International Journal of Computational Intelligence Systems
Vol. **13(1)**, 2020, *pp.* 310–317
DOI: https://doi.org/10.2991/ijcis.d.200310.006; ISSN: 1875-6891; eISSN: 1875-6883
https://www.atlantis-press.com/journals/ijcis/

Regular paper

Hierarchical Bayesian Choice of Laplacian ARMA Models Based on Reversible Jump MCMC Computation

Suparman*

Department of Mathematics Education, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

ARTICLE INFO

Article History

Received 19 Dec 2019
Accepted 20 Feb 2020

Keywords

ARMA time series
Hierarchical Bayesian
Laplacian noise
Reversible jump MCMC

ABSTRACT

An autoregressive moving average (ARMA) is a time series model that is applied in everyday life for pattern recognition and forecasting. The ARMA model contains a noise which is assumed to have a specific distribution. The noise is often considered to have a Gaussian distribution. However in applications, the noise is sometimes found that does not have a Gaussian distribution. The first objective is to develop the ARMA model in which noise has a Laplacian distribution. The second objective is to estimate the parameters of the ARMA model. The ARMA model parameters include ARMA model orders, ARMA model coefficients, and noise variance. The parameter estimation of the ARMA model is carried out in the Bayesian framework. In the Bayesian framework, the ARMA model parameters are treated as a variable that has a prior distribution. The prior distribution for the ARMA model parameters is combined with the likelihood function for the data to get the posterior distribution for the parameter. The posterior distribution for parameters has a complex form so that the Bayes estimator cannot be determined analytically. The reversible jump Markov chain Monte Carlo (MCMC) algorithm was adopted to determine the Bayes estimator. The first result, the ARMA model can be developed by assuming Laplacian distribution noise. The second result, the performance of the algorithm was tested using simulation studies. The simulation shows that the reversible jump MCMC algorithm can estimate the parameters of the ARMA model correctly.

1. INTRODUCTION

An autoregressive moving average (ARMA) is a time series model that is applied to modeling and forecasting in various fields, for example, [1–3]. The ARMA model is used in the field of short-term load system for forecasting [1]. The ARMA model is used in the field of science for forecasting wind speed [2]. The ARMA model is used in the business field for modeling volatility and risk of shares financial markets [3].

This ARMA model contains a noise. This noise is assumed to have a specific distribution. Noise for ARMA models is often considered to have a Gaussian distribution, for example, [4–7]. The ARMA model is used for sequential and non-sequential acceptance sampling [4]. The ARMA model is used to investigate the non-residual residual surges [5]. The ARMA model is used to predict small-scale solar radiation [6]. The ARMA model is used to forecast passenger service charge [7]. In an ARMA model application, the noise sometimes shows that it does not have a Gaussian distribution. Several studies related to ARMA models with non-Gaussian noise can be found in [1–3,8,9]. Estimating ARMA parameters with non-Gaussian noise is investigated using high-order moments [8]. A cumulant-based order determination of ARMA models with Gaussian noise is studied in [9].

Email: suparman.pmat@gmail.com

A Laplacian is a noise investigated by several authors, for example, [10–12]. If x is a random variable with a Laplace distribution then x has a probability function:

$$f\left(x \mid \delta, \beta\right) = \frac{1}{2\beta} exp - \frac{|x-\delta|}{\beta}. \tag{1}$$

Here, δ is a location parameter and $\beta > 0$ is a scale parameter. The Laplacian noise is used to detect body position changes [10]. The feasibility pump algorithm is used to find the sparse representation under Laplacian noise [11]. Laplacian noise is used in human sensory processing [12]. However, the ARMA model that contains Laplacian noise has not been studied. The significant novelty of the proposed study is the use of a Laplacian noise in the ARMA model. The Laplacian noise provides a smaller error variance compared to Gaussian noise. This study has several objectives. The first objective is to develop the new ARMA model by assuming that noise has a Laplacian distribution. The second objective is to estimate the order of the ARMA model. The third objective is to estimate the ARMA model coefficients.

If the ARMA model is compared to the AR model and the MA model, the ARMA model is a more general model than the AR model and the MA model. If the ARMA model is compared to the ARIMA model, the only difference is the integrated part. Integrated refers to how many times it takes to differentiate a series to achieve stationary condition. The ARMA model is equivalent to the

ARIMA model of the same MA and AR orders with no differencing. An ARMA model was chosen instead of the other models such as AR model, MA model, or ARIMA model because the ARMA model can describe a more general class of processes than AR model and MA model. The ARMA model in this study has stationary and invertible properties that are not possessed by the ARIMA model.

This paper consists of several parts. The first part gives an introduction regarding the ARMA model and its application. The second part explains the method used to estimate the ARMA model. The third part presents the results of the research and discussion. The fourth section gives some conclusions and implications.

