Mathematics Interdisciplinary Research **10** (3) (2025) 267 – 294

Original Scientific Paper

Comparative Study of an Ensemble Machine Learning Model Versus Maximum Likelihood Model to Assess Reliability Measures in Right Censored Data Analysis

Faranak Goodarzi⋆ and *Mahsa Soheil Shamaee*

Abstract

This paper explores the estimation of a new power function under Type-II right censoring using two methods: maximum likelihood estimation (MLE) and an ensemble machine learning model based on stacking. The study aims to assess both methods' effectiveness in estimating various reliability measures, such as hazard rate, mean residual life, variance residual life, mean inactivity time, and variance inactivity time. The stacking model integrates five base models, radial basis function neural network, random forest, Support Vector Regression (SVR), Multilayer Perceptron (MLP), and gradient boosting regression trees, with an radial basis function neural network serving as a meta-learner for final predictions. Numerical experiments compare the performance of the stacking model against MLE for Type-II censored data. Results indicate that the stacking model significantly enhances the accuracy of reliability measure predictions, showcasing its potential as a robust tool for reliability analysis in the context of Type-II censoring.

Keywords: Reliability function, Maximum likelihood estimation, Ensemble learning, Stacking model, Type-II censoring.

2020 Mathematics Subject Classification: 92B20, 90B25.

How to cite this article
F. Goodarzi and M. Soheil Shamaee, Comparative study of an ensemble machine learning model versus maximum likelihood model to assess reliability measures in right censored data analysis, *Math. Interdisc. Res.* **10** (3) (2025) 267-294.

⋆Corresponding author (E-mail: f-goodarzi@kashanu.ac.ir)
Academic Editor: Mohammad Amini
Received 28 May 2024, Accepted 28 November 2024
DOI: 10.22052/MIR.2024.254968.1465

1. Introduction

Iqbal et al. in [1] studied a two-parameter new power function (NPF). Cumulative distribution function and probability density function of the NPF model with parameters δ and η given by

$$F(x;\delta,\eta) = 1 - \left(\frac{1-x}{1+\delta x}\right)^{\eta}, \tag{1}$$

and

$$f(x;\delta,\eta) = \frac{\eta(\delta+1)(1-x)^{\eta-1}}{(1+\delta x)^{\eta+1}}, \tag{2}$$

for $0 < x < 1$, $\delta > -1$ and $\eta > 0$. Some of valuable reliability measures are survival function $S(x)$, hazard rate function $h(x)$, cumulative hazard rate function $H(x)$, reversed hazard rate function $r(x)$, Mills ratio $M(x)$, odd function $O(x)$, mean residual lifetime $\delta(x)$, mean inactivity function $\tilde{\delta}(x)$, variance residual life function $\sigma^2(x)$, and variance inactivity time function $\tilde{\sigma}^2(x)$. See [2] for a detailed study on the given reliability measures.

In the following, we introduce the measures and evaluate them for distribution. The survival function is $S(x) = 1-F(x)$. The hazard (failure) rate function (HRF) is the probability that an item has survived time x, given that it has survived to time x and is defined as $h(x) = \frac{f(x)}{S(x)}$. The reversed hazard rate (RHRF) of X is defined as $r(x) = P(X = x|X \leq x) = \frac{f(x)}{F(x)}$. The $r(x)$ in the discrete case is interpreted as the conditional probability that a device fails at age x, given that its lifetime is at most x. The cumulative hazard rate function (CHRF) is defined as

$$H(x) = -\ln S(x) = \int_0^x h(t)dt. \tag{3}$$

It measures the total amount of risk that has been accumulated until x. Indeed, the cumulative hazard rate records the number of times that we would theoretically expect to observe the occurrence of an event. It should be noted that cumulative hazards must be interpreted based on repeated events regardless of whether the event of interest is, because of its very nature, repeatedly observable or not.

The reciprocal of the hazard rate is called the Mills ratio (see [3]). The Mills ratio has several applications in economics, mathematical statistics, and engineering.

The odd function is given as $O(x) = \frac{F(x)}{S(x)}$. A useful reliability measure of X is mean residual life, which is defined as the expectation of the residual life random variable $X_t = (X - t|X > t)$, given by

$$\mu(x) = \frac{1}{S(x)} \int_x^{\infty} S(y)dy.$$

The mean inactivity time (MIT) function, or mean past lifetime, is an important reliability measure that has applications in many fields such as survival analysis, reliability theory, and actuarial studies. Also, MRL is called the mean reversed residual life function, $\tilde{\mu}(x) = E(x - X|X \leq x)$, represents the waiting time elapsed since the failure of an item on condition that this failure had occurred in $(0, x)$ and is defined as:

$$\tilde{\mu}(x) = \frac{1}{F(x)} \int_0^x F(y)dy.$$

When discussing the variance of the residual lifetime X_x, it will be assumed that $E(X^2) < \infty$. The variance residual life (VRL) function is

$$\sigma^2(x) = Var(X - x|X > x) = \frac{2}{S(x)} \int_x^\infty (t - x)S(t)dt - \mu^2(x). \tag{4}$$

Kundu and Nanda in [4] defined the variance of the inactivity time (VIT) as:

$$\tilde{\sigma}^2(x) = Var(x - X|X \leq x) = \tilde{\mu}_2(x) - \tilde{\mu}^2(x), \tag{5}$$

where $\tilde{\mu}_2(x) = E[(x - X)^2|X \leq x] = \frac{2}{F(x)} \int_0^x (x - t)F(t)dt$.

The measures mentioned above, for the NPF model are respectively given as:

$$h(x; \delta, \eta) = \frac{\eta(\delta + 1)}{(1 + \delta x)(1 - x)}, \tag{6}$$

$$H(x; \delta, \eta) = -\log\left[\left(\frac{1 - x}{1 + \delta x}\right)^\eta\right], \tag{7}$$

$$r(x) = \frac{\eta(\delta + 1)(1 - x)^{\eta - 1}}{(1 + \delta x)\left\{(1 + \delta x)^\eta - (1 - x)^\eta\right\}}, \tag{8}$$

$$M(x; \delta, \eta) = \frac{(1 + \delta x)(1 - x)}{\eta(\delta + 1)}, \tag{9}$$

$$O(x; \delta, \eta) = \frac{(1 + \delta x)^\eta - (1 - x)^\eta}{(1 - x)^\eta}, \tag{10}$$

$$\mu(x) = \frac{(1 + \delta x)^\eta}{(1 - x)^\eta}\left(E\sum_{i=0}^\infty D_i(B(2 + i, \eta) - B_x(2 + i, \eta))\right) - x, \tag{11}$$

$$\tilde{\mu}(x) = x - \frac{(1 + \delta x)^\eta}{(1 + \delta x)^\eta - (1 - x)^\eta} E\sum_{i=0}^\infty D_i B_x(2 + i, \eta), \tag{12}$$

$$\sigma^2(x) = \frac{(1+\delta x)^\eta}{(1-x)^\eta}\left\{E\sum_{i=0}^{\infty} D_i\left(B(3+i,\eta) - B_x(3+i,\eta)\right)\right\} - 2x\left(E\sum_{i=0}^{\infty} D_i\left(B(2+i,\eta) - B_x(2+i,\eta)\right)\right) + x^2 - \mu^2(x), \qquad (13)$$

$$\tilde{\sigma}^2(x) = x^2 + \frac{(1+\delta x)^\eta}{(1+\delta x)^\eta - (1-x)^\eta}\left\{E\sum_{i=0}^{\infty} D_i\left(B(3+i,\eta)\right) - 2x\left(E\sum_{i=0}^{\infty} D_i\left(B(2+i,\eta)\right)\right)\right\} - \tilde{\mu}^2(x), \qquad (14)$$

where in Equations (11), (12), (13) and (14), $B(x;\delta,\eta) = \int_0^x y^{\delta-1}(1-y)^{\eta-1}dy$ is the beta function, $E = \eta(1+\delta)$, $D_i = \delta^i\binom{-\eta-1}{i}$, and $\delta > -1$, $\eta > 0$.

2. Related works

Ensemble methods have gained traction in survival analysis due to their ability to improve predictive accuracy and handle the complex data structures often encountered in survival studies. Here's an overview of their applications in this field.

One of the most prominent ensemble methods in survival analysis is random survival forests. This method extends the random forest algorithm to accommodate censored data, making it particularly suitable for survival outcomes. Ishwaran et al. [5] introduced random survival forests, demonstrating its effectiveness in handling right-censored data and providing variable importance measures that can guide clinical decision-making. Ensemble methods can enhance traditional Cox proportional hazards models by combining predictions from multiple models, improving robustness and accuracy. Rane et al. [6] discussed the integration of ensemble techniques, highlighting how combining models can yield better predictive performance than single-model approaches.

Boosting algorithms have also been adapted for survival analysis, allowing for sequential model building that focuses on misclassified instances. Friedman [7] introduced gradient boosting machines, which have been applied to survival data by modifying the loss function to account for censoring. Recent advancements have also seen the integration of ensemble methods with neural networks for survival analysis. Katzman et al. [8] developed a deep learning model called DeepSurv, which uses Cox proportional hazards as a loss function and can be combined with ensemble techniques to improve performance.

Combining multiple survival models into an ensemble can leverage the strengths of different methodologies. Wey et al. [9] demonstrated the effectiveness of multi-model ensembles in survival analysis by combining parametric and non-parametric models to improve prediction accuracy. Ensemble methods are increasingly applied in clinical trials to predict patient outcomes based on complex covariates. Kleinbaum et al. [10] highlighted the potential of using ensemble methods in clinical settings, emphasizing their ability to improve risk stratification and treatment decision-making.

