Thailand Statistician
April 2026; 24(2): 257-276
http://statassoc.or.th
Contributed paper

A Study on unit Gompertz Distribution with Multi-Layer Artificial Neural Network Modeling

Tabassum Naz Sindhu*[a], Andaç Batur Çolak [b], Showkat Ahmad Lone [c], Tahani A Abushal [d], Nevine M Gunaime [c] and Anum Shafiq [e, f]

[a] Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan

[b] Department of Information Systems and Technologies, Niğde Ömer Halisdemir University, Niğde, Türkiye

[c] Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, (Jeddah-M), Riyadh, Saudi Arabia

[d] Department of Mathematics, College of Science, Umm Al-Qura University, Makkah, Saudi Arabia

[e] IT4Innovations, VSB -Technical University of Ostrava, Ostrava, Czech Republic

[f] Center for Theoretical Physics, Khazar University, Baku, Azerbaijan

*Corresponding author; e-mail: anumshafiq@ymail.com; anum.shafiq@vsb.cz; sindhuqau@gmail.com

Received: 27 July 2023
Revised: 4 January 2025
Accepted: 4 August 2025

Abstract

By incorporating the unit Gompertz (UG) distribution into the framework of artificial neural network (ANN) modelling, we present a novel approach in this paper. We include the UG distribution in our ANN modelling framework to improve the precision and interpretability of predictions. Through this integration, we hope to learn new things, strengthen forecast accuracy, and better understand the mechanisms at work in our data. After performing computations for a number of situations using the Hazard Rate Function (HRF), Cumulative Density Function (CDF), Probability Density Function (PDF), and Reliability (R) functions, a data collection has been developed. With this UG distribution and ANN modelling combo, it is expected that the ability to analyze and predict these functions would improve. Two separate artificial neural network models have been created using a total of 32 data sets gathered. The generated multi-layer perceptron network models utilized 15% for model validation, 70% of data for model training, and 15% for model testing. The findings demonstrate that ANNs are highly accurate at predicting the PDF, CDF, HRF, and R functions of the UG model.

Keywords: Unit Gompertz model, multi-layer perceptron, artificial neural network modeling, reliability.

1. Introduction

A continuous model with a bounded domain is required for dealing with the ambiguity of closed phenomena, such as the proportion of a defined feature. The only continuous, bounded model that has been extensively discussed in literature is known as the beta model, and evidence of its veracity dates back to Sir Isaac Newton's letter to Henry Oldenberg in 1676. To simulate the ambiguity of closed phenomena, various continuous models, including closed domains, have been created and

applied over the years in various sectors. For instance, Johnson SB model (Johnson 1949; 1955), Unit-Logistic model (Tadikamalla and Johnson, 1982), Topp-Leone model (Topp and Leone, 1955), Unit-Gamma model (Consul and Jain, 1971; Grassia, 1977; Ghitany et al., 2019; Mazucheli et al. 2019), the generalized exponentiated unit Gompertz distribution (Sindhu et al., 2023), the Kumaraswamy model (Kumaraswamy, 1980), the unit Gumbel Type-II model (Shafiq et al., 2023), the Arcsine model (Arnold and Groeneveld, 1980), the Beta Power model (Cordeiro and dos Santos Brito, 2012), the Reflected Generalized Topp-Leone model (Van Dorp and Kotz, 2006), the mixture of log-Bilal models (Lone et al., 2023), the Exponentiated Kumaraswamy model (Lemonte et al. 2013), the log-Lindley model (Gómez-Déniz et al., 2014), the Exponentiated Topp-Leone model (Pourdarvish et al., 2015), the Reflected Generalized Topp-Leone Power series model (Condino and Domma, 2017), the Marshall–Olkin Extended Kumaraswamy (Castellares and Lemonte, 2016), the Transmuted Kumaraswamy model (Khan et al., 2016), the size biased Kumaraswamy model (Sharma and Chakrabarty, 2016) and the Extended Arcsine model (Cordeiro et al., 2016). It should be noted that the majority of these models involve more than two variables and are centred on a few of the variables within their ranges. it makes estimate challenging. Numerous restricted range models in the literature are produced from regular distributions through mathematical modification.

With regard to actuarial tables and death rates for people, Benjamin Gompertz created the Gompertz model in 1825. Actuaries and Demographers have drawn a lot of attention since then. This model, which is an exponential model in general form, is widely used in real life, particularly in the actuarial and medical fields. It exhibits some positive correlations with a number of well-known models, including the exponential, Weibull, Gumbel, double exponential, and generalised logistic models (see Willekens, 2001). The exponentially escalating rate of failure during the life of a system is a key component of the Gompertz model. Recent studies on the statistical methods and characteristics of this model have been made by a number of academics; for examples, see the references (Waller et al., 2013; Makany, 1991; Rao et al. 1992; Franses, 1994; Chen, 1997; Wu and Lee, 1999).

For modelling the lifetime data, there are numerous methods in the statistics literature. These models, however, are either the result of theoretical consideration, real-world applications, or both. Gamma model, for instance, lacks a clear physical impetus. It has nicer properties and provides better fits because, with just one more parameter, it has a more general mathematical shape than the exponential model. The same reasons hold true for Weibull and a number of other models. The Gompertz model has been widely generalized by researchers. Notable attempts are defined as shifted Gompertz model (Bemmaor, 1992), wrapped generalized Gompertz model (Roy and Adnan, 2012), the generalized Gompertz model (El-Gohary et al., 2013), beta-Gompertz distribution (Jafari et al., 2014), odd generalized exponential Gompertz model (El-Damcese et al., 2015), generalized Gompertz-power series models (Jafari and Tahmasebi, 2016) and beta generalized Gompertz distribution (Benkhelifa, 2017).

Artificial neural networks (ANN) are a popular mathematical tool used in various engineering applications. They are developed through simulating the biological structure of the human brain (Vakili et al., 2016). Because of their ability to model nonlinear complex functions, ANNs are frequently preferred by researchers (Bahiraei et al., 2019). An ANN model was developed from experimental data of ZrO_2/Water nanofluid prepared in five different concentrations by Çolak (2021). The generated multi-layer perceptron (MLP) network model's findings are contrasted with the outcomes of the experiments. The findings demonstrated that created artificial neural network had an average error rate of -0.11% when predicting viscosity of ZrO_2/Water nanofluid. Duan and Li (2021) investigated the temporary heat transfer process of the phase change material (PCM-porous) containing a porous medium by using the ANN method. An experiment conducted in the lab was used to investigate the PCM-porous system's transient heat exchange process. The findings of the experiment demonstrated that the metal foam's porosity has a higher impact on PCM melting and heat exchange than its pore density. An ANN model was created in the study to forecast the metal foam's porosity using the liquid temperature, fraction, and average Nusselt number. The results indicated that the ANN model yielded the most favorable outcomes and exhibited a high level of predictability. Shetty et al. (2021)

employed an ANN model to forecast the thermo-hydraulic performance of a cross-flow circular orifice absorber plate. The values acquired from the experimental investigation were contrasted with the predictions made by the ANN. In order to provide ANN findings, 64 data sets, 45 training samples, 9 testing samples, and 9 verification samples were examined. The hidden layer consists of a Lavenberg Marquardt algorithm with 8 inputs and one output with 11 neurons. ANN models showed an average error of 1.58% based on experimental thermohydraulic efficiency. In their study, Colak et al. (2021) examined the flexural and compressive strength of cement mortars with waste PET aggregates at high temperatures through experimentation. In each concentration, mortar samples were mixed with PET aggregate replacement and heated to 100, 150, 200, 250, 300, 400 °C. Following 1, 2, and 3 hours of waiting at the prescribed temperatures, flexural and compressive strength tests were performed. It was discovered that compressive and flexural strength values decreased with increasing temperature and PET aggregate concentrations in all mixes. Using experimental data, an ANN model was created to predict compressive and flexural strength values. According to the study, ANN can accurately predict flexural and compressive strengths with an average error of -0.51%.

According to an in-depth review of the literature, very few studies have been done on the UG model, which was originally utilized by Muzucheli et al. (2019). However, there have been no studies done on how to model the UG model's properties using ANN. ANN modelling and numerical analysis are both utilized to analyze various UG distribution features. There are no studies that compare and improve lifespan model properties using ANN modelling in the literature. This study, in which the behavior of the PDF, CDF, HRF, and RF have been thoroughly investigated and optimized using ANN modelling, is crucial since it attempts to fill a significant gap in the body of existing literature. This paper has the following layout. In Section 2, our aim is to identify and articulate the fundamental mathematical and statistical traits that underlie the UG model. Section 3 includes Neural Network Modeling. Results and discussions are provided in Section 4. Section 5 concludes with final remarks.

