Jurnal Pendidikan Matematika Vol: 2, No 4, 2025, Page: 1-14

Employing Artificial Intelligence Algorithms to Estimate the Hazard Function of the Inverse Gompertz Distribution with A Practical Application

Qasim Muhammad Jawad Noah Sulaiman*

Department of Mathematics, College of Science, Mohaghegh Ardabili University, Iran.

*Correspondence: Qasim Muhammad Jawad Nuh Sulaiman
Email: asaed1148@gmail.com

Received: 04-06-202
Accepted: 19-07-2025
Published: 26-08-2025

Abstract: *Environmental pollution is one of the most important and serious problems facing humanity today, due to its direct impact on the health of humans and other living organisms. In recent years, an increase in environmental pollution rates has been observed, significantly impacting human health and leading to the emergence of numerous diseases, such as cancer, pneumonia, poisoning, birth defects, and others. Given the importance and seriousness of the issue and its direct impact on human life, this research was conducted to determine the percentage of pollution caused by two of the most important factors in air pollution, CO2. This research was conducted based on the explanatory variables: average temperature, average dew point, average humidity, average wind speed, and the average amount of crude oil used in the refining process. In this research, the risk function of the inverse Gompertz model was estimated using artificial intelligence algorithms, namely the genetic algorithm. These methods were applied to air pollution data obtained from the Central Refineries Company in Baghdad (Dora Refinery), which represents daily measurements of environmental pollution compounds based on time for the period from 2019 to 2025.*

Keywords: *Inverse Gompertz Distribution, Properties, Maximum Likelihood Method, Genetic Algorithm, Risk Function.*

Introduction

The concept of pollution summarizes the various environmental threats to which individuals are exposed and with which they have often become increasingly familiar. Pollution is a state of impurity or uncleanliness, or any process that produces this condition. Pollution is defined as a qualitative or quantitative change in environmental components, provided that this change is outside the natural range of fluctuations of any of the components, resulting in an imbalance in nature that leads to a direct or indirect impact on the ecosystem. Pollution is any change in the natural properties of the environment that causes harm to human life or other organisms. It can also be defined as the addition or introduction of any unfamiliar substance to the environment, resulting in a change in its properties. Air pollution resulting from the use of both types of fuel used to operate generators (gasoline and diesel) and the various combustion products of these fuels... hydrocarbons, sulfur oxides, nitrogen oxides, carbon oxides, particulate matter and heavy elements, especially lead, all of which have various harmful effects on public health, animal and plant life, property and the environment. The most prominent health effects include irritation of the respiratory system, exacerbation of heart disease, allergies, certain eye diseases, and an impact on the physical and mental development of children. The second type of pollution is noise pollution, which negatively affects human health in physical, psychological, and neurological aspects and causes hearing loss, heart disease, atherosclerosis, tumors, immunodeficiency, high blood sugar, and others. Indirect damage includes water pollution by oils and petroleum derivatives leaking from operations, and soil and vegetation pollution resulting from fuel and oil leaks and the disposal of generator maintenance waste.

Numerous classical distributions have been used extensively over the past decades for modelling univariate and bivariate data in areas such as engineering, actuarial, environmental and medical sciences, biological studies, demography, economics, finance, and insurance.

Gompertz and its inverse distributions are of these models. The last two models with their extensions arewidely used in several fields such as reliability, actuarial science, queuing problems, and biological sciences. For example, El-Bassiouny et al. [1–5]. In the statistical literature, there are many distributions have only one parameter, such as exponential, Rayleigh, Lindley, and their inverse and modified versions (univariate and bivariate). These models can be used for modelling lifetime data in various fields because these Models have several desirable properties and nice physical interpretations. For example, Kundu and Raqab [6],Cordeiro et al. [7], Ghitany et al. [8], Merovci [9–11], Elbatal and Elgarhy [12], Adepoju et al. [13], Ahmad et al. [14], Oguntunde et al. [15], Sarhan et al. [16], Oguntunde and Adejumo [17], El-Gohary et al. [18], Merovci and Elbatal [19], Haq [20], El-Morshedy et al. [21–23], Eliwa et al. [24], El-Morshedy and Eliwa [25], Basheer [26], Alizadeh et al. [27], Eliwa and El-Morshedy [28–30], among others.

The research aims to develop a practical framework for estimating the risk function of the Inverse Gompertz distribution and associated environmental pollution data, and to compare the estimation performance between the traditional maximum likelihood method and optimization methods based on the genetic algorithm.

