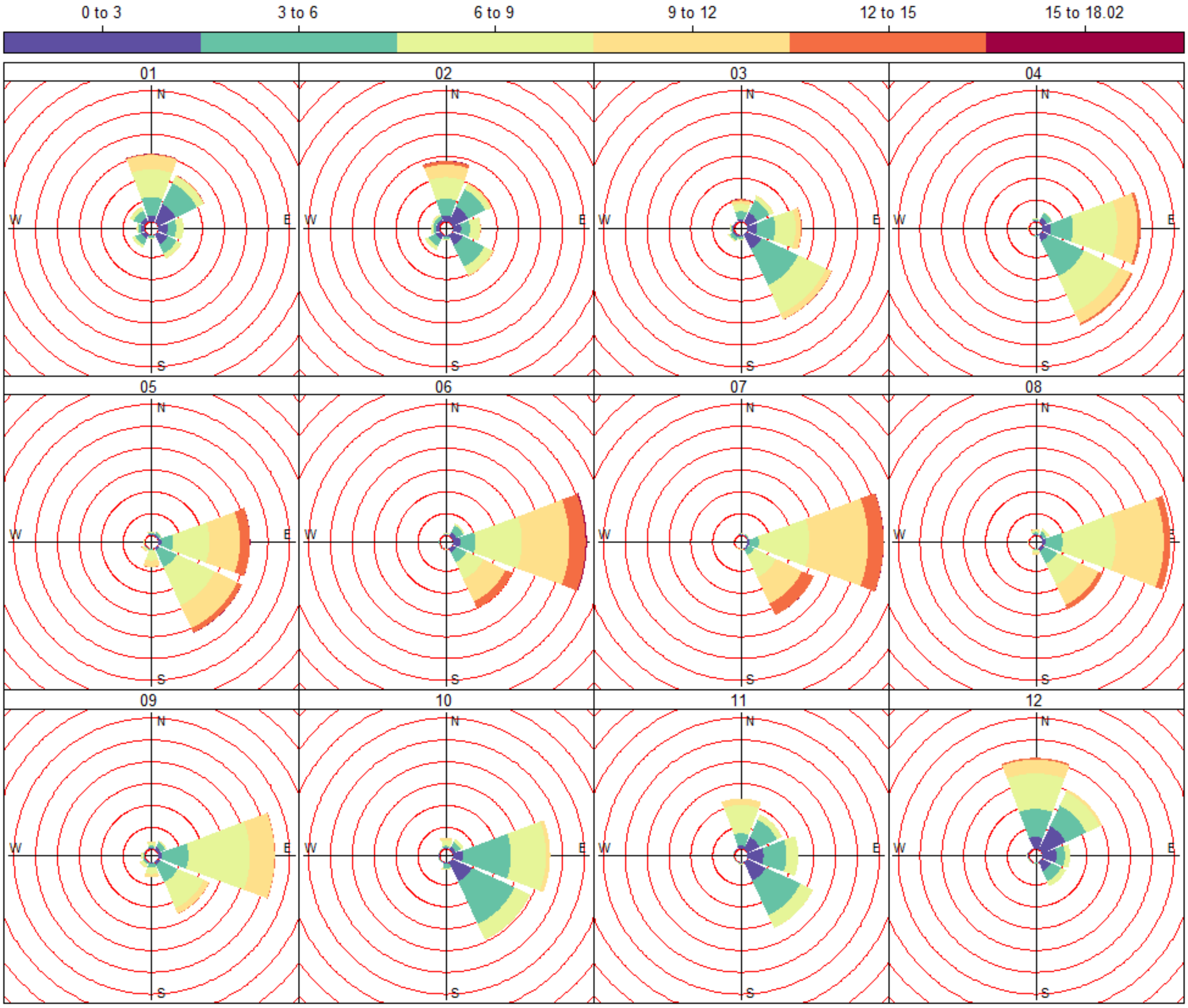



**FIGURE 12.** Monthly wind rose graph of wind data collected at Jhimpir, Pakistan.

The maximum value of wind power density determined by Weibull distribution is 698.65W/m$^2$ in July 2015, 622.80W/m$^2$ in May 2016, 567.33W/m$^2$ in July 2017. For 2018, the maximum value of wind power density is calculated to be 653.50W/m$^2$ in June. The minimum value of 130.3W/m$^2$ is calculated in December 2015, 84.677W/m$^2$ in November 2016, 87.165W/m$^2$ in October 2017, and the minimum of 92.588W/m$^2$ is observed in January 2018. Likewise, maximum and minimum values of wind power density based on Rayleigh distribution are 1021.40W/m$^2$ and 117.15W/m$^2$ in July and December 2015, 905.08W/m$^2$ and 83.681W/m$^2$ in May and November 2016, 768.02W/m$^2$ and 102.37W/m$^2$ in July and October 2017, 924.55W/m$^2$ and 90.417W/m$^2$ in June and January 2018, respectively.

Based on the results given in Table 8, wind power density errors of Weibull and Rayleigh are calculated by using equation 40 for each year. Fig. 10 shows the percentage error measured for the year 2015. Bar graph in the negative direction shows that the estimated wind power density is less than the actual value extracted from the original data. Whereas, bars graphs with positive values shows the overestimation of wind power density estimated by Weibull and Rayleigh distributions. In March 2015, both Weibull and Rayleigh distributions have slightly negative errors which shows the underestimation of wind power density in this month. From Fig. 10, it can be seen that the error of Weibull distribution is less than that of Rayleigh distribution for nine months. For only three months of January, March and April, wind power density based on Rayleigh distribution have smaller errors. The highest error for Weibull is 0.1850 observed in the month of February, whereas for Rayleigh, the highest error is 0.5745 in July. Weibull and Rayleigh have the lowest error of 0.0178 in August and 0.0180 in March, respectively. Overall, the error of wind power density calculated using Weibull is

**TABLE 9.** Percentage of wind direction in different wind speed bins.

| Bin | N | NE | E | SE | S | SW | W | NW |
|---|---|---|---|---|---|---|---|---|
| 0-3 | 2.57% | 1.76% | 3.42% | 5.36% | 0.46% | 0.74% | 0.37% | 0.33% |
