Eur. J. Stat. 5 (2025) 1
doi: 10.28924/ada/stat.5.1

Machine Learning Models in Predicting Failure Times Data Using a Novel Version of the Maxwell Model

Uthumporn Panitanarak[1], Aliyu Ismail Ishaq[2*], Narinderjit Singh Sawaran Singh[3], Abubakar Usman[2], Abdullahi Ubale Usman[4], Hanita Daud[5], Akeem Ajibola Adepoju[6], Ibrahim Abubakar Sadiq[2], Ahmad Abubakar Suleiman[5,6]

[1]Department of Biostatistics, Faculty of Public Health, Mahidol University, 10400, Thailand
uthumporn.pan@mahidol.ac.th

[2]Department of Statistics, Ahmadu Bello University, Zaria. Nigeria
binishaq05@gmail.com, abubakarusman28@gmail.com, abubakarsadiq463@gmail.com

[3]Faculty of Data Science and Information Technology, INTI International University
Persiaran Perdana BBN Putra Nilai, 71800, Malaysia
narinderjits.sawaran@newinti.edu.my

[4]Department of Statistics and Mathematics, Zhejiang Gonshang University, Hangzhou, 310018, China
abdulbista4u@gmail.com

[5]Fundamental and Applied Sciences Department, Universiti Teknologi PETRONAS, Seri Iskandar, 32610, Malaysi
ahmadabubakar31@gmail.com, hanita_daud@utp.edu.my

[6]Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, 713281, Kano. Nigeria
akeebola@gmail.com, ahmadabubakar31@gmail.com

*Corresponding author: binishaq05@gmail.com

ABSTRACT: This work aims to introduce a novel statistical distribution based on Maxwell distribution that can handle both positive and negative data sets with varying failure rates, including decreasing and bathtub-shaped distributions. The novel statistical distribution can be derived via the log transformation approach with an additional exponent parameter, defining the Transformed Log Maxwell (TLMax) distribution. The numerical investigation reveals that the developed TLMax distribution can effectively fit negative and positive data sets. A data set containing failure times for Kevlar 49/epoxy at a pressure of approximately 90% was employed to compare the proposed model against the traditional Maxwell model, and the results obtained indicated that the novel distribution outperformed the comparator. Finally, for the prediction of failure times in the dataset, we employed a machine learning model, including support vector regression (SVR), K-nearest neighbors' regression (KNN),

Received: 21 Nov 2024.

Key words and phrases. Maxwell distribution; Transformed log Maxwell distribution; Moments; Machine learning gradient boosting regression; Process innovation.

linear regression (LR), and gradient boosting regression (XGBoost). The findings indicate that the KNN model demonstrates greater prediction robustness than the other models. Beyond practitioners and researchers, this research holds relevance for professionals in physics and chemistry, where the Maxwell distribution is commonly employed.

1. Introduction

Several conventional distributions were commonly utilized in modeling lifetime data in a variety of domains, including engineering, medicine, biology, demography, economics, finance, insurance, and machine learning [1]. Among these distributions is the so-called Maxwell distribution when studying physics and chemistry. It was established by Maxwell [2] to explain the velocities of molecules in thermal equilibrium. According to statistical mechanics, the Maxwell distribution determines the speed of molecules in thermal equilibrium under particular circumstances. In the kinetic theory of gases, for instance, this distribution explains the distribution of energy and moments among other essential gas features [2]. Aside from potential uses in physics and chemistry, [4] used the Maxwell distribution to represent real-world data for the first time in statistics. In statistics and probability distributions, the Maxwell distribution can be employed to characterize positively skewed data sets with an increasing failure rate.

Many expansions of the Maxwell distribution have recently been investigated by researchers to enable greater performances in simulating real-world occurrences from a variety of applications. For instance, by incorporating a modified Weibull failure rate, the generalized Maxwell distribution was obtained [5]. Several tail features of the generalized Maxwell distribution were investigated in [3]. Other variations include the expansion of the Weibull model by considering Maxwell-G (Max-G) by [4], Max-exponential [5], Max-exponentiated exponential distribution [6], Max -Dagum distribution [7], Max-Mukherjee Islam distribution [8], Max-Lomax distribution [9], Inverse power Max distribution [10], and Max-Burr X distribution [11]. there has been tremendous enthusiasm for expanding or generating well-known distributions to model real data. Several distributions have been employed including the odd beta prime (OBP) G family by [12], A new extended distribution by [13], A New Odd Beta Prime-Burr X by [14], the OBP-logistic by [15], the OBP-Fréchet distributions by [16], and more others.

Machine learning models have emerged as formidable tools in regression analysis, offering a significant enhancement over conventional techniques. Their ability to supplement traditional methods results in more precise and effective assessments [17]. For more reading about machine learning models see [18-21]. In this study for the prediction of failure times in the dataset, using four machine learning models, including support vector regression (SVR), K-nearest neighbors' regression (KNN), linear regression (LR), and gradient boosting regression (XGBoost).

The primary objective of the novel study is to employ the logarithmic transformation technique on the Maxwell distribution, giving rise to a novel and versatile distribution termed the transformed log Maxwell (TLMax) distribution. The formulation for the proposed TLMax distribution resembles to that of the log-logistic distribution, which is obtained from the logistic distribution. This transformation is particularly conducive to elongating the tail of the resulting distribution, as the logarithmic function compresses larger values into smaller ones. Importantly, this method exhibits advantageous properties and features owing to its straightforward applicability [22]. The distributions derived based on the logarithmic approach include log-exponentiated Weibull by [23], the log-Kumaraswamy by [24] and many others.

In this study, we validate the distribution's appropriateness for simulating real-world data through extensive numerical applications on engineering data, the findings affirmatively support its efficacy. Our investigation from the application section highlights why the suggested transformed log-Maxwell model is the best option for skewed data sets, offering a viable alternative across various practical scenarios compared to the famous Maxwell distribution.

The key motivations for introducing the transformed log Maxwell distribution are outlined below:

i. To introduce a transformed log Maxwell distribution that enhances the classical Maxwell distribution. This improvement aims to endow the classical model with the advantage of effectively capturing right-skewed and heavy-tailed data sets, thereby ensuring its competitiveness among contemporary distributions.

ii. To introduce a model with shape capacity to handle different data sets including right-skewed, approximately normal, and so on.

iii. To present novel distribution with adaptable failure functions that can handle bathtub, increasing, decreasing, and many more.

iv. To introduce a model that offers consistent performance over the existing model in contemporary literature.