2. Unit Gompertz Distribution (UGD)

By using $Y = e^{-X}$, if X has two parameters Gompertz distribution, then cdf is presented as

$$\mathcal{F}\left(x;\ b,\ \eta\right) = 1 - \exp\left\{\eta\left(1 - e^{bx}\right)\right\} \qquad b,\ \eta,\ x > 0. \tag{1}$$

We obtain unit Gompertz (UG) distribution (Benkhelifa, 2017) with the pdf

$$f\left(y|\, b,\ \eta\right) = b\eta y^{-b-1} \exp\left\{\eta\left(1 - y^{-b}\right)\right\},\ b,\ \eta > 0, \tag{2}$$

and cdf is

$$F\left(y|\, b,\ \eta\right) = \exp\left\{\eta\left(1 - y^{-b}\right)\right\}, \tag{3}$$

where scale and structure parameters represented as $y \in (0,1)$, and $b, \eta, > 0$, respectively. It is convenient and obvious to note that $F(y|\, b,\ \eta)$ differentiable and increases between 0 and 1 and $\lim_{y\to 0} F(y) = 0$ and $\lim_{y\to 1} F(y) = 1$.

A reliability paradigm analyzes the probability of longevity or failure of a system. The $R(y) = 1 - F(y) = P(Y > y)$ figures out the SF of a r.v. Y. It can be expressed as the likelihood of a system failing within a given time frame. The Reliability (R) function of UG is defined by

$$R(y|\, b,\ \eta) = 1 - \exp\left\{\eta\left(1 - y^{-b}\right)\right\}. \tag{4}$$

The hrf $h(y|\, b, \eta) = f(y|\, b, \eta)/[1 - F(y|\, b, \eta)]$ is a quite helpful tool for lifetime study. The instant failure rate of a random variable Y is probability of a system failing as it has existed to current moment t is given as

$$h(y|\, b,\ \eta) = \frac{\eta b y^{-b-1} \exp\left\{\eta\left(1 - y^{-b}\right)\right\}}{1 - \exp\left\{\eta\left(1 - y^{-b}\right)\right\}}. \tag{5}$$

The odd ratio is defined as

$$\Upsilon_y\left(y|\, b,\ \eta\right)=\frac{R_y\left(y\right)}{h_y\left(y\right)}=\frac{\left(1-\exp\left\{\eta\left(1-y^{-b}\right)\right\}\right)^2}{\eta b y^{-b-1}\exp\left\{\eta\left(1-y^{-b}\right)\right\}}. \tag{6}$$

The cumulative hrf of UG model is obtained as

$$H\left(y\right)=\int_0^y h(t|\, b,\ \eta)dt=-\log\left\{1-\exp\left\{\eta\left(1-y^{-b}\right)\right\}\right\}, \tag{7}$$

where $R_y\left(y\right)$ and $h_y\left(y\right)$ is defined in (4) and (5). Equations (1)-(7) can be conveniently evaluated numerically utilizing computing packages like Maple, MATLAB, Mathematica, Minitab and R.

2.1. Random number generator and related measures

$\mathcal{F}\left(y\right)$ is inverted to generate r.v. Y as mentioned. Suppose ζ be a r.v. and $\zeta\sim U(0,1)$. As a solution of the equation, an observation of Y can then be performed

$$\log\left(\zeta\right)=\eta\left(1-y^{-b}\right),$$

$$Y=Q\left(\zeta,\eta,b\right)=\left(1-\frac{\log\left(\zeta\right)}{\eta}\right)^{-\frac{1}{b}}\sim \text{UG}\left(\eta,b\right). \tag{8}$$

Specifically, by placing $\zeta=(0.25,0.50,0.75)$ in (8), the 1^{st}, 2^{nd}, and 3^{rd} quartiles are attained.

Important quantities of $\text{S}_{skewness}$ and $\text{K}_{kurtosis}$ are provided by $\phi_3=\mu_3/\sigma^3$ and $\phi_4=\mu_4/\sigma^4$, respectively, where μ_κ is the foundational κ^{th} moment and σ is standard deviation. Although for some parameteric values, moments of model cannot occur, relevant quantile-based measures $\text{S}_{kewness}$ and $\text{K}_{urtosis}$ are more appropriate often. For models without moments, such indicators are more reliable and do occur. S_B (Bowley's Skewness) and K_M (Moors' Kurtosis) indicators are defined as

$$S_{B|b,\,\eta}=\frac{Q\left(6/8|\, b,\ \eta\right)+Q\left(2/8|\, b,\ \eta\right)-2Q\left(4/8|\, b,\ \eta\right)}{Q\left(6/8|\, b,\ \eta\right)-Q\left(2/8|\, b,\ \eta\right)}, \tag{9}$$

$$K_{M|b,\,\eta}=\frac{Q\left(7/8|\, b,\ \eta\right)-Q\left(5/8|\, b,\ \eta\right)+Q\left(3/8|\, b,\ \eta\right)-Q\left(1/8|\, b,\ \eta\right)}{Q\left(6/8|\, b,\ \eta\right)-Q\left(2/8|\, b,\ \eta\right)}, \tag{10}$$

where qf is represented by $Q\left(.|\, .\right)$. Therefore, if $S_{B|b,\,\eta}<0$, $S_{B|b,\,\eta}>0$, and $S_{B|b,\,\eta}=0$, then distribution is called left, right skewed, and symmetrical. Respectively, a higher value of $K_{M|b,\,\eta}$ means a strong tail for model and a small finding of $K_{M|b,\,\eta}$ implies a mild tail conversely.

2.2. Moments and some associated indicators and functions

Moments are immensely important for statistical analysis, typically in implementations. The particular parameters which may be utilized to determine the behaviour of a homogeneous data set are named moments. For UG model, we obtain k^{th} ordinary moment $\acute{\rho}_k$ as follows

$$\acute{\rho}_k=E\left(Y^k\right)=\int_0^1 y^k dF\left(y|\, b,\ \eta\right);\ k=1,2,... \tag{11}$$

Initially, some notation is introduced. The upper incomplete gamma function, represented by the symbol $\Gamma\left(\varrho,x\right)$, is defined as

$$\Gamma\left(\varrho,x\right)=\int_x^\infty u^{\varrho-1}e^{-u}du,\ x>0,\ \varrho\in\mathbb{R}. \tag{12}$$

Moreover, the upper incomplete gamma function can be represented in terms of exponential integral function in following manner

$$E_{\hbar}(x) = \int_{1}^{\infty} t^{-\hbar} e^{-tx} dt,\ x > 0,\ \hbar \in \mathbb{R}, \tag{13}$$

$$E_{\hbar}(x) = x^{\hbar-1}\Gamma(1-\hbar, x),\ \hbar,\ x \in \mathbb{R}. \tag{14}$$

The given results presents an equation for k^{th} ordinary moment $\acute{\rho}_k$ of Y in respect of exponential integral function and gamma function. For $b, \eta > 0$, the k^{th} ordinary moment of Y can be converted as

$$\acute{\rho}_k(y|\, b,\ \eta) = \eta \exp(\eta) E_{\frac{k}{b}}(\eta) = \eta^{\frac{k}{b}} \Gamma(1 - k/b,\ \eta) \exp(\eta). \tag{15}$$

Proof: Consider the following integral

$$\acute{\rho}_k(y|\, b,\ \eta) = \int_{0}^{1} y^k b\eta y^{-b-1} \exp\left\{\eta\left(1 - y^{-b}\right)\right\} dy. \tag{16}$$

Allowing $t = x^{-b}$ using the result $-bx^{-b-1}dx = dt$ and we have after some algebraic manipulation

$$\acute{\rho}_k(y|\, b,\ \eta) = \eta \exp(\eta) \int_{\infty}^{1} t^{\frac{-k}{b}} \exp(-\eta t)(-dt). \tag{17}$$

We get, after using (12) and (14) in (16), we obtain the final result. That fills out the proof. □

Specifically, mean of Y is

$$\acute{\rho}_1 = E(Y) = \eta^{\frac{1}{b}} \Gamma(1 - 1/b,\ \eta) \exp(\eta), \tag{18}$$

and facilitate the equation of variance is

$$Var(Y) = \acute{\rho}_2 - (\acute{\rho}_1)^2 = \eta^{\frac{2}{b}} \exp(\eta) \left[\Gamma(1 - 2/b,\ \eta) - \exp(\eta)\{\Gamma(1 - 1/b,\ \eta)\}^2\right]. \tag{19}$$

In model characterization, moment generating function (mgf) is widely used. It is easy to express the mgf of UG model as

$$M(y|\, b,\ \eta) = \sum_{\varsigma=0}^{\infty} \frac{t^{\varsigma}}{\varsigma!} \acute{\rho}_k(y|\, b,\ \eta) = \sum_{\varsigma=0}^{\infty} \frac{t^{\varsigma}}{\varsigma!} \eta^{\frac{k}{b}} \exp(\eta) \Gamma(1 - k/b,\ \eta). \tag{20}$$

As an graphical illustrations, for certain parameteric values, Mean, Variance, $Median|_{b,\eta}$, $S_B|_{b,\eta}$, and $K_M|_{b,\eta}$ of the UG model are given in Figures 7-11. These outcomes indicate that there are important effects of b and η on said measures. We still have $S_B|_{b,\eta} > 0$ so the model is right-skewed with variations for $K_M|_{b,\eta}$.