2. MATERIALS AND METHODS

This paper uses an ARMA model that has Laplacian noise. The parameters used in ARMA model are the order of the ARMA model, the coefficients of the ARMA model, and the variance of the noise. The parameter estimation of the ARMA model is carried out in the Bayesian framework. The first step determines the likelihood function for data. The second step determines the prior distribution for the ARMA model parameters. The reason about the consideration of prior distribution is to improve the quality of parameter estimation. The prior distribution can be determined from previous experiments. The Binomial distribution is chosen as the prior distribution for ARMA orders. The uniform distribution is selected as the prior distribution for the ARMA model coefficient. The inverse Gamma distribution is selected as the prior distribution for the noise parameter. The third step combines the likelihood function for data with the prior distribution to get the posterior distribution. The fourth step determines the Bayes estimator based on the posterior distribution using the reversible jump Markov chain Monte Carlo (MCMC) algorithm [13]. Time series modeling via reversible jump MCMC is a very well studied topic in the literature [14,15]. The power of reversible jump MCMC algorithm is in the fact that it can move between space of varying dimension and not that it is just a simple MCMC method. The fifth step tests the performance of the reversible jump MCMC algorithm by using simulation studies.

3. RESULTS AND DISCUSSION

This section discusses the likelihood function for data, Bayesian approach, reversible jump MCMC algorithm, and simulations.

3.1. Likelihood Function

Suppose that $x_1, \cdots, x_n$ are n data. This data is said to have an autoregressive model if for $t = 1, \cdots, n$ the data satisfies the following equation:

$$x_t = -\sum_{i=1}^{p} \phi_i x_{t-i} + \sum_{j=1}^{q} \theta_j z_{t-j} + z_t \tag{2}$$

The values of p and q are orders for the ARMA model. To abbreviate the mention, the ARMA model that has the order p and q will be written by $ARMA(p, q)$. Given the values of orders p and q, the values $\phi_1, \cdots, \phi_p$ and $\theta_1, \cdots, \theta_q$ express the coefficients of the $ARMA(p, q)$. While the random variables $z_1, \cdots, z_n$ are noise. This noise is assumed to have a Laplace distribution with mean 0 and variance $2\beta^2$. The probability function for the variable z_t is written by the following equation:

$$g\left(z_t \mid \beta\right) = \frac{1}{2\beta} exp - \frac{|z_t|}{\beta}. \tag{3}$$

With a variable transformation between x_t and z_t, the probability function for the variable x_t can be written by

$$f\left(x_t \mid \beta\right) = \frac{1}{2\beta} exp - \frac{\left|\sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{j=1}^{q} \theta_j z_{t-j} + x_t\right|}{\beta}. \tag{4}$$

Suppose that $x = (x_1, \cdots, x_n)$, $\phi^{(p)} = \left(\phi_1, \cdots, \phi_p\right)$ and $\theta^{(q)} = \left(\phi_1, \cdots, \phi_q\right)$. The probability function for x is

$$\begin{aligned} & f\left(x|p, q, \phi^{(p)}, \theta^{(q)}, \beta\right) \\ &= \prod_{t=p+1}^{n} \frac{1}{2\beta} exp - \frac{\left|\sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{j=1}^{q} \theta_j z_{t-j} + x_t\right|}{\beta} \\ &= \left(\frac{1}{2\beta}\right)^{n-p} exp - \frac{1}{\beta} \sum_{t=p+1}^{n} \left|\sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{j=1}^{q} \theta_j z_{t-j} + x_t\right|. \end{aligned} \tag{5}$$

The $ARMA(p, q)$ model is called stationary if and only if the root equation is located outside the unit circle. The stationarity of $ARMA(p, q)$ model refers to the mean, variance, and autocorrelation are all constant over time. Suppose that S_p denotes the stationarity region. This stationary condition is difficult to determine if the order value p is high. To overcome this problem, a transformation is made. Let F be a transformation from $\phi^{(p)} \in S_p$ to $r = \left(r_1, \cdots, r_p\right) \in (-1, 1)^p$ where $r_1, \cdots, r_p$ are functions of partial autocorrelation [16]. With the F transformation, this stationary condition of the $ARMA(p, q)$ model becomes easily determined even though the order value p is high. The $ARMA(p, q)$ model is called stationary if and only if $\left(r_1, \cdots, r_p\right) \in (-1, 1)^p$.

The $ARMA(p, q)$ model is called invertible if and only if the root of equation $01 + \sum_{j=1}^{q} \theta_j B^j = 0$ is located outside the unit circle. The invertibility of $ARMA(p, q)$ model refers to the noises can be inverted into a representation of past observations. The limitation of the $ARMA(p, q)$ model is that if the $ARMA(p, q)$ model is not invertible, the $ARMA(p, q)$ model cannot be used to forecast the future values of the dependent. Suppose I_q denotes an invertible region. This invertibility condition is difficult to determine if the order q value is high. To overcome this problem, a transformation is made. Let G be a transformation from $\theta^{(q)} \in I_q$ to $\rho = \left(\rho_1, \cdots, \rho_q\right) \in (-1, 1)^q$ where $\rho_1, \cdots, \rho_q$ are functions of inverse partial autocorrelation [17]. With the G transformation, the invertibility condition of the $ARMA(p, q)$ model is easily determined even though the order q value is high. The $ARMA(p, q)$ model is called invertible if and only if $\left(\rho_1, \cdots, \rho_q\right) \in (-1, 1)^q$. If the likelihood function for data is expressed in the transformation of F and G, the likelihood function for data can be written by

