J.Natn.Sci.Foundation Sri Lanka 2024 52 (4): 441 - 458
DOI: http://dx.doi.org/10.4038/jnsfsr.v52i4.11928

RESEARCH ARTICLE

Statistics

Maximum likelihood estimation for the two-parameter Maxwell distribution

P Kasap* and AO Faouri
Department of Statistics, Faculty of Science, Ondokuz Mayis University, Atakum, Samsun, Turkey.

Submitted: 08 August 2024; Revised: 13 November 2024; Accepted: 26 November 2024

Abstract: The Maxwell distribution is popular in physics, chemistry and statistical dynamics. Since the estimators obtained using the maximum likelihood method have the desired properties of being efficient, consistent, and asymptotically normal under regularity conditions, this method is a widely used method to estimate the parameters of a probability distribution. Although parameter estimates can be obtained using this method, the derivatives of the log-likelihood equations, known as ML estimation equations, with respect to the parameters do not always have clear solutions. Therefore, numerical methods are used to solve these equations. Various traditional numerical methods for this purpose are well-documented in the literature. Additionally, highly powered algorithms with no required mathematical assumption that improve the computational efficiency like heuristic algorithms can be used to solve these equations. In this article, the maximum likelihood method is applied to estimate the location and scale parameters of the two parameter Maxwell distribution. High-performance heuristic algorithms, such as particle swarm optimization and genetic algorithms, are used and compared with traditional numerical techniques, including Nelder-Mead and Quasi-Newton methods. To show the performance of these techniques, an extensive Monte Carlo simulation study was conducted to compare the efficiencies of maximum likelihood estimators of model parameters concerning bias, mean square error, and deficiency criteria. Simulation results showed that genetic algorithm and particle swarm optimization estimators are more efficient than the other traditional algorithms for estimating the location and scale parameters for the two-parameter Maxwell distribution.

Keywords: Iterative methods, heuristic and traditional algorithms, maximum likelihood, Monte Carlo simulation, two-parameter Maxwell distribution.

INTRODUCTION

The Maxwell distribution is widely used in many scientific fields, including physics, chemistry, and statistical mechanics. This distribution is an appropriate alternative for modelling lifetime data such as flood levels, wind speeds, or failure rates in various application areas like engineering, environment, finance, insurance, and medical science (Pasha *et al.* 2006; Krishna & Malik, 2012; Hernandez, 2017; Prataviera *et al.*, 2020; Omar *et al.* 2021;). James Maxwell proposed the Maxwell distribution (Maxwell, 1860; Maxwell, 1867). It was first used for modelling lifetime data in 1989 (Tyagi & Bhattacharya, 1989a; Tyagi & Bhattacharya, 1989b). They used Bayes estimation method to estimate the scale parameter for the Maxwell distribution.

The Maxwell distribution is considered as a special case of the generalized Rayleigh distribution (Pham-Gia, 1994). In 1998, the Gamma distribution was transformed into the Maxwell distribution, and its scale parameter was estimated by classical and Bayesian methods (Chaturvedi & Rani, 1998). It is also considered a special case of the generalized Weibull distribution (Al-Mutairi & Agarwal, 1999). The parameter of the Maxwell distribution is estimated by the Bayesian method using the modified linear exponential loss function (MLINEX) (Podder & Roy, 2003). The maximum likelihood (ML) estimator, Bayesian estimator, and empirical Bayesian estimator of the Maxwell distribution derived and compared in

* Corresponding author (pelin.kasap@omu.edu.tr; https://orcid.org/0000-0002-1106-710X)

terms of efficiency using Monte Carlo simulation study (Bekker & Roux, 2005). The reliability properties of the Maxwell-Boltzmann distribution under type II censored data were estimated by using the Bayesian and maximum likelihood methods, and the efficiencies of the estimates generated by these estimation methods were compared by performing a Monte Carlo simulation study (Krishna & Malik, 2009). The relationship between ML, moments (MOM) estimators, and the variance of Maxwell distribution is reviewed with theoretical steps. Also, for quadratic loss functions, the minmax (minimax) estimator of the Maxwell distribution is determined using the Lehmann theorem, and the estimators are compared. (Dey and Maiti, 2010). ML estimator and variance-covariance matrix are found for the Maxwell distribution under type 1 censoring data (Kasmi *et al*., 2011). In 2013, Al-Baldawi also made another comparison between some Bayesian estimators using uninformative priors under two different loss functions and ML estimators for the scale parameter of the Maxwell distribution (Al-Baldawi, 2013). Rasheed and Khalifa (2016) used the Bayesian estimation method with a quadratic loss function to estimate the scale parameter of the Maxwell distribution. Many different data sets collected from the literature in 2016 were analyzed using the Maxwell distribution (Hossain & Huerta, 2016). Li used Minmax, Bayes, and ML methods to estimate the scale parameter of the Maxwell-Boltzmann distribution (Li, 2016).

In practice, distributions with only one parameter are limited and hard to fit, so it is better for data analysis to expand the family of distributions by adding an extra parameter using various techniques to give more flexibility (Dey *et al.*, 2016).

The extension of the Maxwell distribution by adding the location parameter provides a new, extended and more flexible type of distribution known as the two-parameter Maxwell distribution. Dey *et al*. (2016) extended this distribution and estimated the location and scale parameters using both classical and Bayesian methods. A new method based on algebraic approximation is used to correct the bias of ML estimator bias for the location and scale parameters of the Maxwell distribution (Maghami & Bahrami, 2020). Arslan *et al*. (2021) obtained the location and scale parameters of the Maxwell distribution using the modified maximum likelihood (MML) method along with other statistical methods for both complete and censored samples, they then compared the performance of the MML estimators with the ML, LS and MOM estimators through a Monte Carlo simulation study. Chowdhury (2022) constructed statistical confidence intervals for the two-parameter Maxwell distribution with ML, MOM and MML methods and the study shows that statistical confidence intervals based on ML estimators offer little improvement over other interval estimators when sample sizes are small.

Among these estimation methods, the ML method is widely used in many studies due to its well-known asymptotic properties such as asymptotic unbiasedness, consistency, efficiency, etc. It is known that ML is an estimation process based on maximizing the likelihood function for the concerned parameters. However, in some cases, the maximum likelihood equations do not have explicit solutions due to their nonlinear functions. In such cases, iterative algorithms can be utilized to the results.

The key objective of this study is to demonstrate the two-parameter Maxwell distribution's relevance in numerous scientific domains and to estimate the location and scale parameters of this distribution. The main problem addressed in this work is the absence of an explicit solution to the likelihood equations for this distribution. To address this, iterative numerical methods, using both heuristic and traditional algorithms, are applied to obtain maximum likelihood estimators.

As it is also seen in the literature, ML estimators can be obtained numerically by using the Newton-Raphson (NR) algorithm to solve the equation systems formed by the partial derivatives of the likelihood function. The main limitation of this algorithm is its reliance on gradient-based search methods (Martinez, 2000). To address this issue, well-known heuristic algorithms such as Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) are often employed. Additionally, other efficient traditional algorithms, including Nelder-Mead (NM) and Quasi-Newton (QN), are also used. Even though the traditional numerical techniques are widely used, they often face challenges such as convergence issues, low computational efficiency, and high costs when applied to non-linear equations. (Chen *et al*., 2023). Iterative methods, while an alternative, can also encounter several issues: (i) failure to converge, (ii) slow convergence, and (iii) convergence to an incorrect root (Barnett, 1966; Puthenpura and Sinha, 1986; Vaughan, 1992). In order to show the high performance of heuristic algorithm estimators, an extensive Monte Carlo simulation study was carried out to compare the ML estimators of GA and PSO with the other ML estimators for NM and QN.

The parameters of various statistical distributions and models, including Weibull, skewed normal, Nakagami, EMLOG, and others, were estimated using the maximum likelihood (ML) method, implemented with genetic algorithms (GA) and/or particle swarm optimization (PSO). (Yalçınkaya *et al*., 2018; Karakoca & Pekgör, 2019; Faouri & Kasap, 2023; Kasap & Faouri, 2024). To the best of our knowledge, this study is considered the first to obtain the ML estimators for the location and scale parameters of the two-parameter Maxwell distribution by using the heuristic algorithms. The rest of the article is organized as follows: In the section Materials and Methods section, the two-parameter Maxwell distribution and its properties are presented. Parameter estimation methodology via heuristic and the other traditional numerical techniques is introduced. In the results and discusion, the Monte Carlo simulation study is presented. Two real-life dataset are implemented as an application of this distribution. In the final section, the conclusion of this work is presented.

MATERIALS AND METHODS

Two-parameter Maxwell distribution

Maxwell distribution can be used to describe the distribution of gas molecules with respect to their speed distributions and relate them to the temperature (Atkins and De Paula, 2011). The probability density function (pdf) of the Maxwell distribution is given by:

$$f(v) = 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} v^2 e^{-\left(\frac{m}{2kT}v^2\right)}, \; v > 0 \qquad ...(1)$$

Where,

T : The temperature of the gas in kelvin
K : The constant value in J/K is 1.380649×10^{-23},
in J/K, which is considered as the proportionality factor that establishes a relationship between the average relative thermal energy of a gas's particles and its thermodynamic temperature.
m : The molecule's weight in kg/mol
v : The speed of molecule

The Maxwell distribution of speeds and its variation with temperature are given in Figure 1.