This is the summary of the remaining portion of the paper: Section 2 discusses the development of the transformed log Maxwell distribution, providing insights into its densities, validity check, and linear representation. The properties associated with the proposed distribution are given in Section 3. Section 4 presents the application to failure times data set, comparing the fitness and flexibility of the new model with its competitors. Section 5 covers the applications of machine learning in predicting failure times data sets. The concluding thoughts are presented in Section 6.

2. Material

2.1. Developing Transformed Log Maxwell Distribution

This section presented and investigated a novel statistical distribution as an alternative to the Maxwell random variable adopting the transforming approach outlined in [22].

As studied in Maxwell [2], the Maxwell random variable's cumulative distribution function (cdf) has been described as:

$$F(t;\beta)=\frac{2}{\sqrt{\pi}}\gamma\left(\frac{3}{2},\frac{t^2}{2\beta^2}\right), \qquad t>0,\ \beta>0 \tag{1}$$

where β represents the scaling parameter. It pdf, or probability density function, connected to equation (1) seems to be obtained as:

$$f(t;\beta)=\frac{\sqrt{2}}{\sqrt{\pi}}\frac{t^2\,\mathrm{e}^{-\frac{t^2}{2\beta^2}}}{\beta^3}, \qquad t>0,\ \beta>0. \tag{2}$$

The suggested distribution can be obtained by transforming $x^{\theta}=\log(t)$ from the Maxwell random variable to form the transformed log Maxwell distribution, here, $\theta>0$ represented by the exponent parameter. For obtaining the pdf of the suggested TLMax distribution, we assess the relationship:

$$f(x;\beta,\theta)=f(t;\beta)\times\left|\frac{dt}{dx}\right|. \quad (3)$$

where $f(t;\beta)$ is the pdf given in equation (2) and $\frac{dt}{dx}=\theta e^{x^{\theta}}x^{\theta-1}$ denotes the derivative of the transformation technique studied by [22]. The pdf of the TLMax distribution with an exponent parameter is now obtained by plugging equation (2) into equation (3) as follows.

$$f(x;\beta,\theta)=\sqrt{\frac{2}{\pi}}\frac{\theta x^{\theta-1}e^{3x^{\theta}}}{\beta^{3}}\mathrm{e}^{-\frac{1}{2\beta^{2}}\left(e^{x^{\theta}}\right)^{2}}, \qquad -\infty<x<\infty,\ \theta>0. \quad (4)$$

2.2. Validity Check

To determine if the novel TLMax distribution is a legitimate probability distribution, the pdf outlined in equation (4) has to satisfy the following specifications:

$$\int_{-\infty}^{\infty} f(x;\beta,\theta)dx=1. \quad (5)$$

To verify this condition, replacing equation (4) with equation (5).

$$\int_{-\infty}^{\infty} f(x;\beta,\theta)dx=\sqrt{\frac{2}{\pi}}\frac{\theta}{\beta^{3}}\int_{-\infty}^{\infty} x^{\theta-1}\mathrm{e}^{-\frac{1}{2\beta^{2}}\left(e^{x^{\theta}}\right)^{2}}e^{3x^{\theta}}dx. \quad (6)$$

Let

$$w=\left(e^{x^{\theta}}\right)^{2}, \text{ and then } dx=\frac{dw}{2\theta x^{\theta-1}e^{2x^{\alpha}}}. \quad (7)$$

Putting equation (7) into equation (6) yields

$$\int_{-\infty}^{\infty} f(x;\beta,\theta)dx=\frac{\sqrt{2}}{2\beta^{3}\sqrt{\pi}}\int_{0}^{\infty} w^{\frac{1}{2}}\mathrm{e}^{-\frac{w}{2\beta^{2}}}dw. \quad (8)$$

Letting,

$$y=\frac{w}{2\beta^{2}}, \text{ and implies } dw=2\beta^{2}dy. \quad (9)$$

When equation (9) is substituted for equation (8), the result is

$$\int_{-\infty}^{\infty} f(x;\beta,\theta)dx=\frac{2}{\sqrt{\pi}}\int_{0}^{\infty} y^{\left(\frac{1}{2}+1\right)-1}\mathrm{e}^{-y}dy$$

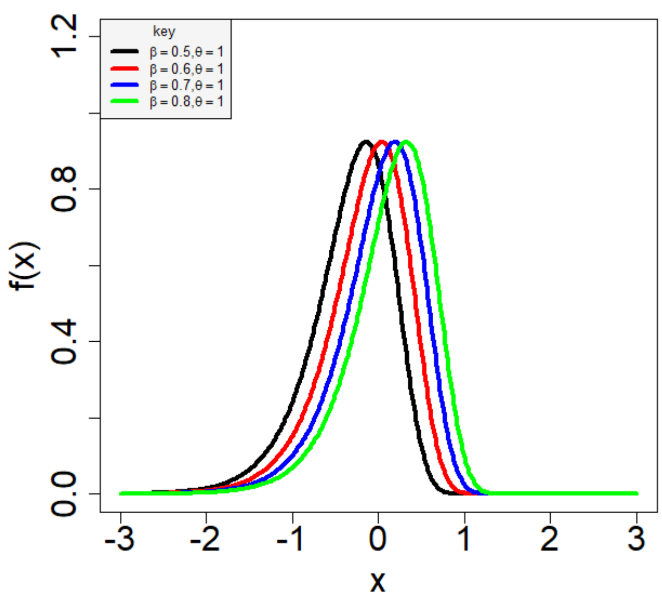
$$=\frac{2}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}+1\right)=1 \quad (10)$$

Equation (10) demonstrates that the TLMax distribution has a valid pdf, implying that the suggested distribution could be utilized to describe real-world occurrences.

