

## Research Article

# On Predictive Modeling for the $\text{Al}_2\text{O}_3$ Data Using a New Statistical Model and Machine Learning Approach

Mahmoud El-Morshedy <sup>1,2</sup>, Zahra Almaspoor <sup>3</sup>, Gadde Srinivasa Rao <sup>4</sup>,  
Muhammad Ilyas,<sup>5</sup> and Afrah Al-Bossly<sup>1</sup>

<sup>1</sup>Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

<sup>2</sup>Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

<sup>3</sup>Department of Statistics, Yazd University, P.O. Box 89175-741, Yazd, Iran

<sup>4</sup>Department of Mathematics and Statistics, University of Dodoma, P.O. Box: 259, Dodoma, Tanzania

<sup>5</sup>Department of Statistics, University of Malakand, Dir (L), Chakdara, Khyber Pakhtunkhwa, Pakistan

Correspondence should be addressed to Zahra Almaspoor; zahra.ferry21@gmail.com

Received 14 March 2022; Accepted 7 July 2022; Published 30 July 2022

Academic Editor: Ravindran Gobinath

Copyright © 2022 Mahmoud El-Morshedy et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this article, we focused on predictive modeling for real data by means of a new statistical model and applying different machine learning algorithms. The importance of statistical methods in various research fields is modeling the real data and predicting the future behavior of data. For modeling and predicting real-life data, a series of statistical models have been introduced and successfully implemented. This study introduces another novel method, namely, a new generalized exponential-X family for generating new distributions. This method is introduced by using the T-X approach with the exponential model. A special case of the new method, namely, a new generalized exponential Weibull model, is introduced. The applicability of the new method is illustrated by means of a real application related to the alumina ( $\text{Al}_2\text{O}_3$ ) data set. Acceptance sampling plans are developed for this distribution using percentiles when the life test is truncated at the pre-assigned time. The minimum sample size needed to make sure that the required lifetime percentile is determined for a specified customer's risk and producer's risk simultaneously. The operating characteristic value of the sampling plans is also provided. The plan methodology is illustrated using  $\text{Al}_2\text{O}_3$  fracture toughness data. Using the same data set, we implement various machine learning approaches including the support vector machine (SVR), group method of data handling (GMDH), and random forest (RF). To evaluate their forecasting performances, three statistical measures of accuracy, namely, root-mean-square error (RMSE), mean absolute error (MAE), and Akaike information criterion (AIC) are computed.

## 1. Introduction

In the class of traditional/classical distributions, the Weibull model is an interesting model. It has been frequently implemented for dealing and modeling data in different sectors. The DF (cumulative distribution function)  $V(x; \phi)$  of the two-parameter Weibull model is

$$V(x; \phi) = 1 - e^{-\sigma x^\kappa}, \dots \dots x \geq 0, \quad (1)$$

where  $\phi = (\kappa, \sigma)$ ,  $\kappa > 0$ , and  $\sigma > 0$ .

The Weibull model and its different generalized/modified variants have been used by researchers for modeling data in numerous sectors. For example, (i) Ghorbani et al. [1] and Moreau [2] applied it to the medical science phenomena; (ii) Zaindin and Sarhan [3], Lai [4], Almalki and Yuan [5], and Singh [6] used it for reliability engineering applications; and (iii) Ahmad et al. [7] studied its applications in the finance sector.

The probability density function (PDF)  $\nu(x; \phi)$  of the Weibull model is

$$v(x; \phi) = \kappa \sigma x^{\kappa-1} e^{-\sigma x^\kappa}, \dots \dots x > 0, \quad (2)$$

with hazard function (HF)  $h(x; \phi)$  given by

$$h(x; \phi) = \kappa \sigma x^{\kappa-1}, \dots \dots x > 0. \quad (3)$$

From equation (3), we can see that the HF of the Weibull model can be (i) constant for  $\kappa = 1$  (in this case, the Weibull captures the properties of the exponential model), (ii) increasing for  $\kappa > 1$  (in this case, the Weibull captures the properties of the Rayleigh model), and (iii) decreasing for  $\kappa < 1$ .

The Weibull distribution provides impressive results when the failure behavior of the data is either increasing, decreasing, or constant (i.e., monotonic behavior). However, in many cases, the HF of the data behaves nonmonotonically [8]. In such scenarios, the Weibull model is not a popular distribution to use. In the literature, numerous authors have addressed different applications in various fields by developing a flexible version of the Weibull distribution with additional parameters (see Pham and Lai [9]; Nadarajah et al. [10]; and Wais [11]).

Here, we introduce a new method, namely, a new generalized exponential- $X$  (for short “NGExp- $X$ ”) family to obtain modified versions of the existing distributions. The NGExp- $X$  family is introduced by implementing the exponential distribution with PDF  $e^{-t}$  and mixing it with the T- $X$  distribution approach [12].

If  $X$  has the NGExp- $X$  family, then its DF  $F(x; \varphi, \delta, \phi)$  is given by

$$F(x; \varphi, \delta, \phi) = 1 - \left( \frac{\varphi[1 - V(x; \phi)]}{\varphi + V(x; \phi)} \right)^\delta, \dots \dots x \in \mathbb{R}, \quad (4)$$

where  $\varphi > 0, \delta > 0$  and  $V(x; \phi)$  is a baseline DF with parameter vector  $\phi$ .

In order to show that  $F(x; \varphi, \delta, \phi)$  is a compact DF, we have the following two propositions.

**Proposition 1.** *For the expression  $F(x; \varphi, \delta, \phi)$  in equation (1), we must prove that  $\lim_{x \rightarrow -\infty} F(x; \varphi, \delta, \phi) = 0$  and  $\lim_{x \rightarrow \infty} F(x; \varphi, \delta, \phi) = 1$ .*

*Proof.*

$$\begin{aligned} \lim_{x \rightarrow -\infty} F(x; \varphi, \delta, \phi) &= \lim_{x \rightarrow -\infty} \left\{ 1 - \left( \frac{\varphi[1 - V(x; \phi)]}{\varphi + V(x; \phi)} \right)^\delta \right\}, \\ &= 1 - \left( \frac{\varphi[1 - V(-\infty; \phi)]}{\varphi + V(-\infty; \phi)} \right)^\delta, \\ &= 1 - \left( \frac{\varphi[1 - 0]}{\varphi + 0} \right)^\delta, \\ &= 1 - 1, \\ &= 0, \end{aligned} \quad (5)$$

And

$$\begin{aligned} \lim_{x \rightarrow \infty} F(x; \varphi, \delta, \phi) &= \lim_{x \rightarrow \infty} \left\{ 1 - \left( \frac{\varphi[1 - V(x; \phi)]}{\varphi + V(x; \phi)} \right)^\delta \right\}, \\ &= 1 - \left( \frac{\varphi[1 - V(\infty; \phi)]}{\varphi + V(\infty; \phi)} \right)^\delta, \\ &= 1 - \left( \frac{\varphi[1 - 1]}{\varphi + 1} \right)^\delta, \\ &= 1 - 0, \\ &= 1. \end{aligned} \quad (6)$$

□

**Proposition 2.** *The DF  $F(x; \varphi, \delta, \phi)$  is differentiable and RC (right continuous).*

*Proof.* By taking the differentiation of equation (1), we get

$$\frac{d}{dx} F(x; \varphi, \delta, \phi) = f(x; \varphi, \delta, \phi). \quad (7)$$

From the proofs of Propositions 1 and 2, it is obvious that the function  $F(x; \varphi, \delta, \phi)$  provided in equation (4) is a valid DF.

For  $x \in \mathbb{R}, \varphi > 0$ , and  $\delta > 0$ , the PDF  $f(x; \varphi, \delta, \phi)$  and HF  $h(x; \varphi, \delta, \phi) = f(x; \varphi, \delta, \phi) / [1 - F(x; \varphi, \delta, \phi)]$  of the NGExp- $X$  family are given by

$$f(x; \varphi, \delta, \phi) = \frac{\delta \varphi^\delta (\varphi + 1) v(x; \phi) [1 - V(x; \phi)]^{\delta-1}}{[\varphi + V(x; \phi)]^{\delta+1}}, \dots \dots v \in \mathbb{R}, \quad (8)$$

and

$$h(x; \varphi, \delta, \phi) = \frac{\delta (\varphi + 1) v(x; \phi) [1 - V(x; \phi)]^{-1}}{[\varphi + V(x; \phi)]}, \dots \dots v \in \mathbb{R}, \quad (9)$$

respectively. □

## 2. A New Generalized Exponential-Weibull Distribution

Let  $X$  be the proposed NGExp-Weibull distribution with parameters  $\varphi > 0, \sigma > 0, \delta > 0$ , and  $\kappa > 0$ , if its DF  $F(x; \varphi, \sigma, \delta, \kappa)$  and PDF  $f(x; \varphi, \sigma, \delta, \kappa)$  are given by

$$F(x; \varphi, \sigma, \delta, \kappa) = 1 - \left( \frac{\varphi e^{-\sigma x^\kappa}}{\varphi + 1 - e^{-\sigma x^\kappa}} \right)^\delta, \dots \dots v \geq 0, \quad (10)$$

and

$$f(x; \varphi, \sigma, \delta, \kappa) = \frac{\delta \kappa \sigma \varphi^\delta (\varphi + 1) x^{\kappa-1} e^{-\delta \sigma x^\kappa}}{[\varphi + 1 - e^{-\sigma x^\kappa}]^{\delta+1}}, \dots \dots v > 0, \quad (11)$$

respectively.

