MODELING WIND SPEED DISTRIBUTIONS USING SKEWED PROBABILITY FUNCTIONS: A MONTE CARLO SIMULATION WITH APPLICATIONS TO REAL WIND SPEED DATA

Kaolee Yang | Department of Mathematics | University of Wisconsin – Eau Claire

Introduction

The fast depletion of non-renewable energy has challenged researchers to look for clean and fuel-efficient sources. Among the bioenergies, wind energy has shown to be clean, fuel-efficient, and cost-effective [1]. Thus, researchers are actively seeking for ways to describe wind speed distribution. The most commonly used distribution is Weibull [2]. However, typical wind speed show skewness and bimodality therefore we focused on flexible skew distributions. We demonstrated the accuracy of each model with application to three datasets and a Monte Carlo simulation.

Methods

Unimodal Skewed Distributions

The probability density function (pdf) of the Weibull, Normal, Gamma, and Lognormal distributions can be found in any standard book on distributions.

Skew Normal Distribution

A skew-normal (SN) distribution has the following probability density function [3]

$$f(x) = 2\varphi(x)\Phi(\alpha x),\ x \in \mathbb{R} \quad (1)$$

Skew-t Distribution

Let $Z \sim SN(\alpha)$ and $W \sim \chi2(\nu)$ be independently distributed. Then, define

$$Y = \frac{Z}{\sqrt{W/\upsilon}} \quad (2)$$

Then the linear transformation $Y = \mu + \sigma X$ has a skew-t (ST) distribution with parameters μ, σ and α.

Flexible Distributions

Alpha Skew Normal Distribution

An alpha skew normal (ASN) distribution has the following pdf

$$f(y;\alpha) = \frac{(1-\alpha y)^2+1}{2+\alpha^2}\phi(y), \quad y \in R. \quad (3)$$

Alpha Skew Logistic Distribution

Let Y be an alpha skew-logistic (ASLG) random variable then Y has the density function

$$f(y;\alpha) = \frac{3((1-\alpha y)^2+1)e^{-y}}{(6+(\alpha^2\pi^2))(1+e^{-y})^2} \quad (4)$$

Alpha Skew Laplace Distribution

An alpha skew Laplace (ASLP) distribution has the pdf

$$\frac{(1-\alpha y)^2+1}{4(1+\alpha^2)}e^{-|y|}, \quad y \in R \quad (5)$$

Bimodal Skew Symmetric Normal

A continuous random variable Y follows a bimodal skew symmetric normal distribution (BSSN) if it has the pdf

$$\psi(y) = \Phi(y) - \frac{y+\mu-2\beta}{1+2\psi(\delta+(\beta-\mu)^2)}\phi(y) \quad (6)$$

Mixture Distributions

Mixture models can be expressed as a combination of densities such that

$$f(y,\psi) = \sum_i^p \pi_i f(x;\theta_i), \quad (7)$$

where $\pi_i \geq 0, i = 1, 2, \ldots, p$ with $\sum_{i=1}^p \pi_i = 1$, are called mixing weights of the pth component of the mixture. Additionally, $\psi = (\pi_1, \pi_2, \ldots, \pi - 1_p, \theta_1, \theta_2, \ldots, \theta_p)$ denotes the vector of parameters of the model.

Applications

We used three datasets to examine the distribution of wind speed data. Two datasets were used to analyze unimodal and bimodal features. The other dataset was used to demonstrate general application to other wind speed data.

Dataset 1 and 2 were obtained from the National Buoy Center at two stations (Station 46014 and 46054) in California. Both datasets were recorded at 5m above sea level at an average of ten minutes. For dataset 1, n = 25130 was used for analysis after deleting missing and N/A values. For dataset 2, n = 15812 was considered for final analysis after the same process.

Dataset 3 (MSP) was obtained from the Minnesota Department of Natural Resources. All data were collected by the National Weather Service and Federal Aviation Administration in miles per hour. We used n = 1827 for analysis.

Application: Dataset 1

When considering unimodal skew distributions, Weibull and the skew-normal provide the best fit. This is evident from their small K-S errors and high R2 value.