| 3-6 | 3.12% | 3.07% | 9.98% | 9.74% | 1.07% | 1.32% | 0.66% | 0.47% |
| 6-9 | 2.82% | 2.32% | 14.63% | 10.54% | 1.22% | 0.92% | 0.49% | 0.35% |
| 9-12 | 1.68% | 0.91% | 8.43% | 6.11% | 0.45% | 0.25% | 0.12% | 0.07% |
| 12-15 | 0.42% | 0.22% | 1.97% | 1.45% | 0.04% | 0.01% | 0.00% | 0.01% |
| >15 | 0.00% | 0.00% | 0.08% | 0.03% | 0.00% | 0.00% | 0.00% | 0.00% |
| Total | 10.61% | 8.29% | 38.52% | 33.24% | 3.23% | 3.25% | 1.65% | 1.22% |

much less than that of Rayleigh, showing that Weibull is best for the investigated site.

In the purpose to select, the best performing empirical distribution comparison of PDF of actual wind speed data with eight empirical distributions is performed. For assessment of goodness-to-fit seven statistical test are performed for complete duration and yearly bases showing Weibull is the best fit distribution for the candidate site. In addition to statistical tests performed, error calculated between wind power density of theoretical and empirical distributions also supports our claim. Which shows that simulation results of Weibull distribution have good match to the actual theoretical distribution of real world data.

### D. WIND DIRECTION

Wind direction plays an important role for deciding the optimal position of wind turbines for wind power production. Table 9 illustrates the percentage of time that wind speed has blown in a particular direction for the complete wind data collected at Jhimpir, Pakistan. Wind speed is divided into six bins with each having a width of 3m/s. Wind direction is investigated in eight directions, i.e., North (N), Northeast (NE), East (E), Southeast (SE), South (S), Southwest (SW), West (W), and Northwest (NW). Overall, the highest percentage of wind direction is observed toward East (38.52%), followed by Southeast (33.24%) and North (10.61%), respectively. For wind direction of East, the bin of 6-9m/s has the highest percentage of 14.63%, followed by bins of 3-6m/s and 9-12m/s with 9.98% and 8.43%. Likewise, bins of 6-9m/s, 3-6m/s and 9-12m/s have 10.54%, 9.72% and 6.11% in the southeastern direction, respectively. It can be seen that for 71.76% of the total time, wind direction is Eastern and Southeastern. This result is good for the investigated site, because this uniformity in wind directions helps in the improvement of losses in wind power production due to small changes in wind direction.

