

Generalized gamma distribution based on the Bayesian approach with application to investment modelling

Amani Idris A. Sayed^{1,2,*} and Shamsul Rijal Muhammad Sabri¹

¹ School of Mathematical Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysia

² Department of Mathematics, Jazan University, Jazan, Saudi Arabia

Received: 11 March 2022 / Accepted: 4 August 2023

Abstract. The Generalized Gamma Distribution (GGD) is one of the most popular distributions in analyzing real lifetime datasets. Estimating the parameters of a high dimensional probability distribution is challenging due to the complexities associated with the resulting objectives function. When traditional estimation techniques fail due to complexity in the model objectives function, other powerful computational approaches are employed. In this study, a Bayesian approach to Generalized Gamma Distribution (GGD) based on Markov Chain Monte-Carlo (MCMC) has been employed to estimate model parameters. This study considers the Bayesian approach to GGD parameters using the Adaptive Rejection Metropolis Sampling (ARMS) technique of random variable generation within the Gibbs sampler. The MCMC approach has been used for estimating the multi-dimensional objectives function distribution. The results of the ARMS were compared to the existing Simulated annealing (SA) algorithm and Method of Moment (MM) based on modified internal rate of return data (MIRR). The performances of various derived estimators were recorded using the Markov chain Monte Carlo simulation technique for different sample sizes. The study reveals that ARMS's performance is marginally better than the existing SA and MA approaches. The efficiency of ARMS does not require a larger sample size as the SA does, in the case of simulated data. The performances of ARMS and SA are similar comparing them to the MM as an initial assumption in the case of real MIRR data. However, ARMS gives an acceptable estimated parameter for the different sample sizes due to its ability to evaluate the conditional distributions easily and sample from them exactly.

Keywords: Modified internal rate of return / Markov chain Monte-Carlo / Maximum likelihood / simulated annealing algorithm / method of moment

1 Introduction

The generalized gamma distribution (GGD) family, which includes the subfamilies of exponential, gamma, Weibull, and lognormal as limiting distribution, has been frequently utilized in many research areas. One of the earlier works on GGD was proposed by Amoroso [1]. In the study, a GGD model was used in fitting the income rates of return. A four GGD proposed by Stacy [2] was in Johnson et al. [3] by setting the location parameter to zero. A study by Mudholkar et al. (1993) pioneered the exponentiated approach for estimating the GGD parameters. A study conducted by Agarwal et al. [4] investigated hazard rates using the GGD. Balakrishnan and Peng [5] utilized the GGD to build a GG frailty model. A different version of

GGD was used by Nadarajah and Gupta [6] in fitting drought data sets. A lifetime survival analysis based on GGD was conducted by Pinho et al. [7]. The flexibility of GGD was proposed by Jaggia [8] which was used in Specification tests based on the heterogeneous generalized gamma model of duration [9,10].

Various approaches have been used by various researchers in estimating the parameters of the GGD. However, estimating the parameters using the numerical methods is challenging one to the difficulties in deriving their values from the mean and variance equations unless one parameter is known. Furthermore, using the maximum-likelihood function to estimate all three parameters simultaneously is a computationally complicated technique since it is impossible to obtain a straightforward differentiation of the log-likelihood function for the shape parameter [11]. Some of the parameter estimation methods include the Least Squares (LS) estimation of GGD

* e-mail: amani.sayed1983@yahoo.com

parameters presented in [12,13]. An alternative method of estimation for GGD parameters based on heuristic optimization approaches was proposed [14,15]. The heuristics approach was found to be effective.

The conventional Inferential statistical methods strongly depend upon the sample size drawn from the population. Often p -value approach is used to accept or reject a hypothesis. The computed p -value can be different with different stopping conditions and are therefore not reliable. Bayesian statistics overcomes these difficulties by assuming prior probabilities of the parameters which are used with the likelihood model to generate posterior probabilities of the parameters. A white paper on Bayesian statistics with its uses and examples for complex problems is provided in [16,17]. Numerous applications of Bayesian statistics in various domains such as physics, atmospheric science, geology, biomedicine and economics have been researched. It has been demonstrated with examples that Bayesian analysis provides proper scientific Inference and rational decision-making [17,18].

A Markov Chain Monte Carlo (MCMC) is used to estimate the posterior distribution of parameters with the help of systematic random sampling in a probabilistic domain. In the straightforward Monte-Carlo approach, samples are independently drawn from the distributions, whereas in MCMC methods, the next sample depends upon the existing sample, hence the name Markov chain. The Monte Carlo method alone cannot track high-dimensionality probabilistic models and the alternative MCMC method provides a more realistic approach. In the MCMC approach, each sample is linked only to the previous sample and remains independent of the sample's prior distribution (Van et al., 2018). An approach to the MCMC method is the Gibbs sampling algorithm. In this algorithm, only one variable is sampled at a time based on the condition of the other variables. An approach of deterministic sweep in Gibbs sampler is demonstrated for complex models involving multi-dimensional distributions as compared to sampling from entire conditional distributions [19]. The work of Geman et al. [20] demonstrated the low state energy concept of Gibbs sampling to an image processing problem. A generalization of the MCMC method is presented by Hastings [21] along with applications to numerical problems. The Gibbs sampling has been used for a variety of problems that include expert systems and image processing. The Gibbs sampler is known to converge when the conditional distributions are compatible. Otherwise, full conditionals are estimated with the assumption that a single complete conditional distribution is only dependent on a "neighbourhood" subset of the variables [22]. Gibbs sampling does not work with joint distribution of variables, it needs to sample a high volume of data based on a conditional probability approach where a sample from one variable is drawn at a time. Therefore, sampling must be efficient to save time. Often, conditional distributions reduce to well-known analytical distributions from which sampling is much easier [23]. If this is not possible, then as an alternative, rejection sampling (RS) and Metropolis-Hastings (MH) algorithms are used to draw samples.

