International Journal of Engineering &Technology, 7 (4) (2018) 6162-6166

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.23127
Research paper

Change-time detection in moving average model based on reversible jump MCMC algorithm

Suparman*

Department of Mathematics Education, University of Ahmad Dahlan, Indonesia
**Corresponding author E-mail: suparman@pmat.uad.ac.id*

Abstract

A piecewise moving average is a model that is more flexible than the moving average model. If the piecewise moving average is used to model data, the model parameters are unknown. The model parameters include the number of models, model coefficients, and white noise variance. This paper discusses the parameter estimation of the piecewise moving average model. Parameter estimation is done using the hierarchical Bayesian approach. Bayes estimators are calculated using the reversible jump Markov chain Monte Carlo algorithm. The performance of the algorithm is tested using synthesis data. The results showed that the reversible jump Markov chain Monte Carlo algorithm can estimate the parameters of the piecewise moving average model well.

Keywords: *Bayes Method; Monte Carlo Method; Piecewise Linear; Parameter Estimation; Signal Detection.*

1. Introduction

A moving average (MA) is a mathematical model that is widely used in many fields. The MA model is used to estimate the natural location of the bee colony [1]. The MA model is used as a noise in the non-linear-finite impulse response (IIR) [2]. If the MA model is fitted to the data, then the parameter of the MA model must be estimated. Many authors have examined the estimation of the MA model parameter, for example Yule-Walker estimator is proposed to estimate the parameter of the MA model [3]. But in many applications, the data have a model that changes at one (several) unknown time. In other words, the model changes from one time interval to another time interval. So the combination of several mathematical models is needed to model this data.

A piecewise function is a mathematical model developed to model the data that has a different pattern from one time interval to another time interval. The piecewise model is widely used in many fields. The piecewise model is used to model longitudinal data [4]. The piecewise model is used to model the virus-immune system [5,6]. The piecewise model is used for the diversity of species in the river [7]. The piecewise model is used to calculate the energy expenditure (EE) in real time from the heart rate [8]. A piecewise linear model to model the degradation of railway geometry with the known breakpoints [9]. The piecewise regression model is used on semiconductors [10]. The piecewise regression model is used for clustering and segmentation [11]. The piecewise regression model is used to detect the divergence of trajectory in groups of children with different developmental phases [12].

In this paper, a piecewise MA model is proposed. Consider a data $y = (y_1, \dots, y_n)$ where n is the number of observations. This data is modeled as an MA model if for $t = 1, \dots, n$ it verifies the following equation:

$$y_t = z_t + \sum_{j=1}^{q_{i,k}} \phi_{i,j,k}^{(q_{i,k})} z_{t-j}, \text{ for } \tau_{i,k} \le t < \tau_{i+1,k}, \qquad (1)$$

Where the quantity of k $(k = 0, \dots, k_{max})$ represents the number of MA models. For $i = 0, \dots, k$, the quantity $\tau_{i,k}$ is the i^{th} change-time of the MA model with the usual convention $\tau_{0,k} = 0$ and $\tau_{k+1,k} = n$. Similarly for $i = 0, \dots, k$, the quantity $q_{i,k}$ and the quantity $\phi_{i,k}^{(q_{i,k})} = \left(\phi_{i,k,1}^{(q_{i,k})}, \dots, \phi_{i,k,q_{i,k}}^{(q_{i,k})}\right)'$ are the MA model order and the parameter of the MA model associated to this time interval. For $i = 0, \dots, k$, the quantity z_t is the Gaussian noise with mean 0 and variance $\sigma_{i,k}^2$ associated with the MA model in this time interval.

