

Bayesian Estimation in Piecewise Constant Model with Gamma Noise by Using Reversible Jump MCMC

Suparman

Department of Mathematics Education, University of Ahmad Dahlan, Indonesia

Received August 3, 2019; Revised September 25, 2019; Accepted February 20, 2020

Copyright©2020 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract A piecewise constant model is often applied to model data in many fields. Several noises can be added in the piecewise constant model. This paper proposes the piecewise constant model with a gamma multiplicative noise and a method to estimate a parameter of the model. The estimation is done in a Bayesian framework. A prior distribution for the model parameter is chosen. The prior distribution for the parameter model is multiplied with a likelihood function for the data to build a posterior distribution for the parameter. Because a number of models are also parameters, a form of the posterior distribution for the parameter is too complex. A Bayes estimator cannot be calculated easily. A reversible jump Monte Carlo Markov Chain (MCMC) is used to find the Bayes estimator of the model parameter. A result of this paper is the development of the piecewise constant model and the method to estimate the model parameter. An advantage of this method can simultaneously estimate the constant piecewise model parameter.

Keywords Bayesian, Gamma Noise, Piecewise Constant, Reversible Jump MCMC

multiplicative noise is also used by several authors, for example [7-10]. The multiplicative noise is used as a measurement error in line transect sampling [7]. An asymptotic Cramer-Rao bound is discussed for frequency estimation in the multiplicative noise [8]. Adsorption of ligands on DNA is considered for an arbitrary filling in the presence of multiplicative noise [9]. A multiplicative noise is used in a segmentation [10].

Noise in a mathematical model is assumed to be of a certain distribution, for example [11-14]. Gaussian additive noise is used in the piecewise constant model [11]. Exponential additive noise is used in an autoregressive model [12-14]. However, in some applications, the data is often modeled following the piecewise constant model with a gamma multiplicative noise. The Gamma is a distribution that is more general than an exponential distribution. The exponential distribution is a particular case of the Gamma distribution. If the piecewise constant model with the Gamma multiplicative noise used to model the data, the model parameters are unknown. The model parameters include a number of constant models, a location of constant model changes, a constant model height, and a noise variance. This study proposes an estimation method of the piecewise constant model that has a Gamma multiplicative noise where the number of constant models is unknown.

1. Introduction

A piecewise constant model is a model used to model data in many fields, for example [1-3]. The piecewise constant model is used for smoothing images of flowers [1]. The piecewise constant model is used for a population size modeling [2], [3]. The piecewise constant model can contain an additive noise or a multiplicative noise. The additive noise is considered by various authors, for example [4-6]. The additive noise is added to a spatial regression model [4]. The additive noise is used in a partially linear functional model[5]. This linear functional model is partly applied to tecator data. The additive noise is used in a log regression model [6]. On the other hand, a

2. Method

A Bayesian framework is adopted to estimate the parameters [15]. A prior distribution for the number of constant models, the location of changes in the constant model, the constant height of the model, and the noise variance are selected. Then this prior distribution is combined with a likelihood function of the data to get a posterior distribution. Based on this posterior distribution, a Bayes estimator for the number of constant models, the location of changes in the constant model, the constant height of the model, and the noise variances are estimated.

A reversible jump Monte Carlo Markov Chain (MCMC) method [16] was proposed to determine the Bayes estimator. The basic idea of the MCMC reversible jump method is a creation of the Markov chain that is recurrent and irreducible such that limit distribution of the Markov chain will be equal to the posterior distribution. Furthermore, a resulting Markov chain is used to calculate estimator for the parameters.

