6. Mixture Distributions for Wind Speed Modelling

Weibull distribution is a widely used distribution to fit wind speed data around the globe. However, there are certain wind data available that are heterogeneous by nature. These heterogeneous data shows a sign of bi-modality (extra hump) or bi-tangentiality (extra bump) in it. These types of wind data are difficult to model with 2-parameter *W.pdf*. As 2-parameter *W.pdf* is a unimodal distribution; therefore, researchers **[7, 10, 13, 16, 69-73, 75-84, 145, 154]** look beyond the 2-parameter continuous distribution, and in that context, the 5-parameter mixture distribution enables the researchers to solve the problem of fitting heterogeneous data. The mixture distribution is a convex linear combination of 2-parameter continuous distribution mixed at different weight proportion **[85, 86]**. According to the literature review carried out at chapter 2, it clearly reveals that no single distribution shows universal acceptance for wind speed assessment for all locations, especially, for Indian climatology, where the latitudinal and longitudinal variations are intensely vast. Moreover, the literature review carried out at chapter 2 (see Table 2.2) have analyzed the mixture distributions considering short-term wind speed data. The behaviour of the mixture distributions for large sample size (long-term data) which is a prime requirement to fit any distribution has rarely been studied. As for long-term data, it is difficult to distinguish between the modalities of the distributions. Hence, keeping in view the shortcomings of other researchers, the present study attempts to check the behaviour of mixture distributions for a large sample of wind speed data. The sites selected for the analyses of mixture modelling are the three onshore locations of India namely, Trivandrum, Ahmedabad and Calcutta (see Table 4.1 for detail description)

Several researchers **[16, 70, 71, 75, 213]** claim that only two-component mixture distribution is suitable to model wind speed data. In this study, the claim of the researchers has been reinvestigated by taking 100,000 random variables between 0 and 1 and estimate wind speed by reverse computing of *W.pdf* taking Weibull shape parameter (k) = 2 and scale parameter (s) = 3. Then the authors have tried to fit the N-component of *W.pdf*, where N = 1, 2 and 3. Table 6.1 shows the estimated parameters of the *W.pdf*. It has been found that the weight percentage of the third component is zero, which implies that only 2-component mixture distribution is sufficient to analyze the given wind speed data.

Table 6.1: Estimated parameters with mixture Weibull distribution.

Parameter	w_1	k_1	s_1	w_2	k_2	s_2	w_3	k_3	s_3
N = 1	1	2.009	3.002	-	-	-	-	-	-
N = 2	0.993	1.991	2.997	0.007	5.709	3.354	-	-	-
N = 3	0.992	1.997	3.003	0.007	2.278	1.865	0.000	8.193	2.012

In this chapter, several types of mixture distribution such as Weibull-Weibull distribution (*WW.pdf*), Gamma-Weibull distribution (*GW.pdf*), Truncated Normal-Weibull distribution (*TNW.pdf*), Truncated Normal-Normal distribution (*TNN.pdf*), Truncated Normal-Gamma distribution (*TNG.pdf*) and proposed Gamma-Gamma distribution (*GG.pdf*) along with *MEP*-distribution (constraint moment = 4) have been compared with conventional 2-parameter Weibull distribution for analysis of both wind speed and wind power density. The data considered are the long-term data. The five parameters of mixture distribution can be estimated using the graphical method (Lavenberg-Marquardt algorithm), Method of Moments and Maximum Likelihood Method (Expectation-maximization algorithm) **[70]**. However, in this chapter, Maximum Likelihood Method using non-linear programming technique has been used to solve the log-likelihood function of the mixture distributions.

The basic assumption before analyzing wind speed data is that the data considered in this study are independent and identically distributed (*i.i.d.*) random variables. This chapter deals with the wind speed after exclusion of calm hours, as calm wind speed does not help in energy production. This non-linear technique requires a good starting point or base point. The starting point for each mixture distribution is separately discussed in subsequent section.

6.1 Mathematical Models

6.1.1 Mixture Weibull distribution (WW.pdf)

The expression for probability density function of the mixture Weibull distribution are given as **[20, 69]**:

$$ff(v;w,k_1,s_1,k_2,s_2) = wf(v;k_1,s_1) + (1-w)f(v;k_2,s_2)$$

$$= w\left\{\frac{k_1}{s_1}\left(\frac{v}{s_1}\right)^{k_1-1}\exp\left[-\left(\frac{v}{s_1}\right)^{k_1}\right]\right\} + (1-w)\left\{\frac{k_2}{s_2}\left(\frac{v}{s_2}\right)^{k_2-1}\exp\left[-\left(\frac{v}{s_2}\right)^{k_2}\right]\right\} \tag{6.1}$$

for $v > 0$, $k_1, s_1, k_2, s_2 > 0$, $0 \le w \le 1$ *and for* $v \le 0$, $ff(v;w,k_1,s_1,k_2,s_2) = 0$.

