

Confidence Intervals for Differences between Percentiles of Two Weibull Distributions: An Application to Wind Speed Data

MANUSSAYA LA-ONGKAEW, SA-AAT NIWITPONG

Faculty of Applied Science, Department of Applied Statistics

King Mongkut's University of Technology North Bangkok, Bangkok, Thailand

1518 Pracharat 1 Road, Wongsawang, Bangsue, Bangkok 10800

THAILAND

Abstract: This study aims to analyze, develop, and compare methods for assessing the difference between percentiles of Weibull distributions. Confidence intervals for the difference between percentiles are constructed to evaluate the accuracy and reliability of group comparisons. The statistical methods considered include the Generalized Confidence Interval (GCI), Bootstrap, and Bayesian approaches, which are used to construct and compare confidence intervals for differences between percentiles. The performance of the confidence interval is evaluated based on coverage probability and average interval width to determine the precision and efficiency of each approach. The result indicated that the Bayesian highest posterior density (Baye-HPD) interval outperforms others. Real-world wind speed data, categorized by location, are employed to demonstrate the practical applicability of the proposed methods.

Key-Words: - Weibull distribution, Percentile, GCI, Bayesian, Bootstrap, Wind speed

Received: April 27, 2026. Revised: May 29, 2026. Accepted: July 4, 2026. Published: September 30, 2026.

1 Introduction

Weibull distribution is recognized as one of the most versatile statistical distributions for modeling asymmetric data. Its flexibility has led to widespread applications across diverse disciplines, including engineering, industrial reliability analysis, actuarial science, and environmental studies, [1], [2], [3], [4], [5]. Existing literature has largely focused on estimating the primary parameters of the distribution, namely the shape and scale parameters, using both point and interval estimation methods. Additional distributional characteristics, such as the mean, variance, and coefficient of variation, have also been subjects of substantial investigation.

Despite substantial progress in parameter estimation, research on the percentiles of the Weibull distribution remains relatively limited, even though percentiles are essential for risk assessment and informed decision-making. Percentiles such as the median or the 90th percentile offer robust indicators of central tendency and data concentration, as they are less sensitive to extreme observations. Higher-order percentiles, including the 95th and 99th percentiles, are particularly important in environmental and public health contexts, where they serve as benchmarks for evaluating extreme or hazardous conditions. Furthermore, examining differences in percentiles across groups provides valuable insights into spatial or temporal variability in phenomena such as wind speed or PM_{2.5} concentrations.

Research on percentile estimation has gained increasing attention in recent years. Several methods

for estimating percentiles of the normal distribution were investigated, [6], including the traditional approaches, Normal Asymptotic, and Lawless. Using simulated samples of various sizes drawn from a standard normal distribution, their findings indicated that the Traditional and Lawless methods outperformed the Normal Asymptotic method, as they yielded shorter interval lengths while maintaining superior coverage probabilities (CPs). Confidence intervals (CIs) for percentiles under the normal distribution were examined, and approaches based on maximum likelihood estimation and Bayesian inference were proposed, [7]. More recently, bootstrap-based CIs for percentiles of the exponential family were developed, [8]. The results demonstrated that when the true parameter lies on or near the boundary of the constrained region, the CPs consistently exceed the nominal level. Generalized confidence intervals, fiducial confidence intervals, and approximate CIs expressible in closed-form equations for interval estimation of the mean and upper percentiles of a lognormal population in the presence of zero-inflated data were proposed, [9]. More recently, CIs for both the difference and the ratio of percentiles of the zero-inflated Birnbaum–Saunders distribution were examined, [10]. Simple interval estimation procedures based on the Method of Variance Estimates Recovery (MOVER) and Bayesian inference were introduced, and the Bayesian approach generally offered superior performance.

Due to the limited number of studies on the difference between percentiles of Weibull distributions, this study explores and compares several methods for

comparing Weibull percentiles between groups. The main focus is on constructing CIs for the difference between percentiles in order to assess how reliable these comparisons are. The methods considered include the Generalized Confidence Interval (GCI), bootstrap, and Bayesian approaches. A simulation study carried out in R programming is used to evaluate the performance of these methods based on CP and average interval width. The methods are then applied to real wind speed data to demonstrate their practical usefulness.

2 Methods

Let $X = (X_1, X_2, \dots, X_i)$ be observations drawn from a two-parameter Weibull distribution with scale parameter α and shape parameter β . The probability density function (pdf) of this distribution is given in Eq. (1).

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x_i}{\alpha}\right)^{\beta-1} \exp \left[-\left(\frac{x_i}{\alpha}\right)^{\beta} \right], \quad x_i > 0, \quad (1)$$

for $i = 1, 2, \dots, n$. To estimate the parameters α and β , the maximum likelihood estimation is applied. The maximum likelihood estimators (MLEs) can be obtained from, [11], they are as follows. The MLE $\hat{\beta}$ of β is solution of the following equation

$$\frac{1}{\hat{\beta}} - \frac{\sum_{i=1}^n \left[x_i^{\hat{\beta}} \ln(x_i) \right]}{\sum (x_i^{\hat{\beta}})} + \frac{1}{n} \sum_{i=1}^n \ln(x_i) = 0 \quad (2)$$

and the MLE $\hat{\alpha}$ of α is given by

$$\hat{\alpha} = \left[\sum_{i=1}^n x_i^{\hat{\beta}} / n \right]^{\frac{1}{\hat{\beta}}}. \quad (3)$$

For solving the Eq. (2), the Newton-Raphson iterative method can be applied. Let $x_1, x_2, \dots, x_n$ be a sample of observations from a $Weibull(\alpha, \beta)$ distribution. Let $z_i = \ln(x_i), i = 1, 2, \dots, n$. [12] showed that the estimator $\hat{\beta}_u = \left(\frac{\pi}{\sqrt{6}} \right) \left(\sqrt{\frac{n-1}{s^2}} \right)$ is asymptotically unbiased estimator, which is the starting value. The following algorithm is used to compute the MLEs:

Algorithm 1

1. Generate $x_1, x_2, \dots, x_n$, let $z_i = \ln(x_i)$.
2. Compute $s_1 = \sum_{i=1}^n z_i$, $\bar{z} = s_1/n$ and $s^2 = \sum_{i=1}^n (z_i - \bar{z})^2$.
3. Compute $\hat{\beta}_u = \left(\frac{\pi}{\sqrt{6}} \right) \left(\sqrt{\frac{n-1}{s^2}} \right)$.
4. For $l = 1$ to number of iterations.

