Thailand Statistician
April 2021; 19(2): 308-316
http://statassoc.or.th
Contributed paper

Reliability Estimation of Three Parameters Gamma Distribution via Particle Swarm Optimization

Shaimaa Waleed Mahmood and Zakariya Yahya Algamal*
Department of Statistics and Informatics, College of Computer science & Mathematics,
University of Mosul, Mosul, Iraq.
*Corresponding author; e-mail: zakariya.algamal@uomosul.edu.iq

Received: 16 January 2020
Revised: 1 February 2020
Accepted: 27 April 2020

Abstract

Reliability analysis is considered as one of the most used approaches in various real data applications. Usually, the reliability function is based on a statistical distribution. The three-parameter gamma continuous distribution is a widely used in the study of reliability. A lot of attention has been considered on theirs parameter estimation. In this paper, a particle swarm optimization (PSO) is proposed to estimate the three-parameter gamma distribution and then to estimate the reliability and hazard functions. The real data results demonstrate that our proposed estimation method is considerably consistent in estimation compared to the maximum likelihood estimation method, in terms of log likelihood and mean time to failure (MTTF).

Keywords: Reliability analysis, swarm optimization algorithm, maximum likelihood estimation method.

1. Introduction

Attention has begun to the object reliability since half century specifically after the Second World War, studies and theories have been rolled to become the object of reliability independent has foundations, theories and applications in the scientific life. Then, recently there have been considerable methods developed in field of reliability engineering in order to help the management in determining the reliability function of their equipment's and combined this function with the maintenance and replacements methods.

The concept of the reliability is ability the device or machine to make the processes without failure. But the concept of reliability is statistically the probability that device or machine work to fulfill a certain work for a span of time until the breakdown has occurred. The study of reliability need to the study breakdowns and stops the machines and equipments which is need to describe the times of machines life, then it get on data set to one machine represent life or system of machines.

The gamma distribution is one from the continuous probability distributions which is use in reliability and other application. It is used as distribution for service times and waiting times (Whitt 2000). The swarm intelligence based techniques have successfully been applied for many engineering optimization problems as these techniques processes search speed in finding optimized result for such applications.

The particle swarm optimization is a population based search algorithm which was inspired by the collective behavior of swarm (Zhu et al. 2009).

In this paper, a three-parameter gamma distribution is used and a particle swarm optimization (PSO) algorithm is employed to estimate the reliability function with application on several datasets. Further, a comparison is made with the maximum likelihood estimation method.

2. The Gamma Distribution

The gamma distribution is an important continuous distributions which is widely used in reliability applications and test life. This distribution consider by Stacy (1962). Can be in two types the one is two parameters and other three parameters. The probability density function for two parameters is (Kirimi et al. 2014)

$$f(t)=\frac{t^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp\left(-\frac{t}{\beta}\right), \quad \alpha>0,\ \beta>0. \tag{1}$$

The three parameters of gamma distribution has probability density function as (Chen and Kotz 2013)

$$f(t)=\frac{(t-\gamma)^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp\left(-\frac{t-\gamma}{\beta}\right), \alpha>0,\ \beta>0,\ -\infty<\gamma<\infty,\ \ t>\gamma, \tag{2}$$

where α represent the shape parameter, β represent the scale parameter, γ represent the location parameter (threshold), and $\Gamma(\cdot)$ is the gamma function. The cumulative distribution function and the expectation are

$$F(t)=\frac{\Gamma_{\frac{(t-\gamma)}{\beta}}(\alpha)}{\Gamma(\alpha)}, \tag{3}$$

and

$$E(t)=\gamma+\alpha\beta. \tag{4}$$

3. Reliability Concepts

3.1. Reliability function

Known as probability of not failure of the machine to time t where $t>0$. The extensive meaning for reliability which scale of performance. From the scientific concept the reliability function has the same mathematical concept for the survived function, which use the concept in the studying be alive in community of living organisms. Let T is a random variable nonnegative as failure time and has probability density function although about cumulative probability function as (Bakar et al. 2002)

$$R(t)=P(T>t), \quad 0<t<\infty, \tag{5}$$

where $R(t)$ is the reliability function, and can be rewrite Equation (5),

$$R(t)=1-P(T\le t)=1-F(t). \tag{6}$$

Then the reliability function for three parameters gamma distribution is

$$R(t)=1-F(t)=1-\frac{\Gamma_{\frac{(t-\gamma)}{\beta}}(\alpha)}{\Gamma(\alpha)}. \tag{7}$$