where k and s are the shape and scale parameters of the Weibull distribution respectively, w is the weight parameter or mixing proportion of the mixture distribution. The expression for cumulative distribution function of the mixture Weibull distribution are given as:

$$FF(v;w,k_1,s_1,k_2,s_2) = wF(v;k_1,s_1) + (1-w)F(v;k_2,s_2)$$

$$= w\left\{1-\exp\left[-\left(\frac{v}{s_1}\right)^{k_1}\right]\right\} + (1-w)\left\{1-\exp\left[-\left(\frac{v}{s_2}\right)^{k2}\right]\right\} \tag{6.2}$$

Taking logarithmic of the likelihood function of Eq. (6.1) yields the following expression:

$$LL = \sum_{i=1}^{n} \ln\left\{wf(v_i;k_1,s_1) + (1-w)f(v_i;k_2,s_2)\right\} \tag{6.3}$$

where v_i is the wind speed at time step i and n is the number of data.

The starting points for WW distribution

Carta, and Ramirez [70] proposed the starting points for *WW.pdf* as given below:

$$k_1 = k_2 = \left(\frac{\sqrt{\sigma^2}}{m}\right)^{-1.086}; s_1 = s_2 = m\left[\Gamma\left(1+\frac{1}{k_1}\right)\right]^{-1} \left[ms^{-1}\right]; w = 0.25,\ 0.5,\ 0.75 \quad (6.4)$$

They suggested that if the system of Eq. (6.1), with the constraints as mentioned below the equation, does not find a solution for w = 0.25, then the algorithm should be programmed in such a way that it successively changes its w to 0.5 and 0.75 and again tried to find the solution.

6.1.2 Mixture Gamma-Weibull distribution (GW.pdf)

The expression for probability density function of Gamma distribution is given as **[7]**:

$$\begin{cases} g(v;\zeta,\beta) = \dfrac{v^{\zeta-1}}{\beta^{\zeta}\Gamma(\zeta)}\exp\left[-\dfrac{v}{\beta}\right], v>0; \zeta,\beta>0 \\ 0, v \le 0 \end{cases} \quad (6.5)$$

where ζ and β are the shape and scale parameters respectively. The unit of scale parameter is same as that of wind speed and $\Gamma(\bullet)$ is the Gamma function. The relevant cumulative Gamma distribution function is given as:

$$G(v;\zeta,\beta) = \int \frac{v^{\zeta-1}}{\beta^{\zeta}\Gamma(\zeta)}\exp\left[-\frac{v}{\beta}\right]dv \quad (6.6)$$

The Gamma distribution function mixed with conventional 2-parameter Weibull distribution function that has been used for wind speed assessment is expressed as:

$$h(v;w,\zeta,\beta,k,s) = wg(v;\zeta,\beta) + (1-w)f(v;k,s) \quad (6.7)$$

The corresponding cumulative distribution function is written as:

$$H(v;w,\zeta,\beta,k,s) = wG(v;\zeta,\beta) + (1-w)F(v;k,s) \quad (6.8)$$

Taking logarithmic of the likelihood function of Eq. (6.7) yields the following expression:

$$LL = \sum_{i=1}^{n} \ln\left\{ wg(v_i;\zeta;\beta) + (1-w) f(v_i;k,s) \right\} \tag{6.9}$$

where v_i is the wind speed at time step *i* and *n* is the number of data.

The starting points for GW.pdf:

The starting points for the *GW.pdf* is given as:

$$\varsigma = \frac{\mu^2}{\sigma^2}; \qquad \beta = \frac{\sigma^2}{\mu}$$
$$k = \left(\frac{\sqrt{s^2}}{m}\right)^{-1.086}; s = m\left[\Gamma\left(1+\frac{1}{\alpha_0}\right)\right]^{-1} \left[ms^{-1}\right];\ w = 0.25, 0.5, 0.75 \tag{6.10}$$

6.1.3 Mixture Truncated Normal-Weibull distribution (TNW.pdf)

The singly Truncated Normal distribution mixed with conventional 2-parameter Weibull distribution has been used to evaluate wind speed data. The advantage of using *TNW.pdf* is that it takes into account the null wind speed. The probability density function and corresponding cumulative distribution function of *TNW.pdf* are expressed as **[16]**:

$$\psi(v; w,\mu,\sigma,k,s) = wq(v;\mu,\sigma) + (1-w) f(v;k,s) \tag{6.11}$$

$$\Psi(v; w,\mu,\sigma,k,s) = wQ(v;\mu,\sigma) + (1-w) F(v;k,s) \tag{6.12}$$

Eq. (6.13) represents the relevant likelihood function:

$$LL = \sum_{i=1}^{n} \ln\left\{ wq(v_i;\mu,\sigma) + (1-w) f(v_i;k,s) \right\} \tag{6.13}$$

