

## RESEARCH ARTICLE

# Wind speed analysis and energy calculation based on mixture distributions in Narakkalliya, Sri Lanka

K.W.G.D.H. Rajapaksha<sup>1</sup> and Kanthi Perera<sup>1,2\*</sup>

<sup>1</sup>Postgraduate Institute of Science, University of Peradeniya, Peradeniya.

<sup>2</sup>Department of Engineering Mathematics, Faculty of Engineering, University of Peradeniya, Peradeniya.

Revised: 15 March 2016; Accepted: 19 May 2016

**Abstract:** Wind is an important resource for generating renewable energy. To effectively estimate wind energy potential, Weibull, lognormal and gamma distribution functions have been widely used. However, when the distribution is bi-modal these conventional distribution functions are not suitable. Since Narakkalliya is affected by both monsoon seasons to varying degrees, the wind speed data show a bi-modal behaviour. A study was conducted at Narakkalliya in 2001 to calculate the energy generated from wind using the conventional Weibull distribution. The objective of this study was to show that energy calculation could be carried out more accurately using mixture distribution. Thus in this study, mixture Weibull distributions, mixture lognormal and Weibull distribution, mixture gamma and Weibull distribution were considered to model the wind speed frequency distribution (WSFD). Parameters of the above mentioned distributions were calculated using maximum likelihood method. The goodness-of-fit of the distributions was compared using K-S error, Chi-square test and root mean square error. Mixture Weibull distribution had the lowest error followed by the mixture lognormal and Weibull distribution. The fitted mixture Weibull distribution and the power curve data for the Narakkalliya wind turbine were used for energy calculation. According to the analysis mixture Weibull estimates 83.2 % of the actual energy while Weibull estimates 75.9 % of the actual.

**Keywords:** Chi-square test, Kolmogorov-Smirnov test, maximum likelihood method, mixture Weibull distribution, power curve data, root mean square error.

## INTRODUCTION

Sri Lanka experiences two distinct monsoonal wind seasons, which are the Northeast (NE) Monsoon from December to February and the Southwest (SW) Monsoon from mid-May to September. The two transitional periods in between are referred to as inter-monsoon

seasons. Strong monsoon winds can be used to generate electricity using modern state-of-the-art wind turbine generators. The annual energy generation potential of a wind power plant is estimated using the wind turbine power curve and the wind speed frequency distribution (WSFD) derived from measured time-series wind speed data - typically over a year. The present study has been carried out to model the WSFD using hourly wind speed data at 40 m mast height for the year 2001 at Narakkalliya. Literature indicates that the most suitable distribution for wind statistics modelling and wind power evaluation is the Weibull distribution (Drobinski & Coulais, 2012). In Sri Lanka a study has been carried out to model the wind speed data and for energy calculation at Narakkalliya using the Weibull distribution (Fernando, 2010). Although this seems to be a reasonable approximation, it is critical to note that the Weibull distribution does not show a high accuracy in modelling WSFD in locations, which are strongly affected by both monsoons, hence displaying a dual peak scenario. The present study shows that energy calculation could be carried out more accurately using mixture distributions.

The most widely used distribution for the characterisation of wind speed is the 2-parameter Weibull distribution (W2) (Manwell *et al.*, 2002). In some wind climates, the simple 1-parameter Rayleigh (RAY) distribution offers a more concise fit to a sample, since it is a special case of the W2 (Lackner *et al.*, 2008). However, having only a single model parameter makes the RAY distribution much less flexible. Despite the widespread acceptance of the W2 and RAY distributions, Carta *et al.* (2008) have noted that under different wind regimes other distributions may fit wind samples better. The other distributions used to characterise wind speed

\* Corresponding author (kanthip@pdn.ac.lk)

include the 3-parameter generalised gamma (GG), 2-parameter gamma (G2), inverse Gaussian, 2-parameter lognormal (LN2), 3-parameter beta, and the bimodal (two component mixture) Weibull model (BIW). Kiss and Janosi (2008) recommended the GG for the ERA-40 dataset (6-hourly mean) after testing the RAY, W2 and LN2 distributions.

In the recent past, mixture distributions were used to estimate the wind energy potential that are quite accurate in describing wind speed characteristics. Jaramillo and Borja (2004) used mixture Weibull distributions (WW) to model bi-modal wind speed frequency distribution. Akpinar & Akpinar (2009) used the mixture normal and Weibull distribution (NW), which is a mixture of truncated normal distribution and conventional Weibull distribution to model wind speeds. Chang (2011) employed mixture gamma and Weibull distribution (GW), which is a combination of gamma and Weibull distribution, and also mixture normal distribution (NN), which is a mixture function of two-component truncated normal distribution for wind speed modelling.