5. Compute $w_i = x_i^{\hat{\beta}_u}$, $s_2 = \sum_{i=1}^n (w_i)$, $s_3 = \sum_{i=1}^n (w_i z_i)$ and $s_4 = \sum_{i=1}^n (w_i z_i^2)$.
6. Compute $F = \left(\frac{1}{\hat{\beta}_u} \right) + \bar{z} - \left(\frac{s_3}{s_2} \right)$.
7. Compute $f = \left(\frac{1}{\hat{\beta}_u^2} \right) + [(s_2 s_4) - (s_3^2)] / (s_2^2)$.
8. Compute $\hat{\beta}_u = \hat{\beta}_u + \left(\frac{F}{f} \right)$.
9. If $(F \leq \text{error tolerance})$ return $\hat{\beta} = \hat{\beta}_u$.
10. End l loop.
11. Compute $\hat{\alpha}$ from Eq. (3).

Since the MLEs of the Weibull parameters do not have closed-form expressions, their exact expectations and variances cannot be obtained analytically. However, under regularity conditions, the MLEs are asymptotically normally distributed. Specifically,

$$(\hat{\alpha}, \hat{\beta}) \xrightarrow{\text{approx}} \mathcal{N}((\alpha, \beta), I(\alpha, \beta)^{-1}),$$

where $I(\alpha, \beta)$ denotes the Fisher Information Matrix.

Consequently, the asymptotic covariance matrix of $(\hat{\alpha}, \hat{\beta})$ can be approximated by the inverse of the Fisher Information Matrix, that is,

$$\text{Var} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} \approx I(\alpha, \beta)^{-1}.$$

These asymptotic results provide the basis for constructing approximate CIs and for applying the Delta method to functions of the Weibull parameters.

According to Eq. (4), the cumulative distribution function (cdf) expresses the probability that the random variable takes a value less than or equal to x_i .

$$F(x) = 1 - \exp \left[-\left(\frac{x_i}{\alpha}\right)^{\beta} \right]. \quad (4)$$

One primary motivation for estimating Weibull parameters is to determine the $(100p)$ th percentile, denoted X_p , which satisfies

$$P(X \leq X_p) = p. \quad (5)$$

Following, [13], setting $F(x) = p$ in Eq. (4) yields the percentile

$$X_p = \lambda = \alpha \left[\ln \left(\frac{1}{1-p} \right) \right]^{1/\beta}. \quad (6)$$

Thus, an estimator of λ can be expressed as

$$\hat{\lambda} = \hat{\alpha} \left[\ln \left(\frac{1}{1-p} \right) \right]^{1/\hat{\beta}}. \quad (7)$$

Let $X_{1,1}, X_{1,2}, \dots, X_{1,n}$ and $X_{2,1}, X_{2,2}, \dots, X_{2,m}$ denote two independent random samples drawn from

Weibull distributions with scale parameters α_1 and α_2 and shape parameters β_1 and β_2 , respectively. Let λ_1 and λ_2 represent the percentile for two populations. Thus, the difference between these percentiles is defined as

$$\hat{\theta} = \hat{\alpha}_1 \left[\ln \left(\frac{1}{1-p_1} \right) \right]^{1/\hat{\beta}_1} - \hat{\alpha}_2 \left[\ln \left(\frac{1}{1-p_2} \right) \right]^{1/\hat{\beta}_2}. \quad (8)$$

2.1 Generalized confidence interval method

The generalized confidence interval (GCI), introduced by, [14], is based on the concept of a generalized pivotal quantity (GPQ). Let $X = (X_1, X_2, \dots, X_n)$ denote a random sample from a distribution depending on a parameter of interest φ and a nuisance parameter γ , and let $x = (x_1, x_2, \dots, x_n)$ represent the observed sample. The GPQ $R(X; x, \varphi, \omega)$ is defined to satisfy two essential conditions: (i) for a fixed observation x , the distribution of $R(X; x, \varphi, \omega)$ is free of unknown parameters; and (ii) the observed value $r(X; x, \varphi, \omega)$ obtained by substituting $X = x$ does not depend on any nuisance parameters. Based on this pivotal structure, a $100(1 - \gamma)\%$ two-sided GCI for φ is given by $[R_\varphi(\gamma/2), R_\varphi(1 - \gamma/2)]$, where $R_\varphi(\gamma/2)$ denotes the $100(\gamma/2)$ th percentile of $R_\varphi(X; x)$.