3. Results and Discussion

Suppose that n represents a number of data and y_1, \dots, y_n represent a data set. These data follow the piecewise constant model if these data satisfy the following mathematical equation:

$$y_t = m_t z_t, t = 1, \dots, n \quad (1)$$

where

$$m_t = \begin{cases} h_1, & \tau_1 < t \leq \tau_2 \\ h_2, & \tau_2 < t \leq \tau_3 \\ \dots & \dots \\ h_{k+1}, & \tau_{k+1} < t \leq \tau_{k+2} \end{cases} \quad (2)$$

with $\tau_1 = 0$ and $\tau_{k+2} = n$. The value of k denotes the number of constant models. The values of $\tau = (\tau_1, \dots, \tau_k)$ state the location of the change in the constant model. The value of $h = (h_1, \dots, h_{k+1})$ expresses the height of the constant model. Here, z_t is assumed to have the Gamma distribution with the parameters $\alpha > 0$ and $\beta > 0$.

3.1. Likelihood Function

The random variable z_t is distributed Gamma so that the probability function for z_t can be written as

$$g(z_t | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} z_t^{\alpha-1} \exp - \beta z_t \quad (3)$$

Here $\alpha = 1$. Suppose that $y = (y_1, \dots, y_n)$. By using variable transformation, a likelihood function for data y is

$$\begin{aligned} f(y | k, \tau, h, \alpha, \beta) &= \prod_{i=1}^{k+1} \prod_{t=\tau_{i-1}+1}^{\tau_i} \frac{\beta^\alpha}{\Gamma(\alpha)} y_t^{\alpha-1} \frac{1}{h_i^\alpha} \exp - \beta \frac{y_t}{h_i} \\ &= \left(\frac{\beta^\alpha}{\Gamma(\alpha)} \right)^n \prod_{i=1}^{k+1} y_t^{(\alpha-1)n_i} \frac{1}{h_i^{\alpha n_i}} \exp - \beta \frac{s_i}{h_i} \end{aligned} \quad (4)$$

where $s_i = \sum_{t=\tau_{i-1}+1}^{\tau_i} (y_t - m_t)$ and $n_i = \tau_i - \tau_{i-1}$ for $i = 1, \dots, k+1$.

3.2. Prior Distribution

To obtain a posterior distribution, a prior distribution must be determined. As in [10], the prior distribution for k is chosen of a Binomial distribution with a parameter $0 < \lambda < 1$. For $k = 0, 1, \dots, k_{max}$

$$\pi(k | \lambda) = \binom{k_{max}}{k} \lambda^k (1 - \lambda)^{k_{max}-k} \quad (5)$$

Where k_{max} states the maximum value of k . A hyperprior distribution for λ is chosen as a uniform

distribution. A prior distribution for τ_1, \dots, τ_k according to ordered statistics.

$$\pi(\tau_1, \dots, \tau_k | k) = \frac{1}{C_{n-2}^{2k+1}} \prod_{i=1}^{k+1} (n_i - 1) \quad (6)$$

A prior distribution for h_1, \dots, h_{k+1} is chosen inversely Gamma distribution with parameter $u > 0$ and $v > 0$.

$$\begin{aligned} \pi(h_1, \dots, h_{k+1} | k, u, v) &= \prod_{i=1}^{k+1} \frac{v^u}{\Gamma(u)} h_i^{-u-1} \exp - v \frac{1}{h_i} \\ &= \left(\frac{v^u}{\Gamma(u)} \right)^{(k+1)} \left(\prod_{i=1}^{k+1} h_i \right)^{-u-1} \\ &\quad \exp - v \sum_{i=1}^{k+1} \frac{1}{h_i} \end{aligned} \quad (7)$$

Here, $u = 1$ and Jeffreys prior distribution are chosen as a hyperprior distribution for v , i.e., $\pi(v) \propto v^{-1}$. Similarly, Jeffreys prior distribution is also selected as a hyperprior distribution for β , i.e., $\pi(\beta) \propto \beta^{-1}$. So the prior distribution for the parameters $(k, \tau, h, \lambda, v, \beta)$ can be written as

$$\begin{aligned} \pi(k, \tau, h, \alpha, \beta) &= C_k^{k_{max}} \lambda^k (1 - \lambda)^{k_{max}-k} \frac{1}{C_{n-2}^{2k+1}} \\ &\quad \prod_{i=1}^{k+1} (n_i - 1) \left(\frac{v^u}{\Gamma(u)} \right)^{(k+1)} \left(\prod_{i=1}^{k+1} h_i \right)^{-u-1} \\ &\quad \exp - v \sum_{i=1}^{k+1} \frac{1}{h_i} \frac{1}{\beta} \frac{1}{v} \end{aligned} \quad (8)$$

