Bayesian Method for Estimating Weibull Parameters for Wind Resource Assessment in the Equatorial Region: a Comparison Between Two-parameter and Three-parameter Weibull Distributions

M. G. M. Khan (khan_mg@usp.ac.fj)
The University of the South Pacific
M. Rafiuddin Ahmed
The University of the South Pacific

Research Article

Keywords: Bayesian method, Weibull distribution, Weibull parameters, Equatorial region, wind resource

Posted Date: May 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-504670/v1

Bayesian method for estimating Weibull parameters for wind resource assessment in the Equatorial region: a comparison between two-parameter and three-parameter Weibull distributions

M.G.M. Khan* and M. Rafiuddin Ahmed

School of Information Technology, Engineering, Mathematics and Physics, The University of the South Pacific, Suva, Fiji.

* Correspondence to khan_mg@usp.ac.fj

Abstract

The two-parameter Weibull distribution has garnered much attention in the assessment of wind energy potential. The estimation of the shape and scale parameters of the distribution has brought forth a successful tool for the wind energy industry. However, it may be inappropriate to use the two-parameter Weibull distribution to accurately characterize wind speed at every location, especially at sites where the frequency of low speed is high, such as the Equatorial region. In this work, for the robustness in wind resource assessment, we first propose a Bayesian approach in estimating Weibull parameters. Secondly, we compare the techniques of wind resource assessment using both two and three-parameter Weibull distributions for different sites in the Equatorial region. The Bayesian inference approach is adopted using Markov Chain Monte Carlo (MCMC) algorithms. Simulation studies conducted in this research confirms that the Bayesian approach seems to be a new robust alternative technique for accurate estimation of Weibull parameters. An appropriate Weibull distribution and the application of the Bayesian approach in estimating distribution parameters were determined using data from six sites in the Equatorial region from 1^{o} N of Equator to 19^{o} South of Equator. Results revealed that a three-parameter Weibull distribution is a better fit for wind data having a greater percentage of low wind speeds (0-1 m/s) and low skewness. However, wind data with a smaller percentage of low wind speeds and high skewness showed better results using a two-parameter Weibull distribution. The results also demonstrate that the proposed Bayesian approach to estimate Weibull parameters is extremely useful in the analysis of wind power potential, as it provides more accurate results while characterizing lower wind speeds.

Nomenclature

2-p	two-parameter Weibull distribution
3-p	three-parameter Weibull distribution
A	scale parameter of Weibull distribution, m/s
AD	Anderson-Darling test
AGL	above ground level
AIC	Akaike information criteria
BAYESIAN	Bayesian estimation method
BIC	Bayesian information criteria
COE	coefficient of efficiency
DIC	deviance information criteria
Γ	gamma function
γ_1	skewness of the wind speed
k	shape parameter of Weibull distribution, dimensionless
KS	Kolmogorov-Smirnov test

log-like	log-likelihood method for the goodness of fit
MAE	mean absolute error
MAPE	mean absolute percentage error
MCMC	Markov Chain Monte Carlo method
MLE	maximum likelihood estimation method
RMSE	root mean square error, m/s
n	number of observations performed
R^2	correlation coefficient
U	wind speed, m/s
$\bar{U}$	mean wind speed, m/s
ρ	air density, kg/m^3
θ	shift of location parameter, m/s
s	standard deviation of wind speed, m/s

1. Introduction

Wind energy has now become one of the world's fastest-growing sources of energy. It is an inexhaustible source of energy with increasing consumptions all around the world. Growing climate change concerns have prompted many developed and developing countries to implement policies that reduce their reliance on non-renewable energy and utilize renewable energy sources such as wind, hydro, and solar instead (Mostafaeipour et al., 2014). However, developing countries encounter several challenges in generating sustainable wind energy. There is a need for reliable wind data and proper assessments of a country's wind energy potential before initiating energy generation projects that would help them meet sustainable development goals outlined by the United Nations.

While climate change is being experienced globally, some regions are getting affected more than the others. Pacific islands countries (PICs), particularly those in the warmer Equatorial region, are more susceptible to its effects. The contribution of the PICs to current global greenhouse gas emissions, according to the UN Permanent Forum on Indigenous Issues (2015), is below 0.03%; yet they are among the first to be affected. It is projected that the people of PICs will be among the first that will need to adapt to climate change or be required to relocate or abandon their traditional homeland. Some islands are already facing the impacts of climate change on their communities, infrastructure, water supply, coastal and forest ecosystems, fisheries, agriculture, and human health. Island states such as Kiribati, Marshall Islands, Tokelau and Tuvalu that are the immediate victims of this phenomenon due to rising sea level. Knowledge of the effects of climate change on PICs should act as a driving force behind the commitment to decrease greenhouse emissions. The PICs, which currently depend heavily on imported oils and their by-products, need to become more energy efficient and self-reliant (Mohanty, 2012). PICs have some of the lowest rates of access to electricity and the prices of electricity are among the highest in the world due to their heavy reliance on high-cost diesel-based generation. Energy security and low-cost energy are becoming increasingly important within the region, which leads to increasing investments in renewable energy technologies. PICs are also some of those most vulnerable to natural disasters. The energy sector can be highly vulnerable to such events, which requires adequate attention to these issues in the design of energy production and distribution infrastructure. This can only be achieved by adopting renewable energy policies. Most of the countries in the region have their national sustainable development plans to achieve United Nations' sustainable development goals (SDGs); for examples, Cook Islands aims to have 100% renewable power generation in near future and Fiji is committed to reducing 30% of its national greenhouse emissions and achieve 99% renewable energy generation by 2030 (VNR Report, 2019).

However, lack of reliable and accurate wind resource data acts as a barrier to an energy-efficient future in the PIC, especially in the smaller developing islands (Kidmo et al., 2015). So far, wind resource assessment has received only limited attention in PICs, and there is a need for further studies on the wind data analysis and accurate wind energy potential assessment. World Bank provides support to PICs through the Sustainable Energy Industry Development Project (SEIDP). In various phases of renewable energy resource mapping, they support the countries to carry out an assessment of solar and wind potential. The objectives of this component are to enhance awareness and knowledge of the potential for renewable technologies (solar and wind) to the governments, power utilities and private sector, and to provide governments with a spatial planning framework to guide investment in the renewable energy sector (PPA 2015).

The utilization of wind power technology to generate energy is slowly increasing in PICs such as New Caledonia, Fiji, Vanuatu, Cook Islands and Samoa. However, there have been little to no attempts to establish wind power in many of these countries. The University of the South Pacific has installed towers of 34 m height, called here as Integrated Renewable Energy Resource Assessment Systems (IRERAS), in Kiribati, Nauru, Niue, Tuvalu, Tokelau, Samoa, Tonga, Fiji, Vanuatu, Solomon and Cook Islands to collect data on wind and solar energies (Gosai, 2014).

In recent years, the Weibull distribution has become a widely accepted tool in determining the potential of wind energy (Indhumathy, 2014). Various parts of the world have widely employed the use of Weibull distributions in the statistical analyses of wind characteristics and wind power density (Corotis et al., 1978). The Weibull shape parameter defines the width of wind distribution. A larger shape parameter indicates that the distribution is narrower, and the peak value is higher. The Weibull scale parameter controls the abscissa scale of the data distribution plot (Chang, 2011). Thus, the Weibull distribution function is comprehensively used for analyzing the wind power potential at a site.

The past researchers have found the two-parameter Weibull distribution to be a useful and practical tool for wind energy estimation. The advantages of two-parameter Weibull distribution include its flexibility, simplicity in parameter estimation, ability to use goodness-of-fit tests on parameters as well as its dependence only on two parameters that can be expressed in closed form.

However, some authors suggested that the distribution is not suited for all wind regimes encountered in nature such as regimes with a high percentage of low wind speeds and bimodal distributions. Therefore, its usage cannot be generalized. To minimize errors, a suitable probability density function must be carefully selected for different wind regimes (Carta et al. 2009; Sukkiramathi and Seshaiah, 2020).

Tuller and Brett (1984) proposed a three-parameter Weibull function in wind analysis and found that it showed better fitness and flexibility than the two-parameter Weibull function. Recently, some authors utilized the three-parameter Weibull distribution and found that it has more flexibility with improved fitness than the two-parameter Weibull distribution in wind energy assessments. Wais (2017) compared the two and three-parameter Weibull distribution to study the most appropriate distribution of wind speed. The results revealed that methods other than the three-parameter Weibull distribution cannot account for cases where the frequency of low wind speed is higher. The author compared the wind speeds for three different sites and found that the three-parameter Weibull distribution performed best when there was a greater frequency of lower wind speeds. Sukkiramathi and Seshaiah (2020) also utilized the three-parameter Weibull distribution for analyzing wind power potential. However, to date, only limited research has been carried out on wind analysis using the three-parameter Weibull distribution.

Furthermore, many estimation methods have been proposed for estimating Weibull parameters. Among these, maximum likelihood estimation (MLE), a popular frequentist technique, has been a widely used method for estimating parameters (Teimouri and Gupta, 2013). Recently, the Bayesian estimation approach has received great attention from many researchers. Among them is Al Omari and Ibrahim (2011). They considered the Bayesian survival estimator for Weibull distribution with censored data. Many authors, including Hossain and Zimmer (2003) and Pandey et al. (2011), did some comparative studies on the estimation of Weibull parameters using complete and censored samples, and Lye et al. (1993) determined the Bayes estimation for the extreme-value reliability function. More recently, Guure et al. (2012) examined the performance of MLE and Bayesian estimators for estimating the two-parameter Weibull failure time distribution. However, it has not been explored the use of the Bayesian technique for modelling wind data and analyzing wind power potential.

The present work aims to compare the two-parameter and three-parameter Weibull distributions to fit wind speed data more accurately at six locations in the Equatorial region, where wind speeds are generally lower. The aim is also to develop a novel approach using the Bayesian method in estimating Weibull parameters. The results from Bayesian technique will be compared with those of the traditional MLE method to determine a more accurate evaluation method of wind speed characteristics.

2. Wind speed data

Wind speed data from six different sites in the Equatorial region were used in the present work, as shown in Table 1.

Table 1: Locations of data collection sites

Sites	Location	Country	Measurement period	Topography
1. Tarawa	Latitude 1° 26' N Longitude 173° 00' E	Kiribati	September 2012 to September 2013	Flat
2. Pentecost	Latitude 15° 41' S Longitude 168° 11' E	Vanuatu	October 2012 to November 2013	Mountainous terrain
3. Rakiraki	Latitude 17° 22' S Longitude 178° 10' E	Fiji	February 2012 to October 2013	Flat
4. Kadavu	Latitude 19° 0' S Longitude 178° 15' E	Fiji	January 2018 to December 2018	Mountainous terrain
5. Rarotonga	Latitude 21° 15' S, Longitude 159° 45' W	Cook Islands	January 2016 to December 2018	Flat
6. Nuku'alofa	Latitude 21° 15' S Longitude 175° 15' W	Tonga	January 2016 to August 2019	Flat

For sites 1, 2 and 3, data were obtained from measurements using 34 m tall towers with the help of sensors described in Table 2. The NRG systems towers, named Integrated Renewable Energy Resource Assessment Systems (IRERAS), with a height of 34 m were used. NRG *SymphoniePlus3* was the data-logger used and it was connected to seven different sensors installed on the tower. The sensors measured wind speed, temperature, pressure, rainfall, solar insolation, humidity, and wind direction. The data were either collected from the SD card in person or sent via the GSM-based network to the data-bank located at the ICT centre of the University of South Pacific at the Laucala Campus, Fiji. The anemometers (serial numbers 179500189054-57, 179500189089-90) have an accuracy of 0.1 m/s and a range of 0.4 to 96 m/s. The wind vane is placed at 30 m AGL. The data were recorded in a time-series format in an RWD file which was later transferred to a Microsoft excel sheet. The wind speed data were measured continuously at an interval of 10 minutes with a cup anemometer at hub heights of 34 m and 20 m, respectively.
For sites 4, 5 and 6, satellite data were downloaded; land data from ERA5 were used in the present work (Ref: https://cds.climate.copernicus.eu/). ERA5 is the fifth generation ECMWF reanalysis for the global climate and weather for the past 4 to 7 decades. Reanalysis combines model data with observations from across the world into a globally complete and consistent dataset using the laws of physics. This principle, called data assimilation, is based on the method used by numerical weather prediction centres, where every so many hours (12 hours at ECMWF) a previous forecast is combined with newly available observations in an optimal way to produce a new best estimate of the state of the atmosphere, called analysis, from which an updated, improved forecast is issued. Reanalysis works in the same way, but at reduced resolution to allow for the provision of a dataset spanning back several decades. Reanalysis does not have the constraint of issuing timely forecasts, so there is more time to collect observations, and when going further back in time, to allow for the ingestion of improved versions of the original observations, which all benefit the quality of the reanalysis product.

Table 2: Specifications of the measurement sensors (Aukitino et al., 2017)

Parameter	**Sensor Type**	**Range**	**Accuracy**
Wind speed	NRG#40C anemometer	0.4-96.0 m/s	0.1 m/s
Wind direction	NRG 200P direction vane	0-360^{o}	N/A
Pressure	NRG BP-20 barometric pressure sensor	15 kPa - 115 kPa	1.5 kPa
Temperature	NRG 110S	-40 ^{o}C – 65 ^{o}C	1.11 ^{o}C

For the measured values, some uncertainties were taken into account such as calibration errors, the terrain of the site that was used, the dynamic over speeding, the error introduced due to wind shear and the inflow angle (Jain, 2016). The measurements in the present work were performed close to the shoreline at a flat terrain. The flow was in the horizontal plane, resulting in a lower uncertainty level. The calibration report for the anemometers used in the present work showed a maximum uncertainty of 0.6% for a wind speed range of 4-7 m/s, which reduced at higher wind speeds. The overall uncertainty in the estimation of wind speed is obtained by taking all the above uncertainties into account (Jain, 2016) and using the relation in equation (1):

$$\varepsilon = \sqrt{\sum_{i=1}^{N} \varepsilon_i^2} \tag{1}$$

where ε_i is each component of uncertainty and N is the number of components of uncertainty. The uncertainties were estimated at 95% confidence level. As per the IEC Standard IEC 61400-12-1 (IEC, 2017), the uncertainty in the measurements was estimated to be approximately 1.74%.

3. Weibull Distribution

Assessment of wind power energy at a site requires knowledge of the appropriate probability distribution of the site's wind speed, as the estimation of wind energy depends on its accuracy.

3.1 Two-parameter Weibull distribution

The two-parameter Weibull probability density functions (pdf) and the cumulative distribution function (cdf) for wind speed (U), respectively are given by

$$f(U)=\frac{k}{A}\left(\frac{U}{A}\right)^{k-1} e^{-\left(\frac{U}{A}\right)^{k}}; \quad U>0, k>0, A>0 \tag{2}$$

and

$$F\left(U\right)=1-e^{-\left(\frac{U}{A}\right)^{k}} \tag{3}$$

where $f\left(U\right)$ is the probability of observing the wind speed, k is the shape parameter and A is the scale parameter (m/s) of the distribution. The parameter k indicates the wind potential and what peak the distribution can reach. Its value ranges between 1 and 3. A lower k value signifies highly variable winds, while constant winds are characterized by a larger k. The parameter A denotes how windy the site under study is and it takes a value proportional to the mean wind speed (Manwell et al., 2010; Sukkiramathi and Seshaiah, 2020).

3.2 Three-parameter Weibull distribution

The three-parameter Weibull pdf and the cdf for wind speed, respectively are given by

$$f\left(U\right)=\frac{k}{A}\left(\frac{U-\theta}{A}\right)^{k-1} e^{-\left(\frac{U-\theta}{A}\right)^{k}}; \quad U>0, k>0, A>1, -\infty<\theta<\infty \tag{4}$$

and

$$F\left(U\right)=1-e^{-\left(\frac{U-\theta}{A}\right)^{k}} \tag{5}$$

where $f\left(U\right)$ is the probability of observing the wind speed, k is the shape parameter, A is the scale parameter (m/s), and θ is the shift or location parameter (m/s) of the distribution. If $\theta=0$, $f\left(U\right)$ and $F\left(U\right)$ become the PDF and CDF of a two-parameter Weibull distribution, respectively.

As the name implies, the shift parameter, θ, shifts the distribution along the abscissa. When $\theta=0$, the distribution starts at U = 0 or at the origin. Whereas, if θ > 0, the distribution starts at the location θ to the right of the origin. If $\theta < 0$, the distribution starts at the location θ to the left of the origin. For the distribution of wind speed, θ provides an estimate of the earliest time-to-start the wind (Tuller and Brett, 1984; Wais, 2017).

4. Methods of Estimating Weibull Parameters

To estimate the Weibull parameters, we propose a Bayesian approach and compare its performance with a popular frequentist approach, the maximum likelihood estimation (MLE) method.

4.1 The Maximum Likelihood Method

4.1.1 Two-parameter distribution

MLE is the most popular technique for deriving estimators (Casella and Berger, 2002; Aukitino et al., 2017; Chaurasiya et al., 2018). If $U_1,...,U_n$ are the wind speed values with the Weibull density function given in (2), the shape parameter (k) and scale parameter (A) are the values that maximize the likelihood function $L\left(k, A \middle| U_1,...,U_n\right) = \Pi_{i=1}^{n} f\left(U_i \middle| k, A\right)$. Then, solving $\partial \ln L / \partial k = 0$ and $\partial \ln L / \partial A = 0$ gives the equation of MLE of the scale parameter A as:

$$A = \frac{1}{n}\sum_{i=1}^{n} U_i^k \tag{6}$$

Finally, equation (7) is used for estimating the shape parameter (k) as:

$$\frac{1}{k} + \frac{1}{n}\sum_{i=1}^{n}\ln\left(U_i\right) - \frac{\sum_{i=1}^{n} U_i^k \ln\left(U_i\right)}{\sum_{i=1}^{n} U_i^k} = 0, \tag{7}$$

which may be solved to obtain the estimate of k using Newton-Raphson method or any other numerical procedure because equation (7) does not have a closed form solution. When k is obtained, the value of A is found from equation (6).

4.1.2 Three-parameter distribution

The likelihood function L for estimating parameters is given by $L\left(k, A, \theta \middle| U_1,...,U_n\right) = \Pi_{i=1}^{n} f\left(U_i \middle| k, A, \theta\right)$. Then, solving $\partial \ln L / \partial k = 0$, $\partial \ln L / \partial A = 0$ and $\partial \ln L / \partial \theta = 0$ gives the equations of MLE of the parameters as follow:

$$\frac{n}{k}\sum_{i=1}^{n}\log\left(\frac{U_i - \theta}{A}\right) - \sum_{i=1}^{n}\left(\frac{U_i - \theta}{A}\right)^k \log\left(\frac{U_i - \theta}{A}\right) = 0 \tag{8}$$

$$\frac{nk}{A} + \frac{k}{A}\sum_{i=1}^{n}\left(\frac{U_i - \theta}{A}\right)^k = 0 \tag{9}$$

$$-\left(k-1\right)\sum_{i=1}^{n}\frac{1}{U_i - \theta} - \frac{k}{A}\sum_{i=1}^{n}\left(\frac{U_i - \theta}{A}\right)^{k-1} = 0 \tag{10}$$

There is no closed form of the solution but the non-linear equations (8) - (10) may be solved by applying some optimization techniques such as Newton-Raphson method or other numerical procedures (Lawless, 2003; Teimouri and Gupta, 2013).

4.1.3 Evaluation of MLE methods

To determine the best model, we can compare the fit of the two MLE methods using different measures of goodness-of-fit (Luceño, 2008; Ramachandran and Tsokos, 2015; Cousineau and Allan, 2015). The most used criteria are:

Log-likelihood (log-like):

If a pdf $f_{\hat{\eta}}(U)$ fitted on the wind speed data and $\hat{\eta}$ is the estimated parameter of the distribution, then the log-likelihood for the goodness of fit is obtained by the following equation:

$$\text{log-like} = \log\left(\prod_{i=1}^{n} f_{\hat{\eta}}(U_i)\right) \tag{11}$$

where U_i is the ith observed wind speed and n is the number of observations in the dataset. A higher value of log-likelihood value indicates a better fitting of the model.

Akaike information criteria (AIC):

If k be the number of distribution parameters to estimate, the AIC is obtained by:

$$\text{AIC} = -2(\text{log-like}) + 2k \tag{12}$$

A lower value of AIC indicates that the model fits the data better. Compared to the log-likelihood, this criterion takes into consideration the parsimony of the model as it includes a penalty term that increases the number of parameters.

Bayesian information criteria (BIC):

This criterion is obtained by:

$$\text{BIC} = -2(\text{log-like}) + k\log(n) \tag{13}$$

Similar to AIC, a lower value of BIC indicates that the model fits the data better. However, BIC provides a stronger penalty than AIC for additional parameters.

