Stat

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Efficient Goodness-of-Fit Tests for the Maxwell–Boltzmann Distribution via Stein-Type Characterization With Applications to (Un)censored Data

Ganesh Vishnu Avhad[1] | Bruno Ebner[2]

[1]Department of Mathematics and Statistics, Indian Institute of Technology, Tirupati, India | [2]Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Correspondence: Bruno Ebner (bruno.ebner@kit.edu)

Received: 4 June 2025 | **Revised:** 7 July 2025 | **Accepted:** 27 July 2025

Keywords: goodness-of-fit test | Maxwell–Boltzmann distribution | right censoring | Stein's method | warp-speed bootstrap

ABSTRACT

Understanding molecular motion in an ideal gas is fundamental to thermodynamics, yet directly measuring individual molecular velocities remains impractical. The Maxwell–Boltzmann (MB) distribution provides a well-established statistical model for describing the distribution of molecular speeds. In this study, we propose efficient goodness-of-fit tests for the MB distribution with unknown parameters, leveraging a novel fixed-point characterization derived from a Stein-type identity. We derive the asymptotic distribution of the proposed test statistic and establish its consistency against a broad class of alternative distributions. Furthermore, we extend the methodology to handle right-censored data, broadening its applicability to real-world scenarios where incomplete observations are common. The performance of our approach is evaluated through extensive Monte Carlo simulations, and its practical utility is demonstrated with applications to empirical data.

1 | Introduction

The Maxwell–Boltzmann (MB) distribution was introduced by James Clerk Maxwell (1860) in the context of the kinetic theory of gases (Rowlinson 2005). Since then, it has been a cornerstone of statistical modelling, especially in applications related to physics and engineering. It describes the distribution of particle speeds in an idealized gas under equilibrium and has been widely used in reliability analysis, survival studies and wireless communication modelling. In the 1870s, Boltzmann further investigated the physical origins of the MB distribution, showing that it arises by maximizing the entropy of the system, a key principle in statistical mechanics. This distribution has many important applications in reliability analysis and life testing, especially when the assumption of a constant failure rate is unrealistic (see, e.g., Bekker and Roux 2005; Chaudhary and Tomer 2018; Panwar and Tomer 2019). Recently, Hossain et al. (2017) and Godase et al. (2022) proposed control charts for the MB distribution. Büchel et al. (2021) studied nuclear quantum statistics in molecular dynamics and showed that the MB velocity distribution naturally appears, supporting a deterministic view of quantum mechanics. For the detailed application of the MB distribution, the interested reader is referred to Maxwell (1860), Peckham and McNaught (1992) and Chattamvelli and Shanmugam (2021). Because the MB distribution is widely used in modelling particle speeds and certain lifetime phenomena, the presence of censoring can significantly affect inference and model adequacy if not properly addressed. Censoring is a common occurrence in lifetime and reliability studies. Therefore, it is crucial to develop goodness-of-fit tests for the MB distribution that can accommodate right-censored observations.

In recent studies, the development of goodness-of-fit tests based on characterizations has gained significant attention.

Because characterizations are a good way to distinguish one family from the others, they are useful in goodness-of-fit testing. Numerous such tests exist for different distribution classes; famous examples are testing for normality, exponentiality or testing the fit to families of positive random variables as the gamma-, Pareto-, Weibull- or Rayleigh-law. Examples for these can be found in Nikitin (2017), Henze et al. (2012), Liebenberg et al. (2022), Allison et al. (2022), Ebner et al. (2023) and Jovanović et al. (2015). Betsch and Ebner (2021) developed fixed-point characterizations for continuous distributions using Stein's type identity. Based on this characterization, several tests have been proposed for different distributions; see e.g., Betsch and Ebner (2019) and Vaisakh et al. (2025) for tests of fit to the gamma-law, Betsch and Ebner (2020) for a test of normality and Xavier et al. (2025) for a test of fit to the inverse Gaussian distribution. Note that analogous characterizations for families of discrete distributions are found in Betsch et al. (2022).

The widespread use of the MB distribution has motivated efforts to develop goodness-of-fit tests for its validation. However, to the best of our knowledge, no test has been specifically designed for the MB distribution. In this paper, we propose new goodness-of-fit tests based on a Stein-type characterization, applicable to both complete and right-censored data. A random variable X follows the MB distribution with scale parameter μ if its probability density function (PDF) is of the form

$$f(x;\mu) = \sqrt{\frac{2}{\pi}}\frac{x^2}{\mu^3}\exp\left(-\frac{x^2}{2\mu^2}\right),\quad x>0,\quad \mu>0,$$

and cumulative distribution function (CDF) is given by

$$F(x;\mu) = \mathrm{erf}\left(\frac{x}{\sqrt{2}\mu}\right) - \sqrt{\frac{2}{\pi}}\frac{x}{\mu}\exp\left(-\frac{x^2}{2\mu^2}\right),\quad x>0,$$

where erf is the error function. In the following, we denote the class of the MB distribution by $\mathscr{M} := \{\mathrm{Max}(\mu): \mu > 0\}$. Note that $X \sim \mathrm{Max}(\mu)$ if and only if $\frac{X}{\mu} \sim \mathrm{Max}(1)$ and hence the family of the MB distributions is a member of the scale family of distributions. Let $X, X_1, X_2, \ldots, X_n$ be real-valued independent and identically distributed (i.i.d.) random variables from the continuous distribution function F on $\mathcal{X} \in \mathscr{R}^p$, for some $p \geq 1$. Based on this sample, we are interested in testing the composite hypothesis

$$H_0: F \in \mathscr{M}$$

against general alternatives.