3.3. Posterior distribution

Let $H_1 = (k, \tau, h)$ and $H_2 = (\lambda, v, \beta)$. A posterior distribution can be written as

$$\begin{aligned} \pi(H_1, H_2 | y) &\propto \frac{\beta^{\alpha n-1}}{(\Gamma(\alpha))^n} \prod_{i=1}^{k+1} y_t^{(\alpha-1)n_i} \frac{1}{h_i^{\alpha n_i}} \\ &\quad \exp - \beta \frac{s_i}{h_i} C_k^{k_{max}} \lambda^k (-\lambda)^{k_{max}-k} \frac{1}{C_{n-2}^{2k+1}} \\ &\quad \prod_{i=1}^{k+1} (n_i - 1) \\ &\quad \frac{v^{u(k+1)-1}}{(\Gamma(u))^{(k+1)}} \left(\prod_{i=1}^{k+1} h_i \right)^{-u-1} \\ &\quad \exp - v \sum_{i=1}^{k+1} \frac{1}{h_i} \end{aligned} \quad (9)$$

3.4. Reversible Jump MCMC

Parameter estimation (H_1, H_2) is carried out using the Gibbs algorithm which consists of two stages, namely: distribution simulation $\pi(H_2 | H_1, y)$ and distribution simulation $\pi(H_1 | H_2, y)$. The distribution simulation $\pi(H_2 | H_1, y)$ can be done using the following distributions,

i.e.:

$$\begin{aligned} \beta &\sim G\left(\alpha n, \sum_{i=1}^{k+1} \frac{s_i}{h_i}\right), \lambda \\ &\sim B(k+1, k_{max} - k + 1), \text{ and } v \\ &\sim G\left(u(k+1), \sum_{i=1}^{k+1} \frac{1}{h_i}\right). \end{aligned}$$

The distribution simulation $\pi(H_1|H_2, y)$ is done by using the reversible jump MCMC algorithm. This algorithm uses 3 transformations, namely: changes in a location of the constant model, birth of the constant model, and death of the constant model.

3.4.1. Change in the Location

The change in the location of the constant model is as follows. Take a location randomly between τ_1, \dots, τ_k . If τ_j is selected, the location τ_j is deleted and replaced by location τ_j^* . Take u randomly according to $U(\tau_{j-1}, \tau_{j+1})$. so that $\tau_j^* = u$. Suppose that $x = (\tau_1, \dots, \tau_j, \dots, \tau_k, h_1, \dots, h_{k+1})$ and $x^* = (\tau_1, \dots, \tau_j^*, \dots, \tau_k, h_1, \dots, h_{k+1})$. The point x^* will replace x with probability

$$\rho(x, x^*) = \min\left\{1, \frac{f(y|x^*) \pi(x^*|k) q(x^*, x)}{f(y|x) \pi(x|k) q(x, x^*)}\right\} \quad (10)$$

If $\tau_j^* = \tau_j$ then a ratio of the likelihood function is

$$\frac{f(y|x^*)}{f(y|x)} = 1 \quad (11)$$

If $\tau_j^* < \tau_j$ then a ratio of the likelihood function is

$$\frac{f(y|x^*)}{f(y|x)} = \left(\frac{h_{j-1}}{h_j}\right)^{(\tau_j - \tau_j^*)} \exp - \beta \sum_{t=\tau_j^*+1}^{\tau_j} y_t \left(\frac{1}{h_j} - \frac{1}{h_{j-1}}\right) \quad (12)$$

