International Journal of
Information
and
Management
Sciences

International Journal of Information and Management Sciences
29 (2018), 235-256. DOI:10.6186/IJIMS.201809_29(3).0001

Inference from Two-Variable Degradation Data Using Genetic Algorithm and Markov Chain Monte Carlo Methods

Jyun-You Chiang, Jianping Zhu, Yu-Jau Lin, Y. L. Lio and Tzong-Ru Tsai

Southwestern University of Finance and Economics, Xiamen University,
Chung Yuan Christian University, University of South Dakota and Tamkang University

Abstract

Two-variable gamma process with generalized Eyring model have been widely used to assess the reliability of reliable products in engineering applications. Because no close forms of the maximum likelihood estimators for the model parameters can be derived and iterative procedure to evaluate the maximum likelihood estimate is very sensitive to the initial input and difficult to control, the analytic genetic algorithm, Gibbs sampling Markov chain Monte Carlo algorithm and Metropolis-Hastings Markov chain Monte Carlo algorithm methods are established and applied to implementing parameter estimation for the gamma process. The performance of those methods are evaluated through simulations. Simulation results show that the Markov chain Monte Carlo method based maximum likelihood estimates outperform the other competitors with smaller bias and mean squared error. The application of the proposed methods was illustrated with a lumen degradation data set of light-emitting diodes.

Keywords: Cumulative exposure model, Gibbs sampling algorithm, Markov chain Monte Carlo, Metropolis-Hastings algorithm.

1. Introduction

1.1 Problem Description

The failure times are difficult to obtain for the reliability inference of highly reliable products that are tested by using conventional life test procedures (see Lim and Yum [10], Peng and Tseng [19]). One approach to overcome the difficulty is to predict the mean time to failure (MTTF) or lifetime percentiles of products based on degradation information, which is measured through life testing under high stress-loading conditions (see Lim and Yum [10], Padgett and Tomlinson [14], Park and Padgett [15], Tsai et al. [22]). The Brownian motion (BM) and geometric Brownian motion (GBM) processes have been widely used to model the degradation of products under stresses over time (see Liao and Tseng [9], Lim and Yum [10], Tsai et al. [22], Tsai et al. [24], Whitmore [27]). However, because the BM and GBM processes can generate negative increments, they are unsuitable to model the degradation processes of products. To overcome the drawback of generating negative increments from the BM and GBM processes, the gamma processes (GPs) can be used to model the degradation processes of highly reliable products subject to an accelerated degradation test (ADT), instead. A GP always exhibits a monotone-increasing accumulative pattern and hence is more suitable to model a degradation process for performing reliability assessment than the BM or GBM process (see Boulanger and Escobar [3], Guan and Tang [6], Park and Padgett [15], Park and Padgett [16], Park and Padgett [17], Peng [18], Tsai et al. [23], Tseng et al. [26]).

To save test cost and time, the degradation test can be conducted with two stress loading variables to accelerate product damage (see Park and Padgett [16], Park and Padgett [17]). In this study, a GP with two stress loading variables, namely the ambient temperature and drive current, was considered to model the damage processes of highly reliable products. We use the term 2ADT-GP to dente the ADT with two stress variables in a GP hereafter. The ambient temperature and drive current are two widely used stress variables for engineering applications. Let X_t denote the cumulative damage path of a highly reliable product and follow the GP with a positive shape coefficient ν_L and a positive scale parameter β, where only the shape coefficient ν_L depends on the stress level L of the degradation test. The product is classified as a failure if X_t passes a given threshold, C, before the termination of the degradation test. Otherwise, the product is classified as a survivor. Let S denote the first passage time of the cumulative damage path over the given threshold, C. The working assumptions (A1) to (A5) of Tsai et al. [25] are also applied in this study for the 2ADT-GP and outlined as follows:

(A1) In the degradation test, k runs of comprising various stress levels are applied to the tested units. Run i represents a combination of levels for the two stress loading variables and is labeled by $L'_i = (L'_{1i}, L'_{2i})$ for $i = 1, 2, \ldots, k$.

(A2) A total of n_i units are allocated to the run i of degradation test, and all these units are subject to the stress loading L'_i.

(A3) The two components of $L'_i,\ i = 1, 2, \ldots, k$, are respectively standardized by

$$L_{1i} = \frac{1/L'_{10} - 1/L'_{1i}}{1/L'_{10} - 1/L'_{1M}} \tag{1.1}$$

for ambient temperature, and

$$L_{2i} = \frac{\log(L'_{2i}) - \log(L'_{20})}{\log(L'_{2M}) - \log(L'_{20})} \tag{1.2}$$

for drive current, where L'_{j0}, and L'_{jM} respectively represent the normal-used stress loading levels and the maximum stress loading levels for the stress loading variable j, $j = 1, 2$. Therefore, $L_{10} = L_{20} = 0$, $L_{1M} = L_{2M} = 1$ and $0 < L_{1i} \leq 1$, $0 < L_{2i} \leq 1$ for $i = 1, 2, \ldots, k$. L_{1i} and L_{2i} are scale-free and increasing functions of L'_{1i}, and L'_{2i}, respectively.

(A4) Let the starting time of the degradation test $t_{ij0} = 0$ and the initial damage of each unit in the life test be $x_{ij0} = 0$. The damage of each surviving unit in run i is measured at times $t_{ij1} < t_{ij2} < ... < t_{ijm_i}$ and labeled by $x_{ij1}, x_{ij2}, \ldots, x_{ijm_i}$, respectively. The damage increment $y_{ijh} = x_{ijh} - x_{ij(h-1)}$ follows a two-parameter gamma distribution with a shape coefficient $\delta_{ijh} = \nu_{L_i}\tau_{ijh}$ and a scale parameter β, where $\tau_{ijh} = t_{ijh} - t_{ij(h-1)}$, for $j = 1, 2, \ldots, q_i$, $h = 1, 2, \ldots, m_i$, and $i = 1, 2, \ldots, k$. Hence, the probability density function (PDF) of the two-parameter gamma distribution is defined by

$$f_A(y_{ijh}; \tau_{ijh}) = \frac{1}{\Gamma(\delta_{ijh})\beta^{\delta_{ijh}}} y_{ijh}^{\delta_{ijh}-1} e^{-y_{ijh}/\beta}, \; y_{ijh} > 0. \tag{1.3}$$

(A5) The parameter ν_{L_i} in the shape parameter δ_{ijh} of the gamma distribution can be expressed in terms of L_{1i} and L_{2i} through the generalized Eyring model (GEM) as

$$\nu_{L_i} = \exp(\gamma_0 + \gamma_1 L_{1i} + \gamma_2 L_{2i} + \gamma_3 L_{1i} L_{2i}), \; i = 1, 2, \ldots, k, \tag{1.4}$$

where $\gamma_0 < 0$, $\gamma_1, \gamma_2 > 0$, and $\gamma_3 \in R$.

The GEM model in (1.4) is a generalized function that includes three widely used single-loading acceleration models for degradation test as special cases, for example, the Arrhenius law model, power law model, and exponential law model, when only either L_1 or L_2 is considered. The precise distribution of S may be too complicated for practical use. Therefore, Park and Padgett provided an approximation procedure in Park and Padgett [15] and showed that the distribution of S could be approximated by the inverse Gaussian distribution if $C_\beta/\sqrt{\nu_L} \gg C_\beta/\nu_L$ (i.e., $\sqrt{\nu_L} \gg 1$), where $C_\beta = (C - x_0)/\beta$. Let $\mu_L = C_\beta/\nu_L$ and $\lambda_L = C_\beta^2/\nu_L$. When μ_L is very large, the approximation is effective even if $\sqrt{\lambda_L}$ is not excessively greater than μ_L (see Park and Padgett [15]). The PDF of the inverse Gaussian distribution is defined by

$$g_S(s; C) \equiv g_S(s; x_0 = 0, C) = \frac{C_\beta/\sqrt{\nu_L}}{\sqrt{2\pi s^3}} \exp\left[-\frac{\nu_L(s - C_\beta/\nu_L)^2}{2s}\right]. \tag{1.5}$$

The PDF of the damage increments observed from the GP can be described by Equation (1.3) for $j = 1, 2, \ldots, n_i$, $h = 1, 2, \ldots, m_i$, and $i = 1, 2, \ldots, k$, where $\delta_{ijh} = \nu_{L_i}\tau_{ijh}$.