Corresponding to  $F(x; \varphi, \sigma, \delta, \kappa)$  and  $f(x; \varphi, \sigma, \delta, \kappa)$ , the SF  $SF(x; \varphi, \sigma, \delta, \kappa) = 1 - F(x; \varphi, \sigma, \delta, \kappa)$ , HF

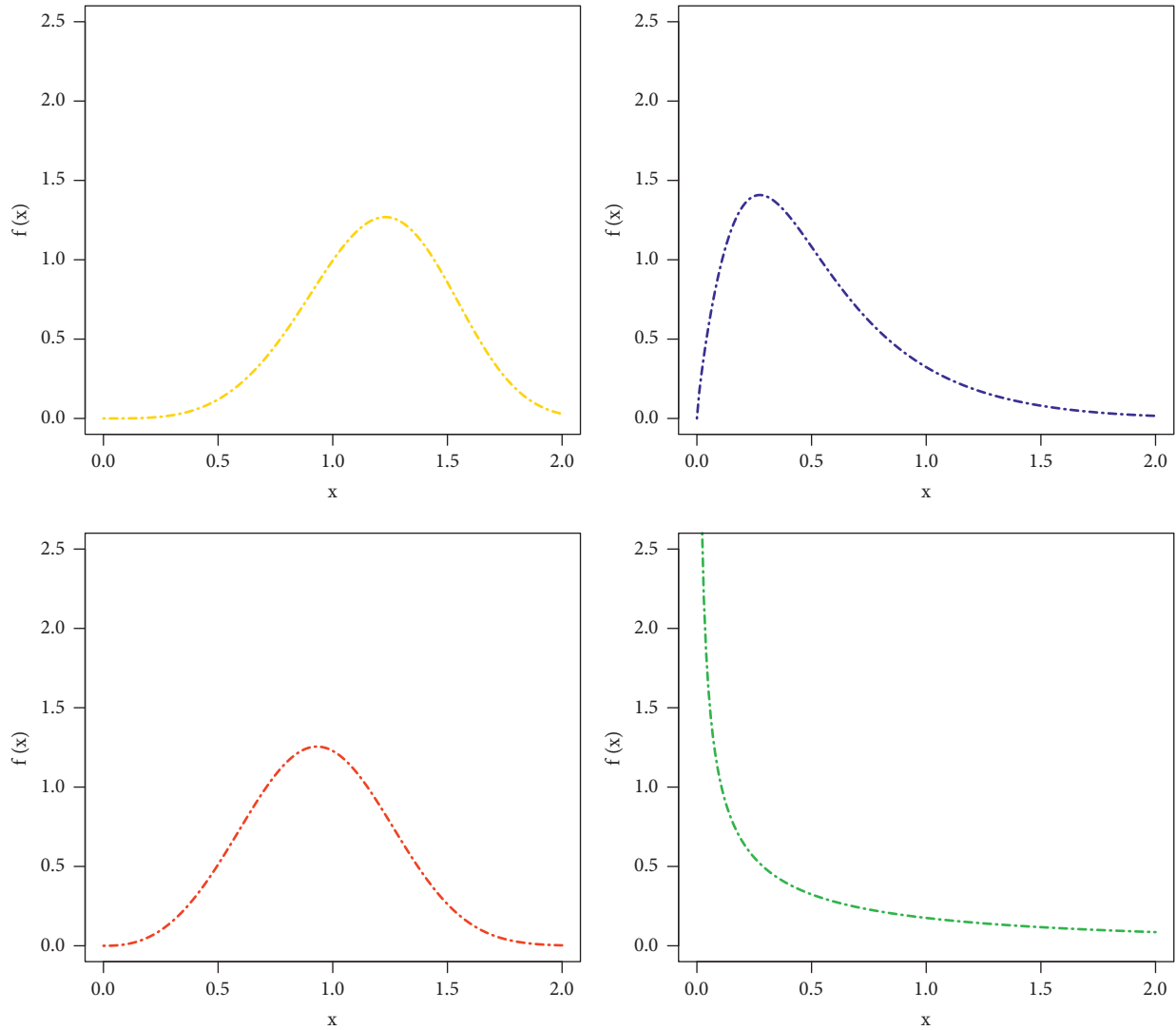


FIGURE 1: A visual behavior of  $f(x; \varphi, \sigma, \delta, \kappa)$  for different values of  $\varphi, \sigma, \delta > 0$ , and  $\kappa > 0$ .

$h(x; \varphi, \sigma, \delta, \kappa) = f(x; \varphi, \sigma, \delta, \kappa)/S(x; \varphi, \sigma, \delta, \kappa)$ , and CHF  $H(x; \varphi, \sigma, \delta, \kappa) = -\log S(x; \varphi, \sigma, \delta, \kappa)$  are given by

$$S(x; \varphi, \sigma, \delta, \kappa) = \left( \frac{\varphi e^{-\sigma x^\kappa}}{\varphi + 1 - e^{-\sigma x^\kappa}} \right)^\delta, \dots x > 0, \tag{12}$$

$$h(x; \varphi, \sigma, \delta, \kappa) = \frac{\delta \kappa \sigma (\varphi + 1) x^{\kappa-1}}{[\varphi + 1 - e^{-\sigma x^\kappa}]}, x > 0,$$

and

$$H(x; \varphi, \sigma, \delta, \kappa) = -\log \left( \frac{\varphi e^{-\sigma x^\kappa}}{\varphi + 1 - e^{-\sigma x^\kappa}} \right)^\delta, \dots x > 0, \tag{13}$$

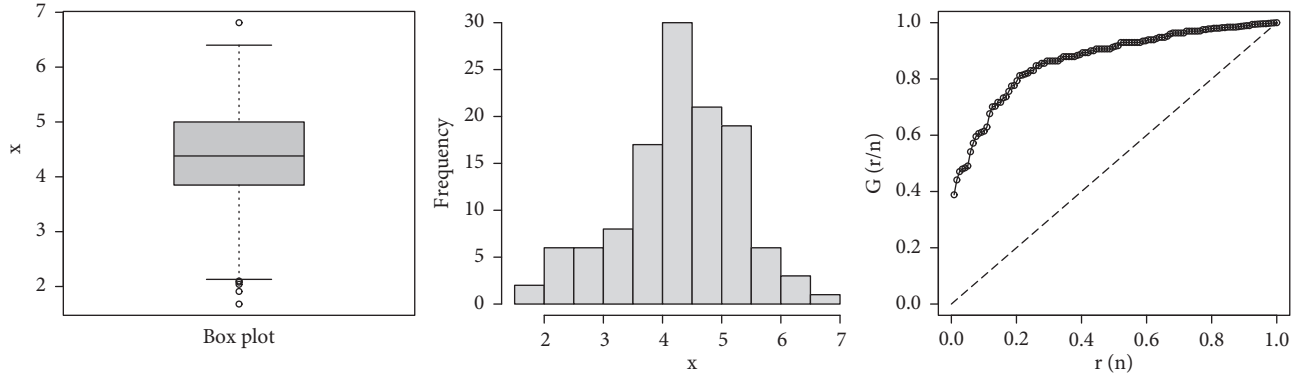
respectively.

Different visual behaviors of  $f(x; \varphi, \sigma, \delta, \kappa)$  for (i)  $\kappa = 4.5, \sigma = 0.1, \delta = 2.5, \varphi = 4.5$  (gold curve), (ii)  $\kappa = 1.8, \sigma = 1, \delta = 1.2, \varphi = 0.5$  (blue curve), (iii)  $\kappa = 3.5, \sigma = 0.4, \delta = 1.8, \varphi = 3.9$  (red curve), and (iv)  $\kappa = 0.5, \sigma = 0.4, \delta = 1.8, \varphi = 3.9$  (green curve) are presented in Figure 1.

