Ibn Al-Haitham Journal for Pure and Applied Sciences
Journal homepage: jih.uobaghdad.edu.iq
PISSN: 1609-4042, EISSN: 2521-3407
IHJPAS. 2026, 39(1)

A New Artificial Intelligence Algorithm to Estimate the Survival Function of the Modified Log-Logistic Distribution

Mohammed A.R. Hayoun [1], **Bayda Atiya Kalaf** [2*] **and Erum Rehman** [3]
[1,2] Department of Mathematics, College of Education for Pure Sciences (Ibn Al-Haitham), University of Baghdad, Baghdad, Iraq
[3] Department of Mathematics, School of Social Sciences, Nazarbayev University, Astana, Kazakhstan
*Corresponding Author

Received: 4/May/2025
Accepted: 3/August/2025
Published: 20/January/2026
doi.org/10.30526/39.1.4167

Abstract

Survival analysis is a fundamental statistical tool for studying the period leading up to a specific event, such as patient death or device failure. Interest in this analysis has grown due to its ability to provide accurate insights into survival rates and survival function estimation in various fields. This paper modifies the Log-logistic Distribution (MLLGD) by adding a shape parameter to enhance its ability to predict more effectively. After that, a new hybrid meta-heuristic algorithm (PSWGO), combining particle swarm optimization with the gray wolf algorithm, is proposed to estimate the survival function. To compare and investigate the performance of the proposed algorithm with three meta-heuristic algorithms (particle swarm optimization, the gray wolf algorithm, and genetic algorithm), and different sample sizes (n = 25, 50, 75, and 100), a simulation study is conducted based on the mean square error (MSE). to obtain the numerical results, MATLAB version 2022a is used. The results showed that the proposed method (PSWGO) provides an accurate and satisfactory estimate for the Survival function, as it has a less mean-square error than the other estimation methods in most cases. MATLAB version 2020a is used to obtain the numerical results

Keywords: Survival function, particle swarm optimization, gray wolf algorithm, mean square error, simulation.

1.Introduction

In recent decades, the world has witnessed a significant qualitative leap in statistical data analysis, especially survival analysis. This type of analysis is a fundamental tool of applied statistics that can be used to examine the time until the occurrence of a particular event, such as the death of a patient in medical studies or the failure of a device in engineering studies[1,2]. Despite its great importance, survival analysis faces several challenges, particularly the complexity of data modeling, especially when the data is non-linear or truncated. These challenges are most evident when using traditional models based on a particular distribution, such as the exponential or Weibull distribution. With the increasing complexity of modern data, the need has arisen to develop more flexible and suitable models that can handle truncated or heterogeneous data[3].

Among the distributions used in survival analysis, the logistic distribution is a practical choice due to its flexibility in dealing with nonlinear data[4]. This distribution enables the model to accommodate various types of data and is therefore more suitable for data centered around a specific value[5].

On the other hand, adding a shape parameter to probability distributions gives them significant flexibility in representing real-world data. This enables the distribution to adapt to data patterns, such as symmetric, skewed, or peaked distributions. This enables distributions to describe complex statistical phenomena more accurately than elemental distributions that lack this parameter[6]. Therefore, the log-logistic distribution was developed by adding a location parameter to achieve this goal. This allows the model to switch from zero centering to centering around a specific location parameter, thereby improving model accuracy and facilitating more straightforward interpretation of statistical results.
However, estimating parameters in nonlinear models is a complex process due to the nonlinear relationship between the variables, so conventional methods are insufficient to solve this problem. Statistical analysis of these estimates requires advanced techniques such as numerical methods and modern optimization algorithms to ensure accurate estimates[7]. Therefore, using meta-heuristic optimization algorithms helps overcome these challenges by providing stochastic search capabilities and efficient solution space exploration. Meta-heuristic algorithms encompass various methods that aim to enhance the survival function's estimation by refining the input parameters' accuracy[8].
The Jackknife approach, which evaluated the Gumbel distribution's parameters, was I troduced[9] metaheuristic method (harmony search algorithm) was used a to estimate the proper probability distribution parameters[10]. Genetic technique (GA) was utilized for estimating the parameters of the mixture's normal distribution by maximizing the log likelihood outcome[11]. For more details, see [12-15]. Additionally, the hybrid approach has several advantages in estimating the survival function. Hybrid algorithms strike an effective balance between exploration and exploitation, enhancing survival function estimation accuracy. Experimental studies have demonstrated that the hybrid approach often outperforms traditional methods and even single algorithms, particularly with complex and variable data[16-19]
The development of hybrid algorithms for survival analysis is a crucial step towards overcoming the challenges posed by the complexity of modern data. Combining different meta-inferential algorithms can obtain more accurate and reliable survival function estimates, contributing to better scientific decisions in medicine and engineering. This approach holds great promise for advancing survival analysis, offering speed, efficiency, and flexibility.

