Int. J. Anal. Appl. (2026), 24:282

International Journal of Analysis and Applications

Classical and Bayesian Estimation of Weibull Distribution Parameters with an Application to Daily Wind Speed Data

Dalia Kamal Alnagar*

Department of Statistics, Faculty of Science, University of Tabuk, Kingdom of Saudi Arabia

**Corresponding author:* d_alnajjar@ut.edu.sa

Abstract. Accurate characterization of the probability distribution of wind speed is essential for preliminary wind resource assessment and renewable energy planning. This paper presents a comparative classical and Bayesian analysis of daily wind speed distributions using 24 years (2001–2024, $n = 8{,}766$ daily observations) of NASA POWER daily mean wind speed data at 10 m above the surface for Tabuk, Saudi Arabia. The two-parameter Weibull distribution, a standard benchmark model in wind-energy studies, is fitted using three classical estimators—maximum likelihood estimation (MLE), the method of moments (MoM), and L-moments—alongside a Bayesian analysis with weakly informative Gamma priors on the shape and scale parameters, in which posterior inference is carried out via a random-walk Metropolis–Hastings Markov chain Monte Carlo (MCMC) algorithm and convergence is assessed using the Gelman–Rubin diagnostic across multiple chains. The MLE and Bayesian posterior-mean estimates of the Weibull parameters agree closely ($\hat{k} = 3.553$, $\hat{\lambda} = 4.235$ m/s), reflecting the dominance of the likelihood in this large sample. A Monte Carlo simulation study further shows that the Bayesian estimator attains a lower mean squared error than the MLE for the shape parameter in small samples ($n \leq 100$). Goodness-of-fit comparisons against Rayleigh, Gamma, and Lognormal alternatives show that the Lognormal and Gamma distributions provide a better formal statistical fit to the Tabuk daily wind speed data than the Weibull distribution, confirming that the Weibull model should not be assumed statistically optimal for every wind regime, even though it remains a useful and interpretable benchmark for wind-energy applications. Using the fitted Weibull model, the mean wind power density at 10 m height is estimated at approximately 43.9 W/m^2, indicating a low near-surface wind resource. The results provide a statistically grounded characterization of the Tabuk wind regime, illustrate the practical trade-offs between classical and Bayesian Weibull estimation, and underscore the importance of benchmarking the Weibull distribution against alternative probability models before drawing conclusions about wind-resource potential.

Received: Jul. 8, 2026.

2020 *Mathematics Subject Classification.* 62F10, 62F15.

Key words and phrases. Weibull distribution; Lognormal distribution; Gamma distribution; Rayleigh distribution; Bayesian inference; maximum likelihood estimation; Markov chain Monte Carlo; wind speed modelling; NASA POWER; wind power density; Tabuk; Saudi Arabia.

https://doi.org/10.28924/2291-8639-24-2026-282
ISSN: 2291-8639

© 2026 the author(s).

1. Introduction

Wind energy is one of the fastest-growing renewable energy sources worldwide, and reliable statistical modelling of wind speed is a prerequisite for preliminary resource assessment, turbine selection, and energy yield prediction. Among the many probability models proposed for wind speed, the two-parameter Weibull distribution has become the de facto standard in the wind energy literature because of its flexibility in capturing a wide range of empirical wind speed shapes, its simple analytical relationship with wind power density, and its parsimony [1,2].

Saudi Arabia has committed to substantial diversification of its energy mix under the Saudi Vision 2030 framework, with wind power identified as a complementary resource to the country's already dominant solar potential, particularly in the northern and north-western regions. Tabuk, located in north-western Saudi Arabia, has been highlighted in national renewable-energy planning owing to its relatively favourable wind and solar conditions compared with other regions of the Kingdom. A statistically rigorous characterization of the local wind speed distribution is therefore of direct practical relevance.

The parameters of the Weibull distribution are conventionally estimated using classical frequentist methods, most commonly maximum likelihood estimation (MLE), the method of moments (MoM), and variants based on L-moments or empirical percentiles [3,4]. These methods provide point estimates and, in the case of MLE, asymptotic standard errors based on the observed Fisher information. An alternative, increasingly popular in engineering and environmental statistics, is Bayesian estimation, which treats the unknown parameters as random variables endowed with prior distributions and updates this prior belief with the observed data likelihood to obtain a posterior distribution [5]. Bayesian methods provide full posterior distributions and credible intervals, allow the incorporation of prior information from related meteorological studies, and may improve finite-sample performance when only limited wind speed records are available.

Although the Weibull distribution is widely used in wind energy applications, it is not necessarily the statistically best-fitting distribution for every wind regime. Alternative probability models such as the Lognormal, Gamma, and Rayleigh distributions are commonly considered in goodness-of-fit comparisons. Therefore, in addition to estimating the Weibull parameters, this study explicitly compares the fitted Weibull model with alternative parametric distributions in order to assess whether the Weibull model is statistically competitive for the Tabuk wind speed data.

The purpose of this paper is threefold. First, we estimate the parameters of the Weibull distribution for the NASA POWER daily mean wind speed series for Tabuk over the period 2001–2024 using three classical estimators, namely MLE, MoM, and L-moments, together with a Bayesian approach implemented via Metropolis–Hastings MCMC. Second, we compare the two estimation paradigms both empirically, using the observed data, and through a Monte Carlo simulation study that evaluates mean squared error (MSE) as a function of sample size. Third, we compare the Weibull distribution with Rayleigh, Gamma, and Lognormal alternatives and translate the fitted

Weibull model into practically relevant wind energy indicators, namely the Weibull mean wind speed and wind power density.

The remainder of the paper is organized as follows. Section 2 reviews related work on Weibull parameter estimation and wind speed distribution modelling. Section 3 describes the data and the classical and Bayesian estimation methodologies. Section 4 presents the empirical results, including descriptive statistics, parameter estimates, goodness-of-fit assessment, comparison with alternative distributions, the simulation study, and the wind energy application. Section 5 concludes the paper and discusses limitations and directions for future research.

2. Related Work

The estimation of Weibull parameters for wind speed data has a long history in the wind engineering literature. Early work proposed systematic methods, including the method of moments, for fitting the Weibull distribution to wind speed frequency distributions for energy applications [1]. A later comparative study examined several classical estimators, including MLE, the modified method of moments, and the graphical least-squares method, and showed that MLE generally provides accurate estimates when high-resolution time-series data are available, whereas the modified moment method performs better with binned frequency data [2]. Subsequent studies have proposed alternative estimators, such as the power density method, the equivalent energy method, and various hybrid procedures, aimed at improving accuracy under specific data conditions [3, 4]. Comparative studies applying several classical estimators to observed wind data generally conclude that no single method dominates uniformly across all sites, thereby motivating continued methodological research [6, 7]. A recent bibliometric review of 96 studies published between 2014 and 2023 confirms this pattern, reporting that maximum likelihood remains the most widely used estimator overall, while several moment-based and empirical methods remain competitive depending on the site characteristics and data resolution [8].