From the visual illustration of  $f(x; \varphi, \sigma, \delta, \kappa)$  in Figure 1, we can see that  $f(x; \varphi, \sigma, \delta, \kappa)$  possess different behaviors. For example, it takes (i) the left-skewed form (gold curve), (ii) the right-skewed (blue curve), (iii) the symmetrical shape (red curve), and (iv) the reverse-J shape (green curve).

### 3. Modeling the Al<sub>2</sub>O<sub>3</sub> Data Set

This section offers a practical illustration of the NGExp-Weibull distribution by analyzing data from the engineering sector. We implement the NGExp-Weibull distribution to analyze the Al<sub>2</sub>O<sub>3</sub> (in the units of MPa m1/2) data set (see Nadarajah and Kotz) [13]. The data set is given by 5.5, 5, 4.9, 6.4, 5.1, 5.2, 5.2, 5, 4.7, 4, 4.5, 4.2, 4.1, 4.56, 5.01, 4.7, 3.13, 3.12, 2.68, 2.77, 2.7, 2.36, 4.38, 5.73, 4.35, 6.81, 1.91, 2.66, 2.61, 1.68, 2.04, 2.08, 2.13, 3.8, 3.73, 3.71, 3.28, 3.9, 4, 3.8, 4.1, 3.9, 4.05, 4, 3.95, 4, 4.5, 4.5, 4.2, 4.55, 4.65, 4.1, 4.25, 4.3, 4.5, 4.7, 5.15, 4.3, 4.5, 4.9, 5, 5.35, 5.15, 5.25, 5.8, 5.85, 5.9, 5.75, 6.25, 6.05, 5.9, 3.6, 4.1, 4.5, 5.3, 4.85, 5.3, 5.45, 5.1, 5.3, 5.2, 5.3, 5.25, 4.75, 4.5, 4.2, 4, 4.15, 4.25, 4.3, 3.75, 3.95, 3.51, 4.13, 5.45, 2.1,

FIGURE 2: A visual sketching of the  $\text{Al}_2\text{O}_3$  data.TABLE 1: The numerical values of  $\hat{\kappa}$ ,  $\hat{\sigma}$ ,  $\hat{\varphi}$ ,  $\hat{\delta}$ ,  $\hat{\eta}_1$ ,  $\hat{\eta}_2$ , and their SEs.

Model	$\hat{\kappa}$	$\hat{\sigma}$	$\hat{\varphi}$	$\hat{\delta}$	$\hat{\eta}_1$	$\hat{\eta}_2$
NGExp-Weibull	4.99058 (0.32556)	0.01412 (0.01143)	0.030667 (0.02670)	5.59645 (NaN)	—	—
Exp-Weibull	3.47923 (0.18832)	0.00696 (0.00236)	—	—	1.76624 (0.26376)	—
Kum-Weibull	2.25647 (0.39147)	0.01436 (0.00793)	—	—	2.64266 (0.49563)	12.99694 (14.37373)

4.6, 3.2, 2.5, 4.1, 3.5, 3.2, 3.3, 4.6, 4.3, 4.3, 4.5, 5.5, 4.6, 4.9, 4.3, 3, 3.4, 3.7, 4.4, 4.9, 4.9, and 5.

Corresponding to the  $\text{Al}_2\text{O}_3$  data, the summary measures are as follows: minimum = 1.680, 1<sup>st</sup> quartile = 3.850, median = 4.380, mean = 4.325, 1<sup>st</sup> quartile = 5.000, maximum = 6.810, variance = 1.037332, range = 5.13, standard deviation = 1.018495, skewness = -0.4167136, and kurtosis = 5.13.

The box plot and histogram of the  $\text{Al}_2\text{O}_3$  data are sketched in Figure 2. Additionally, the corresponding curve of the TTT (total test time) is also displayed in Figure 2.

The NGExp-Weibull model is applied to the  $\text{Al}_2\text{O}_3$  data, and its results are compared with (i) the exponentiated Weibull (Exp-Weibull) model with DF given by  $F(x; \kappa, \sigma, \eta_1) = (1 - e^{-\sigma x^\kappa})^{\eta_1}$ ,  $\nu \geq 0$ , and (ii) Kumaraswamy Weibull (Kum-Weibull) model with DF given by  $F(x; \kappa, \sigma, \eta_1, \eta_2) = 1 - [1 - (1 - e^{-\sigma x^\kappa})^{\eta_1}]^{\eta_2}$ ,  $\nu \geq 0$ .

Furthermore, to figure out a suitable model for the  $\text{Al}_2\text{O}_3$  data, three statistical tests with  $p$ -value are considered. These tests are given by (i) AD (Anderson-Darling) test given by  $AD = -k - 1/k \sum_{a=1}^k (2a-1) [\log V(x_a) + \log \{1 - V(x_{k-a+1})\}]$ , (ii) CM (Cramer-von Mises) test expressed by  $CM = 1/12k + \sum_{a=1}^k [2a - 1/2p - M(x_a)]^2$ , and (iii) Kolmogorov-Smirnov (KS) test obtained as  $KS = \sup_{\nu} [V_k(x) - V(x)]$ .

Corresponding to the  $\text{Al}_2\text{O}_3$  data, the values of  $\hat{\kappa}$ ,  $\hat{\sigma}$ ,  $\hat{\varphi}$ ,  $\hat{\delta}$ ,  $\hat{\eta}_1$ , and  $\hat{\eta}_2$  are provided in Table 1. In the same table, the standard errors (SEs) (numerical values in the parentheses) of  $\hat{\kappa}$ ,  $\hat{\sigma}$ ,  $\hat{\varphi}$ ,  $\hat{\delta}$ ,  $\hat{\eta}_1$ , and  $\hat{\eta}_2$  are also presented.

Corresponding to the  $\text{Al}_2\text{O}_3$  data, the  $p$ -value along with the selected tests CM, AD, and KS of the fitted models is provided in Table 2. Based on the results of CM, AD, KS, and  $p$  value in Table 2, it is clear that the NGExp-Weibull model has the smallest values of CM, AD, and KS, and largest  $p$ -value. The values of these statistics show that the NGExp-Weibull model is the best competitor. In addition to the

TABLE 2: The analytical measures of the fitted models for the  $\text{Al}_2\text{O}_3$  data.

Model	CM	AD	KS	$p$ -value
NGExp-Weibull	0.08555	0.52371	0.07204	0.56730
Exp-Weibull	0.16151	1.00382	0.09560	0.22600
Kum-Weibull	0.12816	0.79910	0.08058	0.42220

numerical illustration, a visual display of the performances of the NGExp-Weibull model is provided in Figure 3.

#### 4. A New Acceptance Sampling Plans

In the usual practice of life testing experiment, the test will be terminated by the pre-assigned time  $t_0$  and the number of failures observed. To establish a lower confidence limit on the mean life/percentile lifetime is the aim of the experiments. To protect the consumer's risk, the test has to establish the definite mean life with a certain probability. There are various methods for testing in the literature of acceptance sampling. Epstein [14] was first considered truncated life tests in the exponential distribution. Several authors described about truncated life tests for various distributions, for example, Soble and Tischendorf [15]; Gupta and Groll [16]; Gupta [17]; Baklizi and El-Masri [18]; Tsai and Wu [19]; Balakrishnan et al. [20]; and Kantam et al. [21].

In fact for life distributions, percentiles furnish more information than the mean life does. When the given life distribution is symmetric, the mean life, the median, and the 50th percentile are alike. Thus, developing acceptance sampling plans based on the mean life is a generalized case of developing the acceptance sampling plans based on percentiles of life distribution. Balakrishnan et al. [20] suggested that the acceptance sampling plans could be used for the quantiles and derived the formulae, whereas Lio et al. [22, 23] established for the acceptance sampling plans for any other percentiles of the Birnbaum-Saunders (BS) and