3.2. Failure function

Probability the machine failure during interval $\{t<T<t+\Delta t\}$ which probability not successful during the same interval and denote to $f(t)$ and it's

$$f\left(t\right)=\lim_{\Delta t\to 0}\frac{P_r(t<T<t+\Delta t)}{\Delta t}. \tag{8}$$

3.3. Hazard function

The mathematical definition for hazard function or failure rate is (Lawless 2003)

$$h(t)=\lim_{\Delta t\to 0}\frac{P_r\left(t<T<t+\Delta t|T>t\right)}{\Delta t}=\frac{f(t)}{R(t)}, \tag{1}$$

where $h(t)$ is the hazard function. Then the hazard function for this distribution and mean time to failure (MTTF) are

$$h(t)=\frac{\frac{(t-\gamma)^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp-\left(\frac{t-\gamma}{\beta}\right)}{1-\frac{\Gamma_{\frac{(t-\gamma)}{\beta}}(\alpha)}{\Gamma(\alpha)}}, \tag{2}$$

and

$$MTTF=E(t). \tag{3}$$

4. Maximum Likelihood Method for Estimation of the Reliability Function

There are many used methods in estimating of the reliability function of which parametric methods as in ML or using unbiased estimator uniformly minimum variance. When use these methods for reliability function estimation you need to knowledge probability distribution for failure models, almost the using of extension distributions in the failure models is (normal, gamma, Weibull and exponential distributions). The other methods for estimating the reliability function are nonparametric methods and can use intelligent techniques in estimate this function. There are many references which explain parametric methods are (Cohen 1965, Li 1984, Pugh 1963), and the maximum likelihood method is one of important parametric estimation methods which aim to make likelihood function for the variables in the end of maximum. Then the likelihood function for this distribution is (Bowman and Shenton 2002)

$$L\left(t;\alpha,\beta,\gamma\right)=\frac{\sum_{i=1}^{n}(t_i-\gamma)^{\alpha-1}}{\beta^{n\alpha}\left(\Gamma(\alpha)\right)^{n}}\exp-\sum_{i=1}^{n}\left(\frac{t_i-\gamma}{\beta}\right). \tag{12}$$

If we take the logarithm, we get on

$$\ln L=\left(\alpha-1\right)\sum_{i=1}^{n}\ln\left(t_i-\gamma\right)-n\alpha\ln\beta-n\ln\Gamma\left(\alpha\right)-\sum_{i=1}^{n}\left(\frac{t_i-\gamma}{\beta}\right). \tag{13}$$

To find the maximum likelihood estimator to α,β,γ parameters derived it (13) with respect to α,β,γ and equal it with zero it get on

$$\frac{\partial\ln L}{\partial\alpha}=\sum_{i=1}^{n}\ln\left(t_i-\gamma\right)-n\ln\beta-n\frac{\Gamma'\left(\alpha\right)}{\Gamma\left(\alpha\right)}, \tag{14}$$

$$\frac{\partial\ln L}{\partial\beta}=\sum_{i=1}^{n}\left(\frac{t_i-\gamma}{\beta^2}\right)-\frac{n\alpha}{\beta}, \tag{15}$$

$$\frac{\partial\ln L}{\partial\gamma}=-(\alpha-1)\sum_{i=1}^{n}\frac{1}{(t_i-\gamma)}+\frac{n}{\beta}. \tag{16}$$

5. Particle Swarm Optimization Algorithm

The particle swarm optimization which is one of the intelligent techniques to solve optimization problems and this algorithm consider by Kennedy and Eberhart (1995). Let an unconstrained maximization problem (Bai 2010)

$$\text{maximize } f(x),\ \ X^l \le X \le X^u,$$

whereas X^l and X^u indicated the lower and upper bounds on X. The steps of the PSO algorithm can be implement as follows:

1) Let the size of the swarm number of particles is N. To less the total number of function evaluations need to find a solution, we must suppose a smaller of the swarm. But in this case you take long time to find the perfect solution. Usually a size of 20 to 30 particles is supposed for the swarm as compromise.