Research Problem

Despite the diversity of age distribution and reliability models, some forms of age/time-to-event data in environmental pollution studies remain difficult to model—especially when a non-uniform risk function (e.g., a rise and then a fall) appears to be present. The Alshenawy distribution (a single-parameter distribution) exhibits properties that make it a good candidate for such data. However, estimating the distribution parameter using traditional numerical methods (Newton–Raphson, e.g.) may encounter difficulties when combined with embedded data or when the likelihood surface is multi-peaked or highly curved. Therefore, we need a more flexible optimization method (the genetic algorithm) and compare its performance with traditional MLE in estimating the parameter and risk function on actual environmental pollution data.

Inverse Gompertz Distribution

Gompertz and its inverse distributions are among the most important models that are widely used in several fields such as reliability, actuarial sciences, environmental problems, and biological sciences. The Inverse Gompertz distribution function for a single sample can be written as follows: [11]

$$f_X(x;\beta) = \frac{1}{x^2}\exp\left(\frac{1}{\beta}\left(1-\exp\left(\frac{\beta}{x}\right)\right)+\frac{\beta}{x}\right),\ \ x>0 \qquad (1)$$

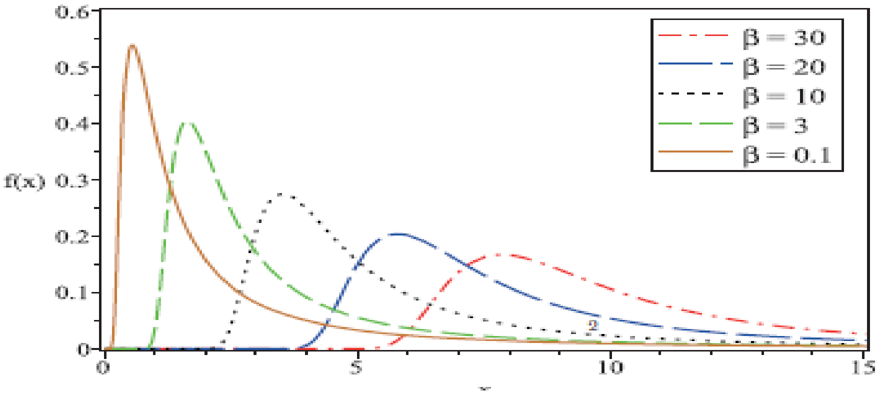

Figure 1. The PDF of A distribution for different values of β.

cumulative distribution function (CDF) is given by:

$$F(x;\varphi) = \exp\left(\frac{1}{\varphi}\left(1-exp\left(\frac{\varphi}{x}\right)\right)\right), x>0, \varphi>0 \quad (2)$$

$where$:

x: represents the ordinal variable (time)

φ: represents the measurement parameter

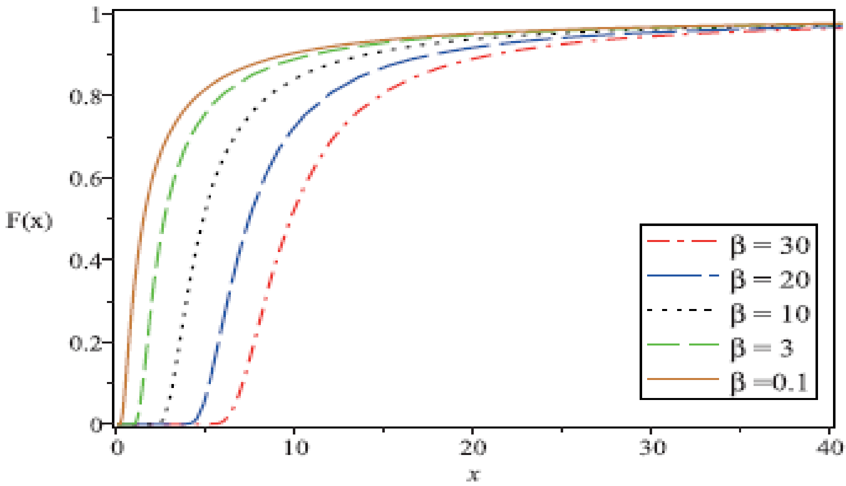

Figure 2. The CDF of A distribution for different values of β.