If $\tau_j^* > \tau_j$ then a ratio of the likelihood function is

$$\frac{f(y|x^*)}{f(y|x)} = \left(\frac{h_j}{h_{j-1}}\right)^{(\tau_j^* - \tau_j)} \exp - \beta \sum_{t=\tau_j+1}^{\tau_j^*} y_t \left(\frac{1}{h_{j-1}} - \frac{1}{h_j}\right) \quad (13)$$

The ratio of the posterior distribution is

$$\frac{\pi(x^*|k)}{\pi(x|k)} = \frac{n_{j-1}^* n_j}{n_{j-1} n_j} \quad (14)$$

The ratio of the instrumental distribution is

$$\frac{q(x^*, x)}{q(x, x^*)} = 1 \quad (15)$$

3.4.2. Birth of the Constant Model

The birth of the constant model is as follows. Take location τ^* randomly between $2, \dots, n-1$. If $\tau^* \in (\tau_j, \tau_{j+1})$ then the height of h_j is deleted and replaced by the height h_j^* and h_{j+1}^* such that

$$(\tau^* - \tau_j) \log(h_j^*) + (\tau_{j+1} - \tau^*) \log(h_{j+1}^*) = (\tau_{j+1} - \tau_j) \log(h_j) \quad (16)$$

Suppose that $x = (\tau_1, \dots, \tau_j, \dots, \tau_k, h_1, \dots, h_j, \dots, h_{k+1})$ and $x^* = (\tau_1, \dots, \tau_j, \tau_j^*, \tau_{j+1}, \dots, \tau_k, h_1, \dots, h_{j-1}, h_j^*, h_{j+1}^*, h_{j+1}, \dots, h_{k+1})$. The point x^* will replace x with probability

$$\rho(x, x^*) = \min\left\{1, \frac{f(y|x^*) \pi(x^*|k) q(x^*, x)}{f(y|x) \pi(x|k) q(x, x^*)}\right\} \quad (17)$$

In this case, the ratio of the likelihood function is

$$\frac{f(y|x^*)}{f(y|x)} = \exp - \beta \left(\frac{s_j^*}{h_j^*} + \frac{s_{j+1}^*}{h_{j+1}^*} - \frac{s_j}{h_j}\right) \quad (18)$$

The ratio of the posterior distribution is

$$\begin{aligned} \frac{\pi(x^*|k)}{\pi(x|k)} &= \frac{k_{max} - k}{k + 1} \frac{\lambda}{1 - \lambda} \frac{(2k + 3)(2k + 2)}{(n - 2k - 2)(n - 2k - 3)} \\ &\frac{n_j^* n_{j+1}^*}{n_j} \frac{v^u}{\Gamma(u)} \left(\frac{h_j^* h_{j+1}^*}{h_j}\right)^{-u-1} \\ &\exp - v \left(\frac{1}{h_j^*} + \frac{1}{h_{j+1}^*} - \frac{1}{h_j}\right) \end{aligned} \quad (19)$$

The ratio of the instrumental distribution is

$$\frac{q(x^*, x)}{q(x, x^*)} = \frac{n}{k+1} \frac{(h_j^* + h_{j+1}^*)^2}{h_j} \quad (20)$$

3.4.3. Death of the Constant Model

The death of the constant model is as follows. Take a location randomly between τ_1, \dots, τ_k . If location τ_{j+1} is selected, then location τ_{j+1} is deleted. The height of h_j and h_{j+1} is also removed and replaced by the height h_j^* such that

$$\begin{aligned} (\tau_{j+1} - \tau_j) \log(h_j) + (\tau_{j+2} - \tau_{j+1}) \log(h_{j+1}) = \\ (\tau_{j+2} - \tau_j) \log(h_j^*) \end{aligned} \quad (21)$$