The application of the Weibull distribution to Saudi wind speed data also has a long history. Early Saudi Arabia-focused work estimated Weibull parameters for several stations across the Kingdom and provided some of the first systematic wind resource characterizations for the region [9]. More recent studies have applied the Weibull distribution to characterize wind resources at sites in the northern and north-western regions geographically close to Tabuk. One study assessed wind energy potential at Sharma, Al-Qurayyat, and Sakaka, located in the neighbouring Tabuk and Al-Jouf regions, using multi-year wind speed data and Weibull-based power density estimates [10]. Another study assessed the wind energy potential of Medina, Saudi Arabia, using Weibull distribution parameters estimated from long-term wind speed records, reporting power density estimates broadly comparable to those obtained for other Saudi sites [11]. Similarly, a long-term 21-year analysis of wind characteristics in the adjacent Al-Jouf region reported broadly comparable wind speed magnitudes and low-to-moderate power densities [12]. These findings are consistent with the near-surface wind resource pattern obtained for Tabuk in the present study.

However, these Saudi Arabia-focused studies rely mainly on classical Weibull estimation and do not explicitly incorporate a Bayesian treatment of parameter uncertainty, which represents an important gap in the regional literature.

Bayesian approaches to Weibull estimation have received comparatively less attention in the wind energy literature, despite being well established in reliability engineering and survival analysis, where the Weibull distribution is widely used to model failure times and lifetimes [5]. Bayesian methods are particularly attractive when historical or expert information is available to construct informative priors, or when uncertainty quantification beyond asymptotic normal approximations is required, for example with moderate sample sizes or skewed posterior distributions. Markov chain Monte Carlo (MCMC) methods, particularly the Metropolis–Hastings algorithm, provide a general and flexible computational framework for Bayesian inference when the posterior distribution is not available in closed form, as is the case for the Weibull distribution under standard prior choices [5]. A related methodological comparison contrasted MLE with Bayesian estimators obtained via both the Lindley approximation and Metropolis–Hastings MCMC under exponential priors, and found that the Bayesian estimators achieved lower mean squared error than MLE for small sample sizes [13]. This finding is consistent with the simulation results reported in Section 4.5 of the present study.

Beyond the choice of estimation method for a given distributional family, a separate strand of the wind engineering literature has examined whether the Weibull distribution itself is the most suitable probability model for wind speed, relative to other commonly used continuous distributions. A broad review of probability models used for wind speed characterization and wind energy assessment noted that, although the Weibull distribution remains the most widely applied model, several alternative distributions can outperform it depending on the local wind regime [14]. Consistent with this observation, comparative studies have examined several candidate distributions, including the Weibull, Lognormal, and Gamma distributions, and found that the relative ranking of distributions by goodness-of-fit varies across sites, with no single distribution providing the best fit uniformly [15, 16]. Other studies have gone further and proposed distributions beyond the standard two-parameter families. For example, the four-parameter Johnson S_B distribution has been proposed for offshore wind speed modelling and shown to outperform the Weibull distribution in terms of fit accuracy in some settings [17]. More recent work has also revisited the adequacy of the Weibull distribution for wind speed statistics and proposed extensions capable of capturing empirical features such as multimodality and heavy tails that the standard Weibull model cannot represent [18]. Recent work has also considered extended Weibull-related models as more flexible alternatives to the standard two-parameter Weibull distribution. For example, the Weibull–Rayleigh model has been studied in terms of its statistical properties, parameter estimation, and applications to industrial data [19]. Such developments highlight the continuing interest in constructing flexible Weibull-type families for applied modelling problems. However, the present study deliberately focuses on the standard two-parameter Weibull distribution because of its interpretability, its established role as a benchmark model in wind-energy assessment, and its direct analytical relationship with wind power density.

Taken together, this body of work indicates that, despite its widespread use, the Weibull distribution is not universally the best-fitting model for wind speed data. This conclusion is consistent with the site-specific comparison reported in Section 4.4 of the present study, where the Lognormal and Gamma distributions are found to outperform the Weibull distribution in terms of AIC and BIC for the Tabuk data set. Nevertheless, the Weibull distribution is retained as a benchmark model for wind-energy interpretation because of its practical advantages, especially its interpretability and its direct relationship with wind power density.

Overall, the existing literature shows that Weibull modelling remains a standard and practically interpretable approach for wind speed analysis, but that the relative performance of different estimation methods, and indeed of the Weibull distribution itself relative to other candidate models, depends on sample size, data resolution, and site-specific wind characteristics. The present paper contributes to this literature by comparing classical and Bayesian Weibull estimation using a long daily NASA POWER wind speed data set for Tabuk, Saudi Arabia, by explicitly benchmarking the Weibull distribution against Rayleigh, Gamma, and Lognormal alternatives, and by quantifying, through simulation, the sample-size conditions under which Bayesian estimation may offer finite-sample advantages over maximum likelihood estimation.

3. Data and Methodology

3.1. **Data.** The data used in this study consist of $n = 8{,}766$ daily mean wind speed observations at 10 m above the surface (WS10M) for Tabuk, Saudi Arabia, covering the period from 1 January 2001 to 31 December 2024. The data were obtained from the NASA Prediction of Worldwide Energy Resources (NASA POWER) database, which provides satellite- and reanalysis-based meteorological variables for user-specified geographic locations. The WS10M series was extracted for Tabuk using latitude 28.383°N and longitude 36.566°E. The series contains no missing values and no non-positive observations, and was therefore used without further imputation or adjustment.

Since the observations are derived from a gridded meteorological product rather than from an on-site meteorological mast, the results should be interpreted as a regional statistical characterization of the wind regime rather than as a site-specific turbine feasibility assessment. Figure 1 shows the full daily time series, and Figure 2 presents the monthly distribution of wind speed. The monthly distribution indicates a mild seasonal pattern, with somewhat higher and more variable wind speeds during the spring months. It should be noted that daily wind speed observations may exhibit serial dependence and seasonal variation. The likelihood-based inference is therefore interpreted as a marginal distributional fit under a working independence assumption, rather than as a full stochastic time-series model. Accordingly, the present analysis models the marginal distribution of daily wind speed over the full study period, rather than the dynamic temporal evolution of the wind speed process. To examine the possible influence of seasonality, the empirical monthly distribution of wind speed is reported and discussed. Future work may extend this analysis by estimating separate Weibull models by month or season.

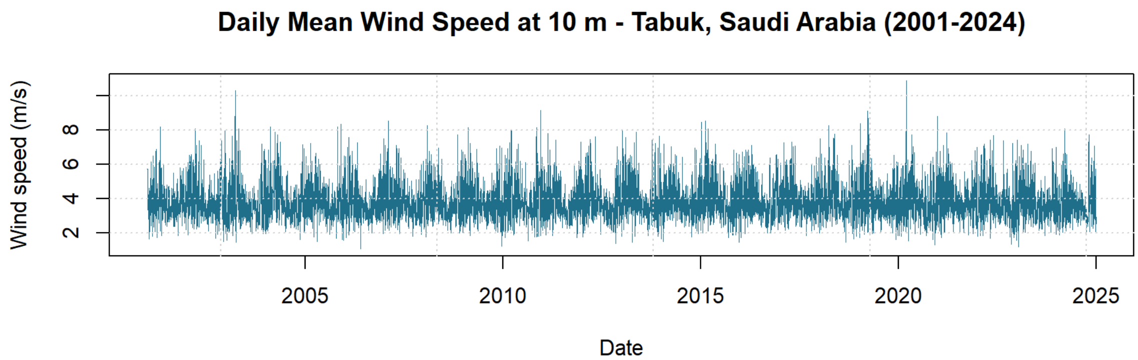

Figure 1. Daily mean wind speed at 10 m height in Tabuk, Saudi Arabia, 2001–2024.