For the Weibull distribution, [15] showed that the distributions of $\frac{\hat{\beta}}{\beta}$ and $\hat{\beta} \ln \left(\frac{\hat{\alpha}}{\alpha} \right)$ do not depend on the scale–shape parameters α and β . Let $\hat{\alpha}^*$ and $\hat{\beta}^*$ denote the MLEs obtained from a sample of size n drawn from a *Weibull*(1, 1) distribution. Then it follows that $\frac{\hat{\beta}}{\beta} \sim \hat{\beta}^*$ and $\hat{\beta} \ln \left(\frac{\hat{\alpha}}{\alpha} \right) \sim \hat{\beta}^* \ln(\hat{\alpha}^*)$, both of which are free of unknown parameters and therefore serve as pivotal quantities. [16] further derived the GPQs for the scale and shape parameters of a Weibull distribution. Let $\hat{\alpha}_0$ and $\hat{\beta}_0$ be the observed MLEs of $\hat{\alpha}$ and $\hat{\beta}$. Then the GPQs for the Weibull parameters are

$$R_\beta = \frac{\beta}{\hat{\beta}} \hat{\beta}_0 = \frac{\hat{\beta}_0}{\hat{\beta}^*}, \quad (9)$$

and

$$R_\alpha = \left(\frac{\alpha}{\hat{\alpha}} \right)^{\frac{\hat{\beta}}{\hat{\beta}_0}} \hat{\alpha}_0 = \left(\frac{1}{\hat{\alpha}^*} \right)^{\frac{\hat{\beta}^*}{\hat{\beta}_0}} \hat{\alpha}_0. \quad (10)$$

For two independent samples X_1 and X_2 , the GPQs for their respective shape parameters β_1 and β_2 are

$$R_{\beta_1} = \frac{\hat{\beta}_{1,0}}{\hat{\beta}_1^*}, \quad (11)$$

and

$$R_{\beta_2} = \frac{\hat{\beta}_{2,0}}{\hat{\beta}_2^*}, \quad (12)$$

where $\hat{\beta}_{1,0}$ and $\hat{\beta}_{2,0}$ are the observed MLEs, and $\hat{\beta}_1^*$ and $\hat{\beta}_2^*$ denote the corresponding MLEs obtained from simulated samples drawn from a *Weibull*(1, 1) distribution. Similarly, the GPQs for the scale parameters are

$$R_{\alpha_1} = \left(\frac{1}{\hat{\alpha}_1^*} \right)^{\frac{\hat{\beta}_1^*}{\hat{\beta}_{1,0}}} \hat{\alpha}_{1,0}, \quad (13)$$

and

$$R_{\alpha_2} = \left(\frac{1}{\hat{\alpha}_2^*} \right)^{\frac{\hat{\beta}_2^*}{\hat{\beta}_{2,0}}} \hat{\alpha}_{2,0}. \quad (14)$$

Since the p th percentile of Weibull distribution is defined as Eq. (5). The GPQ for λ can be obtained by substituting the GPQs R_α and R_β into the functional form of Eq. (6). Therefore, the GPQ for each percentile is

$$R_{\lambda_i} = R_{\alpha_i} \left[\ln \left(\frac{1}{1-p_i} \right) \right]^{1/R_{\beta_i}}, \quad (15)$$

which is free of nuisance parameters. A measure of practical interest is the difference between percentiles of two Weibull populations. The corresponding GPQ is naturally obtained by

$$R_\theta = R_{\lambda_1} - R_{\lambda_2}, \quad (16)$$

Consequently, a $100(1 - \gamma)\%$ generalized CI for θ can be obtained directly from the empirical $\gamma/2$ and $(1 - \gamma/2)$ quantiles of the simulated distribution of R_θ , which is

$$CI_{gci,\theta} = [L_{gci,\theta}, U_{gci,\theta}]. \quad (17)$$

2.2 Bayesian method with gamma prior distribution

Consider two samples $X_1 = (X_{1,1}, \dots, X_{1,n})$ and $X_2 = (X_{2,1}, \dots, X_{2,m})$ arising from Weibull distributions with parameter vectors (α_1, β_1) and (α_2, β_2) , respectively. For Bayesian inference, the posterior distribution is constructed according to

$$\text{posterior} \propto \text{prior} \times \text{likelihood}.$$

Because the Weibull parameters are strictly positive, Gamma priors are assigned to the scale and shape parameters of each population: $\alpha_i \sim \text{Gamma}(v_{2i}, z_{2i})$ and $\beta_i \sim \text{Gamma}(v_{1i}, z_{1i})$, $i = 1, 2$.

Given the sample X_i , the joint posterior density of (α_i, β_i) is expressed as

$$\pi(\alpha_i, \beta_i | X_i) \propto \pi(\alpha_i) \pi(\beta_i) L(\alpha_i, \beta_i | X_i), \quad (18)$$

where the likelihood function for the i th Weibull sample is

$$L(\alpha_i, \beta_i | x_i) = \prod_{j=1}^{n_i} \left[\frac{\beta_i}{\alpha_i} \left(\frac{x_{i,j}}{\alpha_i} \right)^{\beta_i-1} \exp \left\{ - \left(\frac{x_{i,j}}{\alpha_i} \right)^{\beta_i} \right\} \right]. \quad (19)$$

Since the posterior distributions do not admit closed-form solutions, estimation must rely on Markov Chain Monte Carlo (MCMC) techniques. In this study, Gibbs sampling, [17], implemented through the OpenBUGS software is used to approximate the joint posterior distributions of (α_i, β_i) for $i = 1, 2$. Posterior summaries are obtained based on Monte Carlo samples after convergence of the Markov chains. Posterior samples of θ_i are generated directly from the MCMC draws of (α_i, β_i) . The $100(1 - \gamma)\%$ Bayesian credible interval for θ is

$$CI_{Baye, \theta} = [L_{Baye, \theta}, U_{Baye, \theta}], \quad (20)$$

where L_{Baye, θ_i} and U_{Baye, θ_i} denote the equal-tailed bounds and the highest posterior density (HPD) limits obtained from the MCMC sample.