Let $\mathbf{D} = \{(y_{ijh}, \tau_{ijh}),\ i = 1, 2, \ldots, k,\ j = 1, 2, \ldots, n_i,\ h = 1, 2, \ldots, m_i\}$ denote the data set of damage increments observed. The likelihood function for the 2ADT-GP can be presented as

$$L(\Theta; D) \propto \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh})\beta^{\delta_{ijh}}} y_{ijh}^{\delta_{ijh}-1} e^{-y_{ijh}/\beta}, \tag{1.6}$$

where $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) = (\beta, \gamma_0, \gamma_1, \gamma_2, \gamma_3)$. The maximum likelihood estimate (MLE) $\hat{\Theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4, \hat{\theta}_5)$ of $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ is the maximizer of the log-likelihood function, $\log(L(\Theta; \mathbf{D}))$. A summary of the technique terms is given as follows:

Technique term	**Full name**
ACO	ant colony optimization
ADT	accelerated degradation test
2ADT-GP	two stress variables ADT with a gamma process
BM	Brownian motion
CP	crossover probability
GA	genetic algorithm
GA-QN	combination of GA and QN
GBM	geometric Brownian motion
GEM	generalized Eyring model
GP	gamma processes
LED	light emitting diode
L-BFGS-B	limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm
M-H MCMC	Metropolis-Hastings MCMC
MCMC	Markov chain Monte Carlo
MCMCG	MCMC algorithm with using inverse gamma distribution as prior
MCMCU	MCMC algorithm with using independent uniform distributions as prior
MI	maximum number of iterations
MLE	maximum likelihood estimate
MP	mutation probability
MTTF	mean time to failure
MSE	mean square error
PDF	probability density function
PS	population size
PSO	particle swarm optimization
QN	quasi-Newton method
SNR	signal to noise ratio
SSO	swallow swarm optimization

1.2. Motivation and organization

The closed form of the MLE, $\hat{\Theta}$, for the 2ADT-GP cannot be obtained and an iterative procedure, such as the quasi-Newton (QN) method, must be applied to search the numerical values of all components in $\hat{\Theta}$ through the system of nonlinear likelihood equations. Because the terms, $\log(\Gamma(\nu_{L_i}\tau_{ijh}))$ and $\nu_{L_i}\tau_{ijh}$, with parameters of high dimension generate the complexity of the nonlinear likelihood equations, the computation results for the MLEs could not be accurate and hence induce relatively large bias and large mean squared error (MSE) based on our simulation experience. In this paper, an artificial intelligence computing method, namely genetic algorithm (GA) method, the GA-QN that is the method of combining the GA and QN, and the Gibbs sampling Markov chain Monte Carlo algorithm method (Gibbs MCMC), and Metropolis-Hastings Markov chain Monte Carlo algorithm (M-H MCMC) method are employed to search the MLE, $\hat{\Theta}$, of the GP parameter, Θ. The Gibbs MCMC and M-H MCMC methods can also be used to evaluate the Bayesian estimator of Θ. To implement the GA method, the GA parameters must be decided. To avoid choosing the GA parameters subjectively, the Taguchi design method is used for reaching an optimal design selection of the GA method.

We understand other potential soft computing methods besides the GA, for example, the algorithms of particle swarm optimization (PSO), swallow swarm optimization (SSO) or ant colony optimization (ACO) could also work well to optimize a complicated target function. We believe that the PSO-QN, SSO-QN or ACO-QN could works similarly to improve the performance of the QN method to search reliable MLEs of the 2ADT-GP parameters. The major purpose in this paper is to study the efficiency of the combined soft computation and gradient computing methods for searching reliable MLEs of the 2ADT-GP parameters. In this study, we use the GA-QN method for illustration. The performance of the GA-QN and MCMC methods are evaluated through using Monte Carlo simulation. The remainder of this paper is organized as follows: In Section 2, the GA method is addressed, and the Taguchi design method is implemented to reach an optimal GA design. In Section 3, the Gibbs MCMC and M-H MCMC methods are constructed analytically. The performance of the GA, QN, GA-QN and MCMC methods are investigated through simulation study, and the application via the data set of light-emitting diodes (LEDs) is presented in Section 4. Concluding remarks are provided in Section 5.

2. The GA Method

The GA is a widely used artificial intelligence computing method that can help practitioners to obtain high-quality estimates of model parameters for the implementation of system evaluation. The GA, introduced by Holland in 1975, is an evolutionary algorithm that generates solutions to optimize a target function. Many researchers have investigated the applications of the GA for optimization (see Akbari [1], Baudry et al. [2], Ferreira [5], Holland [8], Nicholson [12], Scrucca [20], Ting [21], Zhang et al. [28]). To determine the ranges of GA parameters is an issue for using a GA method. We understand several exiting studies have suggested methods to determine appropriate numerical ranges of the GA parameters. In this study, we would like to provide a feasible and simple method based on the engineering knowledge and computations to overcome this issue and make the GA workable to maximize the target function of maximum likelihood. Typically, engineers can have some knowledge to set up the ranges of the model parameters in an ADT study. Hence, we can make the applications of GA more efficiently by using the knowledge. The implementation of GA is presented in Figure 1 (see Ting [20]). The common termination condition(s) can be one or more combinations of the following conditions,

1. A solution is reached to meet the specific criteria.
2. The given number of iterations is reached.
3. The allocated budget is reached.

4. The highest ranking solution's fitness is reached, or the solutions cannot be improved by successive iterations.

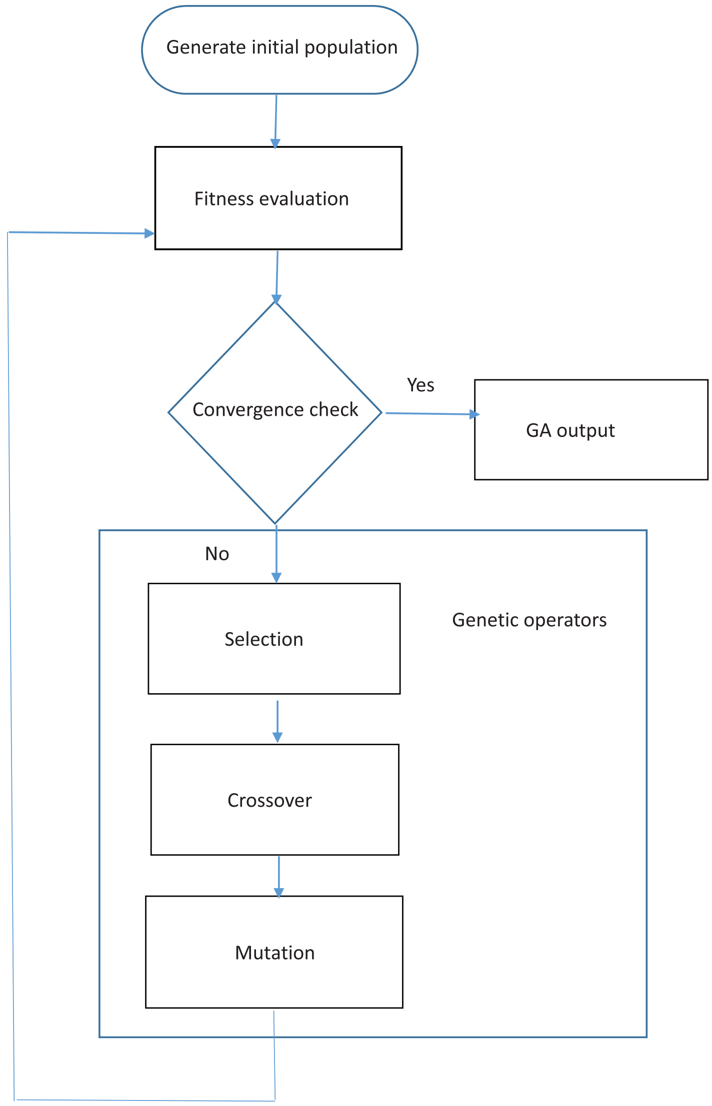

Figure 1: Flowchart of a GA.

When using the GA method, the GA parameters must be subjectively established. No absolute guidelines can be used to select these parameters. In fact, an optimal parameter combination design of GA depends on study cases. The determination of parameters, namely the population size (PS), crossover probability (CP), mutation probability (MP), and maximum number of iterations (MI), depends on the target function to be optimized and the working data sets. The reference ranges for parameters to implement a GA are 50%-100% for the PS, 60%-90% for the CP, and 5%-10% for the MP. The MI depends on the computation time.

Table 1: Factors and their levels for the GA.

Factors	Levels			
	1	2	3	4
PS	50	70	90	100
CP	0.60	0.70	0.80	0.90
MP	0.05	0.07	0.08	0.10
MI	100	150		

To avoid choosing the GA parameters subjectively, the Taguchi design method is used to reach an optimal GA design for the estimation of the GP parameters. Let the PS, CP, MP, and MI of the GA be the factors in the Taguchi design method. Table 1 displays their levels for a $2 \times 4^3 = 128$ factorial design. The L_{32} orthogonal array is used to reduce the total $2 \times 4^3 = 128$ experimental runs to 32 runs to achieve an optimal design for the PS, CP, MP, and MI of the GA. Let $y = -\log(L(\Theta; \mathbf{D}))$, then $y > 0$. The minimum signal to noise ratio (SNR) can be evaluated on the basis of repetitions of y_j for $j = 1, 2, \ldots, n_y$. The SNRs are evaluated by

$$SNR_i = -10 \times \log_{10} \left(\sum_{j=1}^{n_y} \frac{y_{ji}^2}{n_y} \right), \; i = 1, 2, \ldots, 32. \tag{2.1}$$

The optimal parameter combinations of PS, CP, MP and MI in Table 1 are used to implement the GA method for this study. The implementation of using Taguchi design method based on the data will be studied in Section 4.