3.2. **The Weibull Distribution.** A random variable $X > 0$ (here, daily mean wind speed) follows a two-parameter Weibull distribution with shape parameter $k > 0$ and scale parameter $\lambda > 0$ if its probability density function (pdf) is

$$f(x \mid k, \lambda) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{x}{\lambda}\right)^{k}\right], \qquad x > 0, \tag{3.1}$$

with corresponding cumulative distribution function (CDF)

$$F(x \mid k, \lambda) = 1 - \exp\left[-\left(\frac{x}{\lambda}\right)^{k}\right], \qquad x > 0. \tag{3.2}$$

The shape parameter k governs the form of the distribution (values near 2 approximate a Rayleigh distribution, commonly used for wind speed, while larger k indicates a more peaked, less variable distribution), and λ is a scale parameter closely related to the mean wind speed. The mean and variance of the Weibull distribution are

$$E(X) = \lambda\,\Gamma\left(1 + \frac{1}{k}\right), \qquad \mathrm{Var}(X) = \lambda^2\left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma\left(1 + \frac{1}{k}\right)^2\right], \tag{3.3}$$

where $\Gamma(\cdot)$ denotes the gamma function.

3.3. **Classical Estimation.**

3.3.1. *Maximum Likelihood Estimation.* For an independent and identically distributed (i.i.d.) sample $x_1, \ldots, x_n$, the log-likelihood function of the Weibull distribution is

$$\ell(k, \lambda) = n \log k - nk \log \lambda + (k-1)\sum_{i=1}^{n} \log x_i - \sum_{i=1}^{n}\left(\frac{x_i}{\lambda}\right)^{k}. \tag{3.4}$$

The MLEs $(\hat{k}_{\mathrm{MLE}}, \hat{\lambda}_{\mathrm{MLE}})$ maximize Equation (3.4) and were obtained here by direct numerical maximization (Nelder–Mead simplex algorithm), with the scale parameter profiled out in closed form given k. Asymptotic standard errors were computed from the inverse of the observed Fisher information matrix (the numerical Hessian of the negative log-likelihood evaluated at the MLE), and 95% confidence intervals were constructed using the normal approximation $\hat{\theta} \pm 1.96\,\mathrm{SE}(\hat{\theta})$.

3.3.2. *Method of Moments.* The method-of-moments (MoM) estimator equates the sample mean and variance to their theoretical counterparts in Equation (3.3). Writing $\widehat{CV} = s/\bar{x}$ for the sample coefficient of variation, $\hat{k}_{\mathrm{MoM}}$ solves

$$\frac{\Gamma(1+2/k)}{\Gamma(1+1/k)^2} - 1 = \widehat{CV}^2, \tag{3.5}$$

via numerical root-finding, after which $\hat{\lambda}_{\mathrm{MoM}} = \bar{x}/\Gamma(1+1/\hat{k}_{\mathrm{MoM}})$.

3.3.3. *L-Moments.* As a further classical benchmark, the L-moment estimator uses the first two sample L-moments, ℓ_1 (equal to the sample mean) and ℓ_2, and the population relationship $\ell_2/\ell_1 = 1 - 2^{-1/k}$ for the Weibull distribution, solved numerically for $\hat{k}_{\mathrm{LMOM}}$, with $\hat{\lambda}_{\mathrm{LMOM}} = \ell_1/\Gamma(1 + 1/\hat{k}_{\mathrm{LMOM}})$.

3.4. **Alternative Distributions.** To formally benchmark the Weibull distribution against other commonly used wind speed models, three additional two- and one-parameter distributions were fitted to the same daily wind speed series: the Rayleigh, Gamma, and Lognormal distributions. All three were fitted by maximum likelihood.

3.4.1. *Rayleigh Distribution.* The Rayleigh distribution is a special case of the Weibull distribution obtained by fixing the shape parameter at $k = 2$. A random variable $X > 0$ follows a Rayleigh distribution with scale parameter $\sigma > 0$ if its pdf is

$$f(x \mid \sigma) = \frac{x}{\sigma^2} \exp\left[-\frac{x^2}{2\sigma^2}\right], \qquad x > 0, \tag{3.6}$$

with corresponding mean and variance

$$E(X) = \sigma\sqrt{\frac{\pi}{2}}, \qquad \mathrm{Var}(X) = \frac{4-\pi}{2}\,\sigma^2. \tag{3.7}$$

The relationship to the Weibull family follows by setting $k = 2$ and $\lambda = \sigma\sqrt{2}$ in Equation (3.1). For an i.i.d. sample $x_1, \ldots, x_n$, the log-likelihood is

$$\ell(\sigma) = \sum_{i=1}^{n} \log x_i - 2n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2, \tag{3.8}$$

which yields the closed-form MLE

$$\hat{\sigma}^2_{\mathrm{MLE}} = \frac{1}{2n}\sum_{i=1}^{n} x_i^2. \tag{3.9}$$

3.4.2. *Gamma Distribution.* A random variable $X > 0$ follows a Gamma distribution with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$ if its pdf is

$$f(x \mid \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\, x^{\alpha-1} e^{-\beta x}, \qquad x > 0, \tag{3.10}$$

with mean and variance

$$E(X) = \frac{\alpha}{\beta}, \qquad \mathrm{Var}(X) = \frac{\alpha}{\beta^2}. \tag{3.11}$$

The log-likelihood for an i.i.d. sample is

$$\ell(\alpha, \beta) = n\alpha \log \beta - n \log \Gamma(\alpha) + (\alpha - 1) \sum_{i=1}^{n} \log x_i - \beta \sum_{i=1}^{n} x_i. \tag{3.12}$$

No closed-form solution exists for $\hat{\alpha}_{\mathrm{MLE}}$. Profiling out β in closed form given α ($\hat{\beta} = \hat{\alpha}/\bar{x}$) reduces the problem to the single nonlinear equation

$$\log \hat{\alpha} - \psi(\hat{\alpha}) = \log \bar{x} - \overline{\log x}, \tag{3.13}$$

where $\psi(\cdot)$ is the digamma function, solved here by numerical root-finding (Newton–Raphson), after which $\hat{\beta}_{\mathrm{MLE}} = \hat{\alpha}_{\mathrm{MLE}}/\bar{x}$.

3.4.3. *Lognormal Distribution.* A random variable $X > 0$ follows a Lognormal distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ if $\log X \sim N(\mu, \sigma^2)$, with pdf

$$f(x \mid \mu, \sigma) = \frac{1}{x\,\sigma\,\sqrt{2\pi}} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right], \qquad x > 0, \tag{3.14}$$

and mean and variance

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right), \qquad \mathrm{Var}(X) = \left[\exp(\sigma^2) - 1\right]\exp(2\mu + \sigma^2). \tag{3.15}$$

The MLEs follow directly from the log-transformed sample and are available in closed form:

$$\hat{\mu}_{\mathrm{MLE}} = \frac{1}{n}\sum_{i=1}^{n} \log x_i, \qquad \hat{\sigma}^2_{\mathrm{MLE}} = \frac{1}{n}\sum_{i=1}^{n} \left(\log x_i - \hat{\mu}_{\mathrm{MLE}}\right)^2. \tag{3.16}$$

3.4.4. *Model Comparison Criteria.* All four fitted distributions (Weibull, Rayleigh, Gamma, and Lognormal) were compared using the maximized log-likelihood, AIC, BIC, and the one-sample Kolmogorov–Smirnov statistic, as described in Section 3.6 below.