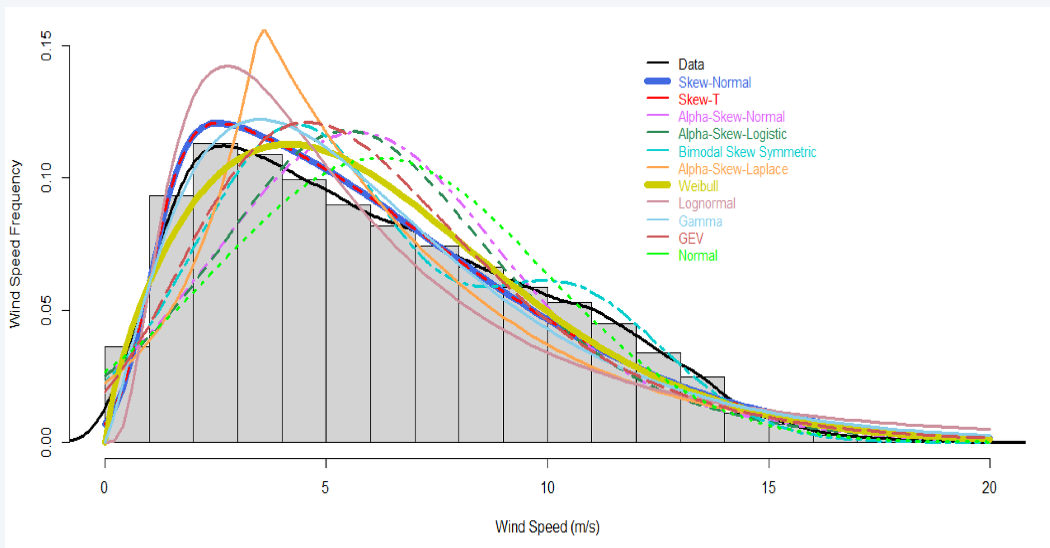

Figure 1: Observed and expected density plot of Station 46014 wind speed data

Fitting mixture distributions to the data, we found that GW provided the closest fit.

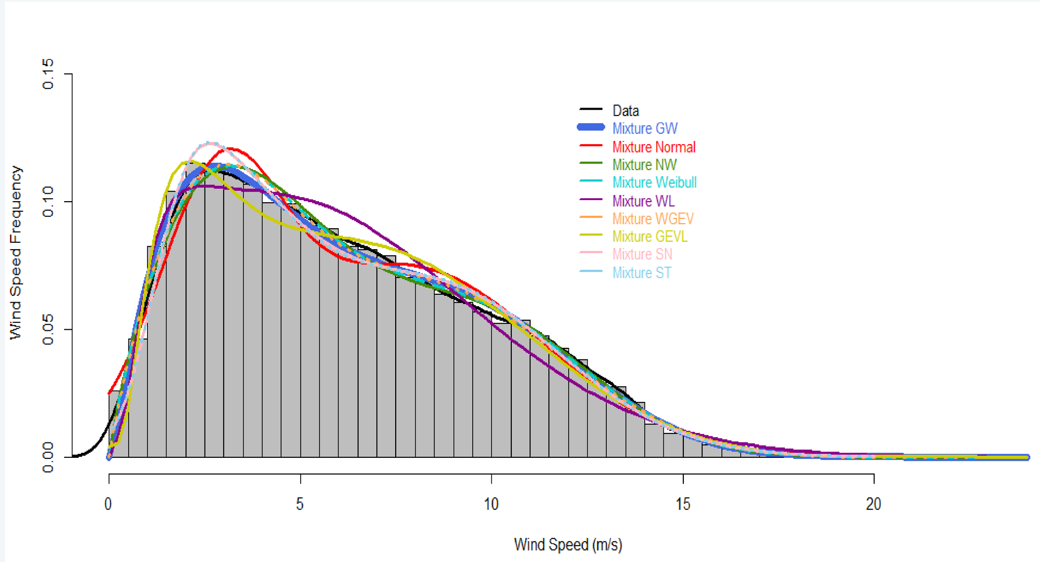

Figure 2: Observed and expected mixture density plot of Station 46014 wind speed data

Application: Dataset 2

Of the univariate distributions, GEV provided the closest fit to the data, followed by BSSN, skew-normal, skew-t, and ASN distributions, which all have similar judgment values.