The starting points for TNW.pdf

Carta, and Ramirez [75] proposed the starting points for *TNW.pdf* as given below:

$$\mu = \frac{2m_2'm - m_3'}{2(m)^2 - m_2'} \left[ms^{-1}\right]; \quad \sigma^2 = \frac{m_3'm - \left(m_2'\right)^2}{2(m)^2 - m_2'} \left[ms^{-1}\right]$$
$$k = \left(\frac{\sqrt{s^2}}{m}\right)^{-1.086}; \quad s = m\left[\Gamma\left(1 + \frac{1}{\alpha_0}\right)\right]^{-1} \left[ms^{-1}\right] \quad ; w = 0.25, 0.5, 0.75 \quad (6.14)$$

$m; m_2'; m_3'$ are the mean, mean of the square, mean of the cube of the wind speed respectively.

6.1.4 Two components Singly Truncated Normal Distribution (TNN.pdf)

The Truncated Normal distribution is the probability distribution of normally distributed random variables, whose values are bounded either from upper or lower (or both) tail. Since the recorded wind speeds are not below zero, the Normal distribution truncated from the lower tail, to analyze the wind speed data. The probability density function of singly Truncated Normal distribution is given as:

$$\begin{cases} q(v;\mu,\sigma) = \dfrac{1}{I(\mu,\sigma)\sigma\sqrt{2\pi}} \exp\left[-\dfrac{(v-\mu)^2}{2\sigma^2}\right], v \geq 0; \sigma > 0 \\ 0, v < 0 \end{cases} \quad (6.15)$$

The corresponding cumulative Truncated Normal distribution is expressed as:

$$Q(v;\mu,\sigma) = \int_0^v \frac{1}{I(\mu,\sigma)\sigma\sqrt{2\pi}} \exp\left[-\frac{(v-\mu)^2}{2\sigma^2}\right] dv \quad (6.16)$$

where μ, and σ are the mean, and standard deviation of the function respectively. The unit of μ and σ are same as that of wind speed (v). $I(\mu,\sigma)$ is the normalization factor. It leads to the integration of the Truncated Normal distribution to one. The simplest way to calculate the normalization factor is to evaluate the difference between cumulative Truncated Normal distribution at infinity and at zero.

The expression for normalization factor are given as:

$$I(\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\int_{0}^{\infty}\exp\left[-\frac{(v-\mu)^2}{2\sigma^2}\right]dv$$

$$I(\mu,\sigma)=Q(\inf;\mu,\sigma)-Q(0;\mu,\sigma) \tag{6.17}$$

The probability distribution function and corresponding cumulative distribution function of two-component singly Truncated Normal distribution are expressed as:

$$r(v;w,\mu_1,\sigma_1,\mu_2,\sigma_2)=wq(v;\mu_1,\sigma_1)+(1-w)q(v;\mu_2,\sigma_2) \tag{6.18}$$

$$R(v;w,\mu_1,\sigma_1,\mu_2,\sigma_2)=wQ(v;\mu_1,\sigma_1)+(1-w)Q(v;\mu_2,\sigma_2) \tag{6.19}$$

Taking logarithmic of the likelihood function of Eq. (6.18) yields the following expression:

$$LL=\sum_{i=1}^{n}\ln\left\{wq(v_i;\mu_1,\sigma_1)+(1-w)q(v_i;\mu_2,\sigma_2)\right\} \tag{6.20}$$

6.1.5 Gamma-Gamma Distribution (GG.pdf)

The idea behind using Gamma distribution is that the curve of Gamma distribution drops off much gradually than the *W.pdf* for $k > 1$ and more quickly for $k < 1$. Therefore, the Gamma distribution also has the fair chance to fit wind speed data. The probability density function and corresponding cumulative distribution function of two-component mixture Gamma distribution are given as:

$$g(v;w,\zeta_1,\beta_1,\zeta_2,\beta_2)=wg_1(v;\zeta_1,\beta_1)+(1-w)g_2(v;\zeta_2,\beta_2) \tag{6.21}$$

$$G(v;w,\zeta_1,\beta_1,\zeta_2,\beta_2)=wG_1(v;\zeta_1,\beta_1)+(1-w)G_2(v;\zeta_2,\beta_2) \tag{6.22}$$

Taking logarithmic of the likelihood function of Eq. (6.21) yields the following expression:

$$LL=\sum_{i=1}^{n}\ln\left\{wg_1(v_i;\zeta_1,\beta_1)+(1-w)g_2(v_i;\zeta_2,\beta_2)\right\} \tag{6.23}$$