Suppose that

$x = (\tau_1, \dots, \tau_j, \tau_{j+1}, \tau_{j+2}, \dots, \tau_k, h_1, \dots, h_j, h_{j+1}, \dots, h_{k+1})$ and $x^* = (\tau_1, \dots, \tau_j, \tau_{j+2}, \dots, \tau_k, h_1, \dots, h_{j-1}, h_j^*, h_{j+2}, \dots, h_{k+1})$. The point x^* will replace x with probability

$$\rho(x, x^*) = \min\left\{1, \frac{f(y|x^*) \pi(x^*|k) q(x^*, x)}{f(y|x) \pi(x|k) q(x, x^*)}\right\} \quad (22)$$

In this case, the ratio of the likelihood function is

$$\frac{f(y|x^*)}{f(y|x)} = \exp - \beta \left(\frac{s_j}{h_j} - \frac{s_j}{h_j} - \frac{s_{j+1}}{h_{j+1}}\right) \quad (23)$$

The ratio of the posterior distribution is

$$\begin{aligned} \frac{\pi(x^*|k)}{\pi(x|k)} &= \frac{k_{max} - k}{k + 1} \frac{\lambda}{1 - \lambda} \frac{(2k + 3)(2k + 2)}{(n - 2k - 2)(n - 2k - 3)} \\ &\frac{n_j^* n_{j+1}^*}{n_j} \frac{v^u}{\Gamma(u)} \left(\frac{h_j^* h_{j+1}^*}{h_j}\right)^{-u-1} \\ &\exp - v \left(\frac{1}{h_j^*} + \frac{1}{h_{j+1}^*} - \frac{1}{h_j}\right) \end{aligned} \quad (24)$$

The ratio of the instrumental distribution is

$$\frac{q(x^*,x)}{q(x,x^*)} = \frac{n}{k+1} \frac{(h_j^*+h_{j+1}^*)^2}{h_j} \quad (25)$$

3.5. Simulation

Performance of the algorithm is tested using a simulation study. Synthetic data is made using the piecewise constant model in equation (1). A value of a piecewise constant model parameter is presented in Table 1 while noise is assumed to have a Gamma distribution with parameter values $\alpha = 5$ and $\beta = 5$. A value of the maximum k is 10.

Table 1. Value of model parameter

Value of k	Value of τ	Value of h
5	$\begin{pmatrix} 40 \\ 80 \\ 120 \\ 170 \\ 200 \end{pmatrix}$	$\begin{pmatrix} 1.5 \\ 1.1 \\ 1.6 \\ 0.8 \\ 0.4 \\ 0.7 \end{pmatrix}$

This synthetic data is presented in Figure 1.

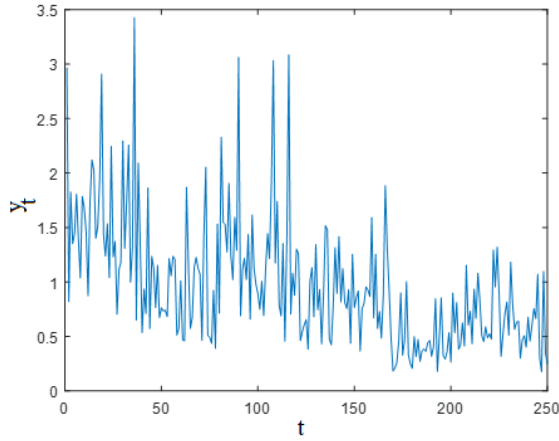


Figure 1. Synthetic data

This synthetic data is used as input to the reversible jump MCMC algorithm. The output is the number of constant models, the location of the model changes, and the height of the constant model.

The algorithm runs as many as 100,000 iterations with a burn-in period of 20,000. A histogram of the number of models is presented in Figure 2.