As shown in [13], for each the i^{th} time interval $(i = 0, \dots, k)$ the MA model of the order $q_{i,k}$ is invertible if $\phi_{i,k}^{(q_{i,k})}$ belongs to

$$I_{q_{i,k}} = \left\{\phi_{i,k}^{(q_{i,k})} \in R^{q_{i,k}} \middle| 1 + \phi_{i,k,1}^{(q_{i,k})} x + \dots + \phi_{i,k,q_{i,k}}^{(q_{i,k})} x^{q_{i,k}} \ne 0, x \in C, |x| \le 1\right\}. \qquad (2)$$

The invertible region $I_{q_{i,k}}$ is difficult to be find if $q_{i,k} > 2$. To solve this problem, a reparameterization of the MA $(q_{i,k})$ model is adopted [13]. The notation MA $(q_{i,k})$ refers to the MA model of order $q_{i,k}$. For a MA$(q_{i,k})$ model, there is a one-to-one transformation $G: \rho_{i,k}^{(q_{i,k})} \in (-1,1)^{q_{i,k}} \to \phi_{i,k}^{(q_{i,k})} \in I_{q_{i,k}}$ where $\rho_{i,k}^{(q_{i,k})} = \left(\rho_{i,k,1}^{(q_{i,k})}, \dots, \rho_{i,k,q_{i,k}}^{(q_{i,k})}\right)'$ is the vector of the first $q_{i,k}$ inverted partial autocorrelations of the MA$(q_{i,k})$ model (see [13] for a definition of the vector $\rho_{i,k}^{(q_{i,k})}$). This reparametrization changes the invertible region from the $\phi_{i,k}^{(q_{i,k})} \in I_{q_{i,k}}$ becomes the $\rho_{i,k}^{(q_{i,k})} < 1$ $(i = 1, \dots, q_{i,k})$.

Let $\theta = (k, \tau^{(k)}, q^{(k)}, \phi^{(k)}, \sigma^{(k)})'$ be a parameter vector. Here $\tau^{(k)} = (\tau_1, \dots, \tau_k)'$, $q^{(k)} = (q_{0,k}, \dots, q_{k,k})'$, $\phi^{(k)} = \left(\phi_{0,k}^{(q_{0,k})}, \dots, \phi_{k,k}^{(q_{k,k})}\right)'$ and $\sigma^{(k)} = (\sigma_{0,k}, \dots, \sigma_{k,k})'$. If the piecewise MA model is used to model the data, the parameter θ must be estimated. Because the number of MA models is also a parameter, the previous methods cannot be used to finding the parameter estimation. This paper proposes a reversible jump MCMC (Markov Chain Monte Carlo) algorithm in [14] to estimate the parameter θ. The MCMC reversible jump algorithm has used for the selection of mathematical models in which the number of mathematical models is assumed to be unknown, for example in [15,16]. The reversible jump MCMC algorithm is used to estimate ARMA (Autoregressive Moving Average) model parameters of unknown model order [15]. The MCMC reversible jump algorithm is implemented to detect piecewise regression models where the number of change-points is unknown [16]. This paper consists of four sections. The first section describes the introduction. The second section presents the research method. The third section explains the results and discussion. The fourth section provides a conclusion.

2. Method

The parameter estimation is done in a Bayesian framework. First, the likelihood function is determined. Second, the prior distribution for the parameter of the piecewise MA model is selected. Third, the posterior distribution is determined by using Bayes theorem. Fourth, since the Bayes estimator for parameters cannot be calculated analytically, the Bayes estimator for the parameter is calculated by using the reversible jump MCMC algorithm. The reversible jump MCMC algorithm is implemented in three stages: (a) The birth of the change-time, (b) the death of the change-time, and (c) the change in the change-time location. Furthermore, for the number of change-time and change-time location known there are three sub-stages, namely: the birth order MA model, the death order MA model, and the change in the value of the MA model coefficient.

3. Results and analysis

The following section describes the results related to likelihood function, Bayesian approach, reversible jump MCMC approach, and simulation.