3. Markov Chain Monte Carlo Methods

In this section, the MCMC methods are established for 2ADT-GP. Assume that the model parameters $\theta_1, \theta_2, \theta_3, \theta_4$, and θ_5 have a joint prior PDF $g_\Theta(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$. The posterior likelihood function is represented as

$$Pr(\Theta; \mathbf{D}) \propto L(\Theta; \mathbf{D}) g_\Theta(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5). \tag{3.1}$$

The analytic form of the marginal posterior distribution in Equation (3.1) is often difficult to obtain, and numerical integration is also difficult to implement for obtaining the marginal posterior distribution. The M-H MCMC method (see Hastings [7], Metropolis et al. [11], Ntzoufras [13]) is applied in this study to determine the MLE $\hat{\Theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4, \hat{\theta}_5)$ by establishing a noninformative prior for the prior density function $g_\Theta(.)$. In this study, the squared error loss function is considered for implementing Bayesian estimation. Because the conjugate type of prior distribution can be found only for θ_1 but no conjugate type of prior distribution for $\theta_2, \theta_3, \theta_4$, and θ_5. The joint prior PDF is given below for this study:

$$g_\Theta(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) = \pi_1(\theta_1)\pi_2(\theta_2)\pi_3(\theta_3)\pi_4(\theta_4)\pi_5(\theta_5), \tag{3.2}$$

where $\pi_1(\theta_1)$ is the PDF of the conjugate-type inverse gamma distribution, and $\pi_j(\theta_j) \propto 1$ for $j = 2, 3, 4$ and 5. Thus, $g_\Theta(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ can be presented as

$$g_\Theta(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) \propto \frac{\eta^\lambda}{\Gamma(\lambda)} \theta_1^{-\lambda-1} e^{-\eta/\theta_1}, \;\; \theta_1, \eta, \lambda > 0. \tag{3.3}$$

Then we can obtain the posterior distribution of Θ, given data $\mathbf{D}$ as

$$Pr(\Theta; \mathbf{D}) \propto \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{\theta_1^{(\delta_{ijh}+\lambda)-1}}{\Gamma(\delta_{ijh})} y_{ijh}^{\delta_{ijh}-1} e^{-(\eta+y_{ijh})/\theta_1}. \tag{3.4}$$

The Bayesian estimates of θ_i, $i = 1, 2, \ldots, 5$, are close to the MLEs for Equation (1.6) if the hyperparameters λ and η are selected to have a big variance of θ_1 according to the PDF in Equation (3.3). Using Equation (3.4), we can obtain the conditional marginal PDF of θ_i, given Θ_{-i}, where $\Theta_{-i} = (\theta_{j_1}, \theta_{j_2}, \theta_{j_3}, \theta_{j_4})$ with the integer $i \notin \{j_1, j_2, j_3, j_4\}$ and $1 \le j_1 < j_2 < j_3 < j_4 \le 5$. The analytic procedure used to obtain the conditional marginal PDF of θ_i is outlined below. To update θ_1, the following conditional marginal posterior density function of θ_1, given $\theta_2, \theta_3, \theta_4$, and θ_5, is used:

$$g_1(\theta_1; \theta_2, \theta_3, \theta_4, \theta_5) = d_1^{-1} \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \theta_1^{-(\delta_{ijh}+\lambda)-1} e^{-(\eta+y_{ijh})/\theta_1}, \tag{3.5}$$

where

$$d_1 = \int_0^\infty \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \theta_1^{-(\delta_{ijh}+\lambda)-1} e^{-(\eta+y_{ijh})/\theta_1} d\theta_1.$$

The conditional marginal posterior distribution of θ_1, given $\theta_2, \theta_3, \theta_4$, and θ_5, in Equation (3.5) is a product of inverse gamma distributions, which have the shape parameters $\delta_{ijh} + \lambda$ and scale parameters $\eta + y_{ijh}$ for $j = 1, 2, \ldots, n_i, h = 1, 2, \ldots, m_i$, and $i = 1, 2, \ldots, k$.

To update θ_j for $j = 2, 3, 4$, and 5, the following conditional marginal posterior distributions are used:

$$g_2(\theta_2; \theta_1, \theta_3, \theta_4, \theta_5) = d_2^{-1} \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh}) \theta_1^{A_0}} y_{ijh}^{A_0}, \tag{3.6}$$

$$g_3(\theta_3; \theta_1, \theta_2, \theta_4, \theta_5) = d_3^{-1} \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh}) \theta_1^{A_1}} y_{ijh}^{A_1}, \tag{3.7}$$

$$g_4(\theta_4; \theta_1, \theta_2, \theta_3, \theta_5) = d_4^{-1} \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh}) \theta_1^{A_2}} y_{ijh}^{A_2}, \tag{3.8}$$

$$g_5(\theta_5; \theta_1, \theta_2, \theta_3, \theta_4) = d_5^{-1} \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh}) \theta_1^{A_3}} y_{ijh}^{A_3}, \tag{3.9}$$

where $A_0 = \tau_{ijh} e^{\theta_2}$, $A_1 = \tau_{ijh} e^{\theta_3 L_{1i}}$, $A_2 = \tau_{ijh} e^{\theta_4 L_{2i}}$, and $A_3 = \tau_{ijh} e^{\theta_5 L_{1i} L_{2i}}$; and

$$d_2 = \int_{-\infty}^{0} \Pi_{i=1}^{k} \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh}) \theta_1^{A_0}} y_{ijh}^{A_0} d\gamma_0,$$

$$d_3 = \int_0^\infty \Pi_{i=1}^k \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh})\theta_1^{A_1}} y_{ijh}^{A_1} d\gamma_1,$$

$$d_4 = \int_0^\infty \Pi_{i=1}^k \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh})\theta_1^{A_2}} y_{ijh}^{A_2} d\gamma_2,$$

and

$$d_5 = \int_{-\infty}^\infty \Pi_{i=1}^k \Pi_{j=1}^{n_i} \Pi_{h=1}^{m_i} \frac{1}{\Gamma(\delta_{ijh})\theta_1^{A_3}} y_{ijh}^{A_3} d\gamma_3,$$

Let $G_1(\theta_1;\Theta_{-1})$, $G_2(\theta_2;\Theta_{-2})$, $G_3(\theta_3;\Theta_{-3})$, $G_4(\theta_4;\Theta_{-4})$ and $G_5(\theta_5;\Theta_{-5})$ denote the conditional cumulative distribution functions, of which the conditional PDFs are defined by Equations (3.5)-(3.9), respectively. The Bayesian method with Gibbs MCMC method can be implemented using Algorithm 1.

Algorithm 1: the Gibbs MCMC for the GP

Step 1: At iteration $i = 0$, randomly generate the Gibbs sampling initial states, $\theta_j^{(0)}$, for $j = 1, 2, \cdots, 5$, from their respective prior distribution.

Step 2: At iteration $i = i + 1$, perform the following steps:

Step 2.1: Generate $\theta_1^{(i)}$ from $G_1(\theta_1; \theta_2^{(i-1)}, \theta_3^{(i-1)}, \theta_4^{(i-1)}, \theta_5^{(i-1)})$;

Step 2.2: Generate $\theta_2^{(i)}$ from $G_2(\theta_2; \theta_1^{(i)}, \theta_3^{(i-1)}, \theta_4^{(i-1)}, \theta_5^{(i-1)})$;

Step 2.3: Generate $\theta_3^{(i)}$ from $G_3(\theta_3; \theta_1^{(i)}, \theta_2^{(i)}, \theta_4^{(i-1)}, \theta_5^{(i-1)})$;

Step 2.4: Generate $\theta_4^{(i)}$ from $G_4(\theta_4; \theta_1^{(i)}, \theta_2^{(i)}, \theta_3^{(i)}, \theta_5^{(i-1)})$;

Step 2.5: Generate $\theta_5^{(i)}$ from $G_5(\theta_5; \theta_1^{(i)}, \theta_2^{(i)}, \theta_3^{(i)}, \theta_4^{(i)})$.

Step 3: Go to Step 2 until $i = N$, where N is a huge number. Then Go to Step 4.

Step 4: Based on squared error loss function, the Bayesian estimate of θ_j is the posterior mean that can be approximated by $\frac{1}{N-N_0}\sum_{i=N_0+1}^{N} \theta_j^{(i)}$, where $N_0(< N)$ chains are used for burn-in for $j = 1, 2, \ldots, 5$.