Kolmogorov-Smirnov (KS):

The Kolmogorov-Smirnov (KS) is also used to test the adequacy of a given theoretical distribution for a given set of wind speed data. The KS test computes the maximum difference between the predicted and observed distribution, and the test statistics D is given by:

$$\text{D} = \max_{1\le i\le n}\left|F_i - \hat{F}_i\right| \tag{14}$$

where $\hat{F}_i$ is the ith predicted cumulative probability from the theoretical cdf and F_i is the empirical probability of the ith observed wind speed.

Anderson-Darling (AD):
For a finite data sample, the Anderson-Darling (AD) test statistics A^2 is defined by:

$$A^2 = -n - s\,, \tag{15}$$

where $s = \sum_{i=1}^{n} \frac{2i-1}{n}\left[\log\left(\hat{F}_i\right) + \log\left(1-\hat{F}_{n-i+1}\right)\right]$.

4.2 The Bayesian Method
A classical frequentist approach such as MLE has certain drawbacks. Most of its properties hold only for large sample size and it requires a symmetric form of sampling distribution. The Bayesian approach, however, is free from such limitations. Moreover, Bayesian simulation tools provide an exact method of inference even if sample size is very small. Thus, in empirical situations where the sample size may be small, Bayesian methods seem to be more suitable over frequentist methods if prior information about the parameters is available.

In this paper, we propose a Bayesian inference approach for modeling of wind speed data. In the Bayesian paradigm, data and prior are combined together to make an inference about the parameters of interest.

The most influential contribution of the Bayesian approach is its modification of the likelihood function into a posterior - a valid probability distribution defined by the classic Bayes' rule. The posterior distribution of wind speed is expressed as:

$$p\left(\eta|U\right) = \frac{p\left(U|\eta\right)p\left(\eta\right)}{p\left(U\right)} \tag{16}$$

where $p\left(\eta|U\right)$ is the posterior distribution of wind speed, $p\left(\eta\right)$ is the prior distribution of unknown parameters $\eta = \left(k, A, \theta\right)$, $p\left(U|\eta\right)$ is the likelihood of wind speed data and $p\left(U\right) = \int p\left(U|\eta\right)p\left(\eta\right)d\eta$. The denominator of equation (16), $p\left(U\right)$, normalizes the posterior distribution, $p\left(\eta|U\right)$. Since it is independent of U, it is often convenient to write the posterior distribution as:

$$p\left(\eta|U\right) \alpha\ p\left(U|\eta\right)p\left(\eta\right), \tag{17}$$

that is, the posterior distribution of the parameters is proportional to the likelihood function times the prior distribution of parameters. While, fitting wind speed data, we use a noninformative uniform prior distribution as we have very little prior knowledge about its model parameters. In

Bayesian computations, a sample of the joint posterior distribution is obtained by using Gibbs sampler to simulate a sample from a Markov Chain Monte Carlo (MCMC). Then, we can calculate the desired values of the posterior.

In this paper, we use the software `JAGS` to fit the model. The R package `R2jags` is used to summarize the posterior inference, which is discussed in more detail in Section 4.2.1.

4.2.1 Bayesian Fitting of Weibull distribution with JAGS

`JAGS`, an acronyms for "Just Another Gibbs Sampler" (Plummer, 2003), accepts a model string written in an R-like syntax that compiles and generates MCMC samples from the model using Gibbs sampling. It is an open-source software written in C++ using GNU compilers and packaging tools - freely available at http://mcmc-jags.sourceforge.net/.

R packages such as `R2jags` or `rjags` allow running `JAGS` models within R on Windows machines for the summarization of posterior inference.

The Bayesian fit of wind speed using `JAGS` – including model specification for a two-parameter Weibull distribution – and code for creating MCMC data are briefly presented as:

Model specification:

```
cat("
    model{
    # Likelihood
    for (i in 1:length(U)){
    p[i] <- dweib(U[i],shape, lambda);
    observed[i] ~ dbern(p[i]);
    }

    # Priors
    shape ~ dunif(0,4)  # After a series of trial and error guesses
    scale ~ dunif(0,100)
    lambda <- pow(1/(scale), shape)
    }
    ", file="weibull_model_2p.txt")
```

MCMC simulation using Gibbs sampler:

In MCMC simulation, we run the Gibbs sampler with the `JAGS` model for 10,000 iterations. We usually discard the first 1,000 iterations, identifying them as the length of a burn-in period, i.e., the point after which a Gibbs sampler is supposed to attain convergence to estimate the Weibull parameter.

In this research, we use `R2jags` (Su and Yajima, 2020), an R package that allows fitting `JAGS` models from within R. The `jags` function and its arguments used for fitting two-parameter Weibull distribution are as follows:

```
jags.fit <- jags(data, inits, parameters.to.save, n.iter=10000,
model.file="weibull_model_2p.txt",n.chains = 3, n.burnin = 1000,
n.thin=5)
```

Similarly, as discussed above, we can perform the Bayesian fit of wind speed with the three-parameter Weibull distribution using the following JAGS model:

```
cat("
    model{
    # Likelihood
    for (i in 1:length(U)){
    p[i] <- dweib(U[i] - shift ,shape, lambda) / 1000
    observed[i] ~ dbern(p[i]);
    }

    # Priors
    shape ~ dunif(0,4)
    scale ~ dunif(0,31)
    shift ~ dunif(-1, 1)
    lambda <- pow(1/(scale), shape)
    }
    ", file="weibull_model_3p.txt")
```

4.2.2 Evaluation of Bayesian Models

The standard likelihood, AIC and BIC statistics as discussed in Section 4.1.3 are not relevant while evaluating Bayesian methods such as MCMC. Instead, Spiegelhalter et al. (2002) suggests that the Deviance Information Criterion (DIC) be used to compare models. The DIC is a generalization of AIC that is based on Deviance statistics:

$$D(\eta) = -2\log f(U/\eta) + 2\log h(U), \tag{18}$$

where $h(U)$ is some standardizing function of the data. The DIC is then defined as:

$$DIC = \bar{D} + p_D , \tag{19}$$

where $\bar{D} = E_{\eta|U}(D)$ is the posterior expectation of deviance and p_D is effective number of parameters that captures the complexity of a model. A smaller value of DIC indicates a better-fitting model.

5. Estimate of wind power and energy

When the wind speed (U) of a site and the frequency distribution $f(U)$ are known, the available wind power and wind energy can be estimated. Let ρ be the density of air, D is the turbine rotor diameter and $A_R = \pi D^2/4$ is the rotor cross sectional area, then the probability of available wind power for a given velocity U is obtained by

$$p(U) = \frac{1}{2}\rho A_R U^3; \quad U > 0. \tag{20}$$

Then, the expected wind power (P) is estimated by

$$\text{Wind power, } P = \int_0^\infty P(U).f(U)dU. \tag{21}$$

Substituting (20), (2) and (4) in (21), the expected wind power for two-parameter and three-parameter Weibull distributions respectively, are determined by

$$P_{2P} = \frac{1}{2}\rho A_R \frac{k}{A^k}\int_0^\infty U^{k+2} e^{-\left(\frac{U}{A}\right)^k} dU \tag{22}$$

and

$$P_{3P} = \frac{1}{2}\rho A_R \frac{k}{A^k}\int_0^\infty U^3.(U-\theta)^{k-1} e^{-\left(\frac{U-\theta}{A}\right)^k} dU \tag{23}$$

If $t_{Year} = 8760$ is the total number of hours in a year, the expected wind energy for two-parameter and three-parameter Weibull distributions respectively, are determined by

$$E_{2P} = P_{2P} \cdot t_{Year} \tag{24}$$

and

$$E_{3P} = P_{3P} \cdot t_{Year} \tag{25}$$

6. Analyzing performance of different estimators

The efficiency and performance of MLE and Bayesian methods for estimating two and three parameter Weibull distributions were determined using different goodness of fit and error measures such as coefficient of determination (R^2), root mean square error (RMSE), coefficient of efficiency (COE), mean absolute error (MAE) and the mean absolute percentage error (MAPE). Arithmetically, these are computed as follows (Azad et al., 2014; Kidmo et al., 2015; Aukitino et al., 2017):

Coefficient of determination (R^2):

It is a statistical measure that gives some information about the goodness of fit of a model, that is, how much the variance of the observed data is explained by the fitted model. It is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(U_i - \hat{U}_i)^2}{\sum_{i=1}^{n}(U_i - \bar{U})^2}, \tag{26}$$

where n is the number of observations, U_i is the ith actual data, $\hat{U}_i$ is the ith predicted data with the Weibull distribution, $\bar{U}$ is the mean of actual data. A higher R^2 value indicates a better fit and $R^2 = 1$ indicates that the regression predictions perfectly fit the data.

Root mean square error (RMSE):
It determines the deviation between the observed and predicted values of wind speed and obtained by

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^{n} \left(U_i - \hat{U}_i \right)^2 \right]^{\frac{1}{2}} \tag{27}$$

A smaller RMSE value normally indicates accurate modeling. The calculated RMSE value approaches zero as the deviation between the observed and predicted values becomes smaller (Indhumathy et al., 2014).

Coefficient of efficiency (COE):
It measures the ratio of deviation of predicted values and actual values from the average of wind speeds. A higher COE value indicates good fitting for the data described. It is expressed as:

$$COE = \frac{\sum_{i=1}^{n} \left(\hat{U}_i - \bar{U} \right)^2}{\sum_{i=1}^{n} \left(U_i - \bar{U} \right)^2}. \tag{28}$$

Mean absolute error (MAE):
The mean absolute error is a measure of the absolute difference between predicted and actual values. A smaller value of MAE indicates improved accuracy. The MAE is mathematically expressed as:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| \hat{U}_i - U_i \right|. \tag{29}$$

Mean absolute percentage error (MAPE):
It is a comparative measure, indicating the error as a percentage of the actual data which helps accurately predict the forecasting method. Like MAE, a lower value of MAPE indicates better accuracy. It is mathematically expressed as:

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{\hat{U}_i - U_i}{U_i} \right|. \tag{30}$$

7. Results

In this section, we present the results of the fitting of two-parameter (2-p) and three-parameter (3-p) Weibull distributions, Further, we present results for the application of MLE and the proposed Bayesian approach for estimating the parameters, as discussed in Sections 3 and 4. To accomplish this, we experiment with wind speed data at six different sites as mentioned in Section 2. Table 3 provides wind speed distributions at these sites. The table shows that the range of speed varies at different sites. The lowest range of wind speed was observed at Site 1 (0-19 m/s) and the highest range was found at Site 3 (0-34 m/s). Some sites tend to have more null speed (0-1 m/s).

Table 3: Frequency Distributions of wind speed at different sites

Wind speed *U*	Site 1 Frequency	Site 2 Frequency	Site 3 Frequency	Site 4 Frequency	Site 5 Frequency	Site 6 Frequency
(0,1]	889	1063	5882	62	209	251
(1,2]	2197	1998	4246	219	773	799
(2,3]	5224	3579	4951	367	1498	1543
(3,4]	7006	5739	5845	641	1928	2390
(4,5]	8721	8077	7134	803	2388	3330
(5,6]	8783	10332	8716	949	2878	4260
(6,7]	7554	9570	10771	1026	2965	4590
(7,8]	5591	7094	11215	1052	2937	3915
(8,9]	3381	4810	9363	851	2833	3502
(9,10]	1706	2701	6496	882	2349	2812
(10,11]	604	1255	4226	723	2016	2023
(11,12]	236	482	2366	437	1411	1205
(12,13]	125	119	1177	311	892	715
(13,14]	42	26	588	215	661	278
(14,15]	11	9	247	99	405	193
(15,16]	3	8	107	54	91	114
(16,17]	1	10	60	7	42	97
(17,18]	4	3	53	9	17	43
(18,19]	2	3	38	6	6	30
(19,20]	0	1	13	6	2	16
(20,21]	0	0	5	7	3	9
(21,22]	0	0	4	15	0	8
(22,23]	0	0	3	8	0	11
(23,24]	0	0	6	2	0	1
(24,25]	0	0	5	5	0	1
(25,26]	0	0	4	2	0	0
(26,27]	0	0	4	0	0	0
(27,28]	0	0	3	1	0	0
(28,29]	0	0	5	0	0	0
(29,30]	0	0	10	0	0	0
Above 30	0	0	4	1	0	0
Total frequency	52080	56879	83554	8760	26304	32136

We fit the 2-p and 3-p Weibull distributions to the recorded wind speed data and estimate the parameters in the distributions using MLE and the Bayesian method. In MLE, the goodness of fit with 2-p and 3-p Weibull distributions are evaluated using the statistical measures AIC, BIC, AD, KS and log-like. Table 4 presents the estimated values of the parameters of 2-p Weibull and 3-p Weibull distributions and various statistical measures determined by the MLE at Site 1.

Table 4: Estimated values of parameters and statistical measures by MLE at Site 1

Distribution	k	A	θ	AIC	BIC	AD	KS	log-like
2-p Weibull	2.564486	6.019568	-	230352.8	230370.6	23.95999	0.021077	-115174.4
3-p Weibull	2.777792	6.438856	-0.3806114	230075.7	230102.3	9.984823	0.013953	-115034.9

In Bayesian estimates, we use uniform prior distributions of parameters to fit wind data. Firstly, we obtain a sample of the joint posterior distribution by simulating a sample from MCMC methods using a Gibbs sampler. Then, we obtain the posterior estimates of parameters by performing 10000 Gibbs sampler iterations and using 1000 burn-in period with five thinning intervals and 3 chains with a sample size of 1800 per chain. Finally, the DIC is used to evaluate the fitting of the two Weibull distributions. Table 5 presents the estimated values of the parameters of two distributions and the summary statistics including the DIC values for Site 1.

Table 5: Estimated values of parameters and summary statistics by Bayesian at Site 1

Distribution	Parameter	Mean	SD	2.5%	97.5%	DIC
2-p Weibull	k	2.564487	0.008762	2.547292	2.581560	949864.6
	A	6.019417	0.010876	5.998086	6.041268	
3-p Weibull	k	2.778195	0.018610	2.742305	2.814923	949587.6
	A	6.439670	0.033898	6.373026	6.507357	
	θ	-0.381144	0.029685	-0.439400	-0.323193	

Figures 1 and 2 show the trace and posterior density plots for 2-p Weibull and 3-p Weibull, respectively. The plots indicate convergence of the Bayesian estimates.

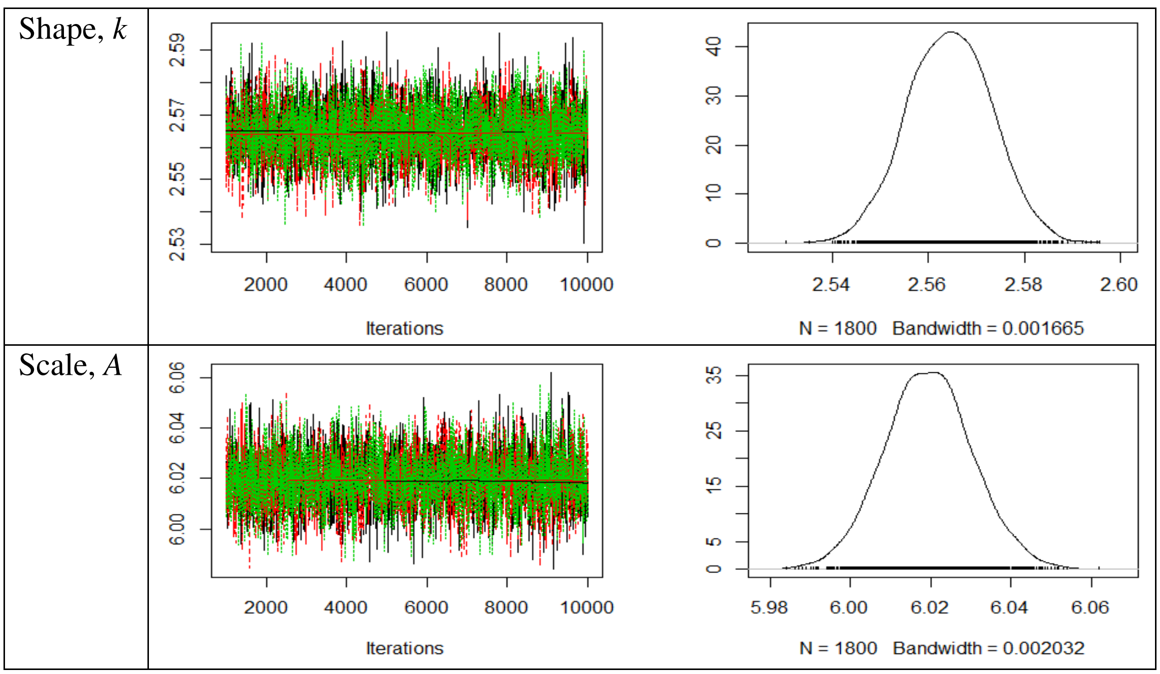


Figure 1: Trace and posterior density plots for Site 1 (2p-Weibull)

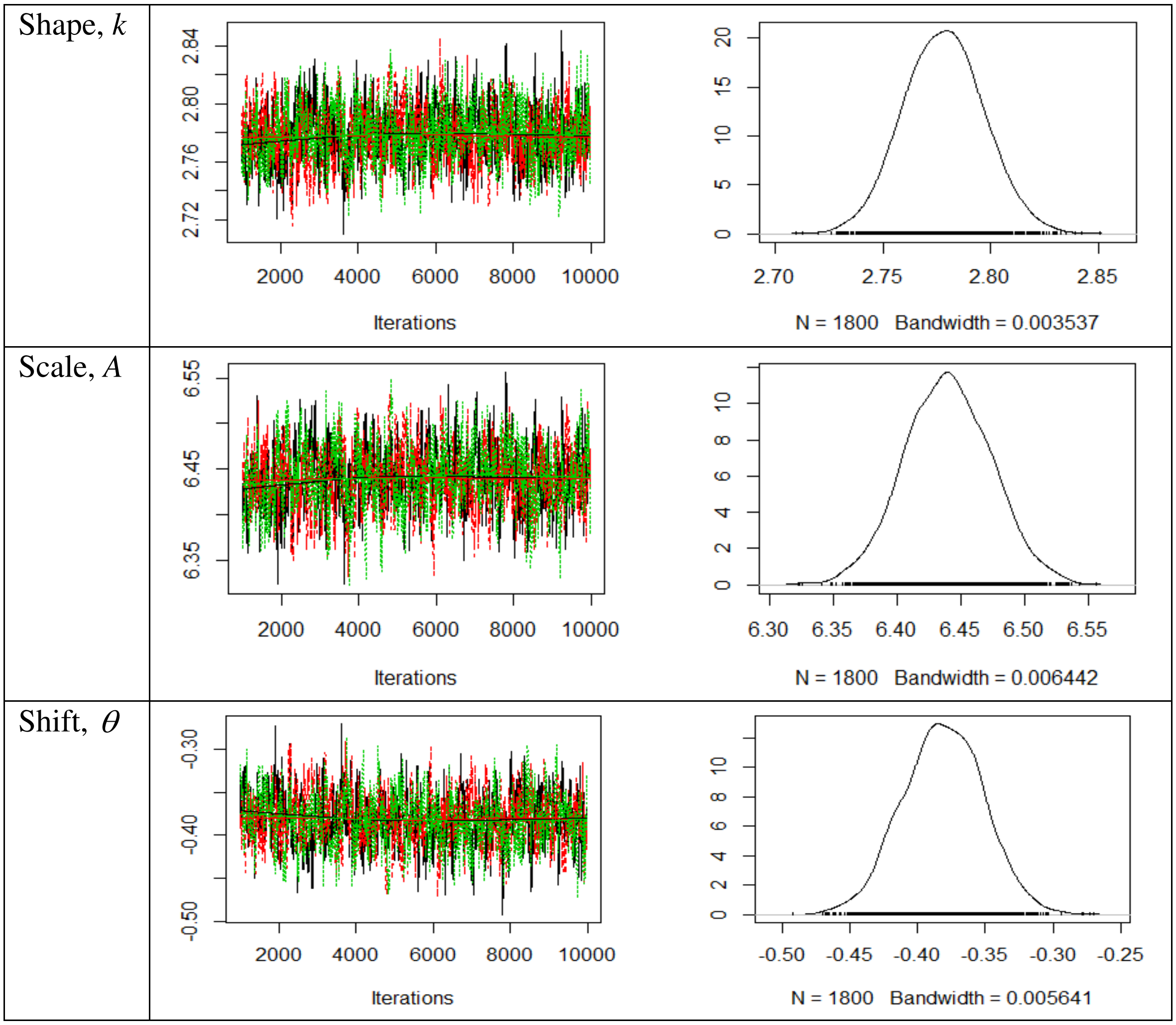


Figure 2: Trace and posterior density plots for Site 1 (3p-Weibull)

The goodness-of-fit criteria and summary statistics presented in Tables 4 and 5 indicate that the **3-p Weibull** distribution fits better than the 2-p Weibull distribution for wind speed at Site 1 as all the goodness-of-fit measures (AIC, BIC, AD, KS and log-like) are smaller in MLE estimate and the DIC is also smaller in Bayesian estimate. Moreover, as shown in Table 5 and Figure 2, the marginal posterior mean of shift parameter θ is -0.381144 and a 95% credible region for the parameter is (-0.439400, -0.323193), which indicates that the value of its shift parameter is non-zero (i.e. $\theta \neq 0$).