The organization of the paper is as follows: Section 2 presents the Stein-based characterization theorem, and we propose a testing procedure for assessing the fit to the family of MB distributions. We also examine the asymptotic properties of the test statistic. In Section 3, we discuss how to modify the test statistic to accommodate right-censored observations. Section 4 employs Monte Carlo simulations to evaluate the finite sample properties of the proposed methods. The practical applicability of the proposed methods is illustrated using real-life datasets in Section 5, followed by conclusions in Section 6.

2 | Fixed Point Characterization and Test Statistic

Using a Stein-type identity, Betsch and Ebner (2021) established a fixed-point characterization for a broad class of absolutely continuous univariate distributions. To be specific, let X be a positive random variable with continuously differentiable density f satisfying $\int_0^\infty x\,|f'(x)|\,dx < \infty$ to match Theorem 1 of Betsch and Ebner (2021). Then it is shown that X is said to follow a distribution function F if and only if,

$$F(t) = \mathbb{E}\left[-\frac{f'(X)}{f(X)}\min(X,t)\right],\quad t>0.$$

We apply this characterization result to the standard MB distribution as follows.

Theorem 1. *Let X be a positive random variable with $E(X) < \infty$. Then, X follows a standard MB distribution Max(1) if, and only if, its distribution function F is of the form*

$$F(t) = \mathbb{E}\left[-\left(\frac{2-X^2}{X}\right)\min(X,t)\right],\quad t>0.$$

Next, to address the invariance of the MB family with respect to rescaling, we take into account the scaled residuals $(Y_1, Y_2, \ldots, Y_n)$, which are defined as $Y_j = \frac{X_j}{\hat{\mu}_n}, j = 1, \ldots, n$. Here, $\hat{\mu}_n = \hat{\mu}_n(X_1, \ldots, X_n)$ represents a consistent estimator of $\mu > 0$. In the sequel, we use the maximum likelihood estimator (MLE) of the parameter μ given by $\hat{\mu}_n = \sqrt{\frac{1}{3n}\sum_{i=1}^n X_i^2}$. For every $b > 0$ the estimator satisfies $\hat{\mu}_n(bX_1, \ldots, bX_n) = b\hat{\mu}_n(X_1, \ldots, X_n)$. It is clear that Y_j for $j = 1, \ldots, n$ are invariant under scale transformation of the data set $X_1, \ldots, X_n$. Therefore, we fix, w.l.o.g., the scale parameter to be $\mu = 1$ in the subsequent analysis.

2.1 | Test Statistic

Motivated by Theorem 1, we introduce the departure measure as

$$\Delta = \int_0^\infty \left(\mathbb{E}\left[\left(\frac{Y^2-2}{Y}\right)\min(Y,t)\right] - F(t)\right)dF(t), \tag{1}$$

where $F(\cdot)$ is the CDF of Max(1). Note that Δ is zero under the null hypothesis H_0, i.e., if $Y \sim$ Max (1). Now, we can express Δ as

$$\Delta = \int_0^\infty \mathbb{E}\left[\left(\frac{Y^2-2}{Y}\right)\min(Y,t) - I(Y\le t)\right]dF(t) = \mathbb{E}\left[\mathbb{E}\left[\left(\frac{Y^2-2}{Y}\right)\min(Y,Z) - I(Y\le Z)\,|\,Y\right]\right],$$

where $Z \sim F$. Hence, a natural empirical version of Δ based on i.i.d. copies $Y_1, \ldots, Y_n$ of Y is given by

$$\widehat{\Delta}_n = \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^{n}\left(\frac{Y_j^2-2}{Y_j}\right)\min(Y_j,Z) - I(Y_j \le Z) \mid Y_j\right]$$
$$= \frac{1}{n}\sum_{j=1}^{n}\left(\frac{Y_j^2-2}{Y_j}\right)[\mathbb{E}[Y_j I(Y_j \le Z) \mid Y_j] + \mathbb{E}[Z I(Y_j > Z) \mid Y_j]] - \mathbb{E}[I(Y_j \le Z) \mid Y_j]$$
$$= \frac{1}{n}\sum_{j=1}^{n}\left(Y_j^2-2\right)\left(1-F(Y_j)\right) + \sqrt{\frac{2}{\pi}}\left(\frac{Y_j^2-2}{Y_j}\right)\left(2-\left(Y_j^2+2\right)\exp\left(-Y_j^2/2\right)\right) + F(Y_j) - 1$$
$$=: \frac{1}{n}\sum_{j=1}^{n}\psi(Y_j),$$

because direct calculations yield $\mathbb{E}[ZI(Y_1 > Z) \mid Y_1] = \sqrt{\frac{2}{\pi}}\left(2-\left(Y_1^2+2\right)\exp\left(-Y_1^2/2\right)\right)$. A composite test for testing the fit to the MB family based on $\widehat{\Delta}_n$ requires to replace the Y_j by the rescaled $Y_{n,j} = X_j/\widehat{\mu}_n$, $j = 1, \ldots, n$, which is possible due to the scale invariance discussed above. A test based on $\widehat{\Delta}_n(Y_{n,1}, \ldots, Y_{n,n})$ then rejects the null hypothesis H_0 for large (or small) values of $\widehat{\Delta}_n$.

