RTM: Interactive estimation tool for modelling real-time wind speed

Volkan Soner Özsoy[1 *], H. Hasan Örkcü[1], Hasan Bal[1]

1 Gazi University, * Corresponding author, volkansoner@gazi.edu.tr

Abstract

Renewable energy sources have been become important in the whole world along with the rapid depletion of energy resources. Potential of the wind energy, one of the most important renewable energy sources, in any region can be estimated using the statistical methods. For modelling, various distributions were used the wind speed data in the related modelling literature. Moreover, in the literature, fitting these distributions was performed via static data. Distributions must be dynamically adapted as the wind speed changes over time. The Real- Time Modelling (RTM) tool is proposed to determine the most appropriate distribution of real-time wind speed in this study. The developed RTM tool for modelling real-time wind speed model can determine the best distribution according to some evaluation criteria. This study show that the developed RTM tool work effectively and efficiently in the real-time wind speed data.

Keywords: Modelling, Real-time, Wind speed.

Citation: Özsoy, V. S., Örkcü, H. H., Bal, H. (2018, October) *RTM: Interactive estimation tool for modelling real-time wind speed.* Paper presented at the Fifth International Management Information Systems Conference.

Editor: H. Kemal İlter, Ankara Yıldırım Beyazıt University, Turkey

Received: August 19, 2018, **Accepted:** October 18, 2018, **Published:** November 10, 2018

IMISC 2018 PAPER

Abstract

Renewable energy sources have been become important in the whole world along with the rapid depletion of energy resources. Potential of the wind energy, one of the most important renewable energy sources, in any region can be estimated using the statistical methods. For modelling, various distributions were used the wind speed data in the related modelling literature. Moreover, in the literature, fitting these distributions was performed via static data. Distributions must be dynamically adapted as the wind speed changes over time. The Real-Time Modelling (RTM) tool is proposed to determine the most appropriate distribution of real-time wind speed in this study. The developed RTM tool for modelling real-time wind speed model can determine the best distribution according to some evaluation criteria. This study show that the developed RTM tool work effectively and efficiently in the real-time wind speed data.

Keywords

Modelling; Real-time; Wind speed.

Introduction

Renewable energy sources should be used more often to leave a more livable world wealth for future generations. It also makes renewable energy very significant because of the need for increased energy and the demolition of energy resources. The most popular renewable energy source is wind energy because it is clean, inexhaustible and low cost. The wind speed is the most important parameter to be used when evaluating the potential of a region (Ren et al., 2014). It is necessary to use wind energy effectively and efficiently to select the best suitable region of the wind power plant as well as to make forecasts for the future. Therefore, it is very important to select the appropriate distribution and estimate the distribution to model the wind speed data.

Weibull seems to be the most popular distribution for modeling the real-time data in the literature. However, in addition, Morgan et al. (2011) other well-known statistical distributions such as Log-normal, Gamma, Kappa and Wakaby have modeled offshore wind speeds. Lee et al. (2012) suggested using the Gumbell distribution for modelling the wind data. Soukissian (2013) investigated the modeling performance of the Johnson SB distribution for such data. Alavi et al. (2016a) considered the Gamma, Rayleigh, Lognormal and Weibull distributions for modelling such speed data belonging to five different regions in

Iran. Shin et al. (2016) showed that the optimal distribution for such modelling in UAE is a Weibull-Extreme value type-distribution. Some distributions such as Log-normal, Gamma, Exponential, Inverse Gauss (Zhou et al., 2010), Mixtures Distributions (Akpinar and Akpinar, 2009; Akdağ et al., 2010; Shin et al., 2016), and some flexible distributions based on MaxEnt and MinxEnt principles (Ramírez and Carta, 2006; Kantar and Usta, 2008) were also applied to common literature. For more detailed information, Carta et al. (2009) is recommended.

In this study, the distributions commonly used in the wind energy literature is used by comparing different model selection criteria on instant real wind speed data provided from Istanbul University Observatory.

The remainder of this paper is organized as follows. Widely-used wind speed distributions and a common information about the wind speed data are provided in Section 2. Description of some estimation methods and the model evaluation criteria are briefly provided in section 3. Then, in Section 4, the analysis results and a discussion are presented. The study is concluded in Section 5.