Similarly, Tables 6-14 present the parameter estimates of 2p-Weibull and 3p-Weibull distributions along with various statistical measures determined by MLE and Bayesian at Sites 2-6.

Table 6: Estimated values of parameters and statistical measures by MLE at Site 2

Distribution	k	A	θ	AIC	BIC	AD	KS	log-like
2 parameter	2.735978	6.535565	-	256112.4	256130.3	101.4772	0.038265	-128054.2
3 parameter	3.233794	7.532297	-0.9220459	255252.9	255279.8	32.0876	0.025796	-127623.5

Table 7: Estimated values of parameters and summary statistics by Bayesian at Site 2

Distribution	Parameter	Mean	SD	2.5%	97.5%	DIC
2 parameter	k	2.736076	0.008940	2.718410	2.753285	256112.2
	A	6.535458	0.010427	6.515002	6.555774	
3 parameter	k	3.231943	0.023096	3.184961	3.273617	255252.4
	A	7.528328	0.044253	7.438378	7.607035	
	θ	-0.918508	0.040458	-0.989749	-0.835880	

Table 8: Estimated values of parameters and statistical measures by MLE at Site 3

Distribution	k	A	θ	AIC	BIC	AD	KS	log-like
2 parameter	1.948074	7.017269	-	433669.2	433687.8	1414.721	0.092759	-216832.6
3 parameter	2.636074	8.804806	-1.546749	429666.9	429694.9	478.7482	0.058134	-214830.5

Table 9: Estimated values of parameters and summary statistics by Bayesian at Site 3

Distribution	Parameter	Mean	SD	2.5%	97.5%	DIC
2 parameter	k	1.947640	0.005556	1.936812	1.958478	433669.2
	A	7.017671	0.013022	6.992180	7.042895	
3 parameter	k	2.635850	0.016118	2.604492	2.667974	429666.8
	A	8.804653	0.044092	8.718409	8.892835	
	θ	-1.546999	0.038766	-1.624995	-1.472661	

Table 10: Estimated values of parameters and statistical measures by MLE at Site 4

Distribution	k	A	θ	AIC	BIC	AD	KS	log-like
2-p Weibull	2.38554	8.445309	-	45434.68	45448.83	6.759304	0.0197348	-22715.34
3-p Weibull	2.401323	8.493057	-0.042536	45435.82	45457.05	6.577964	0.0197392	-22714.91

Table 11: Estimated values of parameters and summary statistics by Bayesian at Site 4

Distribution	Parameter	Mean	SD	2.5%	97.5%	DIC
2-p Weibull	k	2.385596	0.019486	2.347390	2.424420	166458.7
	A	8.446062	0.039889	8.367275	8.525195	
3-p Weibull	k	2.405809	0.026990	2.356128	2.461114	166459.8
	A	8.507469	0.070222	8.377540	8.660753	
	θ	-0.055714	0.051466	**-0.169674**	**0.031573**	

Table 12: Estimated values of parameters and statistical measures by MLE at Site 5

Distribution	k	A	θ	AIC	BIC	AD	KS	log-like
2-p Weibull	2.438405	8.233451	-	135020.5	135036.9	16.38031	0.016043	-67508.26
3-p Weibull	2.569015	8.594684	-0.322758	134971.8	134996.4	10.39361	0.014675	-67482.91

Table 13: Estimated values of parameters and summary statistics by Bayesian at Site 5

Distribution	Parameter	Mean	SD	2.5%	97.5%	DIC
2-p Weibull	k	2.438513	0.011980	2.415006	2.462002	498423.8
	A	8.233617	0.022116	8.190394	8.276777	
3-p Weibull	k	2.405809	0.026990	2.356128	2.461114	498375.1
	A	8.507469	0.070222	8.377540	8.660753	
	θ	-0.324413	0.057377	-0.442912	-0.217255	

Table 14: Estimated values of parameters and statistical measures by MLE at Site 6

Distribution	k	A	θ	AIC	BIC	AD	KS	log-like
2-p Weibull	2.503177	7.843472	-	159616.8	159633.6	22.04708	0.019279	-79806.41
3-p Weibull	2.522584	7.896873	-0.047855	159615.8	159641.0	21.85474	0.018862	-79804.91

Table 15: Estimated values of parameters and summary statistics by Bayesian at Site 6

Distribution	Parameter	Mean	SD	2.5%	97.5%	DIC
2-p Weibull	k	2.502775	0.010741	2.482126	2.524369	603592.3
	A	7.843795	0.018777	7.806656	7.880416	
3-p Weibull	k	2.523740	0.016243	2.492982	2.555752	603591.3
	A	7.900740	0.038140	7.830501	7.978555	
	θ	-0.051032	0.030016	-0.113036	0.004279	

The trace and posterior density plots obtained in Bayesian simulations for 2p-Weibull and 3p-Weibull are presented in Figures 3-12, respectively, for sites 2-6. The plots indicate convergence of Bayesian estimates.

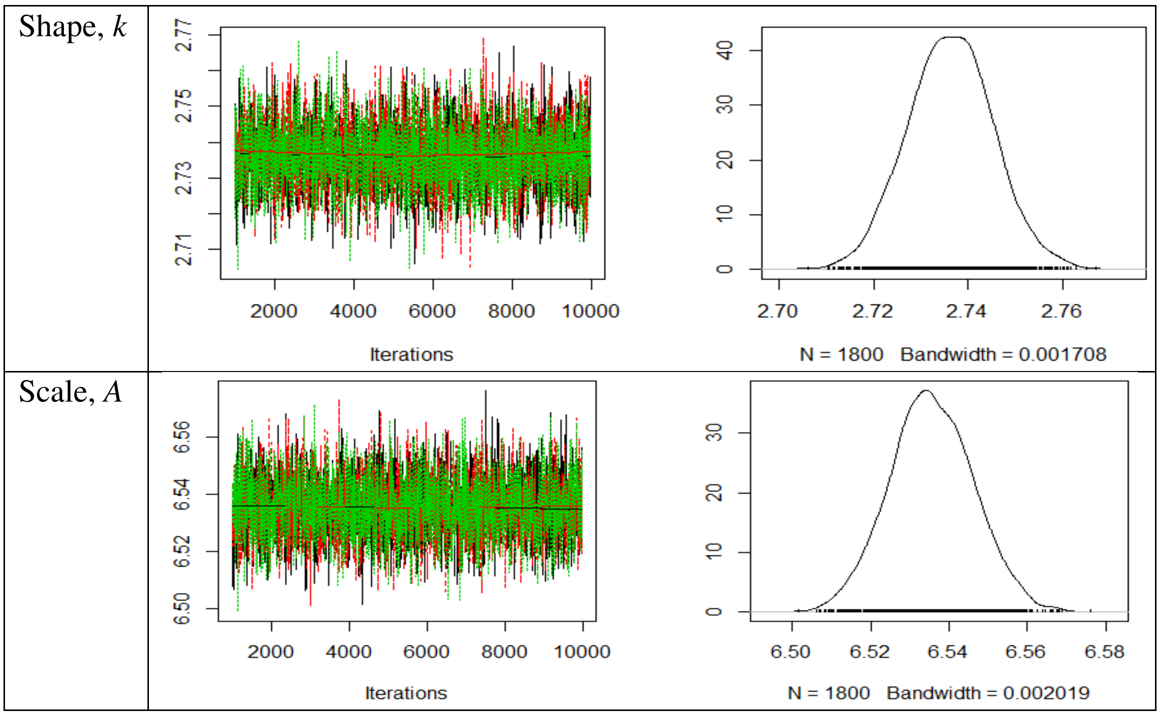
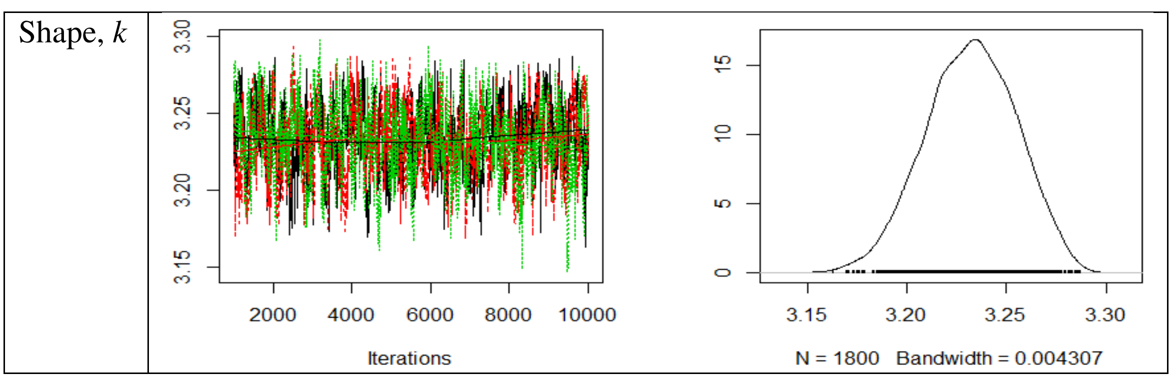


Figure 3: Trace and posterior density plots for Site 2 (2p-Weibull)



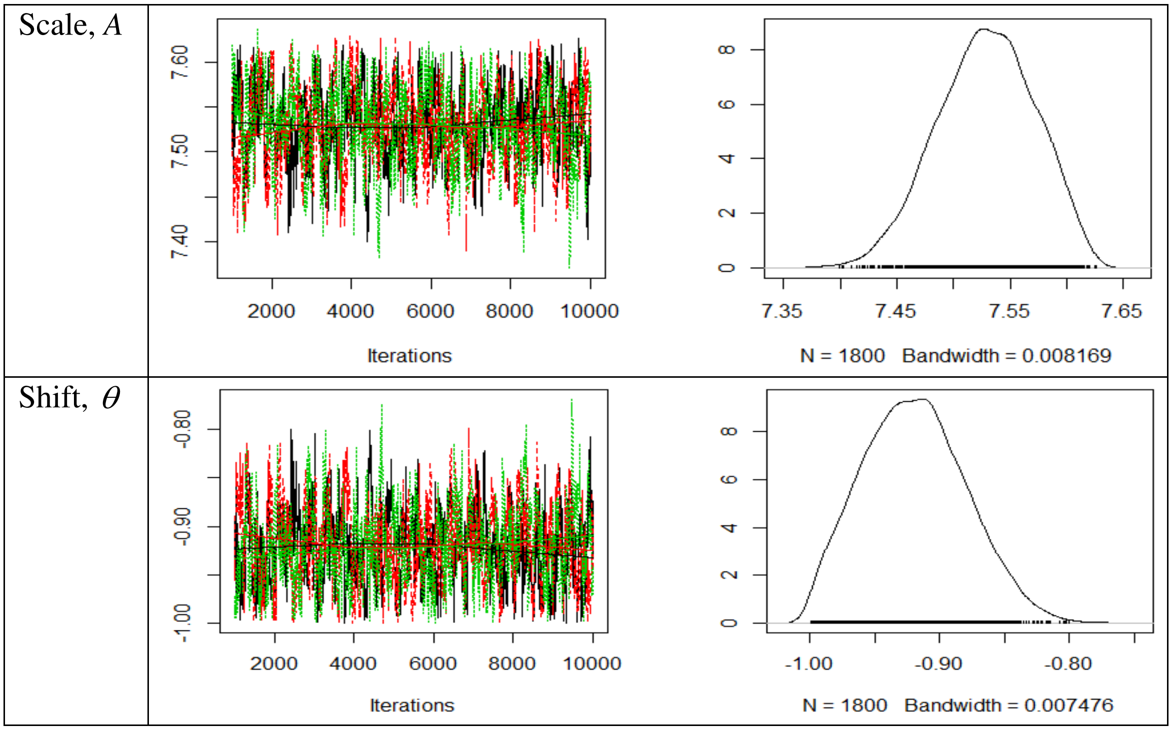


Figure 4: Trace and posterior density plots for Site 2 (3p-Weibull)

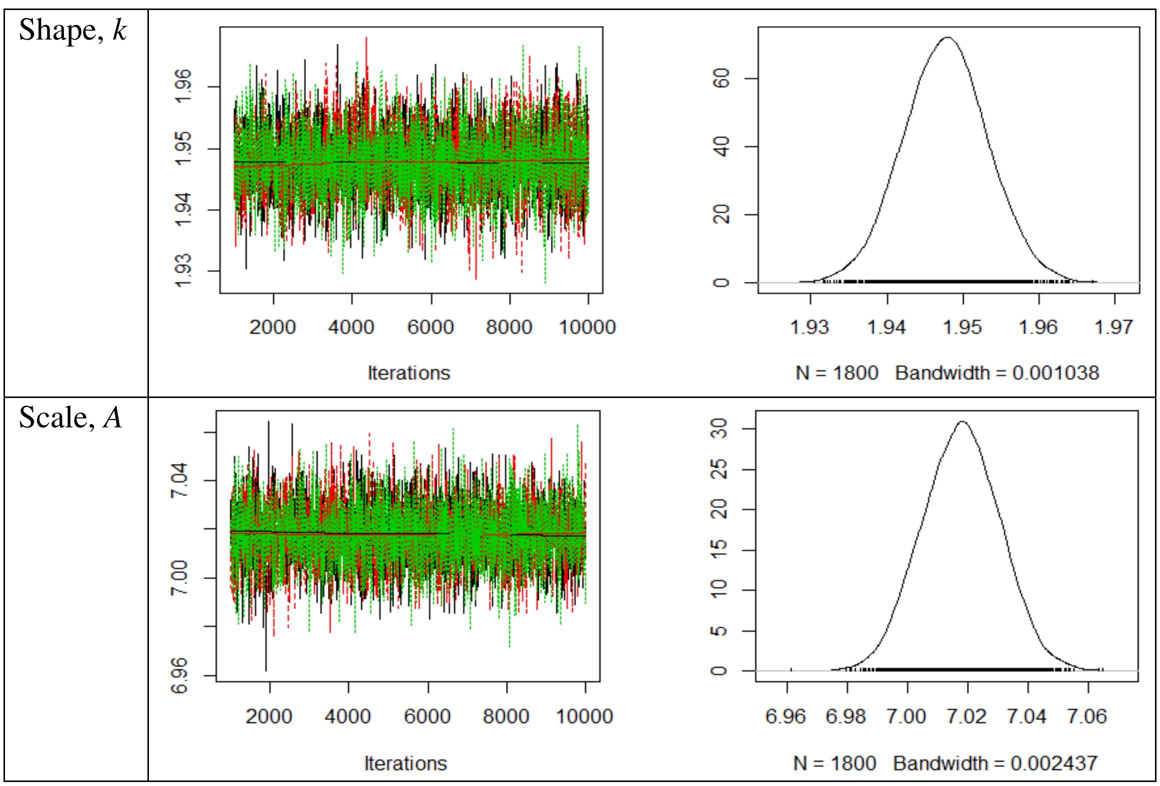
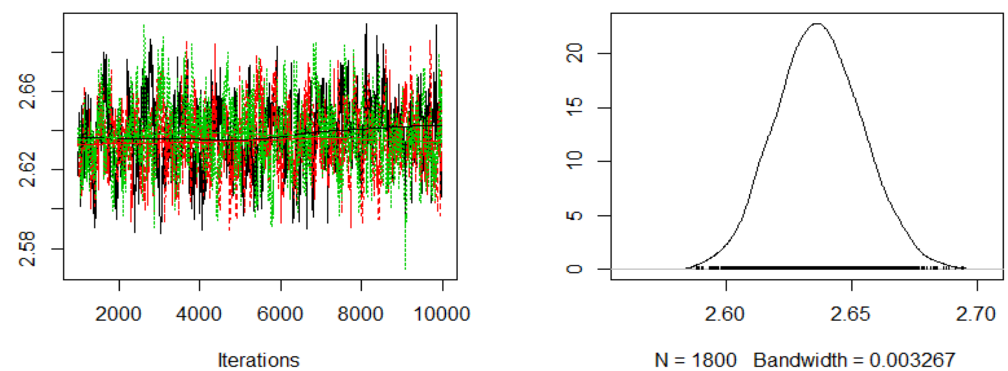
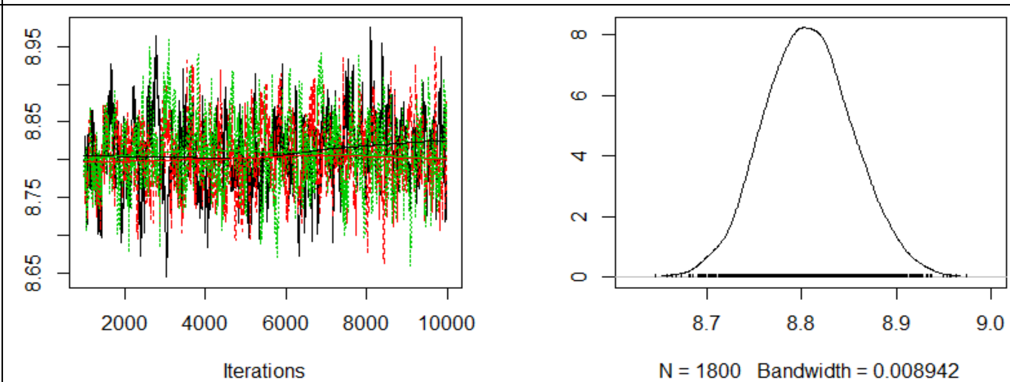


Figure 5: Trace and posterior density plots for Site 3 (2p-Weibull)

Shape, k



Scale, A



Shift, θ

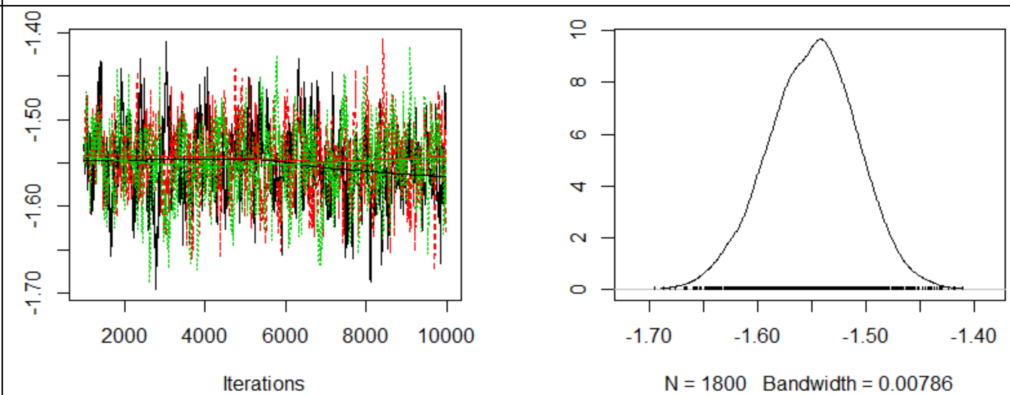
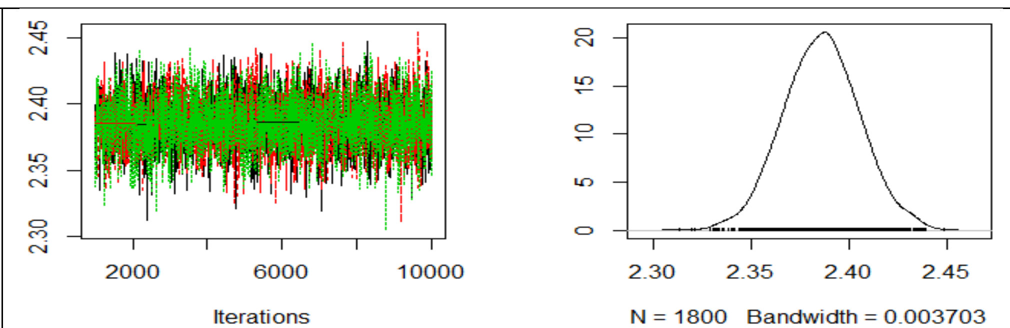


Figure 6: Trace and posterior density plots for Site 3 (3p-Weibull)