Inverse Gompertz Properties

Reliability and Reversed (Hazard) Rate Functions

If $X \sim A(\beta)$, then the reliability function of X is given by

$$R_X(x;\beta) = 1 - \exp\left(\frac{1}{\beta}\left(1 - \exp\left(\frac{\beta}{x}\right)\right)\right), \quad x, \beta > 0. \tag{3}$$

The hazard rate (HR) function of X is:

$$h_s(x_i) = \frac{\exp\left(\frac{\varphi}{x}\right)}{x^2 \exp\left(\frac{1}{\varphi}\left(\exp\left(\frac{\varphi}{x}\right) - 1\right)\right)}, x > 0. \tag{4}$$

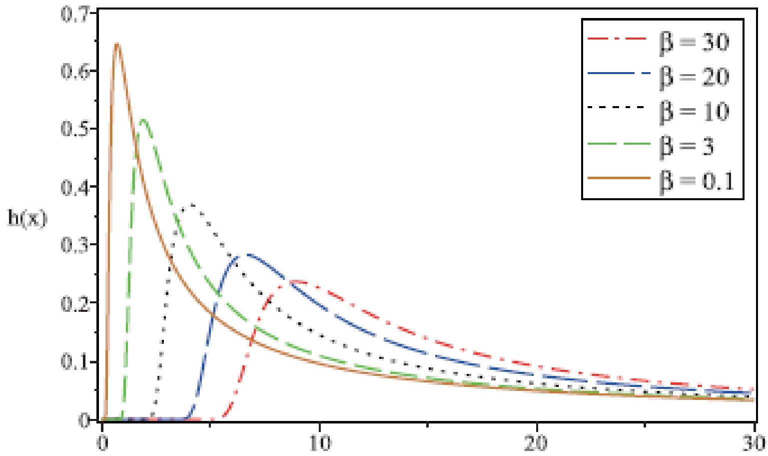

Figure 3. The HR function of the A distribution for different values of β.

Renyi and δ–entropies

If $X \sim A(\beta)$, then the Rényi entropy is defined as

$$\begin{aligned} I_\delta(X) &= \frac{1}{1-\delta} \log \int_0^\infty f_X^\delta(x)\,dx, \quad \delta \in]0, \infty[- \{1\} \\ &= \frac{1}{1-\delta} \log \left(\sum_{k=0}^{\infty} \sum_{m=0}^{k} \right. \\ &\quad \left. \times \frac{(-x_0)^m u^{(k)}}{k!} C_m^k (\delta\beta)^{-\tau} \Gamma(\tau) \right), \quad \tau \notin \{0\} \cup Z^- \end{aligned} \tag{5}$$

Where:

$$u^{(k)} = \frac{d^k}{dx^k}\left(\exp\left(\frac{\delta}{\beta}\left(1 - \exp\left(\frac{\beta}{x}\right)\right) + \frac{2\delta\beta}{x}\right)\right)\Big|_{x=x_0},$$

$$\tau = 2\delta - k + m - 1, \ x_0 \in]0, 0.1[.$$

Equation (5) can be obtained by using Taylor and binomial expansions. The δ–entropy, say $H\delta(X)$, is given as:

$$H_\delta(X) = \frac{1}{\delta - 1} \log\{1 - [(1-\delta) I_\delta(X)]\}.$$

Quantile Function

$$x_q = \frac{\beta}{\ln(1-\beta \ln q)}, \quad 0 < q < 1$$

Mode Function

$$\exp\left(\frac{\beta}{x}\right) - 2x - \beta = 0$$

Median Function

$$x_{0.5} = \frac{\beta}{\ln(1-\beta \ln 0.5)}$$

Parameter Estimation

Maximum Likelihood Method

This method is considered one of the most important estimation methods because of its good properties, including stability, high efficiency, and consistency in some cases. Suppose we have several observations from the distribution with a sample size of n from the Inverse Gompertz distribution. Then, we get the probability function in the following form:

Let us assume $x_1, x_2, \ldots, x_n$represents a random sample of size n from the Inverse Gompertz distribution. Then the probability distribution function is

$$\ell = \prod_{i=1}^{n} f_X(x_i; \varphi) \qquad (6)$$

$$\ell = \prod_{i=1}^{n} \frac{1}{x_i^2} \exp\left(\frac{1}{\varphi}\left(1 - exp\left(\frac{\varphi}{x_i}\right)\right) + \frac{\varphi}{x_i}\right) \qquad (7)$$

$$LL = \varphi \sum_{i=1}^{n} \frac{1}{x_i} - \frac{1}{\varphi} \sum_{i=1}^{n} \left(\exp\left(\frac{\varphi}{x_i}\right) - 1\right) - 2\sum_{i=1}^{n} \ln(x_i) \qquad (8)$$