Note that the test statistic $\widehat{\Delta}_n$ is computationally efficient, as it involves a single summation of a closed-form expression. As we will demonstrate in the following section, the limiting null distribution is straightforward to derive. Consequently, the $\alpha/2$ and $1-\alpha/2$ quantiles of the corresponding normal distribution serve as critical values, identifying unusually small and large values of the test statistic, respectively, while ensuring an asymptotic type I error rate of $\alpha \in (0,1)$. To find the critical region of the proposed test, a Monte Carlo simulation procedure is used, as outlined in Avhad et al. (2025). Specifically, we compute the lower (C_l) and upper (C_u) quantiles such that the probabilities $P\left(\widehat{\Delta}_n < C_l\right) = P\left(\widehat{\Delta}_n > C_u\right) = \alpha/2$, for a table of critical values for different significance levels α, see Table 1. Note that the critical values approach the respective normal quantiles (denoted in the row by '∞') as suggested by Theorem 2.

2.2 | The Limit Distribution Under the Null Hypothesis

Assume w.l.o.g. that $X_1, \ldots, X_n$ are i.i.d. with distribution $Max(1)$ and that $\widehat{\mu}_n$ is a consistent estimator of μ such that there exists a function ℓ for which

$$\sqrt{n}(\widehat{\mu}-1) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\ell(X_j) + o_P(1)$$

holds, where $\mathbb{E}\left(\ell(X_1)\right) = 0$ and $\mathbb{E}\left(\ell(X_1)\right)^2 < \infty$. Note that for the MLE, we have (see Corollary 10.16 in Henze (2024)) by direct evaluation of the score function and Fisher information

$$\ell(x) = \frac{x^2-3}{6}, \qquad (2)$$

so indeed $\mathbb{E}\ell(X_1) = 0$ and $\mathbb{E}\left(\ell(X_1)\right)^2 = \frac{1}{6}$.

Theorem 2. *Under the standing assumptions, we have as $n \to \infty$*

$$\sqrt{n}\widehat{\Delta}_n \xrightarrow{\mathcal{D}} N\left(0, \sigma^2\right),$$

where $\sigma^2 = \mathbb{E}\left[\psi(X_1) - (3+4/\pi)\ell(X_1)\right]^2$. Using the MLE, we get $\sigma^2 \approx 0.24515$.

Proof. By the mean value theorem, we have

$$\sqrt{n}\Delta_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\psi(Y_{n,j}) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\psi(X_j) + \sqrt{n}(\widehat{\mu}_n - 1)\frac{1}{n}\sum_{j=1}^{n}\frac{\partial}{\partial\widehat{\mu}_n}\psi(X_j/\widehat{\mu}_n)\Big|_{\widehat{\mu}_n=\mu^*}$$

where μ^* lies between $\widehat{\mu}_n$ and 1. Because of the law of large numbers and straightforward calculations, the last mean converges

TABLE 1 | Critical values for $\sqrt{n}\widehat{\Delta}_n$ using the MLE at different significance levels α.

	$\alpha = 10\%$		$\alpha = 5\%$		$\alpha = 1\%$	
n	C_l	C_u	C_l	C_u	C_l	C_u
10	−0.6973	0.7746	−0.9055	0.8563	−1.3485	0.9923
20	−0.7523	0.7978	−0.9519	0.9009	−1.3631	1.0815
30	−0.7704	0.8042	−0.9614	0.9169	−1.3552	1.1221
40	−0.7802	0.8104	−0.9694	0.9286	−1.3527	1.1478
50	−0.7857	0.8120	−0.9738	0.9336	−1.3514	1.1571
75	−0.7918	0.8144	−0.9743	0.9434	−1.3431	1.1849
100	−0.7989	0.8142	−0.9771	0.9463	−1.3356	1.2001
150	−0.8040	0.8156	−0.9764	0.9528	−1.3220	1.2112
200	−0.8045	0.8143	−0.9765	0.9549	−1.3229	1.2207
300	−0.8076	0.8176	−0.9774	0.9599	−1.3156	1.2306
500	−0.8103	0.8151	−0.9776	0.9611	−1.3006	1.2407
∞	−0.8144	0.8144	−0.9704	0.9704	−1.2754	1.2754

in probability to $\mathbb{E}\left[\frac{\partial}{\partial \hat{\mu}_n}\psi(X_1)\Big|_{\hat{\mu}_n=1}\right] = -3-4/\pi$, so we have $\sqrt{n}\hat{\Delta}_n = \frac{1}{\sqrt{n}}\sum_{j=1}^n \psi(X_j) - (3+4/\pi)\ell(X_j) + o_P(1)$, hence, by the central limit theorem and Slutsky's lemma, the claim follows with $\sigma^2 = \mathbb{E}\left[\psi(X_1) - (3+4/\pi)\ell(X_1)\right]^2$. Inserting (2) and numerical integration yields the result for the MLE. □

2.3 | Consistency

In this subsection, let $X, X_1, X_2, \ldots$ be i.i.d. positive random variables with $\mathbb{E}X^3 < \infty$. Note that in this case $\mathbb{E}\,|\,\psi(X)\,| < \infty$, and by the law of large numbers combined with the continuous mapping theorem, the third moment assumption also implies that $\hat{\mu}_n$ converges to $\left(\mathbb{E}X^2/3\right)^{1/2} =: \mu_0 > 0$ almost surely as $n \to \infty$. In view of the scale invariance of the test statistic Δ_n, we can again assume w.l.o.g. that $\mu_0 = 1$.