6.1.6 Truncated Normal Gamma Distribution (TNG)

The probability distribution function and corresponding cumulative distribution function of singly Truncated Normal distribution mixed with Gamma distribution are given as:

$$\theta(v; w, \mu, \sigma, \zeta, \beta) = wq(v; \mu, \sigma) + (1-w)g(v; \zeta, \beta) \tag{6.24}$$

$$\Theta(v; w, \mu, \sigma, \zeta, \beta) = wQ(v; \mu, \sigma) + (1-w)G(v; \zeta, \beta) \tag{6.25}$$

Taking logarithmic of the likelihood function of Eq. (6.24) yields the following expression:

$$LL = \sum_{i=1}^{n} \ln\left\{ wq(v_i; \mu, \sigma) + (1-w)g(v_i; \zeta, \beta) \right\} \tag{6.26}$$

6.1.7 Maximum Entropy Principle (MEP.)

According to *MEP.*, as introduced by **[85, 86]** and elaborately discussed by **[88]** that when one has only partial information about a random variate, scalar or vector, one should choose that probability distribution for it, which is consistent with the given information but has otherwise maximum uncertainty associated with it. Based on the concept of entropy as defined by **[87]**, the measure of maximum uncertainty associated with probability density function is defined by the following expression **[15]**:

$$S = -\int_{V_{\min}}^{V_{\max}} f(v) \ln f(v) dv \tag{6.27}$$

where *v* is the wind speed; *f(v)* is the probability density function of wind speed; *S* is the self-information entropy; -ln *f(v)* is the uncertainty of wind speed.

Eqs. (6.28). and (6.29) show the set of constraints used to maximize the information entropy:

$$\int_{V_{\min}}^{V_{\max}} f(v) d(v) = 1 \tag{6.28}$$

$$\int_{V_{\min}}^{V_{\max}} v^n f(v)dv = \overline{v^n} \tag{6.29}$$

The maximization of information entropy is carried out using the method of Lagrange multipliers. Eq. (6.30) shows the construct function $L(v)$ as follow:

$$L(v) = \int_{V_{\min}}^{V_{\max}} f(v)\ln f(v)dv + \lambda_0\left(\int_{V_{\min}}^{V_{\max}} f(v)dv - 1\right) + \sum_{i=1}^{m}\lambda_i\left(\int_{V_{\min}}^{V_{\max}} v^i f(v)dv - \overline{v^i}\right) \tag{6.30}$$

with the partial derivative of $L(v)$ equal to zero, the probability density function $f(v)$ yields:

$$f(v) = \exp\left(-\lambda_0 - \sum_{i=1}^{m}\lambda_i v^i\right) \tag{6.31}$$

where $\lambda_0, \lambda_1, ..., \lambda_m$ are the Lagrangian multipliers; m is the constraint moment. In this study, *MEP.pdf* with constraint moment = 4 has been used for comparison against other mixture distribution. Thus, the probability density function *f(v)* can be evaluated once the value of Lagrangian multiplier (λ_i) can be obtained. For this, Eqs. (6.32)-(6.36) need to be solved simultaneously **[24]**:

$$\int_{V_{\min}}^{V_{\max}} \exp(-\lambda_0 - \lambda_1 V - \lambda_2 V^2 - \lambda_3 V^3 - \lambda_4 V^4)dV = 1 \tag{6.32}$$

$$\int_{V_{\min}}^{V_{\max}} V\exp(-\lambda_0 - \lambda_1 V - \lambda_2 V^2 - \lambda_3 V^3 - \lambda_4 V^4)dV = \overline{V} \tag{6.33}$$

$$\int_{V_{\min}}^{V_{\max}} V^2\exp(-\lambda_0 - \lambda_1 V - \lambda_2 V^2 - \lambda_3 V^3 - \lambda_4 V^4)dV = \overline{V^2} \tag{6.31}$$

$$\int_{V_{\min}}^{V_{\max}} V^3\exp(-\lambda_0 - \lambda_1 V - \lambda_2 V^2 - \lambda_3 V^3 - \lambda_4 V^4)dV = \overline{V^3} \tag{6.32}$$

$$\int_{V_{\min}}^{V_{\max}} V^4\exp(-\lambda_0 - \lambda_1 V - \lambda_2 V^2 - \lambda_3 V^3 - \lambda_4 V^4)dV = \overline{V^4} \tag{6.36}$$

Modified second order Newton-Raphson method has been employed to solve the above set of equations **[214]**. Zero is considered as a starting point or base point.

Eq. (6.32) has been employed as a tool to estimate the cumulative distribution function for *MEP*. analysis.