Yearly analysis of wind direction is also part of this study and wind rose graph for each year is shown in Fig. 11. Frequency of occurrence of particular wind directions for different wind speed bins from 2015 to 2018 is shown in the figure. It can be seen that the frequency of wind direction toward East and Southeast are significantly greater than other directions for each year. In 2015, 2016 and 2018, the most occurring wind direction is towards East, having frequencies of about 35%, 45% and 37%, respectively. The second most occurring direction for these years is towards Southeast. In 2017, wind direction of Southeast has the frequency of about 40%, followed by wind direction of East with the frequency slightly higher than 30%. From Fig. 11, it is obvious that the yearly pattern of wind direction remains almost the same with two most occurring directions of East and Southeast. This observation is very similar to the pattern observed from the whole data shown in Table 9.

Monthly analysis of wind direction for the whole data is also performed and shown in Fig. 12. From March to November, Eastern and Southeastern wind directions have the highest frequency of occurrence. For January, February and December, the frequency of Northern and Northeastern wind directions is higher than other directions.

## VII. CONCLUSION

To meet the day to day demand of electricity and reduce the gap between demand and supply, Pakistan has increased its power generation through both conventional and non-conventional resources of energy. Adding more power through convention resources will put an extra burden on the trembling economy of the country. Pakistan has a good potential of wind energy generation in the Southern part of the country. Using these wind energy resources will not only solve the energy crisis of the country but will also support the economic development in Pakistan. To support wind power development in Pakistan, the assessment of wind power potential of Jhimpir, Sindh, Pakistan is carried out in this study. Detail analysis is also done to find out the best fit parametric distribution using the monthly, yearly or whole data. Results extracted from the original data can be concluded as follows:

1. Monthly mean wind speed is more than 4m/s for all months from January 2015 to July 2018. The maximum monthly mean wind speed is 9.46m/s in July 2015, however the minimum speed is 4.07m/s for November and December 2016. The yearly average wind speed is recorded to be more than 6m/s each year, showing the potential of wind power development in Jhimpir. The highest yearly average wind speed is 6.62m/s in 2015 and the lowest speed is 6.4m/s in 2016.
2. Maximum value of the most probably wind speed (i.e., 13.8m/s) is observed in May 2016 and a minimum of 3.83m/s is observed in December 2016. Likewise, wind speed carrying maximum energy shows a maximum of 17.1m/s in July 2015 and a minimum of 8.93m/s is observed in November of 2016 and 2017.

3. Based on $R^2$, K-S, $\chi^2$, CvM, A-D, AIC and BIC tests, Weibull distribution is found to be the best one among eight parametric distributions, followed by Rayleigh distribution.
4. Wind power density is high during the month of May, June, July and August. The highest value of wind power density, i.e., 648.73W/m$^2$, is observed in January 2015 and the minimum value is observed in October, November, December and January with the least value of 73.672W/m$^2$ in November 2016.
5. Maximum wind power density for Weibull and Rayleigh distributions are 698.65W/m$^2$ and 1021.4W/m$^2$ in July 2015. Minimum wind power density for Weibull and Rayleigh distributions are 84.677W/m$^2$ and 83.681W/m$^2$ in November 2016, respectively.
6. In terms of wind power density errors, smaller errors for Weibull distribution shows that Weibull is the most suitable function for the investigated site.
7. Analysis of wind direction data shows that during the month of January, February and December, the predominate wind direction is Northern. While for the remaining part of the year, Eastern and Southeastern wind directions have the highest occurrence. Likewise, for the complete data and yearly analysis, Eastern and Southeastern wind directions are predominant.