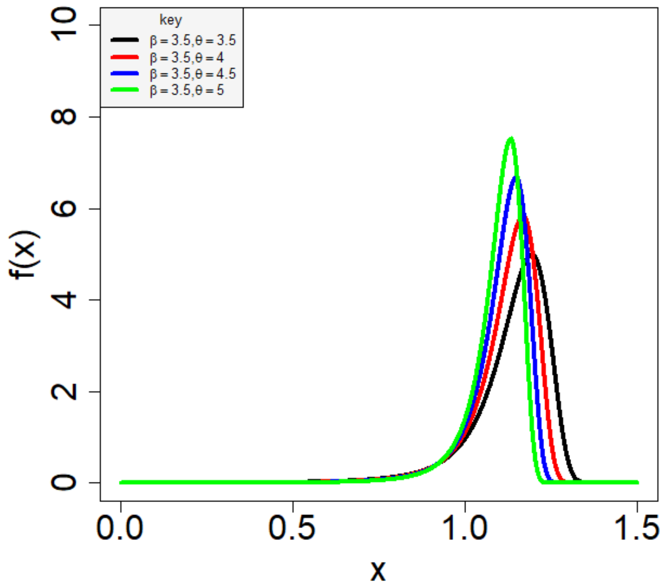
To obtain the CDF of the TLMax distribution, we differentiate equation (4) for x and the results are presented as

$$F(x;\beta,\theta)=\frac{2}{\sqrt{\pi}}\gamma\left(\frac{3}{2},\frac{1}{2\beta^2}\left(e^{x^\theta}\right)^2\right), \qquad -\infty<x<\infty;\ \beta,\theta>0. \tag{11}$$

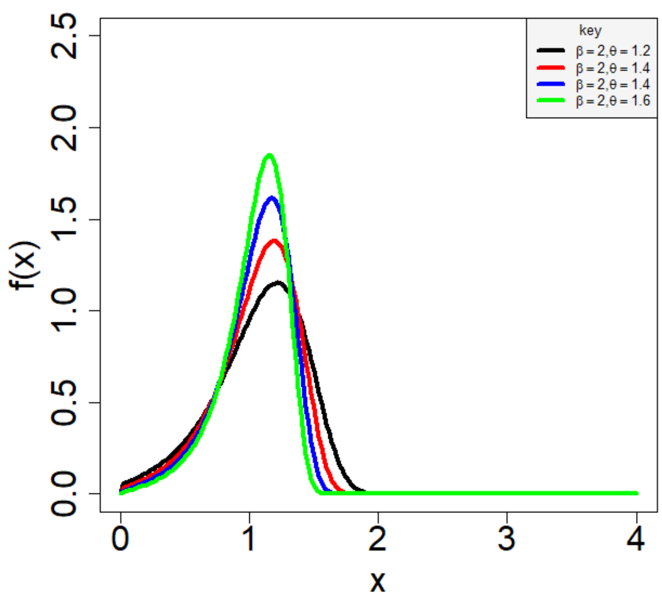
This corresponds to the CDF of the proposed TLMax distribution with the scale (β) and exponent (θ) parameters. The pdf of the TLMax distribution is plotted for a range of parameter values in Figure 1. Various plots for the suggested TLMax distribution are displayed in Figures 1(a), (b), and (c), where the distribution's shapes exhibit patterns of (a) symmetry, (b) left-skewed, as well as (c) right-skewed. This demonstrates that the suggested TLMax distribution, unlike the classic Maxwell distribution, can be used to represent data sets that are left-skewed and right-skewed.



(a)



(b)



(c)

Figure 1. Plots showing the TLMax distribution's pdf varying parameter values.

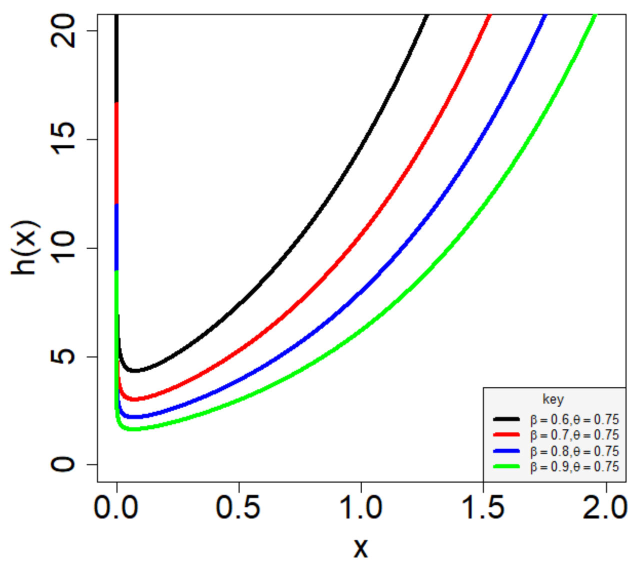
The survival function for the TLMax distribution might be determined by utilizing equation (12) as follows:

$$S(x;\beta,\theta)=1-\frac{2}{\sqrt{\pi}}\gamma\left(\frac{3}{2},\frac{1}{2\beta^2}\left(e^{x^\theta}\right)^2\right),\qquad -\infty<x<\infty;\ \beta,\theta>0. \tag{12}$$

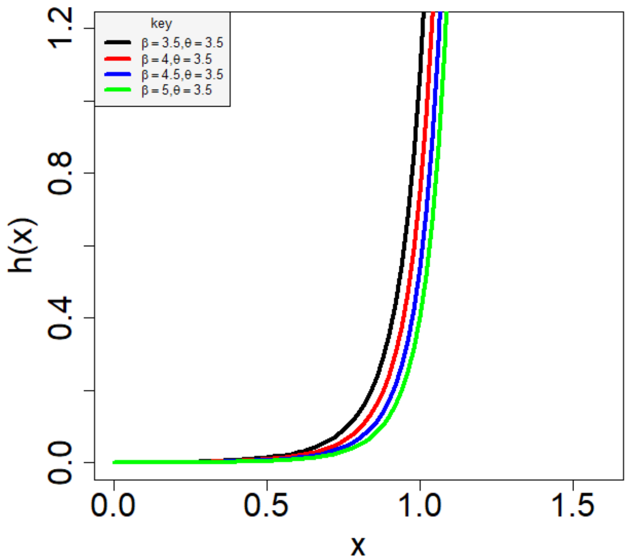
Similarly, the hazard function (hf) can be evaluated using equations (4) and (12), and is expressed as:

$$h(x;\beta,\theta)=\frac{\theta\sqrt{2}e^{3x^\theta}x^{\theta-1}\mathrm{e}^{-\frac{1}{2\beta^2}\left(e^{x^\theta}\right)^2}}{\sqrt{\pi}\beta^3\left\{1-\frac{2}{\sqrt{\pi}}\gamma\left(\frac{3}{2},\frac{1}{2\beta^2}\left(e^{x^\theta}\right)^2\right)\right\}},\qquad -\infty<x<\infty;\ \beta,\theta>0. \tag{13}$$