Shape, k



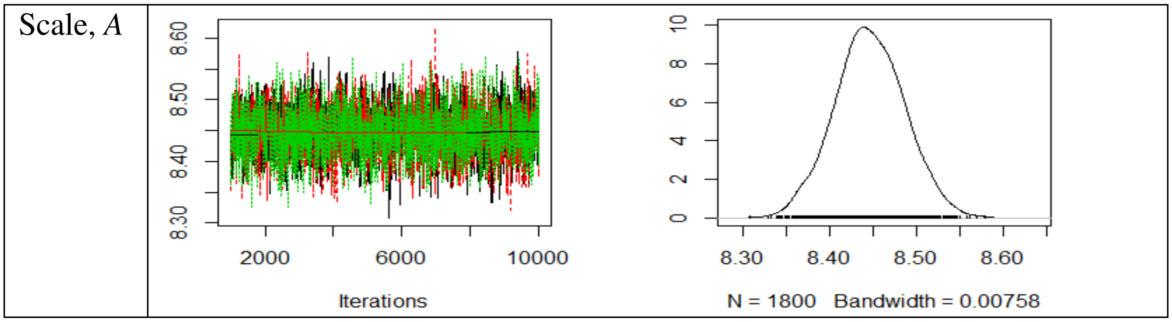


Figure 7: Trace and posterior density plots for Site 4 (2p-Weibull)

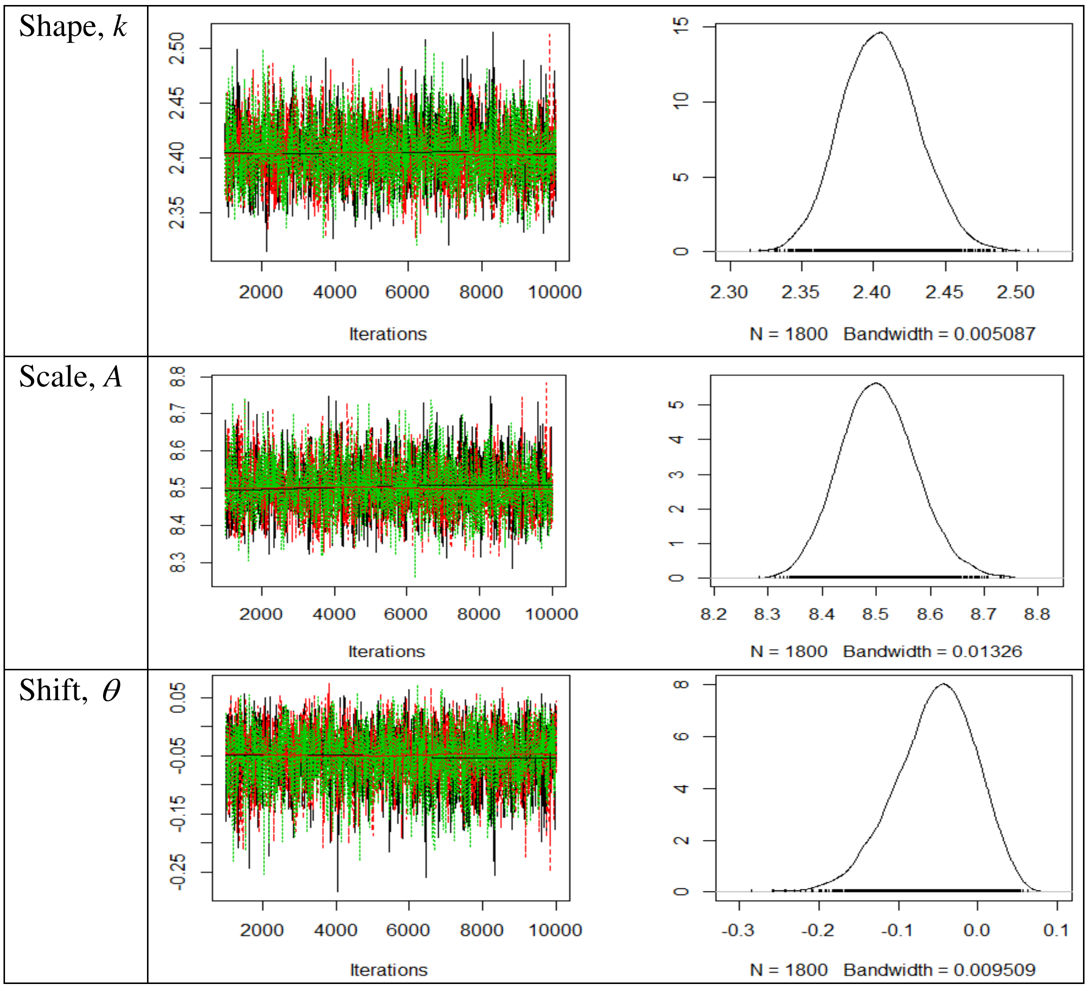


Figure 8: Trace and posterior density plots for Site 4 (3p-Weibull)

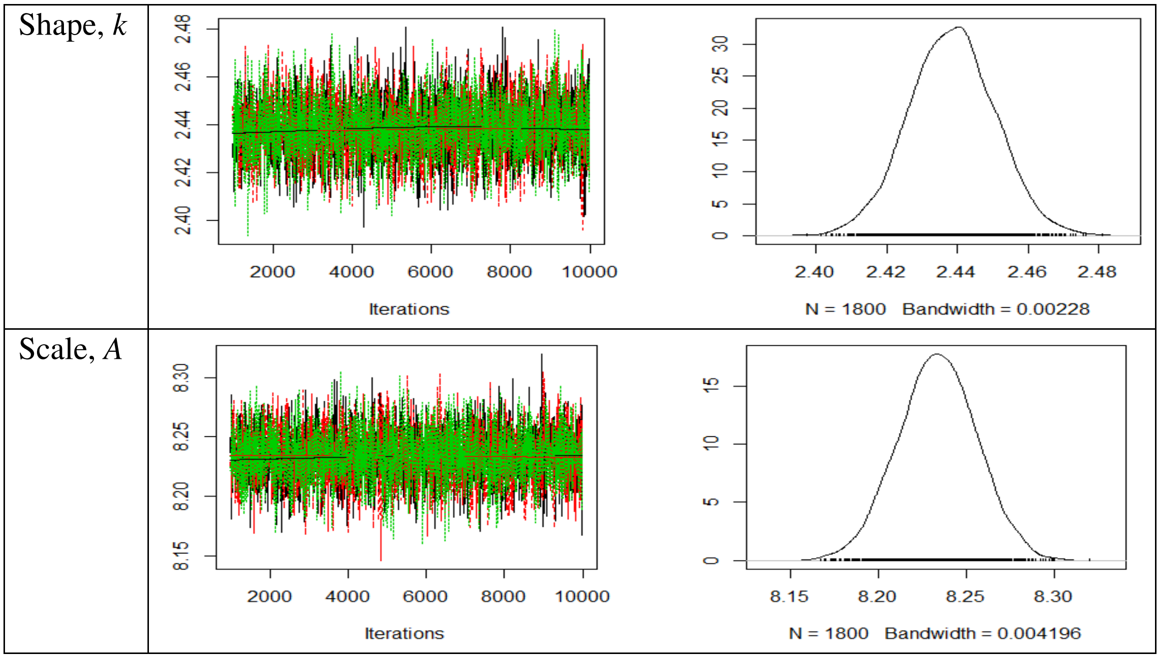


Figure 9: Trace and posterior density plots for Site 5 (2p-Weibull)

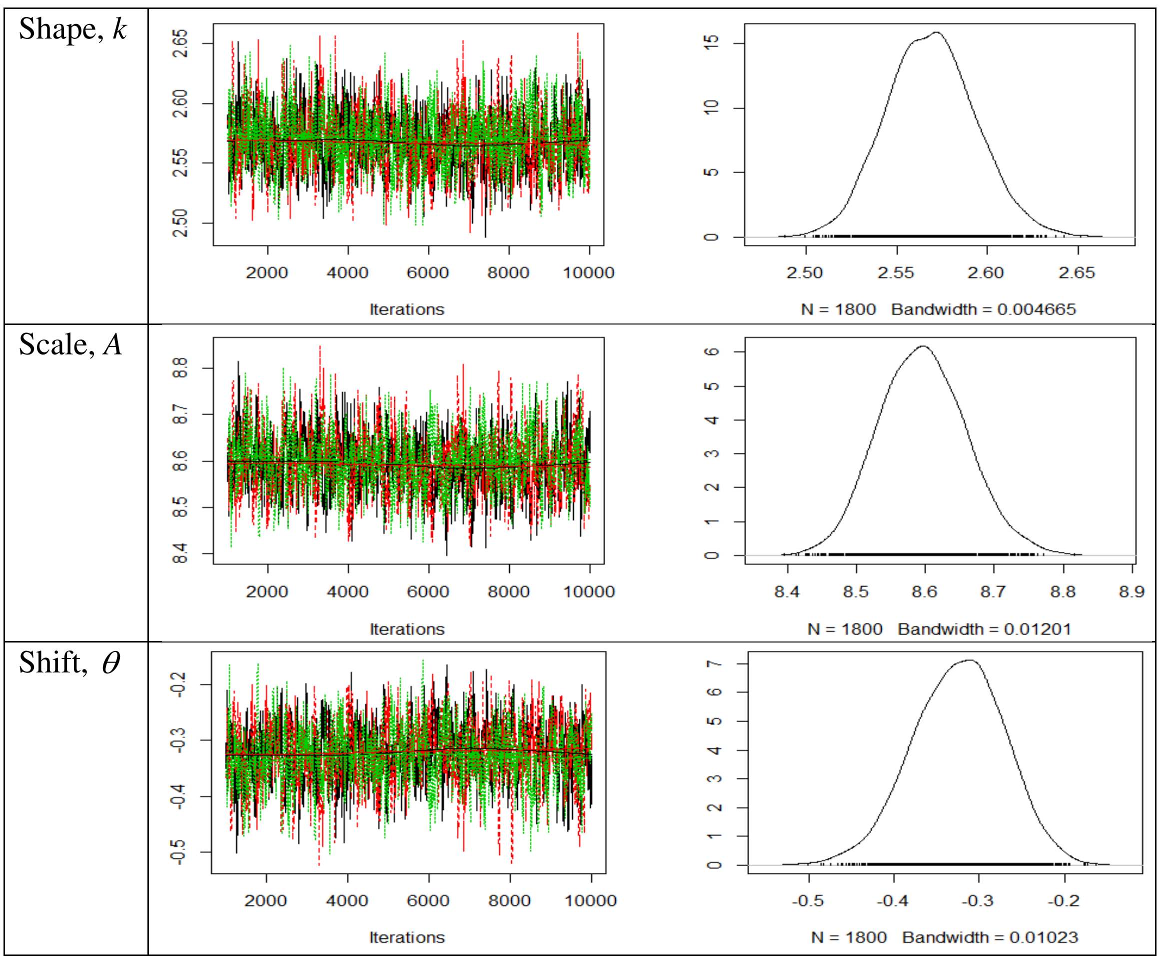


Figure 10: Trace and posterior density plots for Site 5 (3p-Weibull)

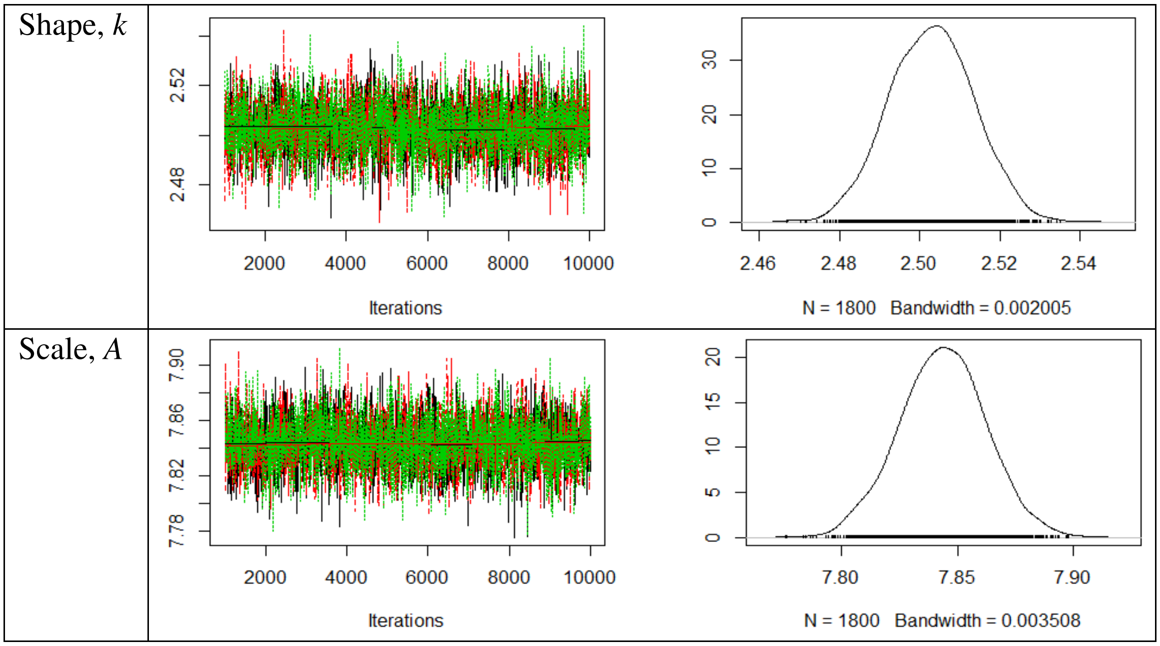


Figure 11: Trace and posterior density plots for Site 6 (2p-Weibull)

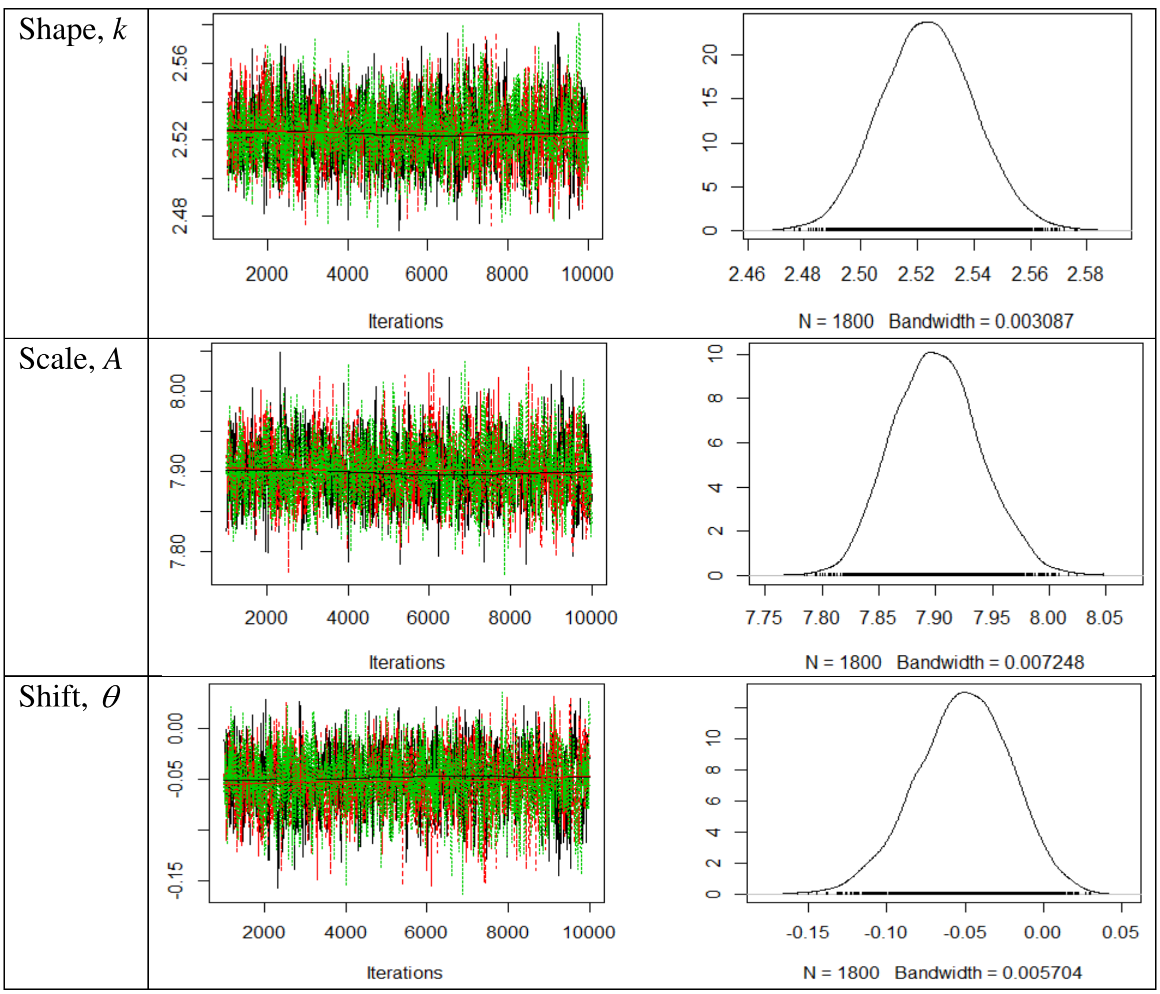


Figure 12: Trace and posterior density plots for Site 6 (3p-Weibull)

The results presented in Tables 6-14 allow us to make the following inferences on the fit of 2-p and 3-p Weibull distributions for the Sites 2-6:

- **Site 2:** The results in Tables 6 and 7 reveal that **3-p Weibull** fits better than the 2-p Weibull. All five measures in MLE as well as the DIC in Bayesian are smaller for the 3-p Weibull distribution. Also, the 95% credible region (-0.989749, -0.835880) of 3-p Weibull indicates that the value of its shift parameter is non-zero (i.e. $\theta \neq 0$).
- **Site 3 and 5:** Like Site 2, the results in Tables 8-9 and 12-13 reveal that **3-p Weibull** fits the recorded wind speed data better than the 2-p Weibull distribution. Moreover, the 95% credible region of 3-p Weibull clearly shows that the shift parameter is $\theta \neq 0$.
- **Site 4:** The results in Tables 10 and 11 indicate that the **2-p Weibull** distribution fits better than the 3-p Weibull distribution at this site as the AIC, BIC and KS are smaller in MLE estimate and the DIC is smaller in Bayesian estimate. Moreover, the marginal posterior mean of θ is -0.055714 and a 95% credible region for the parameter is (-0.169674, 0.031573), which indicates that the value of its shift parameter is zero (i.e. $\theta = 0$).
- **Site 6:** The results in Tables 14 and 15 show a lack of significant difference between the two Weibull distributions while fitting wind speeds, as all MLE and Bayesian estimates are similar in numerical value. Also, the 95% credible region indicates that $\theta = 0$. Thus, **2-p Weibull** may be a better distribution.

8. Discussion

In Section 7, we presented the results for the goodness of fit for wind speed distributions at six different sites. Results showed that the 2-p Weibull distribution was a better fit for wind speeds at Sites 4 and 6. However, the 3-p Weibull distribution was a better fit for wind speeds at all other sites. We performed further investigations to explain the difference between the performance of the two distributions.

By looking into the percentage of lowest wind speed (0-1 m/s) presented in Table 16, the results clearly show that the wind distribution of the sites (Sites 1, 2, 3 and 5) that have high percentage (0.79% - 7.04%) of lower (or closer to null) wind speed perfectly fits 3-p Weibull distribution. The shift parameter is also found in Bayesian simulation significantly $\theta \neq 0$ for these sites. On the other hand, the 2-p Weibull distribution was a better fit for wind distributions at Sites 4 and 6, where the percentage of low wind speed was smaller (< 0.78%). These findings align with the work of Wais (2017).

Moreover, histograms presented in Figures 13-18 for wind distributions at Sites 1 to 6 show different shapes, indicating a variation in skewness. Thus, another reason for fitting a better distribution perhaps the skewness of the wind speed. A skewness is a measure of the asymmetry of the wind speed distribution about its mean. For a sample of *n* values, the skewness (γ_1) is defined as:

$$\gamma_1 = \frac{m_3}{s^3} \tag{30}$$

where, s = standard deviation and $m_3 = \frac{1}{n}\sum_{i=1}^{n}\left(U - \bar{U}\right)^3$ is the third central moment. A normal distribution is symmetrical and has $\gamma_1 = 0$. If γ_1 is negative, the distribution is left skewed whereas a positive γ_1 indicates a right skewed distribution. Since the Weibull distribution is a right skewed, γ_1 is expected to be positive.

Table 16 presents the mean ($\bar{U}$), standard deviation (s) and skewness (γ_1) of wind data at each site. It shows that the wind distribution of Sites 4 and 6 indicate higher skewness compared to Sites 1, 2, 3 and 5.