Theorem 3. *Under the alternatives stated above, we have* $\hat{\Delta}_n \xrightarrow{\mathbb{P}} \mathbb{E}\psi(X) =: \Delta$ *as* $n \to \infty$.

Proof. The same expansion as in the proof of Theorem 2 shows that $\hat{\Delta}_n = \frac{1}{n}\sum_{j=1}^n \psi(X_j) + o_{\mathbb{P}}(1)$ because the last term on the r.h.s. tends to 0 due to the almost sure convergence of $\hat{\mu}_n$ to 1, and because the law of large numbers and the finite third moment condition ensure the boundedness of the second component. Finally, an application of the law of large numbers and standard properties of stochastic Landau symbols (see Theorem 6.15 in Henze (2024)) finishes the proof. □

As a direct consequence of Theorem 3, the test based on $\hat{\Delta}_n$ is consistent against any alternative for which $\Delta \neq 0$.

3 | Test Statistic for Right Censored Data

In this section, we modify the proposed testing procedure for the randomly right-censored data. Let $X_1, X_2, \ldots, X_n$ be a random sample from a non-negative continuous distribution function F and $C_1, C_2, \ldots, C_n$ be censoring random variables with an absolutely continuous distribution function G defined on $\mathbb{R}^+$. We assume that the sequences $\{X_i\}$ and $\{C_i\}$ are independent. Now, for the right-censored data, we observed $Z_i = \min(X_i, C_i)$ and $\delta_i = I(X_i \le C_i)$, $i = 1, 2, \ldots, n$. Using an approach analogous to the one outlined in the previous section, the test statistic will depend on each observed value Z_j only through the transformed quantity $Z_j^* = Z_j^{\hat{\mu}_C}$, for $j = 1, \ldots, n$, where $\hat{\mu}_C$ is the MLE estimator of μ in the right-censored case. The log-likelihood of the MB distribution with parameter μ is obtained as

$$\mathscr{L}(\mu \mid Z_1, \ldots, Z_n, \delta_1, \ldots, \delta_n) = \prod_{i=1}^n \left[f_\mu(Z_i)\right]^{\delta_i}\left[1 - F_\mu(Z_i)\right]^{1-\delta_i}.$$

Taking logs and differentiating with respect to μ, we get the score function as $\frac{\partial}{\partial \mu}\log \mathscr{L}(\mu) = 0$. After simplification, we get

$$\sum_{i=1}^n \delta_i\left(-\frac{3}{\mu} + \frac{Z_i^2}{\mu^3}\right) + \sum_{i=1}^n (1-\delta_i)\frac{\partial}{\partial \mu}\log\left(1 - F_\mu(Z_i)\right) = 0. \tag{3}$$

Because expression (3) does not have a closed-form solution, the MLE $\hat{\mu}_C$ is obtained using numerical optimization techniques.

To develop the test statistic for the right-censored case, we adopt the inverse probability of censoring weighting (IPCW) (Datta et al. (2010)) technique to handle right-censored data. As this approach is based on U-statistics, we use a modified version of the departure measure Δ given in (1). After some algebra, we express Δ as

$$\begin{aligned}\Delta &= \int_0^\infty \int_0^\infty \left(\frac{y^2-2}{y}\right)\min(y,t)dF(y)dF(t) - \frac{1}{2} \\ &= \int_0^\infty \int_0^\infty y\min(y,t)dF(y)dF(t) - 2\int_0^\infty \int_0^\infty \left(\frac{1}{y}\right)\min(y,t)dF(y)dF(t) - \frac{1}{2} \\ &= \mathbb{E}\left[Y_1\min(Y_1,Y_2)\right] - 2\mathbb{E}\left[\frac{Y_1}{Y_2}I(Y_1 < Y_2)\right] - \frac{3}{2}.\end{aligned} \tag{4}$$

Hence, we obtain the U-statistic based test statistic in the right-censored case as

$$\hat{\Delta}_n^C = \frac{2}{n(n-1)}\sum_{i=1}^n \sum_{j<1;j=1}^n \frac{Z_i^*\min\left(Z_i^*, Z_j^*\right)\delta_i\delta_j}{\hat{K}_c(Z_i^*-)\hat{K}_c\left(Z_j^*-\right)} - \frac{2}{n(n-1)}\sum_{i=1}^n \sum_{j<1;j=1}^n \frac{\left(\frac{Z_i^*}{Z_j^*}I\left(Z_i^* < Z_j^*\right) + \frac{Z_j^*}{Z_i^*}I\left(Z_j^* < Z_i^*\right)\right)\delta_i\delta_j}{\hat{K}_c(Z_i^*-)\hat{K}_c\left(Z_j^*-\right)} - \frac{3}{2},$$

where $\hat{K}_c(z^*-) = \prod_{Z_i^* < z^*}\left(1 - \frac{1-\delta_i}{\sum_{l=1}^n I\{Z_l^* \ge Z_i^*\}}\right)$ with $\hat{K}_c(\cdot)$ be the Kaplan–Meier estimator of the survival function $K_c(\cdot)$, and $K_c(z^*-) = P(C > z^*)$ is the survival function of the censoring variable C such that $\hat{K}_c(z^*) > 0$, for each z^* with probability 1. Hence, in the right-censored case, we reject H_0 in favour of H_1 for large (or small) values of $\hat{\Delta}_n^C$.