Ensemble methods have significantly enhanced survival analysis by improving predictive accuracy and handling the complexities of censored data. Their adaptability and robustness make them valuable tools in clinical research and beyond (see [11–14]). In this paper, we present an ensemble machine-learning model specifically designed to approximate the values of functions that assess various aspects of the lifespan of a particular electronic device. We validate the model's effectiveness through experimental studies utilizing the Type-II censored data provided by [1]. Our results are compared with those from [15], demonstrating that our stacking-based machine learning model outperforms the Multilayer Perceptron employed in their work for function approximation.

3. Maximum likelihood estimation (MLE)

MLE is one of the most widely used and important methods in statistics. This method determines parameter values for which the given observations would have the highest probability. Additionally, under some regularity conditions, maximum likelihood estimators are consistent and asymptotically normally distributed (see [16]).

3.1 Learning NPF with censored data

Type-II censoring is a type of right censoring in which the study continues until the failure of the first r individuals, where r is an integer that is predetermined ($r < n$). For testing of equipment life, often experiments involving Type-II censoring are used. Here, all items are simultaneously tested, and the test is terminated when r of the n items have failed. Such an experiment may shorten the test time and save money because it could take a very long time for all components to fail. Noted that here r the number of failures and $n-r$ the number of censored observations are fixed integers and the censoring time t_r, the rth ordered lifetime is random. For a comprehensive overview of common censoring types, including right and left censoring, refer to the work of Lawlese [17].

When the data are Type-II right censored, the likelihood function for this type of censored sample can be acquired as follows (see [18]):

$$L(\eta,\delta) = \frac{n!}{(n-r)!}\prod_{i=1}^{r}\frac{\eta(\delta+1)(1-t_i)^{\eta-1}}{(1+\delta t_i)^{\eta+1}}\left(\frac{1-t_r}{1+\delta t_r}\right)^{\eta(n-r)}, \qquad (15)$$

where $t_1 < t_2 < \ldots < t_r$ and loglikelihood is

$$l(\eta, \delta) = \ln\left\{\frac{n!}{(n-r)!}\right\} + r\ln\eta + r\ln(\delta+1) + (\eta-1)\sum_{i=1}^{r}\ln(1-t_i) \qquad (16)$$
$$-(\eta+1)\sum_{i=1}^{r}\ln(1+\delta t_i) + \eta(n-r)\left\{\ln(1-t_r) - \ln(1+\delta t_r)\right\}.$$

To find the likelihood estimates, we differentiate (16) with respect to the parameters δ and η:

$$\frac{\partial}{\partial\delta}l(\delta,\eta) = \frac{r}{\delta+1} - (\eta+1)\sum_{i=1}^{r}\frac{t_i}{1+\delta t_i} - \eta(n-r)\frac{t_r}{1+\delta t_r} = 0, \qquad (17)$$

$$\frac{\partial}{\partial\eta}l(\delta,\eta) = \frac{r}{\eta} + \sum_{i=1}^{r}\ln(1-t_i) - \sum_{i=1}^{r}\ln(1+\delta t_i) + (n-r)\Big\{\ln(1-t_r) - \ln(1+\delta t_r)\Big\} = 0. \qquad (18)$$

Hence

$$\delta = \left(\frac{1}{r}\sum_{i=1}^{r}\frac{t_i}{1+\delta t_i} + \frac{\sum_{i=1}^{r}\frac{t_i}{1+\delta t_i} + (n-r)\frac{t_r}{1+\delta t_r}}{\sum_{i=1}^{r}\ln(\frac{1+\delta t_i}{1-t_i}) + (n-r)\ln(\frac{1+\delta t_r}{1-t_r})}\right)^{-1} - 1, \qquad (19)$$

and

$$\eta = \frac{\frac{r}{\delta+1} - \sum_{i=1}^{r}\frac{t_i}{1+\delta t_i}}{\sum_{i=1}^{r}\frac{t_i}{1+\delta t_i} + (n-r)\frac{t_r}{1+\delta t_r}}. \qquad (20)$$

We do not have the explicit solutions for MLEs and optimal values of δ and η are not obtained by the last two non-linear equations. The Newton-Raphson algorithm is appropriate for obtaining the kinds of maximum likelihood (ML) estimates.

3.2 Goodness-of-fit test for NPF

Consider the case of right censoring, based on the r smallest order statistics $t_1 < t_2 < \ldots < t_r$, that are observed from a sample of size n, and we want to test whether the underlying distribution is NPF or not. Indeed, this test is equivalent to testing that the observations $u_{1:n}, u_{2:n}, \ldots, u_{r:n}$ where $u_{i:n} = F(t_i)$ come from the standard uniform disribution. Therefore to test H_0 : the sample comes from a population with distribution $F(t;\delta,\eta)$, by the probability integral transformation, it is enough to use the uniformity test. Goodness of fit tests based on empirical distribution functions (EDF) with Type-II censored samples for parametric distributions have been investigated by [19] and [20]. To apply these tests, we must first estimate the parameters. Then some statistics are computed for each sample. The first statistic is the modified Kolmogorov-Smirnov statistic

$$D_{r,n} = \sup_{u \leq u_{r:n}} |F_n(u) - u|. \tag{21}$$

Also, the modified Cramér-von Mises type statistics is given by

$${}^{*}W^2_{r,n} = \frac{1}{12n} + \sum_{i=1}^{r} \Big(u_{i:n} - \frac{2i-1}{2n}\Big)^2. \tag{22}$$

Based on Michael and Schucany's [21] method, we can use Anderson-Darling (AD) test for the NPF distribution. First, we use the ML method for estimating the parameters of the distribution with censored data. Second, we apply the following transformation

$$Z_{i:n} = U_{i:n} \frac{\Big[Beta(r, n-r+1; U_{r:n})\Big]^{1/r}}{U_{r:n}}, \quad i = 1, \ldots, r, \tag{23}$$

where $Beta(r, n-r+1; U_{r:n})$ is the cumulative distribution function of $U_{r:n}$ and $Z_{1:n}$, $Z_{2:n}$, …, $Z_{r:n}$ are distributed as order statistics from a complete sample of size r, from the standard uniform distribution. Then, the modified AD statistics to Type-II right censored data is computed by

$$AD_{r,n} = -r - \frac{1}{r} \sum_{i=1}^{r} \Big[(2i-1) \ln Z_{i:n} + (2r+1-2i) \ln(1 - Z_{i:n})\Big], \tag{24}$$

and is compared with the percentage points given in Stephens [22] for the complete sample case.

4. Architecture of our proposed machine learning model for survival analysis

The stacking method, often referred to as stacked generalization [23, 24], is an ensemble learning technique that combines multiple predictive models to improve overall performance. In this approach, several base models are trained on the same dataset, and their predictions are then used as input features for a higher-level model, known as a meta-learner. This meta-learner learns to weigh the predictions from the base models, effectively capturing their strengths and compensating for their weaknesses. By leveraging the diversity of the base models, stacking can enhance predictive accuracy and robustness, particularly in complex tasks where individual models may struggle. The key advantage of stacking lies in its ability to integrate various algorithms and approaches, leading to more reliable and generalized predictions. In this study, we employed a stacking model to predict the values of the Survival, HRF, CHRF, RHRF, Mills Ratio, Odd, MIT, MRL, VIT, and VRL functions. Our proposed stacking model consists of five base models which are described as follows:

- A radial basis function neural network [25] is a type of artificial neural network that uses radial basis functions as activation functions. It consists of three layers: an input layer, a hidden layer with RBF neurons, and an output layer. The hidden layer neurons compute the distance between the input vector and a set of prototype vectors (centers), applying a Gaussian function to produce outputs sensitive to the input's proximity to these centers. The Gaussian function is defined as follows:

$$f(x) = a \cdot e^{-\frac{(x-b)^2}{2c^2}}, \tag{25}$$

  where a represents the curve's height, b is the center, and c controls the width. This structure allows radial basis function neural networks to effectively model complex, nonlinear relationships in data, making them useful for tasks such as function approximation. The training process involves determining the optimal positions of the centers, height, and width, and adjusting the weights in the output layer, often leading to fast convergence and good generalization capabilities. This base model has 10 neurons in the hidden layer.

- The other base model is a MLP structure with feed-forward and backpropagation. MLP networks, composed of input, hidden, and output layers, stand out as one of the most favored artificial neural network models due to their remarkable learning capabilities [26]. Every layer within MLP is directly connected to the subsequent layer. The Levenberg-Marquardt training algorithm, renowned for its high learning capability [27–29], was favored. The sigmoid function is selected for the hidden layer as the activation function, while the Purelin function is chosen for the output layer. These activation functions are defined as follows:

$$f(x) = \frac{1}{1 + \exp(-x)}, \tag{26}$$

$$Purelin(x) = x. \tag{27}$$

  This base model has 8 neurons in the hidden layer.

- Support Vector Machine (SVM) [30, 31] for function approximation, often referred to as SVR, is a powerful technique that aims to find a function that closely fits a set of data points while maintaining a balance between complexity and accuracy. Instead of merely fitting a line or curve through the data, SVR identifies a hyperplane in a high-dimensional space that best represents the underlying relationship between input features and continuous output values. It introduces a margin of tolerance, defined by a parameter ϵ, within which errors are not penalized, allowing for some flexibility in the model. The focus is primarily on support vectors—data points that lie outside this margin—because they are crucial in shaping the regression function. By utilizing kernel functions, SVR can efficiently model complex, non-linear relationships, making it suitable for various real-world applications where precise function approximation is required. The kernel function employed in our proposed SVR model is Gaussian, and the optimal values for the parameters Box constraint, ϵ, and kernel scale were determined through cross-validation.