$$f\left(x|p,q,r^{(p)},\rho^{(q)},\beta\right)$$
$$=\prod_{t=p+1}^{n}\frac{1}{2\beta}exp-\frac{\left|\sum_{i=1}^{p}F^{-1}\left(r_i\right)x_{t-i}-\sum_{j=1}^{q}G^{-1}\left(\rho_j\right)z_{t-j}+x_t\right|}{\beta}$$
$$=\left(\frac{1}{2\beta}\right)^{n-p}exp-\frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p}F^{-1}\left(r_i\right)x_{t-i}\right.$$
$$\left.-\sum_{j=1}^{q}G^{-1}\left(\rho_j\right)z_{t-j}+x_t\right|. \tag{6}$$

In the probability function for this data, F^{-1} and G^{-1} are inverse transformations for the transformation of F and G.

3.2. Prior and Posterior Distributions

The Binomial distribution with the parameter λ is chosen as the prior distribution for order p. The prior distribution for order p can be written by

$$\pi\left(p|\lambda\right)=C_p^{p_{max}}\lambda^p(1-\lambda)^{p_{max}-p} \tag{7}$$

where the p_{max} value is set. Whereas the Binomial distribution with the parameter μ is chosen as the prior distribution for order q. The distribution of priors for q order can be expressed by

$$\pi\left(q|\mu\right)=C_q^{q}\mu^q(1-\mu)^{q_{max}-q} \tag{8}$$

where the value q_{max} is set. For model order p and q, a Binomial distribution is used because it is a conjugate prior.

The prior distribution for $r^{(p)}$ given by the order p is a uniform distribution at $(-1,1)^p$. The uniform distribution is chosen because it is a conjugate prior. The prior distribution for $r^{(p)}$ if given an order p can be written by

$$\pi\left(r^{(p)}|p\right)=\frac{1}{2^p}. \tag{9}$$

The prior distribution for $\rho^{(q)}$ given by the order q is a uniform distribution at $(-1,1)^q$. The prior distribution for $\rho^{(q)}$ if given an order q can be written by

$$\pi\left(\rho^{(q)}|q\right)=\frac{1}{2^q}. \tag{10}$$

The inverse Gamma distribution with parameter ν is selected as the prior distribution for β. Value u is set, namely: $u=1$. The distribution of prior for β can be expressed by

$$\pi\left(\beta|u,\nu\right)=\frac{\nu^u}{\Gamma(u)}\beta^{-(u+1)}exp-\frac{\nu}{\beta}. \tag{11}$$

For β, an inverse Gamma distribution is used because it is a conjugate prior.

This prior distribution contains a parameter, namely: λ, μ, and ν. Uniform distribution at interval $(0,1)$ is chosen as the prior distribution for λ, namely: $\pi(\lambda)=1$. The uniform distribution at interval $(0,1)$ is chosen as the prior distribution for μ, namely: $\pi(\mu)=1$. The uniform distribution is proposed in previous work [7].

Finally, the Jeffreys distribution is chosen as the prior distribution for ν. The prior distribution for ν can be written by $\pi(\nu)\propto\frac{1}{\nu}$. This is a non-informative prior distribution. Thus, the joint prior distribution for $p,q,r^{(p)},\rho^{(q)},\lambda,\mu,\beta$, and ν can be expressed by

$$\pi\left(p,q,r^{(p)},\rho^{(q)},\lambda,\mu,\beta,\nu\right)$$
$$\propto C_p^{p_{max}}\lambda^p\left(1-\lambda\right)^{p_{max}-p}C_q^{q_{max}}\mu^q\left(1-\mu\right)^{q_{max}-q}\frac{1}{2^{(p+q)}}\frac{\nu^u}{\Gamma(u)}\beta^{-(u+1)}$$
$$exp-\frac{\nu}{\beta}\frac{1}{\beta}\frac{1}{\nu}. \tag{12}$$