3.5. **Bayesian Estimation.** In the Bayesian framework, k and λ are treated as random variables with prior densities $\pi(k)$ and $\pi(\lambda)$. We adopt independent, weakly informative Gamma priors. Throughout this paper, the Gamma distribution is parameterized in terms of shape and rate, so that a Gamma(a, b) prior has mean a/b and variance a/b^2:

$$k \sim \mathrm{Gamma}(a_k, b_k), \qquad \lambda \sim \mathrm{Gamma}(a_\lambda, b_\lambda), \tag{3.17}$$

where the hyperparameters are set to $a_k = 2$, $b_k = 1$, $a_\lambda = 2$, and $b_\lambda = 0.5$. These choices imply prior means of 2 for k and 4 for λ, with prior variances of 2 and 8, respectively, and are centred on plausible wind-speed values while remaining diffuse relative to the information contained in a sample of $n = 8{,}766$ observations.

Combining the priors with the Weibull likelihood, Equation (3.4), via Bayes' theorem yields the joint posterior density

$$\pi(k, \lambda \mid \mathbf{x}) \ \propto \ \pi(k)\, \pi(\lambda)\ \exp\{\ell(k, \lambda)\}, \tag{3.18}$$

which is not available in closed form. We therefore sample from the posterior using a random-walk Metropolis–Hastings (MH) algorithm operating on the log-transformed parameters, $\theta = (\log k, \log \lambda)$, to enforce positivity and to improve the mixing of the chain.

At each iteration, a candidate θ^* is drawn from a bivariate normal proposal centred at the current state, $\theta^* \sim \mathcal{N}(\theta^{(t)}, \mathrm{diag}(\sigma_k^2, \sigma_\lambda^2))$, and accepted with probability

$$\alpha = \min\left\{1,\ \frac{\pi(k^*, \lambda^* \mid \mathbf{x})\, |J(\theta^*)|}{\pi(k^{(t)}, \lambda^{(t)} \mid \mathbf{x})\, |J(\theta^{(t)})|}\right\}, \tag{3.19}$$

where $|J(\theta)| = k\lambda$ is the Jacobian of the log transformation. Three independent chains of 60,000 iterations each were run from dispersed starting values; the first 15,000 iterations of each chain were discarded as burn-in and the remaining draws were thinned by a factor of 5, yielding 9,000 retained draws per chain and 27,000 pooled posterior draws.

Convergence was assessed using the Gelman–Rubin potential scale reduction statistic, $\hat{R}$, computed across the three chains; values close to 1 indicate that the chains have converged to a common stationary distribution [5]. Posterior means, medians, standard deviations, and equal-tailed 95% credible intervals were computed from the pooled post-burn-in, thinned draws.

3.6. **Goodness-of-Fit and Model Comparison.** The fitted Weibull models were compared using the maximized log-likelihood, the Akaike Information Criterion (AIC = $2p - 2\ell$), and the Bayesian Information Criterion (BIC = $p \log n - 2\ell$, where p is the number of estimated parameters). In addition, the one-sample Kolmogorov–Smirnov (KS) statistic and its associated p-value were computed against the fitted distribution function.

To assess whether the Weibull distribution provides a competitive model relative to common alternatives used in wind speed modelling, we also fitted the Rayleigh, Gamma, and Lognormal distributions to the same data set. These distributions were compared using AIC, BIC, and the KS statistic. This comparison is important because formal goodness-of-fit tests may reject the Weibull model in very large samples even when the visual fit is practically acceptable.

3.7. **Simulation Study.** To compare the finite-sample performance of the MLE and the Bayesian posterior-mean estimator independently of the specific Tabuk data set, we conducted a Monte Carlo simulation study. Samples of size $n \in \{30, 50, 100, 200, 500\}$ were generated from a Weibull distribution with parameters fixed at the MLE estimates obtained from the Tabuk data ($k = 3.553$, $\lambda = 4.235$), and, for each sample size, 250 independent replications were drawn. For each replicate, both the MLE and the Bayesian posterior mean (using the same priors as in Equation (3.17) and a shortened MCMC chain for computational efficiency) were computed, and the empirical bias and mean squared error (MSE) of each estimator were evaluated across replications.

4. Results

4.1. Descriptive Statistics.

Table 1 summarizes the descriptive statistics of the daily wind speed series. The mean wind speed over the 24-year period is 3.83 m/s, with a standard deviation of 1.09 m/s. The distribution is moderately right-skewed (skewness ≈ 1.00) with positive excess kurtosis (≈ 1.59), indicating a heavier right tail than the normal distribution. This pattern is consistent with the Weibull distribution being an appropriate candidate model.

Table 1. Descriptive statistics of the daily mean wind speed (m/s) at 10 m height, Tabuk, Saudi Arabia (2001–2024).

Statistic	Value
Number of observations (n)	8766
Mean	3.8295
Standard deviation	1.0889
Variance	1.1856
Coefficient of variation	0.2843
Minimum	1.06
Median	3.64
Maximum	10.89
Skewness	1.0034
Excess kurtosis	1.5911

Figure 2 presents the monthly distribution of daily wind speed. The boxplots indicate a mild seasonal pattern, with generally higher and more variable wind speeds during the spring months, especially from February to April, and relatively lower wind speeds during the summer months.

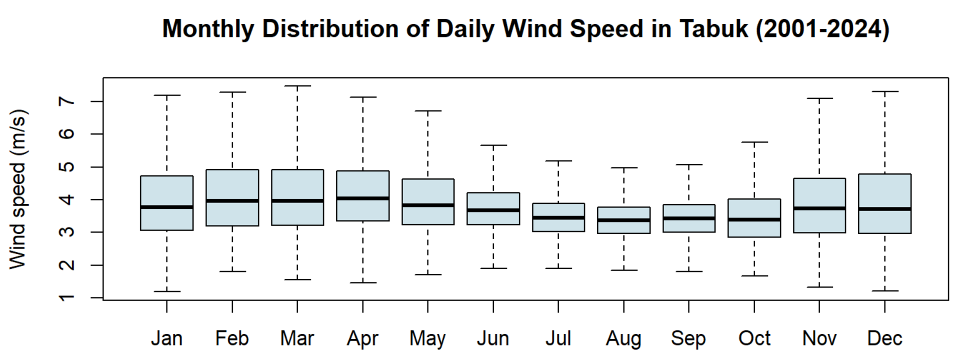

Figure 2. Monthly distribution of daily wind speed in Tabuk, Saudi Arabia (2001–2024).

4.2. **Classical Parameter Estimates.** Table 2 reports the point estimates of k and λ obtained from the three classical methods. The MLE yields $\hat{k} = 3.553$ (SE = 0.027) and $\hat{\lambda} = 4.235$ m/s (SE = 0.014). The method-of-moments and L-moment estimates of the shape parameter are noticeably larger (3.941 and 4.135, respectively), while the scale estimates are very similar across methods, approximately 4.22–4.24 m/s. This pattern, in which the MLE produces a smaller shape estimate than moment-based methods, reflects the differing sensitivity of the three methods to the empirical distribution, particularly in the tails [2].

Table 2. Classical point estimates of the Weibull shape (k) and scale (λ) parameters.