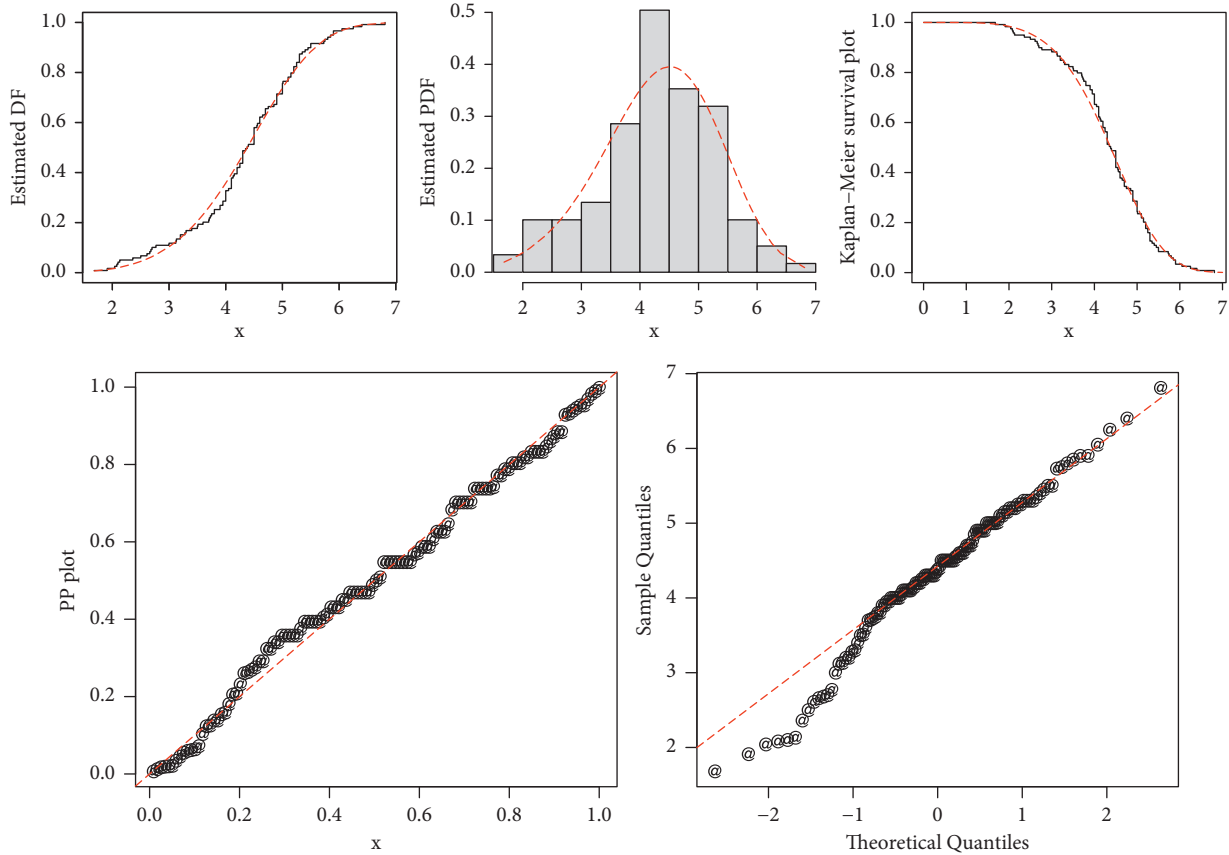


FIGURE 3: Visual performances of the NGExp-Weibull model for the  $Al_2O_3$  data.

Burr type XII models. They have developed the acceptance sampling plans for percentile by replacing the scale parameter by the 100qth percentile.

Rao and Kantam [24] developed acceptance sampling plans from truncated life tests based on the log-logistic and inverse Rayleigh distributions for percentiles. Balamurali et al. [25] developed acceptance sampling plans based on median life for exponentiated half logistic distribution. Rao et al. [26] developed a new acceptance sampling plans based on percentiles for odds exponential log logistic distribution. This persuades to develop the acceptance sampling plans (ASP) based on the percentiles for NGExp-Weibull distribution, and it is a skewed distribution.

This section deals with study of new acceptance sampling plans (ASP) based on new distribution proposed in Section 2. The 100p-th percentile of NGExp-Weibull distribution is given by

$$t_p^k = \frac{1}{\sigma} \log \left( \frac{(1-p)^{1/\delta} (1+\varphi)}{\varphi + (1-p)^{1/\delta}} \right). \quad (14)$$

The expression in equation (14) can also be written as

$$t_p^k = \frac{\tau_p}{\sigma}, \quad (15)$$

where

$$\tau_p^k = -\log \left( \frac{(1-p)^{1/\delta} (1+\varphi)}{\varphi + (1-p)^{1/\delta}} \right). \quad (16)$$

Hence, the quantile  $t_p$  given in equation (16) at a specified value of  $\varphi = \varphi_0$ ,  $\delta = \delta_0$ , and  $k = k_0$ , the quantile  $t_p$  is a function of the scale parameter  $\sigma$ , and at a prespecified value of  $t_p$ , which say  $t_p^0$ , we may obtain the value of  $\sigma$ , which say  $\sigma_0$  as  $\sigma = \tau_p / t_p^k$ . It is remarkable that  $\sigma_0$  could rely on  $\varphi_0$ ,  $\delta_0$ , and  $k_0$  to develop ASP for the NGExp-Weibull distribution determine that  $t_p$  exceeds  $t_p^0$  equivalently  $\sigma$  exceeds  $\sigma_0$ .

At this juncture, the aim of the study was to obtain the minimum sample size needed to acquire percentile life if the life test ended at predestined time  $t_p^0$  and if the amount of nonconformities noticed does not go beyond agreed acceptance number  $c$ . The judgment method is to accept a lot only if the given percentile of the lifetime is recognized with a pre-assigned high probability  $\alpha$  that furnishes security to the consumer. The life testing arrives ended at the time at which  $(c + 1)^{th}$  failure is observed or at quantile of time  $t_p$ , whatsoever is prior.

The chance of accepting lot based on the number of failures is given by

$$L(q) = \sum_{i=0}^c \binom{n}{i} q^i (1-q)^{n-i}, \quad (17)$$

where  $n$  is the sample size,  $c$  is the acceptance number, and  $q$  is the chance of receiving a failure within the life test schedule time  $t_0$ . If the product lifetime follows an NGExp-Weibull distribution, then  $q = F(t_0)$ . Frequently, it would be suitable to define the testing conclusion time  $t_0$  as  $t_0 = \zeta_p^0 t_p^0$  for a constant  $t_p^0$  and the targeted 100p-th lifetime percentile,  $t_p^0$ . Suppose  $t_p$  is the true 100p-th lifetime percentile. Then,  $q$  can be transliterated as:

$$q = 1 - \left( \frac{\varphi \exp\left(-\left(\tau_p(\zeta_p^0)^k / (t_p/t_p^0)^k\right)\right)}{\varphi + 1 - \exp\left(-\left(\tau_p(\zeta_p^0)^k / (t_p/t_p^0)^k\right)\right)} \right)^\delta. \quad (18)$$

In this study, we adopt two points on the OC curve methodology by means of taking into account the both consumer's and producer's risks to obtain the design parameters of the proposed ASP. In this plan, the ratio of percentile lifetime to the lifetime,  $(t_p/t_p^0)$ , is measured the quality intensity of the product. From producer's opinion, the chance of lot acceptance must be at least  $1 - \alpha$  at the acceptable reliability level (ARL), that is,  $q_1$ . Hence, the producer requires that a lot must be accepted at different levels, say  $t_p/t_p^0 = 2, 4, 6, 8, 10$  in equation (18). On the other hand, from consumer's point of view, the chance of rejection of a lot must be at most  $\beta$  at the lot of tolerance reliability level (LTRL),  $q_2$ . Thus, the consumer believes that a lot must be rejected when  $t_p/t_p^0 = 1$ .

Therefore, from equation (18), we get

$$L(q_1) = \sum_{i=0}^c \binom{n}{i} q_1^i (1 - q_1)^{n-i} \geq 1 - \alpha, \quad (19)$$

and

$$L(q_2) = \sum_{i=0}^c \binom{n}{i} q_2^i (1 - q_2)^{n-i} \leq \beta, \quad (20)$$

where  $q_1$  and  $q_2$  are given, respectively, by

$$q_1 = 1 - \left( \frac{\varphi \exp\left(-\left(\tau_p(\zeta_p^0)^k / (t_p/t_p^0)^k\right)\right)}{\varphi + 1 - \exp\left(-\left(\tau_p(\zeta_p^0)^k / (t_p/t_p^0)^k\right)\right)} \right)^\delta, \quad (21)$$

and

$$q_2 = 1 - \left( \frac{\varphi \exp\left(-\tau_p(\zeta_p^0)^k\right)}{\varphi + 1 - \exp\left(-\tau_p(\zeta_p^0)^k\right)} \right)^\delta, \quad (22)$$

where

$$\tau_p = -\log\left(\frac{(1-p)^{1/\delta_0} (1+\varphi_0)}{\varphi_0 + (1-p)^{1/\delta_0}}\right). \quad (23)$$

The design parametric measures for different values of parameters  $\varphi = 2, \delta = 2$ , and  $k = 2$  are constructed. The parameter of the suggested sampling design under the truncated life test at the pre-assigned time  $t_0$  with  $\varphi = 2, \delta = 2$ , and  $k = 2$  is attained for the given producer's

risk  $\alpha = 0.05$  and test termination scheduled time  $t_0$  with  $\zeta_p^0 = 1.0, 1.5, 2.0, 2.5, 3.0$  according to the consumer's confidence levels  $\beta = 0.25, 0.10, 0.05, 0.01$  for 50<sup>th</sup> percentile, and the OC values are determined, and the results are given in Tables 1–3. The design parameter value is tabulated in Tables 3 and 4, for  $\varphi = 2, \delta = 2, k = 2$  and  $\varphi = 0.5, \delta = 1.5, k = 1.5$  with 50th percentiles, while Table 5 displays the design parameter for  $\hat{\varphi} = 0.0307, \hat{\delta} = 5.59645$  and  $\hat{k} = 4.9906$  is the maximum-likelihood estimates from the Al<sub>2</sub>O<sub>3</sub> fracture toughness data set at 50<sup>th</sup> percentile. It is noticed that from Tables 3–5 that the (a) when parametric values are increases the sample size decreases and (b) when the percentile ratio increases the sample size decreases.