2.Methodology and Methods

In this section, we will be introduced how to modify the Log-logistic distribution and propose the hybrid algorithm:

2.1. Modified Log-logistic distribution

Log-logistic distribution is the model of a transformed well known logistic variate[20]. Then, the properties of the Log-logistic distribution were studied and constructed its order statistics[21]. The constructed the acceptance sampling plans for the Log-logistic distribution[22,23]. The probability density function (PDF) of the standard Log-logistic distribution is given as:

$$g(x;\alpha,\beta) = \frac{\beta}{\alpha}\cdot\frac{\left(\frac{x}{\alpha}\right)^{\beta-1}}{\left(1+\left(\frac{z}{\alpha}\right)^{\beta}\right)^{2,}} \qquad x>0 \quad \alpha,\beta>0 \tag{1}$$

When a location parameter μ is added, the distribution is shifted such that it is now centered around $x = \mu$ instead of $x = 0$. Then the PDF and cumulative distribution function (CDF) with a location parameter becomes:

$$f(x;\alpha,\beta,\mu) = \frac{\beta}{\alpha}\cdot\frac{\left(\frac{x-\mu}{\alpha}\right)^{\beta-1}}{\left(1+\left(\frac{x-\mu}{\alpha}\right)^{\beta}\right)^{2,}} \qquad x>\mu,\ \ x>0 \quad \alpha,\beta>0 \tag{2}$$

$$F(x;\alpha,\beta,\mu) = \frac{1}{1+(\frac{x-\mu}{\alpha})^{-\beta,}} \qquad x>\mu,\ \ x>0 \quad \alpha,\beta>0 \tag{3}$$

Thus, the reliability (R) and hazard (h)functions becomes as: $\mathrm{R(x)} = 1 - \mathrm{F(x)}$,

$$\mathrm{R}(x;\alpha,\beta,\mu) = 1 - \frac{1}{1+\left(\frac{x-\mu}{\alpha}\right)^{-\beta}} = \frac{\left(\frac{x-\mu}{\alpha}\right)^{-\beta}}{1+\left(\frac{x-\mu}{\alpha}\right)^{-\beta}}, \qquad (4)$$

And

$$h(x;\alpha,\beta,\mu) = \frac{\frac{\beta}{\alpha}.\frac{\left(\frac{x-\mu}{\alpha}\right)^{\beta-1}}{\left(1+\left(\frac{x-\mu}{\alpha}\right)^{\beta}\right)^{2}},}{\frac{\left(\frac{x-\mu}{\alpha}\right)^{-\beta}}{1+\left(\frac{x-\mu}{\alpha}\right)^{-\beta}},}. \qquad (5)$$