By differentiating Equation (8) with respect to β, and equating it to zero, we get

$$\sum_{i=1}^{n} \frac{1}{x_i} - \frac{1}{(\beta)^2} \sum_{i=1}^{n} \left(\frac{\beta}{x_i} \exp\left(\frac{\beta}{x_i}\right) - \exp\left(\frac{\beta}{x_i}\right) + 1\right) = 0 \qquad (9)$$

It is not possible to get an explicit solution to Equation (9). So, the solution has to be obtained numerically.

$$r^* = \frac{n!}{(n-k)!} \left(R_X(x_k)\right)^{n-k} \prod_{i=1}^{k} f_X(x_i) \qquad (10)$$

If *X*1, *X*2, ., *Xn* is a random sample from the A distribution, and *X*(1), *X*(2), .. , *X*(*k*), *k* ≤ *n* represent the ordered sample obtained from Type II right-censored data, then the log-likelihood function L^* is

$$L^* = \ln\left(\frac{n!}{(n-k)!}\right) + (n-k) \times \ln\left(1 - \exp\left(\frac{1}{\beta}\left(1 - \exp\left(\frac{\beta}{x_k}\right)\right)\right)\right) + \beta\sum_{i=1}^{k}\frac{1}{x_i} - \frac{1}{\beta}\sum_{i=1}^{k}\left(\exp\left(\frac{\beta}{x_i}\right) - 1\right) - 2\sum_{i=1}^{k}\ln x_i. \quad (11)$$

Differentiating Equation (11) with respect to β, we get

$$\frac{\partial}{\partial\beta}L^* = \frac{(n-k)}{1 - \exp\left(\frac{1}{\beta}\left(1 - \exp\left(\frac{\beta}{x_k}\right)\right)\right)}\frac{\partial}{\partial\beta} \times \left(1 - \exp\left(\frac{1}{\beta}\left(1 - \exp\left(\frac{\beta}{x_k}\right)\right)\right)\right) + \sum_{i=1}^{k}\frac{1}{x_i} + \frac{1}{\beta^2}\sum_{i=1}^{k}\left(\exp\left(\frac{\beta}{x_i}\right) - 1\right) - \frac{1}{\beta}\sum_{i=1}^{k}\frac{1}{x_i}\exp\left(\frac{\beta}{x_i}\right). \quad (12)$$

The solution has to be obtained numerically.

Genetic Algorithm Method

A genetic algorithm is a random search method that addresses a problem to reach the best possible results. It revolves around evolutionary techniques based on Dar's theory of evolution, which states that the fittest survive by imitating the work of nature by preserving the good qualities present in the parents' generation and transferring them to the offspring's generation. Its goal is to obtain immediate offspring that have the best qualities of the parents at the very least.

Genetic Algorithm Stages

The genetic algorithm differs according to the different branches of development techniques, but they share the following stages:

1. The beginning: It is represented by a random population of chromosomes (search space) and in other words, it is a set of solutions to the problem
2. Initialization: It is the process of creating the initial generation that includes generating chromosomes randomly according to the size of the community and the nature of the problem
3. The objective function and the evaluation function (comparison): Through it, the chromosomes are evaluated so that each chromosome gives a specific value that represents the extent of efficiency to the extent of its proximity to the solution and the objective function is either maximum or minimum and the evaluation function depends on it
4. Conclusion: The process of generating a new generation of individuals (chromosomes) that are selected through the selection process according to the

principle of survival of the fittest and then performing the hybridization process and the mutation process to produce the children of the next generation.

5. New society: It includes generating a new generation by repeating the following stages until the generation is completed and the generation is represented by the following:
    a. Selection: It is the process of selecting suitable chromosomes from the old generation to form the new generation. It is created according to the values of the evaluation function to have chromosomes with the highest value of the evaluation function in the new generation
    b. Hybridization: Good chromosomes are selected from the first generation. Mating takes place between two chromosomes to form a new generation (children) depending on the chromosomes (mother) by taking the good characteristics from them from its methods of hybridization with one point and two points and others.
    c. Mutation: After creating the new generation (children) and with the possibility of the mutation, the mutation process takes place with random periods in its chromosomal dye, which leads us to maintain the good qualities between the genes in one chromosome and reach the solution faster, and in it, the exchange occurs between the chromosome, and when there is no mutation, the chromosomes (parents) are duplicated directly without the hybridization process.
6. Testing: We test the solution by providing the stopping condition or not. When it is available, the genetic algorithm stops, and the new solution benefits from the formation of the last generation.
7. Stopping criterion: Generations continue to be formed successively to improve the solution examples until the stopping condition is achieved, which depends on the genetic algorithm stopping scale (the optimal solution), and this scale varies according to the problem to be addressed.
8. Termination: The genetic algorithm ends when one of the following factors is present:
    a. Finding the optimal solution
    b. Reaching the required number of generations
    c. Providing a specific value such as calculating the cost of production
    d. Falling into the local minimum value and cannot be exited [9]