Distributions for Wind Speed and Data

Average wind power density estimation can calculated using Eq (1)

$$P_D = (1/2)\rho A \mu_3 \tag{1}$$

where PD is a distributional model, A is the wind turbine blade sweep area (m2), ρ is air density (kg/m3) and μ3 is the third moment of related distribution (Safari and Gasore, 2010; Alavi et al., 2016b).

Let X be a random variable from Birnbaum-Saunders distribution with the shape parameter $\gamma > 0$ and the scale parameter $\beta > 0$. Its probability density function (pdf) is given by Mohammadi et al. (2017)

$$f(x) = \frac{1}{\sqrt{2\pi}}\left(\frac{\sqrt{x/\beta}+\sqrt{\beta/x}}{2\gamma x}\right)e^{\left(-\frac{\left(\sqrt{x/\beta}-\sqrt{\beta/x}\right)^2}{2\gamma^2}\right)} \tag{2}$$

The exponential pdf is given as follows Alavi et al. (2016b):

$$f(x) = \frac{1}{\mu} e^{\frac{-x}{\mu}} \tag{3}$$

Let X be a random variable from the Extreme value distribution with the location parameter μ and the scale parameter σ. Its *pdf* of the distribution is given by Sarkar et al., (2011)

$$f(x) = \sigma^{-1} e^{\left(\frac{x-\mu}{\sigma}\right)} e^{\left(-e^{\left(\frac{x-\mu}{\sigma}\right)}\right)} \tag{4}$$

The gamma pdf is given as follows Morgan et al., (2011):

$$f(x) = \frac{1}{b^a \Gamma(a)} x^{a-1} e^{\frac{-x}{b}} \tag{5}$$

where $\Gamma(.)$ is the Gamma function, a is a shape parameter, b is a scale parameter.

The *pdf* for the generalized Pareto distribution with shape parameter $k \neq 0$, scale parameter σ, and threshold parameter θ, is

$$f(x) = \frac{1}{\sigma} \left(1 + k \frac{(x-\theta)}{\sigma}\right)^{-1-\frac{1}{k}} \tag{6}$$

for $\theta < x$, when $k > 0$, or for $\theta < x < \theta - \sigma/k$ when $k < 0$.

The pdf of the half-normal distribution is

$$f(x) = \sqrt{\frac{2}{\pi} \frac{1}{\sigma}}\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; x \geq \mu \tag{7}$$

where μ and σ are the location parameter and the scale parameter, respectively.

The inverse Gaussian distribution (also known as Wald distribution) has the density function Bardsley (1980),

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} e^{\left(-\frac{\lambda}{2\mu^2 x}(x-\mu)^2\right)} \tag{8}$$

The pdf of the Logistic distribution (Scerri and Farrugia, 1996) is

$$f(x) = \frac{1}{\sigma}\frac{1}{x}\frac{e^{z}}{(1+e^{z})^{2}}; x \geq 0 \tag{9}$$

where $z = \frac{\log(x)-\mu}{\sigma}$.

The *pdf* of the lognormal distribution (Garcia et al., 1998) is

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\left(\frac{-(\ln x-\mu)^{2}}{2\sigma^{2}}\right)}; x > 0. \tag{10}$$

The Nakagami distribution has the density function (Alavi et al., 2016b)

$$f(x) = 2\left(\frac{\mu}{\omega}\right)^{\mu} \frac{1}{\Gamma(\mu)} x^{2\mu-1} e^{\frac{-\mu_x^2}{\omega}};\ x > 0 \tag{11}$$

with shape parameter μ and scale parameter $\omega > 0$.

The normal distribution pdf with first parameter the mean μ, the second parameter the standard deviation σ is (Crutcher and Baer, 1962),

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^{2}}{2\sigma^{2}}} \tag{12}$$

The Rayleigh pdf (Pishgar-Komleh et al., 2015) is

$$f(x) = \frac{x}{b^{2}} e^{\left(\frac{-x^{2}}{2b^{2}}\right)};\ x > 0 \tag{13}$$

The *pdf* of the Rician distribution with noncentrality parameter $s \geq 0$ and scale parameter $\sigma > 0$ is

$$f(x) = I_0\left(\frac{xs}{\sigma^{2}}\right)\frac{x}{\sigma^{2}}\ e^{-\left(\frac{x^{2}+s^{2}}{2\sigma^{2}}\right)} \tag{14}$$

The *pdf* of the t Location scale distrbiution with the location parameter (μ), the scale parameter (σ) and the shape parameter (v) is

$$f(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sigma\sqrt{v\pi}\Gamma\left(\frac{v}{2}\right)}\left(\frac{v+\left(\frac{x-\mu}{\sigma}\right)^2}{v}\right)^{-\left(\frac{v+1}{2}\right)} \tag{15}$$

where $\Gamma(.)$ is the Gamma function.