Owing to the difference in the wind characteristics for each studied site, it is important to select best fit parametric distribution and assess wind power potential for that particular site, Hence, based on our analysis, Jhimpir is very suitable for the development of wind power projects. Weibull distribution function is the most suitable for the investigated site. Technical and economical evaluation by using commercially available wind turbines, Comparing different methods to estimate parameters of parametric distributions and developing non-parametric methods would be further studied in our future work.

## REFERENCES

[1] B. Dudley, "BP statistical review of world energy," Tech. Rep., 2018.

[2] T. R. Ayodele, A. A. Jimoh, J. L. Munda, and J. T. Agee, "Wind distribution and capacity factor estimation for wind turbines in the coastal region of South Africa," *Energy Convers. Manage.*, vol. 64, pp. 614–625, Dec. 2012.

[3] O. Alavi, A. Sedaghat, and A. Mostafaeipour, "Sensitivity analysis of different wind speed distribution models with actual and truncated wind data: A case study for Kerman, Iran," *Energy Convers. Manage.*, vol. 120, pp. 51–61, Jul. 2016.

[4] J. Wang, S. Qin, S. Jin, and J. Wu, "Estimation methods review and analysis of offshore extreme wind speeds and wind energy resources," *Renew. Sustain. Energy Rev.*, vol. 42, pp. 26–42, Feb. 2015.

[5] J. Wang, J. Hu, and K. Ma, "Wind speed probability distribution estimation and wind energy assessment," *Renew. Sustain. Energy Rev.*, vol. 60, pp. 881–899, Jul. 2016.

[6] N. Masseran, "Evaluating wind power density models and their statistical properties," *Energy*, vol. 84, pp. 533–541, May 2015.

[7] J. Wang, X. Huang, Q. Li, and X. Ma, "Comparison of seven methods for determining the optimal statistical distribution parameters: A case study of wind energy assessment in the large-scale wind farms of China," *Energy*, vol. 164, pp. 432–448, Dec. 2018.

[8] O. Alavi, K. Mohammadi, and A. Mostafaeipour, "Evaluating the suitability of wind speed probability distribution models: A case of study of east and southeast parts of Iran," *Energy Convers. Manage.*, vol. 119, pp. 101–108, Jul. 2016.

[9] Q. Hu, Y. Wang, Z. Xie, P. Zhu, and D. Yu, "On estimating uncertainty of wind energy with mixture of distributions," *Energy*, vol. 112, pp. 935–962, Oct. 2016.

[10] T. B. M. J. Ouarda, C. Charron, and F. Chebana, "Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and L-moment ratio diagram methods, with a case study," *Energy Convers. Manage.*, vol. 124, pp. 247–265, Sep. 2016.

[11] N. Aries, S. M. Boudia, and H. Ounis, "Deep assessment of wind speed distribution models: A case study of four sites in Algeria," *Energy Convers. Manage.*, vol. 155, pp. 78–90, Jan. 2018.

[12] D. Mazzeo, G. Oliveti, and E. Labonia, "Estimation of wind speed probability density function using a mixture of two truncated normal distributions," *Renew. Energy*, vol. 115, pp. 1260–1280, Jan. 2018.

[13] 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 Convers. Manage.*, vol. 143, pp. 109–122, Jul. 2017.

[14] A. Keyhani, M. Ghasemi-Varnamkhasti, M. Khanali, and R. Abbaszadeh, "An assessment of wind energy potential as a power generation source in the capital of Iran, tehran," *Energy*, vol. 35, no. 1, pp. 188–201, Jan. 2010.

[15] A. Dabbaghiyan, F. Fazelpour, M. D. Abnavi, and M. A. Rosen, "Evaluation of wind energy potential in province of Bushehr, Iran," *Renew. Sustain. Energy Rev.*, vol. 55, pp. 455–466, Mar. 2016.

[16] S. M. Bina, S. Jalilinasrabady, H. Fujii, and H. Farabi-Asl, "A comprehensive approach for wind power plant potential assessment, application to northwestern Iran," *Energy*, vol. 164, pp. 344–358, Dec. 2018.