6.2 Accuracy Judgment Criteria

To check the most suitable distribution to define the wind speed data for all three locations. Three goodness of fit statistics have been employed. These are coefficient of determination, root mean square error, and Akaike Information Criteria (*AIC*). Eqs. 5.22 and 5.23 have been used to estimate the coefficient of determination, root mean square error respectively. The mathematical expression and discussion about the *AIC* are given as

6.2.1 Akaike Information Criteria (AIC)

The *AIC* is a measure of the relative quality of a statistical distribution for a given set of data. It is used to provides a mean for selection of distribution. The mathematical expression for *AIC* is given as

$$AIC = 2p - 2\ln(L) \tag{6.37}$$

where p is the number of parameters in the statistical model and L is the maximized value of the likelihood function of the estimated model.

6.3 Results and Discussion

Several mixture probability distributions and *MEP*. distribution (constraint moment = 4) have been compared with conventional 2-parameter *W.pdf* for both wind speed and wind power density assessment. Figures 6.1-6.4 show the plot of wind speed probability density function *f(v)*, and wind power density versus hourly mean wind speed (*m/s*) for three stations experiencing different wind climate of India. The cumulative distribution functions of the corresponding probability density function and wind power density have been plotted and referred on the right ordinate (see Figures 6.1- 6.4). Table 6.2 lists the estimated parameters computed for different probability density functions for three stations.

Table 6.2: Estimated parameters for various probability density functions.

Stations	***W.pdf***		***WW.pdf***				
	k	***s (m/s)***	***w***	***k_1***	***s_1 (m/s)***	***k_2***	***s_2 (m/s)***
Calcutta	1.244	2.414	0.884	1.393	2.731	3.495	0.490
Ahmedabad	1.835	3.227	0.974	1.869	3.185	1.702	4.794
Trivandrum	1.359	2.754	0.125	2.115	0.810	1.549	3.079

GW.pdf					
	w	***ζ***	***β (m/s)***	***k***	***s (m/s)***
Calcutta	0.141	6.101	0.077	1.427	2.800
Ahmedabad	0.476	2.162	1.179	2.282	3.557
Trivandrum	0.081	7.494	0.062	1.532	2.994
TNW.pdf					
	w	***μ (m/s)***	***Σ (m/s)***	***k***	***s (m/s)***
Calcutta	0.287	3.136	2.857	1.416	1.757
Ahmedabad	0.072	3.144	0.944	1.774	3.198
Trivandrum	0.038	5.428	0.570	1.386	2.626
TNG.pdf					
	w	***μ (m/s)***	***σ (m/s)***	***ζ***	***β (m/s)***
Calcutta	0.175	4.030	2.557	1.679	1.070
Ahmedabad	0.282	3.395	1.283	2.376	1.115
Trivandrum	0.051	5.428	0.612	1.790	1.313
GG.pdf					
	w	***$ζ_1$***	***$β_1$ (m/s)***	***$ζ_2$***	***$β_2$ (m/s)***
Calcutta	0.153	6.625	0.067	2.038	1.259
Ahmedabad	0.083	4.344	0.154	3.872	0.791
Trivandrum	0.103	7.476	0.0603	2.309	1.188

Table 6.2: Continued

TNN.pdf					
	w	***μ_1***	***σ_1 (m/s)***	***μ_2***	***σ_2 (m/s)***
Calcutta	0.141	0.556	0.000	0.053	3.296
Ahmedabad	0.697	2.267	1.372	3.742	2.077
Trivandrum	0.428	1.257	1.101	2.374	2.653
MEP.pdf					
	λ_0	**λ_1**	**λ_2**	**λ_3**	**λ_4**
Calcutta	1.3346712	-0.1270266	0.1090364	-0.0055848	0.0000810
Ahmedabad	3.0483150	-1.3362548	0.3036938	-0.0144778	0.0001655
Trivandrum	1.7331837	-0.3992600	0.1459743	-0.0069372	0.0000886

Figures 6.1-6.3 show the histograms of wind speed probability density functions. The nature of the graphs, reveals that for long-term data, it is difficult to clearly predict the modality of the histogram. Therefore, all the probability mixture distributions have to be fitted on the measured wind speed probability density histogram and wind power density histogram and checked how well they are suitable to model the given histogram. The *GOF* employed to measure their performances are R^2, *RMSE* and *AIC*. The R^2 and *RMSE* reveals the confidence with which the given distribution fit the observed histogram. However, to be sure about which distribution is the better performer, *AIC* has been performed. The lower the value of *AIC* the better the performer the given distribution is to model the observed wind speed data.

Table 6.3 shows the result for the *GOF* for wind speed and wind power density distributions respectively. From Table 6.3 it is seen that all the distributions have the R^2 value greater than 99% except *MEP.pdf* which performs poor followed by Truncated Normal-Normal distribution for all three locations under study. Table 6.3 reveals that the proposed *GG.pdf* has lowest *AIC* value for all three stations. Thus, it can be concluded that the proposed *GG.pdf* is the suitable distribution of all the mixture distribution to model the wind speed data for all three locations in India. Figures 6.1-6.3 show the observed wind speed histograms of Calcutta, Ahmedabad, and Trivandrum respectively and the best-suited pdf of mixture distribution, i.e., *GG.pdf*.