3.1. Likelihood function

The entire document should be in Times New Roman. The font sizes to be used are specified in Table 1. In lieu of the exact likelihood, an approximation to the likelihood is developed. Let q_{max} be the maximum number of MA model order. Because the noise has a normal distribution, the approximate likelihood takes a following form:

$$f(s|\theta) = \prod_{i=0}^{k}\left(2\pi\sigma_{i,k}^2\right)^{-\frac{1}{2}(\tau_{i+1,k}-\tau_{i,k})}\exp -\frac{1}{2\sigma_{i,k}^2}\sum_{t=\tau_{i,k}+1}^{\tau_{i+1,k}}\left(y_t - \sum_{j=1}^{q_{i,k}} G\left(\rho_{i,k,j}^{(q_{i,k})}\right)\hat{z}_{t-j}\right)^2 \quad (3)$$

Where

$$s = \left(y_{q_{max+1}}, \dots, y_n\right)'$$

and $\hat{z}_1 = \cdots = \hat{z}_{q_{max}} = 0$. The t^{th} residual $(t = q_{max} + 1, \dots, n)$ is calculated by

$$\hat{z}_t = y_t - \sum_{j=1}^{q_{i,k}} G\left(\rho_{i,k,j}^{(q_{i,k})}\right)\hat{z}_{t-j}, \quad \tau_{i,k} \le t < \tau_{i+1,k}, i = 0, \dots, k. \quad (4)$$

3.2. Bayesian approach

A Bayesian approach is adopted in this work. It implies the choice of prior distribution for the parameter. Denote k_{max} as the maximum number of change-time. The k of the change-time location is drawn following a binomial distribution $B(k_{max}, \lambda)$ with parameter $\lambda(0 \le \lambda \le 1)$ and k_{max}.

For the k fixed, the change-time locations are distributed as the even numbered position of the order statistics of 2k+1 points uniformly drawn without repetition in {2...,n}. This choice avoids too small interval. The MA model order $q_{i,k}$ $(i = 0, \dots, k)$ is assumed to follow the binomial distribution $B(q_{max}, \mu)$ with parameter $\mu(0 \le \mu \le 1)$ and q_{max}. For $q_{i,k}$ fixed, the quantity of $\rho_{i,k,j}^{(q_{i,k})}$ is assumed to follow the uniform distribution, i.e. $\rho_{i,k,j}^{(q_{i,k})} \sim U(-1,1)$. For $q_{i,k}$ fixed, the quantity $\sigma_{i,k}^2$ is independent and distributed according to $IG(\frac{\alpha}{2}, \frac{\beta}{2})$ where $\alpha > 0$ and $\beta > 0$. In order to have the robust prior, we consider the hyper parameter vector $(\lambda, \mu, \alpha, \beta)$ to be a random variable. The hyper parameters λ and μ are drawn following the uniform distribution on (0,1), i.e. $\lambda \sim U(0,1)$ and $\mu \sim U(0,1)$. Set $\alpha = 2$. A Jeffreys prior distribution is selected for β. So, the prior distribution of the parameter for θ is given by:

$$\pi(\theta, \xi) = \pi(k|\lambda)\mu\left(\tau^{(k)}|k\right)\pi\left(q^{(k)}|k\right)\pi\left(\rho^{(k)}|k, q^{(k)}\right)\pi\left(\sigma^{(2k)}|\alpha, \beta, k\right)\pi(\lambda)\mu(\beta) \quad (5)$$

Where $\xi = (\lambda, \mu, \beta)$. By the Bayes formula, the posterior distribution can be written by $\pi(\theta, \xi|s) \propto f(s|\theta)\,\pi(\theta, \xi)$. That is the product of the likelihood function in Eq. (3) and the prior distribution in Eq. (5). The Bayesian inference for parameter (θ, ξ) is based on the posterior distribution. This posterior distribution is calculated by Eq. (4). In this case, it is not analytically possible to obtain this quantity. Then the reversible jump MCMC algorithm is applied.

3.3. Reversible jump MCMC

The key idea of the reversible jump MCMC is to build an ergodic Markov chain (θ_1, ξ_1), …, (θ_M, ξ_M) whose equilibrium distribution is the posterior distribution $\pi(\theta, \xi|s)$. This sample generated by the reversible jump Markov chain can be used to estimate all posterior features of interest.