**4.1. Description of the Proposed Plan.** To ensure that, the 50th percentile life of products under inspection is at least 1000 hours  $\beta = 0.25$  under the assumption that the producer wants to enforce a single sampling plan at the percentile ratio  $(t_p/t_p^0)$ . Researcher needs to run this life test for 1000 hours. If the lifetime of the product follows NGExp-Weibull distribution observed from the previous data with  $\varphi = 2, \delta = 2$ , and  $k = 2$ . The best plan from Table 3 for stated demands such as  $\beta = 0.25, \varphi = 2, \delta = 2, k = 2, (t_p/t_p^0)$ , and  $\zeta_p^0 = 1.0$  is attained as  $n = 12$  and  $c = 4$  with the probability of acceptance, which is 0.9651. Most of the researchers are studied earlier based on one point on the operating characteristic curve method for assuring mean or percentile life time under various life distributions. The present study is deal with sampling plans based on two-point approach on the operating characteristic curve approach for ensuring percentile lifetime of the products under NGExp-Weibull distribution.

**4.2. Real Data Illustration.** Here, we consider the suggested ASP application for NGExp-Weibull distribution using Al<sub>2</sub>O<sub>3</sub> fracture toughness (in the units of MPa m<sup>1/2</sup>) data set. The goodness of fit for the given model is shown in Table 2, and to emphasize the goodness of fit, we have proved the visuals in Figure 3. The MLEs of the parameters of NGExp-Weibull distribution for the Al<sub>2</sub>O<sub>3</sub> fracture toughness data set are  $\hat{\varphi} = 0.0307, \hat{\delta} = 5.59645, \hat{k} = 4.9906$ , and  $\hat{\nu} = 1.7443$  and the Kolmogorov–Smirnov test we found that the maximum distance between the data and the fitted of the NGExp-Weibull distribution is 0.07204 with  $p$ -value 0.56730. This shows that Al<sub>2</sub>O<sub>3</sub> fracture toughness data set is well fitted for NGExp-Weibull distribution with estimated parameters, and the plan parameters for these estimated parameters are given in Table 3.

Let us fix that the consumer's risk is at 25% when the true 50<sup>th</sup> percentile is fracture toughness 2 units of MPa m<sup>1/2</sup> and the producer's risk is 5% when the true 50<sup>th</sup> percentile is fracture toughness 4 units of MPa m<sup>1/2</sup>. For  $\hat{\varphi} = 0.0307, \hat{\delta} = 5.59645, \hat{k} = 4.9906$ , the consumer's risk is 25%,  $\zeta_p^0 = 1$ , and  $(t_p/t_p^0) = 2$ . The minimum sample size  $n = 5$  and acceptance number  $c = 1$  are given from Table 3. Thus, the design can be implemented as follows: we select a sample of 5 fracture toughness units, and we will accept the lot when no more than 1 fracture toughness 2 units. Hence, by applying the proposed sampling plan, the fracture toughness lot has

TABLE 3: The plan parameters and OC values for NGExp-Weibull distribution for  $\varphi = 2, \delta = 2$ , and  $k = 2$ .

$\beta$	$(t_p/t_p^0)$	$\zeta_p^0 = 1.0$			$\zeta_p^0 = 1.5$			$\zeta_p^0 = 2$			$\zeta_p^0 = 2.5$			$\zeta_p^0 = 3$		
		$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$
0.25	2	4	12	0.9651	4	7	0.9573	6	8	0.9648	7	8	0.9673	11	12	0.9549
	4	1	5	0.9819	1	3	0.9736	1	2	0.9728	2	3	0.9855	2	3	0.9645
	6	1	5	0.9961	0	1	0.9555	1	2	0.9940	1	2	0.9860	1	2	0.9728
	8	0	3	0.9663	0	1	0.9747	2	1	0.9555	1	2	0.9953	1	2	0.9906
	10	0	3	0.9783	0	1	0.9837	2	2	0.9712	0	1	0.9555	1	2	0.9960
0.10	2	5	17	0.9518	6	11	0.9642	6	11	0.9673	10	13	0.9562	11	12	0.9549
	4	1	7	0.9641	1	4	0.9506	2	4	0.9843	2	2	0.9855	2	3	0.9645
	6	1	7	0.9921	1	4	0.9888	1	3	0.9829	1	2	0.9860	1	2	0.9728
	8	0	4	0.9554	1	4	0.9963	0	2	0.9555	1	2	0.9953	1	2	0.9906
	10	0	4	0.9712	0	2	0.9677	0	2	0.9712	0	1	0.9555	1	2	0.9960
0.05	2	6	21	0.9544	8	15	0.9719	9	13	0.9539	10	12	0.9562	17	19	0.9511
	4	1	8	0.9536	1	4	0.9506	2	4	0.9843	2	4	0.9527	2	3	0.9645
	6	1	8	0.9896	1	4	0.9888	1	3	0.9829	1	2	0.9860	1	2	0.9728
	8	0	8	0.9966	1	4	0.9963	1	3	0.9942	1	2	0.9953	1	2	0.9906
	10	0	5	0.9641	0	3	0.9519	1	3	0.9976	0	1	0.9555	1	2	0.9960
0.01	2	8	30	0.9523	10	20	0.9660	11	16	0.9616	13	16	0.9532	17	19	0.9511
	4	2	14	0.9778	2	7	0.9763	2	5	0.9655	2	4	0.9527	3	5	0.9570
	6	1	11	0.9803	1	6	0.9736	1	3	0.9675	1	3	0.9614	2	4	0.9843
	8	1	11	0.9934	1	6	0.9910	1	3	0.9888	1	3	0.9865	1	3	0.9736
	10	0	7	0.9501	1	6	0.9962	1	3	0.9952	1	3	0.9942	1	3	0.9885

TABLE 4: The plan parameters and OC values for NGExp-Weibull distribution for  $\varphi = 0.5, \delta = 1.5$ , and  $k = 1.5$ .

$\beta$	$(t_p/t_p^0)$	$\zeta_p^0 = 1.0$			$\zeta_p^0 = 1.5$			$\zeta_p^0 = 2$			$\zeta_p^0 = 2.5$			$\zeta_p^0 = 3$		
		$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$
0.25	2	7	18	0.9589	9	16	0.9618	9	13	0.9539	11	14	0.9582	5	18	0.9529
	4	2	7	0.9784	2	5	0.9667	2	4	0.9571	3	5	0.9664	4	6	0.9690
	6	1	5	0.9751	1	3	0.9753	2	4	0.9904	1	2	0.9651	2	3	0.9870
	8	1	5	0.9888	1	3	0.9888	1	3	0.9753	1	2	0.9837	1	2	0.9735
	10	0	2	0.9507	3	3	0.9941	1	3	0.9866	1	2	0.9912	1	2	0.9855
0.10	2	10	28	0.9540	13	24	0.9685	14	21	0.9608	18	24	0.9573	19	23	0.9593
	4	2	9	0.9551	3	8	0.9717	3	6	0.9697	4	7	0.9681	4	6	0.9690
	6	1	7	0.9512	1	4	0.9536	2	5	0.9784	2	4	0.9776	2	4	0.9571
	8	1	7	0.9775	1	4	0.9786	1	3	0.9753	1	3	0.9554	2	4	0.9849
	10	1	7	0.9879	1	4	0.9885	1	3	0.9866	1	3	0.9753	1	3	0.9599
0.05	2	12	35	0.9505	15	29	0.9591	18	28	0.9564	—	—	—	—	—	—
	4	3	13	0.9724	3	9	0.9555	4	9	0.9617	4	7	0.9681	5	8	0.9635
	6	2	11	0.9825	2	7	0.9784	2	6	0.9612	2	5	0.9517	2	4	0.9571
	8	1	8	0.9707	1	5	0.9657	1	4	0.9536	1	3	0.9554	2	3	0.9849
	10	1	8	0.9842	1	5	0.9813	1	4	0.9744	1	3	0.9753	1	3	0.9599
0.01	2	17	52	0.9533	—	—	—	—	—	—	—	—	—	—	—	—
	4	4	19	0.9725	4	13	0.9532	5	12	0.9586	6	11	0.9753	6	10	0.9610
	6	2	14	0.9895	2	9	0.9551	3	8	0.9836	3	7	0.9737	3	6	0.9697
	8	2	14	0.9886	2	7	0.9847	2	7	0.9784	2	6	0.9693	2	5	0.9667
	10	1	11	0.9704	1	7	0.9631	1	5	0.9593	1	4	0.9536	2	5	0.9855

Note. (—) indicates that the parameters cannot be found to satisfy conditions.

been rejected because there is more than one unit before the termination fracture toughness 4 units.