[17] M. A. Saeed, Z. Ahmed, S. Hussain, and W. Zhang, "Wind resource assessment and economic analysis for wind energy development in Pakistan," *Sustain. Energy Technol. Assessments*, vol. 44, Apr. 2021, Art. no. 101068.

[18] B. Wacker, J. V. Seebaß, and J. C. Schlüter, "A modular framework for estimating annual averaged power output generation of wind turbines," *Energy Convers. Manage.*, vol. 221, Oct. 2020, Art. no. 113149.

[19] M. A. Baseer, J. P. Meyer, M. M. Alam, and S. Rehman, "Wind speed and power characteristics for Jubail industrial city, Saudi Arabia," *Renew. Sustain. Energy Rev.*, vol. 52, pp. 1193–1204, Dec. 2015.

[20] M. A. M. Ramli, A. Hiendro, and Y. A. Al-Turki, "Techno-economic energy analysis of wind/solar hybrid system: Case study for western coastal area of Saudi Arabia," *Renew. Energy*, vol. 91, pp. 374–385, Jun. 2016.

[21] A. Ashfaq and A. Ianakiev, "Features of fully integrated renewable energy atlas for Pakistan; wind, solar and cooling," *Renew. Sustain. Energy Rev.*, vol. 97, pp. 14–27, Dec. 2018.

[22] S. F. Khahro, K. Tabbassum, A. M. Soomro, L. Dong, and X. Liao, "Evaluation of wind power production prospective and weibull parameter estimation methods for Babaurband, Sindh Pakistan," *Energy Convers. Manage.*, vol. 78, pp. 956–967, Feb. 2014.

[23] M. M. Aman, G. B. Jasmon, A. Ghufran, A. H. A. Bakar, and H. Mokhlis, "Investigating possible wind energy potential to meet the power shortage in Karachi," *Renew. Sustain. Energy Rev.*, vol. 18, pp. 528–542, Feb. 2013.

[24] I. A. Mirza, N. A. Khan, and N. Memon, "Development of benchmark wind speed for Gharo and Jhimpir, Pakistan," *Renew. Energy*, vol. 35, no. 3, pp. 576–582, Mar. 2010.

[25] Z. H. Hulio, W. Jiang, and S. Rehman, "Techno–economic assessment of wind power potential of Hawke's bay using Weibull parameter: A review," *Energy Strategy Rev.*, vol. 26, Nov. 2019, Art. no. 100375.

[26] Q. Chaudhry, "Wind power potential at Jamshoro-Sindh," Pakistan Meteorological Dept., Tech. Rep., 2006.

[27] WWEA. *Worldwide Wind Statistics by Country*. [Online]. Available: https://library.wwindea.org

[28] P. Drobinski, C. Coulais, and B. Jourdier, "Surface wind-speed statistics modelling: Alternatives to the Weibull distribution and performance evaluation," *Boundary-Layer Meteorol.*, vol. 157, no. 1, pp. 97–123, Oct. 2015.

[29] 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)," *Renew. Sustain. Energy Rev.*, vol. 42, pp. 313–322, Feb. 2015.

[30] S. Ali, S.-M. Lee, and C.-M. Jang, "Statistical analysis of wind characteristics using weibull and Rayleigh distributions in Deokjeok-do Island–Incheon, South Korea," *Renew. Energy*, vol. 123, pp. 652–663, Aug. 2018.

[31] S. F. Khahro, K. Tabbassum, A. M. Soomro, X. Liao, M. B. Alvi, L. Dong, and M. F. Manzoor, "Techno-economical evaluation of wind energy potential and analysis of power generation from wind at Gharo, Sindh Pakistan," *Renew. Sustain. Energy Rev.*, vol. 35, pp. 460–474, Jul. 2014.

[32] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, Jan. 1965.

[33] C. Jung, "High spatial resolution simulation of annual wind energy yield using near-surface wind speed time series," *Energies*, vol. 9, no. 5, p. 344, May 2016.

[34] 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.*, vol. 51, no. 7, pp. 1449–1458, Jul. 2010.

[35] T. H. Soukissian and F. E. Karathanasi, "On the selection of bivariate parametric models for wind data," *Appl. Energy*, vol. 188, pp. 280–304, Feb. 2017.