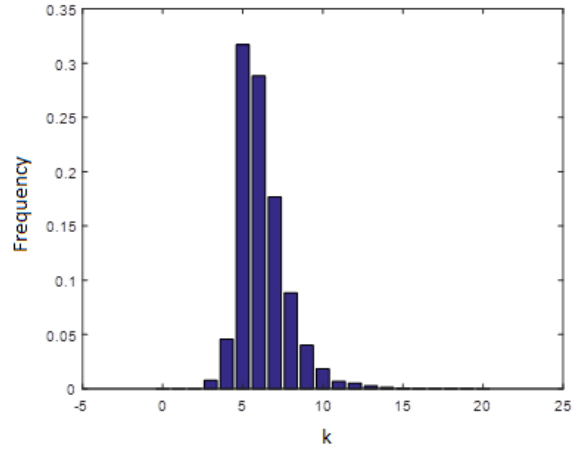


Figure 2. Histogram of the number of models

Figure 2 shows that the maximum value of k is reached at $k = 5$. So the maximum probability estimator for the number of models k is $\hat{k} = 5$. For $\hat{k} = 5$, the estimator for the location of the model change is presented in Table 2. The image of the synthetic data and estimator for a model change location is shown in Figure 3.

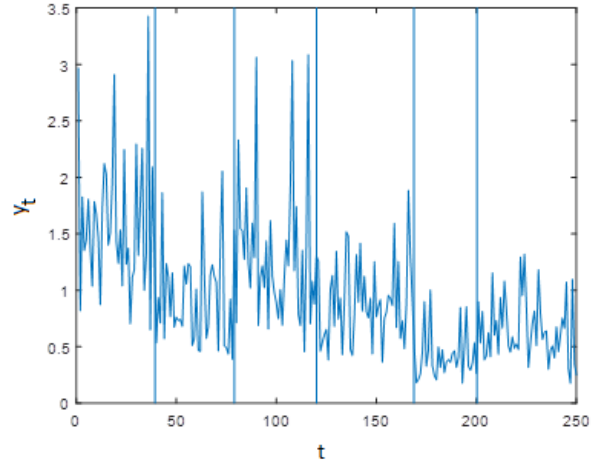


Figure 3. Segmentation of synthetic data

Figure 3 shows that this synthetic data has 6 different models. Finally, for $\hat{k} = 5$, the estimator for the height of the constant model is presented in Table 2.

Table 2. Value of parameter estimation

Estimation of k	Estimation of τ	Estimation of h
5	$\begin{pmatrix} 39.40 \\ 79.02 \\ 120.16 \\ 168.76 \\ 200.33 \end{pmatrix}$	$\begin{pmatrix} 1.61 \\ 0.93 \\ 1.32 \\ 0.89 \\ 0.42 \\ 0.65 \end{pmatrix}$

In Table 1 that is compared with Table 2, the estimated value of the parameter approaches the parameter value. The distance between τ and $\hat{\tau}$ is $|\tau - \hat{\tau}| = 1.73$ while the distance between h and \hat{h} is $|h - \hat{h}| = 0.36$. Synthetic data and an estimator for height are presented in Figure 4.

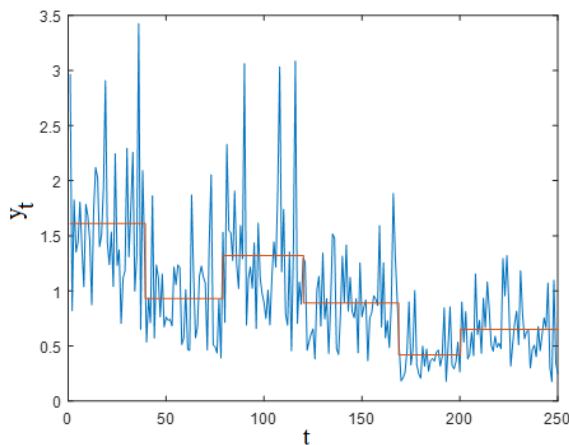


Figure 4. Synthetic data and reconstructed data

Figure 4 shows that the reversible jump MCMC algorithm can be used to estimate the number of constant models, estimate the location of a model change, and the height of a constant model. The simulation studies show that the reversible jump MCMC algorithm can determine the number of models and parameters of a piecewise constant model well.