### 5. Data and Modeling Procedures

The  $Al_2O_3$  forecasting is very crucial and attributes remarkable growth to the country's economy. Knowing its future trajectory is beneficial for the government and policymakers. Therefore, this section provides the multistep ahead forecast of  $Al_2O_3$  by applying different machine

learning algorithms, namely, support vector machine (SVM), group method of data handling (GMDH) neural network, and random forest (RF).

5.1. Support Vector Regression. The SVR method was first introduced by Cortes and Vapnik [27], and to date, it is frequently used for regression and classification problems. The SVR is based on statistical theory and structured risk minimization principle, due to which it circumvents the

TABLE 5: The plan parameters and OC values for NGExp-Weibull distribution for  $\varphi = 0.0307$ ,  $\delta = 5.59645$ , and  $k = 4.9906$ .

$\beta$	$(t_p/t_p^0)$	$c$	$\zeta_p^0 = 1.0$			$\zeta_p^0 = 1.5$			$\zeta_p^0 = 2$			$\zeta_p^0 = 2.5$			$\zeta_p^0 = 3$			
			$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$	$L(q_1)$	$c$	$n$
0.25	2	1	5	0.9950	1	2	0.9749	4	5	0.9688	18	19	0.9546	—	—	—	—	—
	4	0	3	0.9978	1	2	1.0000	0	1	0.9771	1	2	0.9954	1	2	0.9749	—	—
	6	0	3	0.9997	1	2	1.0000	0	1	0.9969	0	1	0.9907	0	1	0.9771	—	—
	8	0	3	0.9999	1	2	1.0000	0	1	0.9993	0	1	0.9978	0	1	0.9945	—	—
	10	0	3	1.0000	1	2	1.0000	0	1	0.9998	0	1	0.9993	0	1	0.9982	—	—
0.10	2	1	7	0.9898	1	2	0.9749	4	5	0.9688	18	19	0.9546	—	—	—	—	—
	4	0	4	0.9971	0	1	0.9945	0	1	0.9771	1	2	0.9954	1	2	0.9749	—	—
	6	0	4	0.9996	0	1	0.9993	0	1	0.9969	0	1	0.9907	0	1	0.9771	—	—
	8	0	4	0.9999	0	1	0.9998	0	1	0.9993	0	1	0.9978	0	1	0.9945	—	—
	10	0	4	1.0000	0	1	0.9999	0	1	0.9998	0	1	0.9993	0	1	0.9982	—	—
0.05	2	1	8	0.9867	1	2	0.9749	4	5	0.9688	18	19	0.9546	—	—	—	—	—
	4	0	5	0.9964	0	1	0.9945	0	1	0.9771	1	2	0.9954	1	2	0.9749	—	—
	6	0	5	0.9995	0	1	0.9993	0	1	0.9969	0	1	0.9907	0	1	0.9771	—	—
	8	0	5	0.9999	0	1	0.9998	0	1	0.9993	0	1	0.9978	0	1	0.9945	—	—
	10	0	5	1.0000	0	1	0.9999	0	1	0.9998	0	1	0.9993	0	1	0.9982	—	—
0.01	2	1	11	0.9750	2	4	0.9860	4	5	0.9688	—	—	—	—	—	—	—	—
	4	0	7	0.9949	0	2	0.9890	0	1	0.9771	1	2	0.9954	1	2	0.9749	—	—
	6	0	7	0.9993	0	2	0.9985	0	1	0.9969	0	1	0.9907	0	1	0.9771	—	—
	8	0	7	0.9998	0	2	0.9997	0	1	0.9993	0	1	0.9978	0	1	0.9945	—	—
	10	0	7	0.9999	0	2	0.9999	0	1	0.9998	0	1	0.9993	0	1	0.9982	—	—

problem of overfitting, and produces an accurate forecast. In practice, it can precisely approximate the linear and nonlinear problems in the real world [28]. The performance of SVR is based on the kernel function to be utilized. Kernel functions are utilized to carry out operations in the input space instead of higher-dimensional space. Various kernel functions are appeared in past studies, namely, linear, polynomial, splines, and sigmoid radial basis functions [29]. RBF has achieved great attention due to its outstanding performance in capturing nonlinear association [30, 31]. The SVR helps to disclose the margin of error, which is acceptable in a model [32, 33]. Let we have a set of training data, having output and input variables. Mathematically, the SVR seeks to find a function that fitted the data more appropriately. In other words, it minimizes the error between output variable and predicted values in the following way:

$$f(\delta) = (\Omega, (\delta)) + \beta, \quad (24)$$

where  $\Omega$  is a weight parameter,  $(\delta)$  represents the nonlinear function, and  $\beta$  is a constant parameter. In equation (25), the loss functions demonstrate the  $\varepsilon$ -insensitive model, where the loss tends to 0, if the difference between actual values and predicted values is less than  $c$ , and is illustrated by the Vapnik's  $\varepsilon$ -insensitive loss function

$$\mathfrak{R}_\varepsilon(f(\delta), P) = \begin{cases} |f(\delta) - P| - \varepsilon & \text{if } |f(\delta) - P| \geq \varepsilon \\ 0 & \text{otherwise} \end{cases}. \quad (25)$$

The problem of SVR is then constructed as the following optimization problem.

$$\min_{\Omega, \beta, \mathfrak{N}_j, \mathfrak{N}_j^*} \frac{1}{2} \Omega + D \sum_{j=1}^h (\mathfrak{N}_j + \mathfrak{N}_j^*), \quad (26)$$

where  $\mathfrak{N}_j^*$  and  $\mathfrak{N}_j$  represent the slack variables describing the lower and upper training errors subject to the error tolerance  $\varepsilon$ , and  $D$  is a positive constant that discovers the extent of penalized loss when a training error occurs. In our case, we set the values of  $D$  and  $E$  are 1 and 0.01, respectively.

$$\Omega = \sum_{m=1}^h (\theta_m - \theta_m^*) Q(\delta_m, \delta). \quad (27)$$

The set of Lagrange multipliers including  $\theta_m$  and  $\theta_m^*$  is utilized for optimization problem solution. Thus, the approximate function is illustrated as follows:

$$P_t = \sum_{m=1}^h (\theta_m - \theta_m^*) Q(\delta_m, \delta) + \omega, \quad (28)$$

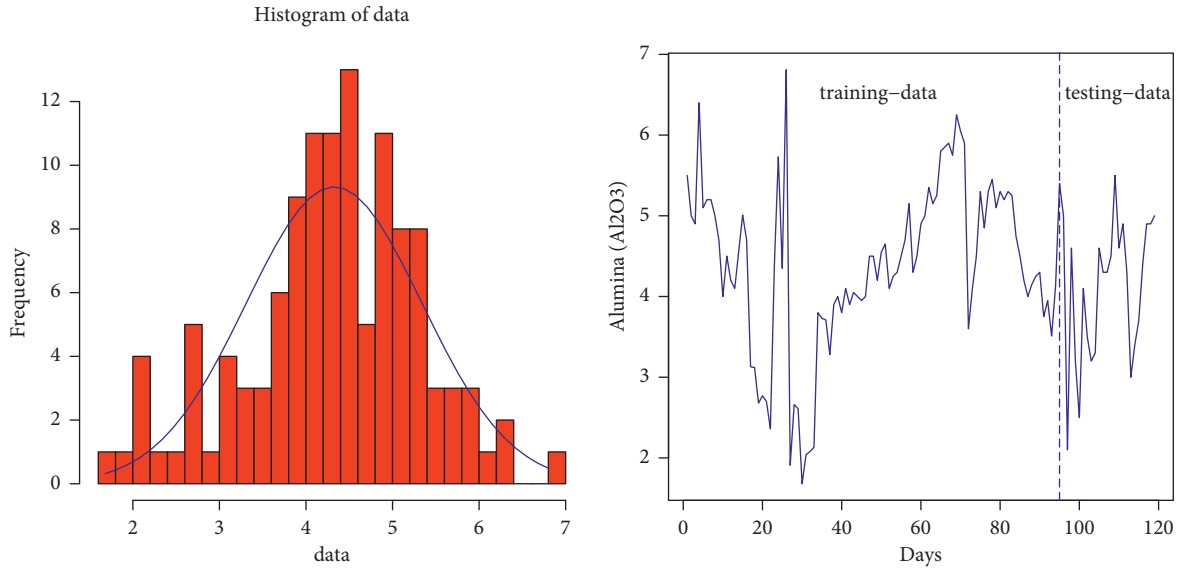
where  $\delta_m$  represents the support vector,  $h$  represents the size of support vector,  $Q(\delta_m, \delta)$  represents the kernel function, and  $\omega$  represents the threshold value. Herein, the radial basis function (RBF) with a parameter  $\pi^2$  is expressed as follows:

$$Q(\delta_m, \delta_n) = \exp\left(-\frac{\|\delta_m - \delta_n\|^2}{2\pi^2}\right), \quad (29)$$

where  $\|\delta_m - \delta_n\|^2$  indicates the Euclidean distance between the two predictors in squared form, and  $\pi^2$  indicates the width of RBF [34]. Hence, in this study, we focus on the RBF kernel function for SVR.