Table 16: Percentage lowest wind speed, mean, SD and best Weibull distribution

Site	% lowest speed (0-1 m/s)	Site mean wind speed ($\bar{U}$)	SD (s)	Skewness (γ_1)	Fitted distribution
1	1.71	5.35	2.22	0.28	3-p Weibull
2	1.87	5.83	2.29	0.09	3-p Weibull
3	7.04	6.29	3.20	0.20	3-p Weibull
4	0.71	7.50	3.31	0.62	2-p Weibull
5	0.79	7.31	3.20	0.22	3-p Weibull
6	0.78	6.97	2.95	0.51	2-p Weibull

Thus, the results reveal that the 3-p Weibull distribution is a better fit for wind data with both a greater frequency of low wind speeds (0-1 m/s) and low skewness, compared to a 2-p Weibull distribution.

To reiterate, this research is aimed at comparing the goodness of fit of both 2-p and 3-p Weibull distributions and to compare the performance of frequentist MLE and Bayesian methods during the estimation of Weibull parameters. Therefore, we conduct a comparison study of the four methods:

1. Fitting of 2-p Weibull distribution with an MLE estimate – MLE.2p
2. Fitting of 3-p Weibull distribution with an MLE estimate – MLE.3p
3. Fitting of 2-p Weibull distribution with a Bayesian estimate – BAYESIAN.2p
4. Fitting of 3-p Weibull distribution with a Bayesian estimate – BAYESIAN.3p

The wind density curves obtained using the four methods are illustrated in Figures 13-18.

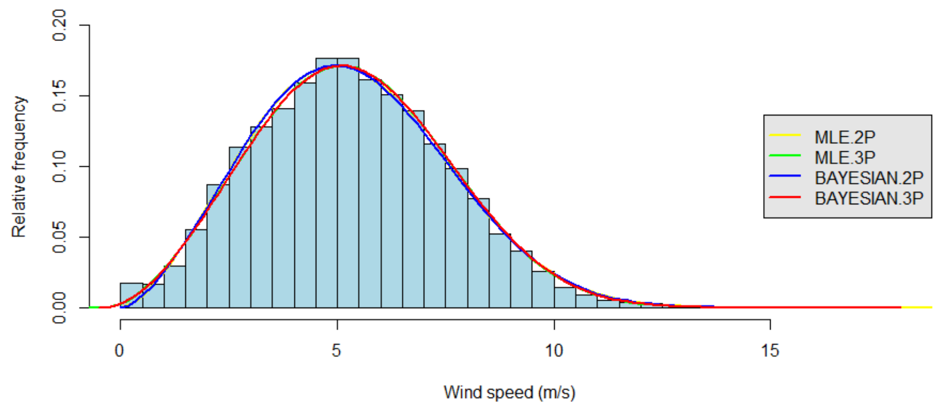


Figure 13: 2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 1.

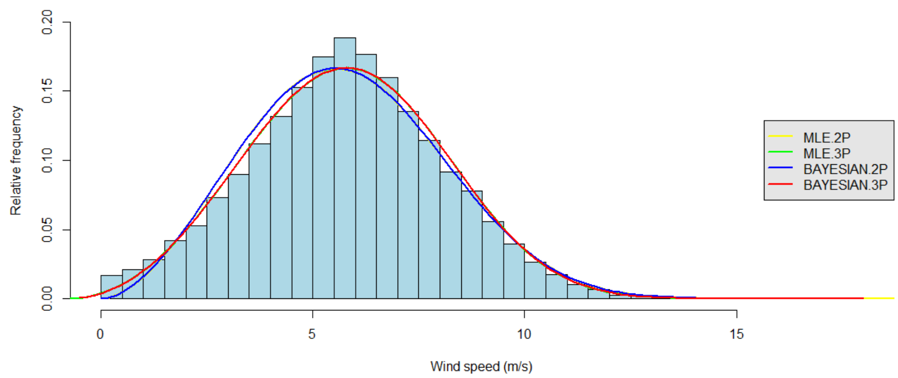


Figure 14: 2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 2.

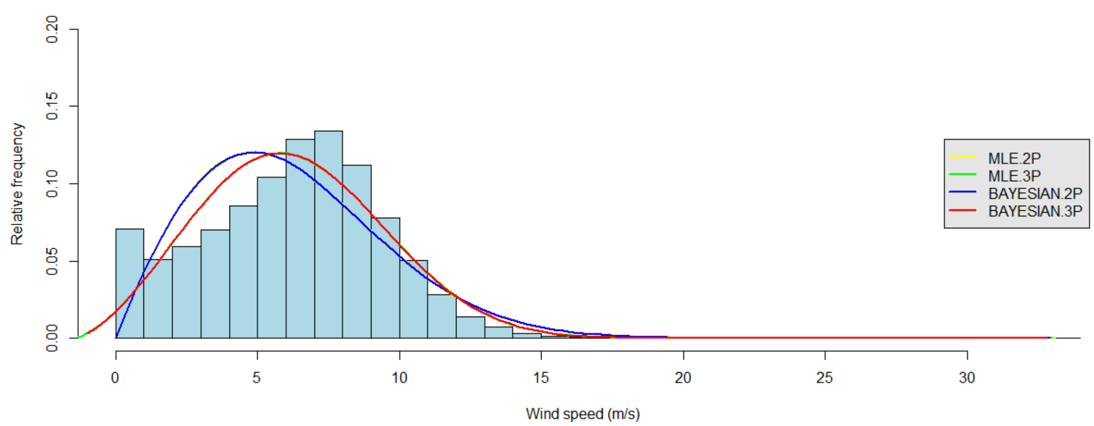


Figure 15: 2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 3.

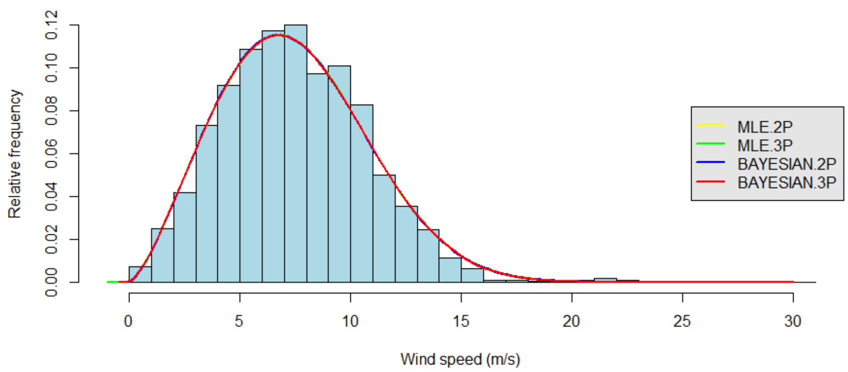


Figure 16: 2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 4.

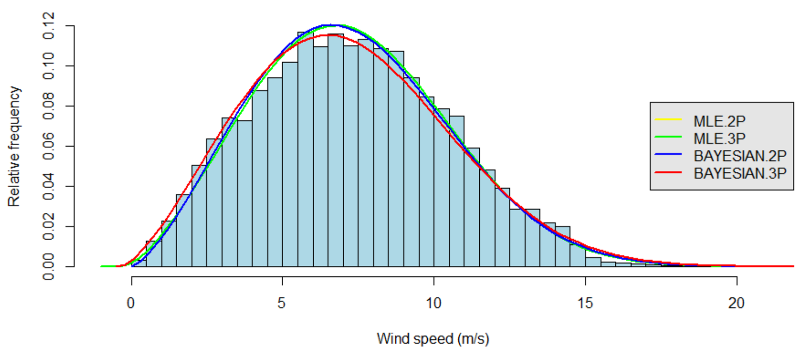


Figure 17: 2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 5.

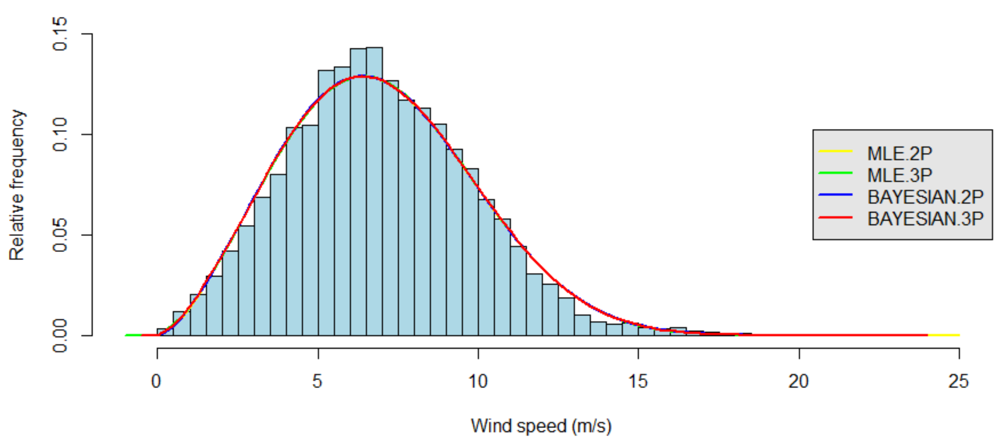


Figure 18: 2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 6.

Figures 13-18 show that for most sites, the density curves of 3-p Weibull distributions are closer to the histograms than the 2-p Weibull distributions. The density curves of Bayesian estimates are also closer to the histograms than MLE estimates. The 3-p Weibull distribution at Site 1 is better than the distribution obtained using the best 2-p Weibull distribution method (Aukitino et al., 2017).

To evaluate the performance of the four methods, we use the five statistical goodness of fit measures discussed in Section 6: R^2, RMSE, COE, MAE and MAPE. The results for wind data at all six sites are presented in Table 17.

Table 17: Goodness-of-fit measures for four methods

Site	Method	R^2	COE	RMSE	MAE	MAPE
1	MLE.2P	0.9982	**0.9936**	0.0983	0.0421	1.3052
	MLE.3P	**0.9983**	1.0011	**0.0975**	**0.0239**	**0.8269**
	BAYESIAN.2P	0.9982	**0.9936**	0.0983	0.0421	1.3053
	BAYESIAN.3P	**0.9983**	1.0011	**0.0975**	**0.0239**	**0.8269**
2	MLE.2P	0.9949	1.0048	0.1655	0.1114	3.4259
	MLE.3P	**0.9971**	**1.0034**	0.1234	0.0661	1.7917
	BAYESIAN.2P	0.9949	1.0048	0.1655	0.1114	3.4259
	BAYESIAN.3P	**0.9971**	1.0041	**0.1227**	**0.0645**	**1.7748**
3	MLE.2P	0.9696	0.9823	0.5675	0.4596	10.7507
	MLE.3P	**0.9755**	1.0681	0.5219	0.3029	7.9253
	BAYESIAN.2P	0.9696	**0.9822**	0.5677	0.4597	10.7516
	BAYESIAN.3P	**0.9755**	1.0679	**0.5216**	**0.3027**	**7.9149**
4	MLE.2P	**0.9787**	1.0512	0.4983	0.1546	1.9778
	MLE.3P	0.9783	1.0523	0.5025	0.1522	1.9013
	BAYESIAN.2P	**0.9787**	**1.0511**	**0.4981**	0.1544	1.9724
	BAYESIAN.3P	0.9783	1.0524	0.5036	**0.1518**	**1.8860**
5	MLE.2P	0.9983	**1.0180**	0.1372	0.1033	2.0063
	MLE.3P	**0.9989**	1.0186	0.1110	0.0788	1.5347
	BAYESIAN.2P	0.9983	**1.0180**	0.1372	0.1033	2.0069
	BAYESIAN.3P	**0.9989**	1.0182	**0.1109**	**0.0785**	**1.5275**
6	MLE.2P	**0.9911**	1.0173	0.2809	**0.1251**	1.9635
	MLE.3P	0.9909	1.0181	0.2852	0.1257	1.9187
	BAYESIAN.2P	**0.9911**	**1.0169**	**0.2806**	0.1253	1.9673
	BAYESIAN.3P	0.9909	1.0180	0.2854	0.1258	**1.9162**

Assessment of wind power energy:

To estimate the turbine productivity, it is necessary to calculate wind power and AEP to be obtained from each fitted model. If the density of the air $\rho = 1.16$ kg/m^3 and the turbine rotor diameter D = 32 m, the expected annual wind power and energy achieved from each site is determined using equations (22)-(25) for both 2-p and 3-p Weibull distributions. Table 18 shows estimated results for wind power and AEP.

The relative error of the estimated power shown in Table 18 is calculated as follows:

$$\text{Relative error (\%), RE} = \frac{\text{Estimated power - Actual power}}{\text{Actual power}} \times 100\%$$

If the RE is close to zero, the method estimates the parameter accurately. However, a positive RE implies an over-estimate and a negative RE implies an under-estimate by the method. From the RE values, we observed that the fitted methods slightly over-estimate the wind power for the Sites 1, 2, 3 and 5 and under-estimate wind power for Sites 4 and 6.

Table 18: Estimated power, relative error in power estimate and energy (Wind speed: 0 m/s – infinity)

Site	Methods	Power (kW)	RE (%)	AEP (kWh)
1	Actual	108457	0.00	950083
	MLE.2P	110306	1.70	966279
	MLE.3P	110051	1.47	964048
	BAYESIAN.2P	110298	1.70	966206
	BAYESIAN.3P	110061	**1.47**	964135
2	Actual	133525	0.00	1169683
	MLE.2P	136039	1.88	1191705
	MLE.3P	135765	1.68	1189301
	BAYESIAN.2P	136030	1.88	1191623
	BAYESIAN.3P	135753	**1.67**	1189195
3	Actual	206184	0.00	1806168
	MLE.2P	220464	6.93	1931263
	MLE.3P	209474	1.60	1834996
	BAYESIAN.2P	220556	6.97	1932074
	BAYESIAN.3P	209458	**1.59**	1834856
4	Actual	322492	0.00	2825032
	MLE.2P	319732	-0.86	2800855
	MLE.3P	319627	-0.89	2799928
	BAYESIAN.2P	319812	**-0.83**	2801557
	BAYESIAN.3P	319591	-0.90	2799619
5	Actual	290717	0.00	2546683
	MLE.2P	291707	0.34	2555350
	MLE.3P	290923	**0.07**	2548486
	BAYESIAN.2P	291715	0.34	2555427
	BAYESIAN.3P	295069	1.50	2584808
6	Actual	249192	0.00	2182919
	MLE.2P	247792	-0.56	2170656
	MLE.3P	247770	-0.57	2170468
	BAYESIAN.2P	247848	**-0.54**	2171151
	BAYESIAN.3P	247808	-0.56	2170799

Observations from results for all six sites are summarized below:

For Site 1, both the **3-p Weibull with MLE and Bayesian estimates fit better** as the R^2 is highest and almost all the other goodness of fit measures (RMSE, MAE and MAPE) in Table 17 are smaller for both 3-p Weibull methods suggesting the use of either of the two methods. However, the power estimation results in Table 18 indicate that the **3-p Weibull with Bayesian estimates is a better fit for wind speed at this site**.

For the Sites 2, 3 and 5, the **3-p Weibull with Bayesian estimates is a better fit** as the R^2 is highest and almost all the other measures (RMSE, MAE, MAPE and COE) are smaller. It also procures the most accurate power, which is very close to the actual power of these sites with smaller relative error. Moreover, the 3-p Weibull with MLE estimates also fits better at Site 5 as it gives the highest R^2 and it produces power with the smallest relative error. Thus, the results indicate that the 3-p Weibull fits the wind speed at these sites better supporting the results discussed in Section 7.

Finally, for the Sites 4 and 6, the **2-p Weibull with Bayesian estimates is a better fit** as the R^2 is highest and the other measures (RMSE and COE) are smaller. It also procures the most accurate power, which is very close to the actual power of the sites with the smallest relative error.

9. Conclusion

Assessing wind energy potential at any site requires knowledge of correct statistical distributions of wind speeds at the given site. Several studies around the world suggest that the two-parameter Weibull distribution is a good fit for wind data and recommend its application in wind resource exploration. However, some sites provide high uncertainty while fitting two-parameter Weibull distributions to wind speed data and warrant the need to explore distributions that characterize wind speed better, such as the three-parameter Weibull distribution. In this study, investigation of wind characteristics and wind energy potential are carried out at different locations in the Equatorial region, of which three sites are in Fiji and one each from Cooks Islands, Tonga and Kiribati, respectively. The wind speed data at these six sites were tested for the best model between the two-parameter and three-parameter Weibull distributions. Furthermore, as there is no unique method that characterizes wind data perfectly, it is also imperative to have the knowledge of the best method of estimation for the parameters of wind speed distribution at a given site. In this study, we introduced a novel approach by using the Bayesian method for estimating parameters of wind speed distributions at the six sites selected for testing the method. Then, a comparison study was conducted for the performance of the proposed Bayesian method with the popular frequentist MLE method. Finally, the results suggest that the three-parameter Weibull distributions should be used in analyzing wind power potential when the distribution has frequent low wind speeds and is less skewed. The results also conclude that the Bayesian approach provides more accurate results while characterizing wind speed and can be proposed as an alternative technique for estimating Weibull parameters. The proposed method can be incorporated in the popular software packages such as WAsP used for wind resource assessment and for planning wind energy projects.

Data Accessibility Statement

Sample data are provided with the manuscript. Full data will be made available upon request and after approval from the respective Governments.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Funds for carrying out this work were provided by Korea International Cooperation Agency (KOICA) under its East-Asia Climate Partnership program. The project number was 2009-00042.

ORCID ID

M.G.M. Khan: https://orcid.org/0000-0001-5400-9703 and M. Rafiuddin Ahmed: https://orcid.org/0000-0002-3514-1327

References:

Al Omari, M. A. and Ibrahim, N. A. (2011) Bayesian survival estimation for Weibull distribution with censored data. *Journal of Applied Sciences*, 11 (2):393-396.

Azad, A.K., Rasul, M.G. and Yusaf, T. (2014) Statistical diagnosis of the best Weibull methods for wind power assessment for agricultural applications. *Energies*, 7:3056–3085. DOI:10.3390/en7053056.

Aukitino, T., Khan, M.G.M., Ahmed, M.R. (2017) Wind energy resource assessment for Kiribati with a comparison of different methods of determining Weibull parameters. *Energy Conversion and Management*, 151:641–660.

Carta, J.A., Ramírez, P. and Velázquez, 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.

Casella, G. and Berger, R.L. (2002) *Statistical inference*. 2nd ed. USA: Thomson Learning Inc.

Chaurasiya, P.K., Ahmed, S. and Warudkar, V. (2018) Comparative analysis of Weibull parameters for wind data measured from met-mast and remote sensing techniques. *Renewable Energy*, 115(C): 1153-1165.

Chang, T.P. (2011) Performance comparison of six numerical methods in estimating Weibull parameters for wind energy application. Applied Energy, 88(1):272-282.

Corotis, R.B., Sigl, A.B. and Klein, J. (1978) Probability models of wind velocity magnitude and persistence. *Solar energy*, 20(6):483-493.

Cousineau, D. and Allan, T. A. (2015) Likelihood and its use in parameter estimation and model comparison. Mesure et évaluation en éducation, 37(3):63-98. DOI:10.7202/1036328ar.

Gosai, A.D. (2014) Wind Energy potential and electricity generation at Rokavukavu, Rakiraki. *MSc thesis*, The University of the South Pacific.

Guure, C.B., Ibrahim, N.A. and Al Omari, M.A. (2012) Bayesian estimation of two-parameter Weibull distribution using extension of Jeffreys' prior information with three loss functions. *Mathematical Problems in Engineering*, 12. DOI:10.1155/2012/589640.

Hossain, A. and Zimmer, W. (2003) Comparison of estimation methods for Weibull parameters: complete and censored samples. *Journal of Statistical Computation and Simulation*, 73 (2): 145–153.

IEC (2017) Wind turbines – part 12-1: power performance measurements of electricity producing wind turbines. Available at: www.iec.ch (accessed 25 November 2020).

Indhumathy, D., Seshaiah, C.V. and Sukkiramathi, K. (2014) Estimation of Weibull parameters for wind speed calculation at kanyakumari in India. *International Journal of Innovative Research in Science*, 3(1):8340-8345.

Jain, P. (2016) Wind Energy Engineering. New York, NY: McGraw Hill.

Kidmo, D.K., Danwe, R., Doka, S.Y. and Djongyang, N. (2015) Statistical analysis of wind speed distribution based on six weibull methods for wind power evaluation in Garoua, Cameroon. Revue des Energies Renouvelables, 18(1):105–125.

Lawless J.F. (2003) *Statistical Models and Methods for lifetime Data*, second ed. John Wiley & Sons, Hoboken, New Jersey.

Luceño, A. (2008) Maximum likelihood vs. maximum goodness of fit estimation of the three-parameter Weibull distribution, *Journal of Statistical Computation and Simulation*, 78(10): 941-949. DOI: 10.1080/00949650701467363.

Lye, L.M., Hapuarachchi, K.P. and Ryan, S. (1993) Bayes estimation of the extreme-value reliability function. *IEEE Transactions on Reliability*, 42(4):641-644. DOI: 10.1109/24.273598.