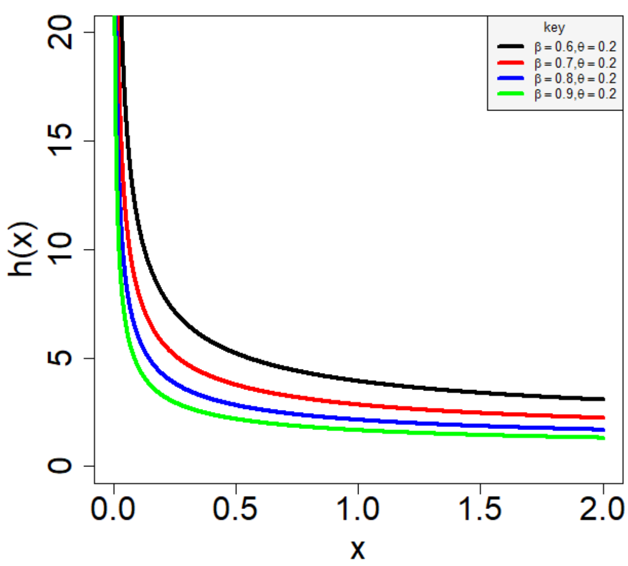
Figure 2 depicts various shapes for the hf of the TLMax distribution. Figures 2(a), (b), and (c) demonstrate the hf for the suggested distribution across parameter values.



(a)



(b)



(c)

Figure 2. Plots showing the TLMax distribution's hazard function varying parameter values.

The hazard shapes for the TLMax distribution follow the pattern of (a) bathtub, (b) increasing, and (c) decreasing failure, as provided in Figure 2(a), (b), as well as (c). This clearly explained that the proposed model might be used to describe real-world occurrences as an alternative to the Maxwell distribution and other existing distributions.

2.3. Statistical Properties

2.3.1. Mixture Representation

Using the exponential expansion, which is defined as:

$$e^{-x}=\sum_{j=0}^{\infty}\frac{(-x)^j}{j!}. \tag{14}$$

After considering the expansion in equation (14), then equation (4) becomes

$$f(x;\beta,\theta)=\sqrt{\frac{2}{\pi}}\frac{\theta x^{\theta-1}e^{3x^{\theta}}}{\beta^3}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!\left(2\beta^2\right)^j}e^{2jx^{\theta}}$$

$$=\sqrt{\frac{2}{\pi}}\frac{\theta x^{\theta-1}}{\beta^3}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!\left(2\beta^2\right)^j}e^{x^{\theta}(3+2j)}. \tag{15}$$

This represents the TLMax distribution's pdf.

2.3.2. Quantile function

For computing the quantile function (qf) of the TLMax distribution, invert equation (11) as follows:

$$\frac{1}{2\beta^2}e^{2x_q{}^{\theta}}=\gamma^{-1}\left(\frac{3}{2},u\Gamma\left(\frac{3}{2}\right)\right), \tag{16}$$

where $x=x_q{}^{\theta}$ and $0<u<1$ is the uniform random variable. Equation (16) can therefore be written as:

$$e^{2x_q{}^{\theta}}=2\beta^2\gamma^{-1}\left(\frac{3}{2},u\Gamma\left(\frac{3}{2}\right)\right).$$

From now on, the qf can be determined as:

$$x_q=\left\{\frac{1}{2}\log\left(2\beta^2\gamma^{-1}\left(\frac{3}{2},u\Gamma\left(\frac{3}{2}\right)\right)\right)\right\}^{\frac{1}{\theta}}. \tag{17}$$

2.3.4. Moments

Assume X has the TLMax distribution whose pdf defined in equation (15), then the moments of X is determined by the relation:

$$E(X^r)=\frac{\theta}{\beta^3}\sqrt{\frac{2}{\pi}}\sum_{j=0}^{\infty}\frac{(-1)^j}{\left(2\beta^2\right)^j j!}\int_{-\infty}^{\infty}x^{\theta+r-1}e^{x^{\theta}(3+2j)}dx$$

$$=\frac{2\theta}{\beta^3}\sqrt{\frac{2}{\pi}}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!\left(2\beta^2\right)^j}\int_{0}^{\infty}x^{r+\theta-1}e^{-x^{\theta}(-(3+2j))}dx. \quad (18)$$

Let,

$m=x^{\theta}\left(-(3+2j)\right)$, then implies $dx=-\frac{dm}{\theta(3+2j)x^{\theta-1}}$. (19)

Equation (18) is obtained by incorporating equation (19)

$$E(X^r)=\frac{-2}{\beta^3(3+2j)}\sqrt{\frac{2}{\pi}}\sum_{j=0}^{\infty}\frac{(-1)^j}{\left(2\beta^2\right)^j j!}\int_{0}^{\infty}\left(\frac{-m}{3+2j}\right)^{\frac{r}{\theta}}e^{-m}dm$$

$$=\frac{2}{\beta^3}\sqrt{\frac{2}{\pi}}\sum_{j=0}^{\infty}\frac{(-1)^j\,\Gamma\left(1+\frac{r}{\theta}\right)}{j!\left(2\beta^2\right)^j\left\{-(3+2j)\right\}^{1+\frac{r}{\theta}}}, \quad (20)$$

For $r=1, 2,\ldots,$ it becomes the first as well as second moments for the TLMax distribution.

2.4. Parameter Estimation

This section shows how the proposed TLMax distribution parameters are determined using maximum likelihood estimators (MLE).