In the RS approach, a theoretical distribution encloses the desired distribution and then the sample is accepted or rejected based on where it lands in the search space. The rejection rate of such methods is often very high leading to enhanced computational effort. An alternative to this is to use Adaptive Rejection Sampling (ARS) wherein the envelop function is defined in log space which helps in making density functions simpler. The ARS method proposed by Gilks et al. [24] has been designed to draw effectively from univariate target densities to raise the acceptance rate. Robert et al. (2004) improved the ARS method. As the name suggests, the ARS method modifies the envelop function based on the rejected sample and is therefore adaptive. In the MH algorithm, samples are drawn from a uniform distribution, a proposal distribution, and a target distribution. Gibbs sampling is a special case of the MH algorithm where the proposed moves are accepted with a probability of 1 ([25]; Liang et al., 2010). The Gibbs sampler ergodicity is preserved by the hybrid technique as proposed by ([25]; Brewer, 1993). Instead of using a single MH step in a Gibbs sampler ([25]; Brewer, 1993). The adaptive Rejection Metropolis Sampling (ARMS) algorithm which combines both ARS and MH algorithms was developed (Robert, 2004). For log-concave distribution functions, the ARMS algorithm provides efficient sampling. ARMS provides superior performance and has been used in numerous inference issues (Liang et al., 2010).

The estimation of gamma distribution parameters is often done using a frequentist approach than using Bayesian methods, thus motivating the present study. Using de Groot's theoretical model, it is demonstrated that conjugate priors exist for the gamma parameters [26]. An exponential family curve in generating the conjugate priors to determine the Bayesian estimates was employed [27]. Gibbs sampling method to estimate the four parameters of the gamma distribution was used [28]. Similarly, Wu et al. [29] estimated three parameters of the lognormal distribution, again using Gibbs sampling. Another study by [30] estimated three parameters of gamma distribution and compared results using Gibbs sampling, ARS, and ARMS methods. A Skew-Generalized Error Distribution (GED) stochastic volatility model was studied by Son et al. [31] where Gibbs sampling and ARMS methods were applied in the estimation of parameters. Bayesian estimates of unknown parameters with inaccurate priors using the Gibbs sampling method were studied by Oh et al. [32] in comparing their results to maximum likelihood estimators (MLEs) and modified moment estimators [28]. A study on Bayesian estimation of the four-parameter gamma distribution considering a non-informative prior was studied in [33]. Where sampling was done using Gibbs sampler, ARS, and ARMS techniques. For the Bayesian estimation of two-parameter gamma distribution by [28] using the Gibbs sampling method. In this study, a Bayesian estimate of generalized gamma distribution (GGD) parameters was carried out using a reference prior distribution, reparameterization and marginalization. The objective of this study is to carry out MCMC-based Bayesian estimation of the parameters of a GGD using Gibbs sampler and ARMS

method and compare the results with the existing Simulated Annealing (SA) and method of moments (MM). The rest of the paper is organized as follows. The next section describes the methodology followed in the present study. GGD, Bayesian Analysis, MCMC, Gibbs sampler and ARMS are elaborated on in this section. The further section enumerates the results and explains the findings. The fourth section provides suggestions for future studies. Finally, conclusions from the study are presented.

2 Methodology

2.1 Generalized gamma distribution

Let $i = 1, 2, \dots, n$, the generalized gamma distribution (GGD) on the random variable X_i with three parameters, the shape parameter α , the scale parameter θ , and the growth parameter γ associated with the given duration $K = 1, 2, \dots, K^*$, be defined in the following Pdf,

$$f_{X_i}(x; \alpha, (1 + \gamma)^{K-1}\theta) = \begin{cases} \frac{1}{\Gamma(\alpha)} \left[\frac{1}{(1 + \gamma)^{K-1}\theta} \right]^\alpha x^{\alpha-1} e^{-\frac{x}{(1 + \gamma)^{K-1}\theta}}, & x > 0 \\ 0, & otherwise. \end{cases} \tag{1}$$

where the gamma $\Gamma(\alpha)$ in equation (1) is defined as follows

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

From equation (1), we can find the mean,

$$E(X) = \alpha[(1 + \gamma)^{k-1}\theta],$$

and the variance,

$$\sigma^2 = \alpha[(1 + \gamma)^{K-1}\theta]^2. \tag{2}$$

From equation (1), we can write the log-likelihood function as follows,

$$l(\alpha, (1 + \gamma)^{K-1}\theta|x_i) = n \alpha \ln((1 + \gamma)^{K-1}\theta) - n \ln(\Gamma(\alpha)) + (\alpha - 1) \sum_{i=1}^n \ln(x_i) - \frac{1}{(1 + \gamma)^{K-1}\theta} \sum_{i=1}^n x_i. \tag{3}$$

2.2 Bayesian analysis and markov chain monte carlo (MCMC)

In this work, the MCMC method employs systematic random samples from the distribution of parameters. The method is different from deterministic maximum likelihood algorithms where each sample is independent of the other. The Markov chain provides connectivity of each sample to

its previous one thus ensuring a systematic and controlled approach to sampling. In Bayesian analysis, the prior distribution of parameters is constructed according to their likelihood function. It is interesting to note that even though the estimator's distribution is unknown, random samples can still be drawn from it. Using the Markov chain, the posterior distribution is characterized by using the gradient information from the random samples [34]. The mean of the random variables is of significance. The random variable and its function are obtained by the probability of the outcome and a logarithm of an odds ratio respectively. To obtain $E[f(y)]$, for example, by drawing y_1, \dots, y_n from the distribution, applying the function to each of these values $f(y_1), \dots, f(y_n)$ and then calculating the mean of these values to find the expectation, which may be written as

$$E[f(y)] \approx \frac{1}{n} \sum_{i=1}^n f(y_i). \tag{4}$$

The Monte Carlo integration is used to get the solution in equation (4). It is a time-consuming process to compute $f(y)$ for a large number of samples. The MCMC method has significant advantages for Bayesian estimation in complex statistical models. It is an excellent tool for accurate statistical modelling in making conclusions about model parameters or predictions. However, Bayesian analysis necessitates integration over potentially high-dimensional probability distributions. Therefore, MCMC implies different techniques to simplify complexity in Bayesian computation [35]. One of these techniques is the Gibbs sampler which will be explained in the next section.