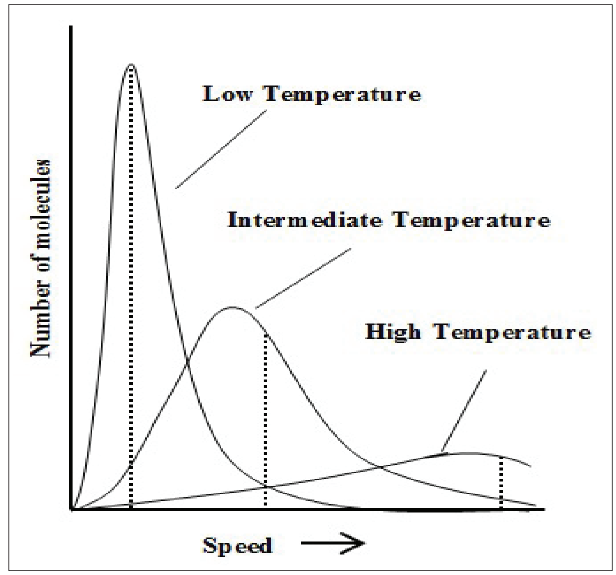

Figure 1: Maxwell distribution of speeds and its variation with temperature.

When the re-parameterization $\sigma = \sqrt{\frac{2kT}{m}}$ is applied and location parameter μ is added to equation (1), the resulting distribution is called the two-parameter Maxwell distribution.

Let X be a random variable that follows a two-parameter Maxwell distribution with parameters μ and σ then the probability density function (pdf) is:

$$f(x;\mu,\sigma) = \frac{4}{\sigma\Gamma(1/2)} \left(\frac{x-\mu}{\sigma}\right)^2 e^{-\left(\frac{x-\mu}{\sigma}\right)^2}, \mu \le x \le \infty, \sigma \ge 0 \qquad ...(2)$$

where,

μ: location parameter and σ: scale parameters

The corresponding cumulative distribution function (cdf) of this distribution is:

$$F(x;\mu,\sigma) = \frac{4}{\Gamma(3/2)} \; \Gamma\left[\left(\left(\frac{x-\mu}{\sigma}\right)^2, 3/2\right)\right] \qquad ...(3)$$

The mean and variance of the random variable X are given by:

$$E(X) = \mu + \frac{2\sigma}{\sqrt{\pi}} \qquad ...(4)$$

$$Var\,(X) = \frac{\sigma^2\,(3\pi - 8)}{2\pi} \qquad ...(5)$$

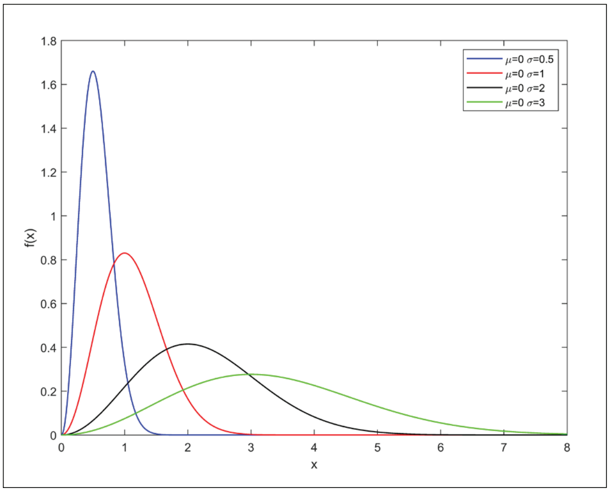

Figure 2: Two-parameter Maxwell distribution for certain values of σ μ = 0.

Maximum likelihood estimation

The maximum likelihood estimates are the values that maximize the likelihood function (Miura, 2011). The log-likelihood function (*LL*) for estimating the unknown parameters μ and σ for the μ two-parameter Maxwell distribution is given μ the follows:

$$L\,L = n ln\left(\frac{4}{\sigma\boldsymbol{\Gamma}(^1/_2)}\right) + 2\sum_{i=1}^{n}\left(\frac{x-\mu}{\sigma}\right) - \sum_{i=1}^{n}\left(\frac{x-\mu}{\sigma}\right)^2 \quad ...(6)$$

The normal equations for estimating the location parameter μ and the scale parameter σ are shown as below:

$$\frac{\partial LL}{\partial \mu} = \frac{-2}{\sigma}\sum_{i=1}^{n}\left(\frac{1}{x_i-\mu}\right) + \frac{2}{\sigma}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right) = 0 \quad ...(7)$$

and

$$\frac{\partial LL}{\partial \sigma} = \frac{n}{\sigma} - \frac{2}{\sigma}\sum_{i=1}^{n}(x_i-\mu) + \frac{2}{\sigma}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 = 0 \quad ...(8)$$

respectively

These normal equations are taken from the partial derivative of the *LL* function for the parameter of interest. Then these equations are equated to zero to find the solutions which represent the resulting ML estimate values of μ and σ. As shown in the normal equations (7) and (8), the functions are nonlinear, and an explicit solution for the likelihood equations cannot be determined. Therefore, we need iterative techniques to find the solutions and to obtain ML estimators of μ and σ. Iterative techniques PSO, GA, NM, and QN are briefly introduced in the following subsections.

Particles swarm optimization (PSO)

PSO is a population-based heuristic optimization technique developed from swarm intelligence and derived from the manual behaviour of bird flocks. It was proposed for the first time in 1995 (Kennedy & Eberhart, 1995). It is also known as an evolutionary self-adaptive search technique. Practically, it is preferred to be used for continuous optimization problems (Kachitvichyanukul, 2012). In this method, each solution is called a particle, and for any set of solutions, is called a population. It can be used in many fields because of its easy implementation, high precision, and fast convergence. The main advantage of PSO is that it can avoid the solution from being trapped in local optima and help to reach the global optimum or very close to it by searching at different points and different regions of the search space (Salehizadeh *et al*., 2009). Each solution (particle) consists of a set of parameters and represents a point in multi-dimensional space. The basic idea of the PSO method can be summarized as a process of continuously moving a swarm of particles in a specific search space concerning certain formulas until finally reaching the optimal solution (Júnior *et al*., 2020). PSO flowchart is shown in figure 3. For the PSO algorithm steps, see Júnior *et al*., 2020; Ren *et al*., 2014; Ab Talib and Mat Darus, 2017.

Genetic algorithm (GA)

GA is the first heuristic random optimization search technique that is inspired by biological evolution operating under natural selection (Whitley, 1994). Like PSO, it is an evolutionary self-adaptive method and it's useful in finding approximately the best solution in optimization problems, especially in discrete optimization (Kachitvichyanukul, 2012). It was first presented by John Holland in the early 1960s (Holland, 1975) and was greatly improved by his student Golberg later on (Golberg, 1989). In GA, each candidate solution represents a chromosome and each set of solutions (chromosomes) represents a population. The *LL* function represents the fitness value for each chromosome in the population, and in every iteration, the fitness value is calculated and evaluated. During the iteration, a certain modification process is applied according to the mechanism of biological evolution to reach the optimal solution at the end. GA flow chart is shown in figure 4; for more details of GA algorithm steps, see (Yalçınkaya *et al*., 2018).

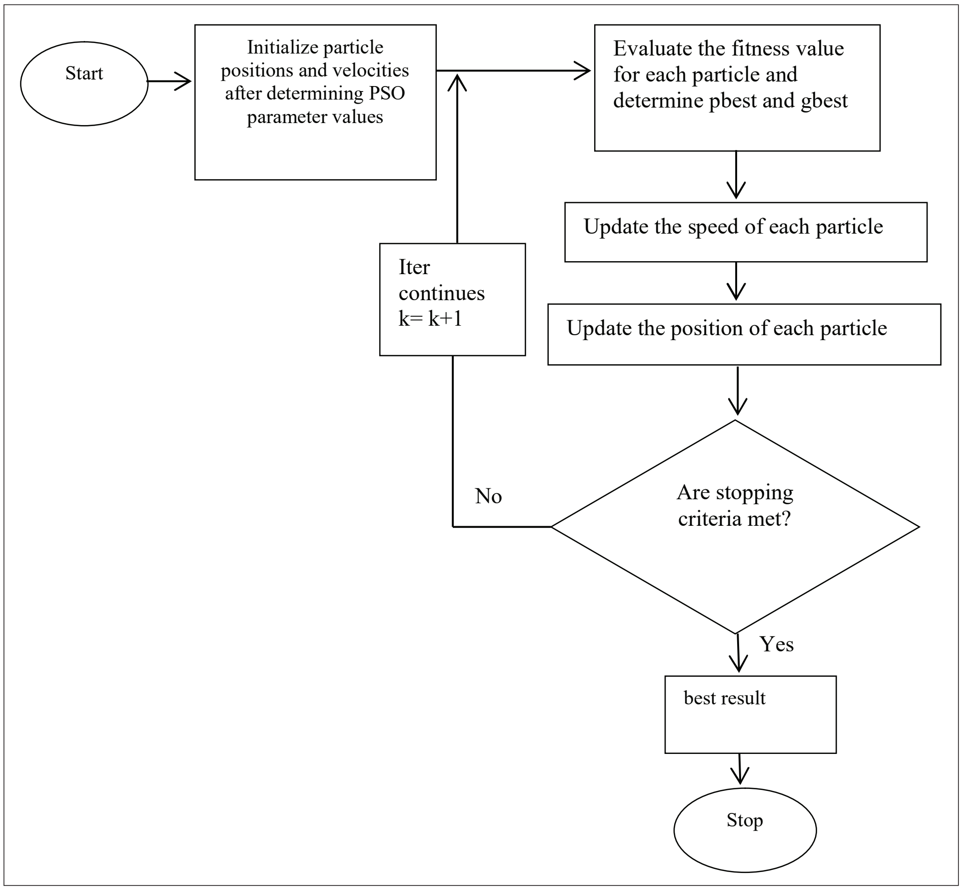

Figure 3: Flow chart of the PSO.