The HPD intervals are computed using the HDInterval package in R, which determines the shortest interval containing $100(1 - \gamma)\%$ of the posterior probability mass. This ensures that all values inside the interval have higher posterior density than those outside, thus capturing the most credible region of the posterior distribution, [18].

2.3 Bootstrap method

The bootstrap method is a nonparametric resampling technique used to approximate the sampling distribution of a statistic by repeatedly resampling the observed dataset with replacement. Consider two independent samples $X_1 = (X_{1,1}, \dots, X_{1,n})$ and $X_2 = (X_{2,1}, \dots, X_{2,m})$ drawn from Weibull distributions with parameter vectors (α_1, β_1) and (α_2, β_2) , respectively. Bootstrap samples X_1^{*b} and X_2^{*b} are generated by sampling with replacement from X_1 and X_2 , each of size n_1 and n_2 , for $b = 1, \dots, B$ iterations.

For each bootstrap replicate, the Weibull MLEs $\hat{\alpha}_i^{*b}$ and $\hat{\beta}_i^{*b}$ are computed. To assess differences between the two Weibull populations, Bootstrap replicates of this difference are obtained as

$$\theta^{*b} = \hat{\lambda}_1^{*b} - \hat{\lambda}_2^{*b}, \quad b = 1, \dots, B.$$

The resulting bootstrap CI $100(1 - \gamma)\%$ for θ is

$$CI_{BS, \theta} = [\theta^*(\gamma/2), \theta^*(1 - \gamma/2)], \quad (21)$$

where $\theta^*(b)$ denotes the empirical p th percentile of the bootstrap sample $\{\theta^{*1}, \dots, \theta^{*B}\}$.

The empirical distribution of θ^{*b} is then used to form the Bootstrap-HPD interval for θ , defined as

$$CI_{BS-HPD, \theta} = [L_{BS-HPD, \theta}, U_{BS-HPD, \theta}], \quad (22)$$

where the HPD limits correspond to the shortest interval containing $100(1 - \gamma)\%$ of the bootstrap distribution of θ^{*b} . This HPD-based bootstrap approach ensures that the resulting interval reflects the region of highest density within the empirical bootstrap distribution, thereby providing an efficient and data-driven CI for the percentile difference.

3 Simulation Results

This study conducts a Monte Carlo simulation using the R programming to construct CIs for the difference between percentiles of two populations. Several methods are considered, including the generalized CI (GCI), the Bayesian approach based on Markov chain Monte Carlo (Baye-MCMC), the Bayesian highest posterior density (Baye-HPD) interval, the bootstrap percentile method, and the bootstrap highest density (BS-HPD) interval. The optimal method is identified as the one that achieves a CP greater than or equal to the nominal confidence level of 0.95 while simultaneously yielding the shortest average interval width. In addition, the Monte Carlo standard error (MCSE) was employed as an additional criterion for evaluation.

The simulation settings are specified as follows. The scale parameters are fixed and identical for both populations, with $(\alpha_1, \alpha_2) = (2, 2)$. The shape parameters are set to $(\beta_1, \beta_2) = (2, 0.5), (2, 0.75), (2, 1)$, and $(2, 2)$. The probabilities are set as $(p_1, p_2) = 0.5$. The sample size combinations considered are $(n, m) = (20, 20), (20, 50), (50, 50), (50, 100)$, and $(100, 100)$. For Bayesian method, all hyperparameters of the Gamma prior are specified as 0.1.

For each scenario, 5,000 replications are generated, together with 2,500 pivotal quantities for the GCI. Moreover, 500 bootstrap resamples are employed, and 20,000 MCMC samples are generated for the Bayesian methods using the Gibbs sampler, with the first 1,000 iterations discarded as burn-in. To assess the convergence of the MCMC procedure, convergence diagnostics are performed. The effective sample sizes (ESS) are computed to evaluate the amount of independent information contained in the MCMC samples. The results indicate that the ESS values are sufficiently large relative to the total number of generated samples (20,000), suggesting that the posterior estimates are reliable. In addition, the Gelman-Rubin diagnostic statistic is calculated to assess the convergence across multiple chains. The

Gelman-Rubin statistic is equal to 1.01, which is close to 1, indicating satisfactory convergence of the MCMC chains. For the bootstrap method, the value was set to 500, as alternative values were examined and found to produce similar results. Because increasing the number of bootstrap resamples beyond 500 resulted in negligible changes in the estimated CPs and interval lengths.

The empirical CPs and average interval widths of the CIs are presented in Table A.1 (Appendix), as well as Figure 1 and Figure 2 respectively. The results show that both the GCI and the Bayesian methods achieve CPs close to the nominal level of 0.95 in many cases. However, when considered jointly with the average interval widths, the Baye-HPD method consistently yields shorter intervals than the GCI, and thus can be regarded as the most prominent method overall. In particular, the Baye-HPD method performs best in the following cases: $(\beta_1, \beta_2) = (2, 0.5)$ for all sample size combinations; $(\beta_1, \beta_2) = (2, 0.75)$ for $(n, m) = (20, 20)$ and $(50, 50)$; $(\beta_1, \beta_2) = (2, 1)$ for $(n, m) = (20, 20)$ and $(100, 100)$; and $(\beta_1, \beta_2) = (2, 2)$ for $(n, m) = (100, 100)$.

Although the bootstrap methods produce the shortest average interval widths, their empirical CPs remain noticeably below the nominal confidence level of 0.95 across most settings. Nevertheless, the CPs of the bootstrap tend to approach 0.95 as (n, m) increase. To further illustrate the performance of the proposed methods, the MCSE was additionally considered and is presented in Table A.2 (Appendix). The results show that the Bayesian-HPD method consistently produces the smallest MCSE, indicating superior efficiency relative to the other methods.