2) Generate the initial population of X in the range $\left[X^l, X^u\right]$ randomly like $X_1, X_2, ..., X_N$, after that for fit the position of j and its velocity in iteration i are indicated $X_j(i)$ and $V_j(i)$. Thus the particle generated initially are indicated $X_1(0), X_2(0), .., X_N(0)$. The vectors $X_j(0), j = 1, 2, \ldots, N$ are called particles. Then evaluate the objective function values corresponding to the particles like $f\left[X_1(0)\right], f\left[X_2(0)\right], ..., f\left[X_N(0)\right]$.

3) Find the velocity of particles. All particles will be moving to the optimal point with velocity. Initially, all particles velocity are supposed to be 0. Let the iteration number $i = 1$.

4) In the i^{th} iteration, we find the two important parameters used by particle j.

a) The best position for the particle.

b) Find the velocity of particle j in the i^{th} iteration as in

$$V_j(i) = V_j(i-1) + c_1 r_1\left[P_{best} - X_j(i-1)\right] + c_2 r_2\left[G_{best} - X_j(i-1)\right],\ j = 1, 2, \ldots, N, \quad (17)$$

so that

$V_j(i)$: As velocity particle in the i^{th} iteration,

c_1, c_2 : As acceleration coefficients and usually take value 2,

r_1, r_2 : As random values in the range 0 to 1,

P_{best} : As best position to the particle swarm,

G_{best} : As best position to the particle includes all swarm.

c) Find the position of the j^{th} particle in the i^{th} iteration,

$$X_j(i) = X_j(i-1) + V_j(i),\ j = 1, 2, \ldots, N. \quad (4)$$

Then, evaluate the objective function values corresponding to the particle as

$$f\left[X_1(i)\right], f\left[X_2(i)\right], \ldots, f\left[X_N(i)\right].$$

5) Test the convergence of the current solution. If the position of all particles converge to itself set of values, the method is supposed to have converged. If the convergence criterion is not got on, Step 4 is repeated with updating the iteration number to be $i = i + 1$, and by computing the new values of P_{best} and G_{best}. The iterative process is continued until all particles converge to the same optimum solution.

6. The Proposed Method

In this section, we propose to use the particle swarm optimization algorithm to estimate the reliability function of three-parameter gamma distribution. The proposed fitness is defined as

$$-\ln L = -\left((\alpha-1)\sum_{i=1}^{n}\ln(t_i-\gamma) - n\alpha\ln\beta - n\ln\Gamma(\alpha) - \sum_{i=1}^{n}\left(\frac{t_i-\gamma}{\beta}\right)\right), \tag{19}$$

where the parameters α,β and γ is the search variable. The aim is to find the values of parameters α,β and γ such that the function $-\ln L$ is minimum.

Let $X_j(i)$ is the position vector of particle j in the multidimensional search space in step (i), where $X_j(i)=(\alpha \;\; \beta \;\; \gamma)$, then the procedures of the solution of PSO are illustrated in the following steps:

1) Let the size of the swarm number of particles is $N=50$.

2) Generate the initial population of X in the range $\left[X^l, X^u\right]$ randomly,

$$X_j(0) = X^l + rand(\bullet)\left(X^u - X^l\right),\; i=1,2,\ldots,N. \tag{20}$$

where $X^l = \left[\alpha^l\,\beta^l\,\gamma^l\right]$ and $X^u = \left[\alpha^u\,\beta^u\,\gamma^u\right]$ are lower and upper bounds of parameters $(\alpha \;\; \beta \;\; \gamma)$, here $0<\alpha<500,\; 0<\beta<500$ and $0<\gamma<2.5$ for Data 1, $0<\gamma<0.25$ for Data 2 and $0<\gamma<1.6667$ for Data 3, and $rand(\bullet)$ random numbers in the range 0 and 1. Then evaluate the objective function values corresponding to the particles as $-\ln L\left[X_1(0)\right], -\ln L\left[X_2(0)\right]$, $.., -\ln L\left[X_N(0)\right]$.