Nelder mead (NM)

The NM simplex algorithm is a popular deterministic direct search method, commonly used in n-dimensional space for finding a local minimum of the objective function in optimization problems. It was devised by John Nelder and Roger Mead in 1965 (Júnior *et al*., 2020). This simplex method denotes a convex hull geometric figure for n-dimensional problems with n+1 vertices, so for two-dimensional problems, the simplex is a triangle with three vertices. The pattern search method of this algorithm is based on ordering the vertices according to their fitness values, with the highest value considered as the worst and replaced with the newly generated vertex, forming a new simplex. This search process is repeated, making a sequence of generated simplexes with different shapes and sizes pass through the main four operations until finally reaching the optimal minimum (or maximum). In this study, the function that needs to be minimized is:

$$f\ (\theta) = -LL\ (\theta), where\ \theta = (\mu, \sigma) \in R \times R^{+}$$

In the implementation of the NM algorithm method, the main four operations are reflection, expansion, contraction, and shrink. Their coefficients are taken as $\alpha = 1$, $\gamma = 2$, $\rho = 1/2$ and $\beta = 1/2$ respectively, in the literature. For the NM algorithm steps (Nelder & Mead, 1965; Kucukdeniz & Esnaf, 2018; Gao & Han, 2012).

Quasi-Newton (QN)

The QN algorithm is the most commonly used numerical technique for finding solutions to nonlinear unconstrained objective functions by using quadratic approximation. This method is based on the regular Newton-Raphson (NR) method. However, in some situations, NR is difficult to apply because in every iteration, a partial derivative must be computed and a linear system must be solved, which demands extra time and costs (Martınez, 2000).

This fact motivated the development of quasi-Newton by Davidon in 1959 (Davidon, 1991). The QN is the same as NR concerning twice differentiability of the objective function; the initial starting point should be close to the true root; and it uses the same search direction. The only difference is that for QN, the gradient of the objective function is used to estimate the inverse Hessian matrix in many different approximation algorithms. More details on that can be found in the literature (Broyden, 1967; Shanno, 1970; Fletcher & Powell, 1963). For the QN algorithm steps, see (Karakoca & Pekgör, 2019).

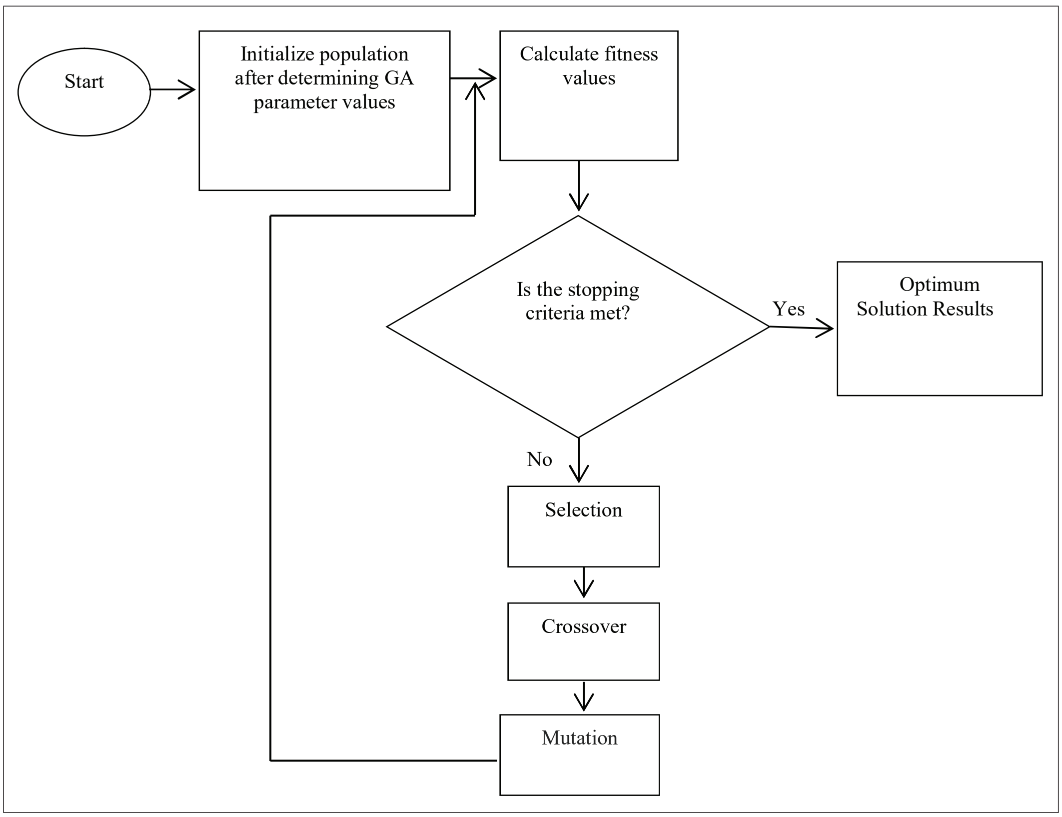

Figure 4: Flow chart of the GA.

Advantages of the GA and PSO algorithm in comparison with other algorithms used in this study.

- GA and PSO are considered as the simplest heuristic algorithm for programming purposes, and it can be easily used for any distribution by implementing the log-likelihood as a fitness function of the model in the programme.
- The efficiency of GA and PSO as a heuristic-based algorithms are better than other traditional algorithms used in this study, according to the simulation results.
- Having a starting point as an initial value that must be close to the optimum solution is critical for traditional iterative algorithms such as NM and QN to successfully converge at the end. However, in GA and PSO, the initial value problem does not exist because they do not depend on it for convergence. Instead, operation of GA and PSO starts with a set of solutions called the population, and it only needs a search space that is considered as an interval containing the global optimum. It is known that defining an interval that includes the solution root is much better than defining a single starting guess value. However, according to the simulation results for this study model, if the starting point for traditional methods is selected very carefully to be very close to the solution, it may need less time to reach the solution, but even if that is the case, the efficiency of the GA and PSO method is

better.

- The GA and PSO methods eliminate the need for mathematical assumptions based on differentiability, as in the Newton method, and instead relies solely on the evaluation of the objective function, demonstrating the method's simplicity.
- The flexibility of the GA and PSO methods allows it to be applied to optimization problems with linear or nonlinear functions, and it can be identified in the discrete, continuous, or mixed search space.

RESULTS AND DISCUSSION

Monte Carlo (MC) simulations study

In this section, an extensive Monte Carlo (MC) simulation study is conducted in which the efficiency of the PSO and GA estimators is compared with the NM, and QN estimators. All computations are conducted by using the following functions: "particleswarm", "ga", "fminsearch", and "fminunc" for programming PSO, GA, NM, and QN respectively in MATLAB 2021a software. The values of the initial conditions for PSO and other methods used in this study are considered to be the default for the fixed parameters; that is, the inertia weight = max {0.1, 1.1}, and the accelerating coefficients c1 = c2 = 1.49. The maximum iteration number until convergence happens is taken to be 400, which is equivalent to the number of estimated parameters (two parameters in this study) multiplied by 200, and the convergence is achieved according to the stopping rule: $\left|\theta^{k+1} - \theta^{k}\right| < \varepsilon$.

Each Monte Carlo (MC) simulation run is executed 2000 times. The location parameter μ and scale σ are considered to be (μ = 0, 1, 2) and (σ = 1, 2) respectively. The sample size is taken as small (n = 10, 20, 30), moderate (n = 50) and large (n = 100, 250, 500). The search space (SS) is selected for both μ and σ parameters to be [0,5]. NM estimates were obtained by taking the starting point values [0.01, 0.01] for all sample sizes, and the same starting point can be used for QN for small and moderate sample sizes only. However, for large sample sizes, this point is not advised, so the point [0.1,0.01] was chosen for the QN method to guarantee convergence. The resulting estimates for location μ and scale σ parameters in the simulations are denoted by $\hat{\mu}$ and $\hat{\sigma}$, respectively. The main criteria used to compare between estimators are Mean, Bias, Mean Square Error (MSE), and Deficiency (Def). Also, the necessary iteration number (Iter) and the required time (Time) in seconds to reach the solution for every sample have been generated in the simulations for all methods. The sum of total estimated values over the total number of MC simulation runs (average) is called the mean; mathematically it can be expressed as:

$$Mean(\hat{\theta}) = \frac{\sum_i^n \hat{\theta}_i}{n} \qquad ...(9)$$

where n is the number of MC simulation runs. The difference between the expected value of the estimator and the true value of the parameter is called Bias. Mathematically, it can be expressed as:

$$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta \qquad ...(10)$$

MSE is one of the main criteria to compare biased estimators according to their efficiencies. Any estimator with the smallest value of MSE is considered the most efficient and the best choice between other estimators, mathematically it can be expressed as:

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + \left(Bias(\hat{\theta})\right)^2 \qquad ...(11)$$

The variance and Standard Error (SE) of an estimator are calculated and expressed mathematically as the following equalities:

$$Var(\hat{\theta}) = \frac{1}{n-1}\sum_{i=1}^n\left(\hat{\theta}_i - Mean\,\hat{\theta}\right)^2 \qquad ...(12)$$

and

$$SE(\hat{\theta}) = \sqrt{Var(\hat{\theta})} \qquad ...(13)$$

respectively,

Deficiency is an essential major of the joint efficiency of the estimators μ and σ, which is defined as the sum of the MSE of the parameter estimators. It can be formulated mathematically as:

$$Def(\hat{\mu}, \hat{\sigma}) = MSE(\hat{\mu}) + MSE(\hat{\sigma}) \qquad ...(14)$$

Iteration is the process of repeating a step continually until convergence occurs. For every method used in this study, the total number of iterations required for the estimation process for each sample was summed up and then divided over the total number of MC simulation runs to calculate the average number of iterations. It has the following mathematical expression:

$$\overline{Iter}(\hat{\theta}) = \frac{\sum_i^n Iter(\hat{\theta})_i}{n} \qquad (15)$$

To calculate the average amount of time needed to estimate the parameters, the total time (measured in seconds) needed for each simulation run is added up and then divided by the total number of simulation runs.

$$\overline{Time}(\hat{\theta}) = \frac{\sum_i^n Time(\hat{\theta})_i}{n} \quad ...(16)$$

The resulting simulated values of Mean, Bias, MSE, and Def for $\hat{\mu}$ and $\hat{\sigma}$ are shown in Tables (1-5) in appendix and the best results are highlighted in bold in these tables. The simulation results show that the heuristic algorithms have better results among the other traditional algorithms. It should be noted that before analyzing the simulation results, it was verified that all of the randomly generated samples have real roots (exitflag = 1).