In practice, WinBUGS or OpenBUGS are helpful for implementing the Bayesian estimation with the MCMC algorithm method. Gibbs sampling is a special case of M-H sampling that always accepts the random value. Hence, practitioners can simulate $N - N_0$ random variables of θ_i sequentially from $g_i(\theta_i; \Theta_{-i})$ (see Ntzoufras [13]). The M-H MCMC algorithm method is implemented to generate the MCMC samples $\{\theta_j^{(i)},\ j = 1, 2, \ldots, 5\}$ for $i = 1, 2, \ldots$ by using the Algorithm 2:

Algorithm 2: the M-H MCMC for GP

Step 1: At iteration $i = 0$, establish the initial states, $\theta_j^{(0)}$ for $j = 1, 2, ..., 5$, which can be generated from the respective prior distribution.

Step 2: Propose transition probabilities $q_j(\theta_j^{(*)}|\theta_j^{(i)})$ for θ_j from the ith state, $\theta_j^{(i)}$, to the $(i+1)$th state, $\theta_j^{(*)}$ for $j = 1, 2, \cdots, 5$.

Step 3: At iteration $i = i + 1$, perform the following Steps:

Step 3.1: Generate $\theta_1^{(*)}$ from $q_1(\theta_1^{(*)};\theta_1^{(i)})$ and $u \sim U(0,1)$. Update $\theta_1^{(i+1)}$ by $\theta_1^{(i+1)} = \theta_1^{(*)}$ if

$$u \le \min\left\{1, \frac{g_1(\theta_1^{(*)};\theta_2^{(i)},\theta_3^{(i)},\theta_4^{(i)},\theta_5^{(i)})}{g_1(\theta_1^{(i)};\theta_2^{(i)},\theta_3^{(i)},\theta_4^{(i)},\theta_5^{(i)})}\frac{q_1(\theta_1^{(*)}|\theta_1^{(i)})}{q_1(\theta_1^{(i)};\theta_1^{(*)})}\right\},$$

and $\theta_1^{(i+1)} = \theta_1^{(i)}$, otherwise.

Step 3.2: Generate $\theta_2^{(*)}$ from $q_2(\theta_2^{(*)};\theta_2^{(i)})$ and $u \sim U(0,1)$. Update $\theta_2^{(i+1)}$ by $\theta_2^{(i+1)} = \theta_2^{(*)}$ if

$$u \le \min\left\{1, \frac{g_2(\theta_2^{(*)};\theta_1^{(i+1)},\theta_3^{(i)},\theta_4^{(i)},\theta_5^{(i)})}{g_2(\theta_2^{(i)};\theta_1^{(i+1)},\theta_3^{(i)},\theta_4^{(i)},\theta_5^{(i)})}\frac{q_2(\theta_2^{(*)}|\theta_2^{(i)})}{q_2(\theta_2^{(i)};\theta_2^{(*)})}\right\},$$

and $\theta_2^{(i+1)} = \theta_2^{(i)}$, otherwise.

Step 3.3: Generate $\theta_3^{(*)}$ from $q_3(\theta_3^{(*)};\theta_3^{(i)})$ and $u \sim U(0,1)$. Update $\theta_3^{(i+1)}$ by $\theta_3^{(i+1)} = \theta_3^{(*)}$ if

$$u \le \min\left\{1, \frac{g_3(\theta_3^{(*)};\theta_1^{(i+1)},\theta_2^{(i+1)},\theta_4^{(i)},\theta_5^{(i)})}{g_3(\theta_3^{(i)};\theta_1^{(i+1)},\theta_2^{(i+1)},\theta_4^{(i)},\theta_5^{(i)})}\frac{q_3(\theta_3^{(*)}|\theta_3^{(i)})}{q_3(\theta_3^{(i)};\theta_3^{(*)})}\right\},$$

and $\theta_3^{(i+1)} = \theta_3^{(i)}$, otherwise.

Step 3.4: Generate $\theta_4^{(*)}$ from $q_4(\theta_4^{(*)};\theta_4^{(i)})$ and $u \sim U(0,1)$. Update $\theta_4^{(i+1)}$ by $\theta_4^{(i+1)} = \theta_4^{(*)}$ if

$$u \le \min\left\{1, \frac{g_4(\theta_4^{(*)};\theta_1^{(i+1)},\theta_2^{(i+1)},\theta_3^{(i+1)},\theta_5^{(i)})}{g_4(\theta_4^{(i)};\theta_1^{(i+1)},\theta_2^{(i+1)},\theta_3^{(i+1)},\theta_5^{(i)})}\frac{q_4(\theta_4^{(*)}|\theta_4^{(i)})}{q_4(\theta_4^{(i)};\theta_4^{(*)})}\right\},$$

and $\theta_4^{(i+1)} = \theta_4^{(i)}$, otherwise.

Step 3.5: Generate $\theta_5^{(*)}$ from $q_5(\theta_5^{(*)};\theta_5^{(i)})$ and $u \sim U(0,1)$. Update $\theta_5^{(i+1)}$ by $\theta_5^{(i+1)} = \theta_5^{(*)}$ if

$$u \le \min\left\{1, \frac{g_5(\theta_5^{(*)};\theta_1^{(i+1)},\theta_2^{(i+1)},\theta_3^{(i+1)},\theta_4^{(i+1)})}{g_5(\theta_5^{(i)};\theta_1^{(i+1)},\theta_2^{(i+1)},\theta_3^{(i+1)},\theta_4^{(i+1)})}\frac{q_5(\theta_5^{(*)}|\theta_5^{(i)})}{q_5(\theta_5^{(i)};\theta_5^{(*)})}\right\},$$

and $\theta_5^{(i+1)} = \theta_5^{(i)}$, otherwise.

Step 4: Go to Step 3 until $i = N$, where N is a huge number. Then go to Step 5.

Step 5: Based on squared error loss function, the Bayesian estimate of θ_j is the posterior mean that can be approximated by $\frac{1}{N-N_0}\sum_{i=N_0+1}^{N}\theta_j^{(i)}$, where $N_0(< N)$ chains are used for burn-in for $j = 1, 2, \ldots, 5$.

4. An Example and Simulations

Tsai et al. [25] proposed an inference method for the GP to evaluate the reliability of transistor outline can-packaged high-power LEDs. The degradation of LEDs was observed using a two-variable accelerated degradation test (ADT) with five stress loading level combinations. The absolute ambient temperature in degree Celsius and drive current in milliamperes (m$\mathring{A}$) were set to be (45°C, 650 m$\mathring{A}$), (60°C, 650 m$\mathring{A}$), (75°C, 450 m$\mathring{A}$), (75°C, 550 m$\mathring{A}$), and (75°C, 650 m$\mathring{A}$). The normal use conditions were at 25°C and 350 m$\mathring{A}$. A KEITHLEY 2430 pulse source current meter with an OL500 integrating sphere and a CAS140B spectroradiometer were used to measure the total light density of the LED source. An LED was classified as a failure if it lost 30% or more light density of its initial light density.

Using the GP model for this light degradation data set of LEDs discussed in Tsai et al. [25], the MLEs, $\hat{\theta}_1 = 0.662$, $\hat{\theta}_2 = -2.902$, $\hat{\theta}_3 = 0.577$, $\hat{\theta}_4 = 0.533$ and $\hat{\theta}_5 = 0.531$, were obtained through the method of Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (L-BFGS-B) QN, which is a modification from the L-BFGS optimization method. The L-BFGS optimization method uses a limited-memory modification for the BFGS QN method to obtain the estimates of model parameters for optimizing a specified target function. The L-BFGS-B method proposed in Byrd et al. [4] extends the L-BFGS method to handle simple box constraints on the model parameters and is hence a popular method for parameter estimation.

The QN methods are easy to use and can efficiently obtain the MLEs of model parameters if MLE determination converges within the parameter space. However, the QN methods are often sensitive to the initial parameter solutions and could not result in proper estimates because of the divergence of iterative procedure during the solution search. In this study, we evaluate the estimation performance of the L-BFGS-B QN, GA, and MCMC methods for obtaining the MLEs of the 2ADT-GP parameters. The model parameters are set up according to the MLEs from the LED example in Tsai et al. [25]; that is, we used $\theta_1 = 0.662$, $\theta_2 = -2.902$, $\theta_3 = 0.577$, $\theta_4 = 0.533$ and $\theta_5 = 0.531$ for the 2ADT-GP parameters to generate the LED damage paths. The simulation framework is given in the Algorithm 3.

Algorithm 3: The simulation framework

Step 1: Generating a simulation run, in which 60 units under the 2ADT-GP with stress loading combinations: (L'_{1i}, L'_{2i})=(25,350), (45, 650), (60, 650), (75, 450), (75,550), and (75, 650) are generated. Each stress loading combination contained 10 test units, and each unit was measured 13 times. The termination time is 26 weeks.

Step 2: Using Taguchi design with Table 2 and Step 1 to identify an optimal GA parameter combination to implement the GA method. Each GA parameter combination of PS, CP, MP and MI in Table 2 with each data set in the simulation run is implemented $n_y = 10$ times to obtain the value of SNR via Equation (2.1) to identify the optimal GA parameter combination. In this step, we repeat the GA evaluation 10 times for obtaining a reliable SNR value.