Method	$\hat{k}$	SE($\hat{k}$)	$\hat{\lambda}$	SE($\hat{\lambda}$)
Maximum Likelihood (MLE)	3.5531	0.0267	4.2351	0.0135
Method of Moments (MoM)	3.9411	–	4.2285	–
L-Moments (LMOM)	4.1347	–	4.2170	–

Table 3 reports 95% asymptotic confidence intervals for the MLEs alongside the corresponding 95% Bayesian credible intervals. The two sets of intervals are very close, both in location and width, reflecting the dominance of the likelihood in this large sample.

Table 3. 95% asymptotic confidence intervals (MLE) and 95% credible intervals (Bayesian) for the Weibull parameters.

Parameter	MLE 95% CI	Bayesian 95% CrI
k (shape)	[3.5009, 3.6054]	[3.5011, 3.6047]
λ (scale)	[4.2086, 4.2616]	[4.2090, 4.2618]

4.3. **Bayesian Estimation Results.** Table 4 summarizes the posterior distributions of the Weibull parameters obtained using the Metropolis–Hastings MCMC algorithm. The posterior means, $\hat{k}_{\text{Bayes}} = 3.552$ and $\hat{\lambda}_{\text{Bayes}} = 4.235$ m/s, are essentially identical to the corresponding MLEs. This result is expected given the large sample size ($n = 8{,}766$), because the likelihood dominates the weakly informative priors.

Table 4. Posterior summary of the Weibull parameters obtained via Metropolis–Hastings MCMC.

Parameter	Posterior mean	Posterior median	Posterior SD	$\hat{R}$
k (shape)	3.5524	3.5520	0.0265	1.00029
λ (scale)	4.2352	4.2353	0.0136	0.99996

Figure 3 shows the retained post-burn-in and thinned MCMC draws for k and λ. The chains fluctuate around the MLE values, with no visible trend or drift. The Gelman–Rubin statistics are very close to one for both parameters, indicating convergence across the independently initialized chains.

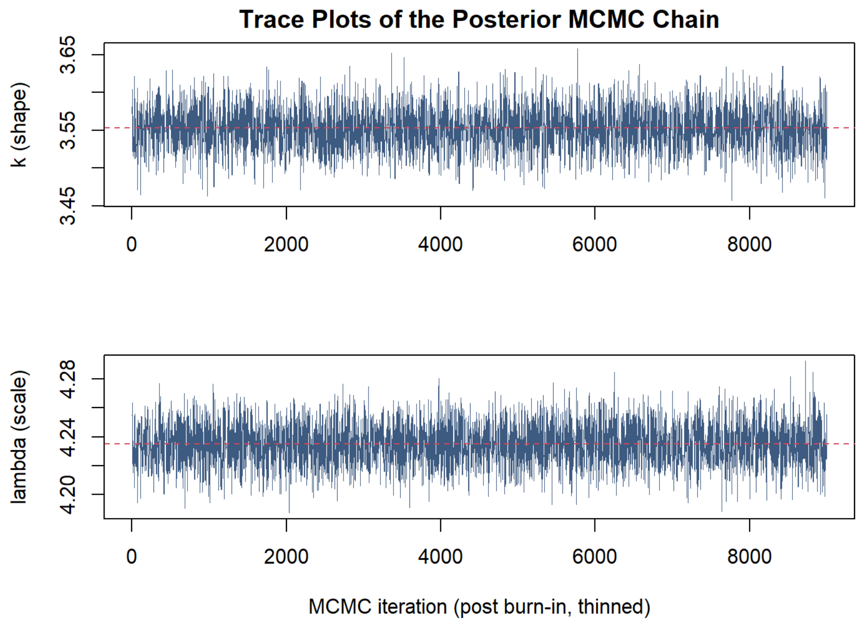

Figure 3. Trace plots of the retained post-burn-in and thinned MCMC draws for the Weibull shape parameter k and scale parameter λ. Dashed lines mark the maximum likelihood estimates.

Figure 4 shows the marginal posterior densities of k and λ together with their 95% credible intervals. Both posterior distributions are approximately symmetric and unimodal, consistent with the large-sample normal approximation underlying the MLE confidence intervals.

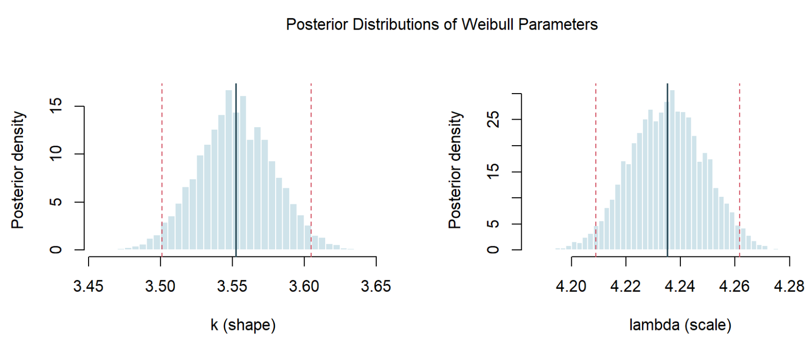

Figure 4. Marginal posterior distributions of the Weibull shape parameter k and scale parameter λ, with posterior means and 95% credible intervals.

4.4. **Goodness of Fit.** Figure 5 overlays the fitted Weibull densities, based on the MLE and Bayesian posterior mean estimates, on the histogram of the observed wind speed data. The two fitted densities are visually almost indistinguishable, reflecting the close agreement between the MLE and Bayesian estimates. Overall, the Weibull model provides a reasonable approximation to the bulk of the empirical distribution, although some departures are visible, particularly in the tails.

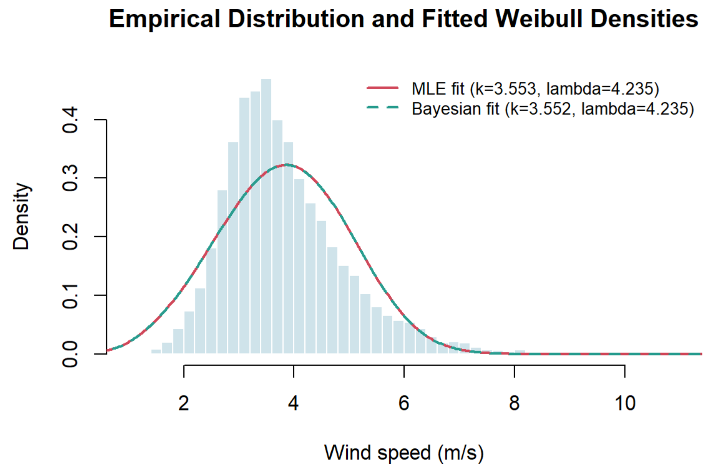

FIGURE 5. Histogram of the observed daily wind speed with fitted Weibull densities obtained from the MLE and Bayesian posterior mean estimates.

Figure 6 compares the empirical and fitted CDFs, while Figure 7 presents the corresponding quantile–quantile (Q–Q) plot. The Q–Q plot indicates close agreement over the central part of the distribution, with mild to moderate departures in the lower and upper tails.