For bias criteria, the results are as follows:

- When $\mu = 0$, $\sigma = 1$, the smallest value of bias belongs to QN for $\hat{\mu}$ and $\hat{\sigma}$ estimators for all sample sizes, and the largest values belong to GA for the small and moderate sample sizes, but for the large sample size, it belongs to NM, as shown in Table 1 in appendix.
- When $\mu = 1$, $\sigma = 1$ the smallest value of bias belongs to GA for $\hat{\mu}$ and $\hat{\sigma}$ estimators for all sample sizes as shown in Table 2 in appendix.
- When $\mu = 1$, $\sigma = 2$, the smallest values of bias belong PSO for $\hat{\mu}$ and $\hat{\sigma}$ estimators for all sample sizes (except when n = 10, 20, the smallest bias values belong to NM and GA, respectively), as shown in Table 3 in appendix.
- When $\mu = 2$, $\sigma = 1$, the smallest values of bias belong to GA for $\hat{\mu}$ and $\hat{\sigma}$ estimators (except when n= 20, the smallest bias values belong to PSO) for all sample sizes as shown in Table 4 in appendix.
- When $\mu = 2$, $\sigma = 2$, the smallest values of bias belong to GA for the $\hat{\mu}$ estimator, and the smallest values of bias belong to PSO for the $\hat{\sigma}$ estimator (except when n = 20, the smallest bias values belong to PSO) as shown in Table 5 in appendix.

The worst bias with the largest values in all cases (except when $\mu = 0$, $\sigma = 1$) belongs to QN for all sample sizes, as shown in Tables 2 - 5 in appendix.

Concerning MSE and Def values for the location estimator $\hat{\mu}$ and scale estimator $\hat{\sigma}$,the results are as the follows

- When $\mu = 0$, $\sigma = 1$ we see that PSO estimators have the best values for all sample sizes. The NM estimator has the worst values among all other estimators, especially at large sample sizes. Also, it's clear that when sample size increases, the MSE values of GA improve to be slightly better than QN and so close to the PSO values as shown in Table 1 in appendix.

  According to the Def criteria, the PSO shows that it has the best performance with the smallest values in comparison with the other estimators for all cases. The results of QN are so close to PSO in the small and moderate sample sizes, but in large sample sizes (n=250, 500), GA outperforms QN to be the second method with the smallest Def values after PSO.
- When $\mu = 1$, $\sigma = 1$ we see that the GA then PSO estimators have the best values for all sample sizes and the NM and QN estimators have the largest values respectively as shown in Table 2 in appendix.
- When $\mu = 1$, $\sigma = 2$, $\mu = 2$, $\sigma = 1$ and $\mu = 2$, $\sigma = 2$ we see that the GA then PSO estimators have the best values for all sample sizes and the NM and QN estimators have the largest values respectively as shown in Tables 3-5. However when $\mu = 2$, $\sigma = 2$ its noticed that In moderate and large sample sizes the MSE as well as the Def values of GA is enhanced with better values than the other algorithms used.
- When $\mu = 2$, $\sigma = 2$ we see that the GA then PSO estimators have the best values for all sample sizes and the NM and QN estimators have the largest values respectively. According to the Def criteria, the GA shows that it has the best performance with the smallest values in comparison with the other estimators for all cases (except when $\mu = 0$, $\sigma = 1$) as shown in Tables 2-5. Also we can notice that when $\mu = 2$, $\sigma = 2$ in moderate and large sample sizes the MSE as well as the Def values of GA is enhanced with noticeably better values than the other algorithms used as illustrated in figures 5-7.

All that can lead us to say that heuristic algorithms represented by GA and PSO are better than traditional algorithms such as NM and QN for estimating the location $\hat{\mu}$ and scale $\hat{\sigma}$ parameters for the two-parameter Maxwell distribution.

Based on the average number of iterations, QN requires the fewest iterations, followed by PSO, GA, and NM, respectively. In terms of the average time needed to reach a solution, NM and QN are faster than PSO and GA. However, it is important to note that both the number of iterations and the time required to reach a solution depend heavily on the choice of the initial point. If a poor initial point is selected, both the iteration count and the solution time may increase significantly. The main issue is that traditional algorithms such as NM and QN may be forced to stop because convergence has not occurred. For these reasons, even though traditional methods are faster in some situations, methods based on

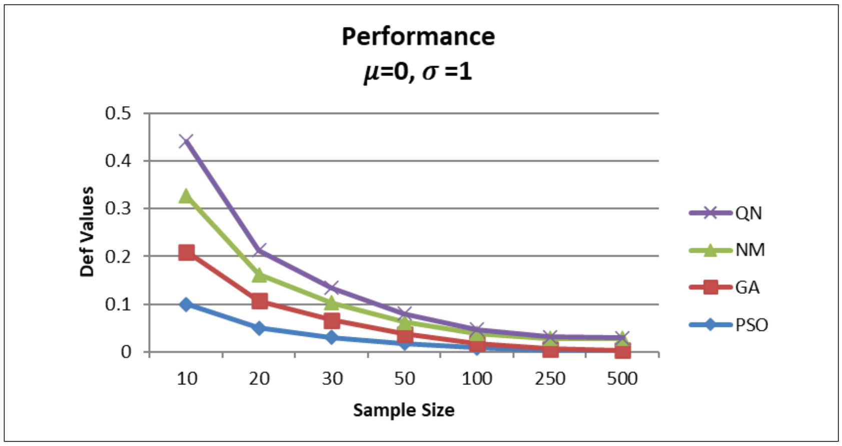

Figure 5: Performance of GA, PSO, NM, and QN according to the Def values when μ=0, σ =1

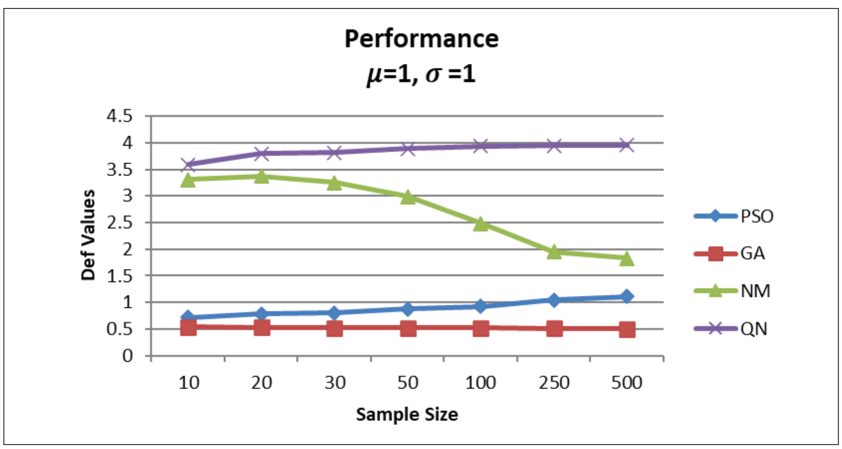

Figure 6: Performance of GA, PSO, NM, and QN according to the Def values when μ=1, σ =1

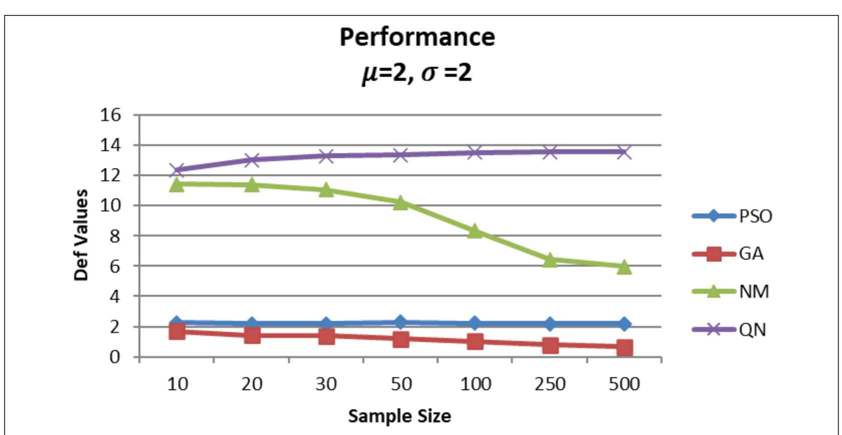

Figure 7: Performance of GA, PSO, NM, and QN according to the Def values when μ=2, σ =2

heuristic algorithms are preferred.

In this study, we can classify the four estimation methods into two groups: heuristic algorithm methods (including PSO and GA) and traditional methods (including NM and QN). When comparing PSO and GA as heuristic-based algorithms used in this study, the GA algorithm shows fewer deficiency values for different values of μ and σ. Additionally, both algorithms require approximately the same time to complete all iterations needed to estimate the model parameters. Simultaneously, when only traditional methods are compared, NM is preferred over the QN algorithm with respect to the deficiency values for different values of μ and σ, while QN is preferred over the NM due to its slightly lower required iteration number and time, even though they almost require the same calculation time. When we compare the two groups, it is clear that heuristic algorithms have the best performance according to Def values regardless of the number of iteration and calculation time, because these criteria are highly dependent on the starting point value for traditional methods. At the same time, in large sample sizes, the GA method outperforms the QN method (when $\mu = 0$, $\sigma = 1$), implying that heuristic algorithms outperform traditional algorithms when large sample sizes are used. Finally, based on the simulation results, it can be said that heuristic algorithms including GA and PSO are more efficient than the other traditional algorithms used in the study for estimating the parameters of the two-parameter Maxwell distribution.