**METHODOLOGY**

The analysis was carried out for Narakkalliya (in the Kalpitiya peninsula) using hourly average wind speed data (m/s) obtained for the year 2001 at 40 m mast height operated by the Ceylon Electricity Board. Frequency distribution was obtained for a range of wind speed from 0 – 18.5 m/s. The width of the wind speed frequency bin was selected as 1 m/s and the bin range was specified as 0.5 – 1.5, 1.5 – 2.5 etc., up to 18.5 m/s. The annual energy production at Narakkalliya was calculated using power curve data obtained from a reputed wind turbine manufacturer.

**Weibull distribution**

In the conventional (two-parameter) Weibull probability density function (pdf) the variable  $v$  is given by:

$$f(v; \alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{v}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{v}{\beta}\right)^\alpha\right]$$

where,  $\alpha$  is the shape parameter and  $\beta$  is the scale parameter.

The Weibull cumulative distribution function (cdf) is given by:

$$F(v; \alpha, \beta) = 1 - \exp\left[-\left(\frac{v}{\beta}\right)^\alpha\right]$$

The parameters were calculated using the maximum likelihood method given by:

$$\alpha = \left[ \frac{\sum_{i=1}^n v_i^\alpha \ln(v_i)}{\sum_{i=1}^n v_i^\alpha} - \frac{\sum_{i=1}^n \ln(v_i)}{n} \right]^{-1}$$

$$\beta = \left( \frac{1}{n} \sum_{i=1}^n v_i^\alpha \right)^{\frac{1}{\alpha}}$$

where,  $v_i$  is the variable value in time step  $i$  and  $n$  are the numbers of nonzero data points.

**Lognormal distribution**

Lognormal distribution is the probability distribution of a random variable  $v$ , whose logarithm is normally distributed. Lognormal pdf is given by:

$$\ln(v; \phi, \lambda) = \frac{1}{v\phi\sqrt{2\pi}} \exp\left[\frac{-(\ln(v) - \lambda)^2}{2\phi^2}\right]$$

Lognormal cdf is written as

$$LN(v; \phi, \lambda) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left[\frac{\ln(v) - \lambda}{\phi\sqrt{2}}\right]$$

where

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

Lognormal parameters  $\lambda$  and  $\Phi$  were estimated using the maximum likelihood method:

$$\lambda = \frac{1}{N} \sum_{i=1}^N \ln(v_i)$$

$$\phi^2 = \frac{1}{N} \sum_{i=1}^N [\ln(v_i) - \lambda]^2$$

**Gamma distribution**

The probability density function of gamma distribution is,

$$g(x; \alpha, \beta) = \frac{v^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \exp\left[-\frac{v}{\beta}\right]$$

The cumulative gamma distribution function is given by:

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

The gamma distribution parameters were estimated using the maximum likelihood method that maximises the logarithm of likelihood function which is given by:

$$LL = \ln \prod_{i=1}^n h(v_i; \alpha, \beta) = \sum_{i=1}^n \ln \{h(v_i; \alpha, \beta)\}$$

**Mixture Weibull distribution (WW)**

The probability density function of a mixture Weibull distribution can be written as:

$$g(v; w, \alpha_1, \beta_1, \alpha_2, \beta_2) = wf(v; \alpha_1, \beta_1) + (1-w)f(v; \alpha_2, \beta_2)$$

Its cumulative distribution function is given by:

$$G(v; w, \alpha_1, \beta_1, \alpha_2, \beta_2) = wF(v; \alpha_1, \beta_1) + (1-w)F(v; \alpha_2, \beta_2)$$

where  $0 \leq w \leq 1$  is a weight parameter; which represents the proportion of component distributions being mixed.