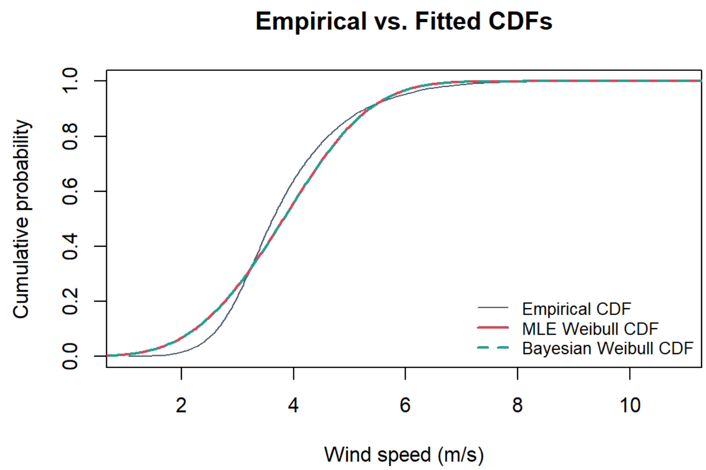

FIGURE 6. Empirical cumulative distribution function versus the fitted Weibull CDFs obtained from the MLE and Bayesian posterior mean estimates.

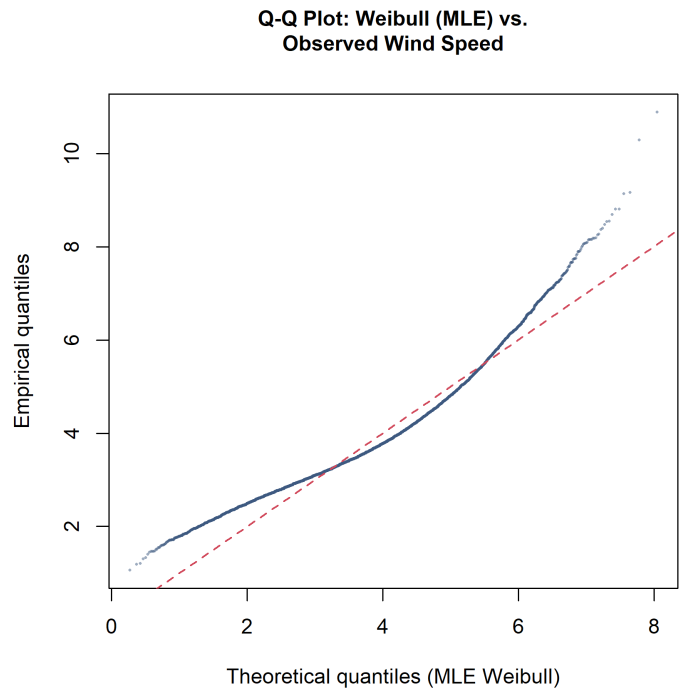

Figure 7. Quantile–quantile plot of the observed wind speed against the fitted Weibull distribution based on the MLE.

Table 5 reports formal goodness-of-fit statistics for the fitted Weibull models obtained under different estimation methods. The MLE achieves the highest log-likelihood and the lowest AIC and BIC among the classical methods, as expected because it is constructed to maximize the likelihood directly. The Bayesian posterior mean produces essentially identical fit statistics, reflecting the close agreement between the Bayesian and maximum likelihood estimates in this large sample.

The Kolmogorov–Smirnov test rejects the null hypothesis of an exact Weibull fit for all estimation methods at conventional significance levels ($p < 0.001$). This result should be interpreted cautiously. Since the parameters are estimated from the data, the reported KS p-values should be interpreted as approximate descriptive measures rather than exact goodness-of-fit probabilities. Moreover, with $n = 8{,}766$ observations, the KS test has very high power to detect even small departures from the fitted parametric distribution. Therefore, the formal rejection does not necessarily imply that the Weibull distribution is practically unsuitable. The visual evidence in Figures 5–7 suggests that the Weibull distribution provides a reasonable approximation to the main body of the wind speed distribution, while some lack of fit remains in the tails.

To assess whether the Weibull distribution is competitive relative to other commonly used models in wind speed analysis, we additionally fitted the Rayleigh, Gamma, and Lognormal distributions to the same data set by maximum likelihood, and compared all four models using the log-likelihood, AIC, BIC, and the one-sample Kolmogorov–Smirnov (KS) statistic (Table 6). Figure 8 overlays the four fitted densities on the histogram of the observed wind speed data.

Table 5. Goodness-of-fit comparison of the fitted Weibull models.

Method	Log-likelihood	AIC	BIC	KS statistic	KS p-value
MLE	-13410.28	26824.55	26838.71	0.0831	5.03e-53
MoM	-13533.01	27070.01	27084.17	0.0899	6.20e-62
LMOM	-13695.95	27395.90	27410.06	0.0915	3.08e-64
Bayes	-13410.28	26824.55	26838.71	0.0831	4.85e-53

Table 6. Goodness-of-fit comparison of the Weibull distribution against alternative parametric models for daily wind speed at Tabuk (2001–2024), ranked by AIC.

Distribution	Parameters (p)	Log-likelihood	AIC	BIC	KS statistic
Lognormal	2	−12549.72	25103.44	25117.60	0.0254
Gamma	2	−12643.41	25290.81	25304.97	0.0438
Weibull	2	−13410.28	26824.55	26838.71	0.0831
Rayleigh	1	−15475.49	30952.98	30960.06	0.2608

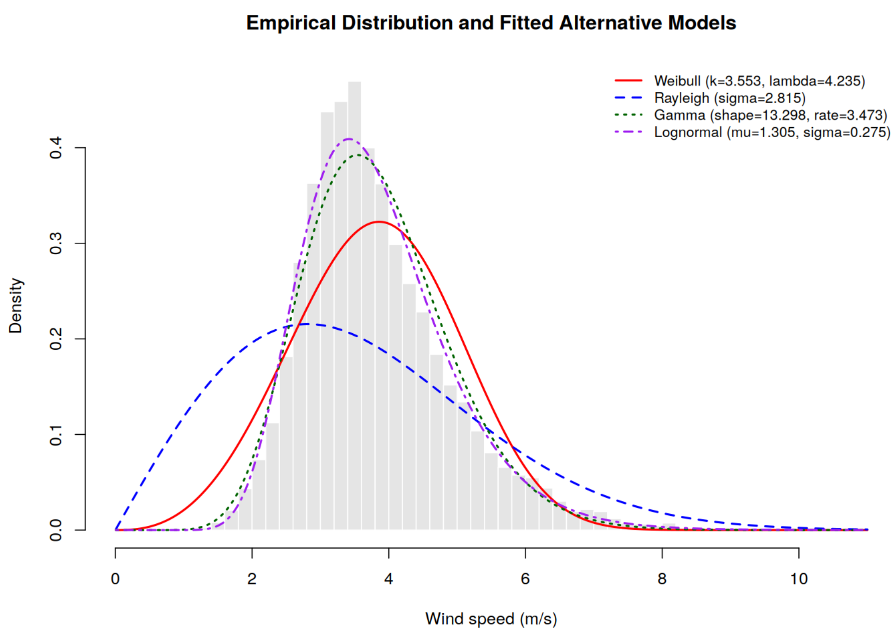

Figure 8. Histogram of the observed daily wind speed with fitted Rayleigh, Weibull, Gamma, and Lognormal densities (maximum likelihood fits).

Table 6 shows that the Lognormal distribution achieves the lowest AIC and BIC among the four candidate models, followed by the Gamma distribution. Both distributions provide a markedly better formal statistical fit to the empirical distribution than the Weibull model, with AIC differences exceeding 1,700 points, well above the conventional threshold (ΔAIC > 10) for decisive support in favour of the better-fitting model. The Rayleigh distribution—a special case of the Weibull distribution with shape parameter fixed at $k = 2$—performs markedly worse than all other candidates, consistent with the estimated Weibull shape parameter ($\hat{k} \approx 3.55$) being substantially larger than 2. This indicates that the Tabuk wind speed distribution is considerably more peaked and less variable than a Rayleigh model would imply.