The *pdf* for the Weibull distribution with shape parameter b and scale parameter a is

$$f(x) = \frac{b}{a}\left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b}. \tag{16}$$

The wind speed data is taken from Department of Astronomy and Space Sciences, İstanbul University, in this study. It contains wind speeds, recorded at 65 m elevation, obtained from Historical Beyazit Tower and Meteorology Station Observatory located at İstanbul in Turkey. The geographical information of the station is Latitude: 41°00'42" north, Longitude: 28°57'55" east.

Estimation Methods and Model Evaluation Criteria

In this subsection, the maximum likelihood method estimates the unknown parameters of the distributions given in the previous section. Because it has good asymptotic statistical properties, the most preferred estimation method than other estimation methods, it is used in this study.

The aim of using a maximum likelihood estimation (MLE) for fitting a distribution model to wind speed set or any data set is to estimate the parameters of the distribution that maximize the likelihood function (Eq. 17) of having the observed data.

$$L = \prod_{i=1}^{n} f(x_i) = f(x_1).f(x_2)\dots f(x_n) \tag{17}$$

The following criteria are used to determine the performances of the modeling of the distributions: Root mean square error (RMSE), used measure of the differences between the real wind speed values and values predicted by related distribution model, Coefficient of determination (R2), Maximum value of the likelihood function (lnL), Akaike information criterion (AIC), Schwarz criterion or Bayesian information criterion (BIC), which are an estimator of the relative quality of statistical models for a given wind speed data, are used in this study.

The formulas of the model selection criteria in this study are illustrated in Table 1.

INSERT TABLE 1 HERE

In Table 1, $x_{(i)}$ is the *i-th* order statistics, $\hat{F}$ is cumulative distribution function, $\bar{\hat{F}}(x_i) = \frac{1}{n}\sum_{i=1}^{n}\hat{F}(x_i)$, n is the number of wind speed data size, *lnL* is the log-value of Eq(17) and K is the number of the estimated parameters.

Analysis and Results

The descriptive statistics of the data measured every 5 minutes are automatically refreshed thanks to the RTM tool developed without any processing by the user. The descriptive statistics of the data taken momentarily from Beyazıt Observatory are shown in Table 2. For example, the statistics for the first five minutes after the start of the RTM tool are given in the second column in Table 2. Re-descriptive statistics are calculated automatically as new data is available after every five minutes and are in the third column in Table 2. As shown in Table 2, it continues in this way.

INSERT TABLE 2 HERE

To better fitting, it is expected that the values of RMSE, AIC and BIC have the lower values, whereas R^2 and lnL have the higher values. To put it more simply, the higher the R^2 and lnL values and the lower the RMSE, AIC and BIC values, the modeling is the better. As the wind speed measurement is made every 5 minutes, the data of 288 wind speeds are obtained in one day. For modeling in this study. A total of 20.160 wind speed data obtained in 70 days was used. The results of modeling performed by the RTM tool are presented in Table 3. Table 3 shows that the lowest RMSE, AIC and BIC values have the Weibull distribution within the distributions used in this study. Likewise, the highest value of R^2 and lnL has the Weibull distribution.

INSERT TABLE 3 HERE

Figure 1 demonstrate the fitting performance of Weibull and Nakagami for data measured in this study, respectively. Weibull distribution seems a better fitting than other distributions in this study. Nakagami distribution provides best fit as the second good performance. The parameters of the Weibull distribution are shape parameter $b = 1.768$, scale parameter $a = 7.3264$, whereas the parameters of the Nakagami distribution are estimated as shape parameter $\mu = 0.83131$ and scale parameter $\omega = 57.0454$. The interface of the developed RTM tool is shown in Figure 2.