Applying the stages of the genetic algorithm in the binary Inverse Gompertz distribution:

We apply the stages of the genetic algorithm in the objective function equation for each method to find estimates of the parameters of the binary Inverse Gompertz distribution according to the following:

Initialization: Forming the chromosome through the values of $\beta_p$that form the genes of the chromosome and that (P=0,1,2,...,p) are within the real numbers.

Initialization: Creating the initial generation by finding an initial value for the genes with the random values of the other set of constraints.

Objective function: The chromosome is evaluated in terms of efficiency to reach the optimal solution by determining the value of β_p Perform the testing process for the chromosome that has a small objective function value by choosing the highest probability for it and finding the evaluation function for it through the following equation

$$Fitness\ Function = \frac{1}{1 + objective\ function}$$

Through the evaluation function formula, we can find the probability of this function (the best values) according to the following mathematical formula:

$$C_{(i)} = \frac{f(i)}{\sum_{i=1}^{n} f(i)} \qquad (13)$$

Since:

$C_{(i)}$:represents the probability of individual i

$f(i)$:Evaluation function of individual i

n: represents the size of observations

In this step, chromosomes that are good in their characteristics are hybridized by mating between each two chromosomes, and one of its criteria is applied, which is organized hybridization based on the probability of hybridization Pc, and this value is compared with the value of the genes for the two chromosomes (parents) to form the new generation of children and the exchange occurs when the value of the gene is greater or equal to the probability value.

The last step that the chromosomes can go through is the mutation process and it also depends on the probability value (Pm) for the parameters of replacing randomly selected genes with a new value that we also obtained randomly according to the following formula: Total genes = Number of genes in the chromosome * Number of population

The Practical Side

Environmental Pollution

Environmental pollution is defined as the occurrence of changes in the ecosystem in a way that leads to the inability of the ecosystem to perform its natural role in getting rid of pollutants, especially organic ones, through natural processes, which leads to a disruption in the ecosystem. In a more precise concept, the meaning of pollution is everything that contributes to changing any element of the environment, whether this element is a living organism such as humans, animals, and plants, or a natural non-living component such as air, water, soil, radiation, etc. The following is a brief presentation of the most important sources of air pollution and the health and environmental damage resulting from these pollutants.

Sources of Air Pollutants

In general, air pollution sources are of two types: industrial and natural sources. Industrial sources are either mobile or fixed sources, including power generation stations, cement and metal factories, and oil and natural gas industry facilities. These sources contribute to increasing the risk of pollution to varying degrees. For example, the oil industry leads to air pollution with sulfur oxides, nitrogen, ammonia, carbon dioxide, hydrogen sulfide, and other gases. Some volatile organic compounds are also released into the air and the surrounding environment at the levels of petroleum and dye factories, etc. Therefore, this research seeks to prove the validity of this common idea by examining pollution levels concerning the volume of oil production specifically through data provided by the Middle Refineries Company.

Stages of Conducting the Practical Side

The first stage: Data collection method

The research included obtaining climate data from the website (weather history for KQTZ) on the Internet, as it is considered one of the best sites that provides historical weather records easily for all countries according to the zip code of each region (city).

As for air pollutant data, it was obtained from the Central Refineries Company, which represents daily measurements of environmental pollution compounds based on the period from September 2019 to December 2025. The approval of the Central Refineries Company (Dura Refinery) was obtained. Thus, this data set tracks the elements causing air pollution through repeated measurements of compounds such as carbon compounds CO_Xsulfur compounds $SO_X$and nitrogen compounds NO_X, in addition to explanatory variables including wind speed, temperatures, and quantities of oil produced, noting that data measurements were taken for the recorded climate factors at the same time in which each pollutant was measured in the refinery

Second Stage: Data Description

Given the data background, the data set consisted of (9) distinctive features, as they were two categorical (date, time) and the response variables were:

Carbon dioxide gas (CO_2): This gas is colorless and odorless and is generated as a result of emissions from the oil refining and production process. This gas can interfere with the blood and hinder its ability to absorb and carry oxygen gas which sometimes leads to death.