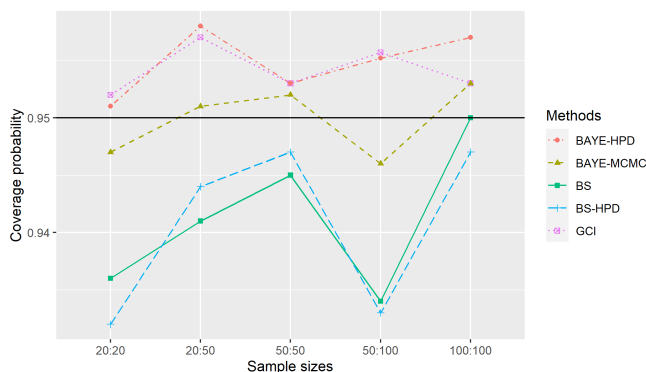


Fig. 1: The performance of the proposed approaches in terms of CP

Source: Created by the authors.

4 Applications

This study aims to propose CI procedures for the difference between percentiles of the Weibull distribu-

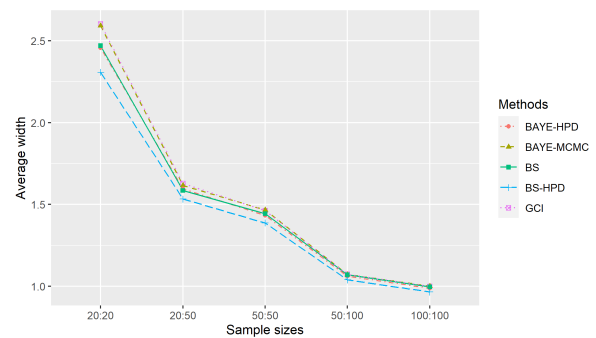


Fig. 2: The performance of the proposed approaches in terms of average interval width

Source: Created by the authors.

tions. To illustrate the proposed methods, wind speed data from Songkhla and Chumphon provinces, located in southern Thailand, are analyzed. Southern Thailand is a coastal region that is frequently exposed to strong winds associated with tropical storms, and it also experiences heavy rainfall and flooding on an annual basis. Higher wind speeds are associated with a greater potential for damage in affected areas. Additionally, another dataset was used to examine the differences in wind speed between surface areas as well as sea and coastal areas. The data were obtained from the Thai Meteorological Department, and the corresponding summary statistics are presented as follows.

4.1 Wind speed data from Chumphon and Songkhla provinces

Based on the data in Table 1, the estimated difference in percentiles, denoted by θ , is equal to 0.7033. The suitability of the distributional assumption for both datasets was assessed using the Akaike Information Criterion (AIC). As shown in Table 2, the Weibull distribution yields the smallest AIC values for both datasets, indicating the best fit among the competing models. Based on the given data, the Anderson-Darling (AD) and Kolmogorov-Smirnov (KS) goodness-of-fit tests were used to examine the distributional assumption. The p-values of the AD and KS statistics for Chumphon Province are 0.811 and 0.804, respectively, while those for Songkhla Province are 0.682 and 0.484, respectively. In addition, the Q-Q plots presented in Figure 3 and Figure 4 further demonstrate that the Weibull distribution provides an adequate fit to the observed data. Table 3 presents the 95% CI estimates for θ obtained using all five methods, which are consistent with and provide further support for the simulation results.

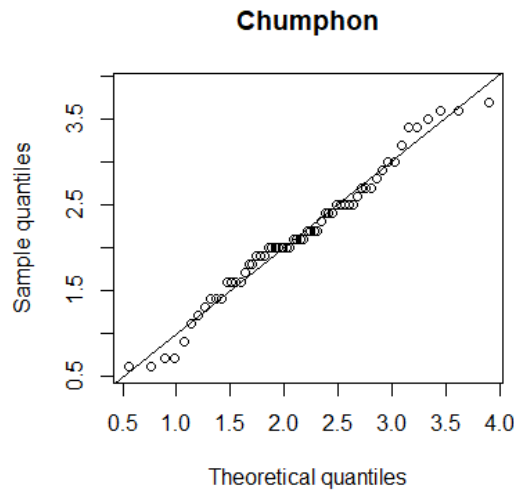


Fig. 3: Weibull Q-Q plot of wind speed measurements in Chumphon Province, comparing empirical quantiles with theoretical Weibull quantiles
Source: Created by the authors.

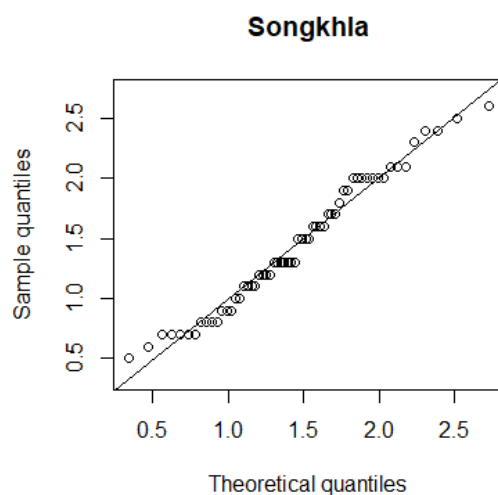


Fig. 4: Weibull Q-Q plot of wind speed measurements in Songkhla Province, comparing empirical quantiles with theoretical Weibull quantiles
Source: Created by the authors.

Table 1: Summary statistics and estimated Weibull parameters for wind speed data from Chumphon and Songkhla provinces

Province	n	α	β
Chumphon	60	2.3964	3.1376
Songkhla	60	1.6187	2.9376

Source: Created by the authors.