Mohanty, M. (2012). New renewable energy sources, green energy development and climate change: Implications to Pacific Island countries. *Management of Environmental Quality: An International Journal*, 23(3):264-274. DOI:10.1108/14777831211217468.

Mostafaeipour, A., Jadidi, M., Mohammadi, K. and Sedaghat, A. (2014) An analysis of wind energy potential and economic evaluation in Zahedan, Iran. *Renewable and Sustainable Energy Reviews*, 30:641-650.

Manwell, J.F., McGowan, J.G. and Rogers, A.L. (2010) *Wind energy explained: theory, design and application*. Wiley.

Pacific Power Association (2015) World Bank Provides Grant for Renewable Energy, available online at http://www.ppa.org.fj/world-bank-provides-grant-for-renewable-energy/

Pandey, B.N., Dwividi, N. and Pulastya, B. (2011) Comparison between Bayesian and maximum likelihood estimation of the scale parameter in Weibull distribution with known shape under linex loss function. *Journal of Scientific Research*, 55:163–172.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20-22, Vienna, Austria, ISSN 1609-395X.*

Ramachandran, K.M. and Tsokos, C.P. (2015) *Mathematical statistics with applications in R*. Elseier.

Spiegelhalter, D.J., Best, N., Carlin, B.P. and van der Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion). *Journal of Royal Statistical Society*, Series B, 64:583-639.

Su, Y. and Yajima, M. (2020). *Using R to run 'JAGS'*. R package version 0.6-1, URL https://cran.r-project.org/web/packages/R2jags.

Sukkiramathi, K. and Seshaiah, C.V. (2020) Analysis of wind power potential by the three-parameter Weibull distribution to install a wind turbine, *Energy Exploration & Exploitation,* 38(1):158-174. DOI: 10.1177/0144598719871628

Teimouri, M. and Gupta, A.K. (2013). On the three-parameter Weibull distribution shape parameter estimation. *Journal of Data Science*, 11:403-414.

Tuller, S.E. and Brett, A.C. (1984). The Characteristics of Wind Velocity that Favor the Fitting of a Weibull Distribution in Wind Speed Analysis, Journal of Climate and Applied Meteorology, 23(1), 124-134.

VNR Report (2019), Voluntary National Review, Republic of Fiji, 38-41

Wais, P. (2017) Two and three-parameter Weibull distribution in available wind power analysis. *Renewable energy*, 103: 15-29.

Figures

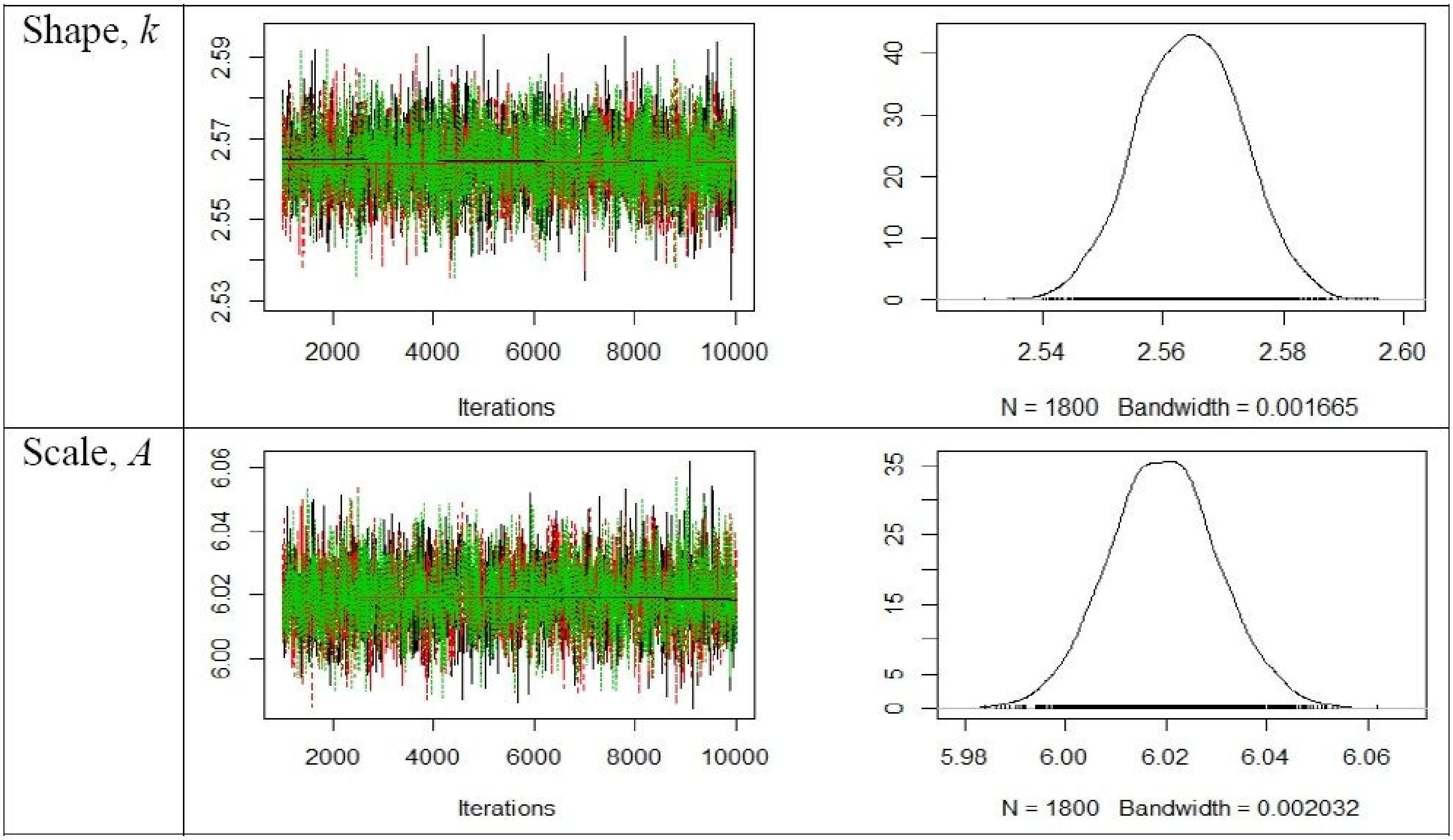


Figure 1

Trace and posterior density plots for Site 1 (2p-Weibull)

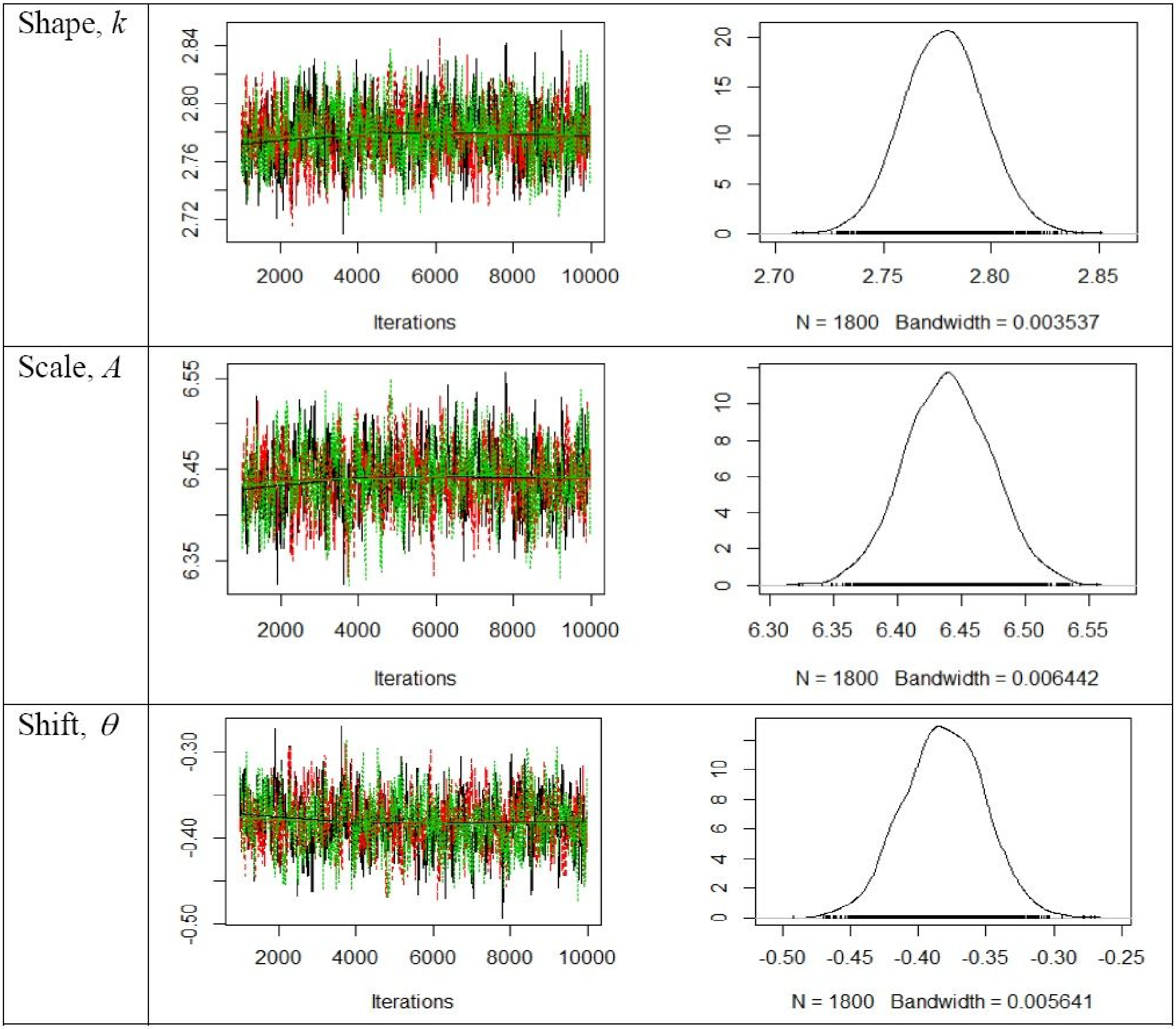


Figure 2

Trace and posterior density plots for Site 1 (3p-Weibull)

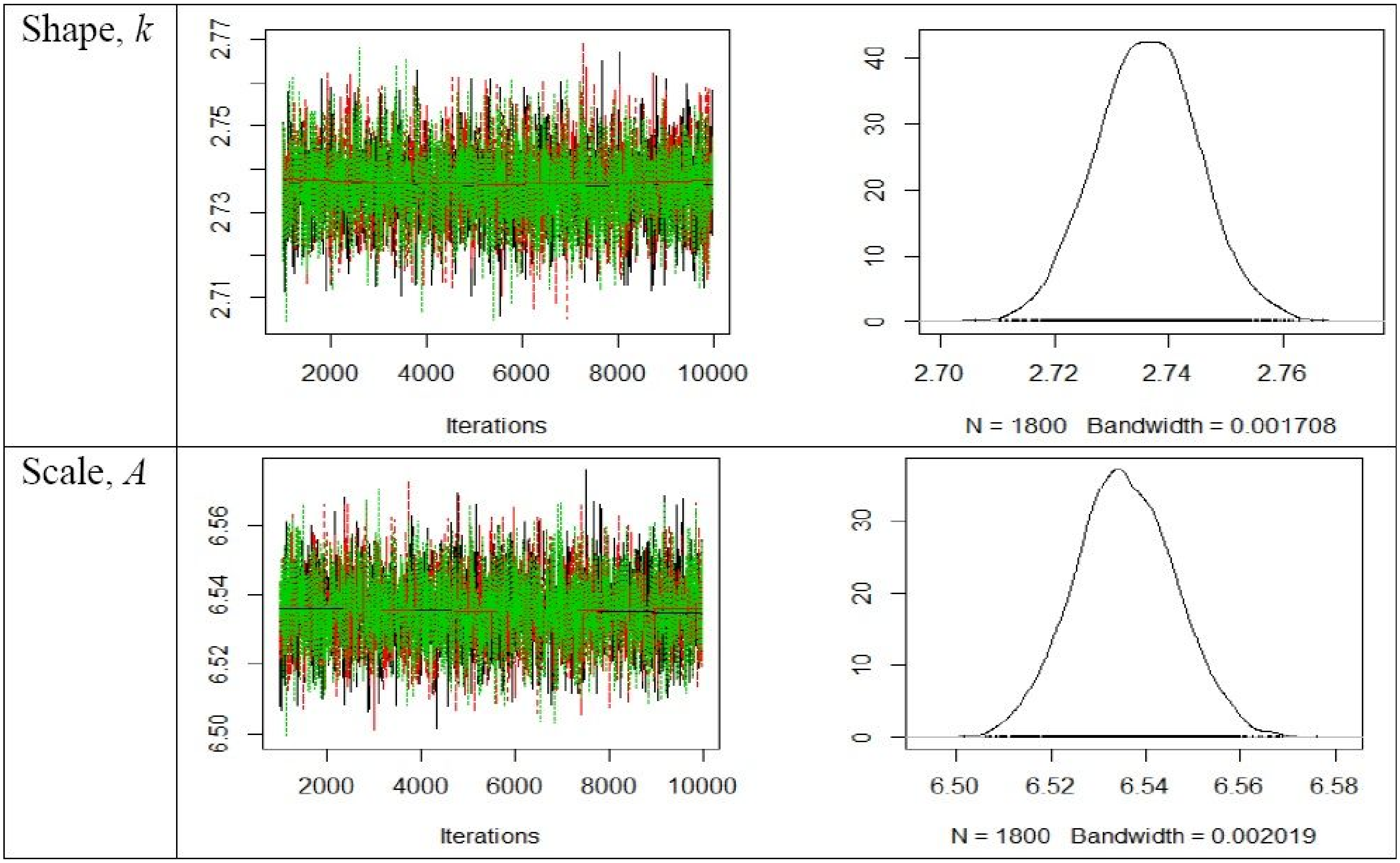


Figure 3

Trace and posterior density plots for Site 2 (2p-Weibull)

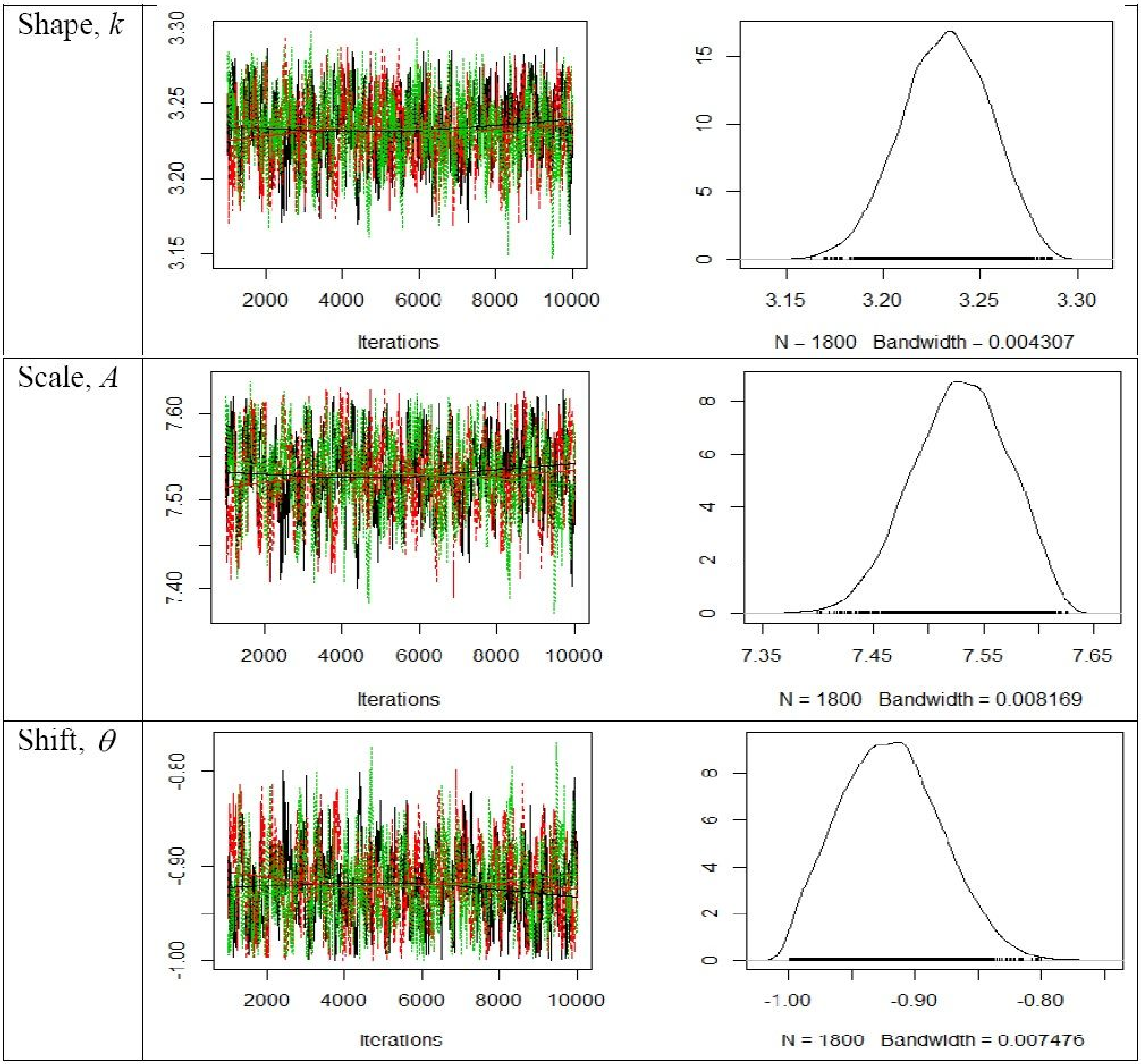


Figure 4

Trace and posterior density plots for Site 2 (3p-Weibull)

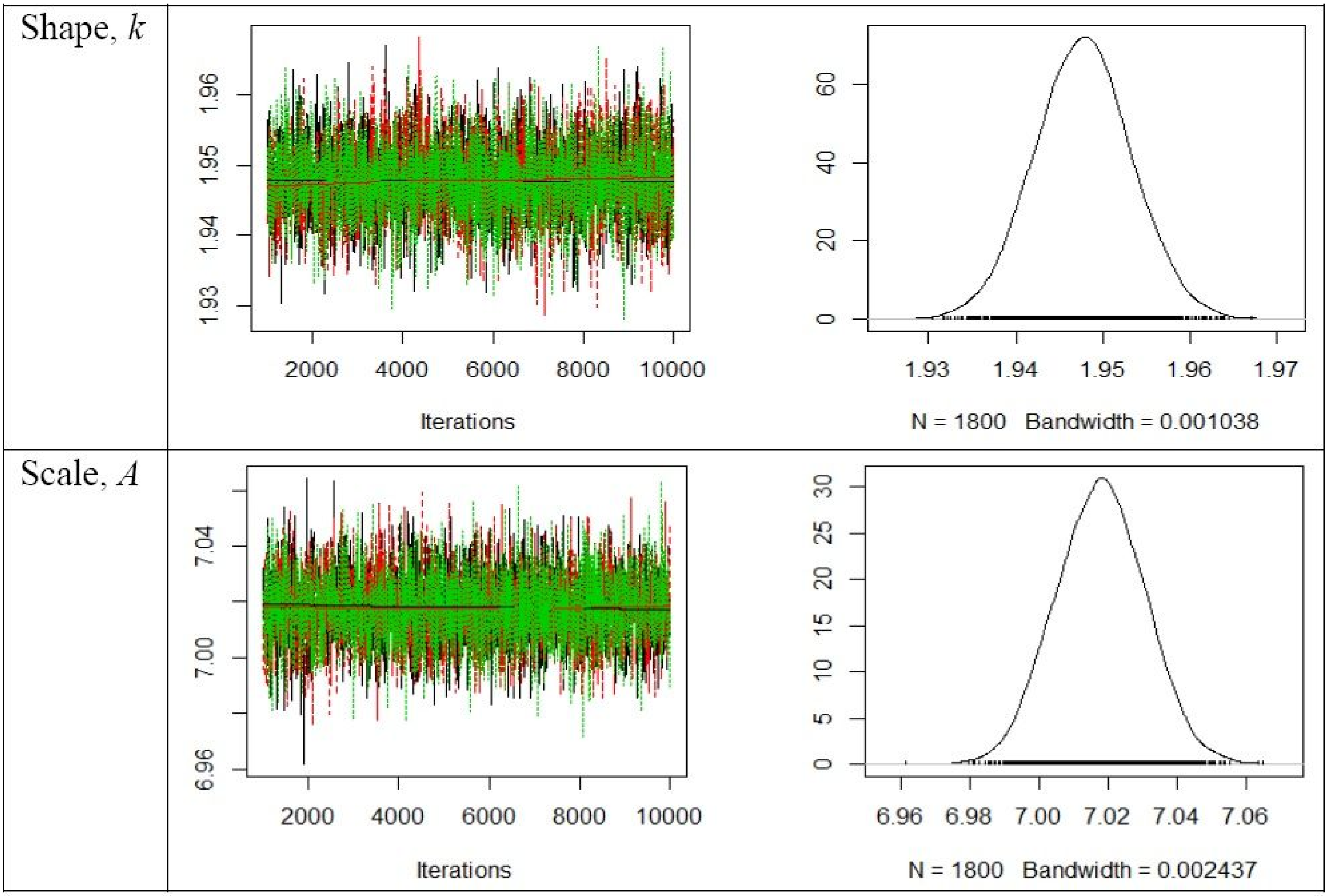


Figure 5

Trace and posterior density plots for Site 3 (2p-Weibull)