According to Theorem 1 presented by Datta et al. (2010), the test statistic under certain regularity conditions tends to follow a Gaussian distribution with zero mean and a complex variance which depends on the unknown censoring distribution, under the null hypothesis; for more details, one can see Cuparić (2021) and the references therein. However, the variance of this limiting distribution is complex and inherently tied to the unknown censoring mechanism. Because we do not impose any specific structure on the censoring distribution, these asymptotic results are not directly applicable for constructing a goodness-of-fit test in the presence of censored data. To address this limitation, we used the parametric warp-speed bootstrap method (Giacomini et al. 2013) to approximate the p-value associated with the test statistic, which allows us to make a decision regarding the null hypothesis.

4 | Simulation Study

In this section, we perform a Monte Carlo simulation study using the statistical software R (R Core Team 2021) to compare the performance in the previous section. A significance level of $\alpha = 0.05$ is used throughout the study. To demonstrate the competitiveness of the proposed tests in both uncensored and censored cases, we compare their empirical type I error rates and powers with those of existing methods. The empirical power of the tests is evaluated against various alternatives such as Rayleigh (R), Gamma (G), Pareto (P), Lévy (L) and chi-squared (χ^2) distributions. Because no existing test is available in the literature for the MB distribution, we compare the performance of the proposed test with classical goodness-of-fit tests, including the Kolmogorov–Smirnov (KS), Cramér–von Mises (CvM) and Anderson–Darling (AD) tests.

4.1 | Uncensored Case

For each alternative, simulations based on 10,000 replicates were performed on the different sample sizes of $n = 25$, 50, 75 and 100. The performance of the newly proposed test, which uses a simulated critical region-based approach ($\widehat{\Delta}_n$) given by the critical values in Table 1, is thoroughly evaluated. The results of the simulation study are reported in Table 2. From Table 2, it can be seen that the proposed tests achieve the desired significance level. The empirical power of the proposed test is reasonably good against various alternatives. The results of the simulation study indicate that the empirical power of the proposed test improves as the sample size increases. For the Gamma, Pareto, Levy and chi-squared distributions, $\widehat{\Delta}_n$ exhibits higher power as compared with other tests. Overall, the newly proposed test performs better in the majority of scenarios evaluated. The proposed test consistently demonstrates high power for small sample sizes, indicating its overall efficiency and robustness.

4.2 | Censored Case

Here, we assess the empirical power of the proposed test, using sample sizes $n = 20$ and 50, for right-censored data. Censoring is introduced via the Koziol–Green model (Koziol and Green 1976). Because the null distribution of the test depends on the parameter μ and the censoring distribution, power is estimated through a bootstrap procedure with $B = 1000$ resamples over 10,000 Monte Carlo iterations. The empirical type I error and power of the proposed test are evaluated for each lifetime distribution with 20% or 40% of censored observations to assess its sensitivity in the presence of incomplete data. To calculate the empirical type I error and the empirical power of the test statistic proposed for right-censored data, we used the bootstrap algorithm described in Bhati et al. (2025). Theoretical justification follows similar arguments to those in Cuparić and Milošević (2022) and Bothma et al. (2024).

The results of empirical type I error and power compression are given in Table 3. From Table 3, we can see that the empirical type I error of the test converges to the specified significance level. The performance of the proposed test is good in terms of empirical power. From Table 3, we observe that the power of the test increases with larger sample sizes, whereas higher levels of censoring lead to a reduction in power. It is well known that censoring reduces the content of the information in the data, as some observations are only partially observed. However, the observed improvement in power under censoring may be due to the use of the bootstrap method along with the structure of the Green–Koziol model. In the presence of moderate to high levels of censoring, the proposed tests continue to demonstrate good power, indicating that the performance remains robust. These findings suggest that the proposed test procedures are effective and reliable even in the presence of censored data.

5 | Analysis of Real Data Sets

In this section, the application of the proposed test procedures is illustrated with two real datasets.

TABLE 2 | Empirical type I error rates and empirical powers at the level 0.05.

Alt.	$\widehat{\Delta}_n$	KS	CvM	AD	$\widehat{\Delta}_n$	KS	CvM	AD
		$n = 25$				$n = 50$		
$M(1)$	0.0496	0.0495	0.0503	0.0534	0.0507	0.0505	0.0455	0.0472
$R(3)$	0.2644	0.1126	0.1563	0.3092	0.4690	0.1283	0.2128	0.4572
$G(2,5)$	0.8620	0.2106	0.3386	0.7407	0.9889	0.2381	0.5325	0.9347
$P(1,2)$	0.9999	0.6349	0.8961	0.9875	1.0000	0.8143	0.9924	0.9998
$L(0,4)$	1.0000	0.8641	0.9923	1.0000	1.0000	0.9837	1.0000	1.0000
$\chi^2(4)$	0.8668	0.2134	0.3448	0.7378	0.9904	0.2478	0.5308	0.9329
		$n = 75$				$n = 100$		
$M(1)$	0.0492	0.0494	0.0511	0.0529	0.0480	0.0467	0.0495	0.0491
$R(3)$	0.6228	0.1387	0.2650	0.5935	0.7398	0.1371	0.3116	0.6713
$G(2,5)$	0.9995	0.2599	0.6559	0.9845	1.0000	0.2707	0.7518	0.9960
$P(1,2)$	1.0000	0.8712	0.9997	1.0000	1.0000	0.9036	1.0000	1.0000
$L(0,4)$	1.0000	0.9899	1.0000	1.0000	1.0000	0.9934	1.0000	1.0000
$\chi^2(4)$	0.9987	0.2612	0.6392	0.9837	1.0000	0.2591	0.7584	0.9960

TABLE 3 | Empirical rejection rates using the test for censored data at the level 0.05.