4. Conclusions

This paper develops a piecewise constant model and its parameter estimation procedure. The piecewise constant model parameter includes the number of constant models, the location of changes in the constant model, the constant model height, and noise variance. The Bayes estimator cannot be formulated explicitly because the number of constant models is a parameter. The reversible jump MCMC method is proposed to estimate these parameters. According to the simulation study, the reversible jump MCMC estimated the piece-wise constant model parameter well.

The reversible jump MCMC algorithm has several advantages. This algorithm can be used to estimate the piecewise constant model parameter that has Gamma multiplicative noise simultaneously. This algorithm can also be used to determine the hyper-parameters that appear

in the prior distribution.

Acknowledgements

The author would like to thank the University of Ahmad Dahlan for providing the grant.

REFERENCES

- [1] X. Pang, S. Zhang, J. Gu, L. Li, B. Liu, and H. Wang, Improved L0 Gradient Minimization Alt L1 Fidelity for Image Smoothing, PLOS ONE, 1-10, 2015.
- [2] D. Zivkovic, M. Steinrucken, Y.S. Song, and W. Stephan, Transition Densities and Sample Frequency Spectra of Diffusion Processes with Selection and Variable Population Size, Genetics, 601-617, 2015.
- [3] J.A. Kamm, J.P. Spence, J. Chan, and Y.S. Song, Two-Locus Likelihoods Under Variable Population Size and Fine-Scale Recombination Rate Estimation, Genetics, 1381-1399, 2016.
- [4] S. Nandy, C.Y. Lim, and T. Maiti, Additive model building for spatial regression, J.R.Statist. Soc.B, 779-800, 2017.
- [5] Y. Hu, S. Feng, and L. Xue, Automatic Variable Selection for Partially Linear Functional Additive Model and Its Application to the Tecator Data Set, Mathematical Problems in Engineering, 1-9, 2018.
- [6] R. Richardson, H.D. Tolley, W.E. Evenson, B.M. and Lunt, Accounting for measurement error in log regression models with applications to accelerated testing, PLOS ONE, 1-13, 2018.
- [7] T. A. Marques, Predicting and Correcting Bias Caused by Measurement Error in Line Transect Sampling Using Multiplicative Error Models, Biometrics, 757-763, 2004.
- [8] Z. Wang and S.S. Abeysekera, Asymptotic Bounds for Frequency Estimation in the Presence of Multiplicative Noise, Journal on Advances in Signal Processing, 7, 1-9, 2007.
- [9] V.K. Andriasyan. Adsorption of Ligands on DNA for Arbitrary Filling in the Presence of Multiplicative Noise, Journal of Contemporary Physics, 48, 243-246, 2013.
- [10] Suparman and M. Doisy, Bayesian Segmentation in Signal with Multiplicative Noise Using Reversible Jump MCMC, Telkomnika, 673-680, 2018.
- [11] E. Punskeya, C. Andrieu, A. Doucet, and W.J. Fitzgerald, Bayesian Curve Fitting Using MCMC With Applications to Signal Segmentation, IEEE Transactions on Signal Processing, 50, 747-758, 2002.
- [12] Z. Shi and H. Aoyama, Estimation of the Exponential Autoregressive time series model by using the genetic algorithm, Journal of Sound and Vibration, 309-321, 1997.
- [13] N. Sad, On Exponential Autoregressive Time Series Models, J. Math, 97-101, 1999.

- [14] L. Larbi and H. Fellag, Robust Bayesian Analysis of an Autoregressive Model with Exponential, Innovations, *Afr. Stat.*, 955-964, 2016.
- [15] G.A. Triantafyllidis, D. Tzovaras, and M.G. Strintzis, A Bayesian Approach for Segmentation in Stereo Image Sequences, *Journal on Applied Signal Processing*, 10, 1116-1126, 2002.
- [16] P.J. Green, Reversible Jump MCMC Computation and Bayesian Model Determination. *Biometrika*, 4, 711-732, 1995.