Table 2: Akaike Information Criterion (AIC) values for several distributions fitted to wind speed data Chumphon and Songkhla provinces

Distribution	Chumphon	Songkhla
Weibull	139.7312	97.63504
Gamma	145.5709	98.16047
Log-normal	153.1028	100.36597
Exponential	213.5767	165.75717
Cauchy	157.6972	131.49684

Source: Created by the authors.

Table 3: (Example 1) 95% CI estimates for θ obtained using all methods

Methods	Lower	Upper	Length
GCI	0.2804	1.0931	0.8127
Baye-MCMC	0.4330	0.9661	0.5331
Baye-HPD	0.4383	0.9703	0.5320
BS	-0.1830	0.7331	0.9161
BS-HPD	-0.2087	0.6209	0.8296

Source: Created by the authors.

4.2 Wind speed data from Surface area and Sea and coastal area

Example 2 presents wind speed data from surface and coastal areas. The summary statistics are reported in Table 4. The estimated difference in percentiles is -1.7961 . The Weibull Q-Q plots, shown in Figure 5 and Figure 6, along with the Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit tests, were used to assess the distributional assumption. The p-values of the AD and KS tests for Surface area are 0.181 and 0.271, respectively, while those for Songkhla Province are 0.101 and 0.058, respectively. These results indicate no evidence to reject the Weibull distribution for both datasets. In addition, the AIC values (Table 5) provide further support that the Weibull distribution offers an adequate fit to the observed data. Table 6 presents the 95% CI estimates for θ obtained using all five methods. The results are consistent with the simulation findings, particularly for small sample sizes, and further confirm the reliability of the proposed approaches.

Table 4: Summary statistics and estimated Weibull parameters for wind speed data from Surface area and Sea and coastal area

Area	n	α	β
Surface	24	2.5736	2.7015
Sea and coastal	36	4.3482	5.0395

Source: Created by the authors.

Table 5: Akaike Information Criterion (AIC) values for several distributions fitted to wind speed data from Surface area and Sea and coastal area

Distribution	Surface area	Sea and coastal area
Weibull	67.4324	105.1392
Gamma	68.3245	118.2662
Log-normal	69.5641	124.4380
Exponential	89.5692	173.4727
Cauchy	68.7864	108.9670

Source: Created by the authors.

Table 6: (Example 2) 95% CI estimates for θ obtained using all methods

Methods	Lower	Upper	Length
GCI	-2.3150	-1.2715	1.0435
Baye-MCMC	-2.3350	-1.3190	1.0160
Baye-HPD	-2.3360	-1.3217	1.0183
BS	-2.3244	-1.2402	1.0842
BS-HPD	-2.2944	-1.2300	1.0644

Source: Created by the authors.

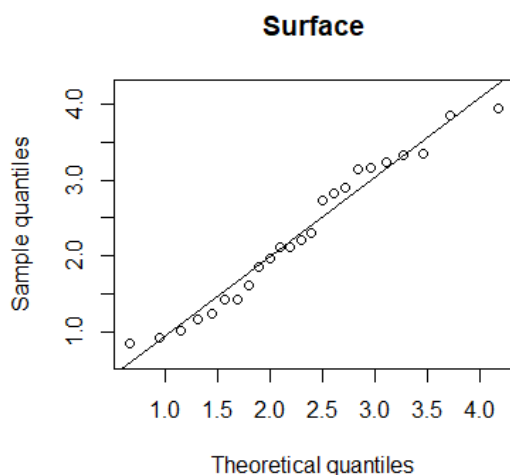


Fig. 5: Weibull Q-Q plot of wind speed measurements in Surface area, comparing empirical quantiles with theoretical Weibull quantiles

Source: Created by the authors.

5 Discussion and Conclusion

Confidence intervals for the difference between percentiles were estimated using the GCI method, the Bayesian approach, the bootstrap approach, and the HPD approach based on the Bayesian as well as the bootstrap. The performance of the CI is evaluated

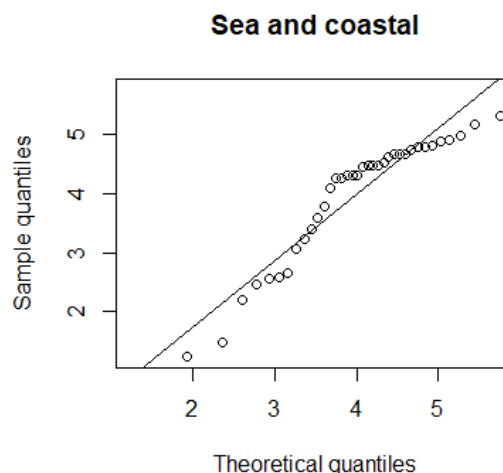


Fig. 6: Weibull Q-Q plot of wind speed measurements in Sea and coastal area, comparing empirical quantiles with theoretical Weibull quantiles

Source: Created by the authors.

based on CP and average interval width to determine the precision and efficiency of each approach. Findings from the results indicated that the Baye-HPD interval was the most effective. However, the GCI method and the Baye-MCMC can be the alternative methods. In many cases, both methods attained CPs close to 0.95. However, in terms of average interval width, they yielded wider intervals compared to the Bayesian-HPD method.

The effect of sample size is also evident in the simulation results. As the sample size increases, the average interval widths decrease across all methods. Although the bootstrap-based methods tend to produce the shortest intervals, their CPs are generally below 0.95 for small to moderate sample sizes. This behavior may be attributed to the skewness of Weibull. In such settings, standard bootstrap intervals may not adequately correct for bias. More advanced bootstrap techniques, such as the bias-corrected and accelerated method, could potentially improve coverage performance and represent a promising direction for future research. Furthermore, numerical stability may become a concern in scenarios with small sample sizes and challenging parameter configurations. Under such conditions, the estimation of Weibull parameters may exhibit increased variability, which can affect the reliability of resulting interval estimates.