Step 3: Based on the optimal GA design that obtained in Step 2, the GA, L-BFGS-B QN and MCMC methods are used to obtain the MLEs of the 2ADT-GP parameters.

Step 4: Repeat Step 1 and Step 3 B times and labeled the MLE of θ_j in the ith simulation run by $\hat{\theta}_j^{(i)}$, $i = 1, 2, \ldots, B$ and $j = 1, 2, \ldots, 5$. The bias and MSE of $\hat{\theta}_j$ are evaluated by $\left(\frac{1}{B}\sum_{i=1}^{B}\hat{\theta}_j^{(i)}\right) - \theta_j$ and $\frac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_j^{(i)} - \theta_j)^2$, respectively for $j = 1, 2, \ldots, 5$. In this study, we adopt $B = 10000$.

Base on the knowledge of the degradation measurements of LED discussed in Park and Padgett [16], the parameter space is suggested as $\Omega = \{\theta_1, \theta_3, \theta_4 > 0, \theta_2 < 0 \text{ and } \theta_5 \in R\}$. Because the parameter θ_2 is related to the MTTF of LEDs under the normal use condition and this type of LED is a highly reliable product, θ_2 should be negative and not close to 0. The parameter θ_1 in the gamma distribution must not be too high. This knowledge facilitates the selection of the initial parameter solutions and the prior distribution when using the L-BFGS-B QN, GA, and MCMC methods.

Using the simulated GP data sets, with the parameters $\theta_1 = 0.662$, $\theta_2 = -2.902$, $\theta_3 = 0.577$, $\theta_4 = 0.533$ and $\theta_5 = 0.531$, to establish the Taguchi L_{32} design, we obtained 10 values of y from GP and the corresponding SNR for each combination of factor levels. Table 2 and Figure 2 summarize these results.

Figure 2 shows the main effect plots for SNRs. The parameters of PS = 100, CP = 0.7, MP = 0.1, and MI = 150 are suggested for using the GA method to obtain the MLEs of $\hat{\theta}_j$, $j = 1, 2, \ldots, 5$ for the 2ADT-GP modeling of the simulated LED data set. On the basis of the normalized formulas for stress levels in (A3), the parameter spaces of θ_3 and θ_4 can be of positive real numbers over small intervals. Moreover, the parameter space of θ_5 can be extended from the parameter spaces of θ_3 and θ_4 to also include the negative part. Hence, consider the parameter space

$$B_1 = \{0 < \theta_1 < 1.5, -4.5 < \theta_2 < 0, 0 < \theta_3, \theta_4 < 3, -3 < \theta_5 < 3\} \tag{4.1}$$

for the 2ADT-GP model with this type LED data set. Using Algorithm 2: M-H MCMC through OpenBUGS package with a non-informative prior inverse Gamma for θ_1, burn-in $N_0 = 1000$ and $N = 10000$ chains to obtain a MCMCG MLEs for θ_j, $j = 1, 2, \ldots, 5$. Please note that the OpenBUGS can decide symmetric transition probabilities to implement the MCMC method. Finally, boxplots for data sets each contains 1000 GA MLEs, QN MLEs, GA-QN MLEs, MCMCG MLEs and MCMCU MLEs, respectively, are generated.

Table 3 displays biases and MSEs for GA MLEs, QN MLEs and GA-QN MLEs. In this study, the initial inputs for searching the MLEs, through using the L-BFGS-B QN method, was randomly generated from the uniform distribution over the parameter space of θ_j, $j = 1, 2, \ldots, 5$. Compared with the GA MLEs in Table 3, the L-BFGS-B QN method was an unstable approach that yielded high MSEs for the estimates of θ_2 to θ_5. Moreover, the bias estimates of the QN MLEs of θ_2 to θ_5 were also higher than those of the GA MLEs. The mean of QN MLEs of θ_5 in 10000 simulation runs are negative,

Table 2: Taguchi L_{32} design and SNRs.

PS	CP	MP	MI	y_1	y_2	y_3	y_4	y_5	y_6
50	0.6	0.05	100	941.922	1015.15	916.864	952.822	999.314	919.111
70	0.6	0.07	100	976.173	933.737	955.192	969.113	944.17	964.909
90	0.6	0.08	100	974.916	950.927	948.661	942.994	932.631	894.729
100	0.6	0.1	100	960.832	941.545	919.696	942.98	937.412	948.339
50	0.7	0.05	100	967.042	934.785	946.898	960.639	958.28	854.763
70	0.7	0.07	100	931.039	934.4	913.039	926.903	962.288	917.379
90	0.7	0.08	100	957.906	886.611	962.214	945.084	895.959	914.5
100	0.7	0.1	100	958.622	935.851	911.972	905.114	955.148	917.728
70	0.8	0.05	100	923.606	1012.57	910.022	936.256	962.264	934.382
50	0.8	0.07	100	995.235	941	963.047	979.953	894.779	912.932
100	0.8	0.08	100	922.478	889.798	900.425	885.879	950.364	945.046
90	0.8	0.1	100	943.006	939.949	912.794	920.925	957.454	934.615
70	0.9	0.05	100	963.385	881.436	930.387	984.857	948.986	910.278
50	0.9	0.07	100	939.611	997.745	963.69	916.826	951.908	935.655
100	0.9	0.08	100	931.261	931.844	973.279	928.076	923.646	914.948
90	0.9	0.1	100	905.884	926.517	928.559	935.173	965.714	926.65
100	0.6	0.05	150	920.87	956.541	898.705	996.758	917.899	926.316
90	0.6	0.07	150	886.139	922.117	938.448	945.65	950.452	949.832
70	0.6	0.08	150	942.36	967.161	923.389	908.16	936.322	910.826
50	0.6	0.1	150	942.879	956.175	949.609	900.205	941.156	906.507
100	0.7	0.05	150	952.917	929.661	936.219	915.545	922.269	904.73
90	0.7	0.07	150	955.409	952.785	921.467	945.344	895.968	964.978
70	0.7	0.08	150	885.627	904.002	969.674	891.222	936.856	902.418
50	0.7	0.1	150	894.707	909.705	910.821	955.033	917.511	910.152
90	0.8	0.05	150	900.643	936.683	895.132	955.935	955.115	956.195
100	0.8	0.07	150	966.335	964.529	910.257	934.316	908.35	919.595
50	0.8	0.08	150	940.544	956.166	924.477	939.998	975.761	944.249
70	0.8	0.1	150	940.513	911.937	874.946	917.261	936.731	907.965
90	0.9	0.05	150	918.969	906.899	945.306	940.268	970.355	947.742
100	0.9	0.07	150	928.483	947.861	924.293	938.196	940.554	952.37

Table 3: (Continued)

PS	CP	MP	MI	y_7	y_8	y_9	y_{10}	SNR
50	0.6	0.05	100	943.97	955.232	964.817	894.869	-59.564
70	0.6	0.07	100	916.532	920.574	938.605	915.584	-59.497
90	0.6	0.08	100	951.162	951.925	973.953	929.674	-59.512
100	0.6	0.1	100	883.455	962.049	927.788	931.987	-59.424
50	0.7	0.05	100	880.853	931.221	1009.26	907.465	-59.426
70	0.7	0.07	100	947.423	955.094	924.528	1018.82	-59.495
90	0.7	0.08	100	945.971	924.306	926.089	918.282	-59.351
100	0.7	0.1	100	906.374	955.818	894.021	888.269	-59.306
70	0.8	0.05	100	899.413	906.396	933.247	952.535	-59.44
50	0.8	0.07	100	957.27	918.571	935.919	865.005	-59.436
100	0.8	0.08	100	952.1	935.82	902.225	927.586	-59.29
90	0.8	0.1	100	895.733	933.263	948.171	930.519	-59.386
70	0.9	0.05	100	970.313	897.97	958.887	945.428	-59.46
50	0.9	0.07	100	937.344	924.883	963.201	913.307	-59.506
100	0.9	0.08	100	950.507	970.927	916.068	966.621	-59.471
90	0.9	0.1	100	921.832	883.787	922.91	945.634	-59.337
100	0.6	0.05	150	938.692	970.77	953.326	941.263	-59.486
90	0.6	0.07	150	973.094	932.044	922.64	975.078	-59.461
70	0.6	0.08	150	916.769	906.009	937.882	915.152	-59.338
50	0.6	0.1	150	904.712	855.527	954.7	945.921	-59.335
100	0.7	0.05	150	944.015	927.151	899.213	957.453	-59.361
90	0.7	0.07	150	926.205	961.607	939.539	938.89	-59.467
70	0.7	0.08	150	930.619	967.797	958.365	918.997	-59.342
50	0.7	0.1	150	892.004	932.851	914.971	882.234	-59.202
90	0.8	0.05	150	907.104	961.365	929.925	952.747	-59.42
100	0.8	0.07	150	933.607	894.608	895.409	924.363	-59.327
50	0.8	0.08	150	968.709	912.943	933.983	946.731	-59.504
70	0.8	0.1	150	916.158	902.493	911.112	892.96	-59.194
90	0.9	0.05	150	944.734	945.122	983.316	913.659	-59.48
100	0.9	0.07	150	888.598	889.195	968.405	877.18	-59.332
50	0.9	0.08	150	951.933	954.936	922.322	901.823	-59.314
70	0.9	0.1	150	885.347	950.314	926.706	905.222	-59.385