By using the Bayes theorem, distribution posterior for $p,q,r^{(p)},\rho^{(q)},\lambda,\mu,\beta$, and ν can be expressed by

$$\pi\left(p,q,r^{(p)},\rho^{(q)},\lambda,\mu,\beta,\nu|x\right)$$
$$\propto\left(\frac{1}{2\beta}\right)^{n-p}exp-\frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p}F^{-1}(r_i)x_{t-i}\right.$$
$$\left.-\sum_{j=1}^{q}G^{-1}(\rho_j)z_{t-j}+x_t\right|$$
$$C_p^{p_{max}}\lambda^p(1-\lambda)^{p_{max}-p}C_q^{q_{max}}\mu^q(1-\mu)^{q_{max}-q}\frac{1}{2^{(p+q)}}\frac{\nu^u}{\Gamma(u)}\beta^{-(u+1)}$$
$$exp-\frac{\nu}{\beta}\frac{1}{\beta}\frac{1}{\nu}$$
$$\propto\left(\frac{1}{2}\right)^{n-p}\left(\frac{1}{\beta}\right)^{n-p-1}exp$$
$$-\frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p}F^{-1}(r_i)x_{t-i}-\sum_{j=1}^{q}G^{-1}(\rho_j)z_{t-j}+x_t\right|$$
$$C_p^{p_{max}}\lambda^p(1-\lambda)^{p_{max}-p}C_q^{q_{max}}\mu^q(1-\mu)^{q_{max}-q}\frac{1}{2^{(p+q)}}\frac{\nu^{u-1}}{\Gamma(u)}\beta^{-(u+1)}$$
$$exp-\frac{\nu}{\beta}. \tag{13}$$

This posterior distribution has a complex form so that the Bayes estimator cannot be determined explicitly. The reversible jump MCMC algorithm was adopted to determine the Bayes estimator.

3.3. Reversible Jump MCMC

The basic idea of the MCMC algorithm is to treat the parameters $p,q,r^{(p)},\rho^{(q)},\lambda,\mu,\beta$, and ν as a Markov chain. To determine the Bayes estimator for parameters, a Markov chain is created by simulation. This Markov chain is designed so that it has a limit distribution that approaches the posterior distribution for parameters $p,q,r^{(p)},\rho^{(q)},\lambda,\mu,\beta$, and ν. This Markov chain is used to find the Bayes estimator. The algorithm consists of two stages, namely; the first stage is to create a Markov chain for $\left(\lambda,\mu,\beta,\nu|p,q,r^{(p)},\rho^{(q)}\right)$ by simulation. The second stage makes the Markov chain for $\left(p,q,r^{(p)},\rho^{(q)}|\lambda,\mu,\beta,\nu\right)$ by simulation.

Distribution for $\left(\lambda,\mu,\beta,\nu|p,q,r^{(p)},\rho^{(q)}\right)$ is the product of the distribution for $\left(\lambda|p,q,r^{(p)},\rho^{(q)}\right)$, the distribution for $\left(\mu|p,q,r^{(p)},\rho^{(q)}\right)$, distribution for $\left(\beta|p,q,r^{(p)},\rho^{(q)}\right)$, and the distribution for $\left(\nu|p,q,r^{(p)},\rho^{(q)}\right)$. The distribution for $\left(\lambda|p,q,r^{(p)},\rho^{(q)}\right)$ is Binomial, the distribution for $\left(\mu|p,q,r^{(p)},\rho^{(q)}\right)$ is Binomial, the distribution for $\left(\beta|p,q,r^{(p)},\rho^{(q)}\right)$ is the Gamma inverse distribution, distribution for $\left(\nu|p,q,r^{(p)},\rho^{(q)}\right)$ is the Gamma distribution. Simulation for the Markov chain $(\lambda,\mu,\beta,\nu|p,q,r^{(p)},\rho^{(q)})$ is as follows [18]:

$$\lambda \sim B\left(p+1,\, p_{max}-p+1\right),$$

$$\mu \sim B\left(q+1,\, q_{max}-q+1\right),$$

$$\beta \sim IG\left(n-p, \nu+\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p}F^{-1}(r_i)x_{t-i} - \sum_{j=1}^{q}G^{-1}(\rho_j)z_{t-j}+x_t\right|\right),$$

$$\upsilon \sim G\left(u, \frac{1}{\beta}\right).$$

The distribution for $(p,q,r^{(p)},\rho^{(q)}|\lambda,\mu,\beta,\nu)$ has a complex form. Simulation for the Markov chain $(p,q,r^{(p)},\rho^{(q)}|\lambda,\mu,\beta,\nu)$ is carried out using the reversible jump MCMC algorithm. The reversible jump MCMC algorithm is an extension of the Metropolis–Hastings algorithm [19,20]. This reversible jump MCMC algorithm uses six types of transformation, namely: birth for order p, death for order p, change for coefficient $r^{(p)}$, birth for order q, death for order q, and change for coefficient $\rho^{(q)}$.