3. Neural Network Modeling

Two alternative ANN models have been created within the parameters of the study to forecast the values of PDF, HRF, CDF, and reliability (R). It is not possible to predict four values using a single ANN model because of different input parameters of parameters examined. Due to this, a separate ANN model has been created to predict HRF and R values, while an ANN model has been created to predict PDF and CDF values. The feed-forward (FF) back-propagation (BP) multi-layer perceptron (MLP) model has been favoured in ANN models because to its great performance (Çolak

et al., 2023). MLP network models include an output layer where predictions are obtained, at least one hidden layer, and an input layer. The basic configuration structure of an MLP network model is depicted in Figure 1.

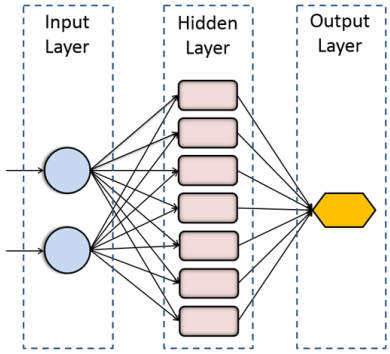


Figure 1 A MLP network model's configuration structure

A neuron is a sort of computational unit found in the hidden layer of an MLP network. The performance of model's prediction depends on how well the data used in ANN modelling are optimised (Çolak, 2021). The optimization model that emerged from studies in the literature has been used to both generated ANN models (Esmaeilzadeh et al., 2021; Shafiq et al., 2023; Rostamian et al., 2017). 15% of the data set was utilised for validation, 15% for testing, and 70% for training. The ANN models were trained using a total of 32 pieces of data; the b, η and y values served as input parameters, while the PDF, CDF, HRF, and R values were predicted at output layer. The number of neurons to be employed in hidden layer of MLP networks is not predetermined by any acknowledged criterion or algorithm (Bonakdari and Zaji, 2016). Due to this, predictive performance of ANN models created with various numbers of neurons has been examined and best performing ANN models have been chosen (Sindhu et al., 2023). The hidden layer of ANN model created to predict PDF and CDF values contains 7 neurons, whereas hidden layer of ANN model created to predict HRF and R values contains 6 neurons. The developed ANN models' fundamental structure is depicted in Figure 2.

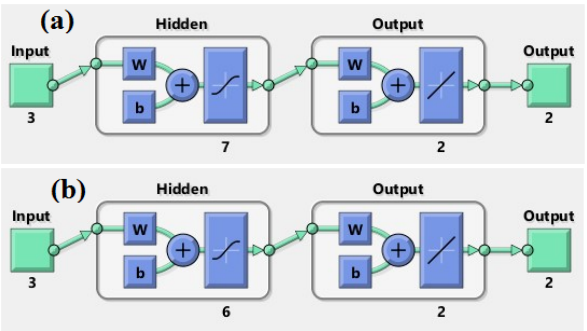


Figure 2 Basic structure of developed ANN models

The effectiveness of various training algorithms employed in ANN models has been examined in literature (Ali et al., 2019). The created ANN models employ high predictive performance Levenberg-Marquardt training technique (Çolak, 2021; Abdul Kareem et al., 2017; Vafaei et al., 2017; Uthansakul et al. 2019; Kujur et al. 2023). In the hidden layer of the MLP network, the Tan-

Sig function serves as the transfer function, while the Purelin function serves as the output layer's transfer function (Shafiq et al., 2022). Below are the transfer functions that were employed.

$$f(x) = \frac{1}{1+e^{-x}}, \quad \text{pureline}(x) = x. \tag{21}$$

The accuracy of ANN models for making predictions has been examined using the coefficient of determination (R), Mean Squared Error (MSE), and error rate metrics. Below are the equations that were employed to calculate the performance parameters.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(X_{\exp(i)} - X_{ANN(i)}\right)^2, \tag{22}$$

$$\tilde{R} = \sqrt{1 - \frac{\sum_{i=1}^{N}\left(X_{\exp(i)} - X_{ANN(i)}\right)^2}{\sum_{i=1}^{N}\left(X_{\exp(i)}\right)^2}}, \tag{23}$$

$$\text{Error Rate}\,(\%) = \left[\frac{X_{\exp} - X_{ANN}}{X_{\exp}}\right] \times 100. \tag{24}$$

Regarding the UG distribution's constraints, it should be mentioned that because of its limited character, it might not offer as much versatility as unbounded distributions when it comes to resembling complex data patterns. Some statistical models might not adequately account for the features of real-world data, such as heavy tails, skewness, or multimodality. As their main goal is to accurately reflect data contained within a predetermined range, bounded probability distributions are frequently insufficient for the purpose of modelling extreme values or outliers. These distributions might not appropriately mimic or depict extreme values. Additionally, a drawback of the constructed ANN model is its capacity to estimate input and output parameters between a range of minimum and maximum values.

4. Results and Discussions

The sketches of PDF and CDF are given in Figures 3 and 4 for various values of pertinent parameters. Figure 3 demonstrate how parameters b and η influence the density of UG and demonstrate flexibility of the PDF forms where small symmetry, skewness, modality and high tails can be calculated. Figure 4 displays the CDF, which shows increasing pattern of CDFs. These figures reflect flexibility of UG distribution. The probability density function (PDF) of this distribution may exhibit a range of shapes, including decreasing, reversed J-shape, unimodal, and positively skewed (as illustrated in Figure 3). Additionally, the hazard rate function of this distribution can have an increasing, upside-down bathtub, and then bathtub shape (as demonstrated in Figure 5). Figures 6 displays reliability function for various values of pertinent parameters.