Although the Lognormal and Gamma distributions provide a better formal statistical fit, the Weibull distribution is retained as a benchmark model for wind-energy interpretation because it remains a standard reference model in wind-energy studies and has a direct closed-form relationship with wind power density. This finding should therefore be interpreted with care. From a purely statistical standpoint, the Lognormal and Gamma distributions describe the central body of the empirical distribution more closely than the Weibull distribution, particularly around the mode (Figure 8). However, the Weibull distribution remains useful for wind-energy assessment for three reasons. First, it is widely used in the wind energy literature, which facilitates direct comparison with the large body of existing regional and international studies discussed in Section 2. Second, it admits a closed-form analytical relationship with the mean wind power density (Equation 4.1), a key practical quantity for wind resource assessment. Third, the Weibull shape parameter k has a direct physical interpretation in terms of the peakedness and variability of the wind regime. We therefore retain the Weibull distribution, estimated via both MLE and Bayesian MCMC, as a benchmark model for the wind-energy assessment, while explicitly noting the statistical evidence in favour of the Lognormal and Gamma alternatives.

4.5. **Simulation Study: MLE versus Bayesian Estimation.** Table 7 and Figure 9 report the Monte Carlo simulation results comparing the MLE and Bayesian posterior-mean estimators across sample sizes $n \in \{30, 50, 100, 200, 500\}$. For small samples ($n \leq 100$), the Bayesian estimator of the shape parameter k exhibits lower MSE than the MLE, reflecting the regularizing effect of the weakly informative Gamma prior in small samples. For the scale parameter λ, the two estimators perform similarly across all sample sizes considered. As n increases, the MSE of both estimators decreases and the difference between them narrows, consistent with the asymptotic equivalence of Bayesian and maximum likelihood estimators under weakly informative priors. These simulation results are conditional on the selected weakly informative Gamma priors and the data-generating parameters chosen from the Tabuk MLE fit.

TABLE 7. Monte Carlo simulation study: mean squared error (MSE) of MLE and Bayesian estimators as a function of sample size n (250 replications; true parameters set equal to the MLE fit of the Tabuk data, $k = 3.553$, $\lambda = 4.235$).

	MSE of $\hat{k}$		MSE of $\hat{\lambda}$	
n	MLE	Bayesian	MLE	Bayesian
30	0.3089	0.2186	0.0468	0.0462
50	0.1815	0.1473	0.0312	0.0310
100	0.0839	0.0726	0.0141	0.0140
200	0.0368	0.0353	0.0077	0.0077
500	0.0161	0.0160	0.0033	0.0033

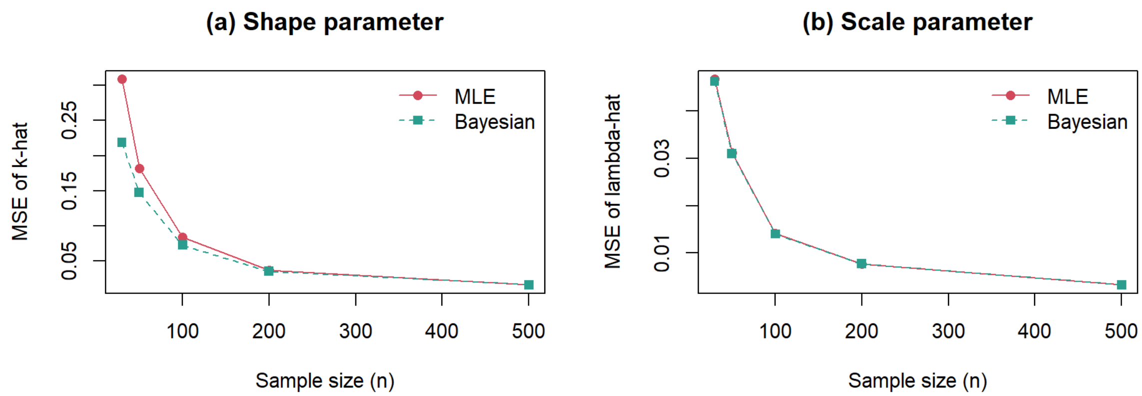

FIGURE 9. Mean squared error of the MLE and Bayesian posterior-mean estimators of the Weibull shape parameter k and scale parameter λ as a function of sample size, based on 250 Monte Carlo replications per sample size.

4.6. **Application: Wind Energy Assessment.** Although the Lognormal and Gamma distributions provide a better formal statistical fit to the daily wind speed data, the Weibull model remains particularly useful for wind-energy applications because it has a simple analytical relationship with mean wind power density. Using the fitted Weibull parameters, the mean wind power density is given by

$$\left(\frac{P}{A}\right)_{\text{Weibull}} = \frac{1}{2}\rho\,\lambda^3\,\Gamma\left(1+\frac{3}{k}\right), \tag{4.1}$$

where ρ denotes air density, λ is the Weibull scale parameter, k is the Weibull shape parameter, and $\Gamma(\cdot)$ is the gamma function.

In addition to the Weibull-based estimate, an empirical wind power density was computed directly from the observed wind speed values as

$$\left(\frac{P}{A}\right)_{\text{emp}} = \frac{1}{2}\rho\,\frac{1}{n}\sum_{i=1}^{n} v_i^3, \tag{4.2}$$

where v_i denotes the observed daily wind speed. This empirical estimate is included as a distribution-free benchmark because it does not depend on assuming that the Weibull distribution is the best-fitting model.

Using the fitted Weibull benchmark model, the mean wind power density at 10 m height was estimated at approximately 43.9 W/m^2. The empirical, distribution-free wind power density computed directly from the observed wind speeds was 43.534 W/m^2, which is very close to the Weibull-based estimates. This agreement suggests that the Weibull benchmark model provides a practically consistent estimate of near-surface wind power density, even though the Lognormal and Gamma distributions provide a better formal statistical fit to the full wind speed distribution. These values indicate a low near-surface wind resource in Tabuk.

Table 8. Wind energy characteristics of Tabuk using empirical and Weibull-based estimates with standard air density $\rho = 1.225$ kg/m^3.

Quantity	Empirical	MLE-based Weibull	Bayesian-based Weibull
Mean wind speed (m/s)	3.8295	3.8136	3.8136
Wind power density (W/m^2)	43.534	43.917	43.922

Several qualifications are important. First, the calculation is based on wind speed at 10 m height and therefore does not directly represent turbine hub-height conditions. Second, the use of standard sea-level air density provides a conventional reference value, but site-specific air density may differ because of local elevation and temperature. Third, a complete wind-energy feasibility assessment would require hub-height wind-speed extrapolation, local air-density adjustment, surface roughness information, turbine power curves, and preferably site-specific mast measurements.

5. Conclusion

This paper presented a comparative classical and Bayesian analysis of daily wind speed distributions using 24 years of NASA POWER daily wind speed data at 10 m height for Tabuk, Saudi Arabia. Within this comparative framework, the two-parameter Weibull distribution was treated as a standard benchmark model in wind-energy studies. Three classical Weibull estimators, namely maximum likelihood estimation, the method of moments, and L-moments, were compared with a Bayesian estimator implemented via Metropolis–Hastings MCMC.