And five explanatory variables: average hourly temperatures in (C^0), average dew point, average hourly humidity in degrees Celsius, average wind speed in (km/h), and average amount of crude oil used in the refining process (m^3/h).

Third Stage: Data Coding

Since the pollution data level was measured for more than two years (28), the number of measurements is relatively small to determine complex paths in the response variables. Accordingly, a logit model was assumed for the sample, according to which the data were converted into binary responses to predict the pollution results. The data were processed based on the maximum permissible limits for air pollutants emitted from combustion sources. The minimum and equal values to the maximum limits are considered non-polluted [0] and the values higher than the maximum limits are considered polluted [1]. The maximum permissible limits for the concentration of each pollutant that is allowed to be released according to the national standards are that CO gas is 500 ug/$\mathrm{m}^3$and SO_2gas is 10 ug/m^3

Stage Four: Hypothesis testing of the general model

The significance of the corresponding parameters is often tested to accept or reject the null hypothesis of less than 0.05, which states that

H_0:The increase in the pollution rate does not depend on an increase in climatic conditions and oil production

H_1: The increase in the pollution rate depends on an increase in climatic conditions and oil production

Stage Five: Data Analysis

The results of the practical side will be presented and analyzed to reach the extent of the suitability of the real data with the Inverse Gompertz model

estimated by conducting tests for the capabilities of the risk function for the Inverse Gompertz model. The following will display the results in the tables that will be analyzed according to the tables as follows:

Table 1: Represents the hazard function estimators and standard error of all explanatory variables for the binary Inverse Gompertz model using the Maximum likelihood method

Estimated parameter	$\hat{h}$(ML	**Standard error SE($\hat{\beta}_i$)**	$\frac{\hat{\beta}}{SE(\hat{\beta}_i)}$	**significations**
Average temperature	4.422	3.471	1.274	**N0N-Sig**
Average *dew poi*	0.005	0.013	0.391	**N0N-Sig**
Average humidity	0.003	0.002	1.310	**N0N-Sig**
Average wind speed	0.137	0.146	2.943	**sig**
amount of crude oil	0.383	0.192	1.999	**Sig**

We note in Table No. (1) the estimates of the risk function and the standard error values for each estimated parameter and the results obtained, which follow the normal distribution with a significance level of (α=0.05) and are compared with the table value of $Z_{\frac{1}{2}(1-0.05)} = \mp 1.96$ and the last column represents the significance of the explanatory variables (less than 0.05) under test is significant. Accordingly, the variables that do not significantly affect the estimated model were removed, i.e., the variables (average temperature, average dew point, and average humidity) have no effect. In contrast, average wind speed and average quantity of oil produced have a significant effect on the gas (CO_2) on its dispersion in large amounts in the air. As for that, the strength of the relationship between the variables through calculating the R^2statistic showed a value equal to 0.62, which means that the average wind speed and average quantity of oil produced explain the differences in the observed responses for the cases of (CO_2) by 62%.

Table 2: Represents the hazard function estimators and standard error of all explanatory variables for the binary Inverse Gompertz model using the genetic algorithm method

Estimated parameter	$\hat{h}$(G.	**Standard error** SE($\hat{\beta}_i$)	$\frac{\hat{\beta}}{SE(\hat{\beta}_i)}$	**significations**
Average temperature	6.942	2.315	3.000	**Sig**
Average *dew poi*	0.021	0.0513	0.581	**N0N-Sig**
Average humidity	0.213	0.0122	1.210	**N0N-Sig**
Average wind speed	0.438	0.584	4.000	**Sig**
amount of crude oil	0.489	0.882	2.879	**Sig**

We note in Table No. (2) the estimates of the risk function and the standard error values for each estimated parameter and the results obtained, which follow the normal distribution with a significance level of (α=0.05) and are compared with the table value of $Z_{\frac{1}{2}(1-0.05)} = \mp 1.96$ and the last column represents the significance of the explanatory variables (less than 0.05) under test is significant. Accordingly, the variables that do not significantly affect the estimated model were removed, i.e., the variables (average dew point and average humidity) have no effect. In contrast, average temperature, average wind speed, and average amount of oil produced have a significant effect on the gas (CO2) on its dispersion in large quantities in the air. As for that, the strength of the relationship between the variables through calculating the R^2 statistic showed a value equal to 0.80, which means that the average temperature, average wind speed, and average amount of oil produced explain the differences in the observed responses for the cases of (CO2) by 80%.