References:

- [1] Mafart, P.; Couvert, O.; Gaillard, S.; Leguérinel, I., On calculating sterility in thermal preservation methods: Application of the Weibull frequency

- distribution model, *International Journal of Food Microbiology*, Vol. 72, No. 1-2, 2002, pp. 107-113, doi: 10.1016/S0168-1605(01)00624-9.
- [2] Chauhan, A.; Saini, R. P., Statistical analysis of wind speed data using Weibull distribution parameters, In *Proceedings of the International Conference on Non-Conventional Energy (ICONCE 2014)*, Kalyani, India, 2014, pp. 160-163, doi: 10.1109/ICONCE.2014.6808712.
- [3] Kidmo, D. K.; Danwe, R.; Doka, S. Y.; Djongyang, N., Statistical analysis of wind speed distribution based on six Weibull methods for wind power evaluation in Garoua, Cameroon, *Journal of Renewable Energies*, Vol. 18, No. 1, 2015, pp. 105-125, doi: 10.54966/jreen.v18i1.491.
- [4] Bidaoui, H.; El Abbassi, I.; El Bouardi, A.; Darcherif, A., Wind speed data analysis using Weibull and Rayleigh distribution functions: Case study of five cities in northern Morocco, *Procedia Manufacturing*, Vol. 32, 2019, pp. 786-793, doi: 10.1016/j.promfg.2019.02.286.
- [5] Abubakar, H.; Danrimi, M. L., A simulation study on the insurance claims distribution using Weibull distribution, *Economic Analysis Letters*, Vol. 2, No. 3, 2023, pp. 67-79, Available at: <https://ideas.repec.org/a/bba/j00004/v2y2023i3p67-79d161.html> (accessed 17 November 2025).
- [6] Arrieta, C.; Souza, R. C.; Calili, R. F., Confidence interval for the 100p-th percentile for measurement error distributions, *Journal of Physics: Conference Series*, Vol. 1044, No. 1, 2018, pp. 012016, doi: 10.1088/1742-6596/1044/1/012016.
- [7] Zhang, Z.; Nadarajah, S., On confidence interval estimation of normal percentiles, *Japanese Journal of Statistics and Data Science*, Vol. 1, 2018, pp. 373-391, doi: 10.1007/s42081-018-0020-8.
- [8] Wang, C.; Marriott, P.; Li, P., A note on the coverage behaviour of bootstrap percentile confidence intervals for constrained parameters, *Metrika*, Vol. 85, 2022, pp. 809-831, doi: 10.1007/s00184-021-00851-0.
- [9] Hasan, M. S.; Krishnamoorthy, K., Confidence intervals for the mean and a percentile based on zero-inflated lognormal data, *Journal of Statistical Computation and Simulation*, Vol. 88, No. 8, 2018, pp. 1499-1514, doi: 10.1080/00949655.2018.1439033.
- [10] Thangjai, W.; Niwitpong, S. A.; Niwitpong, S.; Somkhuean, R., Confidence intervals for function of percentiles of Birnbaum–Saunders distributions containing zero values with application to wind speed modelling, *Modelling*, Vol. 6, No. 1, 2025, pp. 16, doi: 10.3390/modelling6010016
- [11] Cohen, A. C., Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples, *Technometrics*, Vol. 7, No. 4, 1965, pp. 579-588, doi: 10.1080/00401706.1965.10490300.
- [12] Menon, M., Estimation of the shape and scale parameters of the Weibull distribution, *Technometrics*, Vol. 5, No. 2, 1963, pp. 175-182, Available at: <https://www.tandfonline.com/doi/abs/10.1080/00401706.1963.10490073> (accessed 10 November 2025).
- [13] Birgören, B., Estimating confidence lower bounds for Weibull percentiles, *Journal of Materials Science Letters*, Vol. 22, No. 16, 2003, pp. 1121-1124, Available at: 10.1023/A:1025162624698 (accessed 10 November 2025).
- [14] Weerahandi, S., Generalized confidence intervals, *Journal of the American Statistical Association*, Vol. 88, No. 423, 1993, pp. 899-905, doi: 10.1080/01621459.1993.10476355.
- [15] Thoman, D. R.; Bain, L. J.; Antle, C. E., Inferences on the parameters of the Weibull distribution, *Technometrics*, Vol. 11, No. 3, 1969, pp. 445-460, doi: 10.1080/00401706.1969.10490706.
- [16] Krishnamoorthy, K.; Mukherjee, S.; Guo, H., Inference on reliability in two-parameter exponential stress–strength model, *Metrika*, Vol. 65, 2007, pp. 261-270, doi: 10.1007/s00184-006-0074-7.
- [17] Geman, S.; Geman, D., Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, 1984, pp. 721-741, doi: 10.1109/TPAMI.1984.4767596.
- [18] Kruschke, J. K., Chapter 12: Bayesian approaches to testing a point (“null”) hypothesis, in *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*, 2nd ed., Academic Press, Boston, MA, USA, 2015, ISBN: 978-0-12-405888-0, Available at: <https://books.google.co.th/books?id=FzvLAWAAQBAJ> (accessed 12 November 2025).