Alt.	$n=25$				$n=50$			
	$\widehat{\Delta}_n^C$	KS	CvM	AD	$\widehat{\Delta}_n^C$	KS	CvM	AD
				Censoring 20%				
$M(1)$	0.0487	0.0530	0.0524	0.0501	0.0513	0.0458	0.0485	0.0499
$R(3)$	0.3216	0.1788	0.2923	0.3267	0.4155	0.2404	0.4215	0.4675
$G(2,5)$	0.7681	0.2993	0.5428	0.6894	0.8617	0.3665	0.7311	0.8794
$P(1,2)$	0.8496	0.6977	0.7968	0.8671	0.9448	0.8289	0.9602	0.9997
$L(0,4)$	1.0000	0.9962	0.9583	1.0000	1.0000	0.9987	1.0000	1.0000
$\chi^2(4)$	0.7843	0.4382	0.6676	0.8164	0.9324	0.5647	0.8671	0.9881
				Censoring 40%				
$M(1)$	0.0479	0.0453	0.0550	0.0456	0.0491	0.0504	0.0561	0.0526
$R(3)$	0.2843	0.1107	0.2439	0.3088	0.3897	0.2261	0.3846	0.4671
$G(2,5)$	0.5419	0.2657	0.4208	0.5319	0.6840	0.2903	0.6415	0.7861
$P(1,2)$	0.6978	0.5788	0.7623	0.8269	0.7371	0.6993	0.8762	0.6981
$L(0,4)$	0.9206	0.7993	0.8948	0.9533	0.9685	0.7683	0.9887	0.9771
$\chi^2(4)$	0.7438	0.2810	0.5591	0.7622	0.8602	0.4622	0.6451	0.8652

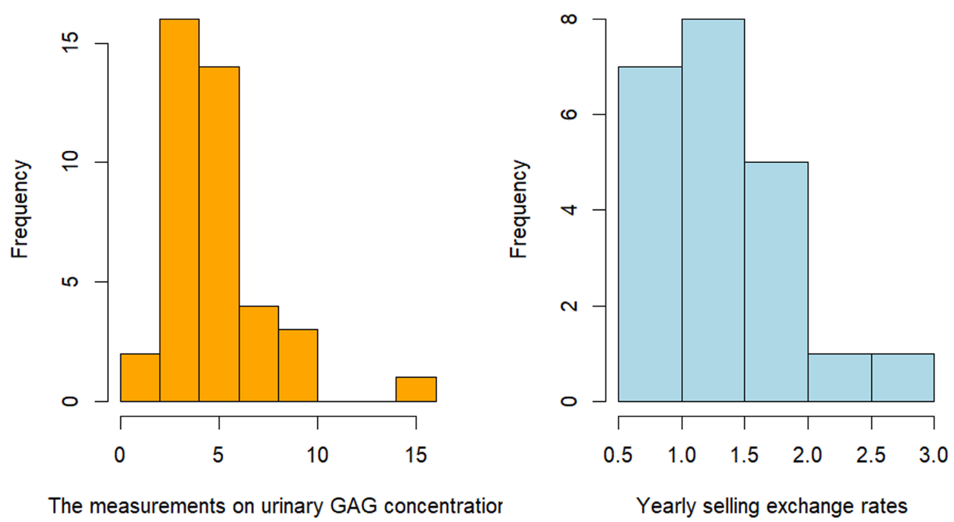

FIGURE 1 | Histogram plots of the real data sets.

5.1 | Uncensored Case: Illustration 1

We consider the data set of measurements of urinary glycosaminoglycan concentration. Godase et al. (2022) examine this dataset within the framework of quality control applications, focusing on the use of SPRT (Sequential Probability Ratio Test) control charts to monitor and assess process performance. Venables and Ripley (2013) provided a dataset with measurements of urinary glycosaminoglycan concentration (in milligrammes per millimole of creatinine) from normal children aged up to 17 years. Analysis using the KS test reveals that these observations follow an MB distribution with a scale parameter $\widehat{\mu}$ equal to 3.2457, where the parameter is estimated from the data using the MLE method.

5.2 | Uncensored Case: Illustration 2

We examine the end-of-year Nigerian Naira to Japanese Yen exchange rates from 1995 to 2016. The data set was studied in Ishaq and Abiodun (2020) for testing the Maxwell–Weibull distribution. We use this dataset to test whether these observations follow the MB distribution. The MLE estimator of μ is $\widehat{\mu} = 0.8895$.

We thus apply the proposed tests of fit to these two data sets to assess their performance on real-world data. The critical values were obtained by applying the tests to the data, while bootstrap p-values were obtained by generating 10,000 samples from the MB distribution using the estimated parameter. The tests were performed on each sample, and the number of times the test

TABLE 4 | Test statistic along with p-value results for real datasets.

	Illustration 1		Illustration 2	
	Test statistic	p-value	Test statistic	p-value
$\widehat{\Delta}_n$	−0.245	0.152	0.298	0.104
KS	0.156	0.279	0.088	0.989
CvM	0.275	0.158	0.027	0.985
AD	1.459	0.186	0.236	0.977

statistic exceeded the critical value was recorded. The bootstrap p-value is obtained by dividing this count by the total number of samples. Figure 1 illustrates that the data sets closely follow an MB distribution. Table 4 presents the test statistics along with the corresponding bootstrap p-values for all the considered data sets. From Table 4, we fail to reject the null hypothesis that the data follows an MB distribution for all the data sets.