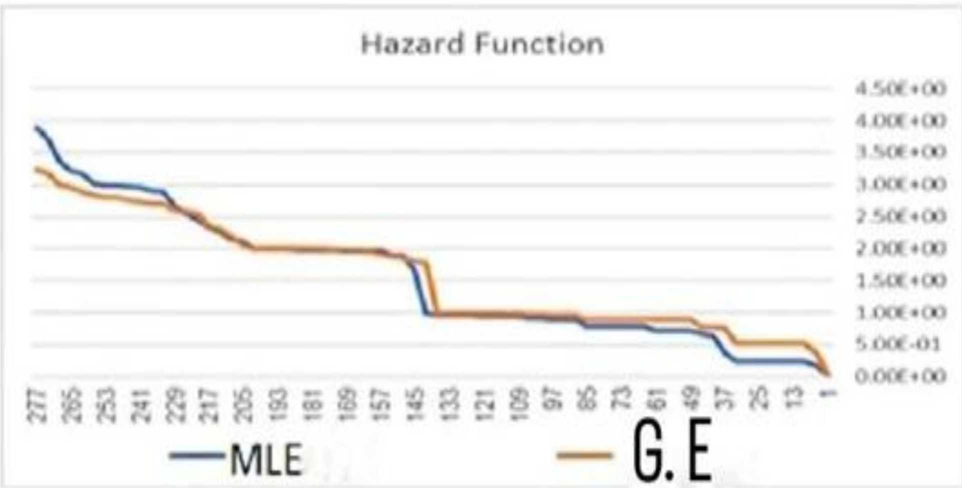

Figure 4. represents the estimation of the risk function using the maximum likelihood method and the genetic algorithm.

We note from Figure (4) that the behavior of the risk function estimates for the two methods appears increasingly for the Inverse Gompertz model.

Conclusions

1. For the Inverse Gompertz model using the genetic algorithm and maximum likelihood method, it was shown through the Z-value test that the factors that had a significant effect of less than 0.05 on the CO2 variable were the average wind speed and the average quantity of oil produced.
2. The average dew point and average humidity factors had no significant effect on the response variable to CO2 for the Inverse Gompertz model using both methods.
3. The average temperature factor had a significant effect on the CO2 variable for the Inverse Gompertz model using the genetic algorithm method.

Recommendations and Future Studies

1. Using the methods presented to estimate the risk function, the inverse Gompertz model can be used to validate data across environmental pollution layers such as soil, water, and radioactive contamination.
2. Strategies for future studies that rely on the conventional estimation methods and genetic algorithms mentioned in the research, and comparing them in the event of outliers in the data.
3. Using types other than those mentioned in the research, such as types that include nitrogen oxides, ammonia, carbon monoxide, and ozone.
4. Adding data to facilitate the development of an optimal prediction model, suggesting that daily data for a period of 10 to 20 years may be used for a more specific model.
5. The Midland Refineries Project should also begin improving data collection processes by continuing to collect data on potential hazards in different climate conditions, such as waste and solids generated through refining processes, to facilitate air path determination in similar future studies.
6. One of the main studies is to examine the ongoing levels of carbon emissions from oil refinery operations relative to the volume of oil produced.
7. Addressing this topic will facilitate the identification of potential carbon pollution from oil refineries, given that environmental variables also perpetuate pollution and stimulate interactions between pollutants.

Reference

Adepoju KA, Chukwu AU, Wang M. The beta power exponential distribution. J Stat Sci Appl. 2014;2:37–46.

Ahmad A, Ahmad SP, Ahmed A. Transmuted inverse Rayleigh distribution: a generalization of the inverse Rayleigh distribution. Math Theory Model. 2014;4(7).

Alizadeh M, Afify AZ, Eliwa MS, Ali S. The odd log-logistic Lindley-G family of distributions: properties, Bayesian and non-Bayesian estimation with applications. Comput Stat. 2019. doi:10.1007/s00180-019-00932-9.

Basheer AM. Alpha power inverse Weibull distribution with reliability application. J Taibah Univ Sci. 2019;13(1):423–432.