APPENDIX

Table A.1: Coverage probabilities (average interval widths in parentheses) of 95% CIs for the differences between percentiles of Weibull distributions

(n, m)	(β_1, β_2)	GCI	Baye-MCMC	Baye-HPD	BS	BS-HPD
(20,20)	(2,0.50)	0.9520 (2.6044)	0.9470 (2.5924)	0.9510 (2.4581)	0.9360 (2.4702)	0.9320 (2.3073)
	(2,0.75)	0.9520 (2.0648)	0.9470 (2.0349)	0.9520 (1.9933)	0.9350 (1.9602)	0.9350 (1.8867)
	(2,1)	0.9550 (1.7774)	0.9520 (1.7416)	0.9530 (1.7258)	0.9430 (1.6742)	0.9390 (1.6298)
	(2,2)	0.9550 (1.2506)	0.9490 (1.2151)	0.9460 (1.2120)	0.9350 (1.1744)	0.9320 (1.1501)
(20,50)	(2,0.50)	0.9570 (1.6281)	0.9510 (1.6166)	0.9580 (1.5952)	0.9410 (1.5857)	0.9440 (1.5331)
	(2,0.75)	0.9490 (1.4313)	0.9460 (1.4125)	0.9450 (1.4058)	0.9370 (1.3895)	0.9330 (1.3565)
	(2,1)	0.9380 (1.2683)	0.9350 (1.2476)	0.9320 (1.2442)	0.9310 (1.2251)	0.9240 (1.1992)
	(2,2)	0.9530 (1.0469)	0.9470 (1.0215)	0.9480 (1.0188)	0.9400 (0.9926)	0.9310 (0.9717)
(50,50)	(2,0.50)	0.9530 (1.4627)	0.9520 (1.4662)	0.9530 (1.4334)	0.9450 (1.4441)	0.9470 (1.3868)
	(2,0.75)	0.9580 (1.2300)	0.9560 (1.2257)	0.9610 (1.2149)	0.9470 (1.2101)	0.9450 (1.1782)
	(2,1)	0.9380 (1.0770)	0.9390 (1.0695)	0.9360 (1.0647)	0.9300 (1.0558)	0.9210 (1.0329)
	(2,2)	0.9370 (0.7768)	0.9370 (0.7686)	0.9340 (0.7668)	0.9270 (0.7610)	0.9140 (0.7451)
(50,100)	(2,0.50)	0.9557 (1.0730)	0.9460 (1.0714)	0.9552 (1.0621)	0.9340 (1.0698)	0.9330 (1.0391)
	(2,0.75)	0.9330 (0.9420)	0.9360 (0.9383)	0.9390 (0.9346)	0.9320 (0.9312)	0.9260 (0.9107)
	(2,1)	0.9460 (0.8466)	0.9440 (0.8416)	0.9420 (0.8393)	0.9410 (0.8350)	0.9320 (0.8176)
	(2,2)	0.9500 (0.6713)	0.9450 (0.6647)	0.9440 (0.6632)	0.9430 (0.6584)	0.9330 (0.6447)
(100,100)	(2,0.50)	0.9530 (0.9998)	0.9530 (1.0006)	0.9570 (0.9878)	0.9500 (0.9962)	0.9470 (0.9647)
	(2,0.75)	0.9480 (0.8542)	0.9490 (0.8531)	0.9460 (0.8487)	0.9420 (0.8474)	0.9330 (0.8278)
	(2,1)	0.9510 (0.7517)	0.9500 (0.7483)	0.9500 (0.7459)	0.9490 (0.7440)	0.9410 (0.7284)
	(2,2)	0.9560 (0.5453)	0.9540 (0.5420)	0.9560 (0.5408)	0.9500 (0.5408)	0.9500 (0.5296)

Source: Created by the authors.

Table A.2: Monte Carlo standard errors (MCSE) of the CPs based on 5000 simulation runs

(n, m)	(β_1, β_2)	GCI	Baye-MCMC	Baye-HPD	BS	BS-HPD
(20,20)	(2,0.50)	0.0030	0.0032	0.0030	0.0035	0.0036
	(2,0.75)	0.0030	0.0032	0.0030	0.0035	0.0035
	(2,1)	0.0029	0.0030	0.0030	0.0033	0.0034
	(2,2)	0.0029	0.0031	0.0032	0.0035	0.0036
(20,50)	(2,0.50)	0.0029	0.0030	0.0028	0.0033	0.0033
	(2,0.75)	0.0031	0.0032	0.0032	0.0034	0.0036
	(2,1)	0.0034	0.0035	0.0036	0.0036	0.0037
	(2,2)	0.0030	0.0032	0.0031	0.0034	0.0036
(50,50)	(2,0.50)	0.0030	0.0030	0.0030	0.0032	0.0032
	(2,0.75)	0.0028	0.0029	0.0027	0.0032	0.0032
	(2,1)	0.0034	0.0034	0.0035	0.0036	0.0038
	(2,2)	0.0034	0.0034	0.0035	0.0037	0.0040
(50,100)	(2,0.50)	0.0029	0.0032	0.0029	0.0035	0.0036
	(2,0.75)	0.0036	0.0035	0.0034	0.0036	0.0037
	(2,1)	0.0032	0.0033	0.0033	0.0033	0.0036
	(2,2)	0.0031	0.0032	0.0033	0.0033	0.0036
(100,100)	(2,0.50)	0.0030	0.0030	0.0029	0.0031	0.0032
	(2,0.75)	0.0031	0.0031	0.0032	0.0033	0.0036
	(2,1)	0.0030	0.0031	0.0031	0.0031	0.0033
	(2,2)	0.0029	0.0030	0.0029	0.0031	0.0031

Source: Created by the authors.

Contribution of individual authors to the creation of a scientific article (ghostwriting policy)

Manussaya La-ongkaew carried out the software development, formal analysis, Resource and investigation, and was responsible for writing the original draft as well as reviewing and editing the manuscript. Sa-Aat Niwitpong developed the methodology, curated the data, conducted the formal analysis and investigation, and provided supervision for the study.

Sources of funding for research presented in a scientific article or scientific article itself

This research was funded by Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, Thailand contract no.sci-690036

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US