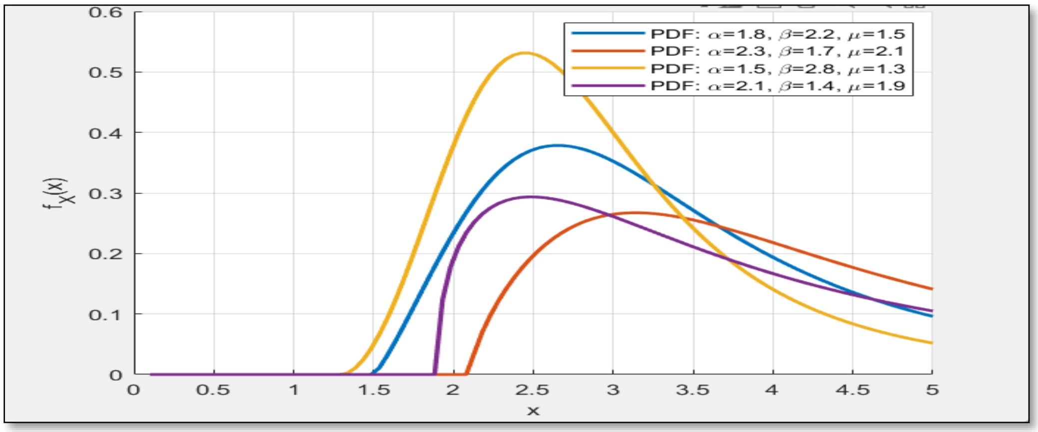

Figure 1. Probability density function for Modified Log-logistic distribution

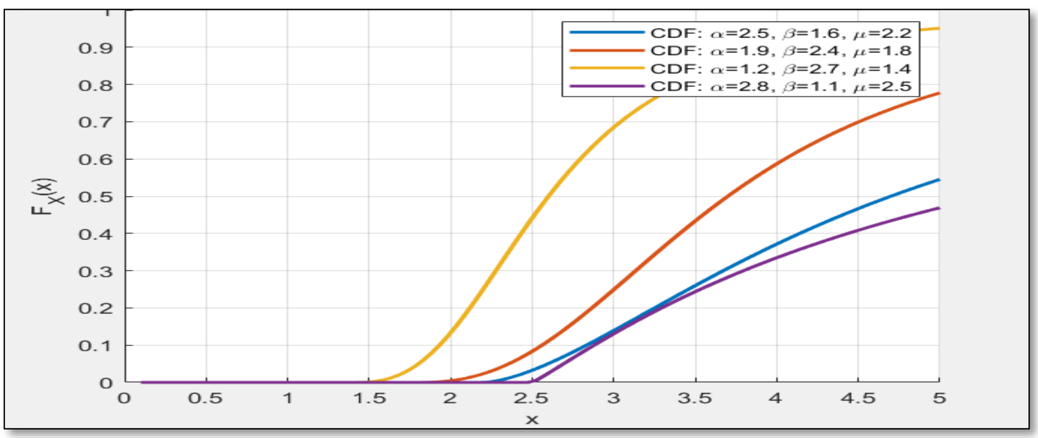

Figure 2: Cumulative distribution function for Modified Log-logistic distribution

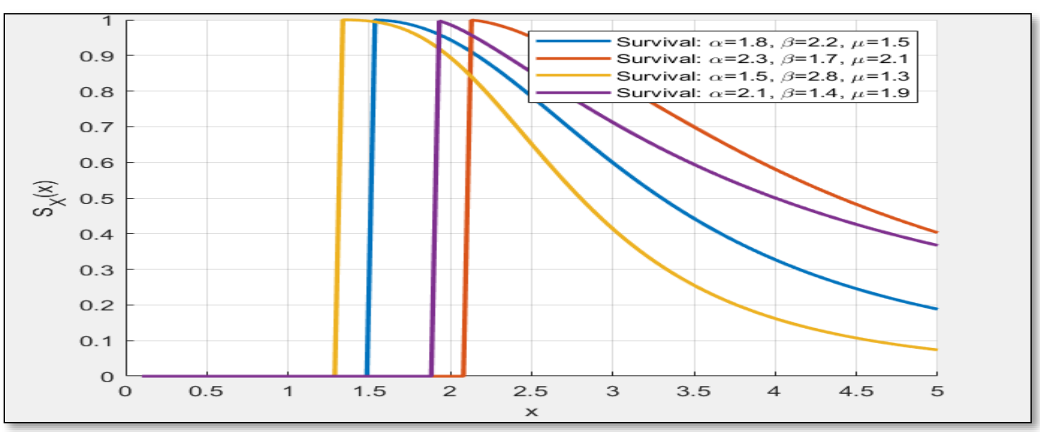

Figure 3. Survival function for the Modified Log-logistic distribution

Some properties of the modified Log-Logistic distribution with location parameter:

- Mode**:** To find the mode of the Log-Logistic distribution.