If the value of k and the value of the MA model order $q^{(k)}$ are known, then the Metropolis-Hasting algorithm can be used to simulate a Markov chain according to this posterior distribution. In this case, because both the value of k and the value of $q^{(k)}$ are unknown, the Markov chain must jump from a model $(k, q^{(k)})$ with parameter $\left(\tau^{(k)}, \rho^{(k)}, \sigma^{(2k)}\right)$ to a another model $(k', q^{(k')})$ with parameter $\left(\tau^{(k')}, \rho^{(k')}, \sigma^{(2k')}\right)$. A solution to this problem of the model selection is proposed in [14]. This model selection is done in two stages: First stage, the reversible jump MCMC algorithm is used to define a transformation so that the Markov chain can jump between the models of different dimension in term of the number of change-time k. In this work, there are 3 transformations, namely: the birth of a change-time, the death of a change-time and the change of the change-time position. The second stage, for the k fixed the reversible jump MCMC algorithm is used to define the jump between models of a different dimension, but in term of the order $q_{i,k}$ $(i = 0, \dots, k)$. For this, there are three transformations, namely: the birth of the MA model coefficient, the death of the MA model coefficient, and the change of the MA model coefficient.

To find the parameter estimation, the reversible jump MCMC is used to simulate a Markov chain distributed according to the posterior distribution $\pi(\theta, \xi|s)$. The samples k_j $(j = 1, \dots, M)$ from the joint posterior distribution $\pi(\theta, \xi|s)$ are collected after a burn-in period. This method provides the marginal distribution of the number of change-time k. The marginal maximum posterior estimator of k can be determined as follows:

$$\hat{k} = argmax\ \hat{p}\left(k_j|s\right), k_j \in \{0, \dots, k_{max}\} \quad (6)$$

If the parameter k has been estimated, then the change-time location and the order can be estimated as follows:

$$\tau^{\hat{k}} = argmax\ \hat{p}\left(\tau_j^{(k)} \middle| k = \hat{k}, s\right), \tau_j^{(k)} \in \{2, \dots, n-1\}$$
(7)

and

$$\hat{q}_{i,k} = argmax\ \hat{p}\left(q_{i,k_j} \middle| k = \hat{k}, s\right), q_{i,k_j} \in \{0, \dots, q_{max}\}$$
(8)

Finally, for $k = \hat{k}$ and $q_{i,k} = \hat{q}_{i,k}$ $(i = 1, \dots, \hat{k})$, the MA model coefficient and the noise variance associated are estimated by the same way (using marginal maximum posterior).

3.4. Simulation

The simulation results are presented for a synthetic signal. The 250 samples of the synthetic signal are generated according to the piecewise MA model in Eq. (1). The value of the number of MA model $k = 1$. The change-time location $\tau^{(1)} = 125$. The MA model order, the MA coefficient and the noise variance for each time interval are summarized in Table 1.

Table 1: Parameter of the Synthetic Signal

i^{th} segment	$\sigma_{i,1}$	$q_{i,1}$	$\phi_{i,1}^{(q_{i,1})}$
0	0.5	1	0.7826
1	1.5	3	(0.52, -0.08, -0.96)

Based on this synthetic signal, the reversible jump MCMC is implemented to find the parameter estimation of the synthetic signal. The MCMC simulation is run for 60.000 iterations, after a burn-in period of 10.000 iterations ($k_{max} = 10, q_{max}$=15). The following is the output of the simulation. The histogram of the marginal a posteriori distribution k is plotted in the Fig. 1.