Applications

For applications 1 and 2 the two-parameter Maxwell distribution is used in this section to model two real data sets. The unknown parameters are estimated via the maximum likelihood method for each dataset and evaluated using the PSO, GA, NM, and QN algorithms. The modelling performance for fitted Maxwell distribution is compared and evaluated between all methods used in this study by using well-known selection criteria including log-likelihood (LL), Akaike Information Criterion (AIC), and corrected AIC (AICc), see (Anderson *et al*., 1998). The method that best fits these criteria is considered to have the lowest values.

Application 1: Aircraft windshield service time data

This dataset, is widely used in many engineering and statistical sectors which is used for method implementation (Tahir *et al*., 2015; Balogun *et al*., 2021). It includes 63 observations about aircraft windshields' service times (1000 hours is the measurement unit used). *Observations on the aircraft windshield service time dataset:*

0.046, 1.436, 2.592, 0.140, 1.492, 2.600, 0.150, 1.580, 2.670, 0.248, 1.719, 2.717, 0.280, 1.794, 2.819, 0.313, 1.915, 2.820, 0.389, 1.920, 2.878, 0.487, 1.963, 2.950, 0.622, 1.978, 3.003, 0.900, 2.053, 3.102, 0.952, 2.065, 3.304, 0.996, 2.117, 3.483,1.003, 2.137, 3.500, 1.010, 2.141, 3.622, 1.085, 2.163, 3.665, 1.092, 2.183, 3.695, 1.152, 2.240, 4.015, 1.183, 2.341, 4.628, 1.244, 2.435, 4.806, 1.249, 2.464, 4.881, 1.262, 2.543, 5.140.

Table 6 in appendix shows the descriptive statistics for this dataset that contain the values of sample size (n), Minimum (Min), First Quartile (1st Qu.), Mean, Mode, Median, Third Quartile (3rd Qu.), Maximum (Max), Variance (S^2), skewness (γ_1), and kurtosis (γ_2) coefficients, respectively.

The parameter estimates corresponding to the selection criteria are given in table 7 in appendix. The results show that the PSO and GA methods have the best performance fitting for the two-parameter Maxwell distributions, in comparison with other traditional methods like NM and QN.

Application 2: Vinyl chloride data

This dataset is concerned with environmental investigation issues and includes 34 observations. The dataset is provided below and the descriptive statistics for this dataset are given in table 8 in the appendix. More details about this dataset can be found in (Bhaumik *et al*., 2009; Shukla, 2019). *Observations of the vinyl chloride dataset:*

5.1, 1.2, 1.3, 0.6, 0.5, 2.4, 0.5, 1.1, 8, 0.8, 0.4, 0.6, 0.9, 0.4, 2, 0.5, 5.3, 3.2, 2.7, 2.9, 2.5, 2.3, 1, 0.2, 0.1, 0.1, 1.8, 0.9, 2, 4, 6.8, 1.2, 0.4, 0.2

Table 9 in appendix shows the parameter estimate results for the various selection criteria. The second application yields the same result as the first, namely that the PSO and GA methods outperform other traditional methods like NM and QN in fitting the two-parameter Maxwell distributions, demonstrating that heuristic algorithms outperform classical and traditional methods. However, when we compare NM and QN, we see that QN performs better in the first application while NM performs better in the second application, which leads us to the conclusion that we cannot rely on a single traditional method for estimating the parameters. Hence, we can claim that GA and PSO are the best methods for estimating the parameters of the two-parameters Maxwell distribution.

CONCLUSION

In this study, the ML method is used to estimate the location and scale parameters of the two-parameter Maxwell distribution. In many cases, it is hard to have a solution for the parameters of interest for complicated nonlinear maximum likelihood equations. For these complicated nonlinear maximum likelihood equations, we need an efficient numerical technique to reach the best estimate values. This is the originality of this study. By conducting an intensive MC simulation study, the PSO, GA, NM, and QN algorithms are selected, and the performance of these algorithm estimators is compared with each other according to bias, MSE, Def, the number of required iterations (*Iter*), and calculation time criteria. Concerning these criteria, the results show that the heuristic algorithms used in this study, which are PSO and GA, respectively have the best performance values with robust efficient solutions regardless of the number of iterations and calculation time due to the fact that the performance criteria of traditional methods' are heavily reliant on the starting point value. However, heuristic algorithm-based methods such as GA and PSO are more efficient than traditional search optimization methods and yield the best results across all sample sizes, so these algorithms are more recommended for the study model.

REFERENCES

Ab Talib, M. H., & Mat Darus, I. Z. (2017). Intelligent fuzzy logic with firefly algorithm and particle swarm optimization for semi-active suspension system using magneto-rheological damper. *Journal of Vibration and Control*, *23*(3), 501-514. https://doi.org/10.1177/1077546315580693

Anderson, D. R., Burnham, K. P., & White, G. C. (1998). Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture-recapture studies. *Journal of Applied Statistics*, *25*(2), 263-282. https://doi.org/10.1080/02664769823250

Arslan, T., Acitas, S., & Senoglu, B. (2021). Parameter estimation for the two-parameter Maxwell distribution under complete and censored samples. *REVSTAT-Statistical Journal*, *19*(2), 237-253. https://doi.org/10.57805/revstat.v19i2.341

Atkins, P., & De Paula, J. (2011). *Physical chemistry for the life sciences*. Oxford University Press, USA.

Balogun, O. S., Arshad, M. Z., Iqbal, M. Z., & Ghamkhar, M. (2021). A new modified Lehmann type–II G class of distributions: exponential distribution with theory, simulation, and applications to engineering sector. *F1000Research*, *10*(483), 483. https://doi.org/10.12688/f1000research.52494.1

Barnett, V. D. (1966). Evaluation of the maximum-likelihood estimator where the likelihood equation has multiple roots. *Biometrika*, *53*(1/2), 151-165. https://doi.org/10.2307/2334061

Bekker, A. J. J. J., & Roux, J. J. J. (2005). Reliability characteristics of the Maxwell distribution: A Bayes estimation study. *Communications in Statistics-Theory and Methods*, *34*(11), 2169-2178. https://doi.org/10.1080/STA-200066424

Bhaumik, D. K., Kapur, K., & Gibbons, R. D. (2009). Testing parameters of a gamma distribution for small samples. *Technometrics*, *51*(3), 326-334. https://doi.org/10.1198/tech.2009.07038

Broyden, C. G. (1967). Quasi-Newton methods and their application to function minimisation. *Mathematics of Computation*, *21*(99), 368-381. https://doi.org/10.2307/2003239

Chen, X., Zhang, K., Ji, Z., Shen, X., Liu, P., Zhang, L., & Yao, J. (2023). Progress and challenges of integrated machine learning and traditional numerical algorithms: taking reservoir numerical simulation as an example. *Mathematics*, *11*(21), 4418.

Davidon, W. C. (1991). Variable metric method for minimization. *SIAM Journal on Optimization*, *1*(1), 1-17. https://doi.org/10.1137/0801001

Dey, S., & Maiti, S. S. (2010). Bayesian estimation of the parameter of Maxwell distribution under different loss functions. *Journal of Statistical Theory and Practice*, *4*(2), 279-287. https://doi.org/10.1080/15598608.2010.10411986

Dey, S., Dey, T., Ali, S., & Mulekar, M. S. (2016). Two-parameter Maxwell distribution: Properties and different methods of estimation. *Journal of Statistical Theory and Practice*, *10*(2), 291-310. https://doi.org/10.1080/15598608.2015.1135090

Faouri, A. O., & Kasap, P. (2023). Maximum likelihood estimation for the nakagami distribution using particle swarm optimization algorithm with applications. *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, *5*(2), 209-218.

Fletcher, R., & Powell, M. J. (1963). A rapidly convergent descent method for minimization. *The computer journal*, *6*(2), 163-168. https://doi.org/10.1093/comjnl/6.2.163

Gao, F., & Han, L. (2012). Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications*, *51*(1), 259-277. http://dx.doi.org/10.1007/s10589-010-9329-3

Golberg, D. E. (1989). Genetic algorithms in search, optimization, and machine learning. *Addion wesley*, *1989*(102), 36.

Holland John, H. (1975). Adaptation in natural and artificial systems. *Ann Arbor: University of Michigan Press*.

Hernandez, H. (2017). Standard Maxwell-Boltzmann distribution: definition and properties. *ForsChem Research Reports*, *2*, 2017-2. https://doi.org/10.13140/RG.2.2.29888.74244

Júnior, S. F. A. X., Xavier, É. F. M., da Silva Jale, J., de Oliveira, T. A., & Sabino, A. L. C. (2020). An application of Particle Swarm Optimization (PSO) algorithm with daily precipitation data in Campina Grande, Paraíba, Brazil. *Research, Society and Development*, *9*(8), e444985841-e444985841 https://doi.org/10.33448/rsdv9i8.5841

Kachitvichyanukul, V. (2012). Comparison of three evolutionary algorithms: GA, PSO, and DE. *Industrial Engineering and Management Systems*, *11*(3), 215-223. https://doi.org/10.7232/iems.2012.11.3.215

Karakoca, A., & Pekgör, A. (2019). Maximum likelihood estimation of the parameters of progressively type-2 censored samples from weibull distribution using genetic algorithm. *Academic Platform-Journal of Engineering and Science*, *7*(2), 189-199. https://doi.org/10.21541/apjes.452564

Kasap, P., & Faouri, A. O. (2024). Comparison of the meta-heuristic algorithms for maximum likelihood estimation of the exponentially modified logistic distribution. *Symmetry*, *16*(3), 259.

Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks* (Vol. 4, pp. 1942-1948). IEEE. https://doi.org/10.1109/ICNN.1995.488968

Krishna, H., & Malik, M. (2012). Reliability estimation in Maxwell distribution with progressively type-II censored data. *Journal of Statistical Computation and Simulation*, *82*(4), 623-641. http://dx.doi.org/10.1080/00949655.2010.550291

Kucukdeniz, T., & Esnaf, S. (2018). Hybrid revised weighted fuzzy c-means clustering with Nelder-Mead simplex algorithm for generalized multisource Weber problem. *Journal of Enterprise Information Management*, *31*(6), 908-924.