Table 6.3: Measure of goodness of fit for wind speed distribution.

GOF	***W.pdf***			***WW.pdf***		
	Calcutta	Ahmed.	Tvm.	Calcutta	Ahmed.	Tvm.
R^2 (%)	99.63	99.96	99.67	99.57	99.97	99.66
RMSE	0.0100	0.0051	0.0112	0.0112	0.0046	0.0117
AIC	5.228x10^5	5.321x10^5	4.469x10^5	5.165x10^5	5.318x10^5	4.451x10^5
	GW.pdf			***TNW.pdf***		
R^2 (%)	99.56	99.97	99.55	99.80	99.96	99.66
RMSE	0.0114	0.0048	0.0135	0.0077	0.0053	0.0120
AIC	5.157x10^5	5.316x10^5	4.438x10^5	5.219x10^5	5.320x10^5	4.457x10^5
	TNG.pdf			***GG.pdf***		
R^2 (%)	99.79	99.97	99.76	99.70	99.97	99.67
RMSE	0.0078	0.0049	0.0100	0.0094	0.0043	0.0116
AIC	5.198x10^5	5.318x10^5	4.440x10^5	**5.120x10^5**	**5.309x10^5**	**4.415x10^5**
	TNN.pdf			***MEP.pdf***		
R^2 (%)	98.39	99.95	99.75	98.27	98.95	98.54
RMSE	0.0217	0.0061	0.0107	0.289	0.0316	0.0298
AIC	5.762 x10^5	5.347 x10^5	4.496 x10^5	5.182x10^5	5.357x10^5	4.456x10^5

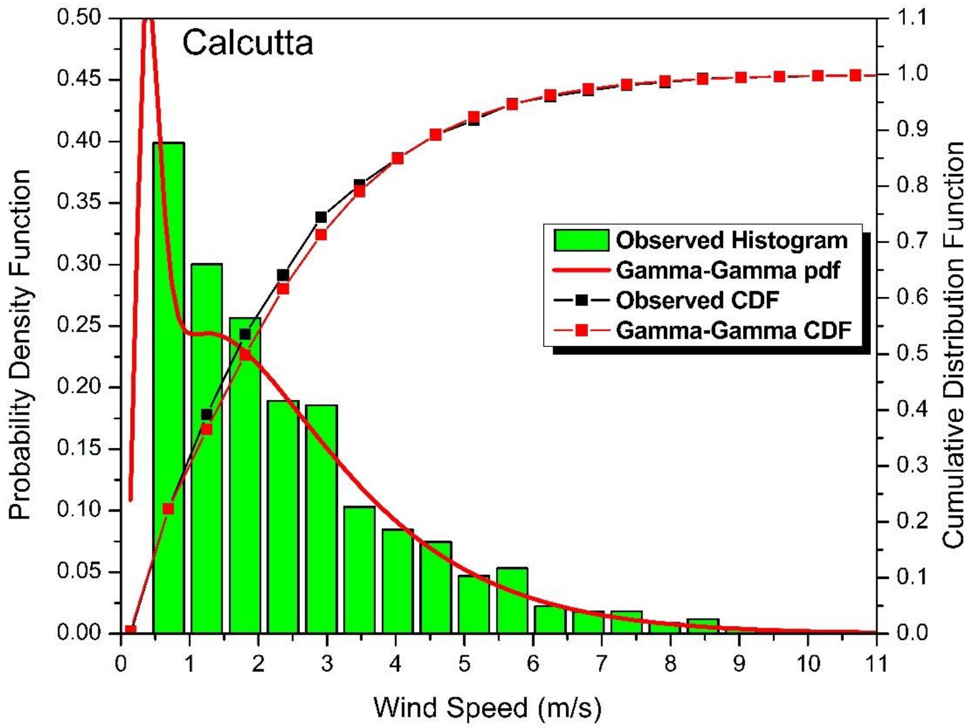

Figure 6.1: Wind speed distributions and corresponding cumulative distribution functions for Calcutta.