2.3 The Gibbs sampler

The Gibbs sampler is a special case of the MH algorithm where the sequence of observations is approximated from a defined multivariate probability distribution. This helps in an approximation of the joint and marginal distribution of unknown parameters. In this work, the methodology proposed by Pang et al (2001) and used by Shang et al. [36] is adopted that estimates the three gamma distribution parameters: shape, scale, and location. Let $p(\delta)$ be the probability distribution of an unknown parameter δ . Then the sequel probability densities of δ are $p(\delta) = F'(\delta)$, where $F(\delta)$ is the cumulative distribution function (CDF) of δ . Imposing another unknown parameter ω , the joint, conditional, and marginal densities are written as $p(\delta, \omega)$, $p(\omega|\delta)$, $p(\omega)$, respectively.

We assumed that $p(\underline{\delta})$ be a probability distribution of the desired parameters where $\underline{\delta} = (\delta_1, \delta_2, \dots, \delta_M)$ then the Gibbs sampler framework is explained in Figure 1. As steps I and II show, the Gibbs sampler requires samples from conditional distributions. This sampling can be done independently using Rejection sampling, which is computationally efficient. Rejection sampling is the technique for sampling independently from a general density $p(\delta)$, which $p(\delta)$ is hard to be analyzed. Moreover, Rejection sampling requires an envelope function $qp(\delta)$ and satisfies $q(\delta) \geq p(\delta)$ for all δ . Samples are drawn from density proportional to q ,

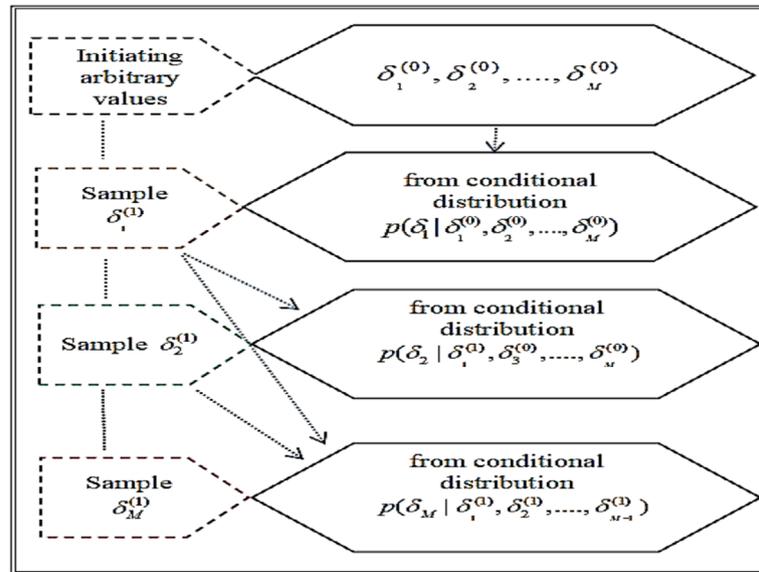


Fig. 1. Gibbs sampler form parameters probability distribution framework.

and each sampled point δ can be accepted or rejected. Alternatively, one can use Adaptive Rejection Metropolis Sampling (ARMS) or the Metropolis method.

2.4 Adaptive rejection metropolis sampling (ARMS)

ARMS is a sampler algorithm used in the MCMC method to form a single variable target distribution indicated by its log density. The rejection distribution is generated using piecewise linear functions. This then envelops the target density function. The ARS sampler works well with log-concave density functions. That is, if $p(\delta)$ is log-concave, the sampler is effective. However, calculating the second derivatives of the log-likelihood function in equation (3) can be difficult due to the presence of nonlinear terms. To account for non-log-concave samples extra MH step is incorporated leading to the ARMS approach. The ARMS algorithm for estimating the three parameter gamma distribution in equation (1) is explained as follows,

- Set $p(\delta)$ as the targeted conditional density and to ease the sampling process, approximate $g(\delta)$ by $p(\delta)$.
- For $i=0, \dots, N-1$, sample a point δ^* from $g(\delta^*|\delta^{(i)})$.
- Randomly sample $r \in Unif(0, 1)$.
- If $r \leq \min[1, \frac{p(\delta^*)g(\delta^{(i)})}{p(\delta^{(i)})g(\delta^*|\delta^{(i)})}]$, then accept $\delta^* = \delta^{(i+1)}$.
- Else, set $\delta^{(i+1)} = \delta^{(i)}$, where $\delta^{(i)}$ the old value in the Markov.

Steps (2) and (3) are repeated until δ is achieved within some pre-specified tolerance value. g is called the proposal distribution with a symmetric random walk, for example, $g(\delta|\omega) = g(\omega|\delta)$, and should be carefully chosen.

3 Cases studies and results

In this study, we utilize the MCMC estimation method to estimate the three parameters of GGD using the Gibbs sampler method's ARMS, explained above, over two

different cases. The ARMS algorithm was run on the Matlab platform to generate α , θ , and γ at every iteration step. The simulation run 15,000 times and we disregarded the first 5000 times as a burn. The ARMS algorithm enables us to obtain the empirical posterior distribution of the parameters under study and to calculate the central tendency.