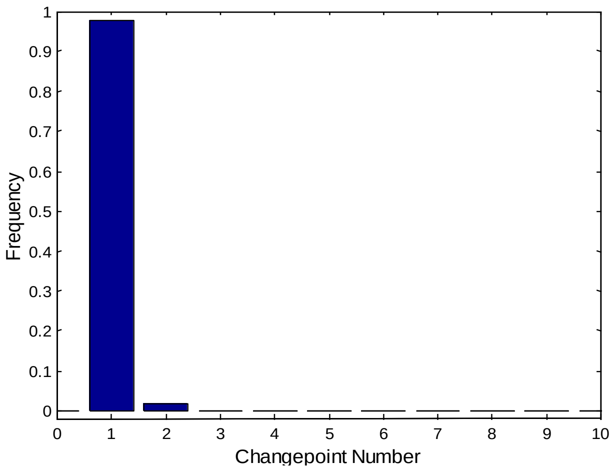

Fig. 1: Histogram of Marginal Posterior Distribution For k.

The estimation of k is the value in $\{0, 1, \dots, 10\}$ at which the frequency is a maximum. Fig. 1 shows that the estimation of the number of the MA models k is $\hat{k} = 1$.
For $\hat{k} = 1$ fixed, the histogram of the conditional marginal posterior distribution of $\tau^{(1)}$ is given in Fig. 2.

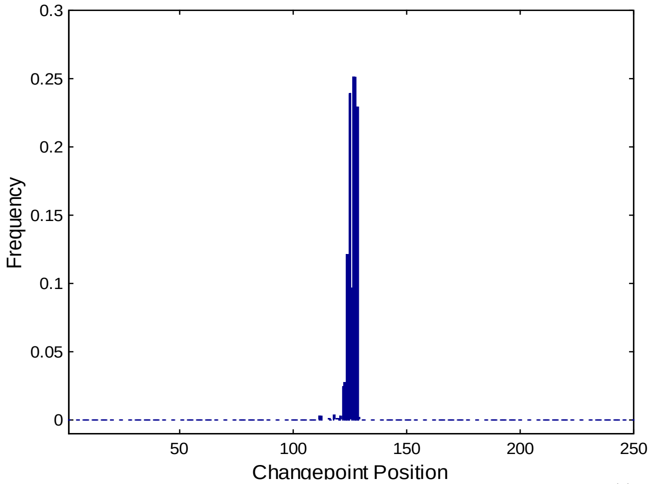

Fig. 2: Histogram of Conditional Marginal Posterior Distribution for $\tau^{(1)}$.

The estimation of the change-time $\tau^{(1)}$ in $\{2, 3, \dots, n-1\}$ is at which the frequency is a maximum. Fig. 2 shows that the estimation of the change-time $\tau^{(1)}$ is $\hat{\tau}^{(1)} = 127$. The estimation of the change-time and the synthetic signal are plotted in Fig. 3.

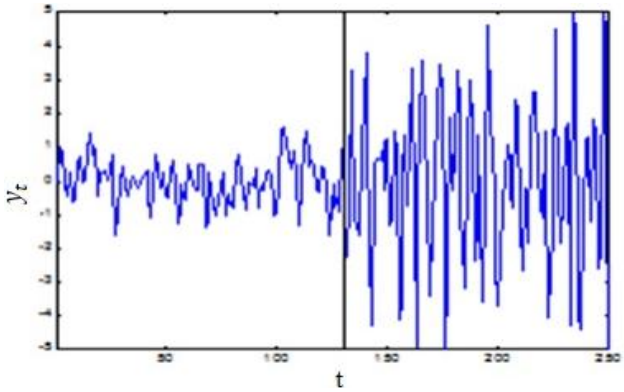

Fig. 3: Estimation of Change-Time and Synthetic Signal.

From Fig. 3, it can see that this synthetic signal is composed of two different MA models.
For $\hat{k} = 1$ fixed, the histogram of the conditional marginal posterior distribution for the MA order $q_{0,1}$ and the order $q_{1,1}$ are given in the Fig. 4 and the Fig. 5 respectively.