INSERT FIGURE 1 HERE

INSERT FIGURE 2 HERE

Concluding Remarks

The wind speed must be measured and modeled continuously as it changes continuously. Therefore, there is a need for a tool capable of continuous modeling for wind forecasters and other implementers. There are criteria such as root mean square error, coefficient of determination, maximum value of the likelihood function, Akaike information criterion and Bayesian information criterion are used to better fitting. This study proposes the RTM tool for automatically selects the best distribution from the most commonly used distributions of wind speed data. Obtained results from this study with the data from the real-time data transfer observer show that RTM tool works efficiently in the wind speed data fitting and the wind speed data can be modelled by the Nakagami distribution. In addition to the wind speed estimation problems, the RTM tool can be used for all parameter estimation problems in real

References

Akdağ, S. A., Bagiorgas, H. S., & Mihalakakou, G. (2010). Use of two-component Weibull mixtures in the analysis of wind speed in the Eastern Mediterranean. Applied Energy, 87(8), 2566-2573.

Akpinar, S., & Akpinar, E. K. (2009). Estimation of wind energy potential using finite mixture distribution models. Energy Conversion and Management, 50(4), 877-884.

Alavi, O., Sedaghat, A., & Mostafaeipour, A. (2016a). Sensitivity analysis of different wind speed distribution models with actual and truncated wind data: a case study for Kerman, Iran. Energy Conversion and Management, 120, 51-61.

Alavi, O., emadi, K., & Mostafaeipour, A. (2016b). Evaluating the suitability of wind speed probability distribution models: A case of study of east and southeast parts of Iran. Energy Conversion and Management, 119, 101-108.

Bardsley, W. E. (1980). Note on the use of the inverse Gaussian distribution for wind energy applications. Journal of Applied Meteorology, 19(9), 1126-1130.

Carta, J. A., Ramirez, P., & Velazquez, S. (2009). A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands. Renewable and Sustainable Energy Reviews, 13(5), 933-955.

Crutcher, H. L., & Baer, L. (1962). Computations from elliptical wind distribution statistics. Journal of Applied Meteorology, 1(4), 522-530.

Garcia, A., Torres, J. L., Prieto, E., & De Francisco, A. (1998). Fitting wind speed distributions: a case study. Solar energy, 62(2), 139-144.

Kantar, Y. M., & Usta, I. (2008). Analysis of wind speed distributions: Wind distribution function derived from minimum cross entropy principles as better alternative to Weibull function. Energy Conversion and Management, 49(5), 962-973.

Lee, B. H., Ahn, D. J., Kim, H. G., & Ha, Y. C. (2012). An estimation of the extreme wind speed using the Korea wind map. Renewable energy, 42, 4-10.

Mohammadi, K., Alavi, O., & McGowan, J. G. (2017). Use of Birnbaum-Saunders distribution for estimating wind speed and wind power probability distributions: A review. Energy Conversion and Management, 143, 109-122.

Morgan, E. C., Lackner, M., Vogel, R. M., & Baise, L. G. (2011). Probability distributions for offshore wind speeds. Energy Conversion and Management, 52(1), 15-26.

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

Ramírez, P., & Carta, J. A. (2006). The use of wind probability distributions derived from the maximum entropy principle in the analysis of wind energy. A case study. Energy Conversion and Management, 47(15), 2564-2577.

Ren, C., An, N., Wang, J., Li, L., Hu, B., & Shang, D., "Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting." Knowledge-Based Systems 56 (2014): 226-239.

Safari, B., & Gasore, J. (2010). A statistical investigation of wind characteristics and wind energy potential based on the Weibull and Rayleigh models in Rwanda. Renewable Energy, 35(12), 2874-2880.

Sarkar, A., Singh, S., & Mitra, D. (2011). Wind climate modeling using Weibull and extreme value distribution. International Journal of Engineering, Science and Technology, 3(5).

Scerri, E., & Farrugia, R. (1996). Wind data evaluation in the Maltese Islands. Renewable energy, 7(1), 109-114.

Shin, J. Y., Ouarda, T. B., & Lee, T. (2016). Heterogeneous mixture distributions for modeling wind speed, application to the UAE. Renewable Energy, 91, 40-52.

Soukissian, T. (2013). Use of multi-parameter distributions for offshore wind speed modeling: the Johnson SB distribution. Applied energy, 111, 982-1000.

Zhou, J., Erdem, E., Li, G., & Shi, J. (2010). Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites. Energy Conversion and Management, 51(7), 1449-1458.