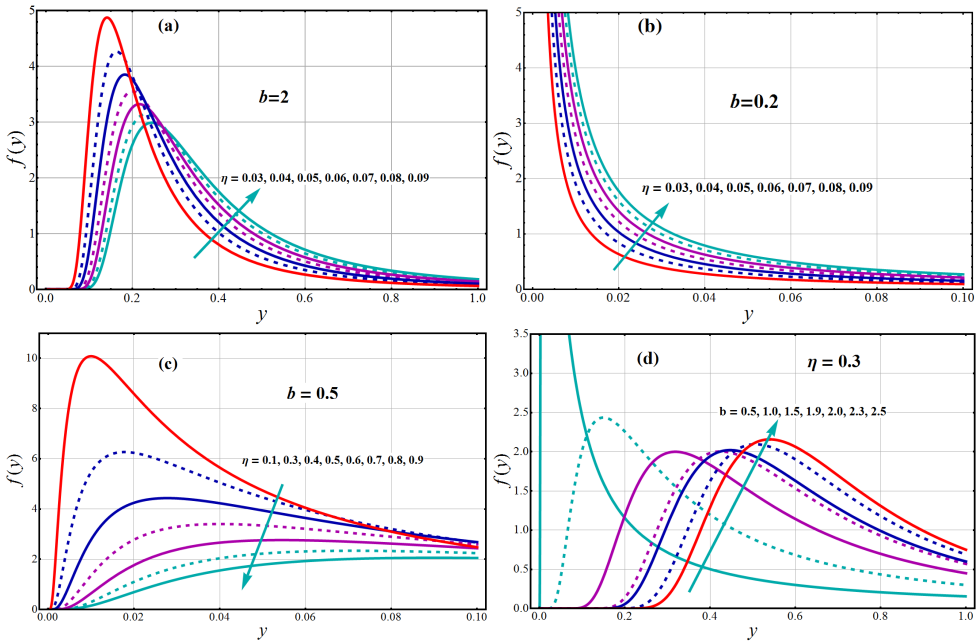


Figure 3 Graphs of pdf of UG at different parameteric values

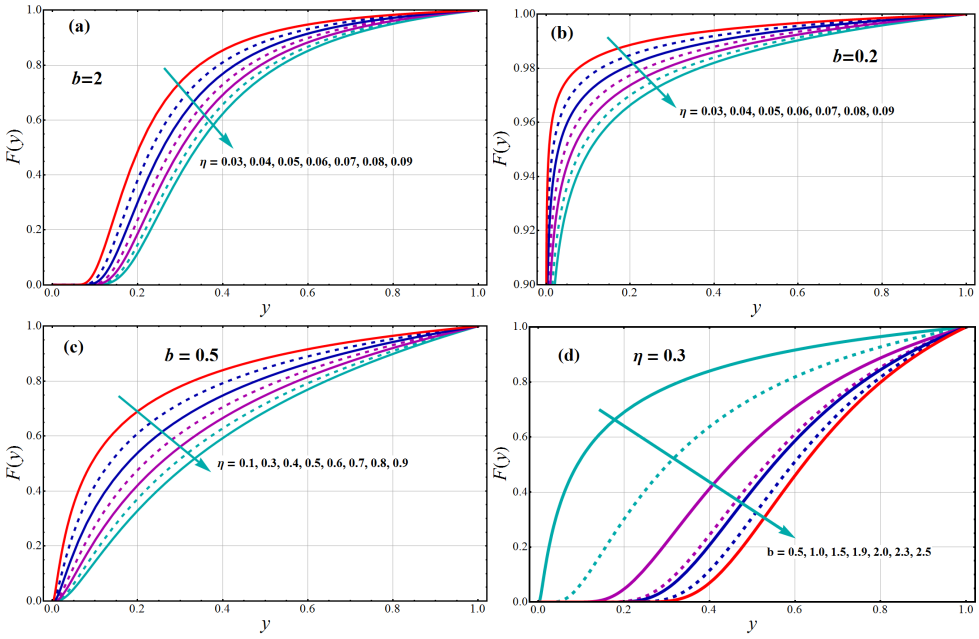


Figure 4 Graphs of cdf of UG at different parameteric values

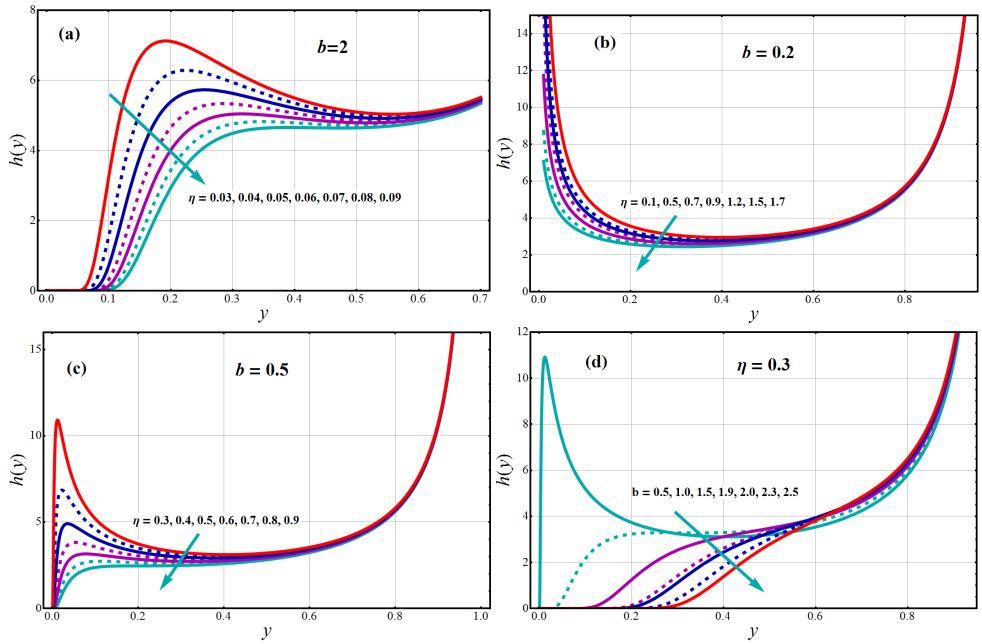


Figure 5 Graphs of hrf of UG model at different parameteric values

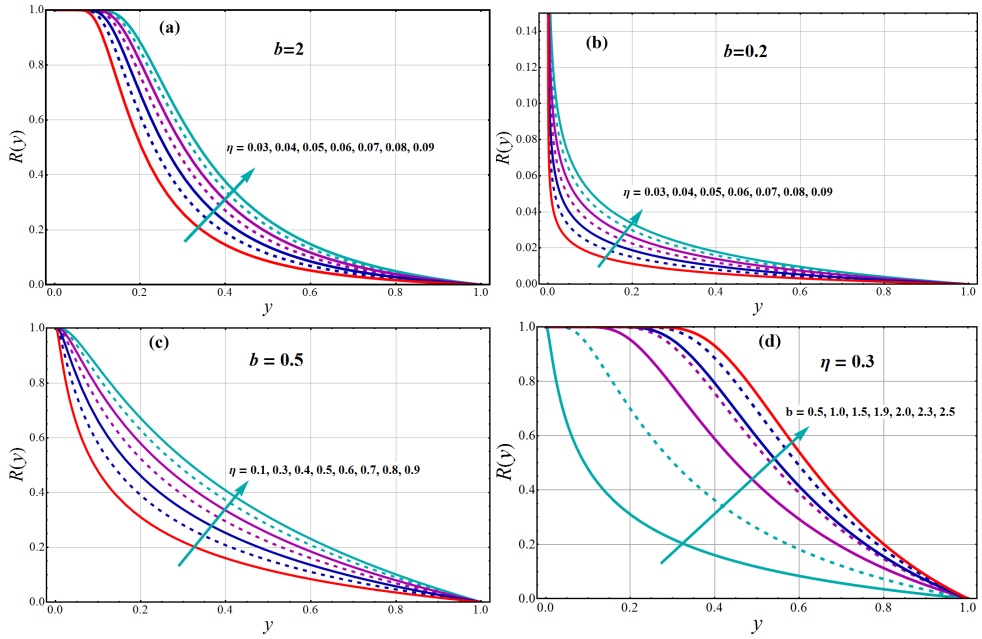


Figure 6 Graphs of Reliability function of UG model at different parameteric values

Figures 7, 8 and 9 display the 3D and isotherm graphs of mean, variance and median behaviour in comparison to UG model and considered parameteric values. It is noted that mean is growing at higher values of b and all levels of η (see Figure 7a). Furthermore, in Figure 7b, it is noted that maximum mean occurs near the higher level of parameters b and η whereas it reduces for smaller values of parameters b and all the level η. Additionally according to Figure 8a, mixed behavior is observed for variance. On the other hand, in Figure 8b, when b lies between 0.5 to 0.6, and maximum level of η, variance approches to zero, while for b lies between 0.4 to 0.5, variance attains higher level for maximum level of η. Figure 9a displays median of UG model as a function of different pertinent parameters. It is observed that for larger levels of both parameters b and η, median is rised, while it reduces for small values of parameters b and η. Moreover when $1 < \eta < 2$, and $3 < b < 5$ median approaches higher level. For lower level of b and η least median is observed (see Figure 9b). In the context of UG model and according to parameteric values, Figures 10 and 11 graphically analyse the behaviour of $S_{B|b,\,\eta}$ and $K_{M|b,\,\eta}$. In Figure 10, positively skewed behavior is observed for UG model for considered parameteric values. For $1 < \eta < 2$, and $3 < b < 5$, Skewness is higher. But for all levels of η and when $3 < b < 5$, minimum Skewness is observed see Figures 10(a-b). Another important property of distribution is its kurtosis, which refers to its degree of peakedness. This is illustrated in Figure 11. The UG model shows leptokurtic and platykurtic behavior. The prefix "platy-" stands for "wide;" as well as having thinner tails and a lower, wider peak around the mean in terms of shape. Lepto-, which means "slender" or "thin," has a similar meaning. A leptokurtic distribution has wider tails and a more acute peak around the mean. The leptokurtic distribution is observed for $1 < \eta < 2$ and $2.5 < b < 5$ and platykurtic distribution is observed for $0.1 < b < 1.5$ and all the levels of η.