3) Find the velocity of particles. All particles will be moving to the optimal point with velocity. Initially, all particles velocity are supposed to be 0. Let the iteration number $i=1$.

4) In the i^{th} iteration, we find the two important parameters used by particle j.

a) The best position for the particle.

b) Find the velocity of particle j in the i^{th} iteration as in

$$V_j(i) = wV_j(i-1) + c_1 r_1\left[P_{best} - X_j(i-1)\right] + c_2 r_2\left[G_{best} - X_j(i-1)\right],\; j=1,2,\ldots,N,$$

where

w : intertia weight,

$V_j(i)$: As velocity particle in the i^{th} iteration,

c_1,c_2 : As acceleration coefficients and usually take value 2,

r_1,r_2 : As random values in the range 0 to 1,

P_{best} : As best position to the particle swarm (the less value of the objective function),

G_{best} : As best position to the particle include all swarm (the less value of the objective function).

The value of inertia weight decreases linearly with the iteration number has been used,

$$w(i) = w_{max} - \left(\frac{w_{max} - w_{min}}{i_{max}}\right) i, \tag{21}$$

where w_{max} and w_{min} are the initial and final values of the inertia weight respectively. The values of w_{max} and w_{min} are usually assumed to be 0.9 and 0.4 respectively, and i_{max} is the maximum number of iterations.

c) Find the position of the j^{th} particle in the i^{th} iteration

$$X_j(i) = X_j(i-1) + V_j(i),\ j = 1,2,...,N.$$

Then, evaluate the objective function values corresponding to the particle as

$$-\ln L[X_1(i)], -\ln L[X_2(i)], ..., -\ln L[X_N(i)].$$

5) Test the convergence of the current solution. If the position of all particles converge to itself set of values, the method is supposed to have converged. If the convergence criterion is not got on, Step 4 is repeated with updating the iteration number to be $i = i+1$, and by computing the new values of P_{best} and G_{best}. The iterative process is continued until all particles converge to the same optimum solution.

7. Application

In this paper, three datasets with different sample size are used, First, the operating times (hours) for the machine. Second, the operating times (hours) of the construction machine first phase (Saffawy and Algamal 2006). Third, the stops times (hours) of the oven (Jassim 2013). To test that if the three datasets follow three parameters gamma distribution according to the hypothesis and the Kolmogorov-Smirnov (K-S) goodness of fit. The null and alternative hypotheses are given by

H_0 : Data follow three parameters gamma distribution

H_1 : Data not follow three parameters gamma distribution.

Depending on the Kolmogorov-Smirnov goodness of fit test, Table1 shows that the used datasets belong to the three parameters gamma distribution under significant level of 0.05.

Table 1 Test the fit of the data to three parameters gamma distribution

Data	n	K-S test	Critical value	p-value
Data 1	29	0.1164	0.2457	0.7847
Data 2	33	0.1438	0.2308	0.4604
Data 3	84	0.0927	0.1461	0.4392

Using maximum likelihood method and the proposed PSO algorithm, Table 2 summarized the parameters estimation values.

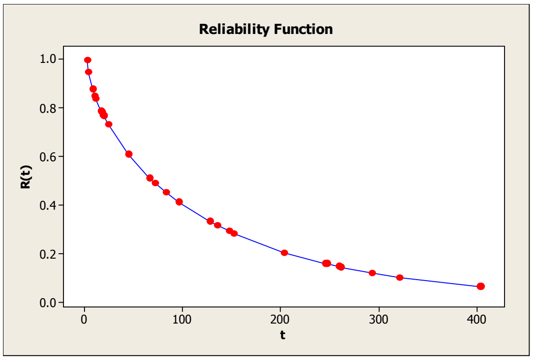
Table 2 ML and PSO estimators of gamma distribution parameters