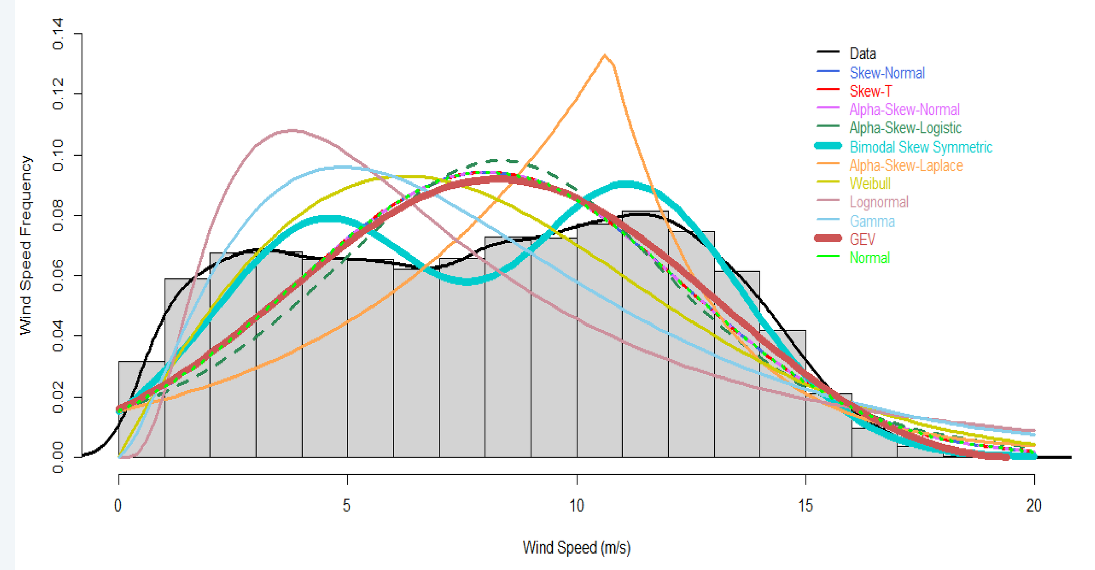

Figure 3: Observed and expected density plot of Station 46054 wind speed data

WGEV and GW provided the closest fit from all mixture models considered.

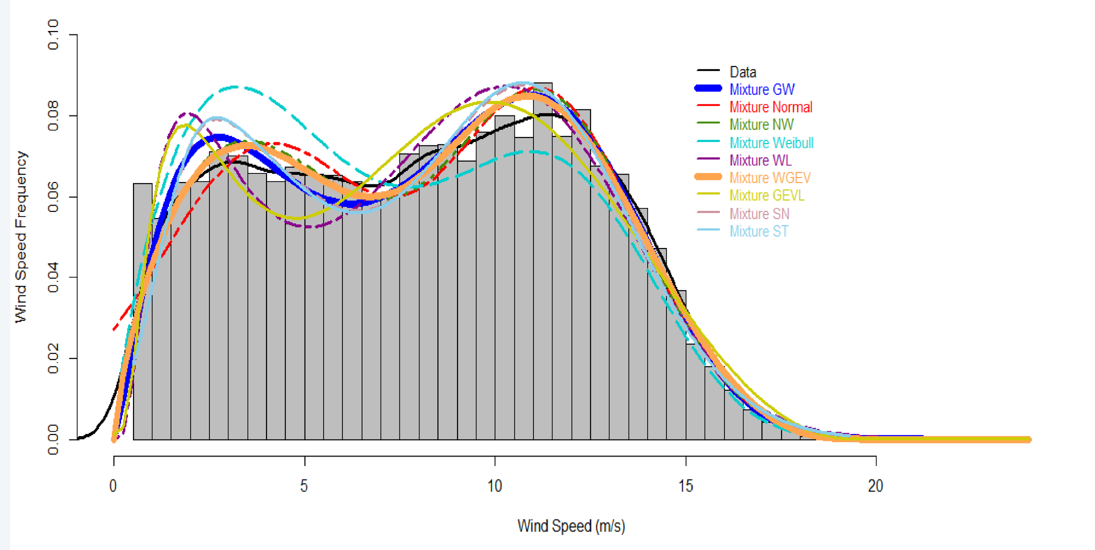

Figure 4: Observed and expected mixture density plot of Station 46054 wind speed data

Application: Dataset 3

Although we do not provide the density plot of the MSP dataset in this poster, gamma, skew-normal and skew-t provided the best fitness evaluation.