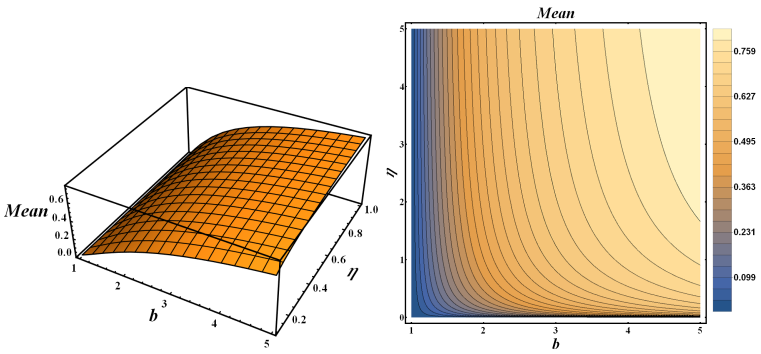


Figure 7 3D and contour curves of mean of UG model

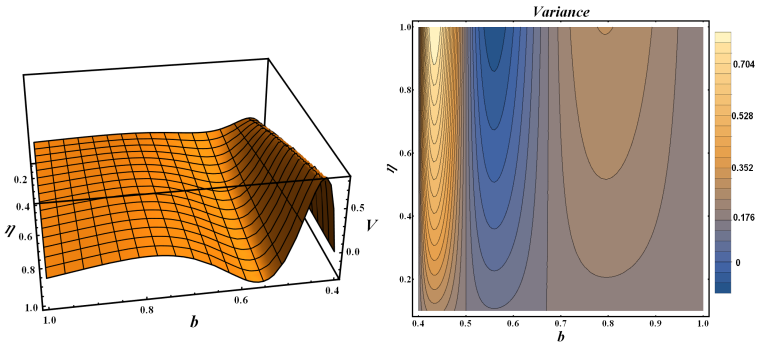


Figure 8 3D and contour graphs of variance of UG distribution

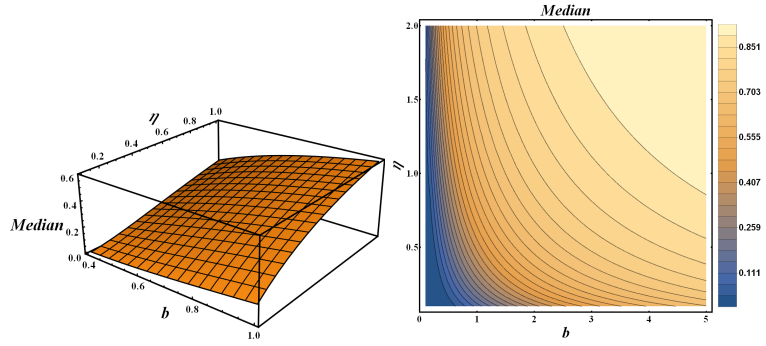


Figure 9 3D and contour graphs for Median of UG model

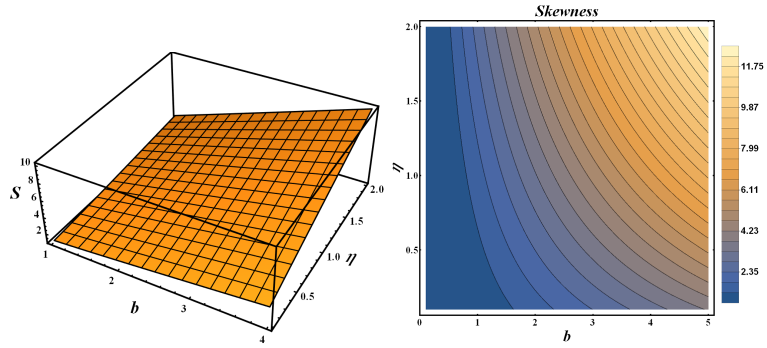


Figure 10 Graphs of Skewness function and contour plot of Skewness function

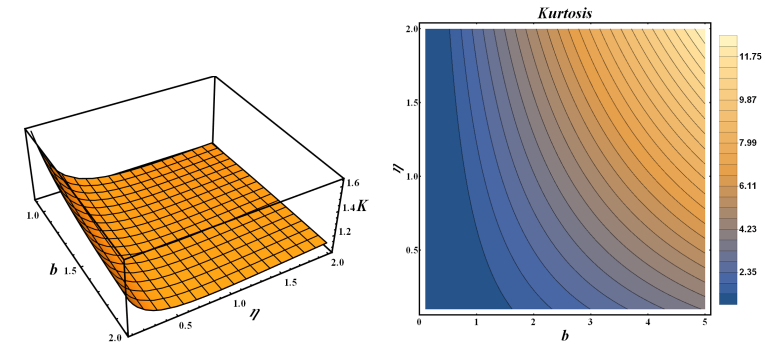


Figure 11 Graphs of Kutosis function and contour plot of Kutosis function

Figure 12 displays the ANN models' training performance. MSE values, that were high at the start of training procedure, began to drop even as training epochs progressed. The training stages of ANN models have been completed with the lowest MSE value and the best validation performance. The optimal completion of the training stages of ANN models is indicated by MSE values that are close to zero. Figure 13 displays the ANN models' training status. The graphs show that the gradient and mu values go less as the epochs get larger. Since the training phase ends after six repetitions of an error, the validation check value equals 6. Analyzing error histograms during ANN model training provides valuable information about model performance. The error histograms illustrate the differences between the predictive values obtained from the ANN model and those obtained from

the target data. Figure 14 shows error histogram graphs where the error values obtained are located close to the zero error line for each data set. In spite of this, the graphs reveal that the numerical values of the errors are also very small. Based on the error histograms, it is evident that ANN models have very low errors during the training stages. The estimated MSE values for ANN models are displayed in Figure 15. The graphs show that the MSE values found for each data point are nearly zero. Another sign that ANN models are grown to their full potential is the fact that MSE values are estimated near to zero. Figure 16 displays the ANN prediction values and target values for each data point utilized to create the ANN models. The data points acquired from the ANN model and the numerical values, which are the target values, are positioned relatively close to one another when the data points' locations are checked. This exact agreement between the data points demonstrates how well the projected values from the ANN model match the target data. This demonstrates how ANN models are created to produce predictions that are highly accurate. On Figure 17, the $y-$axis represents the prediction values from the ANN model, while the $x-$axis represents the target data. Each data point is positioned quite close to the equality line when the data points are evaluated. The distance between the data points and the equation line indicates how close the projected values from the ANN model are to the target data. Figure 18 shows the calculated error rates for each data point. The data points in the graphs are noted to be around the zero error line. The figures, however, reveal that the line depicting the average error lines also follows a similar pattern to the zero error line. R values of 0.99986 and 0.9874 were calculated for ANN models, respectively. Another indication that the constructed ANN models can produce predictions with very low error rates is the values of the R and error rates. Table 1 provides numerical values for the performance parameters of ANN models.