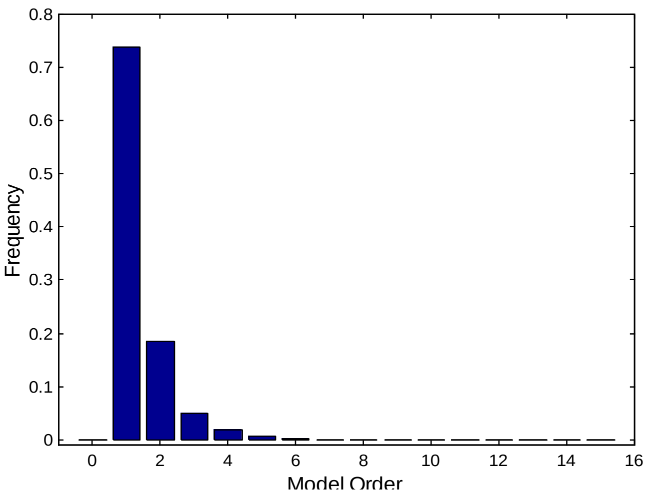

Fig. 4: Histogram of Conditional Marginal Posterior Distribution for $Q_{0,1}$.

The estimation of the MA order $q_{0,1}$ in $\{0, 1, \dots, 15\}$ is at which the frequency is a maximum. Fig. 4 shows that the estimation of MA order $q_{0,1}$ is $\hat{q}_{0,1} = 1$.

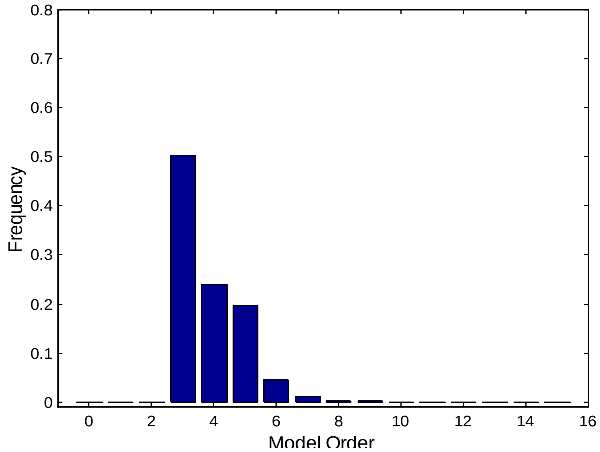

Fig. 5: Histogram of Conditional Marginal Posterior Distribution for $Q_{1,1}$.

The estimation of the MA order $q_{1,1}$ in $\{0, 1, \ldots, 15\}$ is at which the frequency is a maximum. Fig. 5 shows that the estimation of MA order $q_{0,1}$ is $\hat{q}_{1,1} = 3$.

For $\hat{k} = 1$ and $\hat{q}_{0,1} = 1$fixed, the Curve a (Fig. 6) shows the conditional marginal posterior distribution of $\emptyset_{0,1,1}^{(q_{0,1})}$ by using the Gaussian kernel with the standard deviation 0.2. Similarly for $\hat{k} = 1$ and $\hat{q}_{1,1} = 3$, in the same Fig. 6 (Curve b, Curve c, and Curve d) shows the conditional marginal posterior distribution of $\emptyset_{1,1,1}^{(q_{1,1})}$, $\emptyset_{1,1,2}^{(q_{1,1})}$ and $\emptyset_{1,1,3}^{(q_{1,1})}$ by using the same Gaussian kernel with the standard deviation 0.2.

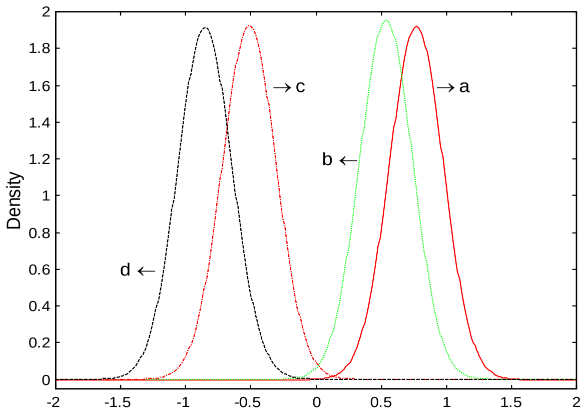

Fig. 6: Histogram of Conditional Marginal Posterior Distribution for the Coefficient of the MA Models.