- Random forest [32] is an ensemble learning method that constructs multiple decision trees during training and outputs the average prediction for regression tasks or the majority vote for classification. For function approximation, it operates by aggregating the predictions from a diverse set of trees, each built on a random subset of the training data and features. This randomness helps to reduce overfitting and increase generalization by capturing various patterns in the data. The final prediction is typically more robust and accurate than that of individual trees, as it mitigates the impact of noise and variance inherent in the training data. Random forest is particularly effective for complex, non-linear relationships in high-dimensional spaces, making it a popular choice for various regression tasks. The proposed random forest model consists of 100 trees.

- Gradient boosting regression trees [7] is an ensemble learning technique that builds a predictive model by combining multiple weak learners, typically decision trees, in a sequential manner. The core idea is to iteratively fit new trees to the residual errors made by the existing ensemble. Initially, a simple model is trained on the data, and subsequent trees are constructed to correct the predictions by minimizing the loss function, often using gradient descent. Each new tree focuses on the errors of the previous trees, effectively refining the overall model. This process continues until a specified number of trees are created or no significant improvement can be made.

For each base model, the parameters lifetime of electronic devices (t), scale (δ), and shape (η) have been designated as input values. At the output layer, the base model predicts the value of the survival, HRF, CHRF, RHRF, Mills ratio, Odd, MIT, MRL, VIT, and VRL functions. The fundamental arrangement illustration of base models are denoted in Figures 1 and 2.

Figure 3 presents our proposed stacking model, in which the meta-model is a radial basis function neural network due to its effectiveness in regression tasks. In our experimental studies, we implemented cross-validation on the base models used in the stacking process to enhance efficiency.

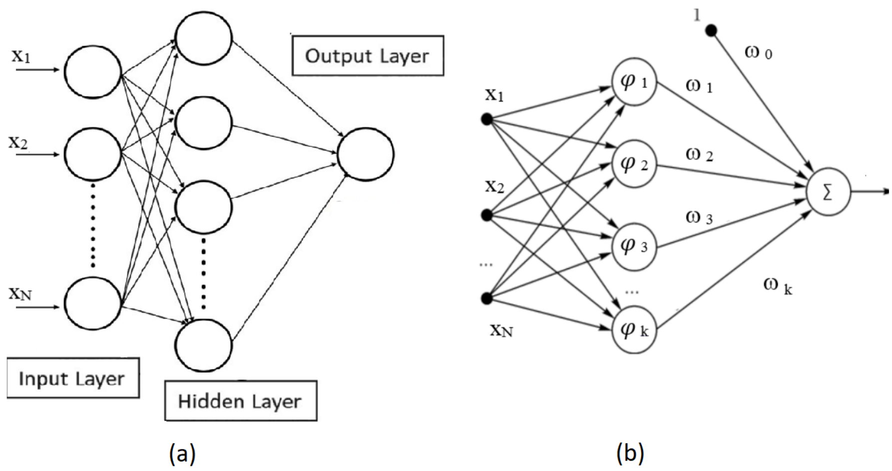

Figure 1: The fundamental arrangement of (a) a MLP, and (b) a radial basis function neural network where $x_1, x_2, ..., x_N$ are input values .

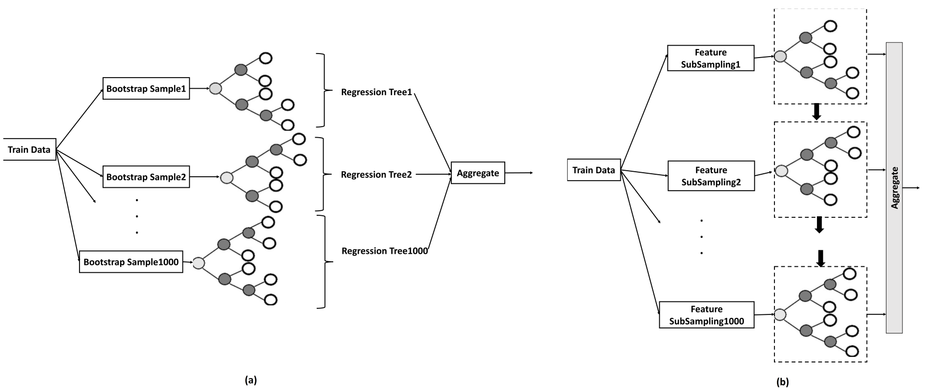

Figure 2: The fundamental arrangement of (a) a random forest, and (b) Gradient boosting regression trees.

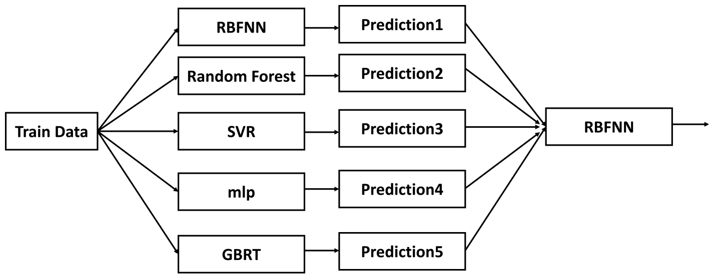

Figure 3: Our proposed stacking model consists of five base models where the meta-model is a radial basis function neural network.

4.1 Effectiveness of the stacking model

Following the design of the stacking model, the subsequent step involves conducting a thorough analysis of learning and training performance. When the training phases of the learning models are not completed optimally, the prediction error rates will be high, rendering the predictions unreliable. Therefore, it's crucial to ensure that the training stage of the model is effectively executed before deriving predictive values from the model.

To assess the learning reliability of the developed stacking model, standard functionality parameters, such as Mean Squared Error (MSE) and correlation coefficient (R), were calculated, followed by a comprehensive analysis of the obtained results. The mathematical expressions used in evaluating the efficiency parameters are provided below.

$$MSE = \frac{1}{N} \sum_{i=1}^{i=N} (X_{targ(i)} - X_{pred(i)})^2, \tag{28}$$

$$R = \sqrt{1 - \sum_{i=1}^{i=N} (X_{targ(i)} - X_{pred(i)})^2 \{\sum_{i=1}^{i=N} (X_{targ(i)})^2\}^{-1}}. \tag{29}$$

5. A real data set analysis

In this section, our objective is to juxtapose the results derived from the MLE technique with those produced by our proposed stacking model. The dataset in question has been meticulously scrutinized by [1]. They explored the longevity (measured in days) of 30 electronic devices, with the dataset comprising values such as 0.020, 0.029, 0.044, 0.057, 0.034, 0.096, 0.139, 0.156, 0.106, 0.164, 0.177, 0.250, 0.326, 0.167, 0.607, 0.672, 0.650, 0.406, 0.736, 0.676, 0.817, 0.838, 0.910, 0.931, 0.946, 0.953, 0.961, 0.981, 0.982, and 0.990. Within this paper, we utilize a subset of the data above, specifically 0.020, 0.029, 0.044, 0.057, 0.034, 0.096, 0.139, 0.156, 0.106, 0.164, 0.177, 0.250, 0.326, 0.167, 0.607, 0.672, 0.650, 0.406, 0.736, 0.676, 0.817, 0.838.

To assess whether these lifespans follow an NPF distribution, we calculated the EDF statistics as follows: $D_{r,n} = 0.0978$, $^*W_{r,n} = 0.0321$, and $AD_{r,n} = 0.349$. Since the value of $AD_{r,n}$ does not exceed the critical point of 0.025 as per Stephens [22], we can conclude that the censored data conforms to an NPF distribution. Our conclusion is further supported by consulting Table 1.0 in [22] and computing the modified versions of $D_{r,n}$ and $^*W_{r,n}$, which yield the same result.

We note that the asymptotic distribution of the MLE of $\boldsymbol{\theta} = (\delta, \eta)$ is $N(\boldsymbol{\theta}, [I(\boldsymbol{\theta})]^{-1})$ where $I(\boldsymbol{\theta}) = -E[H(\boldsymbol{\theta})]$ and $[I(\boldsymbol{\theta})]^{-1}$ is given by

$$[I(\boldsymbol{\theta})]^{-1} = \frac{1}{r}\begin{bmatrix} \eta^2(\eta+1)^2 & -\eta(\eta+1)(\eta+2)(\delta+1) \\ -\eta(\eta+1)(\eta+2)(\delta+1) & \frac{(\eta+1)^2(\eta+2)(\delta+1)^2}{\eta} \end{bmatrix}. \tag{30}$$

Here, $H(\boldsymbol{\theta})$ is the Hessian matrix that is constructed of second derivatives of the log likelihood with respect to the parameters (see [33]). Now, based on the asymptotic normality of estimators, we find the following two-sided 95% confidence intervals for δ and η as follows:

$$\delta \in \left(\hat{\delta} e^{-\frac{z_{0.975}}{\hat{\delta}}\sqrt{\frac{1}{22}\frac{(\hat{\eta}+1)^2(\hat{\eta}+2)(\hat{\delta}+1)^2}{\hat{\eta}}}}, \hat{\delta} e^{\frac{z_{0.975}}{\hat{\delta}}\sqrt{\frac{1}{22}\frac{(\hat{\eta}+1)^2(\hat{\eta}+2)(\hat{\delta}+1)^2}{\hat{\eta}}}}\right), \tag{31}$$

and

$$\eta \in (\hat{\eta} e^{-\frac{z_{0.975}}{\hat{\eta}}\sqrt{\frac{1}{22}\hat{\eta}^2(\hat{\eta}+1)^2}}, \hat{\eta} e^{\frac{z_{0.975}}{\hat{\eta}}\sqrt{\frac{1}{22}\hat{\eta}^2(\hat{\eta}+1)^2}}), \tag{32}$$

where $z_{0.975}$ is the 97.5th percentile of the standard normal distribution and hence $\delta \in (1.5548, 42.7775)$ and $\eta \in (0.1896, 0.5765)$.