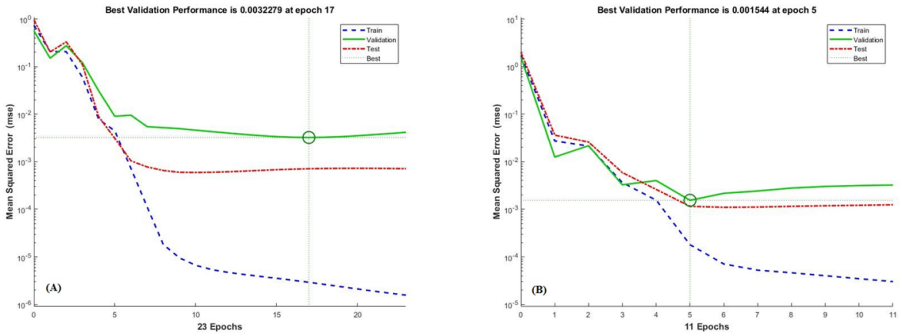


Figure 12 The training performance of ANN models

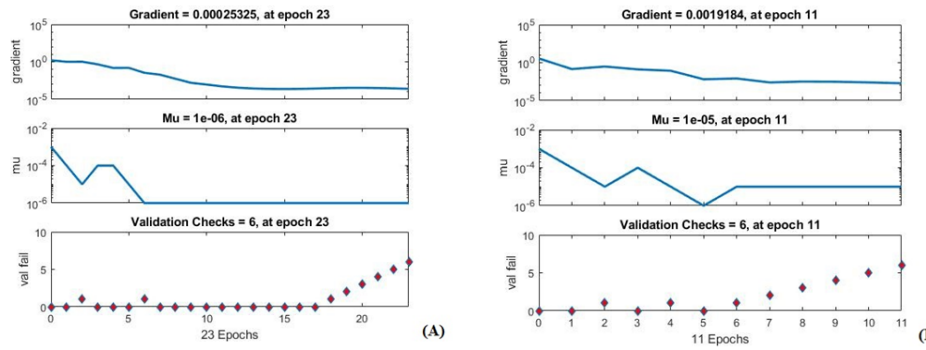


Figure 13 The training status of ANN models

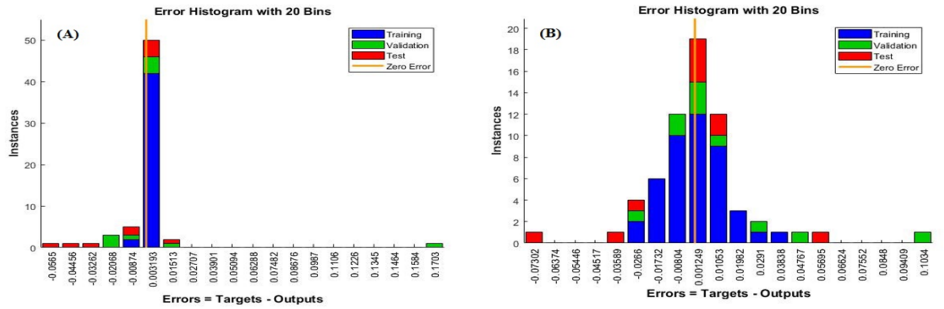


Figure 14 Error histograms for ANN models

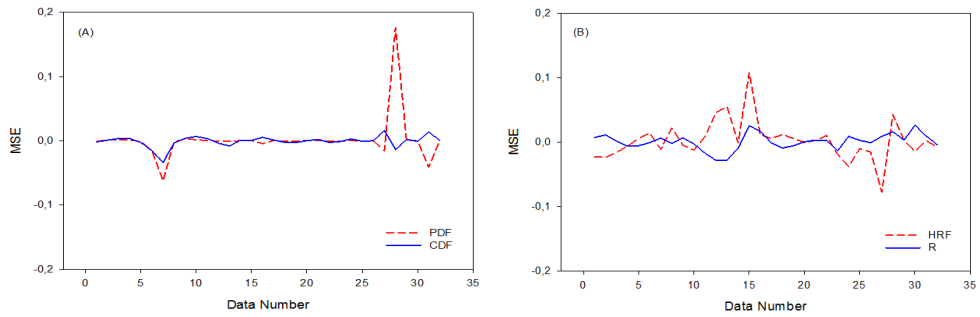


Figure 15 MSE values for ANN models

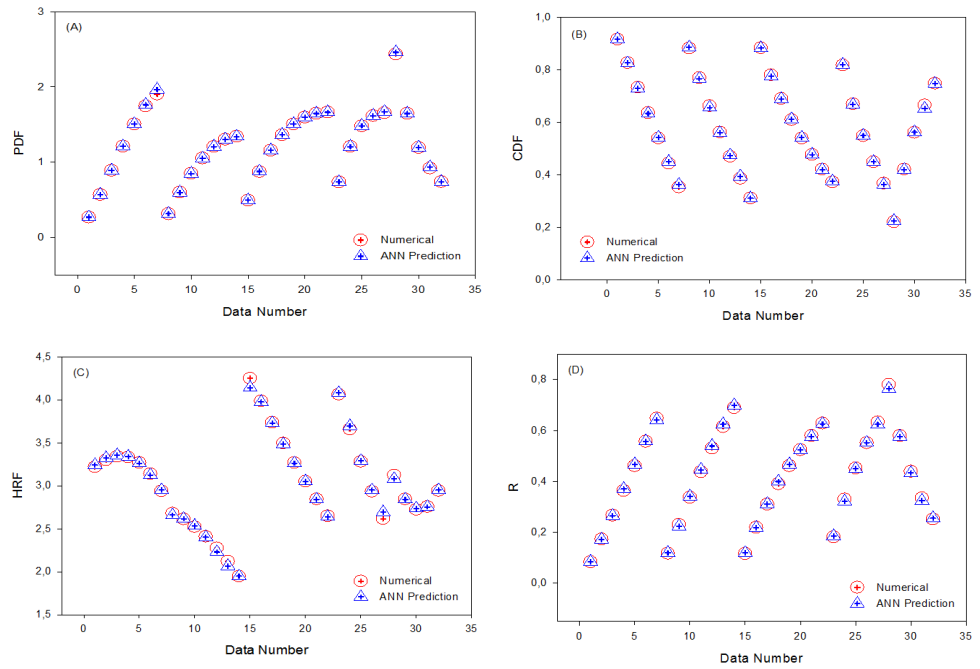


Figure 16 ANN prediction values and target values according to data number

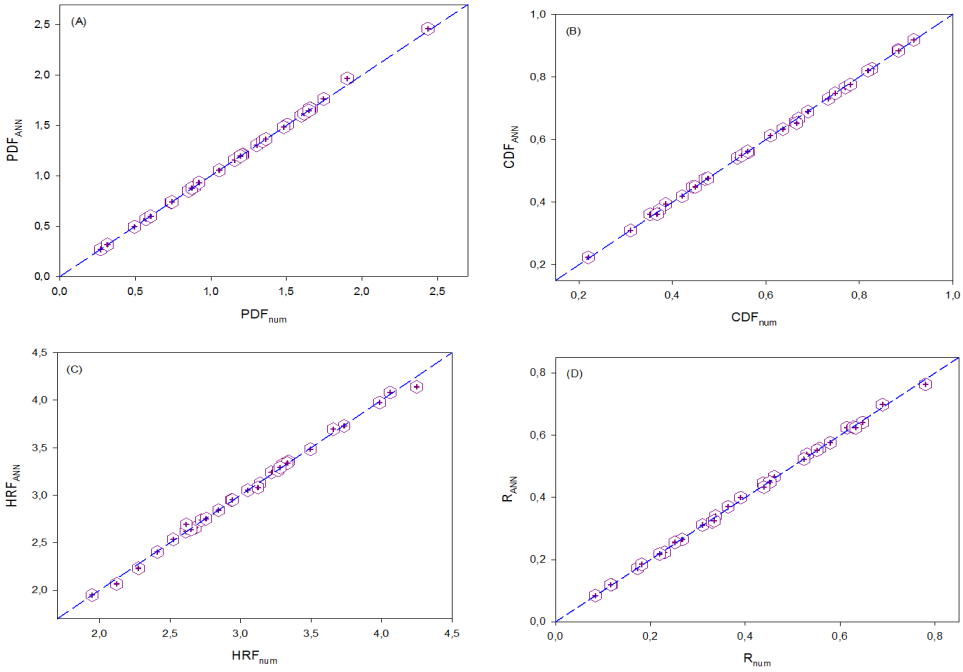


Figure 17 ANN predictions vs target data

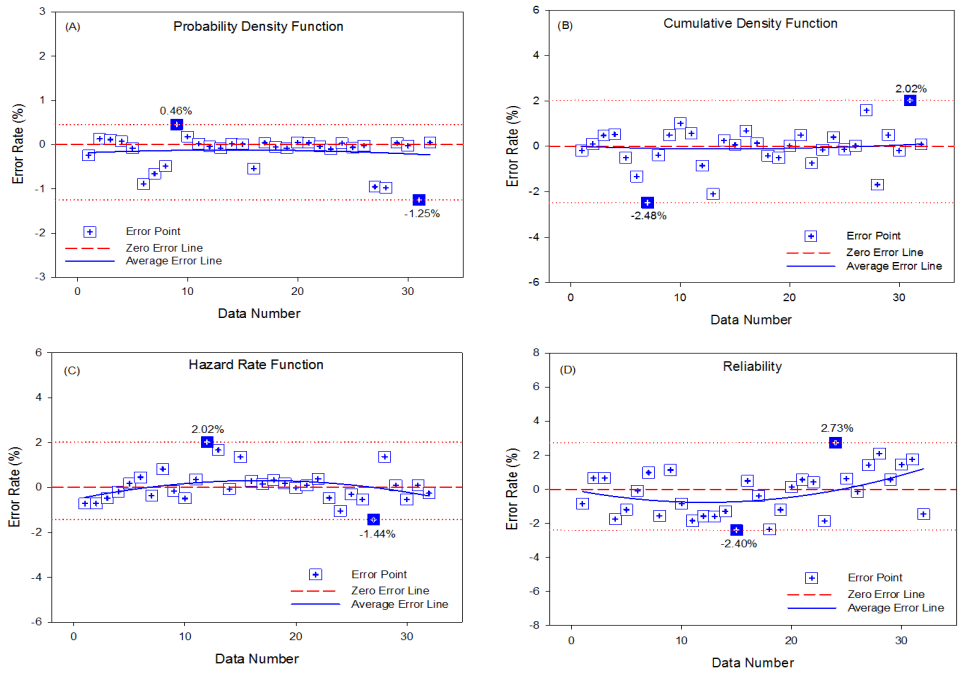


Figure 18 Error rates according to data number

Table 1 Performance parameters of the ANN models

	MSE	$\tilde{R}$	Average Error Rate (%)
Probability Density Function	1.18×10^{-03}		-0.17
		0.99874	
Cumulative Density Function	-8.77×10^{-04}		-0.07
Hazard Rate Function	2.37×10^{-03}		0.07
		0.99986	
Reliability	8.13×10^{-04}		-0.21

Table 2 Numerical values of PDF and CDF

	b	η	y	PDF		CDF	
				Numerical	ANN Prediction	Numerical	ANN Prediction
	0.1	0.5	0.2	0.269103	0.269757041	0.916393	0.918057679
	0.2			0.570567	0.569809773	0.827071	0.826183685
	0.3			0.891207	0.89018775	0.733206	0.729701787
Scenario-I	0.4			1.21161	1.21073153	0.636464	0.633102181
	0.5			1.50656	1.507723042	0.539003	0.541720249
	0.6			1.74694	1.762392432	0.443408	0.449291072
	0.7			1.90339	1.915855297	0.352542	0.361281295
	0.1	1.3	0.4	0.314413	0.315929772	0.882721	0.886159686
	0.2			0.601104	0.598350413	0.769925	0.766082296
	0.3			0.850673	0.849128413	0.66279	0.656090248
Scenario-II	0.4			1.05482	1.054628304	0.562418	0.559197858
	0.5			1.20704	1.207531867	0.469785	0.473785109
	0.6			1.30328	1.304262087	0.38569	0.393774782
	0.7			1.34245	1.342229875	0.310712	0.309925189
	0.5	0.1	0.2	0.494019	0.493955633	0.883727	0.883148401
		0.2		0.873155	0.8778491	0.780974	0.775629319
		0.3		1.157450	1.156963695	0.690168	0.689197108
Scenario-III		0.4		1.363820	1.364553999	0.609920	0.612432062
		0.5		1.506560	1.507723042	0.539003	0.541720249
		0.6		1.597670	1.596895542	0.476332	0.476257778
		0.7		1.647220	1.646511905	0.420947	0.418862475
		0.8		1.663650	1.664279315	0.372003	0.374765369
	1.2	0.1	0.4	0.737341	0.738090464	0.818501	0.819745566
		0.2		1.207030	1.206646195	0.669943	0.667231699
Scenario-IV		0.3		1.481930	1.482803576	0.548349	0.549085582
		0.4		1.617280	1.617596916	0.448824	0.448715837
		0.5		1.654680	1.670429735	0.367363	0.361537538
	0.5	0.7	0.1	2.436250	2.45995802	0.220117	0.223839667
			0.2	1.647220	1.646511905	0.420947	0.418862475
Scenario-V			0.3	1.194970	1.195294833	0.561008	0.562047315
			0.4	0.921103	0.932577807	0.665779	0.65232233
			0.5	0.740781	0.740448868	0.748301	0.747652164