Data	$\hat{\alpha}_{ML}$	$\hat{\beta}_{ML}$	$\hat{\gamma}_{ML}$	$\hat{\alpha}_{PSO}$	$\hat{\beta}_{PSO}$	$\hat{\gamma}_{PSO}$
Data 1	0.6187	201.2300	2.4900	0.6008	208.9065	2.49000
Data 2	0.8702	109.7400	0.2400	0.8824	110.9388	0.2400
Data 3	0.9104	1.1941	0.1666	0.9396	1.1929	0.1666

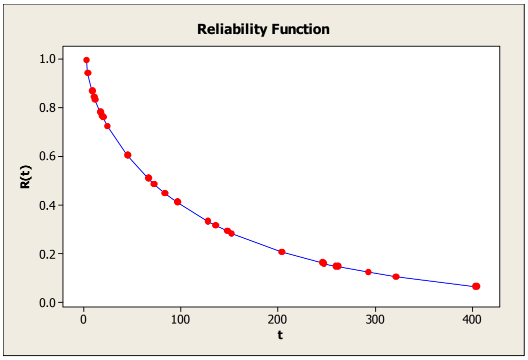
In Table 3, the −log likelihood value is reparation for both ML and the PSO over the three datasets. As we can see from Table 3 the −log likelihood values in proposed PSO is less than the −log likelihood values in ML, then the PSO algorithm is the best comparing with ML in parameters estimation. To find of the reliability function and the hazard function for three parameters gamma distribution of the three datasets the Figures 1 and 2 show them using ML and PSO estimators.

Table 3 −log likelihood values of gamma distribution

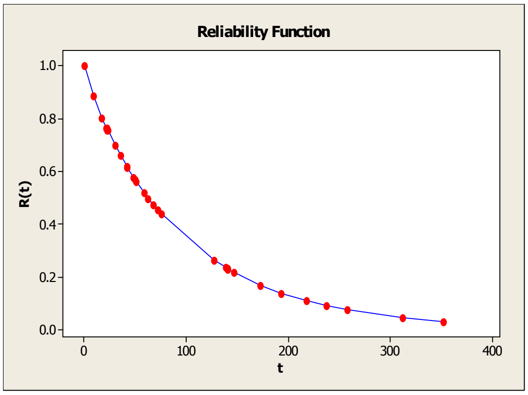
Data	ML	Proposed PSO
Data 1	166.0613	166.0518
Data 2	184.1038	184.0928
Data 3	93.5425	93.4790



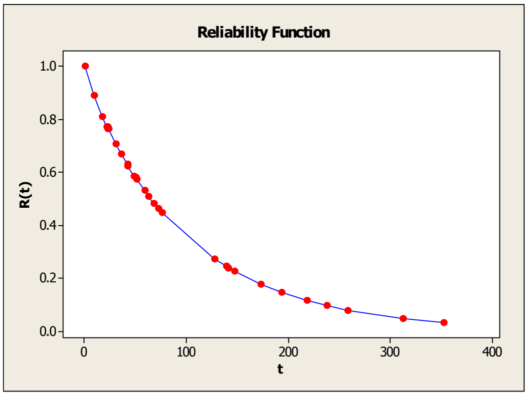
a) ML (Data 1)



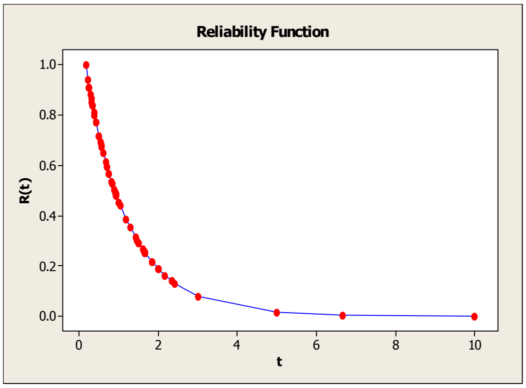
b) PSO (Data 1)