[36] C. Jung, D. Schindler, J. Laible, and A. Buchholz, "Introducing a system of wind speed distributions for modeling properties of wind speed regimes around the world," *Energy Convers. Manage.*, vol. 144, pp. 181–192, Jul. 2017.

[37] N. Masseran, A. M. Razali, and K. Ibrahim, "An analysis of wind power density derived from several wind speed density functions: The regional assessment on wind power in Malaysia," *Renew. Sustain. Energy Rev.*, vol. 16, no. 8, pp. 6476–6487, Oct. 2012.

[38] J.-Y. Shin, T. B. M. J. Ouarda, and T. Lee, "Heterogeneous mixture distributions for modeling wind speed, application to the UAE," *Renew. Energy*, vol. 91, pp. 40–52, Jun. 2016.

[39] I. Usta and Y. M. Kantar, "Analysis of some flexible families of distributions for estimation of wind speed distributions," *Appl. Energy*, vol. 89, no. 1, pp. 355–367, Jan. 2012.

[40] T. B. M. J. Ouarda and C. Charron, "On the mixture of wind speed distribution in a nordic region," *Energy Convers. Manage.*, vol. 174, pp. 33–44, Oct. 2018.

**MUHAMMAD ARMOGHAN KHAN** received the B.E. degree in electrical engineering from the National University of Sciences and Technology, Islamabad, Pakistan, in 2012, and the M.S. degree in power system and its automation from North China Electrical Power University, Beijing, China, in 2014. He is currently pursuing the Ph.D. degree in electrical engineering with Xi'an Jiaotong University, Xi'an, China. From 2014 to 2016, he was associated with the Department of Electrical Engineering, National University of Sciences and Technology, as a Lecturer. His research interests include power system planning, integration of renewable energy resources, and forecasting.

**YAO ZHANG** (Member, IEEE) received the Ph.D. degree from Xi'an Jiaotong University, in 2016. From 2017 to 2019, he was a Research Associate with The University of Tennessee, Knoxville (UTK). He is currently an Associate Professor with the School of Electrical Engineering, Xi'an Jiaotong University. His research interests include power system operation and planning and renewable energy forecasting and integration. He received the IEEE Award for his outstanding performance at Global Energy Forecasting Competition, in 2014.

**JIANXUE WANG** (Senior Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1999, 2002, and 2006, respectively. He is currently a Professor with the School of Electrical Engineering, Xi'an Jiaotong University. His current research interests include power system planning and operation, microgrid planning and operation, and electricity market.

**JINGDONG WEI** received the B.S. degree in electrical engineering from the Shenyang University of Technology, Shenyang, China, in 2014. He is currently pursuing the Ph.D. degree in electrical engineering with Xi'an Jiaotong University, Xi'an, China. He is also a Visiting Ph.D. Student with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, USA. His research interests include multi-energy system operation and planning, multi-energy microgrid planning, and integrated energy system dispatch.

**MUHAMMAD ALI RAZA** received the B.S. degree in electrical engineering from COMSATS University Islamabad, Lahore, Pakistan, in 2009, and the M.S. degree in electrical engineering from the National University of Computer and Emerging Sciences, Lahore, in 2012. He is currently pursuing the Ph.D. degree from Xi'an Jiaotong University. He is also associated with the Electrical Engineering Department, COMSATS University, as a Lecturer. His research interests include image processing, computer vision, machine learning, and deep learning.

**AITIZAZ AHMAD** received the B.E. degree in electrical engineering from North China Electrical Power University, Beijing, China, in 2012. He is currently working as the Manager operations with Three Gorges Wind Farms, Sindh, Pakistan. His research interests include operation and planning of power systems and integration of renewable energy.

**YIPING YUAN** received the B.E. degree in electrical engineering from Zhengzhou University, Zhengzhou, China, in 2014, and the M.S. degree in electrical engineering from Chongqing University, Chongqing, China, in 2017. He is currently pursuing the Ph.D. degree in electrical engineering with Xi'an Jiaotong University, Xi'an, China. His research interests include renewable power system dispatching and planning and probabilistic estimation.