Consider random variables X_i for $i=1,\ldots,n$ presents the random sample of size n with observed values x_i which was drawn from the TLMax distribution. Suppose $\varpi=(\beta,\theta)^T$ be the $p\times 1$ vector parameter, for determining the MLE of the parameter ϖ, equation (4) can be utilized for presenting the likelihood function.

$$\ell\left(x_i/\varpi\right)=\left(\frac{\theta}{\beta^3}\sqrt{\frac{2}{\pi}}\right)^n\prod_{i=1}^{n}\left(x_i^{\theta-1}e^{3x_i^{\theta}}\,\mathrm{e}^{-\frac{1}{2\beta^2}\left(e^{x_i^{\theta}}\right)^2}\right). \quad (21)$$

We can obtain the log-likelihood function for the TLMax distribution by employing equation (21) as:

$$\ell=nl\log\left(\frac{\theta}{\beta^3}\sqrt{\frac{2}{\pi}}\right)+(\theta-1)\sum_{i=1}^{n}l\log\left(x_i\right)+3\sum_{i=1}^{n}\left(x_i^{\theta}\right)-\frac{1}{2\beta^2}\sum_{i=1}^{n}\left(e^{2x_i^{\theta}}\right). \quad (22)$$

Henceforth, the estimates of the parameters β, and θ can be determined by setting the results to zero after partially differentiating equation (22) about the parameter β, and θ.

3. Methods

3.1. Machine Learning Applications in Predicting Failure Times

In the preceding section, we analyzed the dataset concerning the failure times of Kevlar 49/epoxy strands under 90% pressure, as investigated by [26]. This section shifts focus to the application of predictive machine learning algorithms on the same dataset. The primary objective aims to assess how well these machine learning models can forecast failure times.

The machine learning models employed include K-nearest neighbors regression (KNN), support vector Regression (SVR), linear regression (LR), and gradient boosting regression (XGBoost). The dataset has been split into training and testing sets to aid in the evaluation of the model. Specifically, 80% of the data is allocated for model fitting, while the remaining 20% is reserved for the comparison of models, following the approach outlined in [25].

Detailed explanations of each technique used in the modeling process are provided below:

3.1.1. Linear Regression (LR):

LR aims to find the best-fitting linear relationship between the input feature X as well as the output Y [26]. The mathematical form of LR can be described as:

$$y = \beta_0 + \beta_1 . X_1 + \beta_2 . X_2 + ... + \beta_n . X_n + \in, \tag{23}$$

where β_0 is the intercept, $\beta_1, \beta_2, ..., \beta_n$ represents the coefficients, and $\in$ is the error term.

3.1.2. Support Vector Regression (SVR):

SVR might be a regression approach that uses support vectors to determine which hyperplane best matches the data [27]. This equation may be written as:

$$y = \sum_{i=1}^{N} \left(\alpha_i - \alpha_i^* \right) . K \left(X_i , X \right) + b, \tag{24}$$

where y is the predicted output, α_i, α_i^* are Lagrange multipliers, $K\left(X_i, X\right)$ is the kernel function, and b is the bias term.

3.1.3. K-Nearest Neighbors Regression (KNN):

KNN predicts through averaging the outputs of k nearest neighbors [28]. The KNN can be described as:

$$\hat{y} = \frac{1}{k}\sum_{i=1}^{k} y_i, \tag{25}$$

where $\hat{y}$ is the predicted output and y_i is the output of the ith nearest neighbor.

3.1.4. Gradient Boosting Regression (XGBoost):

XGBoost develops an additive model in stages [29]. The XGBoost can be expressed as:

$$\hat{y} = \sum_{i=1}^{N} \alpha_i \,.\, f_i(X), \tag{26}$$

where, $\hat{y}$ is the predicted output, N is the number of boosting stages, α_i is the contribution of the ith stage, $f_i(X)$ is the prediction of the ith weak learner.

The effectiveness of all predictive models is assessed by employing standard performance metrics computed from a testing dataset. From a statistical standpoint, predictive errors serve as more appropriate criteria for evaluating prediction capability and selecting the optimal model [28]. Mean squared error (MSE), mean absolute error (MAE), coefficient of determination (R-squared), and explained variance are often used metrics.

4. Results

4.1. Application to Failure Times Data Set

This portion presents an application for the data set relating to the failure times of Kevlar 49/epoxy strands with pressure at 90%, which is reported in [30].

The box plot, kernel density, histogram, and violin plot are all shown in Figure 3 as descriptive data plots. The data has positive skewness, as evidenced by the histogram and kernel density plots, and the box and violin plots show the existence of outliers, indicating that the data set contains some extreme observations.