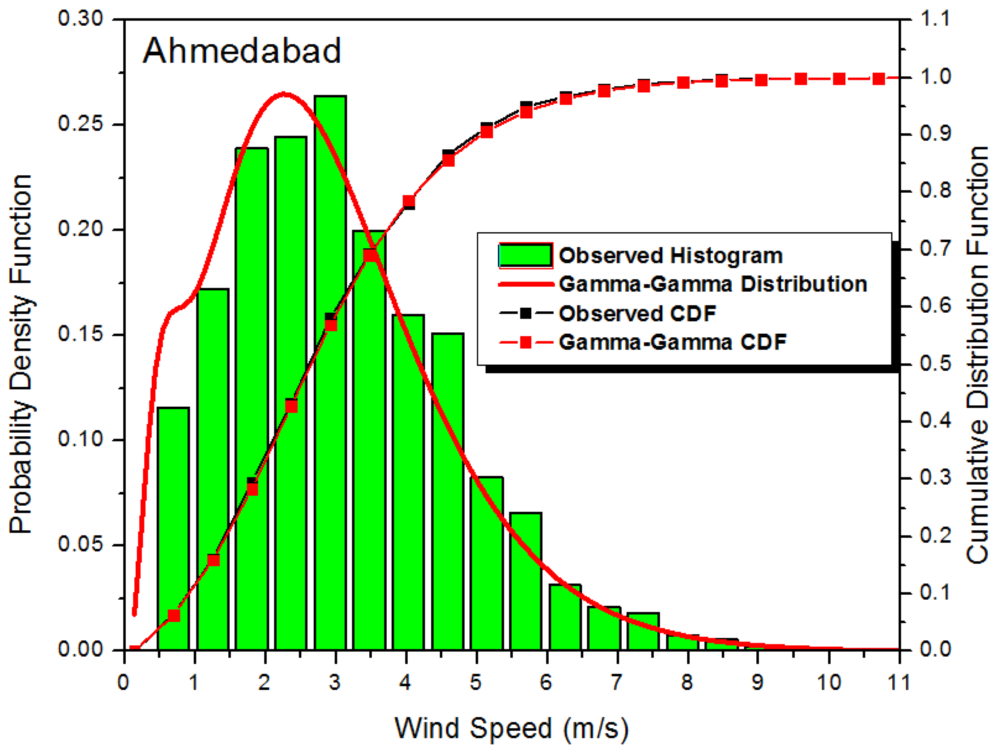

Figure 6.2: Wind speed distributions and corresponding cumulative distribution functions for Ahmedabad.

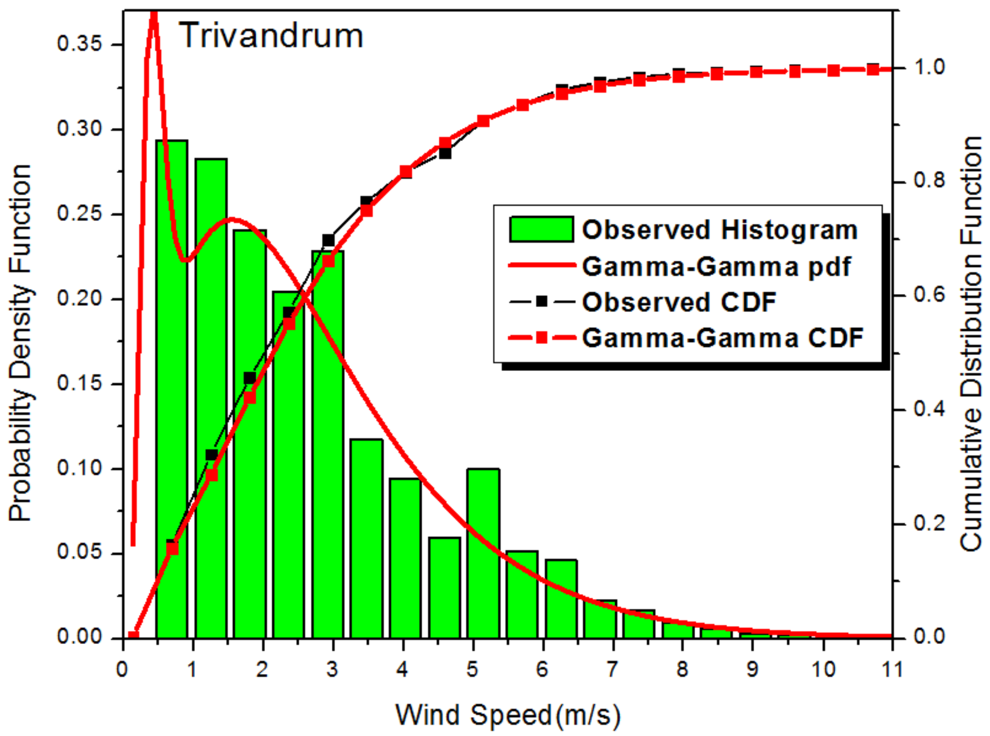

Figure 6.3: Wind speed distributions and corresponding cumulative distribution functions for Trivandrum.

Figure 6.4 shows the wind power density histogram of wind speed data with the best suitable mixture distribution. The proposed *GG.pdf* found to be in good agreement with the wind power density histogram of all three stations. The R^2 value of *GG.pdf* has been mentioned inside the figures itself. Table 6.4 shows the comparison of various distributions based on percentage error in wind power density. As seen from this table that proposed *GG.pdf* shows minimum percentage error in wind power density. Thus it can be concluded from the comparison of different distributions that *GG.pdf* is suitable to model all three stations of India.