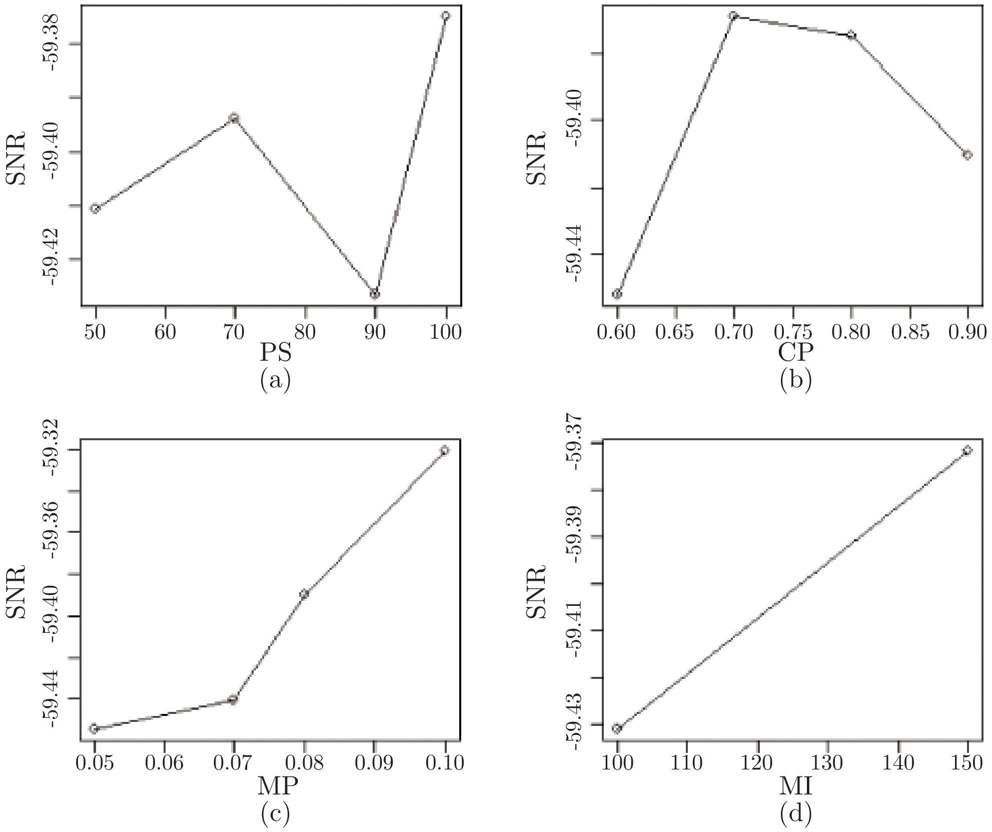

Figure 2: Main effect plots for SNRs.

which has an opposite sign of $\theta_5 = 0.531$. After carefully checking all simulation outputs, we found that some QN MLEs used the boundary values of their solution spaces in many simulation runs. These simulation runs can be considered a type of divergence because of the absence of proper MLE solutions in their parameter spaces. Because the L-BFGS-B QN method is sensitive to the initial parameter solutions, improper initial inputs could also be one reason for the improper MLEs produced by the L-BFGS-B QN method. In practice, selecting suitable initial parameter solutions for the L-BFGS-B QN method is difficult. On the basis of this simulation study, we found that the only strength of the L-BFGS-B QN method over the GA method is that the QN-MLE of θ_1 had smaller bias and MSE than those of the GA-MLE of θ_1. The GA-MLE of θ_1 underestimated the true θ_1.

It could be a good idea to use the GA MLEs as the initial parameter inputs to search the QN MLEs. That is, we implement the L-BFGS-B QN method with two steps. The first step is to find the GA MLEs, then go to the second step to use the GA MLEs as initial parameter inputs to search the MLEs of model parameters based on the L-BFGS-B QN method. In this study, the QN MLEs using the GA MLEs as initial solutions are labeled as GA-QN MLEs. Table 3 shows that the GA-QN MLEs significantly improve

Table 4: Bias and MSEs of the MLEs.

	θ_1	θ_2	θ_3	θ_4	θ_5
GA MLEs					
mean	0.0073	-2.9198	0.5924	0.5844	0.5817
bias	-0.6547	-0.0178	0.0154	0.0514	0.0507
MSE	0.4287	0.0892	0.0488	0.0526	0.0637
QN MSEs					
mean	0.7620	-2.5791	1.3128	1.2963	-0.6063
bias	0.1000	0.3229	0.7358	0.7663	-1.1373
MSE	0.3089	1.9902	1.6762	1.7074	5.8705
GA-QN MSEs					
mean	0.6590	-2.8889	0.5746	0.5310	0.5363
Bias	-0.0030	0.0131	-0.0024	-0.0020	0.0053
MSE	0.0026	0.0288	0.0141	0.0146	0.0332

the drawbacks of QN MLEs and perform better than the QN MLEs and GA MLEs with smaller bias and MSE.

In fact, the parameter space B_1 is obtained based on the knowledge of the LED discussed in Tsai et al. [25]. However, the L-BFGS-B QN method cannot identify effective solutions over B_1. The GA method estimates the parameter θ_1 poorly, with a large bias and MSE. The GA method identifies optimal solutions over B_1 using an artificial intelligence computing method without assuming initial parameter solutions. The performance of the GA method depends on the number of parameters, the complexity of the target function, and the size of the parameter space. The parameters PS = 100, CP = 0.7, MP = 0.1 and MI = 150, which are obtained from the Taguchi design method, can be used in the GA method to reduce the subjectivity of parameter selection. Compared with the L-BFGS-B QN method, the GA method requires more computation time to obtain the MLEs for parameter estimation. According to Table 3, compared with the L-BFGS-B QN method, the GA method exhibits higher performance for obtaining the MLEs of the model parameters except θ_1.

If the knowledge of the parameter space is insufficient, the L-BFGS-B QN method is expected to work relatively poorly and introduce divergence in more simulation cases. For example, considering the following parameter space:

$$B_2 = \{0 < \theta_1 < 1.5, -4.5 < \theta_2 < 0, 0 < \theta_3, \theta_4 < 5, -5 < \theta_5 < 5\}. \qquad (4.2)$$

In B_2, we assume insufficient knowledge of the parameters θ_3, θ_4, and θ_5 such that a wide range of these three parameters is introduced for searching the MLEs of the 2ADT-GP model parameters on basis of the LED data set. To implement the MCMC method, we consider the noninformative prior distributions of the gamma with $\eta = 0.001$ and $\lambda = 0.0006$ for θ_1, uniform distribution over $(-4.5, 0)$ for θ_2, uniform distribution over (0, 5) for θ_3 and θ_4, and uniform distribution over $(-5, 5)$ for θ_5. The MLEs $\hat{\theta}_1$ to $\hat{\theta}_5$ are obtained through the M-H MCMC method. The estimation performance of the L-BFGS-B QN, GA, GA-QN and M-H MCMC methods with the parameter space B_2

are compared on the basis of 1000 simulation runs. Figures 3-7 display the simulation results in boxplots.

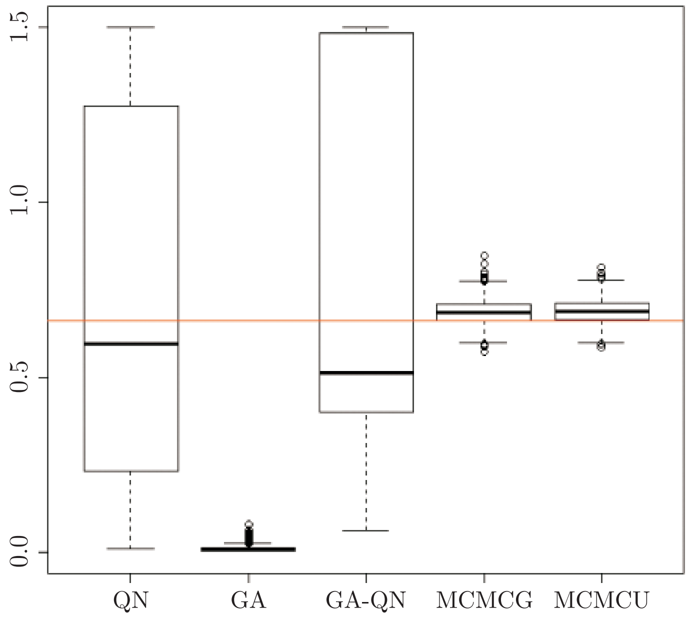

Figure 3: The boxplots of the MLEs of θ_1.