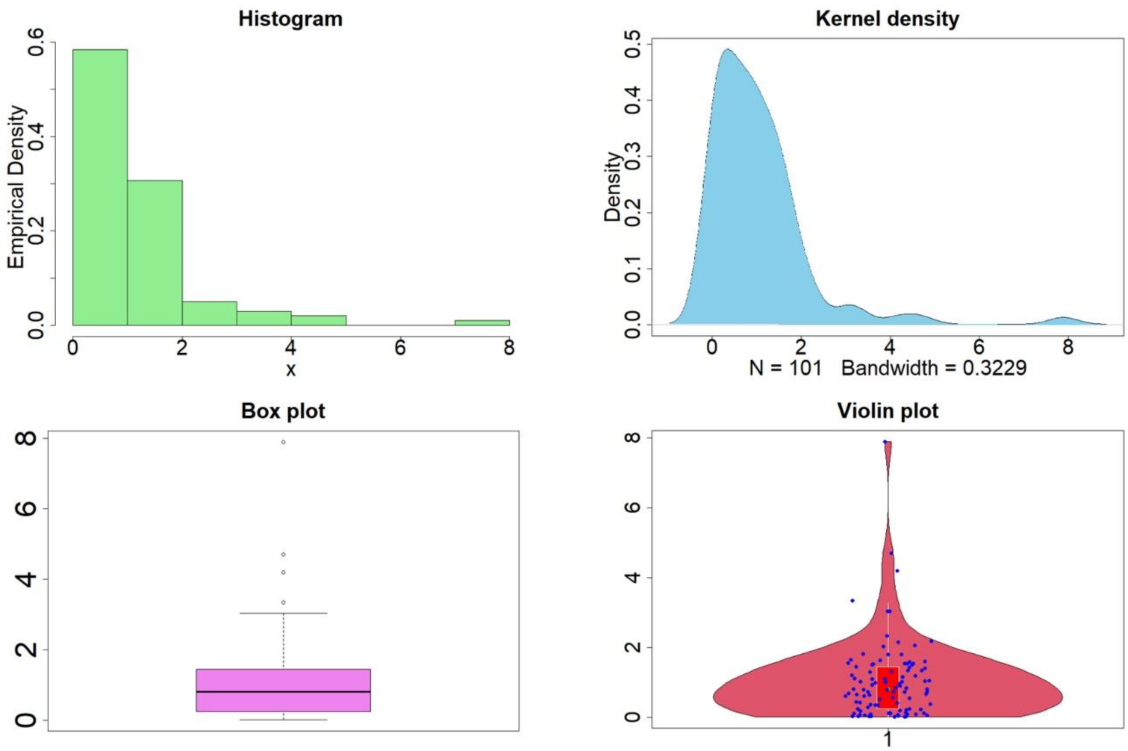


Figure 3. Descriptive plots of failure times data.

The suggested TLMax distribution will be evaluated and compared to the conventional Maxwell distribution employing the Failure Times data set. The information criteria selection model will be used to choose the optimal distribution towards competing distributions based on the minimum value of these criteria including the Hannan Quinn Information Criterion (HQIC), the Bayesian Information Criterion (BIC), the Corrected Akaike Information Criterion (CAIC), and Akaike Information Criterion (AIC). The log-likelihood (LL) value will also be considered in determining the optimum distribution by taking the highest value between.

Table 1 displays the estimated parameters, information criteria, as well as LL values, this table illustrates that the transformed log Maxwell distribution gave the lowest values of those criteria when compared to the competing one, and the proposed distribution provided the highest value when compared to the Maxwell random variable. It demonstrates that modifying the Maxwell model in terms of the Modelling Failure times data could be more effective.

Table 1: Results for the Transformed Log Maxwell distribution against the Maxwell distribution using failure times data.

Model	Estimate	AIC	HQIC	BIC	CAIC	LL
Max	$\beta = 0.8528$	531.3231	532.3818	533.9382	531.3635	-264.6616
TLMax	$\beta = 1.5854$ $\theta = 0.3167$	230.9678	233.0851	236.1980	231.0902	-113.4839

4.2. Machine Learning Empirical Results

The four standard performance metrics failure times data set are reported in Table 2. We can notice that MSE and MAE for KNN and XGBoost models are substantially smaller and higher in R-squared and Explained Variance than LR and SVR. Therefore, it can be concluded that predictions via the KNN and XGBoost models tend to perform better than the rival models' counterparts in terms of prediction.

Table 2: The performance metrics using the failure times data set

Criteria	**LR**	**SVR**	**KNN**	**XGBoost**
MSE	0.1218	0.0367	0.0021	0.0052
MAE	0.2866	0.0995	0.0287	0.0333
R-Squared	0.8214	0.9462	0.997	0.9923
Explained Variance	0.8244	0.9462	0.9971	0.9935

Figure 4 shows a visual depiction of the performance metrics for various machine learning techniques. Such graphical depiction allows for the easy identification of the top accuracy scores achieved by various model methods. Plots in Figure 4 depict that the KNN model continues to be an effective choice for predicting failure times. Furthermore, Figure 5 also shows the performance of all models, and supports the output of Figure 4. The observations highlight the ability of machine learning models to properly anticipate failure times, emphasizing their importance in addressing challenges related to real-life data.

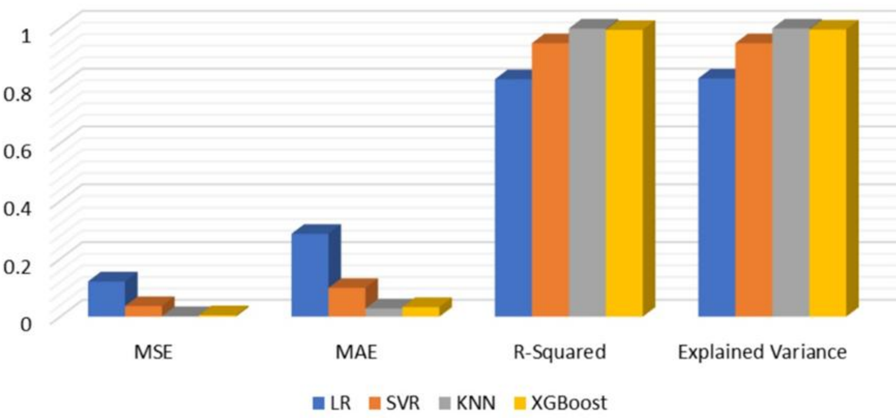


Figure 4. Machine learning algorithms' comparative performance measures.