5.3 | Censored Case: Illustration 1

To demonstrate the applicability of the proposed method for testing the MB distribution under random censoring, we consider the survival data reported in Bjerkedal (1960). The data consist of survival times (in days) of guinea pigs injected with varying doses of tubercle bacilli. Given the high susceptibility of guinea pigs to human tuberculosis, they were chosen for this particular biomedical study. We analyse the dataset provided by Yadav et al. (2023), which includes artificially introduced random right censoring. Specifically, 72 censoring time points were generated uniformly over the range of observed survival times, with the random seed set to 100 to ensure reproducibility. This results in a censored sample consisting of 56 exact failure times and 16 censored observations. We apply the test discussed in Section 3, and the test statistic value obtained was $\widehat{\Delta}_n^C = -1.4743$, with a corresponding p-value of 0.197. The p-value is greater than the significance level 0.05, so we fail to reject the null hypothesis that the data follow an MB distribution.

6 | Concluding Remarks

In this paper, based on a fixed-point type Stein characterization, we proposed novel goodness-of-fit tests for the MB distribution in the presence and absence of censoring. Notably, to the best of our knowledge, no existing goodness-of-fit test for the family of MB distributions has been explored in the literature. Asymptotic properties of the proposed test have been derived. For both cases, our Monte Carlo simulation results show that the proposed tests maintain well-controlled type I error rates and exhibit good power across different alternative distributions. Finally, we illustrated our proposed test procedure using real-life data sets.

We close the article by pointing out further research directions. Firstly, more theory could be derived under fixed alternatives in the uncensored case, showing that $\sqrt{n}\left(\widehat{\Delta}_n - \Delta\right)$ converges to a centred limit normal distribution with variance depending on the underlying alternative distribution. We leave this point open for further research because we do not see a direct impact for practitioners. Secondly, it would be natural to consider a weighted L^2 distance-based test statistic in analogy to (1) within this framework for both complete and censored observations. This approach would lead to a universally consistent procedure in the uncensored case, and theory can be derived in a Hilbert-space framework due to the L^2 structure, for details, see Neuhaus (1979), and for new theoretical developments under alternatives, we refer to Baringhaus et al. (2017). Thirdly, one could investigate how selecting an alternative parameter estimation method affects the power of the tests (see Drost et al. 1990). In our case, we expect that MLE will yield the highest test power.

Author Contributions

Ganesh Vishnu Avhad: writing, methodology, simulations, data examples, editing. **Bruno Ebner:** writing, methodology, editing, reviewing.

Acknowledgements

The authors would like to thank the editor-in-chief, associate editor and two reviewers for their constructive comments that led to an improvement of the article. Open Access funding enabled and organized by Projekt DEAL.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

Allison, J., B. Milošević, M. Obradović, and M. Smuts. 2022. "Distribution-Free Goodness-of-Fit Tests for the Pareto Distribution Based on a Characterization." *Computational Statistics* 37, no. 1: 403–418.

Avhad, G. V., A. Lahiri, and S. K. Kattumannil. 2025. "A Family of Non-Parametric Tests for the Class of Log-Symmetric Distributions." *Journal of Applied Statistics* 1–16.

Baringhaus, L., B. Ebner, and N. Henze. 2017. "The Limit Distribution of Weighted L^2-Goodness-of-Fit Statistics Under Fixed Alternatives, With Applications." *Annals of the Institute of Statistical Mathematics* 69, no. 5: 969–995.

Bekker, A., and J. Roux. 2005. "Reliability Characteristics of the Maxwell Distribution: A Bayes Estimation Study." *Communications in Statistics—Theory and Methods* 34, no. 11: 2169–2178.

Betsch, S., and B. Ebner. 2019. "A New Characterization of the Gamma Distribution and Associated Goodness of Fit Tests." *Metrika* 82, no. 7: 779–806.

Betsch, S., and B. Ebner. 2020. "Testing Normality via a Distributional Fixed Point Property in the Stein Characterization." *Test* 29, no. 1: 105–138.

Betsch, S., and B. Ebner. 2021. "Fixed Point Characterizations of Continuous Univariate Probability Distributions and Their Applications." *Annals of the Institute of Statistical Mathematics* 73: 31–59.

Betsch, S., B. Ebner, and F. Nestmann. 2022. "Characterizations of Non-Normalized Discrete Probability Distributions and Their Application in Statistics." *Electronic Journal of Statistics* 16, no. 1: 1303–1329.

Bhati, D., A. Batsidis, and S. Khandelwal. 2025. "New Goodness of Fit Tests for the Pareto Distribution Using Stein's Characterization for Uncensored and Random Right Censored Data." *Communications in Statistics—Theory and Methods* 1–25.

Bjerkedal, T. 1960. "Acquisition of Resistance in Guinea Pies Infected With Different Doses of Virulent Tubercle Bacilli." *American Journal of Hygien* 72, no. 1: 130–148.

Bothma, E., J. Allison, and I. Visagie. 2024. "On the Effect of the Kaplan-Meier Estimator's Assumed Tail Behavior on Goodness-of-Fit Testing." *Communications in Statistics—Simulation and Computation*: 1–27.

Büchel, R. C., D. A. Rudolph, and I. Frank. 2021. "Deterministic Quantum Mechanics: The Role of the Maxwell–Boltzmann Distribution." *International Journal of Quantum Chemistry* 121, no. 7: e26555.

Chattamvelli, R., and R. Shanmugam. 2021. *Maxwell Distribution*, 253–261. Springer International Publishing.