Table 3 Numerical values of HRF and Reliability

	b	η	y	HRF		Reliability	
				Numerical	ANN	Numerical	ANN
	0.1	0.5	0.2	3.21869	3.241708464	0.0836065	0.084323864
	0.2			3.29943	3.323080756	0.172929	0.171800923
	0.3			3.34043	3.356258189	0.266794	0.265060716
Scenario-I	0.4			3.33284	3.338795613	0.363536	0.369893714
	0.5			3.26805	3.262288815	0.460997	0.466529512
	0.6			3.13863	3.124340429	0.556592	0.557116692
	0.7			2.93979	2.950696378	0.647458	0.641119391
	0.1	1.3	0.4	2.68091	2.658774292	0.117279	0.119115253
	0.2			2.61265	2.616826449	0.230075	0.227462712
	0.3			2.52268	2.535095627	0.337210	0.340068458
Scenario-II	0.4			2.41057	2.401841657	0.437582	0.445656478
	0.5			2.27651	2.230555365	0.530215	0.538589725
	0.6			2.12153	2.08615961	0.614310	0.624066771
	0.7			1.94759	1.948846136	0.689288	0.698265249
	0.5	0.1	0.2	4.24879	4.190775594	0.116273	0.11906067
		0.2		3.98653	3.975082099	0.219026	0.217950612
		0.3		3.73572	3.729935188	0.309832	0.311030078
Scenario-III		0.4		3.49627	3.484407664	0.39008	0.399236161
		0.5		3.26805	3.262288815	0.460997	0.466529512
		0.6		3.05091	3.051414319	0.523668	0.523005296
		0.7		2.84468	2.842021418	0.579053	0.575777864
		0.8		2.64913	2.638977789	0.627997	0.625351458
	1.2	0.1	0.4	4.06250	4.08135228	0.181499	0.184882423
		0.2		3.65703	3.695125335	0.330057	0.32103015
Scenario-IV		0.3		3.28114	3.291099153	0.451651	0.448858845
		0.4		2.93424	2.949879422	0.551176	0.551963502
		0.5		2.61553	2.653195361	0.632637	0.623611388
	0.5	0.7	0.1	3.12387	3.081004817	0.779883	0.763614119
			0.2	2.84468	2.842021418	0.579053	0.575777864
Scenario-V			0.3	2.72207	2.736648231	0.438992	0.432639739
			0.4	2.75597	2.753398589	0.334221	0.328360753
			0.5	2.94312	2.950697173	0.251699	0.255368516

This part scrutinizes the consequences of pertinent parameters on PDF, RF, HRF and CDF for different values of related parameters are displayed in Tables 2 and 3 for different scenarios . Here we have considered V scenarios : (i) b varies while $\eta = 0.5$ and $y = 0.2$, (ii) b varies while $\eta = 1.3$ and $y = 0.4$, (iii) η varies while $b = 0.5$ and $y = 0.2$, (iv) η varies while $b = 1.2$ and $y = 0.4$ and (v) y varies while $b = 0.5$ and $\eta = 0.7$. It is observed that in scenarios I and II the similar behaviour of PDF is noted i.e. PDF are increasing function of b by taking $\eta = 0.5, 1.3$ and $y = 0.2, 0.4$ while it declines for CDF under both scenarios (see Table 2). Moreover for scenarios III and IV are related to variation in η by taking other parameters fixed. It is noted in these scenarios that CDF enhances with increment in η by keeping fixed $b = 0.5, 1.2$ and $y = 0.2, 0.4$. Also, the monotone decreasing behavior is found in CDF for higher b. In case V, CDF improves when y fluctuates and b and η are constant.

The PDF behavior is observed when b changes from 0.1 to 0.7 while other parameters are kept constant. The monotone increasing behavior is found in PDF for higher values of b.Moreover for scenario II the same behavior is observed. For scenarios III and IV are related to variation in η by taking other parameters fixed the monotone increasing behavior is noticed. Whereas when y varies while $b = 0.5$ and $\eta = 0.7$ are kept fixed it declines. Further, the solutions of the numerical results are in good agreement with the ANN prediction outputs.

The numerical values and ANN prediction results of RF and HRF are reported in Table 3. The RF behavior is observed when b changes from 0.1 to 0.7 while other parameters are kept constant. The monotone increasing behavior is found in RF for higher values of b.Moreover for scenario II the same behavior is observed. For scenarios III and IV are related to variation in η by taking other parameters fixed the monotone increasing behavior is noticed. Whereas when y varies while $b = 0.5$ and $\eta = 0.7$ are kept fixed it declines.Additionally, there is good agreement between the numerical results' solutions and the ANN prediction outputs.

The above scenarios are also considered for HRF and reported in Table 3. It is analyzed that mixed behavior is noted for HRF when b changes from 0.1 to 0.7 while other parameters are kept constan but it declines when b varies while $\eta = 1.3$ and $y = 0.4$. As far as scenarios III and IV are concerned, It is also observed that HRF reduces for higher values of η. Moreover, for the last scenario similar performance is noted in case of HRF (see scenario V). The outcomes of numerical modeling and ANN model are in very close agreement.

5. Concluding Remarks

The present study has provided evidence for the effectiveness of utilizing Artificial Neural Networks (ANN) in combination with the Unit Gompertz distribution for the purpose of forecasting the attributes of said distribution. All features of the model have been researched and explained in detail via ANN. This study adds to the existing literature in the field of statistical modeling and machine learning, providing new opportunities for further investigation and real-world implementations across several disciplines. As the refinement and expansion of this approach persist, it is expected that there will be additional progress in the field of predictive modeling and data analysis. For the UG model, two different ANN models have been developed to estimate the PDF, CDF, HRF and R values. 22 of the data used in the MLP network model developed with 32 data sets produced were used for training, 5 for validation and 5 for testing. The main findings as a result of this study are as follows:

- Maximum mean occurs near the higher level of parameters b and η while it reduces for small values of b and all levels of η.
- The UG model shows leptokurtic and platykurtic behavior.
- The probability density function of the UG model may exhibit features like decreasing, unimodal, reversed J-shaped, and positively skewed, while its hazard rate function may have a shape that increases initially, followed by a declines, and then reaches a bathtub shape.
- For higher levels of both parameters b and η; median is raised, while it reduced for small values.
- Positively skewed behavior is observed for UG model for considered parametric values.
- The findings demonstrated that ANNs are a highly effective technique for forecasting PDF, CDF, HRF, and R values.
- For the ANN model developed to predict the PDF and CDF values, the R value has been obtained as 0.99874. The network model has been able to predict the PDF value with an average error of -0.17% and the CDF value with an average error of -0.07%.
- For the network model developed for the prediction of HRF and Reliability values, the R value has been calculated as 0.99986. Average error rates have been found to be 0.07% and -0.21%, respectively.
- These findings demonstrate that ANNs are the best instruments for predicting the particular traits of UG models.

Author Contributions: TNS: Writing – review & editing, Writing – original draft, Validation, Formal analysis, Visualization, Validation. **ABÇ:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **SAL:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **TAA:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **NMG:** Writing – review & editing, Writing – original draft, Investigation,

Formal analysis. **ASh:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Data Availability Statement: The data that support the findings of this study are available with in the manuscript.

Acknowledgment

This article has been produced with the financial support of the European Union under the REFRESH – Research Excellence For REgion Sustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition.

References

Abdul Kareem FA, Shariff AM, Ullah S, Garg S, Dreisbach F, Keong LK, Mellon N. Experimental and neural network modeling of partial uptake for a carbon dioxide/methane/water ternary mixture on 13X zeolite. Energy Technol. 2017; 5(8): 1373-1391.

Ali A, Abdulrahman A, Garg S, Maqsood K, Murshid G. Application of artificial neural networks (ANN) for vapor-liquid-solid equilibrium prediction for CH4-CO2 binary mixture. Greenh Gases Sci Technol. 2019; 9(1): 67-78.

Arnold BC, Groeneveld RA. Some properties of the arcsine distribution. J Am Stat Assoc. 1980; 75(369): 173-175.

Bahiraei M, Heshmatian S, Moayedi H. Artificial intelligence in the field of nanofluids: A review on applications and potential future directions. Powder Technol. 2019; 353: 276-301.

Benkhelifa L. The beta generalized Gompertz distribution. Appl Math Model. 2017; 52: 341-357.

Bemmaor AC. Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity. In Research traditions in marketing 1992; 201-229. Springer, Dordrecht.

Bonakdari H, Zaji AH. Open channel junction velocity prediction by using a hybrid self-neuron adjustable artificial neural network. Flow Meas Instrum. 2016; 49: 46-51.

Castellares F, Lemonte AJ. On the Marshall–Olkin extended distributions. Commun Stat Theory. 2016; 45(15): 4537-4555.

Chen Z. Parameter estimation of the Gompertz population. Biom J. 1997; 39(1): 117-124.