3.4. Birth of the Order *p*

Suppose that $w = (p,q,r^{(p)},\rho^{(q)})$ is the old Markov chain and $w^* = \left(p^*,q,r^{*(p^*)},\rho^{(q)}\right)$ is a new Markov chain. The birth of the order p changes the order p and the coefficient $r^{(p)}$ but does not change the order q and the coefficient $\rho^{(q)}$. The new Markov chain $\left(p^*,q,r^{*(p^*)},\rho^{(q)}\right)$ is defined as follows: (a) first determine the order $p^* = p+1$, (b) the second takes randomly a at the interval $(-1,1)$ and then determines $r^{*(p^*)} = \left(r^{(p)},a\right)$. The old Markov chain will be replaced by a new Markov chain with probability

$$\eta_p^{AR}(w,w^*) = min\left\{1, \frac{f(x|w^*)}{f(x|w)}\frac{\pi(p^*)}{\pi(p)}\frac{\pi\left(r^{*(p^*)}\middle|p^*\right)}{\pi\left(r^{(p)}\middle|p\right)}\frac{q(w^*,w)}{q(w,w^*)}\right\} \tag{14}$$

where $\frac{f(x|w^*)}{f(x|w)}$ is the ratio of the likelihood function, $\frac{\pi(p^*)}{\pi(p)}$ is the ratio between the prior distribution for order p and p^*, $\frac{\pi(r^*|p^*)}{\pi(r|p)}$ is the ratio between the posterior distribution for $r^{*(p^*)}$ and $r^{(p)}$, and $\frac{q(w^*,w)}{q(w,w^*)}$ is the ratio between the distribution of w and w^*. In contrast, the old Markov chain will remain with the probability $1-\eta_p^{AR}(w,w^*)$. Following is the calculation for the probability function ratio, the ratio between the prior distribution for order p and p^*, the ratio between the posterior distribution for $r^{*(p^*)}$ and $r^{(p)}$, and the ratio between the distribution of w and w^*.

The ratio for likelihood function can be stated by

$$\frac{f\left(x|p^*,q,r^{*(p^*)},\rho^{(q)}\right)}{f\left(x|p,q,r^{(p)},\rho^{(q)}\right)} = \frac{exp - \frac{1}{\beta}\sum_{t=p^*+1}^{n}\left|\sum_{i=1}^{p^*}F^{-1}\left(r_i^*\right)x_{t-i} - \sum_{j=1}^{q}G^{-1}(\rho_j)z_{t-j}+x_t\right|}{exp - \frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p}F^{-1}(r_i)x_{t-i} - \sum_{j=1}^{q}G^{-1}(\rho_j)z_{t-j}+x_t\right|}\left(\frac{1}{2\beta}\right). \tag{15}$$

The ratio between the prior distribution for order p and p^* can be expressed by

$$\frac{\pi(p^*)}{\pi(p)} = \frac{p_{max}-p}{p+1}\frac{\lambda}{1-\lambda}. \tag{16}$$

The ratio between posterior distribution for order p and p^* can be written by

$$\frac{\pi\left(r^{*(p^*)}\middle|p^*\right)}{\pi\left(r^{(p)}\middle|p\right)} = \frac{1}{2}. \tag{17}$$

While the ratio between the distribution of w and w^* depends on the value of a. If $a < 0$, the ratio between the distribution of w and w^* can be written by

$$\frac{q(w^*,w)}{q(w,w^*)} = \frac{1}{a+1}. \tag{18}$$

If $a > 0$, the ratio between the distribution of w and w^* can be written by

$$\frac{q(w^*,w)}{q(w,w^*)} = \frac{1}{1-a}. \tag{19}$$

3.5. Death of the Order *p* + 1

Death of the order $p+1$ is the opposite of the birth of order p. Let $w = (p,q,r^{(p+1)},\rho^{(q)})$ be the old Markov chain and $w^* = \left(p^*,q,r^{*(p^*)},\rho^{(q)}\right)$ is a new Markov chain. The death of the order $p+1$ changes the order $p+1$ and the coefficient $r^{(p+1)}$ but does not change the order and the coefficient $\rho^{(q)}$. The new Markov chain $\left(p^*,q,r^{*(p^*)},\rho^{(q)}\right)$ is defined as follows: (a) first determine the order $p^* = p$, (b) the second determines $r^{*(p^*)} = r^{(p)}$. The old Markov chain will be replaced with a new Markov chain with probability

$$\delta_p^{AR}(w,w^*) = min\left\{1, \frac{1}{\eta_p^{AR}(w,w^*)}\right\}. \tag{20}$$

In contrast, the old Markov chain will remain with the probability $1-\delta_p^{AR}(w,w^*)$.

3.6. Change of the Coefficient $r^{(p)}$

Let $w = \left(p, q, r^{(p)}, \rho^{(q)}\right)$ be the old Markov chain and $w^* = \left(p^*, q, r^{*\left(p^*\right)}, \rho^{(q)}\right)$ is a new Markov chain. The change in the coefficient $r^{(p)}$ does not change the order p but changes the coefficient $r^{(p)}$. The change in the coefficient $r^{(p)}$ also does not change both the order q and the coefficient $\rho^{(q)}$. The new Markov chain $\left(p^*, q, r^{*\left(p^*\right)}, \rho^{(q)}\right)$ is defined as follows: (a) first determine the order $p^* = p$, (b) second takes randomly $i \in \{1, \cdots, p\}$, (c) third takes randomly b at the interval $(-1, 1)$ and then determines $r^{*\left(p^*\right)} = \left(r_1^*, \cdots, r_i^* = b, \cdots, r_p^*\right)$. The old Markov chain will be replaced by a new Markov chain with probability

$$\varsigma_p^{AR}(w, w^*) = min\left\{1, \frac{f\left(x|w^*\right)}{f\left(x|w\right)} \frac{q\left(w^*, w\right)}{q\left(w, w^*\right)}\right\}. \tag{21}$$