$\frac{df(x)}{dx} = 0$, from **Equation 2,** we get:

$$\frac{df(x)}{dx} = \frac{\left(\frac{\beta}{\alpha^2}\right)(\beta-1)\left(1+\left(\frac{x-\mu}{\alpha}\right)^{\beta}\right)^2\left(\frac{x-\mu}{\alpha}\right)^{\beta-2} - 2\left(\frac{\beta^2}{\alpha^2}\right)\left(\frac{x-\mu}{\alpha}\right)^{2(\beta-1)}\left(1+\left(\frac{x-\mu}{\alpha}\right)^{\beta}\right)}{\left(1+\left(\frac{x-\mu}{\alpha}\right)^{\beta}\right)^4}$$

$$\left(\frac{\beta}{\alpha^2}\right)\left(1+\left(\frac{x-\mu}{\alpha}\right)^{\beta}\right)\left(\frac{x-\mu}{\alpha}\right)^{\beta-1}\left(\left(1+\left(\frac{x-\mu}{\alpha}\right)^{\beta}\right)(\beta-1)\left(\frac{x-\mu}{\alpha}\right)^{-1} - 2\beta\left(\frac{x-\mu}{\alpha}\right)^{\beta-1}\right) = 0$$

$$x_{mod\,e} = \alpha\left(\frac{\beta-1}{\beta+1}\right)^{\frac{1}{\beta}} + \mu$$

- Median**:** The median of a random variable X of the p.d.f for the modified Log-Logistic distribution can be obtained as follows:

$$F(x;\alpha,\beta,\mu) = \frac{1}{2}$$

$$1+\left(\frac{x-\mu}{\alpha}\right)^{-\beta} = 2 \;\rightarrow\; \frac{1}{\left(\frac{x-\mu}{\alpha}\right)^{\beta}} = 1 \;\rightarrow\; \frac{x-\mu}{\alpha} = 1^{\frac{1}{\beta}} = 1 \;\rightarrow\; x_{median} = \alpha + \mu$$

- Mean

$$E(x) = \int_0^{\infty} x f(x;\alpha,\beta,\mu)dx$$

$$E(x) = \int_0^{\infty} x \frac{\frac{\beta}{\alpha}\left(\frac{x-\mu}{\alpha}\right)^{\beta-1}}{\left(1+\left(\frac{x-\mu}{\alpha}\right)^{\beta}\right)^2} dx$$

$$E(x) = \alpha B\left(1+\frac{1}{\beta}, 1-\frac{1}{\beta}\right) + \mu$$

- Moment-generating function

$$E(e^{tx}) = \int_0^{\infty} e^{tx} f(x;\alpha,\beta,\mu)dx$$

$$= \sum_{n=0}^{\infty} \frac{(\alpha t)^n}{n!} B\left(1+\frac{n}{\beta};\, 1-\frac{n}{\beta}\right)$$

2.2. Hybrid Meta Heuristic Algorithm

The performance of optimization algorithms depends on the type of problems they face. Each algorithm has its advantages and disadvantages that make it effective in some cases and less efficient in others. Their performance may even vary within the same problem class. According to the "No Free Lunch Theorem" proposed[24], no algorithm is better than all others in all scenarios. Based on this theoretical concept, it becomes necessary to explore new strategies that combine the advantages of multiple algorithms to achieve a more balanced and efficient performance. In this context, we have developed a new hybrid algorithm that integrates the particle swarm algorithm (PSO) and the gray wolf algorithm (GWO)[25,26], utilizing both exploration and exploitation elements in a complementary manner. The PSO algorithm begins with a rapid, random search for initial solutions, ensuring broad coverage of the solution space. The GWO algorithm then optimizes these solutions accurately, utilizing the best values found. The results are combined with carefully calculated weights for a more efficient hybrid solution. To improve the accuracy of the final solution, the PSO algorithm is applied again in the final phase. The efficiency of the hybrid algorithm was tested by calculating the survival function and graphing the results. Comparisons with traditional algorithms have shown that the hybrid algorithm outperforms the algorithm when it comes to arriving at optimal solutions quickly and accurately, which improves its feasibility in advanced optimization applications. The hybrid algorithm is built according to the following steps:

Step 1: Exploration using PSO: The algorithm begins by randomly initializing a set of particles within the search space. The initial position and velocity of each particle are determined. The particles move according to the PSO equations to explore various solutions, ensuring that initial solutions are thoroughly examined.
Step 2: Optimization using GWO: The best solutions are identified after an initial round of random search using PSO. The algorithm moves to the GWO optimization phase, where solutions are clustered around the best wolves (Alpha, Beta, Delta). The positions are adjusted based on the gray wolf's movement toward the leader (Alpha), leading to progressively improved solutions.
Step 3: Iterative Updating: The hybridization steps are repeated for a specified number of iterations or until a stopping condition is reached. The locations are updated in each round based on the PSO and GWO rules. PSO benefits from local optimization, while GWO benefits from focusing on the optimal solution.