Çolak AB. An experimental study on the comparative analysis of the effect of the number of data on the error rates of artificial neural networks. Int J Energy Res. 2021; 45(1): 478-500.

Çolak AB. Experimental study for thermal conductivity of water-based zirconium oxide nanofluid: Developing optimal artificial neural network and proposing new correlation. Int J Energy Res. 2021; 45(2): 2912-2930.

Çolak AB. A novel comparative analysis between the experimental and numeric methods on viscosity of zirconium oxide nanofluid: Developing optimal artificial neural network and new mathematical model. Powder Technol. 2021; 381: 338-351.

Çolak AB, Akçaözoğlu K, Akçaözoğlu S, Beller G. Artificial Intelligence Approach in Predicting the Effect of Elevated Temperature on the Mechanical Properties of PET Aggregate Mortars: An Experimental Study. Arab J Sci Eng. 2021; 46(5): 4867-4881.

Çolak AB, Sindhu TN, Lone SA, Shafiq A, Abushal TA. Reliability study of generalized Rayleigh distribution based on inverse power law using artificial neural network with Bayesian regularization. Tribol Int. 2023; 185: 108544.

Condino F, Domma F. A new distribution function with bounded support: the reflected generalized Topp-Leone power series distribution. Metron. 2017; 75(1): 51-68.

Consul PC, Jain GC. On the log-gamma distribution and its properties. Stat Hefte. 1971; 12(2): 100-106.

Cordeiro GM, Dos Santos Brito R. The beta power distribution. Braz J Probab Stat. 2012; 26(1): 88-112.

Cordeiro GM, Lemonte AJ, Campelo AK. Extended arcsine distribution to proportional data: Properties and applications. Stud Sci Math Hung. 2016; 53(4): 440-466.

Duan J, Li F. Transient heat transfer analysis of phase change material melting in metal foam by experimental study and artificial neural network. J Energy Storage. 2021; 33: 102160.

El-Damcese MA, Mustafa A, El-Desouky BS, Mustafa ME. The odd generalized exponential gompertz. 2015; arXiv preprint arXiv:1507.06400.

El-Gohary A, Alshamrani A, Al-Otaibi AN. The generalized Gompertz distribution. Appl Math Model. 2013; 37(1-2): 13-24.

Esmaeilzadeh F, Teja AS, Bakhtyari A. The thermal conductivity, viscosity, and cloud points of bentonite nanofluids with n-pentadecane as the base fluid. J Mol Liq. 2020; 300: 112307.

Franses PH. Fitting a Gompertz curve. J Oper Res Soc. 1994; 45(1): 109-113.

Ghitany ME, Mazucheli J, Menezes AFB, Alqallaf F. The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval. Commun Stat Theory. 2019; 48(14): 3423-3438.

Gómez-Déniz E, Sordo MA, Calderín-Ojeda E. The Log–Lindley distribution as an alternative to the beta regression model with applications in insurance. Insur Math Econ. 2014; 54: 49-57.

Grassia A. On a family of distributions with argument between 0 and 1 obtained by transformation of the gamma and derived compound distributions. Aust J Stat. 1977; 19(2): 108-114.

Jafari AA, Tahmasebi S, Alizadeh M. The beta-Gompertz distribution. Rev Colomb Estad. 2014; 37(1): 141-158.

Jafari AA, Tahmasebi S. Gompertz-power series distributions. Commun Stat Theory. 2016; 45(13): 3761-3781.

Johnson NL. Systems of frequency curves generated by methods of translation. Biometrika. 1949; 36(1/2): 149-176.

Johnson NL. Systems of frequency curves derived from the first law of Laplace. Trab Estad. 1955; 5(3): 283-291.

Khan MS, King R, Hudson IL. Transmuted kumaraswamy distribution. Stat Transit New Ser. 2016; 17(2): 183-210.

Kujur A, Raza Z, Khan AA, Wechtaisong C. Data Complexity Based Evaluation of the Model Dependence of Brain MRI Images for Classification of Brain Tumor and Alzheimer's Disease. IEEE Access. 2022; 10: 112117-112133.

Kumaraswamy P. A generalized probability density function for double-bounded random processes. J Hydrol. 1980; 46(1-2): 79-88.

Lemonte AJ, Barreto-Souza W, Cordeiro GM. The exponentiated Kumaraswamy distribution and its log-transform. Braz J Probab Stat. 2013; 27(1): 31-53.

Lone SA, Sindhu TN, Anwar S, Hassan MK, Alsahli SA, Abushal TA. On Construction and Estimation of Mixture of Log-Bilal Distributions. Axioms. 2023; 12(3): 309.

Makany R. A theoretical basis for Gompertz's curve. Biom J. 1991; 33(1): 121-128.

Mazucheli J, Menezes AFB, Dey S. Improved maximum-likelihood estimators for the parameters of the unit-gamma distribution. Commun Stat Theory. 2018; 47(15): 3767-3778.

Mazucheli J, Menezes AF, Dey S. Unit-Gompertz distribution with applications. Statistica. 2019; 79(1): 25-43.

Pourdarvish A, Mirmostafaee SMTK, Naderi K. The exponentiated Topp-Leone distribution: Properties and application. J Appl Environ Biol Sci. 2015; 5(7S): 251-256.

Rao BR, Damaraju CV, Rao PFR, Damaraju CV. New better than used and other concepts for a class of life distributions. Biom J. 1992; 34(8): 919-935.

Rostamian SH, Biglari M, Saedodin S, Esfe MH. An inspection of thermal conductivity of CuO-SWCNTs hybrid nanofluid versus temperature and concentration using experimental data, ANN modeling and new correlation. J Mol Liq. 2017; 231: 364-369.

Roy S, Adnan MAS. Wrapped generalized Gompertz distribution: An application to Ornithology. J Biom Biostat. 2012; 3(6): 153-156.

Shafiq A, Çolak AB, Sindhu TN, Lone SA, Alsubie A, Jarad F. Comparative study of artificial neural network versus parametric method in COVID-19 data analysis. Results Phys. 2022; 38: 105613.

Shafiq A, Çolak AB, Sindhu TN, Lone SA, Abushal TA. Modeling and survival exploration of breast carcinoma: A statistical, maximum likelihood estimation, and artificial neural network perspective. Artif Intell Life Sci. 2023; 4: 100082.

Shafiq A, Sindhu TN, Hussain Z, Mazucheli J, Alves B. A flexible probability model for proportion data: Unit Gumbel type-II distribution, development, properties, different method of estimations and applications. Aust J Stat. 2023; 52(2): 116-140.

Sharma D, Chakrabarty TK. On size biased Kumaraswamy distribution. 2016; arXiv preprint arXiv:1609.09278.

Shetty SP, Nayak S, Kumar S, Karanth KV. Thermo-hydraulic performance prediction of a solar air heater with circular perforated absorber plate using Artificial Neural Network. Therm Sci Eng Prog. 2021; 23: 100886.

Sindhu TN, Çolak AB, Lone SA, Shafiq A. Reliability study of generalized exponential distribution based on inverse power law using artificial neural network with Bayesian regularization. Qual Reliab Eng Int. 2023.

Sindhu TN, Shafiq A, Hussain Z. Generalized exponentiated unit Gompertz distribution for modeling arthritic pain relief times data: classical approach to statistical inference. J Biopharm Stat. 2023; 33(2): 155-180.

Tadikamalla PR, Johnson NL. Systems of frequency curves generated by transformations of logistic variables. Biometrika. 1982; 69(2): 461-465.

Topp CW, Leone FC. A family of J-shaped frequency functions. J Am Stat Assoc. 1955; 50(269): 209-219.

Uthansakul P, Anchuen P, Uthansakul M, Khan AA. Estimating and synthesizing QoE based on QoS measurement for improving multimedia services on cellular networks using ANN method. IEEE Trans Netw Serv Manag. 2019; 17(1): 389-402.

Vafaei M, Afrand M, Sina N, Kalbasi R, Sourani F, Teimouri H. Evaluation of thermal conductivity of MgO-MWCNTs/EG hybrid nanofluids based on experimental data by selecting optimal artificial neural networks. Physica E. 2017; 85: 90-96.

Vakili M, Khosrojerdi S, Aghajannezhad P, Yahyaei M. A hybrid artificial neural network-genetic algorithm modeling approach for viscosity estimation of graphene nanoplatelets nanofluid using experimental data. Int Commun Heat Mass Transf. 2017; 82: 40-48.

Van Dorp JR, Kotz S. Modeling income distributions using elevated distributions on a bounded domain. In Distribution models theory 2006; 1-25.

Waller NG, Yonce LJ, Grove WM, Faust D, Lenzenweger MF. A Paul Meehl reader: Essays on the practice of scientific psychology. Routledge; 2013.

Willekens F. Gompertz in context: the Gompertz and related distributions. In Forecasting Mortality in Developed Countries. Springer; 2001: 105-126.

Wu JW, Lee WC. Characterization of the mixtures of Gompertz distributions by conditional expectation of order statistics. Biom J. 1999; 41(3): 371-381.