$$LL = \ln(v_i; w, \alpha_1, \beta_1, \alpha_2, \beta_2) = \ln \prod_{i=1}^n \left\{ w \left\{ \frac{\alpha_1}{\beta_1} \left( \frac{v_i}{\beta_1} \right)^{\alpha_1-1} \exp \left[ - \left( \frac{v_i}{\beta_1} \right)^{\alpha_1-1} \right] \right\} + (1-w) \left\{ \frac{\alpha_2}{\beta_2} \left( \frac{v_i}{\beta_2} \right)^{\alpha_2-1} \exp \left[ - \left( \frac{v_i}{\beta_2} \right)^{\alpha_2-1} \right] \right\} \right\} = \sum_{i=1}^n \ln \left\{ w \left\{ \frac{\alpha_1}{\beta_1} \left( \frac{v_i}{\beta_1} \right)^{\alpha_1-1} \exp \left[ - \left( \frac{v_i}{\beta_1} \right)^{\alpha_1-1} \right] \right\} + (1-w) \left\{ \frac{\alpha_2}{\beta_2} \left( \frac{v_i}{\beta_2} \right)^{\alpha_2-1} \exp \left[ - \left( \frac{v_i}{\beta_2} \right)^{\alpha_2-1} \right] \right\} \right\}$$

**Mixture lognormal and Weibull distribution (LW)**

The probability density function of the mixture distribution comprising Weibull and lognormal distributions is:

$$u(v; w, \lambda, \phi, \alpha, \beta) = wl(v; \lambda, \phi) + (1-w)f(v; \alpha, \beta)$$

Its cumulative distribution function is given as

$$U(v; w, \lambda, \phi, \alpha, \beta) = wL(v; \lambda, \phi) + (1-w)F(v; \alpha, \beta)$$

$$LL = \ln(v_i; w, \lambda, \phi, \alpha, \beta) = \ln \prod_{i=1}^n \left\{ w \left\{ \frac{1}{v_i \phi \sqrt{2\pi}} \exp \left[ - \frac{(\ln(v_i) - \lambda)^2}{2\phi^2} \right] \right\} + (1-w) \left\{ \frac{\alpha}{\beta} \left( \frac{v_i}{\beta} \right)^{\alpha-1} \exp \left[ - \left( \frac{v_i}{\beta} \right)^{\alpha-1} \right] \right\} \right\} = \sum_{i=1}^n \ln \left\{ w \left\{ \frac{1}{v_i \phi \sqrt{2\pi}} \exp \left[ - \frac{(\ln(v_i) - \lambda)^2}{2\phi^2} \right] \right\} + (1-w) \left\{ \frac{\alpha}{\beta} \left( \frac{v_i}{\beta} \right)^{\alpha-1} \exp \left[ - \left( \frac{v_i}{\beta} \right)^{\alpha-1} \right] \right\} \right\}$$

**Mixture gamma and Weibull distribution (GW)**

The probability density function of the mixture gamma and Weibull distribution are given by the following equation.

$$h(v; k, c, \alpha, \beta, w) = wg(v; k, c) + (1-w)f(v; \alpha, \beta)$$

The relevant likelihood function is,

$$LL = \sum_{i=1}^n \ln \{ wg(v; k, c) + (1-w)f(v; \alpha, \beta) \}$$

**Judgement criteria**

In order to check the accuracy of fitting a theoretical probability density function with the observed data, three types of statistical errors were considered as judgement criteria, which are shown below. Basically the smaller the errors, the better the fit is.

**Kolmogorov-Smirnov test**

The Kolmogorov-Smirnov test (K-S) is defined as:

$$Q = \max |T(v) - O(v)|$$

where  $T(v)$  is the cumulative distribution function calculated from a theoretical function; similarly  $O(v)$  is calculated from observed data.

**Chi-square error**

Chi-square error is used to assess whether the observed probability differs from the predicted probability. Chi-square error is given by:

$$x^2 = \sum_{i=1}^n \left[ \frac{1}{y_{ic}} (y_i - y_{ic})^2 \right]$$

where,  $y_i$  is the actual value at time stage  $i$ ,  $y_{ic}$  is the value computed from correlation expression for the same stage, and  $n$  is the number of data.