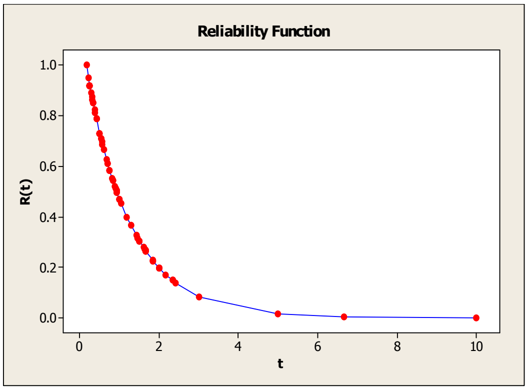
c) ML (Data 2)



d) PSO (Data 2)

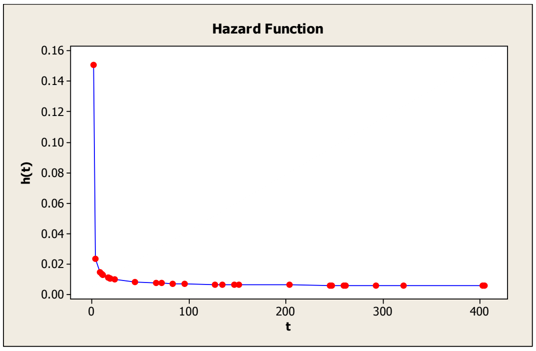


e) ML (Data 3)

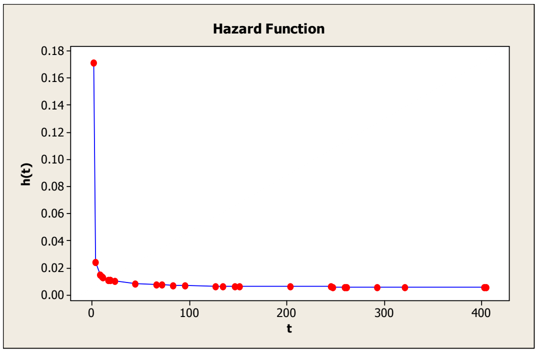


f) PSO (Data 3)

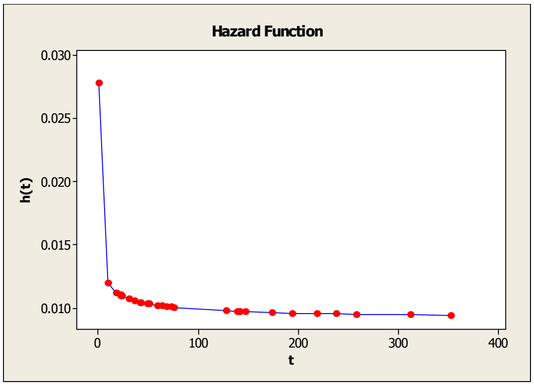
Figure 1 The reliability function using ML and PSO estimators



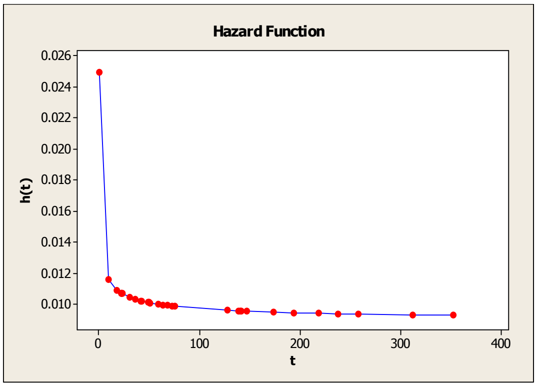
a) ML (Data 1)



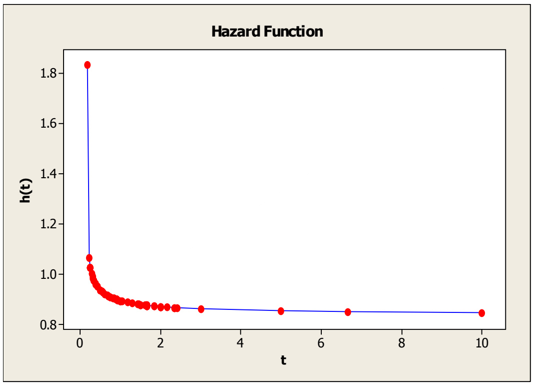
b) PSO (Data 1)