Table 1 presents the mean squared error of our proposed stacking model compared to the MLP model proposed in [15] for approximating the specified functions. It is evident that the MSE of our proposed stacking model consistently outperforms that of the MLP across all functions. Correlation coefficient (R) for both models for approximating the specified functions equals one, making direct comparison unfeasible.

Tables 2 to 11 showcase the patterns of various functions related to the lifetime of 22 electronic devices, including the hazard rate (HRF), RHRF, CHRF, Mills ratio, odd function, survival function (SF), mean residual life (MRL), mean inactivity time (MIT), variance residual life (VRL), and variance inactivity time (VIT). These functions were derived through numerical estimation using the maximum likelihood method and predictions made by the stacking model. The results obtained through numerical methods align closely with those predicted by the stacking model, highlighting the effectiveness of stacking predictions in the analysis and assessment of reliability metrics within survival studies.

Table 1: Comparison of the MSE between our proposed stacking model and the MLP model presented in [15].

	Stacking	MLP
Mills Ratio	**3.66E-11**	2.83E-08
Cumulative Hazard Rate Function (CHRF)	**1.56E-08**	1.92E-07
Hazard Rate Function (HRF)	**1.51E-08**	8.22E-08
Odd Function	**6.09E-07**	2.20E-06
Survival Function	**2.49E-09**	8.64E-09
Reversed Hazard Rate Function (RHRF)	**0.048285**	0.059294
Mean Residual Life (MRL) Function	**9.59E-11**	1.89E-09
Mean Inactivity Time (MIT) Function	**4.01E-12**	9.59E-09
Variance Inactivity Time (VIT) Function	**6.28E-14**	3.63E-08
Variance Residual Life (VRL) Function	**2.96E-12**	7.10E-11

Table 2: Predicted values of hazard rate function.

Observation	$\hat{\delta}$	$\hat{\eta}$	Hazard rate function			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			2.6554	2.6555	-0.0002	-4.59E-06
0.029			2.5210	2.5209	0.0025	6.34E-05
0.034			2.4531	2.4532	-0.0047	-1.16E-04
0.044			2.3300	2.3300	0.0022	5.22E-05
0.057			2.1912	2.1911	0.0028	6.09E-05
0.096			1.8779	1.8780	-0.0012	-2.24E-05
0.106			1.8159	1.8158	0.0033	6.07E-05
0.139			1.6477	1.6476	0.0019	3.21E-05
0.156			1.5783	1.5783	-0.0016	-2.52E-05
0.164			1.5489	1.5490	-0.0029	-4.44E-05
0.167			1.5384	1.5384	-0.0033	-5.06E-05
0.177			1.5051	1.5052	-0.0046	-6.92E-05
0.25			1.3280	1.3281	-0.0078	-1.03E-04
0.326			1.2274	1.2271	0.0258	3.17E-04
0.406			1.1820	1.1821	-0.0091	-1.08E-04
0.607			1.2943	1.2947	-0.0258	-3.34E-04
0.65			1.3725	1.3723	0.0117	1.60E-04
0.672			1.4240	1.4239	0.0073	1.04E-04
0.676			1.4343	1.4342	0.0078	1.12E-04
0.736			1.6373	1.6373	-0.0009	-1.50E-05
0.817			2.1584	2.1585	-0.0042	-9.02E-05
0.838			2.3849	2.3849	0.0010	2.39E-05

Table 3: Predicted values of reversed hazard rate function.

Observation	$\hat{\delta}$	$\hat{\eta}$	RHRF			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			45.5745	45.5745	0.0000	0.00E+00
0.029			30.3027	30.3027	0.0000	0.00E+00
0.034			25.3567	25.3567	0.0000	-1.07E-14
0.044			18.9009	18.9009	0.0000	0.00E+00
0.057			13.9789	13.9789	0.0000	-7.11E-15
0.096			7.4597	7.4454	0.1918	1.43E-02
0.106			6.6015	6.6276	-0.3946	-2.61E-02
0.139			4.7093	4.4381	5.7574	2.71E-01
0.156			4.0742	4.4789	-9.9332	-4.05E-01
0.164			3.8259	3.9984	-4.5082	-1.72E-01
0.167			3.7397	3.7825	-1.1437	-4.28E-02
0.177			3.4760	3.0476	12.3250	4.28E-01
0.25			2.2580	1.5328	32.1187	7.25E-01
0.326			1.6380	1.6890	-3.1183	-5.11E-02
0.406			1.2769	1.4664	-14.8393	-1.89E-01
0.607			0.8892	0.9742	-9.5537	-8.50E-02
0.65			0.8577	0.9333	-8.8160	-7.56E-02
0.672			0.8469	0.9190	-8.5219	-7.22E-02
0.676			0.8453	0.9123	-7.9282	-6.70E-02
0.736			0.8372	0.9043	-8.0135	-6.71E-02
0.817			0.8856	0.9714	-9.6884	-8.58E-02
0.838			0.9155	1.0154	-10.9115	-9.99E-02

Table 4: Predicted values of cumulative hazard rate function.

Observation	$\hat{\delta}$	$\hat{\eta}$	CHRF			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			0.0566	0.0567	-0.1939	-1.10E-04
0.029			0.0799	0.0799	0.0542	4.33E-05
0.034			0.0923	0.0924	-0.0961	-8.87E-05
0.044			0.1162	0.1160	0.1828	2.13E-04
0.057			0.1456	0.1455	0.0890	1.30E-04
0.096			0.2245	0.2247	-0.0703	-1.58E-04
0.106			0.2430	0.2432	-0.0655	-1.59E-04
0.139			0.3000	0.3001	-0.0270	-8.10E-05
0.156			0.3274	0.3274	0.0169	5.54E-05
0.164			0.3399	0.3399	0.0139	4.73E-05
0.167			0.3446	0.3445	0.0162	5.57E-05
0.177			0.3598	0.3597	0.0262	9.42E-05
0.25			0.4626	0.4626	-0.0079	-3.67E-05
0.326			0.5593	0.5592	0.0131	7.30E-05
0.406			0.6553	0.6554	-0.0196	-1.28E-04
0.607			0.8984	0.8983	0.0117	1.05E-04
0.65			0.9556	0.9555	0.0164	1.56E-04
0.672			0.9864	0.9863	0.0093	9.20E-05
0.676			0.9921	0.9924	-0.0335	-3.32E-04
0.736			1.0837	1.0837	0.0000	-1.51E-07
0.817			1.2347	1.2346	0.0064	7.85E-05
0.838			1.2823	1.2823	-0.0038	-4.85E-05

Table 5: Predicted values of Mills ratio.

Observation	$\hat{\delta}$	$\hat{\eta}$	Mills ratio			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			0.3766	0.3766	-0.0016	-6.02E-06
0.029			0.3967	0.3967	0.0051	2.03E-05
0.034			0.4076	0.4077	-0.0029	-1.19E-05
0.044			0.4292	0.4292	-0.0007	-3.05E-06
0.057			0.4564	0.4564	-0.0004	-1.98E-06
0.096			0.5325	0.5325	0.0010	5.56E-06
0.106			0.5507	0.5507	0.0007	3.66E-06
0.139			0.6069	0.6069	-0.0006	-3.59E-06
0.156			0.6336	0.6336	-0.0008	-4.76E-06
0.164			0.6456	0.6456	-0.0005	-3.00E-06
0.167			0.6500	0.6500	-0.0004	-2.52E-06
0.177			0.6644	0.6644	-0.0001	-5.47E-07
0.25			0.7530	0.7530	0.0001	5.81E-07
0.326			0.8147	0.8147	-0.0001	-7.69E-07
0.406			0.8460	0.8460	0.0002	1.31E-06
0.607			0.7726	0.7726	-0.0008	-6.31E-06
0.65			0.7286	0.7286	0.0006	4.46E-06
0.672			0.7023	0.7023	0.0007	4.69E-06
0.676			0.6972	0.6972	0.0006	4.18E-06
0.736			0.6108	0.6108	0.0001	5.02E-07
0.817			0.4633	0.4633	-0.0002	-7.87E-07
0.838			0.4193	0.4193	0.0000	8.96E-09

Table 6: Predicted values of odd function.