3. Simulation study

The simulation process was conducted according to specific steps that ensured the highest possible accuracy, based on the Mean Squared Error (MSE) criteria. The simulation was performed by generating each case using 1,000 iterations to ensure the reliability and accuracy of the results. Different sample sizes were tested to investigate the method's performance, including 25, 50, 75, and 100. This variation in sample size aims to determine the extent to which estimation accuracy is affected by data overload, as the method's efficiency can vary depending on the amount of available information.
Step1: Initialize all the parameters of PSO,GWO, GA , and PSGWO algorithms then used the objective function $nLn(\beta) - nLn(\alpha) + \beta\sum_{i=1}^{n} Ln(x_i + \mu) - \beta\sum_{i=1}^{n} Ln(\alpha) - \sum_{i=1}^{n} Ln(x_i + \mu) + \sum_{i=1}^{n} Ln(\alpha) - 2\sum_{i=1}^{n} Ln\left(1 + \left(\frac{x_i+\mu}{\alpha}\right)^{\beta}\right)$
Step 2: Generate random samples as $u_1, u_2, \dots, u_n$ which are follows the continuous uniform distribution defined on the interval (0,1). Then transform it to a random sample follows Modified Log-logistic Distribution using cdf as follow:

$$F(x,\alpha,\beta) = \frac{1}{1+(\frac{x_i+\mu}{\alpha})^{-\beta}} \Rightarrow Ui = \frac{1}{1+(\frac{x_i+\mu}{\alpha})^{-\beta}} \Rightarrow \frac{x_i+\mu}{\alpha} = (\frac{1-U_i}{Ui})^{\frac{1}{-\beta}} \Rightarrow x_i = \alpha(\frac{Ui}{1-Ui})^{\frac{1}{\beta}} - \mu$$

PSO, GWO, GA and hybrid algorithms were used to find the best initial estimate of the parameters $[\beta, \mu, \alpha]$.
3- Based on L=1000 trials The MSE (mean square error) between the real and estimated values is calculated from $MSE = \frac{1}{L}\sum_{i=1}^{L}(\hat{S}_i - S)^2$.

4. Results and Discussion

In this study, a hybrid optimization method (PSGWO), which combines the particle swarm optimization (PSO) algorithm and the gray wolf optimization algorithm, was used to estimate the survival function of the modified log-logistic distribution. This method aims to improve the accuracy of estimating the distribution parameters by reducing the mean squared error (MSE) compared to the three standard meta-heuristic algorithms (PSO, GWO, and GA).
Tables 1-3 present the simulation results to measure the effectiveness of the estimation methods (PSO, GWO, GA, and PSGWO) when using fixed values (β, μ , α) and at four different sample sizes (25, 50, 75, and 100). This comparison aims to evaluate the performance of each method based on the mean squared error (MSE) for each case. These results demonstrate that the hybrid approach (PSGWO) outperforms the three algorithms in all scenarios, regardless of sample size. This superior performance is due to the benefits of both PSO, GA, and GWO.
It is worth noting that combining the PSO algorithm with MLE enhances estimation accuracy and reduces the mean square error. This is because PSO provides random searches and

multidirectional exploration, which iteratively improves the results until an optimal value is reached. In comparison, MLE has difficulty estimating optimally when parameters are nonlinearly related or interact in complex ways.

These results highlight the importance of utilizing PSOMLE in applications that necessitate precise survival function estimation, such as medical studies and reliability analysis in engineering systems. For example, in clinical studies aimed at evaluating the efficacy of a new drug, significant errors in estimating the survival function can lead to erroneous conclusions about treatment effectiveness. Therefore, using PSOMLE can provide higher accuracy in such estimates, ensuring more reliable results. In addition, applying this method in engineering fields, such as reliability analysis of devices and systems, can contribute to more accurate predictions of failure periods, which helps in developing more effective maintenance plans and reducing operational costs.