Cordeiro GM, Cristino CT, Hashimoto EM, Ortega EM.The beta generalized Rayleigh distribution. Stat Pap. 2013;54:133–161.

E. Demir , O¨. Akkus , " An Introductory Study on How the Genetic Algorithm Works in the Parameter Estimation of Binary Logit Model", IJS:BAR, pp.162-180 , 2015.

Elbatal I, Elgarhy M. Statistical properties of Kumaraswamy quasi Lindley distribution. Int J Math Trends Technol. 2013;4:237–246.

El-Gohary A, El-Bassiouny AH, El-Morshedy M. Inverse flexible Weibull extension distribution. Int J Comput Appl. 2015;115:46–51.

Eliwa MS, El-Morshedy M, Afify AZ. The odd Chen generator of distributions: properties and estimation methods with applications in medicine and engineering. J Natl Sci Found Sri Lanka. 2019. To appear.

Eliwa MS, El-Morshedy M. Bivariate Gumbel-G family of distributions: statistical properties, Bayesian and non-Bayesian estimation with application. Ann Data Sci. 2018. doi:10.1007/s40745-018-00190-4.

Eliwa MS, El-Morshedy M. Discrete flexible distribution for over-dispersed data: statistical and reliability properties with estimation approaches and applications. J Appl Stat. 2020b. To appear.

El-Morshedy M, El-Bassiouny AH, El-Gohary A. Exponentiated inverse flexible Weibull extension distribution. J Stat Appl Prob. 2017;6(1):169–183.

El-Morshedy M, Eliwa MS, El-Gohary A, Khalil AA. Bivariate exponentiated discrete Weibull distribution: statisticalproperties, estimation, simulation and applications. Math Sci. 2019b. doi:10.1007/s40096-019-00313-9

El-Morshedy M, Eliwa MS, Nagy H. A new two-parameter exponentiated discrete Lindley distribution: properties, estimation and applications. J Appl Stat. 2019a. doi:10.1080/02664763.2019.1638893.

El-Morshedy M, Eliwa MS. The odd flexible Weibull-H family of distributions: properties and estimation with applications to complete and upper record data. Filomat. 2019;33(9). To appear.

Ghitany ME, Al-Mutairi DK, Balakrishnan N, Al-Enezi LJ.Power Lindley distribution and associated inference.Comput Stat Data Anal. 2013;64:20–33.

Haq MA. Kumaraswamy exponentiated inverse Rayleigh distribution. Math Theory Model. 2016;6(3):93–104.

I. Yousef ,"Some methods for estimating the conditional logistic regression model in the case of longitudinal data and their application in environmental pollution " A master's thesis in statistics submitted to the College of Administration and Economics at the University of Baghdad , 2017.

Kundu D, Raqab MZ. Generalized Rayleigh distribution different methods of estimation Comput Stat Data Anal. 2005;49:187–200.

M.S. Ahmed , "Forecasting of Environmental Pollution Using Box-Genghis Models for Al-Waziriya Station" Research Diploma in Statistics presented to the College of Administration and Economics at the University of Baghdad , 2015.

Merovci F, Elbatal I. Weibull-Rayleigh distribution: theory and applications. Appl Math Inf Sci. 2015;9(4):2127–2137.

Merovci F. Transmuted exponentiated exponential distribution.Math Sci Appl E-Notes. 2013a;1(2):112–122.

Merovci F. Transmuted Rayleigh distribution. Austrian J Stat. 2013c;42(1):21–31.

Oguntunde PE, Adejumo AO. Transmuted inverse exponential distribution. Int J Adv Stat Prob. 2015;3(1):1–7.

Oguntunde PE, Babatunde OS, Ogunmola AO. Theoretical analysis of the Kumaraswamy inverse exponential distribution. Int J Stat Appl. 2014;4(2):113–116.

R. Alshenawy (2020) A new one parameter distribution: properties and estimation with applications to complete and type II censored data, Journal of Taibah University for Science, 14:1, 11-18, DOI: 10.1080/16583655.2019.1698276 \
S.A Al-Rudaini, "The use of the genetic algorithm in estimating the parameters of the binary logistic regression model with practical application" A master's thesis in statistics submitted to the College of Administration and Economics at the University of Baghdad ,2019
S.Sefian , M. Benbouziane ,"Portfolio Selection Using Genetic Algorithm" , MPRA , pp. 1-12 , 2012 .
Sarhan M, Lotfi T, Hamilton DC. A new lifetime distribution and its power transformation. J Probab Stat. 2014. article ID 532024.