In this change in coefficient, the likelihood function ratio can be written by

$$\frac{f\left(x|p^*, q, r^{*\left(r^*\right)}, \rho^{(q)}\right)}{f\left(x|p, q, r^{(p)}, \rho^{(q)}\right)} = \frac{exp - \frac{1}{\beta}\sum_{t=p^*+1}^{n}\left|\sum_{i=1}^{p^*} F^{-1}\left(r_i^*\right) x_{t-i} - \sum_{j=1}^{q} G^{-1}(\rho_j) z_{t-j} + x_t\right|}{exp - \frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p} F^{-1}(r_i) x_{t-i} - \sum_{j=1}^{q} G^{-1}(\rho_j) z_{t-j} + x_t\right|}. \tag{22}$$

While the ratio between the distribution of w and w^* can be expressed by

$$\frac{q\left(w^*, w\right)}{q\left(w, w^*\right)} = \left(\frac{(1+b)(1-b)}{(1+r_i)(1-r_i)}\right)^{1/2}. \tag{23}$$

3.7. Birth of the Order q

Let $w = \left(p, q, r^{(p)}, \rho^{(q)}\right)$ be the old Markov chain and $w^* = \left(p, q^*, r^{(p)}, \rho^{*\left(q^*\right)}\right)$ is a new Markov chain. The birth for order q changes the order q and the coefficient $\rho^{(q)}$ but does not change the order p and the coefficient $r^{(p)}$. The new Markov chain $\left(p, q^*, r^{(p)}, \rho^{*\left(q^*\right)}\right)$ is defined as follows: (a) first determines the order $q^* = q + 1$, (b) both take randomly c at the interval $(-1, 1)$ and then determine $\rho^{*(q)} = \left(\rho^{(q)}, c\right)$. The old Markov chain will be replaced by a new Markov chain with probability

$$\eta_q^{MA}(w, w^*) = min\left\{1, \frac{f\left(x|w^*\right)}{f\left(x|w\right)} \frac{\pi\left(q^*\right)}{\pi\left(q\right)} \frac{\pi\left(\rho^{*\left(q^*\right)}\middle| q^*\right)}{\pi\left(\rho^{(q)}\middle| q\right)} \frac{q\left(w^*, w\right)}{q\left(w, w^*\right)}\right\} \tag{24}$$

where $\frac{f\left(x|w^*\right)}{f\left(x|w\right)}$ is the ratio for the likelihood function, $\frac{\pi\left(q^*\right)}{\pi\left(q\right)}$ is the ratio between the prior distribution for order q and q^*, $\frac{\pi\left(\rho^*|q^*\right)}{\pi\left(\rho|q\right)}$ is the ratio between the posterior distribution for $\rho^{*\left(q^*\right)}$ and $\rho^{(q)}$, and $\frac{q\left(w^*, w\right)}{q\left(w, w^*\right)}$ is the ratio between the distribution of w and w^*. In contrast, the old Markov chain will remain with the probability $1 - \eta_q^{MA}(w, w^*)$. The following is a calculation for the likelihood function ratio, the ratio between the prior distribution for order q and q^*, the ratio between the posterior distribution for $\rho^{*\left(q^*\right)}$ and $\rho^{(q)}$, and the ratio between the distribution of w and w^*.

The ratio for the likelihood function can be stated by

$$\frac{f\left(x|p, q^*, r^{(p)}, \rho^{*\left(q^*\right)}\right)}{f\left(x|p, q, r^{(p)}, \rho^{(q)}\right)} = \frac{exp - \frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p} F^{-1}(r_i) x_{t-i} - \sum_{j=1}^{q} G^{-1}\left(\rho_j^*\right) z_{t-j} + x_t\right|}{exp - \frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p} F^{-1}(r_i) x_{t-i} - \sum_{j=1}^{q} G^{-1}(\rho_j) z_{t-j} + x_t\right|} \left(\frac{1}{2\beta}\right). \tag{25}$$

The ratio between the prior distribution for order q and q^* can be expressed by

$$\frac{\pi\left(q^*\right)}{\pi(q)} = \frac{q_{max} - q}{q+1} \frac{\mu}{1-\mu}. \tag{26}$$

The ratio between posterior distribution for order q and q^* can be written by

$$\frac{\pi\left(\rho^{*\left(q^*\right)}\middle| q^*\right)}{\pi\left(\rho^{(q)}\middle| q\right)} = \frac{1}{2}. \tag{27}$$