**Root mean square error**

Root mean square error (RMSE) provides a term-by-term comparison of the actual deviation between observed probabilities and predicted probabilities. A lower value of RMSE indicates a better distribution.

$$RMSE = \left[ \frac{1}{n} \sum_{i=1}^n (y_i - y_{ic})^2 \right]^{\frac{1}{2}}$$

**Calculation of annual energy production**

The Primary purpose of deriving the WSFD is to calculate the annual energy production of the wind turbine. The annual energy production is calculated using power curve data obtained from a reputed wind turbine.

$$AEP = \sum E_i = \sum f_i \cdot P_i$$

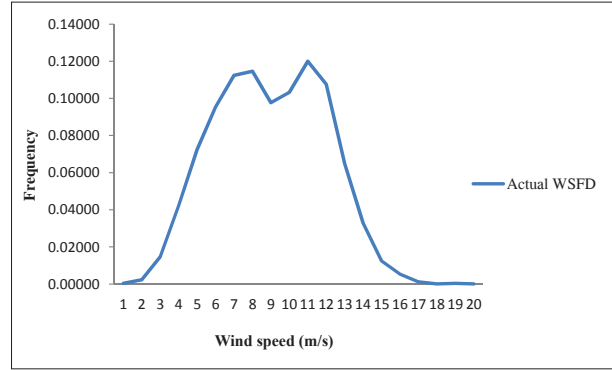
where,

$AEP$  = annual energy production,  $E_i$  = energy per wind speed class  $i$ ,  $f_i$  = frequency of wind speed class  $i$ , and  $P_i$  = power curve data in a wind speed class  $i$ .

**RESULTS AND DISCUSSION**

The line graph of WSFD for the year 2001 at 40 m mast height for Narakkalliya (Figure 1) shows that there is a bi-modal behaviour, signifying the influence of both Southwest and Northeast Monsoons in the local wind climate.

The most common distributions, Weibull, lognormal and gamma, which have been used in the literature were fitted to model wind speed data at Narakkalliya. Table 1 shows the maximum likelihood parameter estimates of these distributions.



**Figure 1:** Line graph of WSFD in year 2001 at 40 m for Narakkalliya

**Table 1:** Maximum likelihood parameter estimates of Weibull, lognormal and gamma distributions

	Parameter	Value
Weibull	$\alpha$	0.00093
	$\beta$	7.4741
Lognormal	$\phi$	1.9068
	$\lambda$	0.52649
Gamma	$k$	5.9051
	$c$	1.2737

Figure 2 shows the Weibull fit diagnostics for hourly average wind speed data. The first measure is the probability density plot of the Weibull distribution (Figure 2a). It can be clearly seen that the Weibull distribution does not fit well with the actual data.

The second measure is the Q-Q plot (Figure 2c). According to the Q-Q plot the sample quantiles match with the theoretical quantiles only at first quarter; at the second quarter the sample quantiles does not match with the theoretical quantiles.

The third measure is the cumulative distribution function (Figure 2b), which shows the difference between the theoretical cumulative distribution and the empirical cumulative distribution. The cumulative distribution function also shows that the conventional Weibull distribution is not a suitable distribution to represent the data. Similarly, Figure 2d also confirms the inapplicability of the conventional Weibull distribution to model the wind speed data at Narakkalliya.

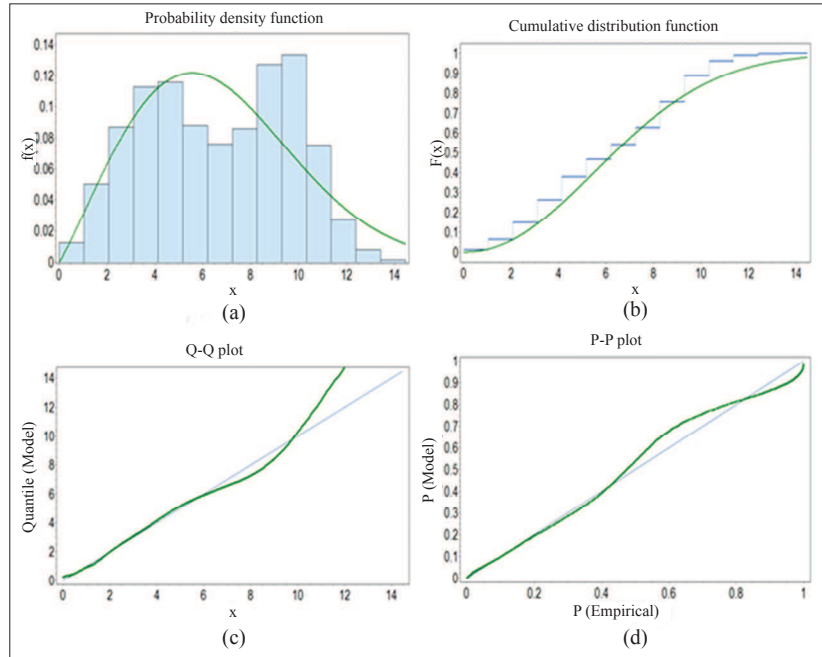


Figure 2: Weibull fit diagnostics for hourly average wind speed data

Similarly, it can be shown that lognormal and gamma distributions are also not suitable distributions to model the wind speed data at Narakkaliya. Thus when describing a bi-modal behaviour the accuracy drops considerably when using the above mentioned distributions. Therefore, the necessity in accurately modelling bi-modal wind speed distributions becomes a requirement. Thus the suitability of mixture distributions were considered in this study.