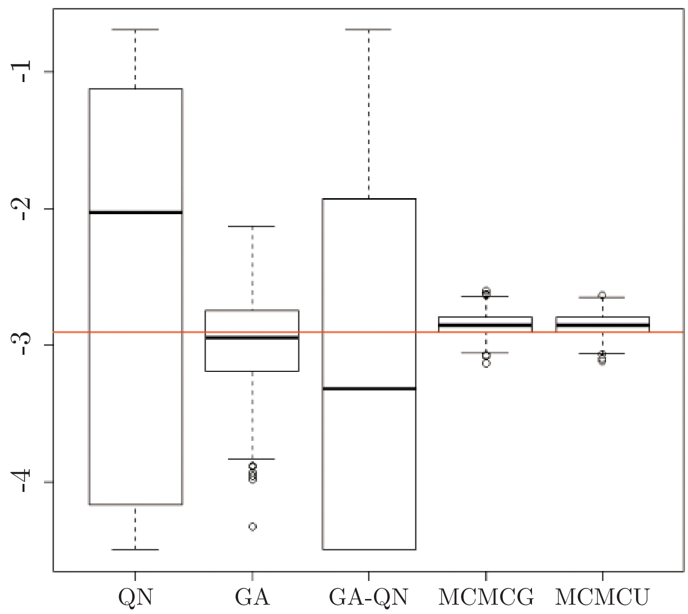

Figure 4: The boxplots of the MLEs of θ_2.

According to Figures 3-7, the GA and L-BFGS-B QN methods perform poorly if

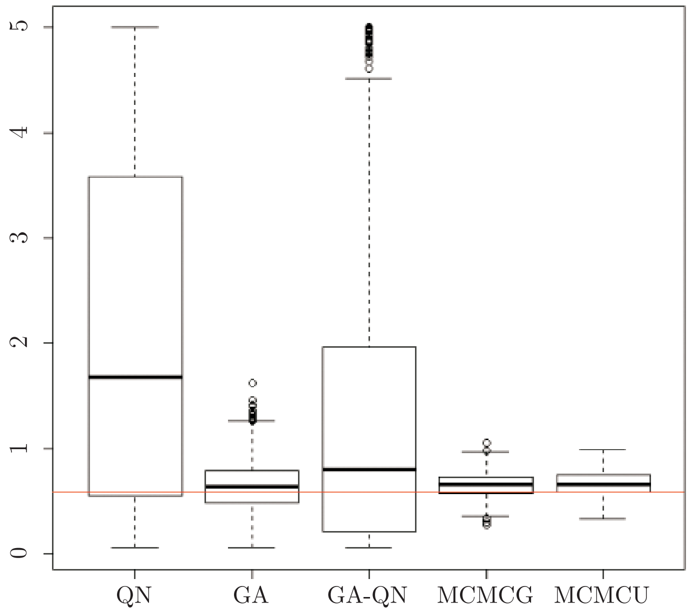

Figure 5: The boxplots of the MLEs of θ_3.

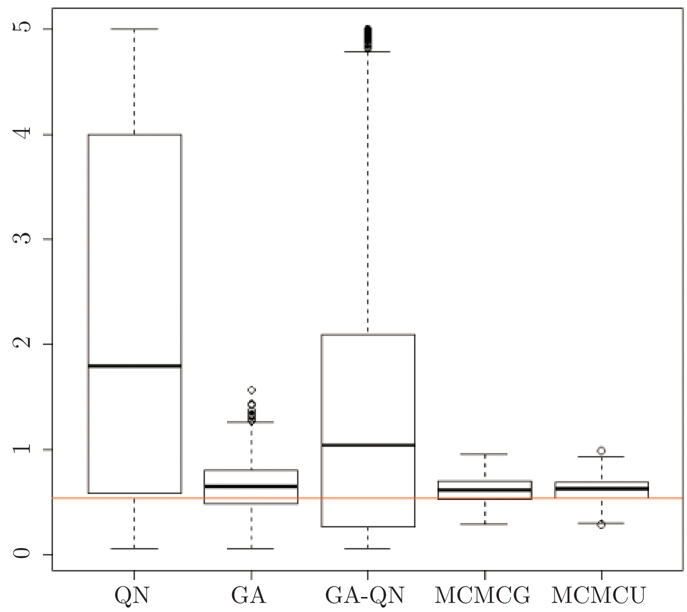

Figure 6: The boxplots of the MLEs of θ_4.

the parameter space B_2 is used to search the MLEs of the GP model parameters. As expected, divergence is found for many simulation cases in using the L-BFGS-B QN method under the parameter space B_2, and the L-BFGS-B QN method leads to larger bias and MSEs. Compared the QN MLEs with the GA MLEs in Figures 3-7, the bias

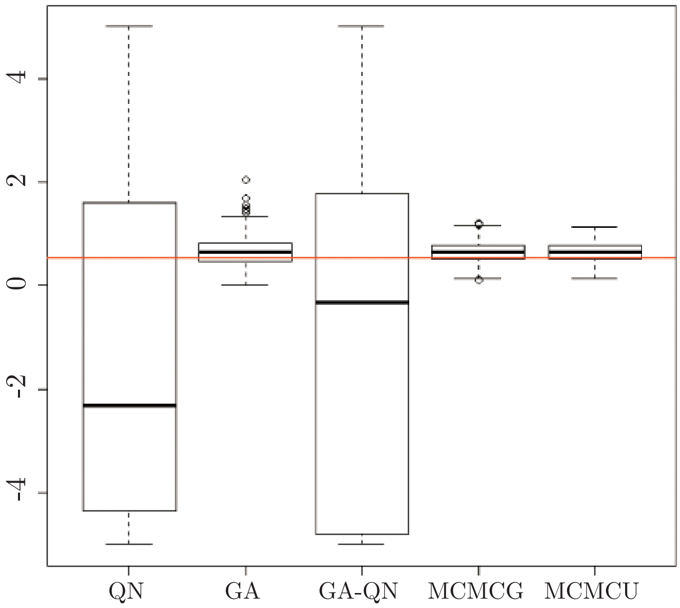

Figure 7: The boxplots of the MLEs of θ_5.

and MSEs of the GA MLEs are more stable, but the GA method fails to provide a good estimate for the parameter θ_1. The bias and MSEs are inflated when the parameter space B_2 is used for the GA method. The GA method cannot work well in the parameter space B_2, and then the L-BFGS-B QN method using the GA MLEs as initial parameter solutions only slightly improves the estimation performance. Carefully checking Figures 3-7, we find the GA-QN MLEs have smaller bias and MSEs than those of the QN MLEs. But most of the GA-QN MLEs of θ_5 are negative, and this fact indicates that the GA-QN MLE cannot be a good estimate of the parameter $\theta_5 = 0.531$ when the parameter space B_2 is applied.

The M-H MCMC method is stable when the parameter space B_2 is used to search the MLEs of the GP model parameters. In Figures 3-7, MCMCG MLE represents the M-H MCMC, which uses the gamma distribution with $\eta = 0.001$ and $\lambda = 0.0006$ as the noninformative prior distribution of θ_1 and the $\pi_j(\theta_j) \propto 1$ as the noninformative prior distributions for θ_j, $j = 2, 3, 4$, and 5. The means of the MCMCG MLEs of θ_j's over 1000 simulation runs are close to the true parameter values, and the bias and MSEs are smaller. Similar estimation results can be obtained if other noninformative prior distributions are used. For illustration, we assume that the noninformative prior of θ_1 is $U(0, 1.5)$ and use $\pi_j(\theta_j) \propto 1$ as the noninformative prior distributions for θ_j, $j = 2, 3, 4$, and 5. For brevity, only the estimation results are shown in Figures 3-7. The analytic M-H MCMC with the new noninformative prior distributions can be derived using a similar procedure as that outlined in Section 3. Let MCMCU MLE denote the M-H MCMC MLE, which uses $U(0, 1.5)$ as the prior distribution for θ_1. According to Figures 3-7, even with the noninformative prior uniform distribution for θ_1, the estimation results are still reliable and have small bias and MSEs. Moreover, the means of the MLEs over 1000 simulation runs are close the true parameters. The results in Figures 3-7 indicate that the MCMC method is less sensitive to the choice of the parameter space, and that it exhibits higher performance for implementing parameter estimation for GP model than the GA, L-BFGS-B QN, and GA-QN methods.

On the basis of the simulation results, we found that the M-H MCMC method can be used to obtain reliable MLEs of the 2ADT-GP parameters if we have knowledge on the parameter space. The M-H MCMC method based MLE performs better than the GA, L-BFGS-B QN method and GA-QN method with small bias and MSE under sharing the same knowledge on the parameter space. The subjective of using M-H MCMC method is to set up the prior distribution. In the simulation study, we tried two different noninformative prior distributions and obtained similar estimation results. Figures 3-7 show that both M-H MCMC estimates, which were obtained by using different noninformative prior distributions, performs better than the other competitors. We hence conclude that the M-H MCMC method is more reliable to obtain the estimates of the 2ADT-GP parameters than the GA, L-BFGS-B QN and GA-QN methods.