Observation	$\hat{\delta}$	$\hat{\eta}$	Odd function			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			0.0583	0.0586	-0.5052	-2.94E-04
0.029			0.0832	0.0830	0.1858	1.55E-04
0.034			0.0967	0.0965	0.2451	2.37E-04
0.044			0.1233	0.1231	0.1660	2.05E-04
0.057			0.1567	0.1568	-0.0141	-2.21E-05
0.096			0.2517	0.2523	-0.2165	-5.45E-04
0.106			0.2751	0.2756	-0.1808	-4.97E-04
0.139			0.3499	0.3498	0.0285	9.97E-05
0.156			0.3874	0.3870	0.1007	3.90E-04
0.164			0.4048	0.4044	0.1097	4.44E-04
0.167			0.4114	0.4109	0.1083	4.46E-04
0.177			0.4330	0.4326	0.0846	3.66E-04
0.25			0.5881	0.5904	-0.3811	-2.24E-03
0.326			0.7494	0.7474	0.2572	1.93E-03
0.406			0.9257	0.9265	-0.0848	-7.85E-04
0.607			1.4556	1.4549	0.0462	6.72E-04
0.65			1.6003	1.6015	-0.0784	-1.26E-03
0.672			1.6815	1.6814	0.0012	2.02E-05
0.676			1.6968	1.6961	0.0439	7.44E-04
0.736			1.9557	1.9557	-0.0035	-6.88E-05
0.817			2.4373	2.4373	0.0002	5.27E-06
0.838			2.6049	2.6049	-0.0001	-1.70E-06

Table 7: Predicted values of survival function.

Observation	$\hat{\delta}$	$\hat{\eta}$	Survival function			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			0.9449	0.9450	-0.0010	-9.82E-06
0.029			0.9232	0.9231	0.0056	5.13E-05
0.034			0.9118	0.9118	-0.0024	-2.21E-05
0.044			0.8903	0.8903	-0.0004	-3.24E-06
0.057			0.8645	0.8646	-0.0073	-6.28E-05
0.096			0.7989	0.7989	0.0013	1.01E-05
0.106			0.7843	0.7842	0.0093	7.33E-05
0.139			0.7408	0.7407	0.0098	7.25E-05
0.156			0.7208	0.7208	-0.0002	-1.67E-06
0.164			0.7118	0.7119	-0.0044	-3.12E-05
0.167			0.7085	0.7086	-0.0044	-3.11E-05
0.177			0.6978	0.6979	-0.0106	-7.40E-05
0.25			0.6297	0.6297	-0.0027	-1.67E-05
0.326			0.5716	0.5716	0.0061	3.49E-05
0.406			0.5193	0.5192	0.0104	5.42E-05
0.607			0.4072	0.4074	-0.0343	-1.40E-04
0.65			0.3846	0.3846	0.0005	2.11E-06
0.672			0.3729	0.3729	0.0120	4.46E-05
0.676			0.3708	0.3708	0.0150	5.56E-05
0.736			0.3383	0.3383	-0.0003	-1.07E-06
0.817			0.2909	0.2909	0.0038	1.12E-05
0.838			0.2774	0.2774	-0.0059	-1.65E-05

Table 8: Predicted values of MRL function.

Observation	$\hat{\delta}$	$\hat{\eta}$	MRL function			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			0.4965	0.4965	-0.0018	-9.15E-06
0.029			0.4991	0.4991	-0.0003	-1.70E-06
0.034			0.5003	0.5003	-0.0007	-3.34E-06
0.044			0.5023	0.5023	0.0016	8.10E-06
0.057			0.5041	0.5041	0.0018	8.93E-06
0.096			0.5049	0.5049	0.0001	6.12E-07
0.106			0.5042	0.5042	-0.0001	-2.65E-07
0.139			0.4999	0.4999	-0.0001	-4.12E-07
0.156			0.4965	0.4965	-0.0020	-1.00E-05
0.164			0.4947	0.4947	0.0002	1.21E-06
0.167			0.4940	0.4940	0.0004	2.14E-06
0.177			0.4915	0.4915	0.0009	4.32E-06
0.25			0.4679	0.4679	0.0003	1.51E-06
0.326			0.4357	0.4357	-0.0021	-9.12E-06
0.406			0.3957	0.3956	0.0035	1.40E-05
0.607			0.2766	0.2766	-0.0060	-1.67E-05
0.65			0.2486	0.2486	-0.0050	-1.26E-05
0.672			0.2340	0.2340	0.0021	4.94E-06
0.676			0.2314	0.2313	0.0107	2.47E-05
0.736			0.1907	0.1907	0.0001	1.02E-07
0.817			0.1340	0.1340	-0.0122	-1.63E-05
0.838			0.1190	0.1190	0.0094	1.11E-05

Table 9: Predicted values of MIT function.

Observation	$\hat{\delta}$	$\hat{\eta}$	MIT function			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			0.0103	0.0103	0.0110	1.14E-06
0.029			0.0151	0.0151	0.0085	1.28E-06
0.034			0.0179	0.0179	0.0113	2.02E-06
0.044			0.0234	0.0234	-0.0206	-4.81E-06
0.057			0.0307	0.0307	-0.0126	-3.87E-06
0.096			0.0536	0.0536	0.0059	3.17E-06
0.106			0.0596	0.0596	0.0062	3.68E-06
0.139			0.0799	0.0799	0.0020	1.59E-06
0.156			0.0906	0.0906	0.0000	-2.79E-08
0.164			0.0957	0.0957	-0.0015	-1.47E-06
0.167			0.0976	0.0976	-0.0018	-1.72E-06
0.177			0.1039	0.1039	-0.0025	-2.56E-06
0.25			0.1513	0.1513	0.0012	1.77E-06
0.326			0.2018	0.2018	-0.0001	-2.57E-07
0.406			0.2556	0.2556	0.0001	1.74E-07
0.607			0.3898	0.3898	-0.0002	-6.76E-07
0.65			0.4176	0.4176	0.0003	1.08E-06
0.672			0.4317	0.4317	0.0000	2.04E-08
0.676			0.4342	0.4342	-0.0001	-5.86E-07
0.736			0.4714	0.4714	0.0000	-1.69E-09
0.817			0.5181	0.5181	0.0001	2.83E-07
0.838			0.5292	0.5292	0.0000	-2.32E-07

Table 10: Predicted values of VRL function.

Observation	$\hat{\delta}$	$\hat{\eta}$	VRL function			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			0.1247	0.1247	0.0005	6.83E-07
0.029			0.1218	0.1218	-0.0011	-1.39E-06
0.034			0.1202	0.1202	-0.0010	-1.16E-06
0.044			0.1171	0.1171	0.0018	2.06E-06
0.057			0.1130	0.1130	0.0018	2.08E-06
0.096			0.1014	0.1014	-0.0017	-1.75E-06
0.106			0.0985	0.0985	-0.0032	-3.15E-06
0.139			0.0895	0.0895	-0.0002	-1.46E-07
0.156			0.0851	0.0851	-0.0033	-2.79E-06
0.164			0.0830	0.0830	0.0018	1.48E-06
0.167			0.0823	0.0823	0.0036	2.93E-06
0.177			0.0798	0.0798	0.0040	3.17E-06
0.25			0.0634	0.0634	-0.0024	-1.52E-06
0.326			0.0490	0.0490	-0.0030	-1.49E-06
0.406			0.0365	0.0365	0.0034	1.24E-06
0.607			0.0145	0.0145	-0.0083	-1.21E-06
0.65			0.0113	0.0113	-0.0054	-6.14E-07
0.672			0.0098	0.0098	0.0085	8.33E-07
0.676			0.0096	0.0095	0.0154	1.47E-06
0.736			0.0062	0.0062	-0.0084	-5.21E-07
0.817			0.0029	0.0029	0.0075	2.15E-07
0.838			0.0022	0.0022	-0.0183	-4.09E-07

Table 11: Predicted values of VIT function.

Observation	$\hat{\delta}$	$\hat{\eta}$	VIT function			
t	8.1555	0.3306	MLE	Stacking model	Deviation	Difference
0.02			0.0000	0.0000	-1.4331	-4.77E-07
0.029			0.0001	0.0001	0.1072	7.50E-08
0.034			0.0001	0.0001	0.1926	1.85E-07
0.044			0.0002	0.0002	0.1545	2.49E-07
0.057			0.0003	0.0003	0.0279	7.52E-08
0.096			0.0008	0.0008	-0.0410	-3.12E-07
0.106			0.0009	0.0009	-0.0332	-3.08E-07
0.139			0.0016	0.0016	-0.0073	-1.16E-07
0.156			0.0020	0.0020	0.0150	2.98E-07
0.164			0.0022	0.0022	0.0042	9.20E-08
0.167			0.0023	0.0023	0.0051	1.15E-07
0.177			0.0026	0.0026	0.0072	1.86E-07
0.25			0.0051	0.0051	-0.0001	-2.62E-09
0.326			0.0086	0.0086	0.0005	4.18E-08
0.406			0.0134	0.0134	-0.0006	-8.54E-08
0.607			0.0306	0.0306	0.0006	1.78E-07
0.65			0.0355	0.0355	-0.0020	-7.10E-07
0.672			0.0382	0.0382	0.0009	3.42E-07
0.676			0.0387	0.0387	0.0004	1.69E-07
0.736			0.0469	0.0469	0.0000	4.97E-09
0.817			0.0602	0.0602	0.0001	5.73E-08
0.838			0.0642	0.0642	-0.0001	-5.57E-08

In Figures 4 and 5, each data point used to construct the stacking model is associated with 22 distinct output values. Upon inspection of Figures 4 and 5, it is clear that the output values generated by the stacking model precisely match the target values for each data point. The alignment of the target values with the predicted values on the graph line indicates the exceptional accuracy of the results produced by the developed Stacking model. This seamless correspondence between the target and predicted values underscores the stacking model's capacity to reliably forecast the parameters crucial for assessing the dependability of electrical components.