The first mixture distribution is the mixture Weibull distribution, which is the most common distribution used to describe bi-modal behaviour in wind speed data. Table 2 shows the maximum likelihood parameter estimates of the mixture Weibull distribution.

The line graphs of the probability density function of mixture Weibull derived WSFD and the actual WSFD, and the line graph of mixture Weibull derived cumulative WSFD and the actual cumulative WSFD (Figure 3)

Table 2: Maximum likelihood parameter estimates of mixture Weibull distribution

	Parameters				
	w	$\alpha_1$	$\alpha_2$	$\beta_1$ (m/s)	$\beta_2$ (m/s)
Values	0.458341	2.666	5.782	4.38	9.6392

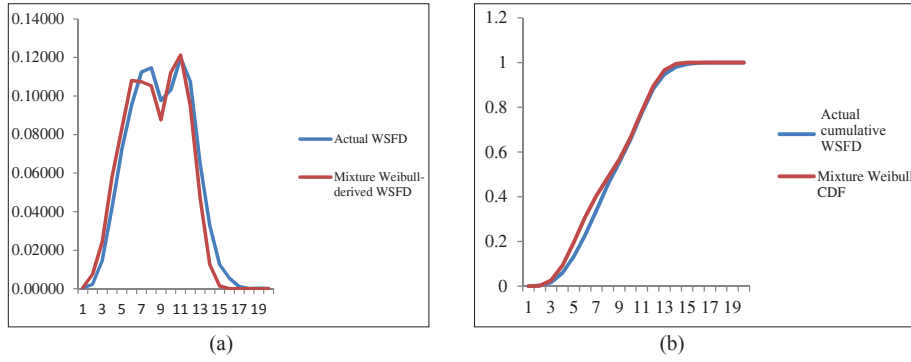
show that the mixture Weibull distribution captures both peaks and fits well with the actual data. There is a slight displacement between wind speeds 6 – 9 m/s and 14 – 18.5 m/s.

To see whether a more accurate distribution can be obtained to represent the actual data, a mixture of lognormal and Weibull distributions and a mixture of gamma and Weibull distributions were considered.

The second mixture distribution is the mixture lognormal and Weibull distribution. Table 3 shows the maximum likelihood parameter estimates of the mixture distribution.

The line graphs of the probability density function of mixture lognormal and Weibull distribution and actual WSFD, and the cumulative distribution function of mixture lognormal and Weibull distribution and actual WSFD (Figure 4), clearly show that the displacement between the theoretical distribution and actual data is larger compared with the mixture Weibull distribution. Thus the mixture lognormal and Weibull distribution does not fit well with the Narakkaliya wind speed data.

The third mixture distribution is the mixture gamma and Weibull distribution. Table 4 shows the maximum likelihood parameter estimates of the mixture gamma and Weibull distribution.



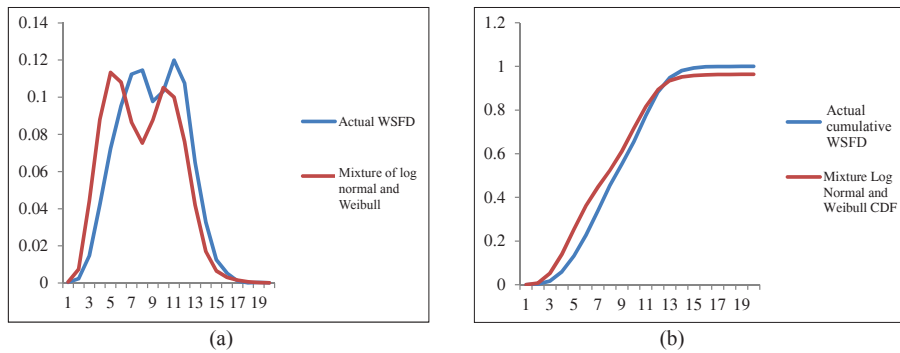
**Figure 3:** (a) Probability density function and (b) cumulative distribution function of mixture Weibull-derived WSFD and actual WSFD