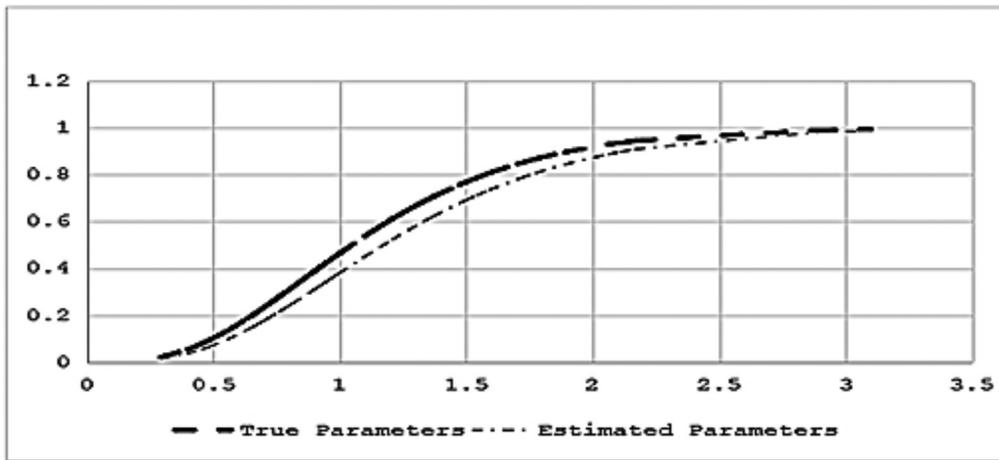
3.1 Case 1: Simulated study

The first case study is based on a simulated study conducted by Sayed et al. [37] with the assumption of three parameters of generalized gamma distribution (GGD). In the study, the data sets were constructed for $K = 2,3,4$ with respective sample sizes of $N = 50,100,200,500$. The three parameters GGD were then estimated using the Simulated Annealing (SA) algorithm, which was observed to be an efficient method. Table 1 shows the result of using the ARMS method of estimating the parameters of GGD and compares this result with the SA algorithm in [38].

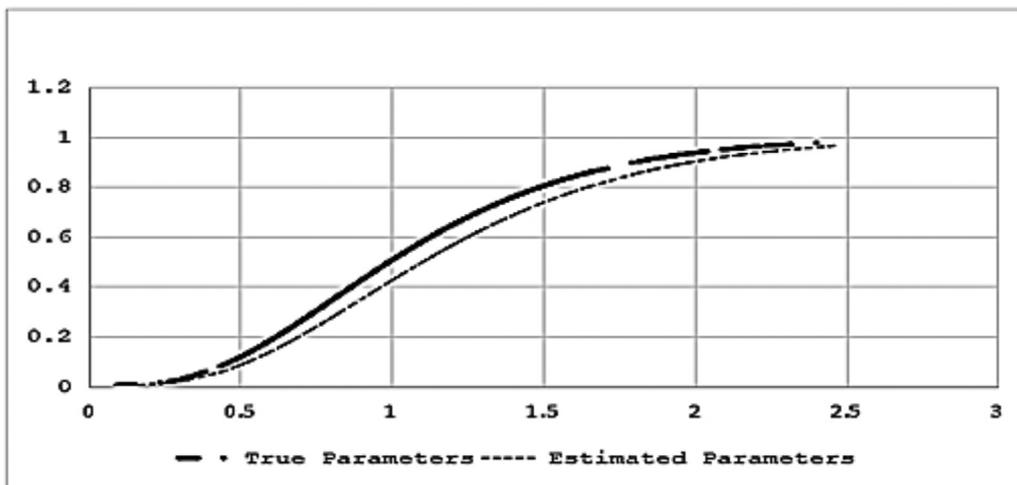
Table 1 shows the computational analysis of ARMS performance in comparison with SA in estimating the parameters of GGD based on different sample sizes. The value of SD is close to zero which indicates that most of the parameter's estimations are close to the mean estimate therefore we can take the mean as the best estimation for the three parameters in most cases. Also, the comparison between the mean estimations of the ARMS has the potential to get slightly closer to the true parameters than the SA estimation in some cases, especially in the small sample size. However, as the sample sizes increases, the performance of both ARMS and SA displayed similar trends. Figure 2 displays the trend in Gamma CDF based on the true parameters. It can be seen that the mean of the ARMS estimated parameter is closer to the true parameters in all cases. For more detail on SA performance refer to [37,39]. Finally, we computed the mean absolute error

Table 1. The result of GGD parameters estimates via ARMS and SA methods.

Parameters	True parameters	ARMS				SA			
		Mean	Median	Mode	SD	Mean	Median	Mode	SD
α	4	4.090	4.091	4.068	0.0093	4.111	3.978	3.879	0.2570
θ	0.3	0.328	0.327	0.297	0.0175	0.330	0.329	0.323	0.0061
γ	-0.05	-0.050	-0.049	-0.072	0.0096	-0.051	-0.052	-0.054	0.0024
α	4	4.116	4.133	3.786	0.1499	4.274	4.286	4.463	0.1444
θ	0.3	0.324	0.309	0.179	0.0672	0.265	0.265	0.267	0.0031
γ	-0.05	-0.051	-0.051	-0.053	0.0007	-0.052	-0.052	-0.054	0.0024
α	4	4.030	4.024	4.075	0.0895	4.152	4.137	4.301	0.1423
θ	0.3	0.312	0.294	0.299	0.0796	0.294	0.295	0.295	0.0054
γ	-0.05	-0.045	-0.045	-0.081	0.0027	-0.051	-0.052	-0.053	0.0024
α	4	4.031	4.031	3.984	0.0288	4.032	4.057	3.918	0.1090
θ	0.3	0.339	0.347	0.292	0.0213	0.314	0.314	0.315	0.0043
γ	-0.05	-0.049	-0.05	-0.072	0.0116	-0.052	-0.052	-0.054	0.0025