In Figures 6 and 7, the x-axis displays the target data, while the y-axis showcases the output values generated by the stacking model. This visual representation aims to provide a comprehensive assessment of the model's predictive precision by analyzing the closeness of data points to the line of zero error. Upon scrutinizing the data points depicted in Figures 6 and 7, it is evident that the majority align precisely with the zero error line. Notably, these data points consistently fall within a $\pm 10\%$ error margin, emphasizing the exceptional accuracy and dependability of the results produced by the stacking model. To verify the precision of the forecasts generated by the stacking model, it is crucial to assess the relative

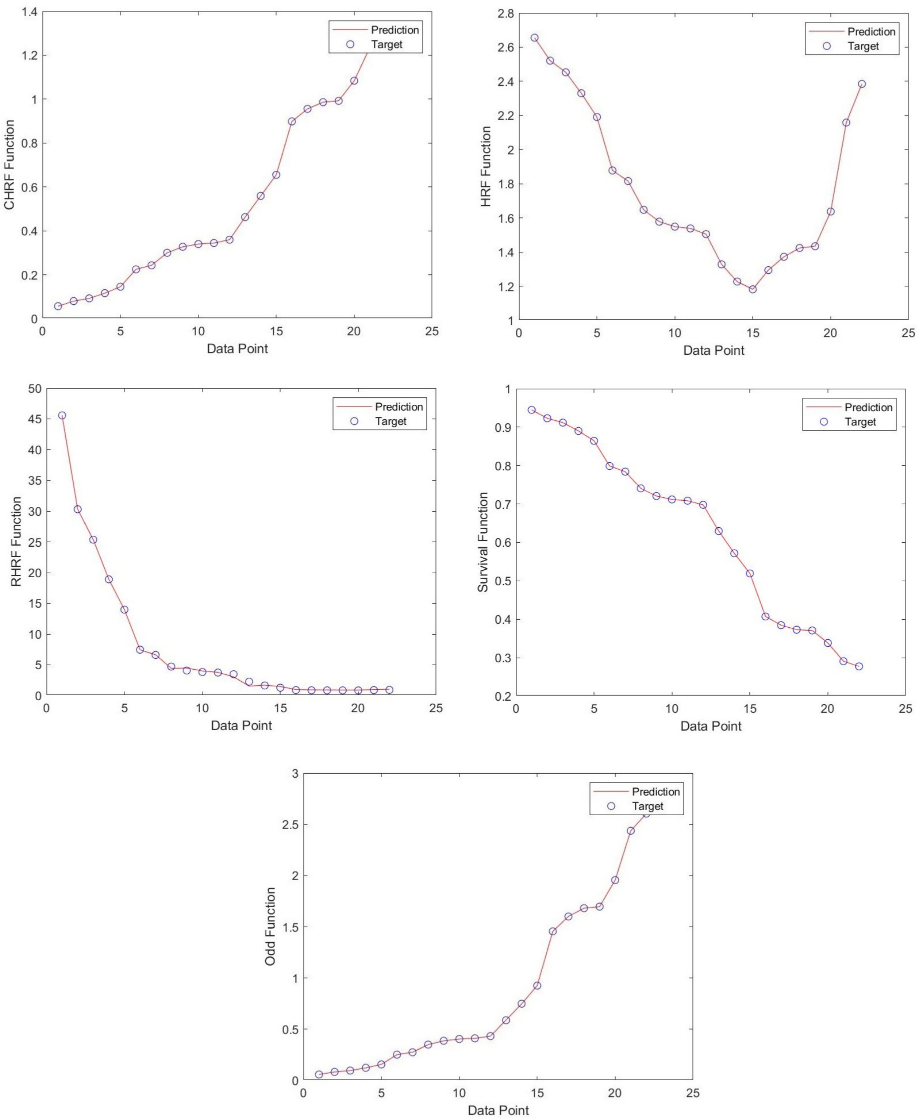

Figure 4: Output values corresponding to data employed in training the stacking model (1).

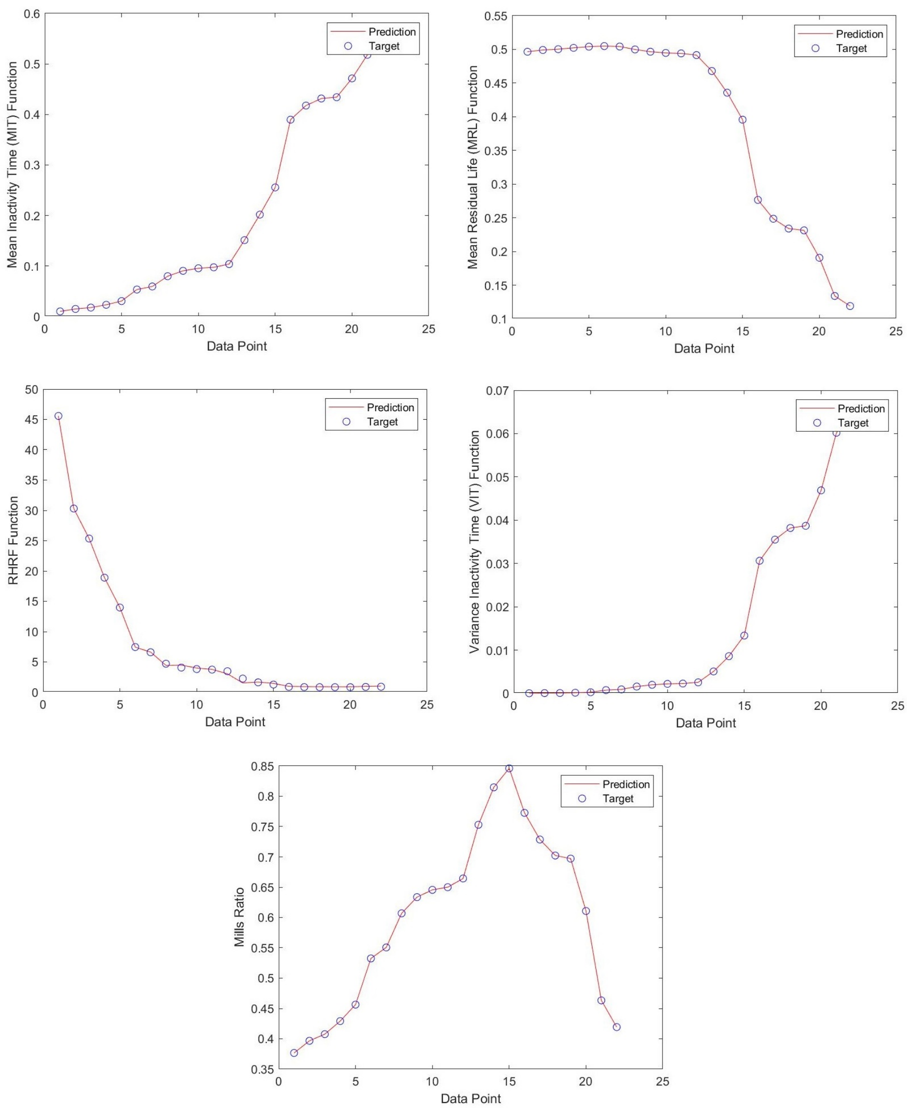

Figure 5: Output values corresponding to data employed in training the stacking model (2).

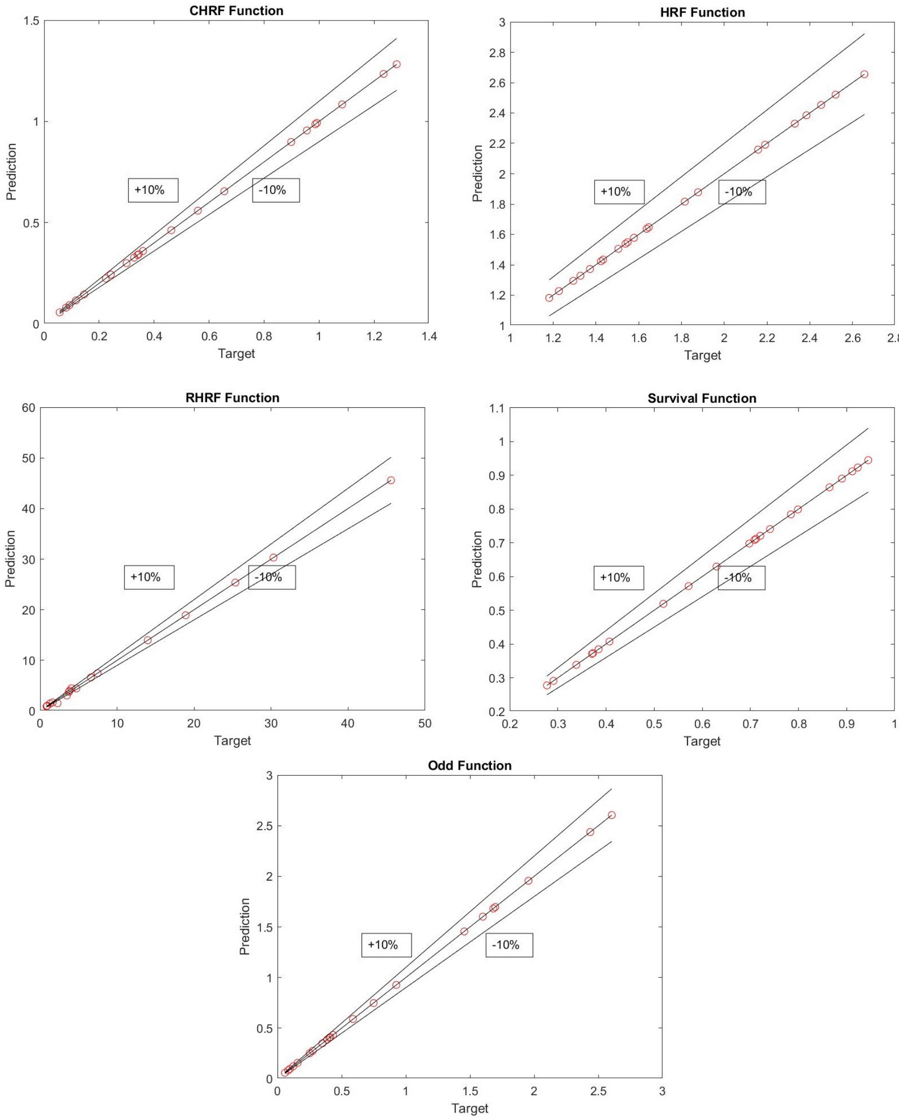

Figure 6: The target and predicted values (1).