Finally, for $\hat{k} = 1$, the histogram of the marginal posterior distribution of $\sigma_{0,1}$ and $\sigma_{1,1}$ are given in Fig. 7 and Fig. 8.

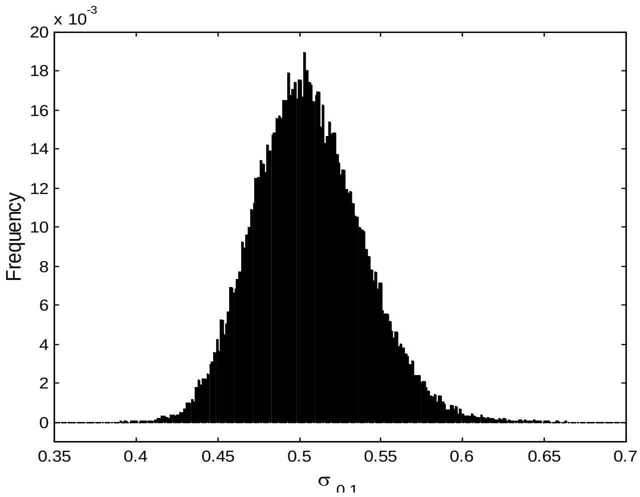

Fig. 7: Histogram of Conditional Marginal Posterior Distribution for $\Sigma_{0,1}$.

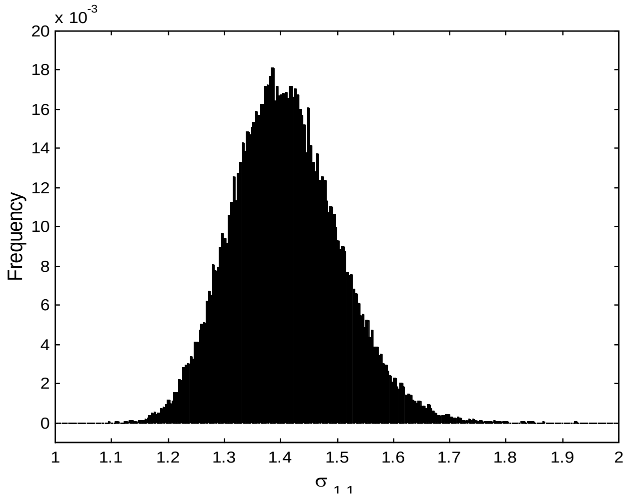

Fig. 8: Histogram of Conditional Marginal Posterior Distribution for $\Sigma_{1,1}$.

4. Conclusion

This paper extends the moving average model into a piecewise moving average model. Bayesian approach is used to estimate model parameters. Because the Bayes estimator cannot be calculated analytically, the MCMC reversible jump algorithm is proposed to calculate the Bayes estimator. This algorithm can estimate the model parameters well. The model parameters include the number of models, model coefficients, and white noise variance can be calculated simultaneously. The advantage of this algorithm is that the resulting piecewise moving average model is invertible.

Acknowledgement

The author would like to thank the University of Ahmad Dahlan in Indonesia for providing a research grant. Not to forget, the author also wants to thank the reviewers who have given suggestions so that this paper becomes more qualified.

References

[1] Nakamura, H.; and Toquenaga, Y. (2002). Estimating Colony Locations of Bumble Bees with Moving Average Model. Ecological Research, 17(1), 39-48. https://doi.org/10.1046/j.1440-1703.2002.00461.x.

[2] Wang, Y.; and Ding, F. (2017). Iterative Estimation for a Non-Linear IIR Filter with Moving Average Noise by Means of the Data Filtering Technique. IMA Journal of Mathematical Control & Information, 34(3), 745-764. https://doi.org/10.1093/imamci/dnv067.

[3] Dimitriou-Fakalou, C. (2011). Yule-Walker Estimation for the Moving Average Model. International Journal of Stochastic Analysis, 1-20. https://doi.org/10.1155/2011/151823.