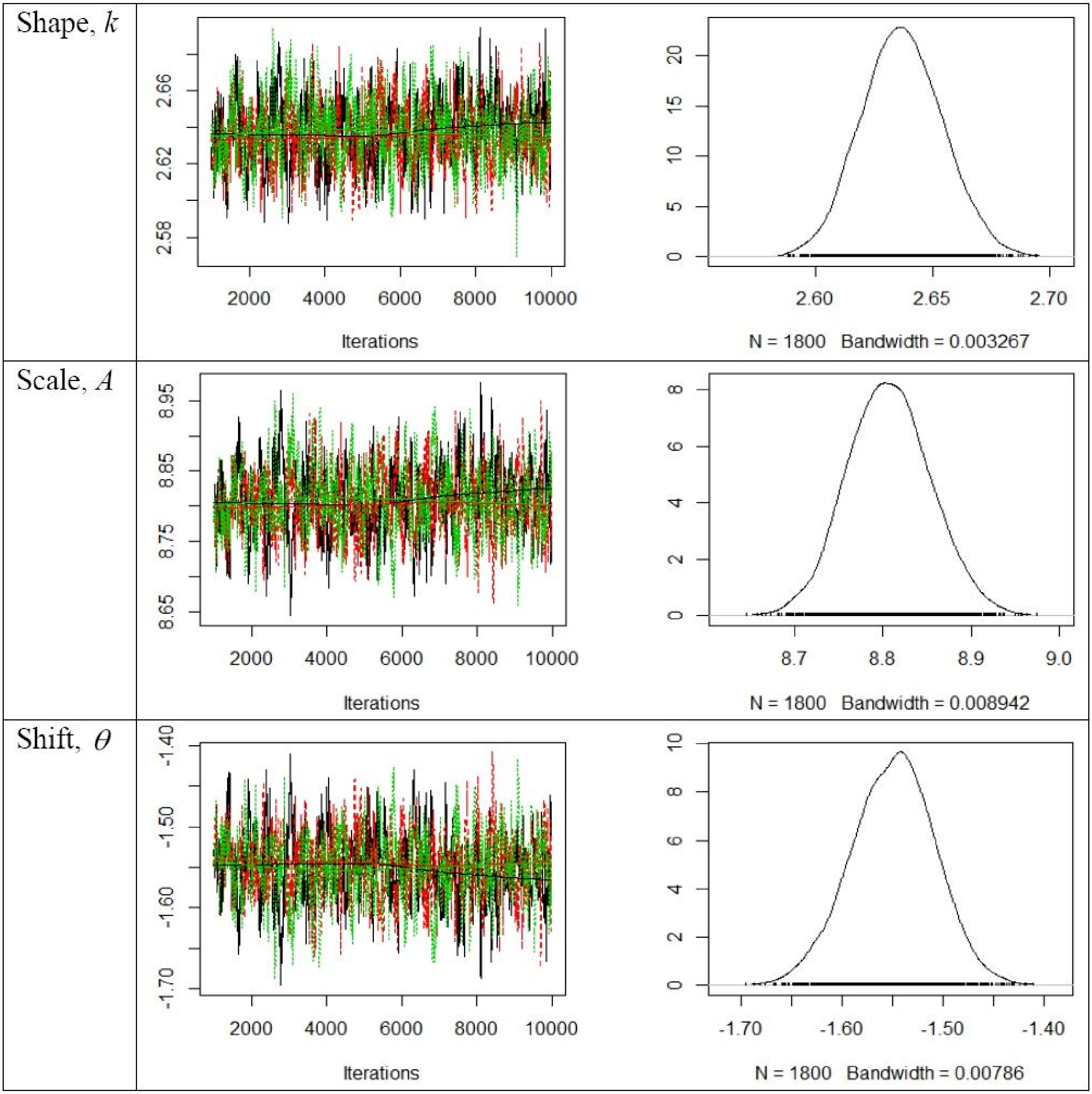


Figure 6

Trace and posterior density plots for Site 3 (3p-Weibull)

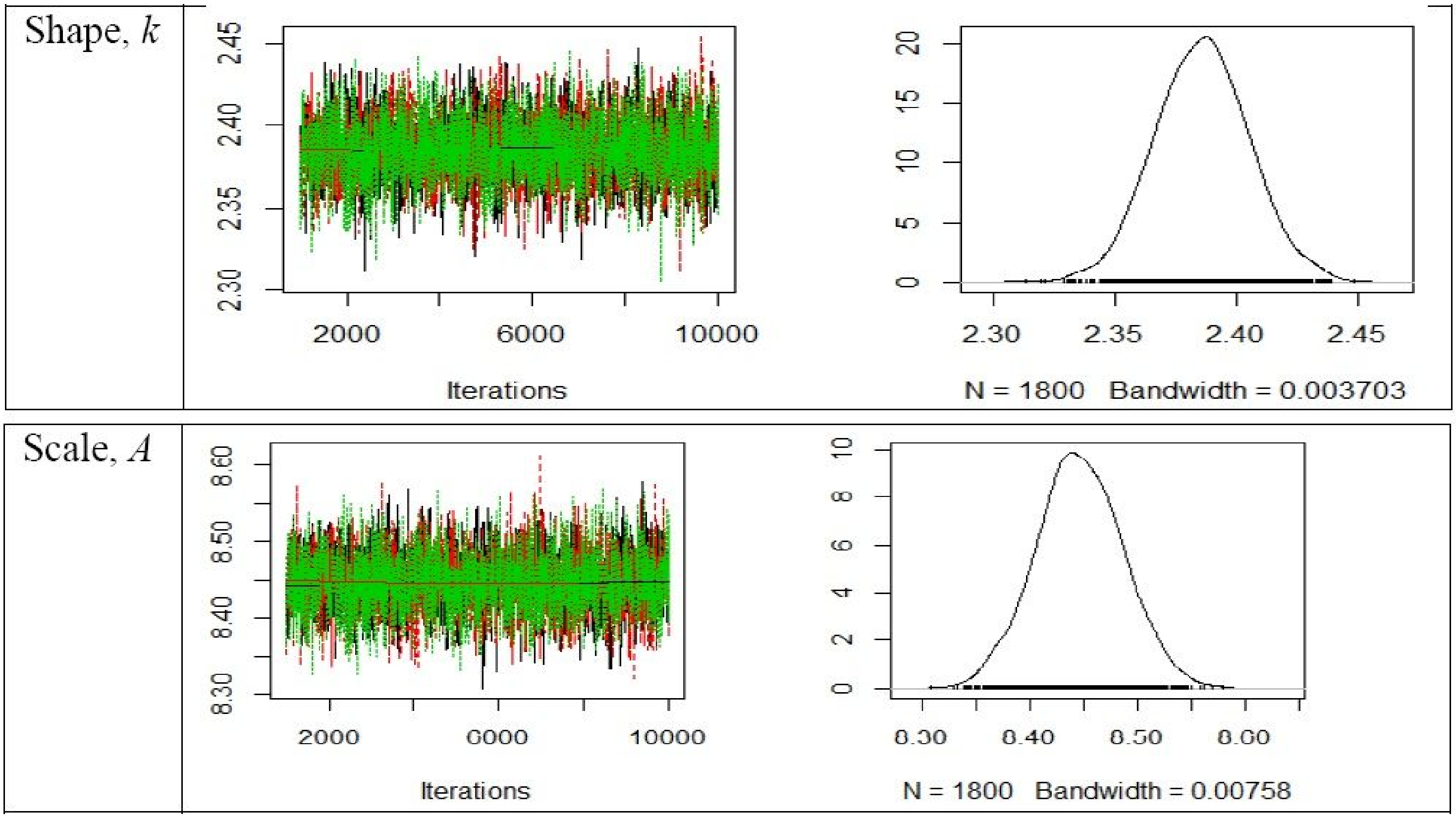


Figure 7

Trace and posterior density plots for Site 4 (2p-Weibull)

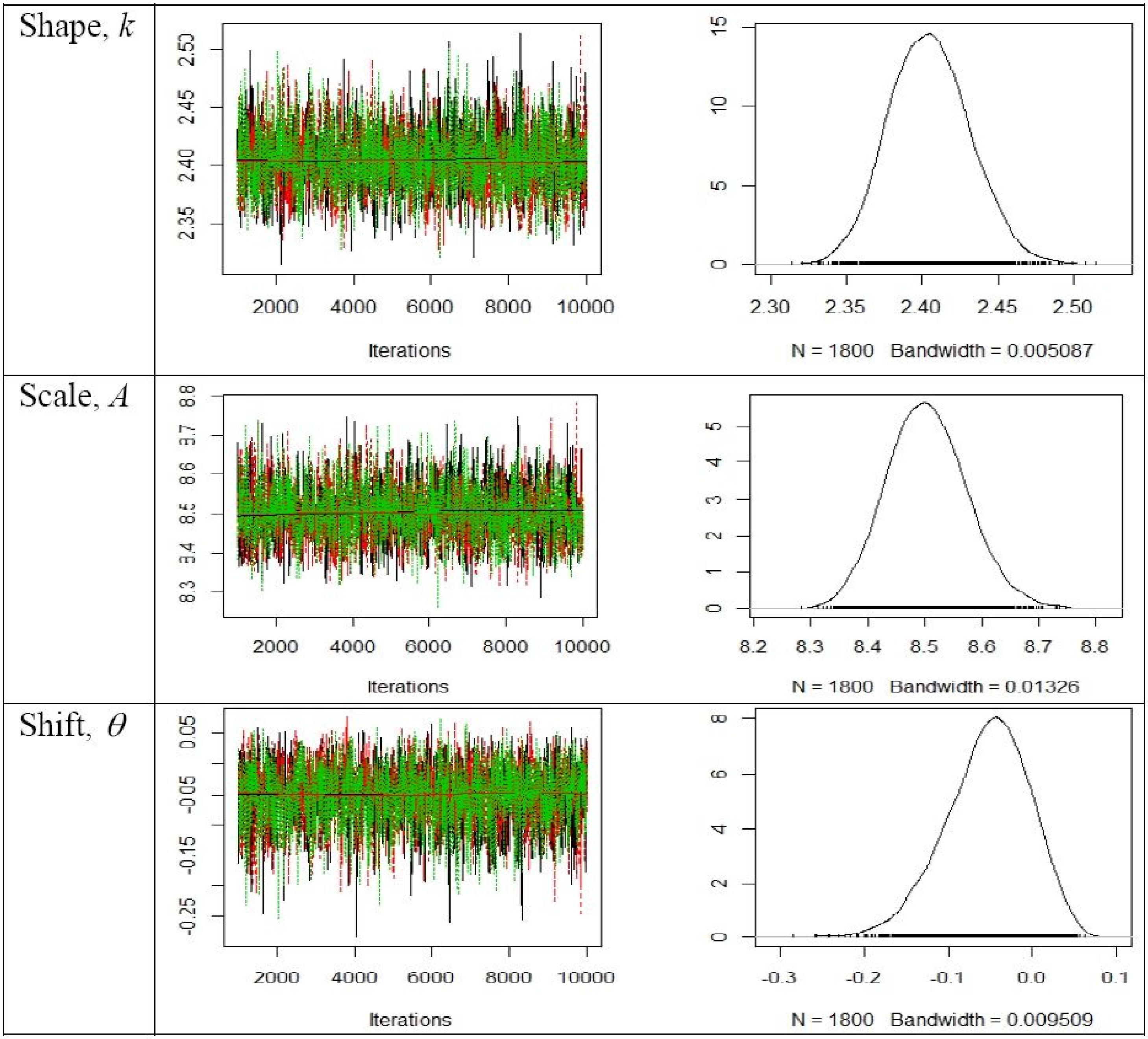


Figure 8

Trace and posterior density plots for Site 4 (3p-Weibull)

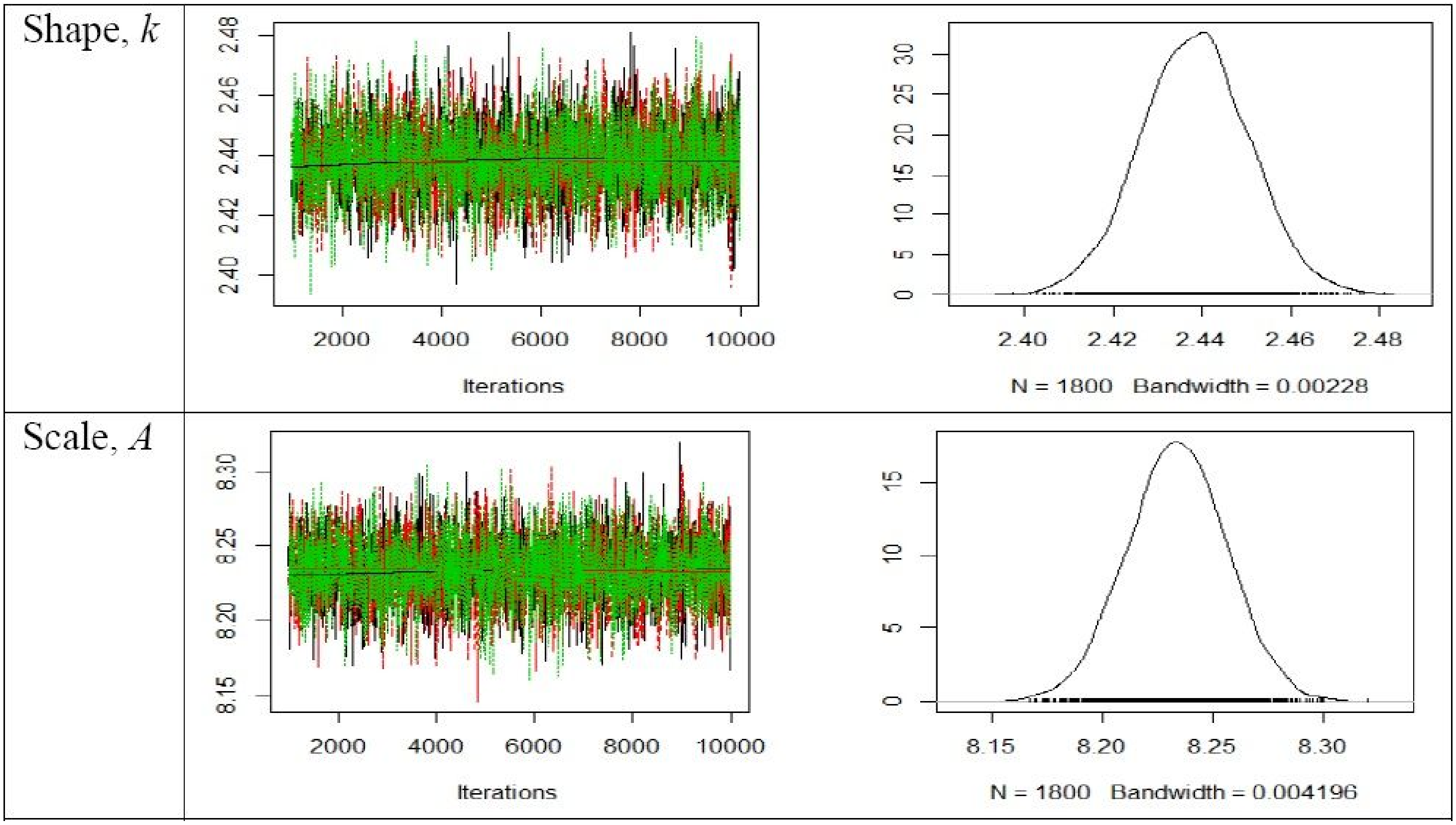


Figure 9

Trace and posterior density plots for Site 5 (2p-Weibull)

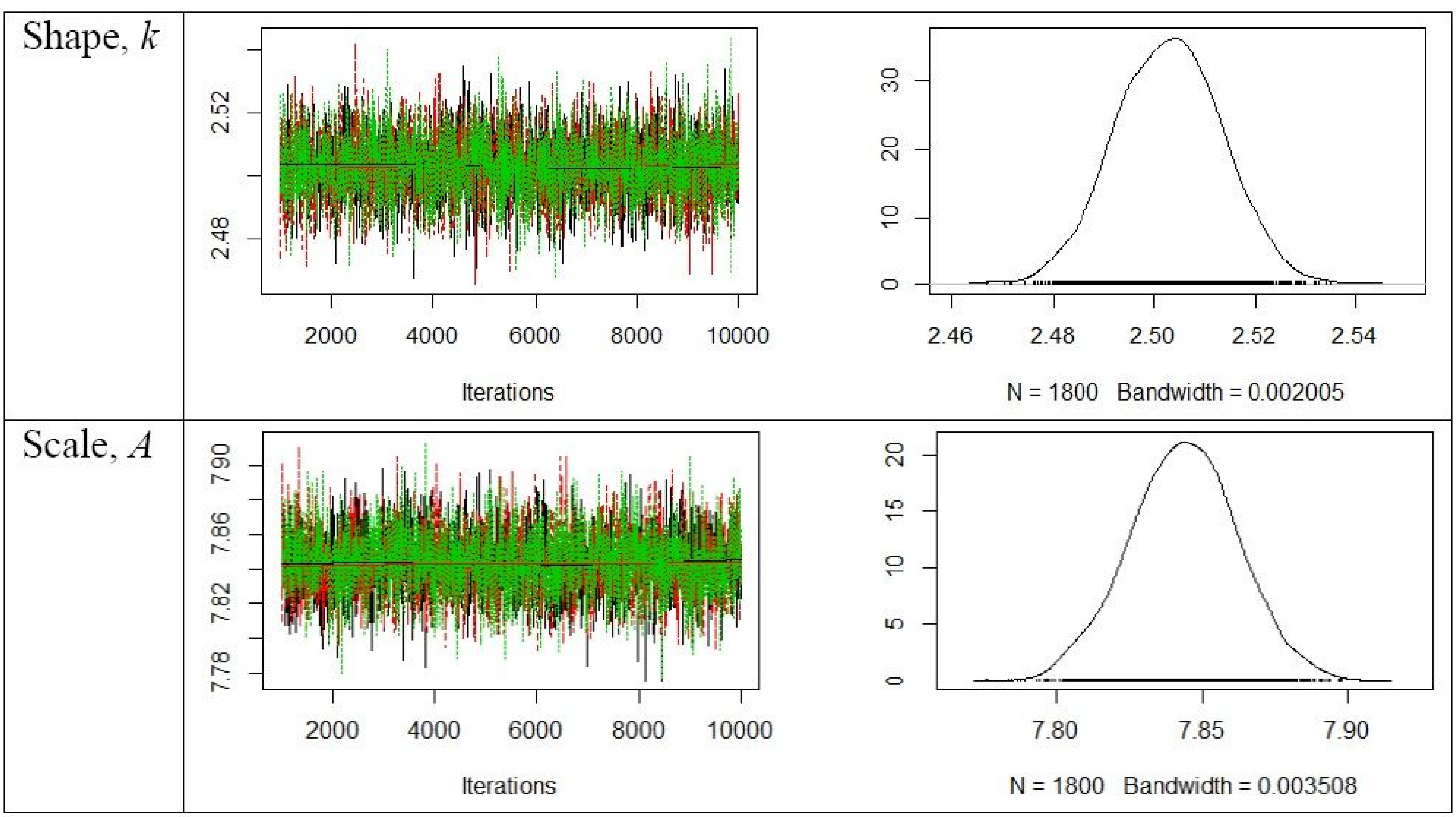


Figure 11

Trace and posterior density plots for Site 6 (2p-Weibull)