Maghami, M. M., & Bahrami, M. (2020). Improved maximum likelihood estimation of parameters in the Maxwell distribution. *Journal of Statistical Modelling: Theory and Applications*, *1*(2), 99-113.

Martınez, J. M. (2000). Practical quasi-Newton methods for solving nonlinear systems. *Journal of computational and Applied Mathematics*, *124*(1-2), 97-121. https://doi.org/10.1016/S0377-0427(00)00434-9

Maxwell, J. C. (1860). V. Illustrations of the dynamical theory of gases.—Part I. On the motions and collisions of perfectly elastic spheres. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 19(124), 19-32. https://doi.org/10.1080/14786446008642818

Maxwell, J. C. (1867). IV. On the dynamical theory of gases. *Philosophical transactions of the Royal Society of London*, (157), 49-88. https://doi.org/10.1098/rstl.1867.0004

Miura, K. (2011). An introduction to maximum likelihood estimation and information geometry. *Interdisciplinary*

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The computer journal*, *7*(4), 308-313. https://doi.org/10.1093/comjnl/7.4.308

Omar, M. H., Arafat, S. Y., Hossain, M. P., & Riaz, M. (2021). Inverse Maxwell Distribution and Statistical Process Control: An Efficient Approach for Monitoring Positively Skewed Process. *Symmetry*, *13*(2), 189. https://doi.org/10.3390/sym13020189

Pasha, G. R., Aslam, M., & Javed, M. (2006). Deviation of the Variances of Classical Estimators and Negative Integer Moment Estimator from Minimum Variance Bound with Reference to Maxwell Distribution. *Pakistan Journal of Statistics and Operation Research*, 145-150. https://doi.org/10.18187/pjsor.v2i2.98

Prataviera, F., Ortega, E. M., & Cordeiro, G. M. (2020). A new bimodal Maxwell regression model with engineering applications. *Applied Mathematics & Information Sciences*, *14*, 817-31. http://dx.doi.org/10.18576/amis/140509

Puthenpura, S., & Sinha, N. K. (1986). Modified maximum likelihood method for the robust estimation of system parameters from very noisy data. *Automatica*, *22*(2), 231-235. https://doi.org/10.1016/0005-1098(86)90085-3

Ren, C., An, N., Wang, J., Li, L., Hu, B., & Shang, D. (2014). Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting. *Knowledge-based systems*, *56*, 226-239. https://doi.org/10.1016/j.knosys.2013.11.015

Salehizadeh, S. M. A., Yadmellat, P., & Menhaj, M. B. (2009, March). Local optima avoidable particle swarm optimization. In *2009 IEEE Swarm Intelligence Symposium* (pp. 16-21). IEEE. https://doi.org/10.1109/SIS.2009.4937839

Shanno, D. F. (1970). Conditioning of quasi-Newton methods for function minimization. *Mathematics of computation*, *24*(111), 647-656. https://doi.org/10.1090/S0025-5718-1970-0274029-X

Shukla, K. K. (2019). A comparative study of one parameter lifetime distributions. *Biometrics & Biostatistics International Journal*, *8*(4), 111-123. https://doi.org/10.15406/bbij.2019.08.00280

Tahir, M. H., Cordeiro, G. M., Mansoor, M., & ZUBAİR, M. (2015). The Weibull-Lomax distribution: properties and applications. *Hacettepe Journal of Mathematics and Statistics*, *44*(2), 455-474. https://doi.org/10.15672/HJMS.2014147465

Tyagi, R. K., & Bhattacharya, S. K. (1989a). Bayes estimation of the Maxwell's velocity distribution function. *Statistica*, *29*(4), 563-567.

Tyagi, R. K., & Bhattacharya, S. K. (1989b). A note on the MVU estimation of reliability for the Maxwell failure distribution. *Estadistica*, *41*(137), 73-79.

Vaughan, D. C. (1992). On the Tiku-Suresh method of estimation. *Communications in Statistics-theory and Methods*, *21*(2), 451-469. https://doi.org/10.1080/03610929208830788

Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and computing*, *4*(2), 65-85. https://doi.org/10.1007/BF00175354

Yalçınkaya, A., Şenoğlu, B., & Yolcu, U. (2018). Maximum likelihood estimation for the parameters of skew normal distribution using genetic algorithm. *Swarm and Evolutionary Computation*, *38*, 127-138. https://doi.org/10.1016/j.swevo.2017.07.007

APPENDIX

Table 1: Simulated Mean, Bias, Variance, MSE, and Def values when μ=0, σ=1.

	μ=0,σ=1	Mean	Bias	Variance	MSE	Mean	Bias	Variance	MSE	Def	$\overline{Iter}$	$\overline{Time}$
10	PSO	0.1564	0.1564	**0.0301**	**0.0546**	0.8711	-0.1289	**0.0287**	**0.0453**	**0.0999**	48	0.1644
	NM	0.1229	0.1229	0.0493	0.0644	0.8955	-0.1045	0.0425	0.0534	0.1178	64	0.0058
	QN	**0.1048**	**0.1048**	0.0509	0.0619	**0.9084**	**-0.0916**	0.0438	0.0522	0.1141	**5**	**0.0053**
20	PSO	0.1027	0.1027	**0.0155**	**0.0260**	0.9154	-0.0846	**0.0164**	**0.0236**	**0.0496**	48	0.1644
	GA	0.1130	0.1130	0.0184	0.0312	0.9078	-0.0922	0.0173	0.0258	0.0570	77	0.2124
	NM	0.0804	0.0804	0.0237	0.0302	0.9321	-0.0679	0.0206	0.0252	0.0554	64	0.0058
	QN	**0.0542**	**0.0542**	0.0239	0.0268	**0.9512**	**-0.0488**	0.0219	0.0243	0.0511	**5**	**0.0053**
30	PSO	0.0800	0.0800	**0.0096**	**0.0160**	0.9394	-0.0606	**0.0105**	**0.0142**	**0.0302**	47	0.1797
	GA	0.0885	0.0885	0.0119	0.0197	0.9333	-0.0667	0.0115	0.0159	0.0357	76	0.2217
	NM	0.0687	0.0687	0.0151	0.0198	0.9480	-0.0520	0.0138	0.0165	0.0363	63	0.0058
	QN	**0.0410**	**0.0410**	0.0148	0.0165	**0.9684**	**-0.0316**	0.0139	0.0149	0.0314	**4**	**0.0054**
50	PSO	0.0572	0.0572	**0.0057**	**0.0090**	0.9533	-0.0467	**0.0060**	**0.0082**	**0.0172**	46	0.3535
	GA	0.0632	0.0632	0.0069	0.0109	0.9489	-0.0511	0.0067	0.0093	0.0202	76	0.4302
	NM	0.0595	0.0595	0.0103	0.0138	0.9517	-0.0483	0.0085	0.0108	0.0247	62	0.0121
	QN	**0.0237**	**0.0237**	0.0091	0.0097	**0.9782**	**-0.0218**	0.0081	0.0086	0.0182	4	**0.0103**
100	PSO	0.0372	0.0372	**0.0028**	**0.0042**	0.9715	-0.0285	**0.0032**	**0.0040**	**0.0082**	47	0.2236
	GA	0.0415	0.0415	0.0034	0.0051	0.9683	-0.0317	0.0036	0.0046	0.0097	76	0.2606
	NM	0.0695	0.0695	0.0072	0.0120	0.9477	-0.0523	0.0055	0.0082	0.0203	61	0.0076
	QN	**0.0120**	**0.0120**	0.0044	0.0045	**0.9904**	**-0.0096**	0.0042	0.0043	0.0088	**4**	**0.0060**
250	PSO	0.0213	0.0213	**0.0009**	**0.0014**	0.9831	-0.0169	**0.0011**	**0.0014**	**0.0027**	46	0.2905
	GA	0.0234	0.0234	0.0011	0.0016	0.9815	-0.0185	0.0013	0.0016	0.0033	75	0.3153
	NM	0.0920	0.0920	0.0049	0.0134	0.9307	-0.0693	0.0033	0.0081	0.0215	59	0.0098
	QN	**0.0066**	**0.0066**	0.0017	0.0017	**0.9941**	**-0.0059**	0.0016	0.0016	0.0034	**5**	**0.0064**
500	PSO	0.0141	0.0141	**0.0004**	**0.0006**	0.9893	-0.0107	**0.0006**	**0.0007**	**0.0013**	45	0.3739
	GA	0.0151	0.0151	0.0005	0.0007	0.9886	-0.0114	0.0006	0.0007	0.0015	74	0.3734
	NM	0.1059	0.1059	0.0044	0.0156	0.9211	-0.0789	0.0027	0.0089	0.0245	57	0.0148
	QN	**0.0039**	**0.0039**	0.0009	0.0009	**0.9969**	**-0.0031**	0.0008	0.0008	0.0017	4	**0.0071**

Table 2: Simulated Mean, Bias, Variance, MSE, and Def values when μ=1, σ=1.