5. Conclusion

In this paper, the GA method, L-BFGS-B QN method, GA-QN method and two MCMC methods are suggested to implement parameter estimation for the 2ADT-GP model with GEM. The 2ADT-GP model has been widely used to evaluate the reliability of highly reliable products. The Taguchi design method is used to reduce the subjectivity of parameter selection when implementing the GA method. The analytic Gibbs MCMC method and the M-H MCMC method are implemented for the GP model. Both MCMC algorithm methods can be easily implemented through the WinBUGS or OpenBUGS packages. Intensive simulations are conducted to evaluate the performance of the aforementioned parameter estimation methods for the GP model on the basis of the lumen degradation data set of LEDs.

The simulation results show that the L-BFGS-B QN method may fail to determine the MLEs of the GP model parameters because of a divergence problem. This divergence is mainly caused by the use of improper initial parameter solutions to determine the MLEs of the 2ADT-GP parameters. The GA method is an artificial intelligence computing method that avoids this divergence problem and exhibits higher performance than the L-BFGS-B QN method. However, the GA method cannot provide a reliable MLE for the scale parameter of the gamma distribution. If the GA MLEs are suitable, combining the GA and L-BFGS-B QN methods can be a good estimation method for inferring on the reliability of 2ADT-GP model. The L-BFGS-B QN, GA and GA-QN methods cannot identify effective parameter solutions if the practitioners have insufficient knowledge to choose the parameter space. In an ADT study, the engineers often keep knowledge to roughly set up the ranges of the model parameters. Using engineering knowledge to set up the ranges of the model parameters can make the application of GA more efficiently. The Taguchi design method is used to set up the GA parameters for a specific data set. The Taguchi method is easy to implement and the users can also consider using the numerical ranges suggested from literature studies to calibrate the ranges of the GA parameters.

The MCMC method can overcome the drawbacks of the L-BFGS-B QN, GA and GA-QN methods, and the MCMC method is less sensitive to the choice of the parameter space. Hence, the MCMC method is recommended for searching the MLEs of the 2ADT-GP model parameters. The two MCMC algorithms proposed in this study are reliable for obtaining the MLEs of the 2ADT-GP model parameters, and both algorithms can be easily applied through the WinBUGS or OpenBUGS packages. Additional studies should extend the proposed estimation procedures to other stochastic processes for two-variable ADTs.

Acknowledgements

This study was supported by a grant from the Ministry of Science and Technology, Taiwan MOST 106-2221- E-032-038-MY2.

References

[1] Akbari, Z. (2011). *A multilevel evolutionary algorithm for optimizing numerical functions*, International Journal of Industrial Engineering Computations, Vol.2, 419-430.

[2] Baudry, B., Fleurey, F., Jzquel, J. M. and Le Traon, Y. (2005). *Automatic test case optimization: a bacteriologic algorithm*, IEEE Software, Vol.22, 76-82.

[3] Boulanger, M. and Escobar, L. A. (1994). *Experimental design for a class of accelerated degradation tests*, Technometrics, Vol.36, 260-272.

[4] Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C. (1995). *A limited memory algorithm for bound constrained optimization*, SIAM Journal on Scientific Computing, Vol.16, 1190-1208.

[5] Ferreira, C. (2001). *Gene expression programming: a new adaptive algorithm for solving problems*, Complex Systems, Vol.13, 87-129.

[6] Guan, Q. and Tang, Y.-C. (2013). *Optimal design of accelerated degradation test based on gamma process models*, Chinese Journal of Applied Probability and Statistics, Vol.29, 213-224.

[7] Hastings, W. K. (1970). *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, Biometrika, Vol.57, 97-109.

[8] Holland, J. H. (1992). Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, MIT Press, second edition. (First edition, University of Michigan Press, 1975).

[9] Liao, C.-M. and Tseng, S.-T. (2006). *Optimal design for step-stress accelerated degradation tests*, IEEE Transactions on Reliability, Vol.55, 59-66.

[10] Lim, H. and Yum, B.-J. (2011). *Optimal design of accelerated degradation tests based on Wiener process models*, Journal of Applied Statistics, Vol.38, 309-325.

[11] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller A. H. and Teller, E. (1953). *Equation of state calculations by fast computing machines*, Journal of Chemical Physics, Vol.21, 1087-1092.

[12] Nicholson, M. (1998). Genetic algorithms and grouping problems, John Wiley & Sons.

[13] Ntzoufras, I. (2009). Bayesian Modeling Using WinBUGS. Hoboken, NJ: John Wiley & Sons.

[14] Padgett, W. J. and Tomlinson, M. A. (2004). *Inference from accelerated degradation and failure data based on Gaussian process models*, Lifetime Data Analysis, Vol.10, 191-206.

[15] Park, C. and Padgett, W. J. (2005). *Accelerated degradation models for failure based on geometric Brownian motion and gamma process*, Lifetime Data Analysis, Vol.11, 511-527.

[16] Park, C. and Padgett, W. J. (2006). *Stochastic degradation models with several accelerating variables*, IEEE Transactions on Reliability, Vol.55, 379-390.

[17] Park, C. and Padgett, W. J.(2007). *Cumulative damage models for failure with several accelerating variables*, Quality Technology & Quantitative Management, Vol.4, 17-34.

[18] Peng, J.-Y. (2012). *A note on optimal allocations for the second elementary symmetric function with applications for optimal reliability design*, Naval Research Logistics, Vol.59, 278-284.

[19] Peng, C.-Y. and Tseng, S.-T. (2013). *Statistical lifetime inference with skew-Wiener linear degradation models*, IEEE Transactions on Reliability, Vol.62, 338-350.

[20] Scrucca, L. (2013). *GA: A package for Genetic Algorithm in R*, Journal of Statistical Software, Vol.53, 1-37.

[21] Ting, C.-K. (2005). *On the mean convergence time of multi-parent genetic algorithms without selection*, Advances in Artificial Life, Vol.3630, 403-412.

[22] Tsai, T.-R., Lin, C.-W., Sung, Y.-L., Chou, P.-T., Chen C.-L. and Lio, Y. L. (2012). *Inference from lumen degradation data under Wiener diffusion process*, IEEE Transactions on Reliability, Vol.61, 710-718.

[23] Tsai, C.-C., Tseng, S.-T. and Balakrishnan, N. (2013). *Optimal burn-in policy for highly reliable products using gamma degradation process*, IEEE Transactions on Reliability, Vol.60, 338-350.

[24] Tsai, T.-R., Lio, Y. L. and Jiang, N. (2014). *Optimal decision on the accelerated degradation test plan under the Wiener process*, Quality Technology & Quantitative Management, Vol.11, 461-470.

[25] Tsai, T.-R., Sung, W-Y, Lio, Y. L., Chang, S. and Lu, J.-C. (2016). *Optimal two-variable accelerated degradation test plan for gamma degradation processes*, IEEE Transactions on Reliability, Vol.65, 459-468.

[26] Tseng, S.-T., Balakrishnan N. and Tsai, C.-C. (2011). *Optimal step-stress accelerated degradation test plan for gamma degradation process*, IEEE Transactions on Reliability, Vol.58, 234-245.

[27] Whitmore, G. A. (1995). *Estimating degradation by a Wiener diffusion process subject to measurement error*, Lifetime Data Analysis, Vol.1, 307-319.

[28] Zhang, J., Chung, H., and Lo, W. L. (2007). *Clustering-based adaptive crossover and mutation probabilities for genetic algorithms*, IEEE Transactions on Evolutionary Computation, Vol.11, 326-335.

School of Statistics, Southwestern University of Finance and Economics, Chengdu 610074, China.

E-mail: 2981104486@qq.com

Major area(s): Process optimization using correlated variables, quality control, reliability analysis.

School of Management and Data-Mining Research Center, Xiamen University, Xiamen 361005, China.

E-mail: xmjpzhu@xmu.edu.cn

Major area(s): Data analysis and modelling, data science and business intelligence, application of mathematical statistics, econometric model and application.

Department of Applied Mathematics, Chung Yuan Christian University, Chung Li City 32023, Taiwan.

E-mail: yujaulin@gmail.com

Major area(s): Bayesian statistics, time series, stochastic process.

Department of Mathematical Sciences, University of South Dakota, Vermillion, SD 57069, USA.

E-mail: Yuhlong.Lio@usd.edu

Major area(s): Reliability, smooth estimation, survival analysis, statistical process control.

(Corresponding author) Department of Statistics, Tamkang University, Tamsui District, New Taipei City 25137, Taiwan.

E-mail: tzongru@gms.tku.edu.tw

Major area(s): Quality control, reliability analysis.

(Received March 2017; accepted July 2018)