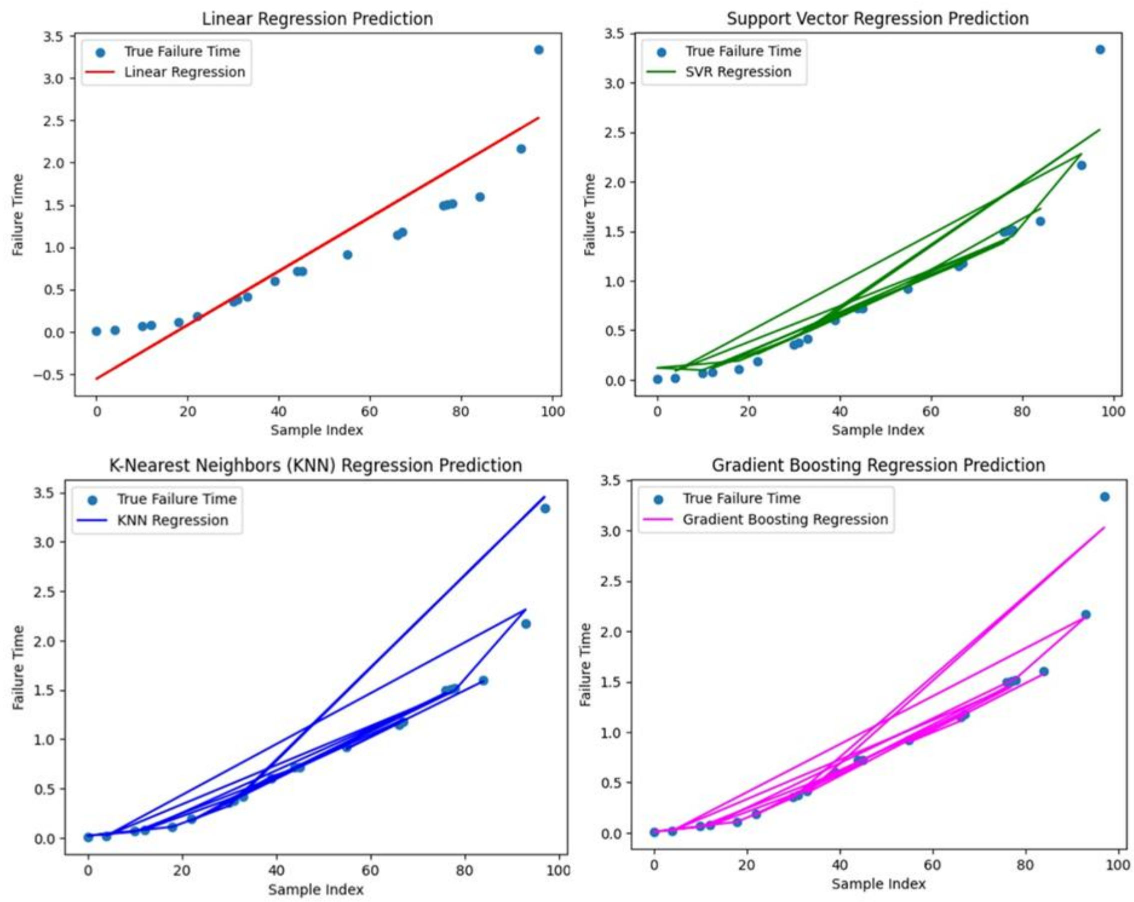


Figure 5. Comparison of LR, SVR, KNN, and XGBoost.

5. Conclusion

This paper introduces a novel statistical model, the Transformed Log Maxwell (TLMax) distribution, for analyzing failure times in the field of physics. The study derived formulations and estimators for the TLMax distribution, which was then compared with classical Maxwell models. The results indicate that the TLMax model emerged as the most effective competitor for handling failure time data in physics. Furthermore, the research delved into the predictive capabilities of machine learning models, it was obtained from the results that the KNN model demonstrated superior predictive performance across all criteria, suggesting its efficacy for policymakers in predicting failure times in physics.

Future research involves the generalization of the distribution to other forms and exploring its applications in regression analysis across various domains.

Declaration of competing interest: The authors state that they have not discovered any competing financial or personal ties that might have influenced the research disclosed in this study.

Acknowledgments: The authors would like to express their heartfelt gratitude to the Faculty of Data Science and Information Technology at INTI International University in Malaysia for providing state-of-the-sart research support to conduct this research.

Availability of Data: Any data that supports the findings of this study is included in the article.

Competing interests: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] A.I. Ishaq, A. Usman, M. Tasi'u, A.A. Suleiman, A.G. Ahmad, A New Odd F-Weibull Distribution: Properties and Application of the Monthly Nigerian Naira to British Pound Exchange Rate Data, in: 2022 International Conference on Data Analytics for Business and Industry (ICDABI), IEEE, Sakhir, Bahrain, 2022: pp. 326–332. https://doi.org/10.1109/ICDABI56818.2022.10041527.

[2] H. Krishna, P.S. Pundir, Discrete maxwell distribution, InterStat, 3 (2007), 1-15.

[3] J. Huang, S. Chen, Tail Behavior of the Generalized Maxwell Distribution, Commun. Stat. - Theory Methods 45 (2016), 4230–4236. https://doi.org/10.1080/03610926.2014.917678.

[4] A.I. Ishaq, A.A. Abiodun, The Maxwell–Weibull Distribution in Modeling Lifetime Datasets, Ann. Data Sci. 7 (2020), 639–662. https://doi.org/10.1007/s40745-020-00288-8.
[5] U.A. Abdullahi, A.A. Suleiman, A.I. Ishaq, A. Usman, A. Suleiman, The Maxwell–Exponential Distribution: Theory and Application to Lifetime Data, J. Stat. Model. Anal. 3 (2021), 65-80.
[6] A.I. Ishaq, A.A. Abiodun, U. Panitanarak, Modelling Nigerian Naira to CFA Franc Exchange Rates with Maxwell-Exponentiated Exponential Distribution, in: 2021 International Conference on Data Analytics for Business and Industry (ICDABI), IEEE, Sakheer, Bahrain, 2021: pp. 342–347. https://doi.org/10.1109/ICDABI53623.2021.9655900.
[7] A. Ishaq, A. Abiodun, On the Developments of Maxwell-Dagum Distribution, J. Stat. Model.: Theory Appl. 2 (2021), 1-23. https://doi.org/10.22034/jsmta.2021.2565.
[8] A.I. Ishaq, A.A. Abiodun, J.Y. Falgore, Bayesian Estimation of the Parameter of Maxwell-Mukherjee Islam Distribution Using Assumptions of the Extended Jeffrey's, Inverse-Rayleigh and Inverse-Nakagami Priors under the Three Loss Functions, Heliyon 7 (2021), e08200. https://doi.org/10.1016/j.heliyon.2021.e08200.
[9] A.A. Abiodun, A.I. Ishaq, On Maxwell–Lomax Distribution: Properties and Applications, Arab J. Basic Appl. Sci. 29 (2022), 221–232. https://doi.org/10.1080/25765299.2022.2093033.
[10] H.S. Al-Kzzaz, M.M.E. Abd El-Monsef, Inverse Power Maxwell Distribution: Statistical Properties, Estimation and Application, J. Appl. Stat. 49 (2022), 2287–2306. https://doi.org/10.1080/02664763.2021.1899143.
[11] K. Koobubpha, T. Panitanarak, P. Domthong, U. Panitanarak, The Maxwell-Burr X Distribution: Its Properties and Applications to the COVID-19 Mortality Rate in Thailand, Thailand Statistician, 21 (2023), 421-434.
[12] A.A. Suleiman, H. Daud, M. Othman, A.I. Ishaq, R. Indawati, M.L. Abdullah, A. Husin, The Odd Beta Prime-G Family of Probability Distributions: Properties and Applications to Engineering and Environmental Data, Comput. Sci. Math. Forum, 7(2023), 20. https://doi.org/10.3390/IOCMA2023-14429.
[13] A.K. Fulment, S.R. Gadde, J.K. Peter, The Odd Log-Logistic Generalized Exponential Distribution: Application on Survival Times of Chemotherapy Patients Data, F1000Research 11 (2022), 1444. https://doi.org/10.12688/f1000research.127363.1.
[14] J.T. Eghwerido, L.C. Nzei, A.E. Omotoye, F.I. Agu, The Teissier-G Family of Distributions: Properties and Applications, Math. Slovaca 72 (2022), 1301–1318. https://doi.org/10.1515/ms-2022-0089.
[15] A.A. Suleiman, H. Daud, N.S.S. Singh, M. Othman, A.I. Ishaq, R. Sokkalingam, A Novel Odd Beta Prime-Logistic Distribution: Desirable Mathematical Properties and Applications to Engineering and Environmental Data, Sustainability 15 (2023), 10239. https://doi.org/10.3390/su151310239.
[16] H.M. Alshanbari, A.M. Gemeay, A.A.-A. Hosni El-Bagoury, S.K. Khosa, E.H. Hafez, A.H. Muse, A Novel Extension of Fréchet Distribution: Application on Real Data and Simulation, Alexandria Eng. J. 61 (2022), 7917–7938. https://doi.org/10.1016/j.aej.2022.01.013.