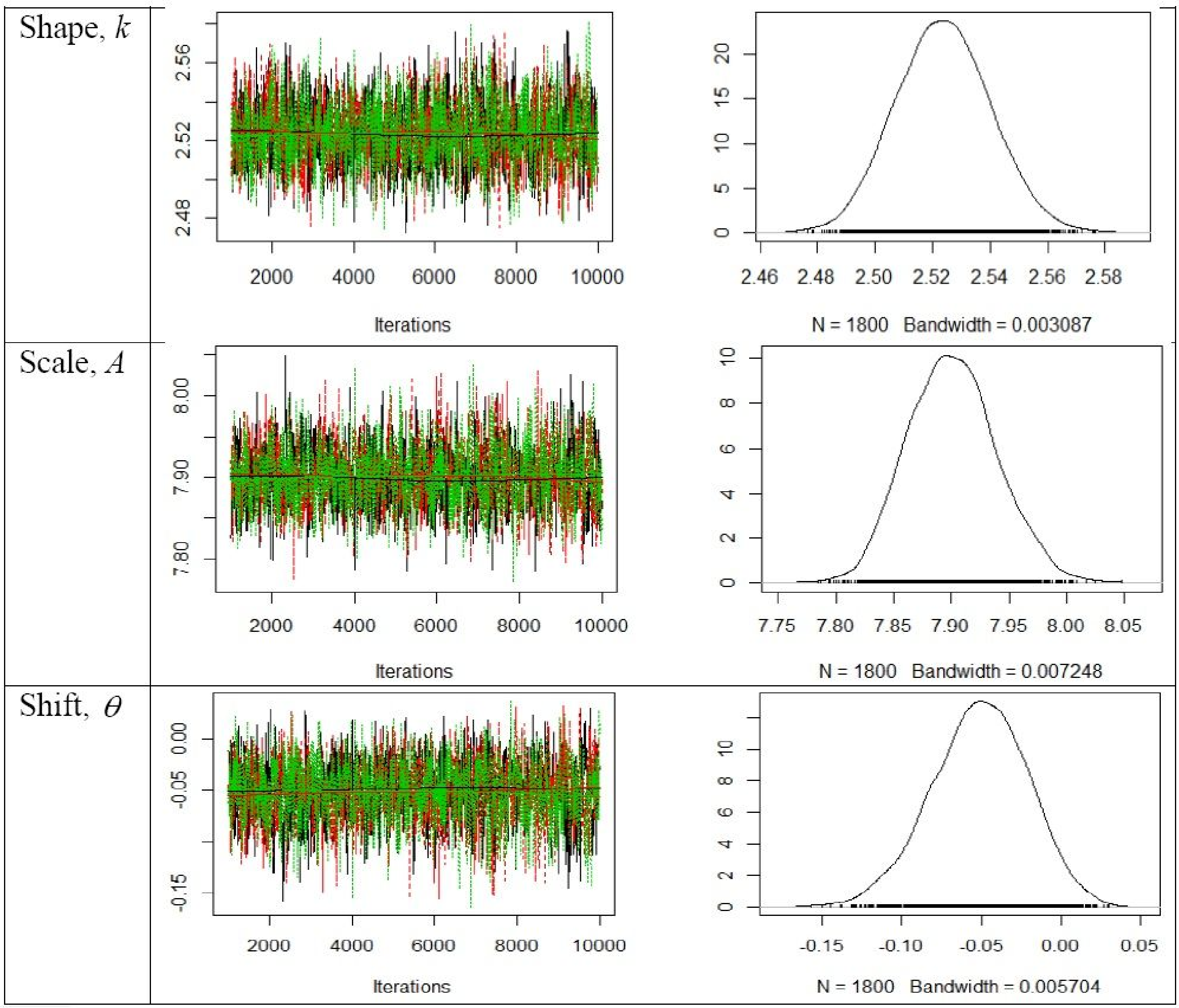


Figure 12

Trace and posterior density plots for Site 6 (3p-Weibull)

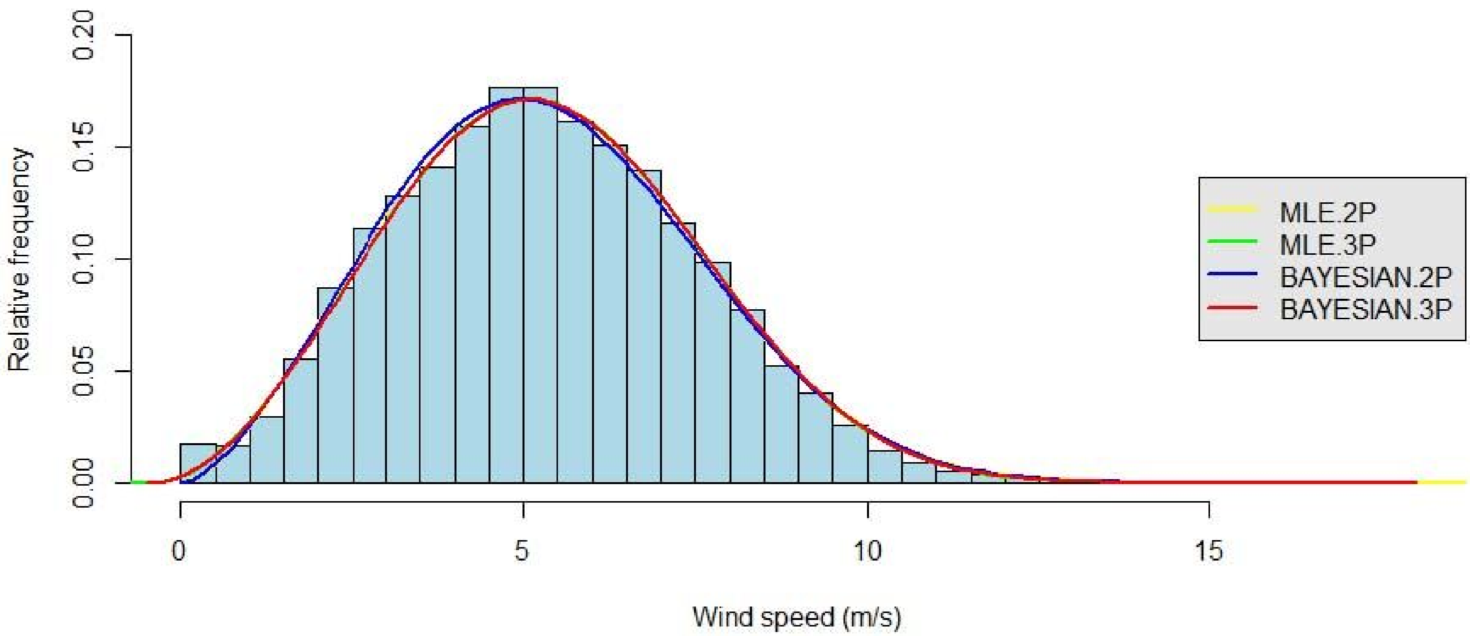


Figure 13

2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 1.

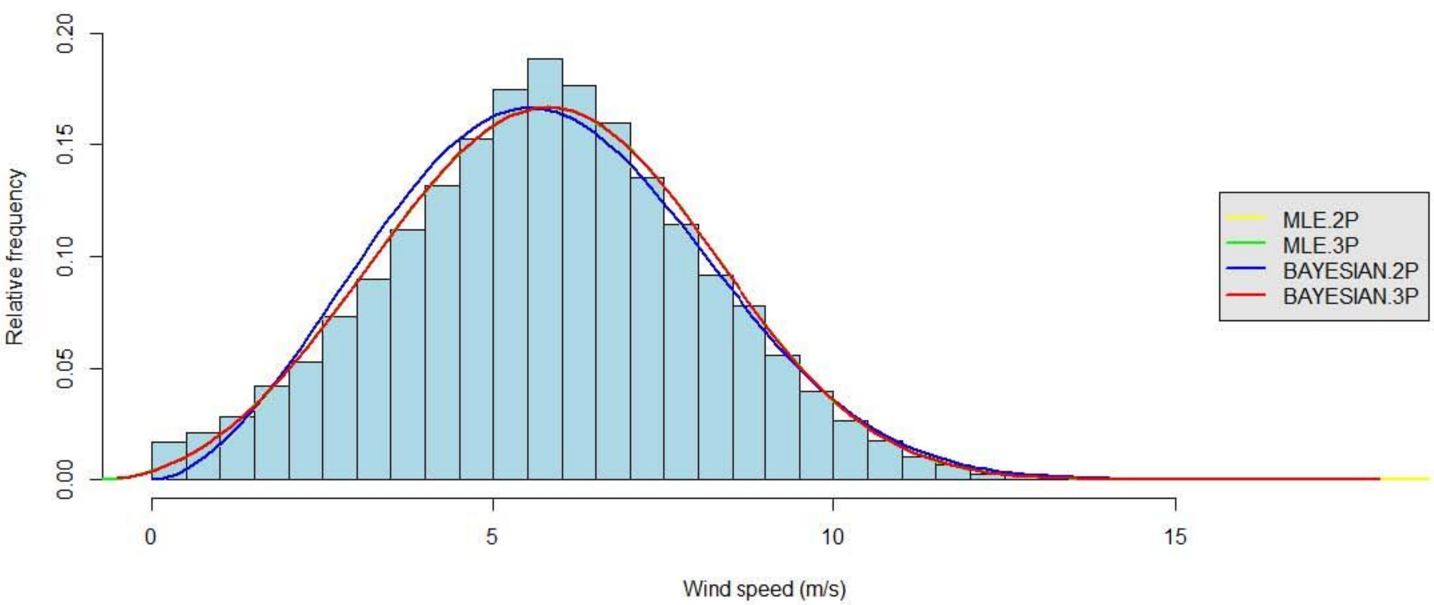


Figure 14

2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 2.

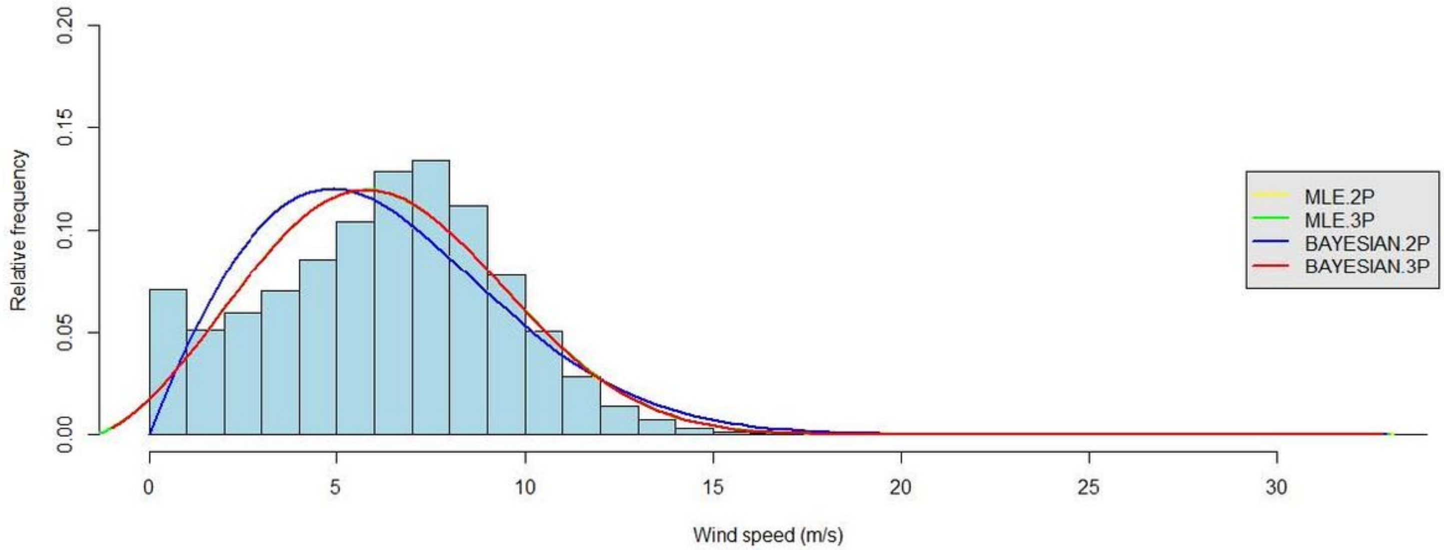


Figure 15

2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 3.

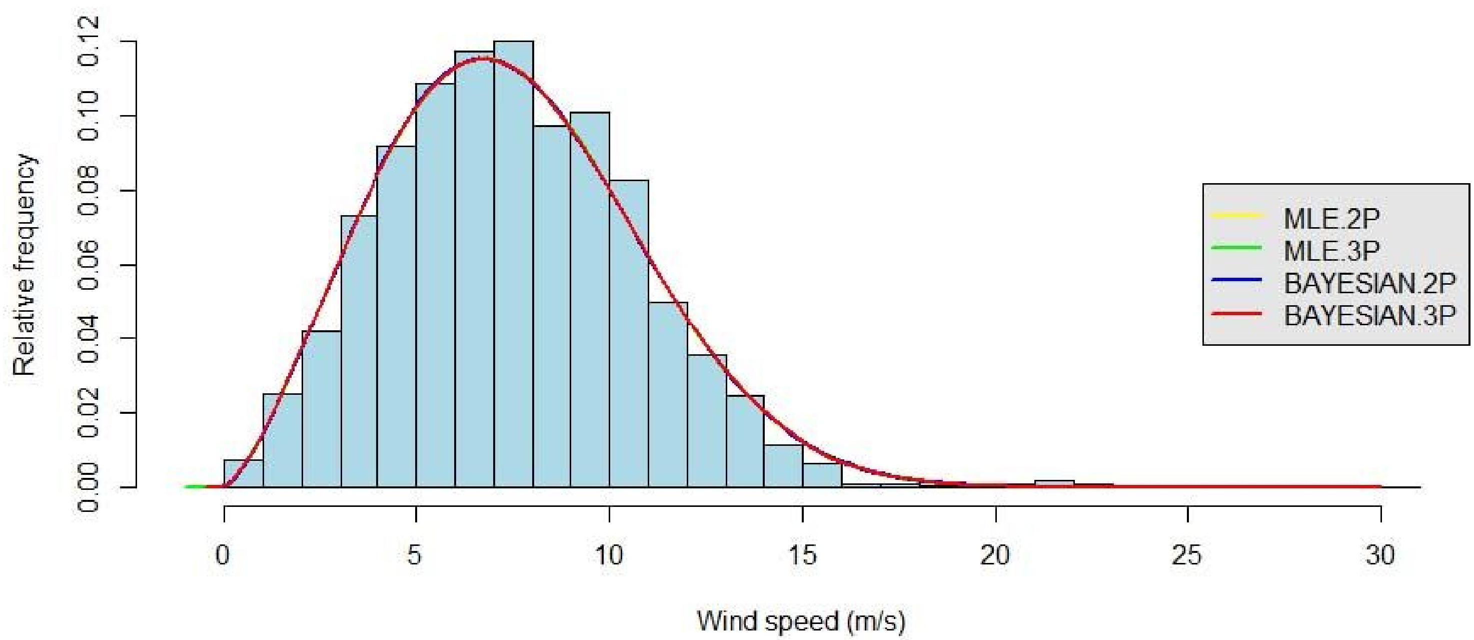


Figure 16

2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 4.

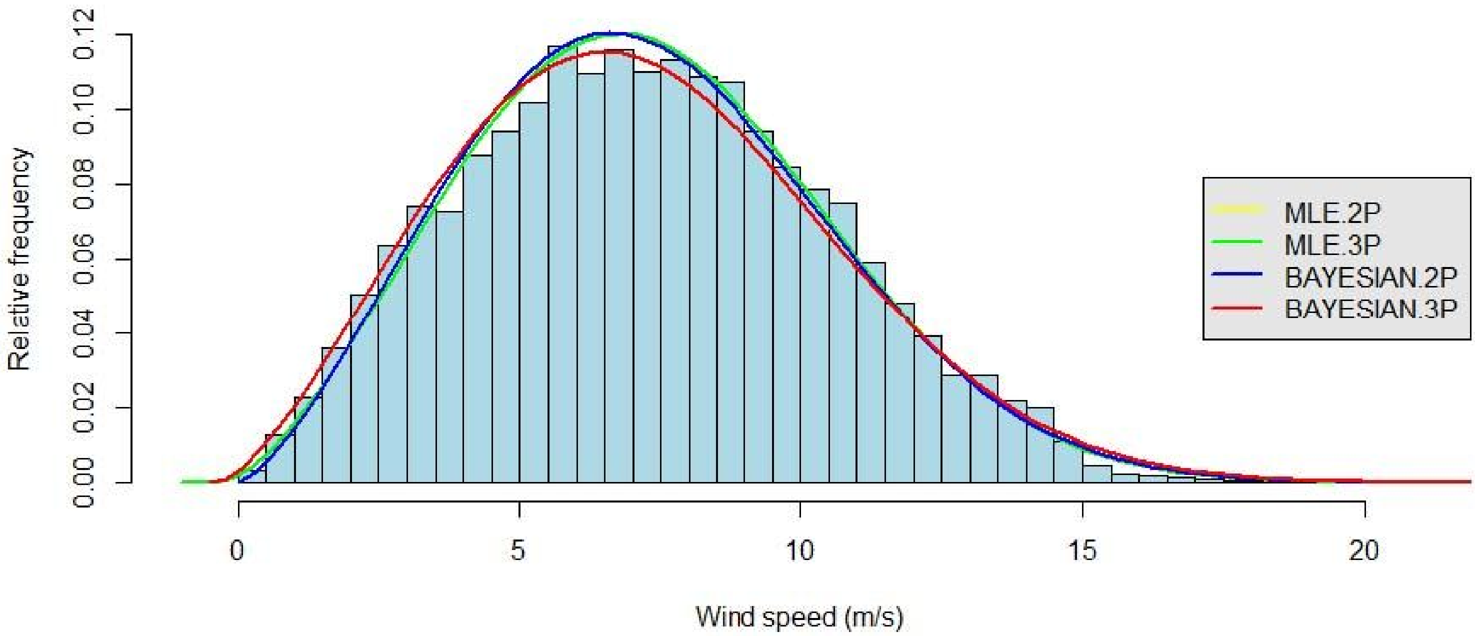


Figure 17

2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 5.

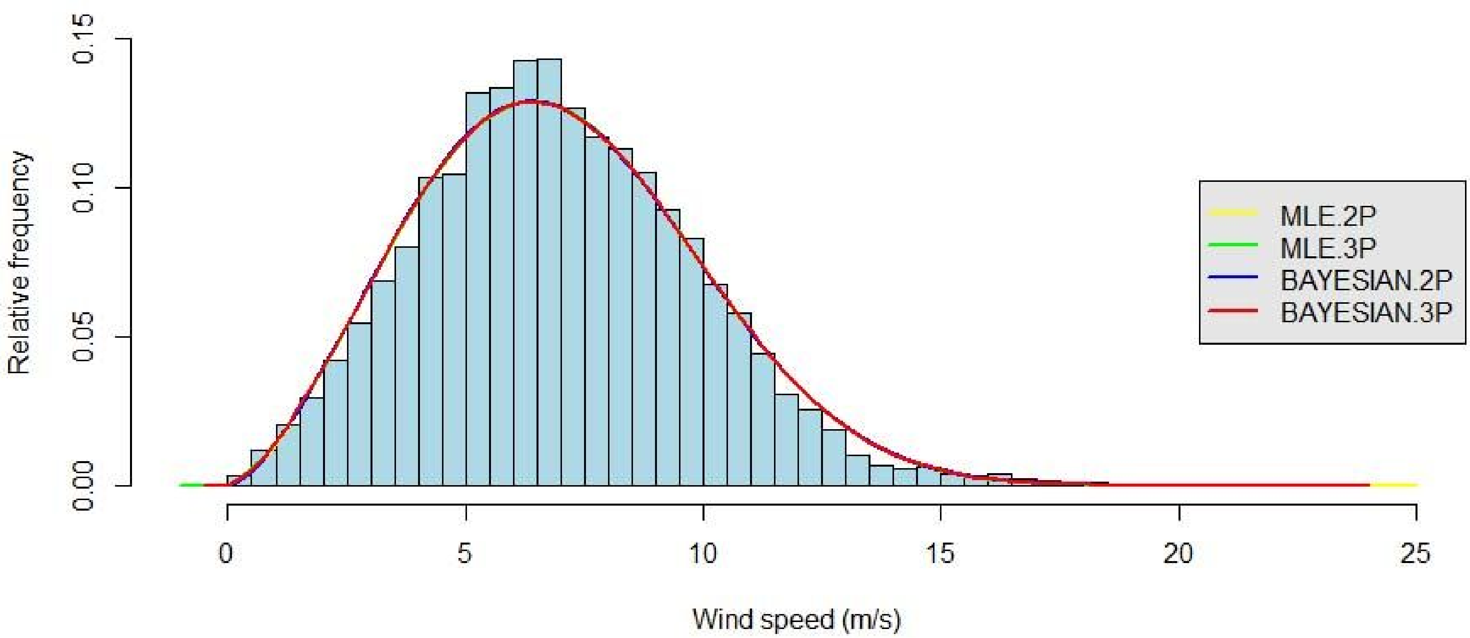


Figure 18

2-p and 3-p Weibull curves by four methods and histogram of the observed wind speeds at Site 6.