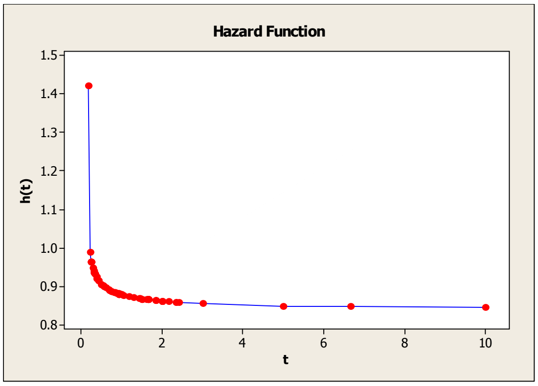
c) ML (Data 2)



d) PSO (Data 2)



e) ML (Data 3)



f) PSO (Data 3)

Figure 2 The hazard function using ML and PSO estimators

We note that the reliability function and the hazard function in the PSO is the best comparing with the reliability function and the hazard function in the ML. Furthermore, the (MTTF) for three datasets to both methods was calculated and reported in Table 4.

Table 4 MTTF values to estimate methods

Data	ML	PSO
Data 1	126.9970	128.0010
Data 2	95.7336	98.1324
Data 3	1.2536	1.2874

We note that MTTF values for three datasets in PSO are higher than ML. This indicates that the PSO algorithm is the best comparing with ML.

8. Conclusions

We note that MTTF values for three datasets in PSO are higher than ML. This indicates that the PSO algorithm is the best comparing with ML.

Acknowledgements

The authors are very grateful to the University of Mosul/College of Computer Sciences and Mathematics for their provided facilities, which helped to improve the quality of this work. We also thank the anonymous reviewers and editor for their helpful comments that led to an improved paper.

References

Bai Q. Analysis of particle swarm optimization algorithm. Comp Info Sci. 2010; 3(1): 180-183.

Bakar D, Bridges D, Hunter R, Johnson G, Krupa J, Murphy J, Sorenson K. Guidebook to decision – making methods. Washington: US Department of Energy; 2002.

Bowman O, Shenton, K. Problems with maximum likelihood estimation and the 3 parameter gamma distribution. Stat Comp Simul. 2002; 72(5): 391-401.

Chen WS, Kotz S. The Riemannian structure of the three-parameter gamma distribution. Appl Math. 2013; 4: 514-522.

Cohen AC. Maximum likelihood estimation in the Weibull distribution based on complete and censored samples. Technometrics. 1965; 7(4): 579-588.

Jassim OR. Estimation of the parameters of the geometric stochastic process and a related process with an application. MS [Thesis]. Iraq: University of Mosul; 2013.

Kennedy J, Eberhart R. Particle swarm optimization. In: The International Conference on Neural Networks; Australia: Perth; 1995. p. 1942-1948.

Kirimi E, Ouko A, Kipkoech CW. Modified moment estimation for a two parameter gamma distribution. IOSR J Math. 2014; 10(6): 42-50.

Lawless JF. Statistical model and methods for life time data. New York: John Wiley & Sons; 2003.

Li TF. Empirical Bayes approach to reliability estimation for the exponential distribution. IEEE Trans Reliab. 1984; 33: 233-236.

Pugh EL. The best estimation of reliability in exponential case. Oper Res. 1963; 11: 57-65.

Saffawy SY, Algamal, Z. The use of maximum likelihood and Kaplan-Meir method to estimate the reliability function: an application on Babylon tires factory. Tanmiyat Al-Rafidain. 2006; 82(28): 9-20.

Stacy EW. A generalization of the gamma distribution. Ann Math Stat. 1962; 33(3): 1187-1192.

Whitt W. The impact of a heavy-tailed service-time distribution upon the M/GI/s waiting-time distribution. Queueing Syst. 2000; 36(1-3): 71-87.

Zhu H, Pu C, Eguchi K, Gu J. Euclidean particle swarm optimization. In: the Second International Conference on Intelligent Networks and Intelligent Systems; 2009 Nov 1; 2009. p. 669-672.