**Table 3:** Maximum likelihood parameter estimates of mixture lognormal and Weibull distribution

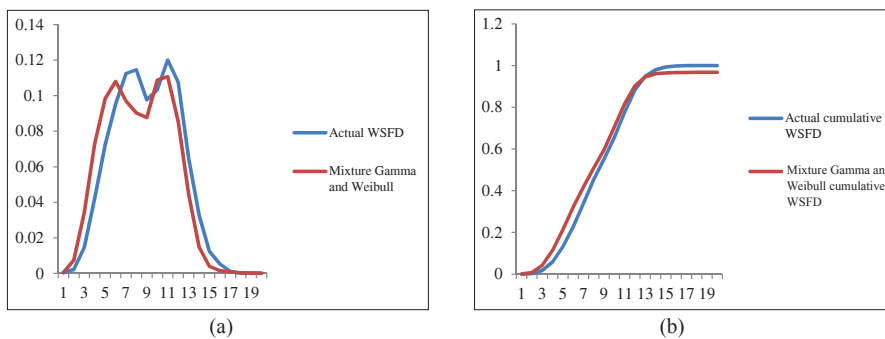
	Parameters				
	w	$\alpha$	$\beta$	$\lambda$	$\Phi$
Values	0.5978	3.7812	11.411	1.1627	0.7333

**Table 4:** Maximum likelihood parameter estimates of mixture gamma and Weibull distribution

	Parameters				
	w	k	c	$\alpha$	B
Values	0.4399	2.3507	1.9608	4.4201	11.9500



**Figure 4:** (a) Probability density function and (b) cumulative distribution function of lognormal and Weibull-derived WSFD and actual WSFD

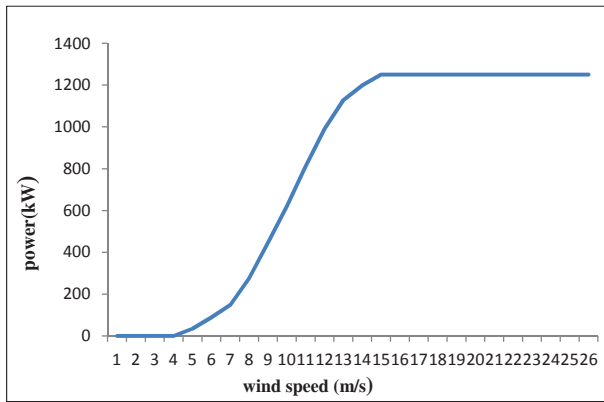


**Figure 5:** (a) Probability density function and (b) cumulative distribution function of mixture gamma and Weibull-derived WSFD and actual WSFD



**Table 5:** Statistical errors for different distribution functions

Function	Max error	RMSE	Chi-square error
Weibull	0.10541	0.01806	0.19421
Lognormal	0.1293	0.0242	0.0601
Gamma	0.10262	0.01883	0.03043
Mixture Weibull	0.02245	0.00732	0.05861
Mixture lognormal and Weibull	0.03965	0.00964	0.07464
Mixture Gamma and Weibull	0.03122	0.00826	0.06789

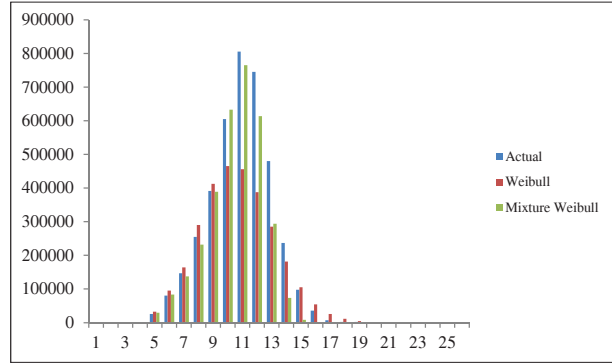


**Figure 6:** Line graph of power curve data for wind turbine

From the line graphs of probability density function of mixture gamma and Weibull distribution and actual WSFD (Figure 5), it can be seen that this distribution fits better than the mixture lognormal and Weibull distribution.

Table 5 shows the statistical errors for the different distribution functions. According to the table conventional distributions have larger errors compared to the mixture distributions. Among the distributions mixture Weibull distribution has the lowest K-S error,  $\chi^2$  error and RMSE, followed by the mixture gamma and Weibull distribution. Thus the mixture Weibull distribution was used for energy calculation.