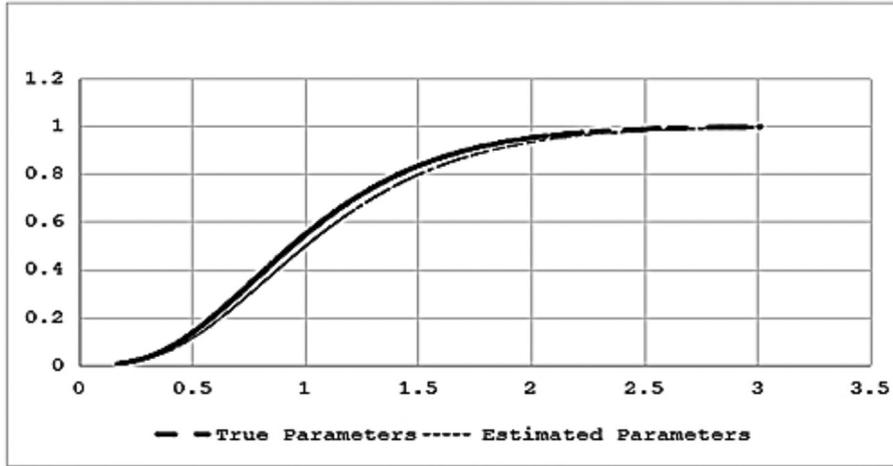


(a)

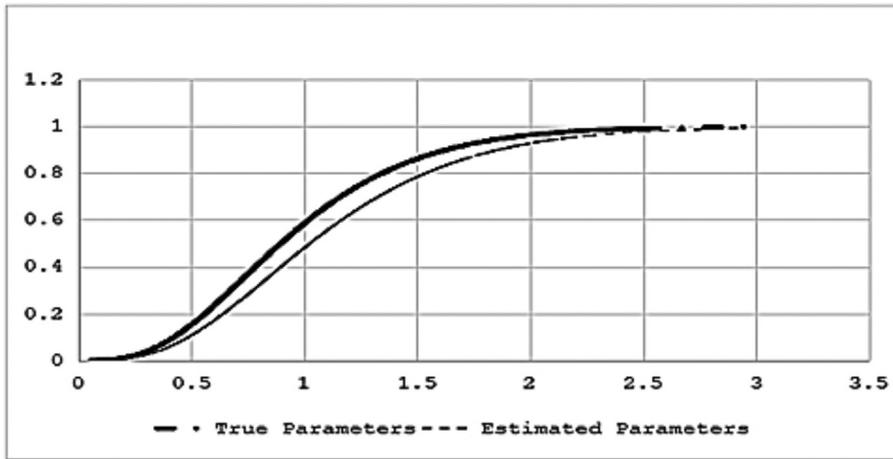


(b)

Fig. 2. a) CDF of the true and estimated values of the GGD via ARMS for 50 sample size; b) CDF of the true and estimated values of the GGD via ARMS for 100 sample size; c) CDF of the true and estimated values of the GGD via ARMS for 200 sample size; d) CDF of the true and estimated values of the GGD via ARMS for 500 sample size.



(c)



(d)

Fig. 2. (Continued).

(MAE) and root mean square error (RMSE) from a set of 100 ARMS estimations for each parameter $\hat{\alpha}$, θ and γ for each data sample using equations (5) and (6) respectively as follows,

$$MAE = \frac{1}{3} \left(\left| \hat{\alpha} - \alpha \right| + \left| \hat{\theta} - \theta \right| + \left| \hat{\gamma} - \gamma \right| \right), \tag{5}$$

$$RMSE_{V_i} = \sqrt{\frac{\sum (\hat{V}_i - V)^2}{n}}. \tag{6}$$

Box-Plot to compare it to the mean absolute error from 100 sample size. The SA estimated estimations for each parameter as presented in Figure 3. It is clear from Figure 3 that, the MAE of the ARMS estimated parameters are less spread and closer to zero than the MAE of the SA estimated parameters in most cases. Also, we notice that working with the ARMS gives a good result for the different sizes of data samples, unlike the SA which works better with larger

data samples. Therefore, using the ARMS to estimate the three-parameter of the generalized gamma distribution is more efficient compared with SA.

3.2 Case 2: MIRR study

The second case deals with the modified internal rate of return (MIRR) data presented in [38]. Sayed et al. [37] calculated the MIRR of 62 companies based on the Malaysian property sector from 2008 to 2019 then, they set up eight investment period $K=1, 2, 3, \dots, 8$; to construct eight data set of different sizes. In this work, we distribute the MIRR using three parameters gamma distribution and we utilize the method of the moment with fixed growth parameter $\gamma = -0.0001$ to build an initial assumption about the parameters to estimate them using the Simulated Annealing algorithm to compare the result of using the SA algorithm against ARMS algorithm.

Table 2 shows the result of using the ARMS method in parameter estimation over the three parameters GGD in equation (1) and compares this result to the fixed γ and SA

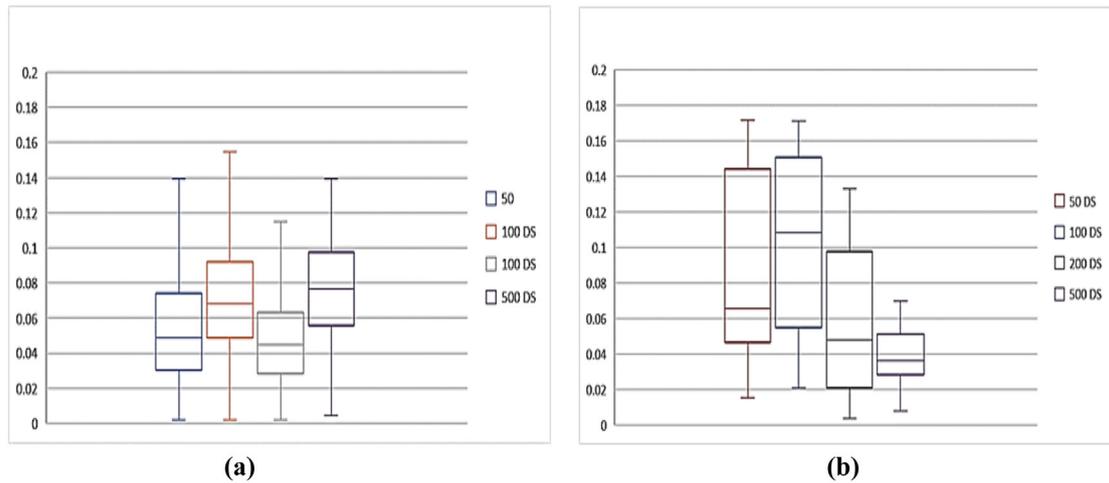


Fig. 3. Comparison of different estimation methods via the the box plots of sample size (50, 100, 200, 500).