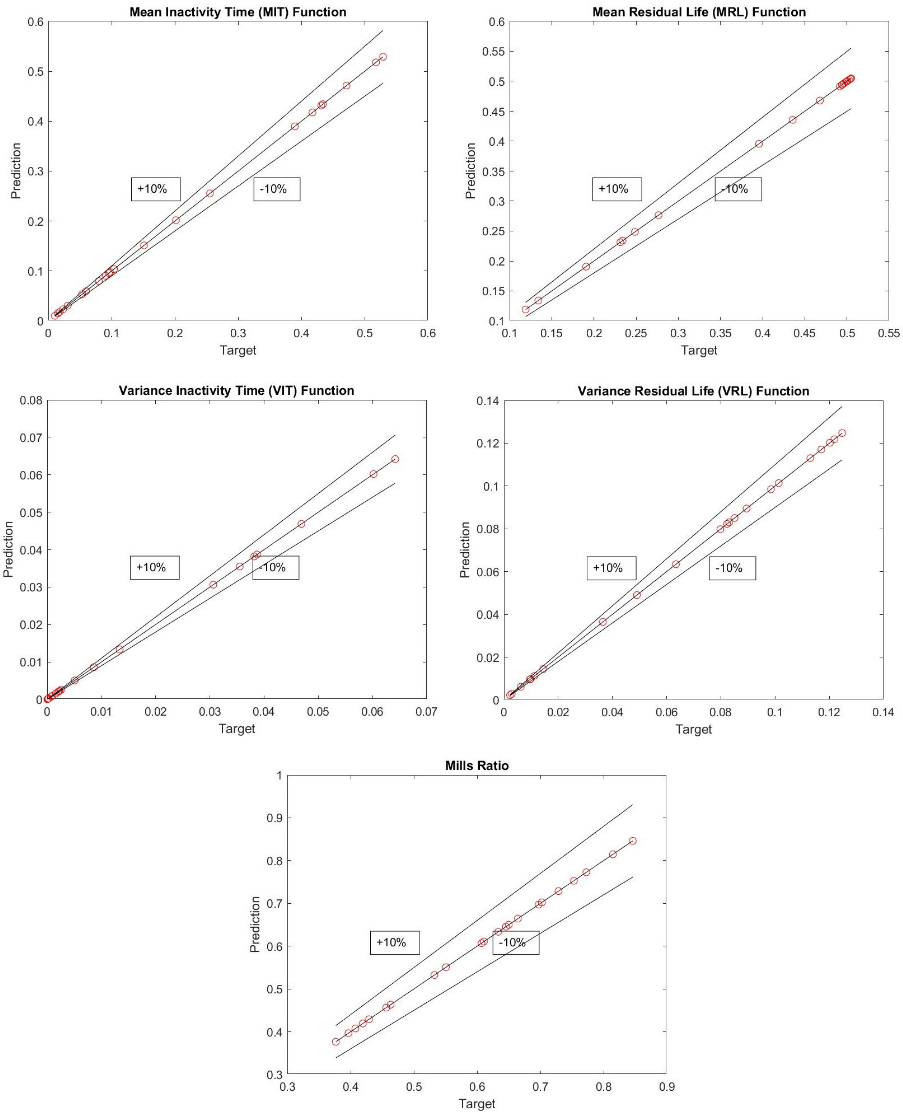

Figure 7: The target and predicted values (2).

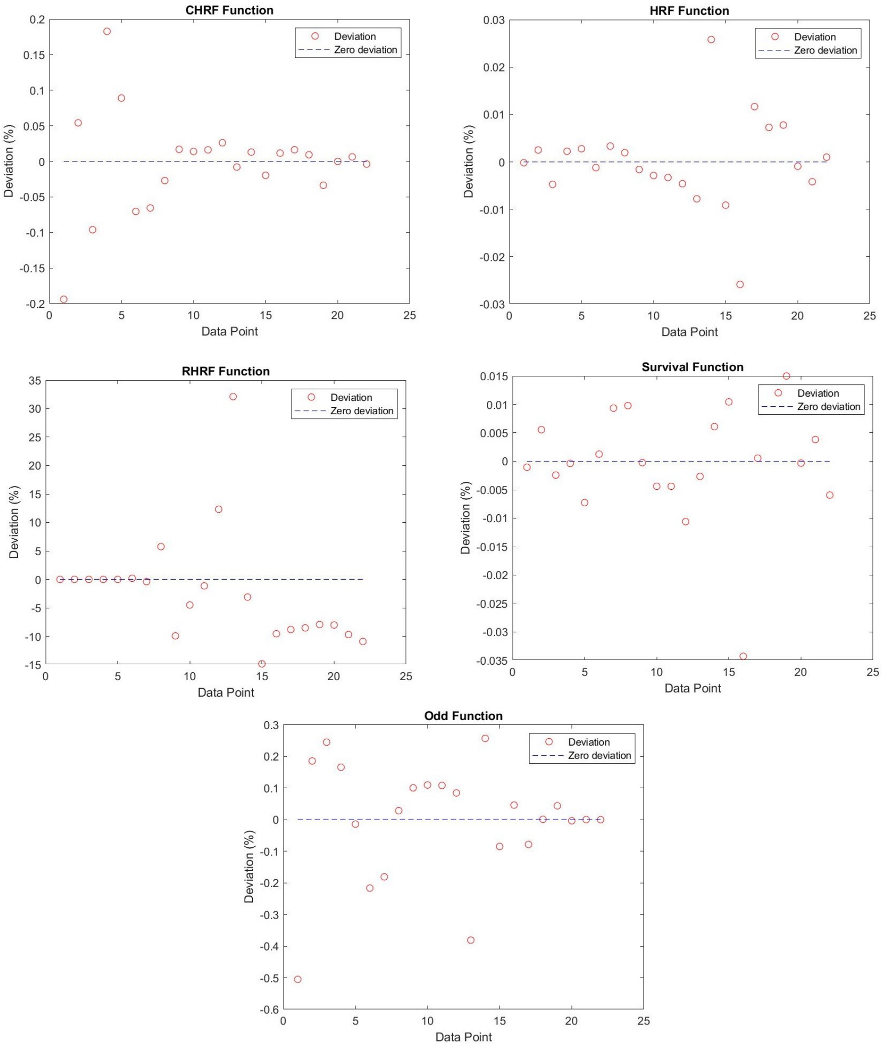

Figure 8: The calculated deviation values for each data point (1).