[4] Diallo, T.M.O.; and Morin, A.J.S. (2015). Power of Latent Growth Curve Models to Detect Piecewise Linear Trajectories. Structural Equation Modeling, 22(3), 449-460. https://doi.org/10.1080/10705511.2014.935678.

[5] Tang, B.; Xiao, Y.; Sivalganathan, S.; and Wu, J. (2017). A Piecewise Model of Virus-Immune System with Effector Cell-Guided Therapy. Applied Mathematical Modelling, 47, 227-248. https://doi.org/10.1016/j.apm.2017.03.023.

[6] Tang, B.; Xiao, Y.; and Wu, J. (2016). A Piecewise model of Virus-Immune System with Two Thresholds. Mathematical Biosciences, 278, 63-76. https://doi.org/10.1016/j.mbs.2016.06.003..

[7] Xu, Z-H.; Yin, X-A.; Zhang, C.; and Yang, Z.F. (2016). Piecewise Model for Species-discharge Relationships in Rivers. Ecological Engineering, 96, 208-213. https://doi.org/10.1016/j.ecoleng.2015.12.024.

[8] Hector, M.R-U.; Garnotel, M.; Doron, M.; Jallon, P.; Charpen-Tier, G.; Franc, S.; Huneker, E.; Simon, C.; and Bonnet, S. (2017). An Original Piecewise Model for Computing Energy Expenditure from Accelerometer and Heart Rate Signals. Physiological Measurement, 38(8), 1-1. https://doi.org/10.1088/1361-6579/aa7cdf.

[9] Soleimanmeigouni, I.; Xiao, X.; Ahmadi, A.; Xie, M.; Nissen, A.; and Kumar, U. (2018). Modeling the Evaluation of Ballasted Railway Track Geometry by Two-Level Piecewise Model. Structure & Infrastructure Engineering: Maintenance, Management, Life-Cycle Design & Performance, 14(1), 33-45. https://doi.org/10.1080/15732479.2017.1326946.

[10] Hong, A.; and Chen, A. (2012). Piecewise Regression Model Construction with Sample Efficient Regression Treen (SERT) and Applications to Semiconductor Yield Analysis. Journal of Process Control, 22(7), 1307-1317. https://doi.org/10.1016/j.jprocont.2012.05.017.

[11] Chamroukhi, F. (2016). Piecewise Regression Mixture for Simultaneous Functional Data Clustering and Optimal Segmentation. Journal of Classification, 33(3), 374-411. https://doi.org/10.1007/s00357-016-9212-8..

[12] Buscot, M-J.; Wotherpoon, S.S.; Magnussen, C.G.; Juonala, M.; Sabin, M.A.; Burger, D.P.; Lehtimaki, T.; Viikari, J.S.A.; Hutri-Kahonen, N.; Raitakari, O.T.; and Thomson, R.J. (2017). Bayesian Hierarchical Piecewise Regression Models: A Tool to detect Trajectory Divergence between Groups in Long-Term Observational Studies. BMC Medical Research Methodology, 17, 1-15. https://doi.org/10.1186/s12874-017-0358-9.

[13] Bhansali, R.J. (1983). The inverse partial correlation function of a time series and its applications. J.Multivar. Anal., 13, 310-327. https://doi.org/10.1016/0047-259X(83)90029-5.

[14] Green, P.J. (1995). Reversible Jump MCMC Computation and Bayesian Model Determination. Biometrika, 82(4), 711-732. https://doi.org/10.1093/biomet/82.4.711.

[15] Suparman, (2015). New Segmentation of Piecewise Linear Regression Models Using Reversible Jump MCMC Algorithm. Journal of Computer Technology and Application, 6(1), 14-18. https://doi.org/10.17265/1934-7332/2015.01.003.

[16] Suparman; and Doisy, M. (2014). Hierarchical Bayesian of ARMA Models Using Simulated Algorithm. Journal of Telkomnika, 12(1), 87-96. https://doi.org/10.12928/telkomnika.v12i1.12.