Table 2. The result of estimated via ARMS, MM and SA with a fixed value of γ .

Parameters		ARMS				$MM\gamma = -0.0001$	SA
		Mean	Median	Mode	SD		
α	K = 1 N = 620	2.9148	2.9129	2.8885	0.0120	2.9861	2.9299
θ		0.3788	0.3790	0.3586	0.0075	0.3752	0.3709
γ		-0.000141	-0.000120	-0.000418	0.000108	-0.0001	-0.000098
α	K = 2 N = 558	9.7151	9.7120	9.6778	0.0185	9.8058	9.6941
θ		0.1136	0.1148	0.0817	0.0113	0.1094	0.1071
γ		-0.000095	-0.000083	-0.000190	0.000044	-0.0001	-0.000085
α	K = 3 N = 496	16.8995	16.9024	16.8638	0.0162	16.6495	17.3998
θ		0.0611	0.0598	0.0493	0.0068	0.0639	0.0610
γ		-0.000098	-0.000105	-0.000180	0.000048	-0.0001	-0.000098
α	K = 4 N = 434	22.3679	22.3642	22.3420	0.0164	22.4730	22.5085
θ		0.0557	0.0547	0.0413	0.0063	0.0471	0.0473
γ		-0.000050	-0.000049	-0.000122	0.000032	-0.0001	-0.000099
α	K = 5 N = 372	31.3936	31.4017	31.3427	0.0239	31.5448	32.4969
θ		0.0393	0.0389	0.0287	0.0043	0.0330	0.0321
γ		-0.000127	-0.000133	-0.000229	0.000056	-0.0001	-0.000099
α	K = 6 N = 310	43.1545	43.1585	43.1186	0.0151	43.1858	43.6111
θ		0.0274	0.0272	0.0196	0.0029	0.0238	0.0235
γ		-0.000065	-0.000066	-0.000150	0.000042	-0.0001	-0.000099
α	K = 7 N = 248	56.6156	56.6171	56.5889	0.0084	56.5871	57.2080
θ		0.0198	0.0196	0.0146	0.0019	0.0180	0.0177
γ		-0.000057	-0.000048	-0.000145	0.000037	-0.0001	-0.000099
α	K = 8 N = 186	69.9174	69.9220	69.8854	0.0165	69.8467	70.2149
θ		0.0160	0.0158	0.0125	0.0015	0.0145	0.0143
γ		-0.000069	-0.000066	-0.000155	0.000047	-0.0001	-0.000099

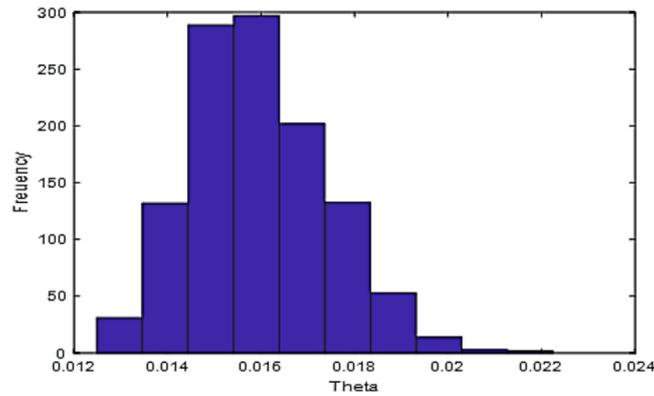


Fig. 4. The frequency of the empirical posterior distribution of the estimated θ for $K=8$.

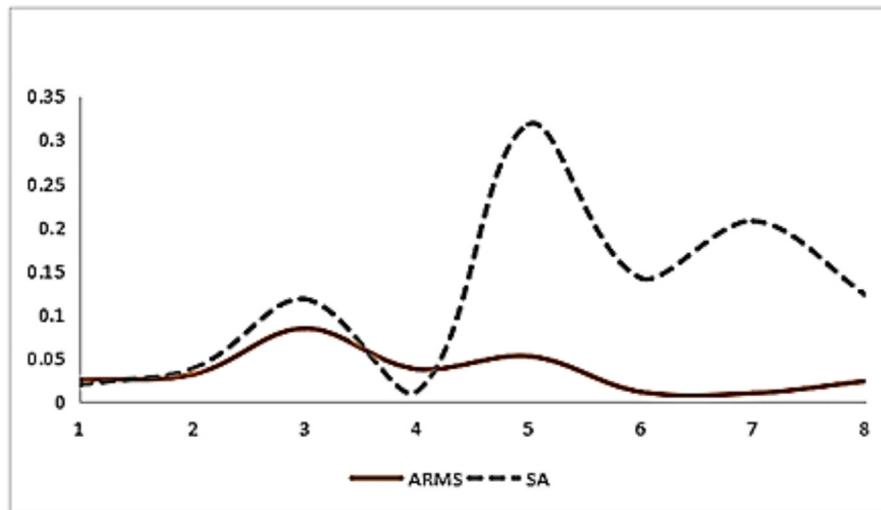


Fig. 5. The MAE for comparing ARMS versus SA estimation methods.

algorithm outcome. It can be observed based on the mean and standard deviation of the ARMS method that all estimated parameters seem to be quite low and close to zero. However, in $K=8$ we can see that the mean fails to give a good estimate for the parameter θ and by looking at the parameter empirical posterior distribution frequency we found it to be skewed to the right as shown in Figure 4 hence, we can take the median to be the best estimation for θ in this case.