n	Algoritma $\mu=1, \sigma=1$	$\hat{\mu}$ Mean	Bias	Variance	MSE	$\hat{\sigma}$ Mean	Bias	Variance	MSE	Def	$\overline{Iter}$	$\overline{Time}$
10	PSO	0.4500	-0.5500	0.2898	0.5923	1.1968	0.1968	0.0854	0.1242	0.7165	43.3835	0.2431
	GA	**0.6094**	**-0.3906**	0.3007	**0.4533**	**1.1014**	**0.1014**	0.0792	**0.0894**	**0.5427**	77.2170	0.2891
	NM	-0.5045	-1.5045	0.2256	2.4892	1.8483	0.8483	0.1043	0.8239	3.3131	64.3280	0.0057
	QN	-0.5808	-1.5808	0.1926	2.6916	1.9026	0.9026	**0.0850**	0.8998	3.5914	4.3280	0.0049
20	PSO	0.3875	-0.6125	0.2657	0.6409	1.2507	0.2507	0.0792	0.1421	0.7830	42.8830	0.2514
	GA	**0.6024**	**-0.3976**	0.2843	**0.4424**	**1.1237**	**0.1237**	0.0765	**0.0918**	**0.5342**	76.3650	0.3005
	NM	-0.5507	-1.5507	0.1074	2.5121	1.8947	0.8947	0.0592	0.8598	3.3719	63.3000	0.0060
	QN	-0.6612	-1.6612	0.0596	2.8191	1.9746	0.9746	0.0282	0.9781	3.7972	4.1475	0.0050
30	PSO	0.3779	-0.6221	0.2664	0.6534	1.2609	0.2609	0.0811	0.1492	0.8026	42.6250	0.2670
	GA	**0.6226**	**-0.3774**	0.2872	**0.4296**	**1.1168**	**0.1168**	0.0801	**0.0937**	**0.5233**	77.2630	0.3260
	NM	-0.5242	-1.5242	0.1002	2.4234	1.8788	0.8788	0.0564	0.8287	3.2521	62.1020	0.0061
	QN	-0.6712	-1.6712	0.0368	2.8297	1.9857	0.9857	0.0173	0.9889	3.8185	4.3190	0.0053
50	PSO	0.3169	-0.6831	0.2440	0.7106	1.3019	0.3019	0.0737	0.1649	0.8755	41.8845	0.2575
	GA	**0.6312**	**-0.3688**	0.2880	**0.4240**	**1.1182**	**0.1182**	0.0819	**0.0959**	**0.5199**	76.6295	0.3225
	NM	-0.4551	-1.4551	0.1161	2.2332	1.8298	0.8298	0.0673	0.7558	2.9890	59.2225	0.0068
	QN	-0.6908	-1.6908	0.0208	2.8795	2.0015	1.0015	0.0101	1.0131	3.8926	6.3730	0.0057
100	PSO	0.2844	-0.7156	0.2344	0.7465	1.3264	0.3264	0.0736	0.1801	0.9266	41.6075	0.2813
	GA	**0.6477**	**-0.3523**	0.2927	**0.4169**	**1.1144**	**0.1144**	0.0889	**0.1020**	**0.5188**	77.7595	0.3571
	NM	-0.3236	-1.3236	0.1245	1.8763	1.7356	0.7356	0.0712	0.6122	2.4886	53.9095	0.0096
	QN	-0.7010	-1.7010	0.0106	2.9038	2.0114	1.0114	0.0049	1.0278	3.9316	4.9035	0.0062
250	PSO	0.1828	-0.8172	0.1703	0.8382	1.3906	0.3906	0.0561	0.2087	1.0469	40.4855	0.3234
	GA	**0.6684**	**-0.3316**	0.2987	**0.4086**	**1.1082**	**0.1082**	0.0962	**0.1079**	**0.5165**	78.1880	0.4149
	NM	-0.1810	-1.1810	0.1013	1.4962	1.6320	0.6320	0.0557	0.4552	1.9513	48.7935	0.0118
	QN	-0.7060	-1.7060	0.0039	2.9142	2.0167	1.0167	0.0018	1.0355	3.9497	4.9205	0.0065
500	PSO	0.1259	-0.8741	0.1239	0.8879	1.4256	0.4256	0.0419	0.2230	1.1109	39.7900	0.3949
	GA	**0.6793**	**-0.3207**	0.2936	**0.3964**	**1.1025**	**0.1025**	0.0956	**0.1062**	**0.5026**	78.7870	0.5216
	NM	-0.1494	-1.1494	0.0900	1.4111	1.6094	0.6094	0.0492	0.4205	1.8316	47.2120	0.0151
	QN	-0.7077	-1.7077	0.0020	2.9183	2.0188	1.0188	0.0010	1.0390	3.9573	5.4170	0.0069

Table 3: Simulated Mean, Bias, Variance, MSE, and Def values when μ=1, σ=2.

n	Algoritma $\mu = 1, \sigma = 2$	$\hat{\mu}$ Mean	Bias	Variance	MSE	$\hat{\sigma}$ Mean	Bias	Variance	MSE	Def	$\overline{Iter}$	$\overline{Time}$
10	PSO	**1.3783**	**0.3783**	0.2818	0.4249	1.5433	-0.4567	0.1424	0.3510	0.7759	48.1750	0.2485
	GA	1.4286	0.4286	0.1916	**0.3753**	1.5089	-0.4911	0.1019	**0.3431**	**0.7184**	69.7605	0.2671
	NM	0.3223	-0.6777	1.0498	1.5092	**2.2860**	**0.2860**	0.5129	0.5947	2.1038	68.4380	0.0060
	QN	0.2524	**-0.7476**	1.0595	1.6184	2.3367	0.3367	0.5200	0.6334	2.2518	5.4490	0.0058
20	PSO	**1.3193**	**0.3193**	0.1612	0.2631	1.5975	-0.4025	0.0813	**0.2433**	0.5064	47.7480	0.2906
	GA	1.3799	0.3799	0.0873	**0.2317**	**1.5542**	**-0.4458**	0.0452	0.2440	**0.4757**	69.1095	0.3026
	NM	0.0136	-0.9864	0.5655	1.5385	2.5386	0.5386	0.2964	0.5865	2.1250	65.4665	0.0061
	QN	-0.0872	-1.0872	0.5567	1.7387	2.6143	0.6143	0.2894	0.6668	2.4055	5.0940	0.0053
30	PSO	**1.3017**	**0.3017**	0.1312	0.2223	**1.6081**	**-0.3919**	0.0704	0.2240	0.4462	47.4780	0.3168
	GA	1.3582	0.3582	0.0612	**0.1895**	1.5671	-0.4329	0.0345	**0.2219**	**0.4114**	69.3300	0.3255
	NM	-0.1420	-1.1420	0.2893	1.5935	2.6577	0.6577	0.1545	0.5870	2.1805	63.2035	0.0058
	QN	-0.2555	-1.2555	0.2629	1.8392	2.7437	0.7437	0.1364	0.6894	2.5286	5.1630	0.0053
50	PSO	**1.2734**	**0.2734**	0.1119	0.1866	**1.6297**	**-0.3703**	0.0621	0.1993	0.3859	47.7395	0.3423
	GA	1.3320	0.3320	0.0353	**0.1455**	1.5870	-0.4130	0.0205	**0.1910**	**0.3365**	69.1030	0.3428
	NM	-0.2055	-1.2055	0.1187	1.5720	2.7102	0.7102	0.0655	0.5699	2.1419	61.6825	0.0066
	QN	-0.3449	-1.3449	0.0989	1.9077	2.8165	0.8165	0.0515	0.7182	2.6258	5.2395	0.0055
100	PSO	**1.2695**	**0.2695**	0.0703	0.1430	**1.6279**	**-0.3721**	0.0389	0.1773	0.3203	47.4200	0.3721
	GA	1.3121	0.3121	0.0188	**0.1162**	1.5969	-0.4031	0.0113	**0.1738**	**0.2899**	68.8375	0.3629
	NM	-0.1998	-1.1998	0.0550	1.4944	2.7033	0.7033	0.0331	0.5277	2.0222	58.7150	0.0079
	QN	-0.3829	-1.3829	0.0299	1.9424	2.8435	0.8435	0.0148	0.7263	2.6687	5.8740	0.0060
250	PSO	**1.2516**	**0.2516**	0.0548	0.1181	**1.6406**	**-0.3594**	0.0295	0.1587	0.2768	47.6015	0.4519
	GA	1.2876	0.2876	0.0089	**0.0916**	1.6143	-0.3857	0.0052	**0.1540**	**0.2456**	69.1105	0.4229
	NM	-0.0970	-1.0970	0.0340	1.2373	2.6269	0.6269	0.0215	0.4146	1.6519	53.0555	0.0112
	QN	-0.3864	-1.3864	0.0106	1.9328	2.8491	0.8491	0.0058	0.7267	2.6595	5.8620	0.0064
500	PSO	**1.2404**	**0.2404**	0.0511	0.1089	**1.6481**	**-0.3519**	0.0275	0.1513	0.2602	47.3030	0.8843
	GA	1.2773	0.2773	0.0050	**0.0819**	1.6212	-0.3788	0.0028	**0.1463**	**0.2282**	69.2405	0.7616
	NM	-0.0343	-1.0343	0.0201	1.0898	2.5785	0.5785	0.0125	0.3472	1.4370	49.3285	0.0210
	QN	-0.3907	-1.3907	0.0057	1.9398	2.8516	0.8516	0.0033	0.7286	2.6685	5.9590	0.0075

Table 4: Simulated Mean, Bias, Variance, MSE, and Def values when μ=2, σ=1.