Chaudhary, S., and S. K. Tomer. 2018. "Estimation of Stress–Strength Reliability for Maxwell Distribution Under Progressive Type-II Censoring Scheme." *International Journal of System Assurance Engineering and Management* 9: 1107–1119.

Cuparić, M. 2021. "Asymptotic Properties of Inverse Probability of Censored Weighted U-Empirical Process for Right-Censored Data With Applications." *Statistics* 55, no. 5: 1035–1057.

Cuparić, M., and B. Milošević. 2022. "New Characterization-Based Exponentiality Tests for Randomly Censored Data." *Test* 31, no. 2: 461–487.

Datta, S., D. Bandyopadhyay, and G. A. Satten. 2010. "Inverse Probability of Censoring Weighted U-Statistics for Right-Censored Data With an Application to Testing Hypotheses." *Scandinavian Journal of Statistics* 37, no. 4: 680–700.

Drost, F. C., W. C. M. Kallenberg, and J. Oosterhoff. 1990. "The Power of EDF Tests of Fit Under Non-Robust Estimation of Nuisance Parameters." *Statistics and Risk Modeling* 8, no. 2: 167–182.

Ebner, B., A. Fischer, N. Henze, and C. Mayer. 2023. "Goodness-of-Fit Tests for the Weibull Distribution Based on the Laplace Transform and Stein's Method." *Annals of the Institute of Statistical Mathematics* 75, no. 6: 1011–1038.

Giacomini, R., D. N. Politis, and H. White. 2013. "A Warp-Speed Method for Conducting Monte Carlo Experiments Involving Bootstrap Estimators." *Econometric Theory* 29, no. 3: 567–589.

Godase, D. G., S. B. Mahadik, and A. C. Rakitzis. 2022. "The SPRT Control Charts for the Maxwell Distribution." *Quality and Reliability Engineering International* 38, no. 4: 1713–1728.

Henze, N. 2024. *Asymptotic Stochastics. An Introduction With a View Towards Statistics. Translated From the German, Volume 10 of Mathematics Study Resources*. Springer.

Henze, N., S. G. Meintanis, and B. Ebner. 2012. "Goodness-of-Fit Tests for the Gamma Distribution Based on the Empirical Laplace Transform." *Communications in Statistics—Theory and Methods* 41, no. 9: 1543–1556.

Hossain, M. P., M. H. Omar, and M. Riaz. 2017. "New V Control Chart for the Maxwell Distribution." *Journal of Statistical Computation and Simulation* 87, no. 3: 594–606.

Ishaq, A. I., and A. A. Abiodun. 2020. "The Maxwell-Weibull Distribution in Modeling Lifetime Datasets." *Annals of Data Science* 7, no. 4: 639–662.

Jovanović, M., B. Milošević, Y. Y. Nikitin, M. Obradović, and K. Y. Volkova. 2015. "Tests of Exponentiality Based on Arnold–Villasenor Characterization and Their Efficiencies." *Computational Statistics & Data Analysis* 90: 100–113.

Koziol, J. A., and S. B. Green. 1976. "A Cramer-von Mises Statistic for Randomly Censored Data." *Biometrika* 63, no. 3: 465–474.

Liebenberg, S. C., J. Ngatchou-Wandji, and J. S. Allison. 2022. "On a New Goodness-of-Fit Test for the Rayleigh Distribution Based on a Conditional Expectation Characterization." *Communications in Statistics—Theory and Methods* 51, no. 15: 5226–5240.

Maxwell, J. C. 1860. "II. Illustrations of the Dynamical Theory of Gases." *London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 20, no. 130: 21–37.

Neuhaus, G. 1979. "Asymptotic Theory of Goodness of Fit Tests When Parameters Are Present: A Survey." *Series Statistics* 10, no. 3: 479–494.

Nikitin, Y. Y. 2017. "Tests Based on Characterizations, and Their Efficiencies: A Survey." *Acta et Commentationes Universitatis Tartuensis de Mathematica* 21: 3–24.

Panwar, M., and S. K. Tomer. 2019. "Robust Bayesian Analysis of Lifetime Data From Maxwell Distribution." *Austrian Journal of Statistics* 48, no. 1: 38–55.

Peckham, G. D., and I. J. McNaught. 1992. "Applications of Maxwell-Boltzmann Distribution Diagrams." *Journal of Chemical Education* 69, no. 7: 554.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Rowlinson, J. 2005. "The Maxwell–Boltzmann Distribution." *Molecular Physics* 103, no. 21–23: 2821–2828.

Vaisakh, K. M., E. P. Sreedevi, and S. K. Kattumannil. 2025. "A New Goodness of Fit Test for Gamma Distribution With Censored Observations." *Communications in Statistics—Simulation and Computation* 54, no. 1: 91–103.

Venables, W. N., and B. D. Ripley. 2013. *Modern Applied Statistics With S-PLUS*. Springer Science & Business Media.

Xavier, T., K. Vaisakh, and E. Sreedevi. 2025. "Goodness-of-Fit Tests for Inverse Gaussian Distribution in the Presence and Absence of Censoring." *Journal of Statistical Computation and Simulation* 95, no. 5: 1010–1032.

Yadav, C. P., J. Kumarb, and M. Panwarc. 2023. "Survival Analysis of Random Censoring With Inverse Maxwell Distribution: An Application to Guinea Pigs Data." *Electronic Journal of Applied Statistical Analysis* 16, no. 2: 382–409.