**5.2. Random Forest.** Breiman [35] developed a nonparametric approach known as random forest (RF). The development of the RF approach is based on decision tree algorithms; however, this is the modified form of classification and regression trees (CART). The RF is, therefore, to



FIGURE 4: Histogram and line chart of  $\text{Al}_2\text{O}_3$  data.

be utilized for handling both types of issues, classification, and regression. In general, the RF is a part of the supervised learning class and its output is based on decision trees' forest. The forecast for the RF approach is basically obtained by taking the mean of various trees. A large number of trees in the forest can help to improve the forecasting accuracy and prevent the problem of overfitting. The RF algorithm utilizes the popular procedure of bagging, also known as bootstrap aggregation, in order to train the tree learners. The random sample is taken repeatedly ( $M$  time) with replacement of the training set and estimates the trees to each repeated sample [32]. To estimate the RF, we utilize three hyperparameters including number of trees, number of nodes, and sample repetition. The number of trees and nodes is used as 500 and 3, respectively.

**5.3. Group Method of Data Handling (GMDH).** The GMDH neural network (GMDH-NN) was initially introduced by Ivakhnenko [36] in analyzing complex systems, which incorporates an output and a set of inputs. The core aim of GMDH-NN is simply to formulate a function in a feed-forward network based on second degree transfer function. The productive input variables, the set of neurons and layers within a hidden variable, and the optimal model framework are established automatically in the GMDH algorithm [37]. In our study, we select these parameters through error and trial approach, followed by Peng et al. [38]. The mapping amid target and input variables is carried out via GMDH-NN, and a nonlinear function is so-called Volterra series, given in equation (30) as follows:

$$\hat{P} = \theta_0 + \sum_{m=1}^h \theta_m \gamma_m + \sum_{m=1}^h \sum_{n=1}^h \theta_{m,n} \gamma_m \gamma_n + \sum_{m=1}^h \sum_{n=1}^h \sum_{s=1}^h \theta_{m,n,s} \gamma_m \gamma_n \gamma_s + \dots \quad (30)$$

For two variables, the Volterra series can be described in terms of second-degree polynomial in equation (31) as

$$M(y_m, y_n) = \theta_0 + \theta_1 y_1 + \theta_2 y_2 + \theta_3 y_1^2 + \theta_4 y_2^2 + \theta_5 y_1 y_2, \quad (31)$$

where  $M(y_m, y_n)$  denotes an output of the model,  $y_i$  ( $i = 1, 2$ ) denotes the input variables, and the corresponding weight is shown by  $\theta$ .

The network neurons are recursively connected to each other through the partial quadratic equation, which reveals the nexus while estimating the unknown parameters using the training data. The key aim of GMDH is to find out the unknown coefficients provided in equation (31), as they demonstrate the least difference amid the actual data and forecasted values. The unknown parameters are computed utilizing the regression tool [39, 40]. Thus, under the rule of principle of least square error, the parameters of each quadratic equation are optimized in the following way:

$$E = \frac{\sum_{j=1}^N (P_j - W_j)^2}{N}. \quad (32)$$

We seek to reduce the squared difference ( $E$ ) between predicted valued and actual values; in order to achieve an accurate forecast, we have

$$P = B\theta, \quad (33)$$

where  $P$  is the response variable,  $\theta$  is the vector of unknown parameters to be estimated from the data at hand, and  $B$  is computed as follows:

$$B = \begin{bmatrix} 1 & \theta_{1p} & \theta_{1q} & \theta_{1p}^2 & \theta_{12} & \theta_{1p}\theta_{1q} \\ 1 & \theta_{2p} & \theta_{2q} & \theta_{2p}^2 & \theta_{12} & \theta_{2p}\theta_{2q} \\ 1 & \theta_{3p} & \theta_{3q} & \theta_{3p}^2 & \theta_{1q}^2 & \theta_{3p}\theta_{3q} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \theta_{Np} & \theta_{Nq} & \theta_{Np}^2 & \theta_{Nq}^2 & \theta_{Np}\theta_{Nq} \end{bmatrix}. \quad (34)$$

TABLE 6: The values of the error metrics.

Criteria	SVR	RF	GMDH
RMSE	0.070	0.090	0.233
MAE	0.031	0.064	0.159
AIC	-243.2	-128.8	-114.2

**5.4. Out-of-Sample  $Al_2O_3$  Forecasting.** The  $Al_2O_3$  data are utilized in this work to assess the predictive capability of different ML algorithms. The data consist of 119 observations. For estimation and prediction, the data set is decomposed into two parts: for model estimation, we use the data points from 1 to 95, and 96 to 119 for evaluating the models' multistep ahead out-of-sample predictive accuracy utilizing the expanding window methods. This study adopts popular three statistical measures, namely, root-mean-square error (RMSE), mean absolute error (MAE), and Akaike information criterion (AIC) to evaluate the forecasting accuracy of ML algorithms. The smaller the values of RMSE, MAE, and AIC, the better the forecast. Mathematically, they can be, respectively, illustrated as

$$MAE = \text{mean}(|\hat{P}_t - P_t|), RMSE = \sqrt{\text{mean}(\hat{P}_t - P_t)^2}, \quad (35)$$

$$AIC = n \log\left(\frac{sse}{n}\right) + n * k,$$

where  $n$  shows the number of observations,  $sse$  indicates the sum of squared error,  $k$  indicates the number of parameters,  $P_t$  represents actual value of  $Al_2O_3$ , and  $\hat{P}_t$  represents the predicted values of  $Al_2O_3$ .

Figure 4 indicates the  $Al_2O_3$  data trend, which does not follow any particular pattern and is highly uncertain. The  $Al_2O_3$  data are plotted in Figure 4 where the vertical blue dotted line separates the estimation (80 percent) and out-of-sample forecasting (20 percent) parts. The accuracy measures for the  $Al_2O_3$  data are reported in Table 6.

The accuracy measures for the  $Al_2O_3$  data are reported in Table 6. The RMSE and MAE values for ML algorithms include GMDH, SVR, and RF. It can be observed that SVR beats the rival counterparts. The RMSE and MAE values associated with SVR are 0.07 and 0.031, the RF resulted in 0.09 and 0.064, and the GMDH resulted in 0.233 and 0.159, respectively. In addition, according to AIC, the best model is SVR among the all.

## 6. Concluding Remarks

In this piece of study, a new generalized exponential-X family of distributions is well thought-out as a real-life data model. The attempts in this paper lead to another approach to developing a new statistical model. Employing the proposed model, a new modification of the Weibull model called a new generalized exponential Weibull model is studied. The effectiveness of the new generalized exponential Weibull model is shown by considering the fracture toughness of the  $Al_2O_3$  data set. A statistical product control application of the developed model is also studied by developing the new single acceptance sampling plan based on

the NGExp-Weibull distribution. The plan parameters are determined such that both consumer's risk and producer's risk satisfy simultaneously. Some tables are given for industrial application purposes. The developed plan is exemplified by the  $Al_2O_3$  fracture toughness data set, which is well fitted for the proposed NGExp-Weibull. Furthermore, using the same data set, we implemented several ML algorithms including SVR, GMDH, and RF. The out-of-sample forecast accuracy was assessed using three statistical measures of accuracy, namely, the RMSE, MAE, and AIC. After the analysis, we found that the SVR produces a more efficient forecast than the competitor counterparts. The findings clearly reveal that the predictive power of the SVR method is superior in a class of ML algorithms while predicting the  $Al_2O_3$  data.