In addition, applying this method in engineering fields, such as reliability analysis of devices and systems, can contribute to more accurate predictions of failure periods, which helps in developing more effective maintenance plans and reducing operational costs.

Table 1. MSE values of $\hat{S}_{MLLGD}$ when $\beta = 0.1$, $\mu =$ 0.5, $\alpha = 0.8$

n	PSO	GWO	GA	PSWGO
25	6.9953e-08	6.9953e-08	1.9659e-07	1.7876e-09
50	2.7607e-07	2.7607e-07	3.8507e-05	1.8872e-07
75	9.8395e-07	2.1236e-07	7.4692e-05	6.8345e-08
100	1.2679e-07	1.2679e-07	1.5081e-05	1.1141e-08

Table 2. MSE values of $\hat{S}_{MLLGD}$ when $\beta =$ 1, $\mu = 0.1$, α =*2*

n	PSO	GWO	GA	PSWGO
25	3.4170e-14	3.4170e-14	2.7971e-05	2.0282e-08
50	6.6620e-13	6.6620e-13	6.4076e-05	1.4275e-07
75	1.2798e-11	1.2798e-11	3.1735e-05	1.0033e-07
100	1.6362e-12	1.6362e-12	2.4446e-05	2.0441e-07

Table 3. MSE values of $\hat{S}_{MLLGD}$ when $\beta =$ 1.5, $\mu =$ 2, $\alpha = 0.5$

n	PSO	GWO	GA	PSWGO
25	6.3215e-07	4.0170e-07	2.8996e-05	2.1989e-08
50	1.9951e-07	1.9951e-07	4.4248e-05	1.4848e-08
75	1.0310e-06	4.1310e-07	6.7700e-05	4.1669e-07
100	3.9906e-06	7.6375e-07	1.7219e-04	3.3690e-08

5. Conclusion

In this paper, a Log-logistic Distribution was modified by adding a shape parameter to improve the accuracy of estimates by adapting the model to data centered around a specific value, thereby increasing the model's ability to predict more effectively. Additionally, a new hybrid meta-heuristic algorithm was proposed to estimate the survival function of the modified log-logistic distribution. We used three standard metaheuristic algorithms to compare the results based on a simulation study and the mean square error criterion. Based on the results obtained, hybrid algorithm offers an optimal balance between exploration and exploitation, resulting in more accurate and reliable solutions for survival function estimation. The study recommends adopting these innovative methods in medical and industrial applications to achieve optimal results. Future studies could further improve the performance of these algorithms by incorporating additional technologies or improving search strategies. Ultimately, this research represents a crucial step toward developing more efficient and flexible methodologies in the field of survival analysis, thereby contributing to more informed and better scientific and practical decision-making.

Acknowledgment

Our researcher extends his Sincere thanks to the editor and members of the preparatory committee of the Ibn AL-Haitham Journal of Pure and Applied Sciences.

Conflict of Interest

There are no conflicts of interest.

Funding

There is no funding for the article.