Table 6.4: Comparison of various distributions based on percentage error in wind power density

Distributions	**Error in Wind Power Density (%)**		
	Calcutta	Ahmedabad	Trivandrum
W.pdf	10.397	6.7785	8.036
WW.pdf	11.046	5.7362	5.706
TNW.pdf	5.1183	6.6101	5.1621
GW.pdf	11.357	5.5375	7.1538
TNN.pdf	5.446	6.066	3.4636
TNG.pdf	7.947	4.7663	4.0695
GG.pdf	**4.783**	**3.9065**	**3.289**
MEP.pdf	6.3523	8.8765	7.4689

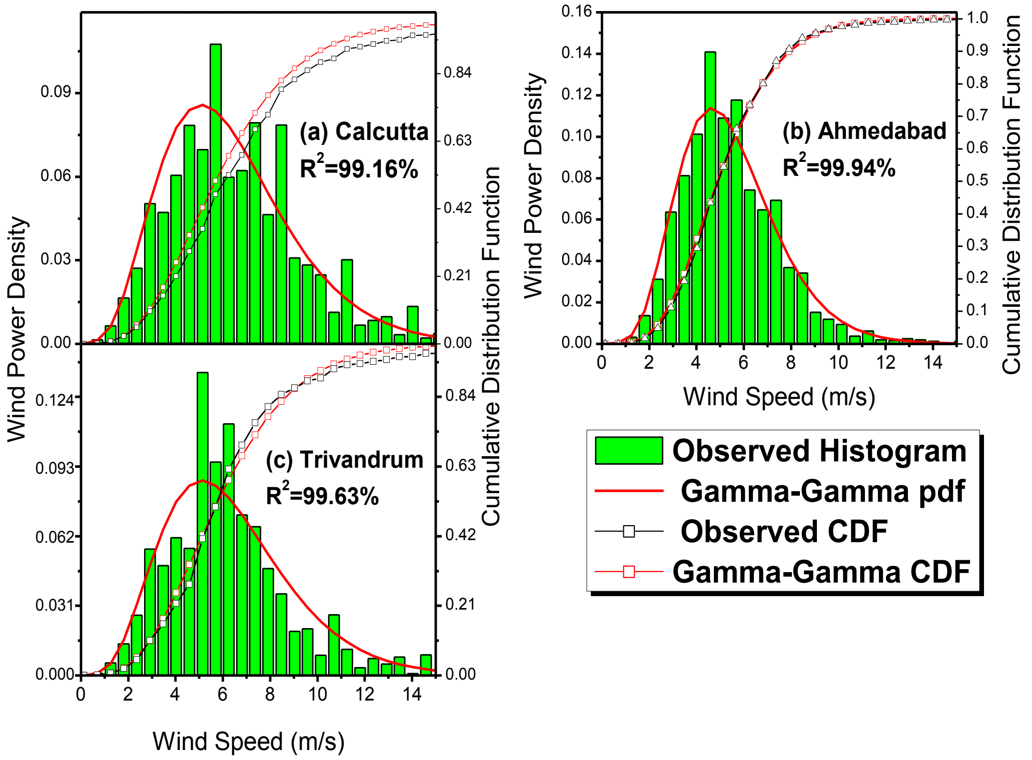

Figure 6.4: Wind power density distributions and corresponding cumulative distribution functions for (a) Calcutta (b) Ahmedabad (c) Trivandrum.

The *MEP.pdf* shows poor results for all three stations in both wind speed and wind power density assessment. Therefore, *MEP* could not be a suitable alternative for analyzing wind speed data, especially, for Indian climate. There are certain limitations associated with *MEP* pdf too **[19, 20]** (a) The cumulative distribution function of *MEP.pdf* variable cannot be expressed in closed form. (b) The *MEP.pdf* has no scale structures. (c) The estimation of the parameters and functions (e.g. variance, modes, median etc.) and the standard error of the *MEP.pdf* is more laborious. (d) There is no specific *GOF* available for *MEP.pdf*.

6.4 Summary

In this chapter, the long-term hourly mean wind speed data of three locations namely Calcutta, Ahmedabad and Trivandrum were analyzed by the mixture distribution and *MEP.pdf* (constraint moment = 4). The base distribution for comparison is 2-parameter *W.pdf*. The suitability of the distribution was judged based on R^2, *RMSE*, *AIC*. From the above discussions, it has been summarize that the proposed *GG.pdf* shows the lowest *AIC* value and are the suitable distribution to model wind speed data for all three locations in India. This distribution performs equally well is estimating the wind power density. Thus, *GG.pdf* can be proved to be a suitable alternative to conventional 2-parameter *W.pdf* to assess the wind speed data for Indian climatology.