## Data Availability

The data set is given in the main body of the paper.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## References

- [1] N. Ghorbani, J. Yazdani-Charati, K. Anvari, and N. Ghorbani, "Application of the Weibull accelerated failure time model in the determination of disease-free survival rate of patients with breast cancer," *Iranian Journal of Health Sciences*, vol. 4, no. 2, pp. 11-18, 2016.
- [2] V. H. Moreau, "Using the Weibull distribution to model COVID-19 epidemic data," *Model Assisted Statistics and Applications*, vol. 16, no. 1, pp. 5-14, 2021.
- [3] M. Zaindin and A. M. Sarhan, "Modified Weibull distribution," *Applied Sciences*, vol. 11, pp. 123-136, 2009.
- [4] C. D. Lai, "Constructions and applications of lifetime distributions," *Applied Stochastic Models in Business and Industry*, vol. 29, no. 2, pp. 127-140, 2013.
- [5] S. J. Almalki and J. Yuan, "A new modified Weibull distribution," *Reliability Engineering & System Safety*, vol. 111, pp. 164-170, 2013.
- [6] B. Singh, "An additive Perks-Weibull model with Bathtub-shaped hazard rate function," *Communications in Mathematics and Statistics*, vol. 4, no. 4, pp. 473-493, 2016.
- [7] Z. Ahmad, E. Mahmoudi, and G. Hamedani, "A class of claim distributions: properties, characterizations and applications to insurance claim data," *Communications in Statistics - Theory and Methods*, vol. 51, no. 7, pp. 2183-2208, 2020.
- [8] C. D. Lai and M. Xie, *Stochastic Ageing and Dependence for Reliability*, p. 420, Springer, New York, 2006.
- [9] H. Pham and C. D. Lai, "On recent generalizations of the Weibull distribution," *IEEE Transactions on Reliability*, vol. 56, no. 3, pp. 454-458, 2007.
- [10] S. Nadarajah, G. M. Cordeiro, and E. M. M. Ortega, "The exponentiated Weibull distribution: a survey," *Statistical Papers*, vol. 54, no. 3, pp. 839-877, 2013.
- [11] P. Wais, "A review of Weibull functions in wind sector," *Renewable and Sustainable Energy Reviews*, vol. 70, pp. 1099-1107, 2017.
- [12] A. Alzaatreh, C. Lee, and F. Famoye, "A new method for generating families of continuous distributions," *Metron*, vol. 71, no. 1, pp. 63-79, 2013.

- [13] S. Nadarajah and S. Kotz, "Strength modeling using Weibull distributions," *Journal of Mechanical Science and Technology*, vol. 22, no. 7, pp. 1247–1254, 2008.
- [14] B. Epstein, "Truncated life tests in the exponential case," *The Annals of Mathematical Statistics*, vol. 25, no. 3, pp. 555–564, 1954.
- [15] M. Soble and J. Tischendorf, "Acceptance sampling with new life test objectives," in *Proceedings of the 5th National Symposium on Reliability and Quality Control*, pp. 108–118, Philadelphia, Pennsylvania, 1959.
- [16] S. S. Gupta and S. S. Gupta, "Gamma distribution in acceptance sampling based on life tests," *Journal of the American Statistical Association*, vol. 56, no. 296, pp. 942–970, 1961.
- [17] S. S. Gupta, "Life test sampling plans for normal and log-normal distributions," *Technometrics*, vol. 4, no. 2, pp. 151–175, 1962.
- [18] A. Baklizi and A. E. Q. El Masri, "Acceptance sampling based on truncated life tests in the Birnbaum Saunders model," *Risk Analysis*, vol. 24, no. 6, pp. 1453–1457, 2004.
- [19] T. R. Tsai and S. J. Wu, "Acceptance sampling based on truncated life tests for Generalized Rayleigh distribution," *Journal of Applied Statistics*, vol. 33, no. 6, pp. 595–600, 2006.
- [20] N. Balakrishnan, V. Leiva, and J. Lopez, "Acceptance sampling plans from truncated life tests based on the generalized birnbaum-Saunders distribution," *Communications in Statistics - Simulation and Computation*, vol. 36, no. 3, pp. 643–656, 2007.
- [21] R. R. L. Kantam, K. Rosaiah, and G. S. Rao, "Acceptance sampling based on life tests: log-logistic model," *Journal of Applied Statistics*, vol. 28, no. 1, pp. 121–128, 2001.
- [22] Y. L. Lio, T. R. Tsai, and S. J. Wu, "Acceptance sampling plans from truncated life tests based on the birnbaum-saunders distribution for percentiles," *Communications in Statistics - Simulation and Computation*, vol. 39, no. 1, pp. 119–136, 2009.
- [23] Y. Lio, T. R. Tsai, and S. J. Wu, "Acceptance sampling plans from truncated life tests based on the Burr type XII percentiles," *Journal of the Chinese Institute of Industrial Engineers*, vol. 27, no. 4, pp. 270–280, 2010.
- [24] G. S. Rao and R. R. L. Kantam, "Acceptance sampling plans from truncated life tests based on the log-logistic distributions for percentiles," *Economic Quality Control*, vol. 25, no. 2, pp. 153–167, 2010.
- [25] S. Balamurali, M. Aslam, and M. S. F. Nezhad, "An acceptance sampling plan under frechet distribution assuring median life," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 06, no. 24, pp. 4519–4523, 2013.
- [26] G. S. Rao, K. Rosaiah, K. Kalyani, and D. C. U. Sivakumar, "A new acceptance sampling plans based on percentiles for odds exponential log logistic distribution," *The Open Statistics & Probability Journal*, vol. 7, no. 1, pp. 45–52, 2016.
- [27] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [28] S. M. Awan, Z. A. Khan, M. Aslam, W. Mahmood, and A. Ahsan, "Application of NARX based FFNN, SVR and ANN Fitting models for long term industrial load forecasting and their comparison," in *Proceedings of the IEEE International Symposium on Industrial Electronics*, pp. 803–807, IEEE, Hangzhou, China, July 2012.
- [29] P. S. Yu, T. C. Yang, S. Y. Chen, C. M. Kuo, and H. W. Tseng, "Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting," *Journal of Hydrology*, vol. 552, pp. 92–104, 2017.
- [30] S. Ghosh, "SVM-PGSL coupled approach for statistical downscaling to predict rainfall from GCM output," *Journal of Geophysical Research*, vol. 115, no. 22, Article ID 22102, 2010.
- [31] D. Raje and P. P. Mujumdar, "A comparison of three methods for downscaling daily precipitation in the Punjab region," *Hydrological Processes*, vol. 25, no. 23, pp. 3575–3589, 2011.
- [32] N. Bibi, I. Shah, A. Alsubie, S. Ali, and S. A. Lone, "Electricity spot prices forecasting based on ensemble learning," *IEEE Access*, vol. 9, pp. 150984–150992, 2021.
- [33] M. H. D. M. Ribeiro, R. G. da Silva, V. C. Mariani, and L. d. S. Coelho, "Short-term forecasting COVID-19 cumulative confirmed cases: perspectives for Brazil," *Chaos, Solitons & Fractals*, vol. 135, Article ID 109853, 2020.
- [34] C. J. Lu, T. S. Lee, and C. C. Chiu, "Financial time series forecasting using independent component analysis and support vector regression," *Decision Support Systems*, vol. 47, no. 2, pp. 115–125, 2009.
- [35] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [36] A. G. Ivakhnenko, "Polynomial theory of complex systems," *IEEE transactions on Systems, Man, and Cybernetics*, no. 4, pp. 364–378, 1971.
- [37] I. Ebtehaj, H. Bonakdari, A. H. Zaji, H. Azimi, and F. Khoshbin, "GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs," *Engineering Science and Technology, an International Journal*, vol. 18, no. 4, pp. 746–757, 2015.
- [38] Z. Peng, F. U. Khan, F. Khan et al., "An Application of Hybrid Models for Weekly Stock Market Index Prediction: Empirical Evidence from SAARC Countries," *Complexity*, vol. 2021, Article ID 5663302, 2021.
- [39] R. Walton, A. Binns, H. Bonakdari, I. Ebtehaj, and B. Gharabaghi, "Estimating 2-year flood flows using the generalized structure of the Group Method of Data Handling," *Journal of Hydrology*, vol. 575, pp. 671–689, 2019.
- [40] A. Gholami, H. Bonakdari, I. Ebtehaj, S. Shaghghi, and F. Khoshbin, "Developing an expert group method of data handling system for predicting the geometry of a stable channel with a gravel bed," *Earth Surface Processes and Landforms*, vol. 42, no. 10, pp. 1460–1471, 2017.