References

1. In J, Lee DK. Survival analysis: Part II—applied clinical data analysis. Korean J Anesthesiol. 2019;72(5):441–457. https://doi.org/10.4097/kja.19183
2. Akbarinasab M, Arabpour ARM, Abbas M. Truncated log-logistic family of distributions. J Biostat Epidemiol. 2019;5(2):137–47. https://doi.org/10.18502/jbe.v5i2.2345
3. Bidaoui H, El Abbassi I, El Bouardi A, Darcherif A. Wind speed data analysis using Weibull and Rayleigh distribution functions: case study—five cities northern Morocco. Procedia Manuf. 2019;32:786–793. https://doi.org/10.1016/j.promfg.2019.02.286
4. Ibrahim A, Kalaf BA. Estimation of the survival function based on the log-logistic distribution. Int J Nonlinear Anal Appl. 2022;13(1):127–41. https://doi.org/10.22075/ijnaa.2022.5466
5. Ahmed IK, Taha TA, Salman AN. Single stage shrinkage estimation methods for reliability function of generalized exponential distribution using simulation. Iraqi J Sci. 2024:1531–40. https://doi.org/10.24996/ijs.2024.65.3.29
6. Casella G, Berger RL. Statistical Inference. 2nd ed. Boca Raton, FL: CRC Press; 2024. https://doi.org/10.1201/9781003456285
7. Zeng HB, Liu XG, Wang W, Xiao SP. New results on stability analysis of systems with time-varying delays using a generalized free-matrix-based inequality. J Frankl Inst. 2019;356(13):7312–7321. https://doi.org/10.1016/j.jfranklin.2019.03.029. .
8. Gandomi AH, Yang XS, Alavi AH. Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. Eng Comput. 2013;29(1):17–35. https://doi.org/10.1007/s00366-011-0241-y
9. Ali OA. Proposed algorithm for Gumbel distribution estimation. Aust J Basic Appl Sci. 2017;11(16):17–24.
10. Yoon S, Jeong C, Lee T. Application of harmony search to design storm estimation from probability distribution models. J Appl Math. 2013;2013:932943. https://doi.org/10.1155/2013/932943
11. Shin JY, Heo JH, Jeong C, Lee T. Meta-heuristic maximum likelihood parameter estimation of the mixture normal distribution for hydro-meteorological variables. Stoch Environ Res Risk Assess. 2014;28(2):347–58. https://doi.org/10.1007/s00477-013-0753-7
12. Özsoy VS, Örkçü HH. Estimating the parameters of nonlinear regression models through particle swarm optimization. Gazi Univ J Sci. 2016;29(1):187–199.
13. Yang F, Ren H, Hu Z. Maximum likelihood estimation for three-parameter Weibull distribution using evolutionary strategy. Math Probl Eng. 2019;2019:6281781. https://doi.org/10.1155/2019/6281781
14. Yonar A, Yapici N. Evaluation and comparison of metaheuristic methods to estimate the parameters of gamma distribution. Nicel Bilim Derg. 2022;4(1):96–119.
15. Yonar A, Yapici N. A novel differential evolution algorithm approach for estimating the parameters of gamma distribution: an application to the failure stresses of single carbon fibres. Hacet J Math Stat. 2020;49(5):1493–1514.
16. Akbari R, Mohammadi A, Ziarati K. A novel bee swarm optimization algorithm for numerical function optimization. Commun Nonlinear Sci Numer Simul. 2010;15(10):3142–55. https://doi.org/10.1016/j.cnsns.2009.11.003
17. Karaboga D, Basturk B. Artificial bee colony (ABC) optimization algorithm for solving constrained optimization problems. In: Proc Int Fuzzy Syst Assoc World Congr; 2007. p. 789–798.
18. Kiani F, Seyyedabbasi A, Aliyev R, Gülle M U, Başyıldız H, Shah MH . Adapted-RRT: novel hybrid method to solve three-dimensional path planning problem using sampling and metaheuristic-based

algorithms. Neural Comput Appl. 2021;33(22):15569–15599. https://doi.org/10.1007/s00521-021-06179-0
19. Han J, Seo Y. Mobile robot path planning with surrounding point set and path improvement. Appl Soft Comput. 2017;57:35–47. https://doi.org/10.1016/j.asoc.2017.03.035.
20. Bain LJ. Analysis for the linear failure rate life-testing distribution. Technometrics. 1974;16(4):551–559.
21. Ragab A, Green J. On order statistics from the log-logistic distribution and their properties. Commun Stat Theory Methods. 1984;13(21):2713–2724.
22. Kantam RRL, Rosaiah K, Rao GS. Acceptance sampling based on life tests: log-logistic model. J Appl Stat. 2001;28(1):121–128.
23. Kantam RRL, Rao GS, Sriram B. An economic reliability test plan: log-logistic distribution. J Appl Stat. 2006;33(3):291–296.
24. Wolpert DH, Macready WG. No free lunch theorems for optimization. IEEE Trans Evol Comput. 1997;1(1):67–82. https://doi.org/10.1109/4235.585893
25. Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. Adv Eng Softw. 2014;69:46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007
26. Liu J, Wei X, Huang H. An improved grey wolf optimization algorithm and its application in path planning. IEEE Access. 2021;9:121944–121956. https://doi.org/10.1109/access.2021.3108973