Figures and Tables

Table 1: Model Selection Criteria

Criteria	**Formulas**
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{F}\left(x_{(i)}\right)-\frac{i}{n+1}\right)^2}$
R^2	$1-\frac{\sum_{i=1}^{n}\left(\hat{F}\left(x_{(i)}\right)-\frac{i}{n+1}\right)^2}{\sum_{i=1}^{n}\left(\hat{F}(x_i)-\bar{\hat{F}}(x_i)\right)^2}$
lnL	$\ln\left(\prod_{i=1}^{n} f(x_i)\right)$
AIC	$-2\log L+2K$
BIC	$-2\log L+K\log n$

Table 2: Descriptive Statistics for Iterative Data

Statistics	**The first 5 minutes**	**The second 5 minutes**	**The third 5 minutes**	**The fourth 5 minutes**
Min	1.6	1.6	1.6	1.6
Max	25.7	25.7	25.7	25.7
Mean	11.765	11.7728	11.7835	11.7970

Variance	39.8702	39.8321	39.8236	39.8533
Skewness	1.8245	1.8252	1.8244	1.8244
Kurtosis	-0.0358	-0.0389	-0.0424	-0.0457
n	555	556	557	558

Table 3. Descriptive Statistics for Real Time Wind Speed Data

Distributions	**RMSE**	**R^2**	**AIC**	**BIC**	**lnL**
Birnbaum Saunders	0.06184	0.95733	108401.8	108417.6	-54198.9
Exponential	0.11099	0.73481	115808.9	115816.8	-57903.5
Extreme Value	0.07367	0.91953	117549.3	117565.1	-58772.7
Gamma	0.04755	0.97547	107819.4	107835.2	-53907.7
Gen. Extreme Value	0.04988	0.97368	109182.8	109206.5	-54588.4
Generalized Pareto	0.07619	0.89937	111361.1	111384.9	-55677.6
Half Normal	0.07593	0.90895	110791.9	110807.7	-55393.9
Inverse Gaussian	0.06604	0.95175	108917.9	108933.8	-54457
Logistic	0.05205	0.96975	112318.9	112334.7	-56157.4
Loglogistic	0.05323	0.96837	110222.5	110238.3	-55109.2
Lognormal	0.05826	0.96252	109000	109015.8	-54498
Nakagami	0.04405	0.97858	107737.9	107753.7	-53866.9
Normal	0.05522	0.96566	111487.6	111503.5	-55741.8
Rayleigh	0.05938	0.96487	108219.1	108227	-54108.5
Rician	0.0594	0.96483	108221.1	108236.9	-54108.5
T Location Scale	0.05522	0.96566	111489.6	111513.4	-55741.8
Weibull	**0.04402**	**0.97869**	**107700.1**	**107716**	**-53848.1**