n	Algoritma $\mu=2, \sigma=1$	$\hat{\mu}$ Mean	Bias	Variance	MSE	$\hat{\sigma}$ Mean	Bias	Variance	MSE	Def	$\overline{Iter}$	$\overline{Time}$
10	PSO	0.1709	-1.8291	0.2147	3.5604	1.7970	0.7970	0.1078	0.7430	4.3035	41.4535	0.2603
	GA	**0.2539**	**-1.7461**	0.3019	**3.3508**	**1.7421**	**0.7421**	0.1172	**0.6679**	**4.0187**	79.9150	0.3025
	NM	-0.5220	-2.5220	0.1917	6.5520	2.2890	1.2890	0.1270	1.7886	8.3406	65.9750	0.0060
	QN	-0.5396	-2.5396	0.1839	6.6335	2.3015	1.3015	0.1242	1.8180	8.4515	5.5580	0.0053
20	PSO	0.0516	-1.9484	0.0412	3.8374	**1.8842**	**0.8842**	0.0429	0.8246	4.6620	40.4100	0.1811
	GA	**0.1124**	**-1.8876**	0.1035	**3.6666**	1.8446	0.8446	0.0558	**0.7692**	**4.4357**	82.2830	0.2526
	NM	-0.6053	-2.6053	0.0817	6.8690	2.3591	1.3591	0.0600	1.9072	8.7762	66.7145	0.0062
	QN	-0.6337	-2.6337	0.0685	7.0048	2.3797	1.3797	0.0525	1.9561	8.9608	4.7425	0.0058
30	PSO	0.0274	-1.9726	0.0171	3.9082	1.9142	0.9142	0.0271	0.8629	4.7711	40.2390	0.1829
	GA	**0.0675**	**-1.9325**	0.0432	**3.7776**	**1.8874**	**0.8874**	0.0345	**0.8221**	**4.5997**	83.0015	0.2547
	NM	-0.6148	-2.6148	0.0664	6.9035	2.3810	1.3810	0.0463	1.9535	8.8570	67.1320	0.0061
	QN	-0.6591	-2.6591	0.0439	7.1147	2.4131	1.4131	0.0339	2.0307	9.1455	4.6910	0.0053
50	PSO	0.0126	-1.9874	0.0088	3.9586	1.9228	0.9228	0.0167	0.8682	4.8268	40.2900	0.2006
	GA	**0.0324**	**-1.9676**	0.0172	**3.8888**	**1.9089**	**0.9089**	0.0189	**0.8450**	**4.7338**	85.5275	0.2769
	NM	-0.6277	-2.6277	0.0550	6.9598	2.3891	1.3891	0.0358	1.9653	8.9250	66.2955	0.0071
	QN	-0.6815	-2.6815	0.0248	7.2150	2.4281	1.4281	0.0194	2.0590	9.2740	4.7130	0.0059
100	PSO	0.0036	-1.9964	0.0003	3.9861	1.9304	0.9304	0.0076	0.8732	4.8594	40.1990	0.2143
	GA	**0.0167**	**-1.9833**	0.0044	**3.9378**	**1.9212**	**0.9212**	0.0087	**0.8573**	**4.7951**	86.6155	0.3022
	NM	-0.5835	-2.5835	0.0727	6.7474	2.3591	1.3591	0.0423	1.8894	8.6367	64.9055	0.0075
	QN	-0.6894	-2.6894	0.0130	7.2459	2.4363	1.4363	0.0098	2.0726	9.3186	6.2960	0.0063
250	PSO	0.0023	-1.9977	0.0023	3.9930	1.9343	0.9343	0.0036	0.8766	4.8696	40.1365	0.2761
	GA	**0.0037**	**-1.9963**	0.0009	**3.9862**	**1.9332**	**0.9332**	0.0034	**0.8742**	**4.8604**	88.9490	0.3809
	NM	-0.4700	-2.4700	0.1113	6.2123	2.2792	1.2792	0.0609	1.6974	7.9096	61.1335	0.0096
	QN	-0.6978	-2.6978	0.0048	7.2827	2.4455	1.4455	0.0040	2.0933	9.3761	6.1940	0.0066
500	PSO	0.0016	-1.9984	0.0018	3.9955	1.9334	0.9334	0.0021	0.8734	4.8688	40.1100	0.3472
	GA	**0.0020**	**-1.9980**	0.0021	**3.9942**	**1.9332**	**0.9332**	0.0020	**0.8729**	**4.8671**	90.6585	0.4587
	NM	-0.3702	-2.3702	0.1265	5.7445	2.2050	1.2050	0.0685	1.5206	7.2651	57.5255	0.0131
	QN	-0.7000	-2.7000	0.0025	7.2927	2.4458	1.4458	0.0021	2.0925	9.3852	6.1360	0.0072

Table 5: Simulated Mean, Bias, Variance, MSE, and Def values when μ=2, σ=2.

n	Algoritma $\mu = 2, \sigma = 2$	$\hat{\mu}$				$\hat{\sigma}$						
		Mean	Bias	Variance	MSE	Mean	Bias	Variance	MSE	Def	$\overline{Iter}$	$\overline{Time}$
10	PSO	1.1104	-0.8896	1.1697	1.9611	**1.9193**	**-0.0807**	0.2730	0.2795	2.2406	43.0340	0.2512
	GA	**1.4029**	**-0.5971**	1.0497	**1.4063**	1.7596	-0.2404	0.2120	**0.2699**	**1.6761**	73.5635	0.2893
	NM	-0.9880	-2.9880	0.5801	9.5086	3.2982	1.2982	0.2516	1.9368	11.4454	68.2965	0.0061
	QN	-1.1206	-3.1206	0.4806	10.2190	3.3928	1.3928	0.1904	2.1302	12.3492	5.3160	0.0050
20	PSO	1.1329	-0.8671	1.1462	1.8980	**1.9464**	**-0.0536**	0.2733	0.2762	2.1742	42.6110	0.2623
	GA	**1.5250**	**-0.4750**	0.9274	**1.1530**	1.7326	-0.2674	0.1925	**0.2641**	**1.4171**	72.4415	0.3032
	NM	-1.0106	-3.0106	0.3514	9.4148	3.3391	1.3391	0.1806	1.9739	11.3888	66.4270	0.0065
	QN	-1.2463	-3.2463	0.1494	10.6878	3.5092	1.5092	0.0565	2.3343	13.0221	6.6040	0.0055
30	PSO	1.1288	-0.8712	1.1566	1.9156	**1.9620**	**-0.0380**	0.2804	0.2818	2.1974	42.4375	0.2665
	GA	**1.5645**	**-0.4355**	0.9138	**1.1035**	1.7286	-0.2714	0.2022	**0.2758**	**1.3793**	71.7285	0.3079
	NM	-0.9650	-2.9650	0.3626	9.1539	3.3104	1.3104	0.2008	1.9179	11.0718	65.2705	0.0068
	QN	-1.2858	-3.2858	0.0883	10.8849	3.5425	1.5425	0.0339	2.4133	13.2982	5.2700	0.0053
50	PSO	1.0947	-0.9053	1.1844	2.0039	**1.9896**	**-0.0104**	0.2978	0.2979	2.3018	42.4375	0.2800
	GA	**1.7020**	**-0.2980**	0.7911	**0.8799**	1.6668	-0.3332	0.1821	**0.2931**	**1.1730**	71.3795	0.3242
	NM	-0.8444	-2.8444	0.4063	8.4966	3.2248	1.2248	0.2250	1.7251	10.2217	62.8565	0.0078
	QN	-1.2953	-3.2953	0.0572	10.9159	3.5524	1.5524	0.0198	2.4297	13.3455	5.1490	0.0067
100	PSO	1.1525	-0.8475	1.1959	1.9141	**1.9742**	**-0.0258**	0.3152	0.3158	2.2299	42.3335	0.3019
	GA	**1.8120**	**-0.1880**	0.6786	**0.7139**	1.6246	-0.3754	0.1682	**0.3092**	**1.0231**	70.9385	0.3466
	NM	-0.5677	-2.5677	0.4418	7.0349	3.0277	1.0277	0.2424	1.2986	8.3335	57.0310	0.0087
	QN	-1.3174	-3.3174	0.0260	11.0313	3.5743	1.5743	0.0105	2.4891	13.5204	5.3150	0.0057
250	PSO	1.1798	-0.8202	1.1820	1.8548	**1.9645**	**-0.0355**	0.3259	0.3271	2.1819	42.6730	0.3734
	GA	**1.9462**	**-0.0538**	0.4637	**0.4666**	1.5559	-0.4441	0.1218	**0.3190**	**0.7856**	69.8560	0.4079
	NM	-0.2944	-2.2944	0.3233	5.5878	2.8282	0.8282	0.1779	0.8638	6.4516	52.2095	0.0125
	QN	-1.3237	-3.3237	0.0110	11.0576	3.5789	1.5789	0.0051	2.4981	13.5558	5.6160	0.0063
500	PSO	1.1791	-0.8209	1.1786	1.8524	**1.9669**	**-0.0331**	0.3318	0.3329	2.1853	42.0765	0.4807
	GA	**2.0160**	**0.0160**	0.3264	**0.3266**	1.5192	-0.4808	0.0904	**0.3216**	**0.6482**	69.2880	0.5222
	NM	-0.2240	-2.2240	0.2717	5.2178	2.7751	0.7751	0.1449	0.7457	5.9634	50.8045	0.0160
	QN	-1.3241	-3.3241	0.0087	11.0587	3.5776	1.5776	0.0095	2.4983	13.5569	5.7860	0.0067

Table 6: The descriptive statistics for aircraft windshield service time data.

n	*Min*	1st Qu.	Mean	Mode	Median	3rd Qu.	*Max*	S^2	γ_1	γ_2
63	0.0460	1.1070	2.0853	0.0460	2.0650	2.8197	5.1400	1.5506	0.4396	2.7326

Table 7: The parameter estimates, *LL*, AIC, AICc, values for **a**ircraft windshield service time data**.**

Method	$\hat{\mu}$	$\hat{\sigma}$	-LL	AIC	AIC_C
PSO	0	1.9789	**115.4526**	**234.9052**	**235.1052**
GA	0	1.9789	**115.4535**	**234.9070**	**235.1070**
NM	0.3638	1.7301	125.3900	254.7800	254.9800
QN	0.0413	2.0310	122.7079	249.4158	249.6158

Table 8: The descriptive statistics for the vinyl chloride data.

n	*Min*	1^{st} Qu.	Mean	Mode	Median	3^{rd} Qu.	*Max*	S^2	γ_1	γ_2
34	0.1000	0.5000	1.8794	0.4000	1.1500	2.500	8.000	3.8126	1.6037	5.0054

Table 9: The parameter estimates, LL, AIC, AICc, values for the vinyl chloride data.

Method	$\hat{\mu}$	$\hat{\sigma}$	*-LL*	AIC	AIC_C
PSO	0	2.1959	**97.3287**	**198.6574**	**199.0445**
GA	0	2.1959	**97.3290**	**198.6580**	**199.0451**
NM	0.2638	2.0511	113.6900	271.3800	271.7671
QN	0.0679	1.0473	197.6675	399.3350	399.7221