In the comparison between the performance of the ARMS method to the SA method which is built on the assumption of $MM_\gamma = -0.0001$. We find that both the ARMS method and SA method give similar results as they both provide a closer estimation of the $MM_\gamma = -0.0001$. However, the ARMS method seems to give a slightly closer estimation than the SA in some cases.

The numerical results of the MAE and RMSE measures provided in Tables 1 and 2 are displayed graphically corresponding to each table in Figure 5 and 6. The

performance comparison based on MAE using equation (5) in Figure 5 reveals that both ARMS and SA have similar error trends based on the predicted values and used $MM_\gamma = -0.0001$ as actual values. However, the performance of ARMS is slightly better than SA when the sample size is small. This emphasizes the efficiency of the ARMS technique. Figure 6 displays the trend of RMSE which was computed using equation (6) of the predicted variance that was calculated using Equation (2) with the ARMS and SA estimated parameters and we used the variance calculated using $MM_\gamma = -0.0001$ as the actual variance, to observe the ARMS potential impact on the model variance by comparison with the SA which is known to be effective in minimizing the model variance. In Figure 6 we can see both methods have ARMS close to zero however, the SA ARMS seem to have a better result as we can notice it gets closer to zero as K gets larger hence, the SA method has more potential in minimizing the variance than the ARMS method.

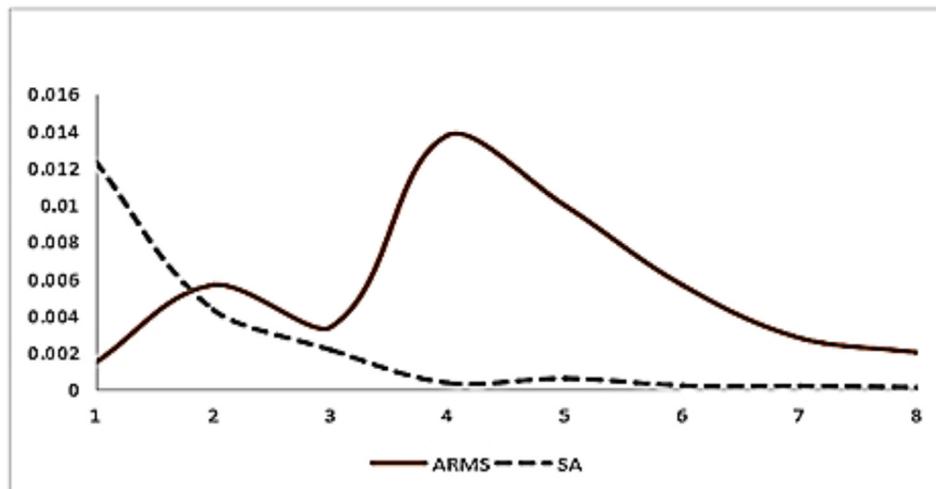


Fig. 6. The RMSE for eight years investment period.

4 Conclusion

This study focused on the Markov Chain Monte-Carlo (MCMC)-based Bayesian estimation of the Generalized Gamma Distribution (GGD) utilizing the Adaptive Rejection Metropolis Sampling (ARMS) technique of random variable generation within the Gibbs sampler algorithm based on two case studies involving simulated data from [14,37,38] and actual data of Modified Internal Rate of Return (MIRR) from [37–39]. The results of applying the ARMS method are compared with those of the application of the SA algorithm in the case of simulated data and with those of the Method of Moment (MM) and SA applications for the real MIRR data. It is observed that in the case of simulated data, the performance of the ARMS technique is marginally superior to that of the SA method. In the case of actual MIRR data, the performances of ARMS and SA applications compared to the MM, as an initial assumption about the parameters is similar; however, ARMS bears the advantage of avoiding the optimization procedure, adding to that its ability to deal with data samples of smaller sizes efficiently. This study suggests using the MCMC method for estimating the parameters of complex multi-dimensional probability distributions belonging to the same exponential family.

Future study

We suggest using the MCMC method to estimate the parameters of multi-dimensional distributions from the same exponential family. Additionally, the Bayes Prediction derivation, which aims to provide an estimate of the posterior predictive density function of future experiment observations based on information from an informative experiment, can yield significant results.

References

1. L. Amoroso, Ricerche intorno alla curva dei redditi, *Ann. Mat. Pura Appl.* **2**, 123–159 (1925)
2. E.W. Stacy, A generalization of the gamma distribution, *Ann. Math. Stat.* **33**, 1187–1192 (1962)
3. N.L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions* (Wiley, New York, 1994), Vol. 2
4. S.K. Agarwal, J.A. Al-Saleh, Generalized gamma type distribution and its hazard rate function, *Commun. Stat. –Theory Methods.* **30**, 309–318 (2001)
5. N. Balakrishnan, Y. Peng, Generalized gamma frailty model, *Stat. Med.* **25**, 2797–816 (2006)
6. S. Nadarajah, A.K. Gupta, Statistical tools for drop size distributions: moments and generalized gamma. *Math. Comput. Simul.* **74**, 1–7 (2007)
7. L.G. Pinho, G.M. Cordeiro, J.S. Nobre, The gamma-exponentiated Weibull distribution, *J. Stat. Theory Appl.* **11**, 379–395 (2012)
8. S. Jaggia, Specification tests based on the heterogeneous generalized gamma model of duration: with an application to Kennan’s strike data, *J. Appl. Econom.* **6**, 169–180 (1991)
9. M. Khodabina, A. Ahmadabadib, Some properties of generalized gamma distribution, *Math. Sci.* **4**, 9–28 (2010)
10. J. Kiche, O. Ngesa, G. Orwa, On generalized gamma distribution and its application to survival data, *Int. J. Stat. Probab.* **8**, 85–102 (2019)
11. O. Gomès, C. Combes, A. Dussauchoy, Parameter estimation of the generalized gamma distribution, *Math. Comput. Simul.* **79**, 955–963 (2008)
12. B. Lagos Álvarez, G. Ferreira, M. Valenzuela Hube, A proposed reparametrization of gamma distribution for the analysis of data of rainfall-runoff driven pollution, *Proyecciones (Antofagasta)* **30**, 415–439 (2011)
13. R. Vani Lakshmi, V.S. Vaidyanathan, Three-parameter gamma distribution: estimation using likelihood, spacings and least squares approach, *J. Stat. Manag. Sys.* **19**, 37–53 (2016)