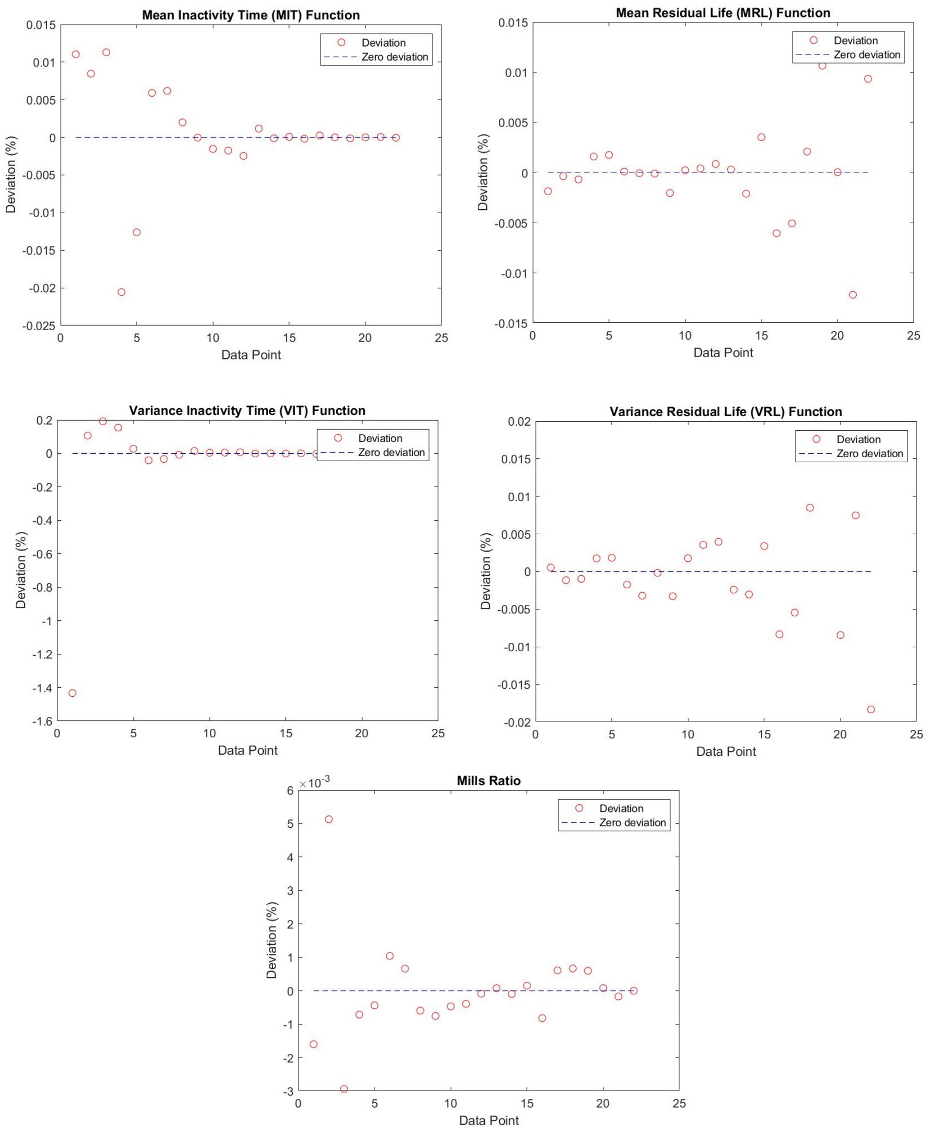

Figure 9: The calculated deviation values for each data point (2).

variances between the outputs produced by the stacking model and the real data. This involves calculating the proportional differences between the expected values and the forecasted values for each data point using the prescribed formula [15]:

$$Deviation(\%) = \left\{ \frac{X_{targ} - X_{pred}}{X_{targ}} \right\} \times 100. \tag{33}$$

The deviation values for individual data points are visually presented in Figures 8 and 9. Upon examination of these values, it is evident that they cluster closely around the zero deviation line. Additionally, the graph illustrating the mean deviation exhibits a comparable trend to that of the zero deviation line. The findings from the analysis of these deviation values confirm that the constructed Stacking can effectively predict ten distinct parameters used to assess the dependability of electrical components, resulting in minimal and satisfactory deviations.

6. Conclusion

This research compared the performance of MLE and a stacking ensemble model in estimating a new power function under Type-II right censoring. Our findings indicate that the stacking model, which integrates predictions from five diverse base models, demonstrates superior accuracy in predicting key reliability measures, such as hazard rate and mean residual life, when dealing with censored data. This suggests that stacking represents a robust and versatile approach for reliability analysis, particularly in handling complex relationships and non-linear patterns within data, potentially outperforming traditional MLE methods. The results highlight the potential of ensemble learning for enhancing reliability assessments in the presence of censored data.

Future research will focus on expanding the scope and applicability of the stacking model. We plan to incorporate a wider range of base models, particularly those designed for handling censored data, within the stacking framework. Furthermore, we aim to refine the model's performance through advanced hyperparameter optimization techniques for both the base models and the meta-learner. To validate its practical relevance, the stacking model will be applied to real-world datasets across diverse domains involving censored data. We will also conduct sensitivity analysis to understand the impact of different censoring schemes and data characteristics on the model's performance. Finally, the stacking approach will be compared with other established ensemble methods, such as bagging and boosting, to assess its relative effectiveness for estimating reliability measures under Type-II right censoring. Through these future directions, we aim to contribute to a more comprehensive understanding of stacking ensembles in survival analysis, ultimately fostering more accurate and reliable predictions in practical applications involving censored data.

Conflicts of Interest. The authors declare that they have no conflicts of interest regarding the publication of this article.

References

[1] M. Z. Iqbal, M. Z. Arshad, G. Özel and O. S. Balogun, A better approach to discuss medical science and engineering data with a modified Lehmann type II model, *F1000Research* **10** (2021) #823.

[2] N. U. Nair, P. G. Sankaran and N. Balakrishnan, *Quantile-Based Reliability Analysis*, Basel: Birkhäuser, 2013.

[3] J. P. Mills, Table of the ratio: area to bounding ordinate for any portion of normal curve, *Biometrika* **18** (1926) 395 – 400, https://doi.org/10.1093/biomet/18.3-4.395.

[4] C. Kundu and A. K. Nanda, Some reliability properties of the inactivity time, *Comm. Statist. Theory Methods* **39** (2010) 899 – 911, https://doi.org/10.1080/03610920902807895.

[5] H. Ishwaran, U. B. Kogalur, E. H. Blackstone and M. S. Lauer, Random survival forests, *Ann. Appl. Stat.* **2** (2008) 841 – 860.

[6] N. Rane, S. P. Choudhary and J. Rane, Ensemble deep learning and machine learning: applications, opportunities, challenges, and future directions, *Stud. Med. Heal. Sci.* **5** (2024) 18 – 41, https://doi.org/10.48185/smhs.v1i2.1225.

[7] J. H. Friedman, Greedy function approximation: a gradient boosting machine, *Ann. Statist.* **29** (2001) 1189 – 1232.

[8] J. L. Katzman, U. Shaham, A. Cloninger, J. Bates, T. Jiang and Y. Kluger, DeepSurv: personalized treatment recommendation via a Cox proportional hazards deep neural network, *BMC Med. Res. Methodol* **18** (2018), https://doi.org/10.1186/s12874-018-0482-1.

[9] A. Wey, J. Connett and K. Rudser, Combining parametric, semi-parametric, and non-parametric survival models with stacked survival models, *Biostatistics* **16** (2015) 537 – 549.

[10] D. G. Kleinbaum and M. Klein, *Survival Analysis: A Self-Learning Text*, Springer New York, NY, 2012.

[11] C. Sun, H. Li, R. E. Mills and Y. Guan, Prognostic model for multiple myeloma progression integrating gene expression and clinical features, *Gigascience* **8** (2019) #giz153, https://doi.org/10.1093/gigascience/giz153.

[12] P. G. Ginestet, E. E. Gabriel and M. C. Sachs, Survival stacking with multiple data types using pseudo-observation-based-AUC loss, *J. Biopharm. Stat.* **32** (2022) 858 – 870, https://doi.org/10.1080/10543406.2022.2041655.

[13] J. Shen, S. Wang, H. Sun, J. Huang, L. Bai, X. Wang and Z. Tang, A novel non-negative bayesian stacking modeling method for cancer survival prediction using high-dimensional omics data, *BMC Med. Res. Methodol.* **24** (2024) #105, https://doi.org/10.1186/s12874-024-02232-3.

[14] J. H. McVittie, D. B. Wolfson, V. Addona and Z. Li, Stacked survival models for residual lifetime data, *BMC Med. Res. Methodol.* **22** (2022) #10, https://doi.org/10.1186/s12874-021-01496-3.

[15] A. B. Çolak, T. N. Sindhu, S. A. Lone, M. T. Akhtar and A. Shafiq, A comparative analysis of maximum likelihood estimation and artificial neural network modeling to assess electrical component reliability, *Qual. Reliab. Eng. Int.* **40** (2024) 91 – 114, https://doi.org/10.1002/qre.3233.

[16] G. Casella and R. L. Berger, *Statistical Inference*, Second edition, Duxbury Press, 2002.

[17] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, John Wiley and Sons, New York, 2003.

[18] B. C. Arnold, N. Balakrishnan and H. N. Nagaraja, *A First Course in Order Statistics*, New York, John Wiley, 1992.

[19] R. B. D'Agostino and M. A. Stephens, *Goodness-of-Fit Techniques*, Marcel Dekker, New York, 1986.

[20] C.-T. Lin, Y.-L. Huang and N. Balakrishnan, A new method for goodness-of-fit testing based on type-II right censored samples, *IEEE Trans. Reliab.* **57** (2008) 633 – 642, https://doi.org/10.1109/TR.2008.2005860.

[21] J. R. Michael and W. R. Schucany, A new approach to testing goodness of fit for censored samples, *Technometrics* **21** (1979) 435 – 441.

[22] M. A. Stephens, EDF statistics for goodness of fit and some comparisons, *J. Am. Stat. Assoc.* **69** (1974) 730 – 737.

[23] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction,* Springer New York, NY, 2009.

[24] D. H. Wolpert, Stacked generalization, *Neural networks* **5** (1992) 241 – 259.

[25] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.

[26] L. V. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*, Pearson Education India, 2006.

[27] K. Levenberg, A method for the solution of certain non-linear problems in least squares, *Quart. Appl. Math.* **2** (1944) 164 – 168.

[28] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math.* **11** (1963) 431 – 441.

[29] J. Nocedal and S. J. Wright, *Numerical optimization*, New York, NY, Springer New York, 1999.

[30] A. J. Smola and B. Schölkopf, A tutorial on support vector regression, *Stat. Comput.* **14** (2004) 199 – 222.

[31] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, 1995.

[32] L. Breiman, Random forests, *Mach. Learn.* **45** (2001) 5 – 32.

[33] H. Rinne, *The Weibull Distribution: A Handbook*, CRC Press: Boca Raton, 2009.

Faranak Goodarzi
Department of Statistics,
Faculty of Mathematical Science,
University of Kashan,
Kashan, I. R. Iran
e-mail: goodarzi_kh@yahoo.com

Mahsa Soheil Shamaee
Department of Computer Science,
Faculty of Mathematical Science,
University of Kashan,
Kashan, I. R. Iran
e-mail: soheilshamaee@kashanu.ac.ir