While the value of the ratio between the distribution of w and w^* depends on the value c. If $c < 0$, the ratio between the distribution of w and w^* can be written by

$$\frac{q\left(w^*, w\right)}{q\left(w, w^*\right)} = \frac{1}{c+1}. \tag{28}$$

If $c > 0$, the ratio between the distribution of w and w^* can be written by

$$\frac{q\left(w^*, w\right)}{q\left(w, w^*\right)} = \frac{1}{1-c}. \tag{29}$$

3.8. Death of the Order $q + 1$

The death of order $q + 1$ is the opposite of the birth of order q. Let $w = \left(p, q, r^{(p)}, \rho^{(q+1)}\right)$ be the old Markov chain and $w^* = \left(p, q^*, r^{(p)}, \rho^{*\left(q^*\right)}\right)$ is a new Markov chain. The death of order $q+1$ changes the order $q + 1$ and the coefficient $\rho^{(q+1)}$ but does not change the order p and the coefficient $r^{(r)}$. The new Markov chain $\left(p, q^*, r^{(p)}, \rho^{*\left(q^*\right)}\right)$ is defined as follows: (a) first determines the order $q^* = q$, (b) the second determines $\rho^{*\left(q^*\right)} = \rho^{(q)}$. The old Markov chain will be replaced by a new Markov chain with probability

$$\delta_q^{MA}(w, w^*) = min\left\{1, \frac{1}{\eta_q^{MA}(w, w^*)}\right\}. \tag{30}$$

In contrast, the old Markov chain will remain with the probability $1 - \delta_q^{MA}(w, w^*)$.

3.9. Change of the Coefficient $\rho^{(q)}$

Let $w = (p, q, r^{(p)}, \rho^{(q)})$ be the old Markov chain and $w^* = \left(p, q^*, r^{(p)}, \rho^{*(q^*)}\right)$ is a new Markov chain. The change in coefficient $\rho^{(q)}$ does not change the order of q but changes the coefficient $\rho^{(q)}$. The change in the coefficient $\rho^{(q)}$ also does not change both the order p and the coefficient $r^{(p)}$. The new Markov chain $\left(p, q^*, r^{(p)}, \rho^{*(q^*)}\right)$ is defined as follows: (a) first determines the order $q^* = q$, (b) second take randomly $j \in \{1, \cdots, q\}$, (c) all three take randomly d at the interval $(-1, 1)$ and then determine $\rho^{*(q^*)} = (\rho_1^*, \cdots, \rho_j^* = d, \cdots, \rho_q^*)$. The old Markov chain will be replaced by a new Markov chain with probability

$$\varsigma_q^{MA}(w, w^*) = min\left\{1, \frac{f(x|w^*)}{f(x|w)} \frac{q(w^*, w)}{q(w, w^*)}\right\}. \tag{31}$$

In this change in coefficient, the ratio for the likelihood function can be written by

$$\frac{f\left(x|p, q^*, r^{(p)}, \rho^{*(q^*)}\right)}{f(x|p, q, r^{(p)}, \rho^{(q)})} = \frac{exp - \frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p} F^{-1}(r_i)x_{t-i} - \sum_{j=1}^{q} G^{-1}(\rho_j^*) z_{t-j} + x_t\right|}{exp - \frac{1}{\beta}\sum_{t=p+1}^{n}\left|\sum_{i=1}^{p} F^{-1}(r_i)x_{t-i} - \sum_{j=1}^{q} G^{-1}(\rho_j) z_{t-j} + x_t\right|}. \tag{32}$$

While the ratio between the distribution of w and w^* can be expressed by

$$\frac{q(w^*, w)}{q(w, w^*)} = \left(\frac{(1+d)(1-d)}{(1+\rho_j)(1-\rho_j)}\right)^{1/2}. \tag{33}$$

4. SIMULATIONS

The performance of the reversible jump MCMC algorithm is tested using simulation studies. The basic idea of simulation studies is to make a synthesis time series with a predetermined parameter. Then the reversible jump MCMC algorithm is implemented in this synthesis time series to estimate the parameter. Furthermore, the value of the estimation of this parameter is compared with the value of the actual parameter. The reversible jump MCMC algorithm is said to perform well if the parameter estimation value approaches the actual parameter value.

4.1. First Simulation

One-synthetic time series is made using Equation (2). The length of this time series is 250. The parameters of the ARMA model are presented in Table 1.

Synthetic time series data with ARMA (2, 3) model are presented in Figure 1. This synthetic time series contains noise that has a Laplace distribution.

This synthetic time series is used as input for the reversible jump MCMC algorithm. The algorithm runs as many as 100,000 iterations with a 25,000 burn-in period. The output of the reversible jump MCMC algorithm is a parameter estimation for the synthetic time series model. The histogram of order p for the synthetic ARMA model is presented in Figure 2.

Figure 2 shows that the maximum order of p is reached at value 2. This histogram shows that the order estimate for p is $\hat{p} = 2$. Also, the histogram of order q for the synthetic ARMA model is presented in Figure 3.