14. H. Abubakar, S.R.M. Sabri, A simulation study on modified Weibull distribution for modelling of investment return, *Pertanika J. Sci. Technol.* **29**, 2767–2790 (2021)
15. V.S. Özsoy, M.G. Ünsal, H.H. Örkücü, Use of the heuristic optimization in the parameter estimation of generalized gamma distribution: comparison of GA, DE, PSO and SA methods, *Comput. Stat.* **35**, 1895–1925 (2020)
16. S.M. Stigler, *The history of statistics: the measurement of uncertainty before 1900* (Harvard Uni. Press, 1986)
17. M.J. Zyphur, F.L. Oswald, Bayesian estimation and inference: a user's guide, *J. Manag.* **41**, 390–420 (2015)
18. H. Abubakar, S.R.M. Sabri, Weibull distribution for claims modelling: a Bayesian approach. in: 2022 *International Conference on Decision Aid Sciences and Applications (DASA) 2022* Mar 23, IEEE, pp. 108–112
19. K. Łatuszynski, G.O. Roberts, J.S. Rosenthal, Adaptive Gibbs samplers and related MCMC methods, *Ann. Appl. Probab.* **23**, 66–98 (2013)
20. S. Geman, D. Geman Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 721–741 (1984)
21. W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika.* **57**, 97–109 (1970)
22. A.E. Gelfand, A.F. Smith, Sampling-based approaches to calculating marginal densities, *J. Am. Stat. Assoc.* **85** (410), 398–409 (1990)
23. W.R. Gilks, Derivative-free adaptive rejection sampling for Gibbs sampling, *Bayesian Stat.* **4** (2), 641–649 (1992)
24. W.R. Gilks, P. Wild, Adaptive rejection sampling for Gibbs sampling, *J. R. Stat. Soc., C: Appl. Stat.* **41**, 337–348 (1992)
25. L. Martino, J. Read, D. Luengo, Independent doubly adaptive rejection Metropolis sampling within Gibbs sampling. *IEEE Trans. Signal Process.* **63**, 3123–3138 (2015)
26. R.B. Miller, Bayesian analysis of the two-parameter gamma distribution, *Technometrics.* **22**, 65–69 (1980)
27. B. Pradhan, D. Kundu, Bayes estimation and prediction of the two-parameter gamma distribution, *J. Stat. Comput. Simul.* **81**, 1187–1198 (2011)
28. S.K. Upadhyay, M. Peshwani, Full posterior analysis of three parameter lognormal distribution using Gibbs sampler, *J. Stat. Comput. Simul.* **71**, 215–230 (2001)
29. C.W. Wu, M.H. Shu, T.Y. Huang, B.M. Hsu, Comparisons of frequentist and Bayesian inferences for interval estimation on process yield, *J. Oper. Res. Soc.* **9**, 1–2 (2021)
30. N. Cappuccio, D. Lubian, D. Raggi, MCMC Bayesian estimation of a skew-GED stochastic volatility model, *Stud. Nonlinear Dyn. Econom.* **8**, (2004)
31. Y.S. Son, M. Oh, Bayesian estimation of the two-parameter Gamma distribution, *Commun. Stat. –Simul. Comput.* **35**, 285–293 (2006)
32. M.R. Oh, K.S. Kim, W.H. Cho, Y.S. Son, Bayesian parameter estimation of the four-parameter gamma distribution, *Commun. Stat. Appl. Methods.* **14**, 255–266 (2007)
33. P.L. Ramos, F. Louzada, Bayesian reference analysis for the generalized gamma distribution, *IEEE Commun. Lett.* **22**, 1950–1953 (2018)
34. J. Tang, B. Fan, L. Xiao, S. Tian, F. Zhang, L. Zhang, D. Weitz, A new ensemble machine-learning framework for searching sweet spots in shale reservoirs, *SPE J.* **26**, 482–497 (2021)
35. G. Lee, W. Kim, H. Oh, B.D. Youn, N.H. Kim, Review of statistical model calibration and validation—from the perspective of uncertainty structures, *Struct. Multidiscip. Optim.* **60**, 1619–1644 (2019)
36. X. Shang, H.K. Ng, On parameter estimation for the generalized gamma distribution based on left-truncated and right-censored data, *Comput. Math. Methods.* **3**, e1091 (2021)
37. A.I.A. Sayed, S.R.M. Sabri, A simulation study on the simulated annealing algorithm in estimating the parameters of generalized gamma distribution, *Sci. Technol. Indones.* **27**, 84–90 (2022)
38. A.I.A. Sayed, S.R.M. Sabri, Transformed modified internal rate of return on gamma distribution for long term stock investment modelling, *J. Manag. Inf. Decis. Sci.* **25**, 1–7 (2022)
39. H. Abubakar, S.R. Sabri, Incorporating simulated annealing algorithm in the Weibull distribution for valuation of investment return of Malaysian property development sector, *Int. J. Simul. Multidiscip. Des. Optim.* **12**, 22 (2021)

Cite this article as: Amani Idris A. Sayed, Shamsul Rijal Muhammad Sabri, Generalized gamma distribution based on the Bayesian approach with application to investment modelling, *Int. J. Simul. Multidisci. Des. Optim.* **14**, 10 (2023)