[17] M. Othman, R. Indawati, A.A. Suleiman, M.B. Qomaruddin, R. Sokkalingam, Model Forecasting Development for Dengue Fever Incidence in Surabaya City Using Time Series Analysis, Processes 10 (2022), 2454. https://doi.org/10.3390/pr10112454.

[18] M. Bouamar, M. Ladjal, Evaluation of the Performances of ANN and SVM Techniques Used in Water Quality Classification, in: 2007 14th IEEE International Conference on Electronics, Circuits and Systems, IEEE, Marrakech, 2007: pp. 1047–1050. https://doi.org/10.1109/ICECS.2007.4511173.

[19] A. Danades, D. Pratama, D. Anggraini, D. Anggriani, Comparison of Accuracy Level K-Nearest Neighbor Algorithm and Support Vector Machine Algorithm in Classification Water Quality Status, in: 2016 6th International Conference on System Engineering and Technology (ICSET), IEEE, Bandung, 2016: pp. 137–141. https://doi.org/10.1109/ICSEngT.2016.7849638.

[20] N. Nasir, A. Kansal, O. Alshaltone, F. Barneih, M. Sameer, A. Shanableh, A. Al-Shamma'a, Water Quality Classification Using Machine Learning Algorithms, J. Water Process Eng. 48 (2022), 102920. https://doi.org/10.1016/j.jwpe.2022.102920.

[21] L. Li, J. Qiao, G. Yu, L. Wang, H.-Y. Li, C. Liao, Z. Zhu, Interpretable Tree-Based Ensemble Model for Predicting Beach Water Quality, Water Res. 211 (2022), 118078. https://doi.org/10.1016/j.watres.2022.118078.

[22] A.I. Ishaq, A.A. Suleiman, A. Usman, H. Daud, R. Sokkalingam, Transformed Log-Burr III Distribution: Structural Features and Application to Milk Production, Eng. Proc. 56 (2023), 322. https://doi.org/10.3390/ASEC2023-15289.

[23] E.M. Hashimoto, E.M.M. Ortega, V.G. Cancho, G.M. Cordeiro, The Log-Exponentiated Weibull Regression Model for Interval-Censored Data, Comput. Stat. Data Anal. 54 (2010), 1017–1035. https://doi.org/10.1016/j.csda.2009.10.014.

[24] A.I. Ishaq, A.A. Suleiman, H. Daud, N.S.S. Singh, M. Othman, R. Sokkalingam, P. Wiratchotisatian, A.G. Usman, S.I. Abba, Log-Kumaraswamy Distribution: Its Features and Applications, Front. Appl. Math. Stat. 9 (2023), 1258961. https://doi.org/10.3389/fams.2023.1258961.

[25] M. Qi, G.P. Zhang, An Investigation of Model Selection Criteria for Neural Network Time Series Forecasting, Eur. J. Oper. Res. 132 (2001), 666–680. https://doi.org/10.1016/S0377-2217(00)00171-5.

[26] S.F. Salleh, A.A. Suleiman, H. Daud, M. Othman, R. Sokkalingam, K. Wagner, Tropically Adapted Passive Building: A Descriptive-Analytical Approach Using Multiple Linear Regression and Probability Models to Predict Indoor Temperature, Sustainability 15 (2023), 13647. https://doi.org/10.3390/su151813647.

[27] M.H.D.M. Ribeiro, R.G. Da Silva, V.C. Mariani, L.D.S. Coelho, Short-Term Forecasting COVID-19 Cumulative Confirmed Cases: Perspectives for Brazil, Chaos Solitons Fractals 135 (2020), 109853. https://doi.org/10.1016/j.chaos.2020.109853.

[28] A.A. Suleiman, A.K. Yousafzai, M. Zubair, Comparative Analysis of Machine Learning and Deep Learning Models for Groundwater Potability Classification, Eng. Proc. 56 (2023), 249.

https://doi.org/10.3390/ASEC2023-15506.

[29] A.L. Catajan Jr, A.C. Fajardo, J.S. Limbago, Classification of Water Quality Index in Laguna de Bay Using XGBoost, in: 2023 20th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE, Phitsanulok, Thailand, 2023: pp. 403–408. https://doi.org/10.1109/JCSSE58229.2023.10202029.

[30] R. Al-Aqtash, C. Lee, F. Famoye, Gumbel-Weibull Distribution: Properties and Applications, J. Mod. Appl. Stat. Methods 13 (2014), 201–225. https://doi.org/10.22237/jmasm/1414815000.