The empirical results showed that the MLE and Bayesian posterior-mean estimates of the Weibull shape and scale parameters were nearly identical in this large sample, with $\hat{k} \approx 3.55$ and $\hat{\lambda} \approx 4.24$ m/s. This close agreement reflects the dominance of the likelihood over the weakly informative priors when the sample size is large. However, the Monte Carlo simulation study showed that the Bayesian estimator can achieve lower mean squared error than the MLE for the Weibull shape parameter in small samples, particularly when $n \leq 100$. This finding supports the use of Bayesian methods as a useful complement to classical estimation in data-limited wind resource studies.

The goodness-of-fit comparison with alternative distributions showed that the Lognormal and Gamma models provide a better formal statistical fit to the Tabuk daily wind speed data than the Weibull distribution. This result confirms that the Weibull model should not be assumed a priori to be the statistically optimal distribution for every wind regime. Rather, its suitability should be assessed empirically against competing probability models. In the present case, the Weibull distribution remains useful as a benchmark model because of its parsimony, its standard role in the wind-energy literature, the interpretability of its shape and scale parameters, and its direct analytical relationship with wind power density.

Using the fitted Weibull benchmark model, the mean wind power density at 10 m height was estimated at approximately 43.9 W/m^2. The empirical, distribution-free wind power density computed directly from the observed wind speeds was 43.534 W/m^2, which is very close to the Weibull-based estimates. This agreement suggests that the Weibull benchmark model provides a practically consistent estimate of near-surface wind power density, even though the Lognormal and Gamma distributions provide a better formal statistical fit to the full wind speed distribution. These values indicate a low near-surface wind resource in Tabuk. A more detailed technical assessment would require hub-height wind-speed extrapolation, local surface roughness information, air-density adjustment, turbine power curves, and preferably site-specific mast measurements.

Overall, the study provides an updated statistical characterization of the wind speed regime in Tabuk and illustrates the practical trade-offs between statistical goodness of fit and practical interpretability in wind-energy applications. The findings emphasize the importance of benchmarking the Weibull distribution against alternative models before drawing conclusions about local wind-resource potential. Future research may extend this work by considering seasonal distributional models, informative Bayesian priors based on regional meteorological evidence, hub-height wind-speed modelling, and broader probability frameworks incorporating Lognormal, Gamma, and other flexible distributions.

Data Availability: The daily wind speed data used in this study are publicly available from the NASA POWER database. The processed data set and analysis code are available from the corresponding author upon reasonable request.

Conflicts of Interest: The author declares that there is no conflict of interest regarding the publication of this paper.

References

[1] C.G. Justus, W.R. Hargraves, A. Mikhail, D. Graber, Methods for Estimating Wind Speed Frequency Distributions, J. Appl. Meteorol. 17 (1978), 350–353. https://doi.org/10.1175/1520-0450(1978)017<0350:mfewsf>2.0.co;2.

[2] J.V. Seguro, T.W. Lambert, Modern Estimation of the Parameters of the Weibull Wind Speed Distribution for Wind Energy Analysis, J. Wind Eng. Ind. Aerodyn. 85 (2000), 75–84. https://doi.org/10.1016/s0167-6105(99)00122-1.

[3] S.A. Akdağ, A. Dinler, A New Method to Estimate Weibull Parameters for Wind Energy Applications, Energy Convers. Manag. 50 (2009), 1761–1766. https://doi.org/10.1016/j.enconman.2009.03.020.

[4] I. Usta, An Innovative Estimation Method Regarding Weibull Parameters for Wind Energy Applications, Energy 106 (2016), 301–314. https://doi.org/10.1016/j.energy.2016.03.068.

[5] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, D.B. Rubin, Bayesian Data Analysis, 3rd ed., Chapman and Hall/CRC, 2013. https://doi.org/10.1201/b16018.

[6] A.N. Celik, A Statistical Analysis of Wind Power Density Based on the Weibull and Rayleigh Models at the Southern Region of Turkey, Renew. Energy 29 (2004), 593–604. https://doi.org/10.1016/j.renene.2003.07.002.

[7] H. Teimourian, M. Abubakar, M. Yildiz, A. Teimourian, A Comparative Study on Wind Energy Assessment Distribution Models: A Case Study on Weibull Distribution, Energies 15 (2022), 5684. https://doi.org/10.3390/en15155684.

[8] A. Bulut, O. Bingöl, Weibull Parameter Estimation Methods on Wind Energy Applications - a Review of Recent Developments, Theor. Appl. Climatol. 155 (2024), 9157–9184. https://doi.org/10.1007/s00704-024-05184-2.

[9] S. Rehman, T.O. Halawani, T. Husain, Weibull Parameters for Wind Speed Distribution in Saudi Arabia, Sol. Energy 53 (1994), 473–479. https://doi.org/10.1016/0038-092x(94)90126-m.

[10] A.H. Alanazi, F.A. Almutlaq, Assessment of Wind Energy Potential at Three Prime Locations in Saudi Arabia: Analysis of Sharma, Qurayyat and Sakaka Sites, Am. J. Nat. Sci. 5 (2024), 30–48. https://doi.org/10.47672/ajns.2074.

[11] K.S. AlQdah, R. Alahmdi, A. Alansari, A. Almoghamisi, M. Abualkhair, M. Awais, Potential of Wind Energy in Medina, Saudi Arabia Based on Weibull Distribution Parameters, Wind Eng. 45 (2021), 1652–1661. https://doi.org/10.1177/0309524x211027356.

[12] M.I. Faraz, Strategic Analysis of Wind Energy Potential and Optimal Turbine Selection in Al-Jouf, Saudi Arabia, Heliyon 10 (2024), e39188. https://doi.org/10.1016/j.heliyon.2024.e39188.

[13] E. Köksal Babacan, S. Kaya, Comparison of Parameter Estimation Methods in Weibull Distribution, Sigma J. Eng. Nat. Sci. 38 (2020), 1609–1621.

[14] J. Wang, J. Hu, K. Ma, Wind Speed Probability Distribution Estimation and Wind Energy Assessment, Renew. Sustain. Energy Rev. 60 (2016), 881–899. https://doi.org/10.1016/j.rser.2016.01.057.

[15] V. Sohoni, S. Gupta, R. Nema, A Comparative Analysis of Wind Speed Probability Distributions for Wind Power Assessment of Four Sites, Turk. J. Electr. Eng. Comput. Sci. 24 (2016), 4724–4735. https://doi.org/10.3906/elk-1412-207.

[16] H.E. Akyuz, H. Gamgam, Statistical Analysis of Wind Speed Data with Weibull, Lognormal and Gamma Distributions, Cumhuriyet Sci. J. 38 (2017), 68–76. https://doi.org/10.17776/csj.358773.

[17] T. Soukissian, Use of Multi-Parameter Distributions for Offshore Wind Speed Modeling: The Johnson S_B Distribution, Appl. Energy 111 (2013), 982–1000. https://doi.org/10.1016/j.apenergy.2013.06.050.

[18] P. Lencastre, A. Yazidi, P.G. Lind, Modeling Wind-Speed Statistics beyond the Weibull Distribution, Energies 17 (2024), 2621. https://doi.org/10.3390/en17112621.

[19] M. Kamal, M. Alam, O.A. Alqasem, E.I. Mohamed, M.I. Habadi, D.K. Alnagar, R. Aldallal, A.S. Alghamdi, Estimation and Statistical Properties of the Weibull–Rayleigh Model with Applications to Industrial Data, Alexandria Eng. J. 128 (2025), 556–570. https://doi.org/10.1016/j.aej.2025.05.001.