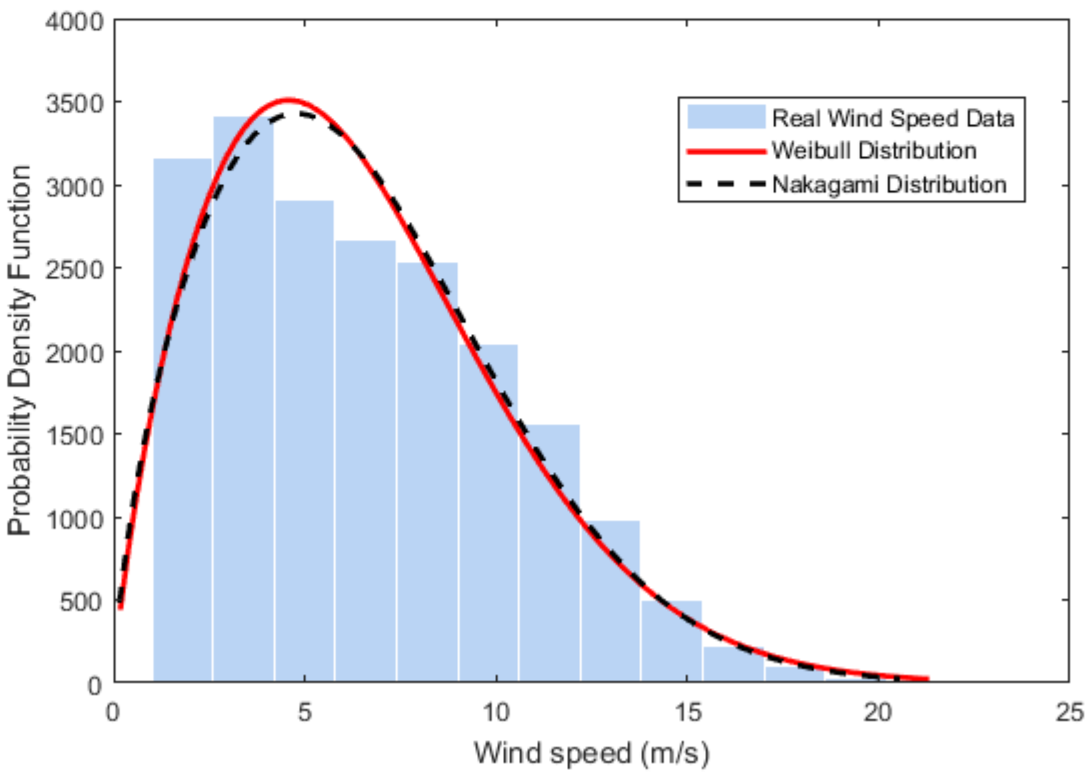

Figure 1: Histogram of wind speed data and *pdf* graphs of the best distributions

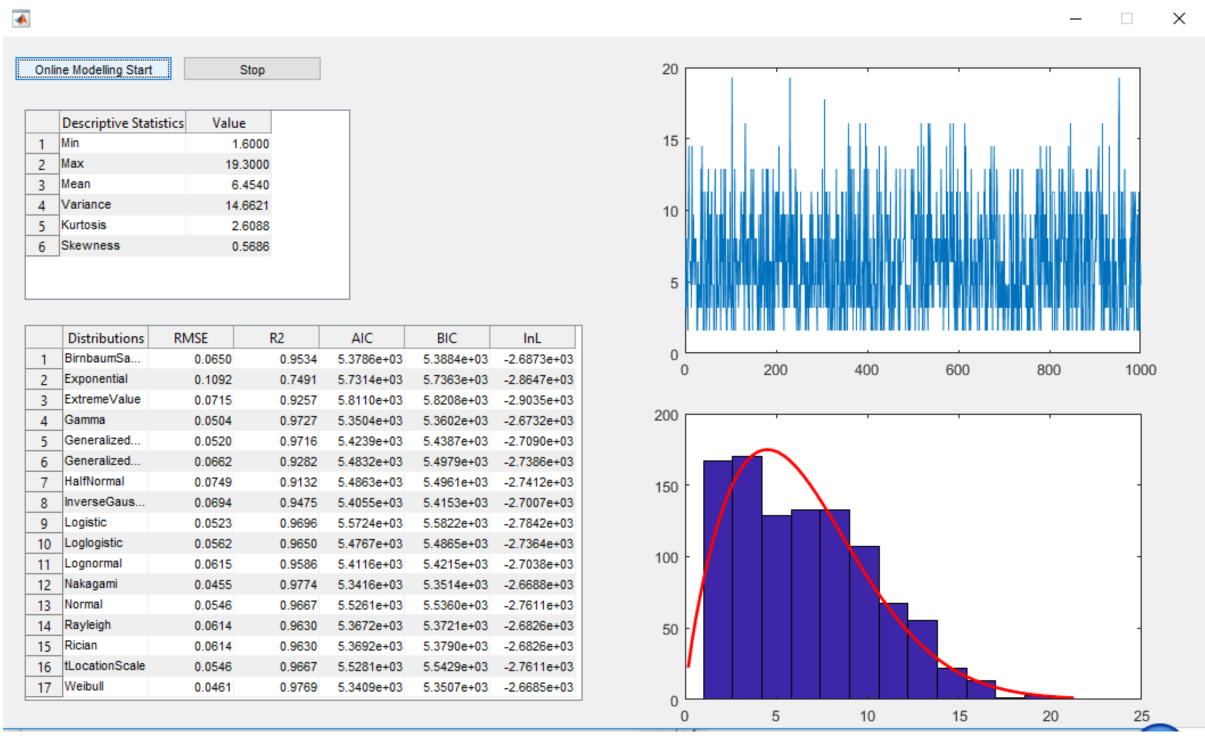

Figure 2: Interface of developed RTM tool