Monte Carlo Simulation

Since Weibull and skew-normal demonstrated accuracy and appropriateness in modeling skewed data, we proceeded to conduct an extensive simulation to study the performance of different estimation methods. They include the methods of moments (MOM), maximum likelihood (MLE), least squares (LS), weighted least squares (WLS) and probability weighted moments (PWM). Performance of each method was compared based on their mean square error (MSE), described by the following equation:

$$MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right] = Var(\hat{\theta}) + Bias^2(\hat{\theta}) \quad (8)$$

The location parameter μ, scale parameter σ and shape parameter α were set to 0, 1 and 0.9 without loss of generality throughout the study.

From our results, we found that WLS performed the best for both cases of the Weibull and skew-normal distributions. This was especially true for larger sample sizes of n. Following WLS, MLE performed well for Weibull and LS performed competitively for skew-normal. **Table 1** provides simulated results for Weibull.

n		MSE for $\alpha = 0.9$	MSE for $\sigma = 1$	MSE for $\mu = 0$
	MOM	0.0025*	0.1966	0.0770
	MLE	0.0042	0.0037*	0.0022*
10	*LS*	0.0039	0.0044	0.0032
	WLS	0.0037	0.0044	0.0030
	PWM	0.0281	0.1530	0.0261
	MOM	0.0023*	0.1086	0.0416
	MLE	0.0043	0.0038*	0.0018*
20	*LS*	0.0036	0.0043	0.0025
	WLS	0.0034	0.0039	0.0019
	PWM	0.0182	0.0826	0.0117
	MOM	0.0025*	0.0441	0.0174
	MLE	0.0040	0.0036	0.0012
50	*LS*	0.0032	0.0036	0.0015
	WLS	0.0030	0.0033*	0.0007*
	PWM	0.0094	0.0346	0.0041
	MOM	0.0024	0.0240	0.0091
	MLE	0.0036	0.0034	0.0008
100	*LS*	0.0029	0.0032	0.0008
	WLS	0.0023*	0.0028*	0.0003*
	PWM	0.0056	0.0176	0.0019
	MOM	0.0021	0.0083	0.0033
	MLE	0.0027	0.0028	0.0002
500	*LS*	0.0017	0.0018	0.0002
	WLS	0.0010*	0.0014*	0.0000*
	PWM	0.0018	0.0040	0.0004

Table 1: Simulated mean square error values for Weibull distribution

Conclusion

The purpose of this research was to present the most accurate and practical distributions for modeling wind speed data. We found that for both unimodal and bimodal skewed data, skew normal families of distribution clearly provided the best fit for wind speed data. For the mixture models, mixture of skew-normal families are not clear winner but the results are very competitive with the ones obtained from conventional mixture distributions.

For the Monte Carlo simulation, we found that WLS performed the best across most parameters and sample sizes. MLE and LS followed behind.

We are hopeful that the framework presented in this paper will provide us better understanding of the distribution of wind speed data. Finally, we had to delete lots of missing values for dataset 1 and 2. For future research, some other skewed models can be chosen that can incorporate the missing values for more accurate results.

Acknowledgements

We thank the McNair Post Baccalaureate Program, Office of Research and Sponsored Programs and the UWEC Mathematics Department for their support of our project. We also thank the National Buoy Center and the Minnesota Department of Natural Resources for facilitating the data collection. Additionally, we thank Learning and Technology Services for printing this poster.

References

[1] Burton, T., Sharpe, D., Jenkins, N., & Bossanyi, E. (2011). Historical development. In *Wind Energy Handbook* (2nd ed.). John Wiley & Sons.

[2] Ouarda, T., Charron, C., Shin, J., Marpu, P., Al-Mandoos, A., Tamimi, M., . . . Al Hosary, T. (2015). Probability distributions of wind speed in the UAE. *Energy Conversion and Management,93*, 414-434.

[3] Azzalini, A. (1985). A class of distributions which includes the normal ones. Scand. J. Statist., 12, 171-178.

The Power of AND

Advisor: Dr. Mohammad Aziz