Energy calculation was carried out using the power curve data obtained for Narakkalliya wind turbine. Figure 6 shows the line graph of power curve data of the wind turbine considered for the Narakkalliya wind regime. According to Figure 6 the cut-in wind speed for the turbine is 3 m/s and the cut-off wind speed is 25 m/s.



**Figure 7:** Energy production of actual and mixture Weibull distribution

The rated output power is 1250 kw and the rated output speed is 14 m/s.

Energy calculation was carried out using the conventional Weibull and mixture Weibull distributions. The mixture Weibull distribution estimates 83.2 % of the actual while the conventional Weibull estimates 75.9 % of the actual. Figure 7 shows the energy produced by the wind turbine for actual data, conventional Weibull and mixture Weibull distributions. According to Figure 7 energy produced by the wind turbine using the mixture Weibull distribution has a similar pattern to the energy produced by the actual data until a wind speed of 11 m/s and after that the actual energy production is slightly higher than in the mixture Weibull energy production. However, it can be clearly seen that the mixture Weibull gives a better approximation than the conventional Weibull distribution.

**CONCLUSION**

This study has shown that for describing a bi-modal behaviour, the accuracy drops considerably when using conventional distributions, confirming the inability to describe the bi-modal behaviour. Thus the applicability of the mixture distributions was compared. Mixture Weibull distribution has the lowest error followed by the mixture gamma and Weibull distribution.

According to the analysis, mixture Weibull distribution estimates 83.2 % of the actual energy while conventional Weibull distribution estimates 75.9 % of the actual energy. Thus this study confirms that the mixture Weibull distribution is the best distribution to be used to calculate the wind energy production for Narakkalliya wind regime.

## Acknowledgement

The authors acknowledge the support provided by the Sustainable Energy Authority (SEA) of Sri Lanka in obtaining wind speed data, and Mr. Sunith Fernando for his insight provided to carry out this research.

---

## REFERENCES

1. Akpınar S. & Akpınar E.K. (2009). Estimation of wind energy potential using finite mixture distribution models. *Energy Conversion and Management* **50**(4): 877 – 884. DOI: <http://dx.doi.org/10.1016/j.enconman.2009.01.007>
2. Carta J.A., Ramirez P. & Velazquez S. (2008). Influence of the level of fit of a density probability function to wind-speed data on the WECS mean power output estimation. *Energy Conversion and Management* **49**(10): 2647 – 2455. DOI: <http://dx.doi.org/10.1016/j.enconman.2008.04.012>
3. Chang T.P. (2011). Estimation of wind energy potential using different probability density functions. *Applied Energy* **88**(5): 1848 – 1856. DOI: <http://dx.doi.org/10.1016/j.apenergy.2010.11.010>
4. Drobinski P. & Coulais C. (2012). Is the Weibull distribution really suited for wind statistic modelling and wind power evaluation. *Journal of Physics: Conference Series* **753**: 5 – 8.
5. Fernando S. (2010). Use of Weibull density function to estimate wind speed frequency distribution. *Annual Publication of PHES of the Department of Electrical and Electronics Engineering*, pp. 19 – 24, Department of Electrical and Electronics Engineering, University of Peradeniya.
6. Jaramillo O.A. & Borja M.A. (2004). Wind speed analysis in La Ventosa, Mexico: a bimodal probability distribution case. *Renewable Energy* **29**(10): 1613 – 1630. DOI: <http://dx.doi.org/10.1016/j.renene.2004.02.001>
7. Kiss P. & Janosi I.M. (2008). Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe. *Energy Conversion and Management* **49**(8): 2142 – 2151. DOI: <http://dx.doi.org/10.1016/j.enconman.2008.02.003>
8. Lackner M.A., Rogers A.L. & Manwell J.F. (2008). Uncertainty analysis in MCP-based wind resource assessment and energy production estimation. *Journal of Solar Energy Engineering - Transactions of the ASME: AIAA 45<sup>th</sup> Aerospace Sciences Meeting and Exhibition*, volume 130, pp. 5 – 8. DOI: <http://dx.doi.org/10.1115/1.2931499>
9. Manwell J.F., McGowan J.G. & Rogers A.L. (2002). *Wind Energy Explained: Theory, Design and Application*. John Wiley & Sons Ltd., West Sussex, UK. DOI: <http://dx.doi.org/10.1002/0470846127>