Figure 3 shows that the maximal order of q is reached at value 3. This histogram shows that the order estimation for q is $\hat{q} = 3$. Given the values $\hat{p} = 2$ and $\hat{q} = 3$, the estimation of the ARMA (2, 3) model coefficients are presented in Table 2. The estimation of noise variance is also shown in Table 2.

Table 1 The parameter value for synthetic ARMA (2, 3).

(p,q)	$\phi^{(2)}$	$\theta^{(3)}$	β
(2, 3)	$\begin{pmatrix}-0.1921\\-0.4467\end{pmatrix}$	$\begin{pmatrix}-0.2364\\0.4377\\-0.3565\end{pmatrix}$	1

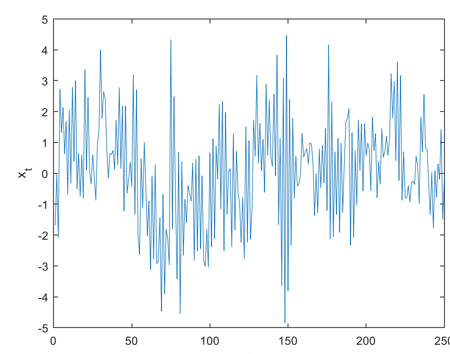

Figure 1 Synthetic time series with ARMA (2, 3) model.

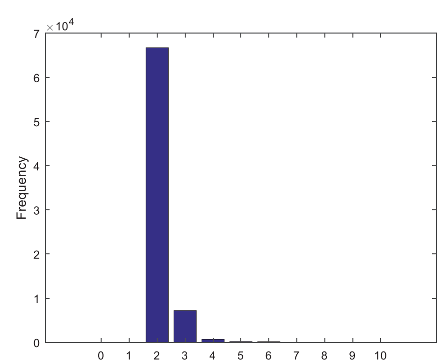

Figure 2 Histogram for order p of the synthetic ARMA (2, 3) model.

If Table 2 is compared with Table 1, the parameter estimation of the ARMA model approaches the actual parameter value.

4.2. Second Simulation

Other one-synthetic time series is made using Equation (2). The length of this time series is also 250, but the parameters are different. The parameters of the ARMA model are presented in Table 3

In this second simulation, the parameters of synthetic ARMA model are taken differently. Synthetic ARMA (4, 2) time series data are presented in Figure 4. This synthetic time series data contains noise that also has a Laplace distribution.

This synthetic time series is used as input for the reversible jump MCMC algorithm. The algorithm runs as many as 100,000 iterations with a 25,000 burn-in period. The output of the reversible jump MCMC algorithm is a parameter estimation for the synthetic time series model. The histogram of order p for the synthetic ARMA model is presented in Figure 5.

Figure 5 shows that the maximum order of p is reached at 4. This histogram shows that the order estimate for p is $\hat{p} = 4$. Also, the histogram of order q for the synthetic ARMA model is presented in Figure 6.

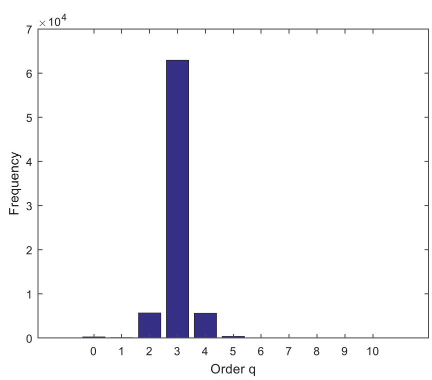

Figure 3 Histogram for order q of the synthetic ARMA (2, 3) model.

Table 2 Estimation of the parameter for the ARMA (2, 3) model.

$(\hat{p},\hat{q})$	$\hat{\phi}^{(2)}$	$\hat{\theta}^{(3)}$	$\hat{\beta}$
(2, 3)	$\begin{pmatrix} -0.2530 \\ -0.5732 \end{pmatrix}$	$\begin{pmatrix} -0.3267 \\ 0.3372 \\ -0.3765 \end{pmatrix}$	1.0058

Table 3 The parameter value for synthetic ARMA (4, 2).

(p,q)	$\phi^{(4)}$	$\theta^{(2)}$	β
(4, 2)	$\begin{pmatrix} 1.4773 \\ 1.1530 \\ 1.4550 \\ 0.9752 \end{pmatrix}$	$\begin{pmatrix} -0.5266 \\ 0.2662 \end{pmatrix}$	1

Figure 6 shows that the maximal order of q is reached at value 2. This histogram shows that the order estimation for q is $\hat{q} = 2$. Given the value of $\hat{p} = 4$ and $\hat{q} = 2$, the estimation of the ARMA (4, 2) model coefficients are presented in Table 4. The estimation of noise variance is also shown in Table 4.

If Table 4 is compared with Table 3, the parameter estimation of the ARMA model approaches the actual parameter value. The first simulation and the second simulation show that the reversible jump MCMC algorithm can estimate the parameter of the ARMA model correctly.

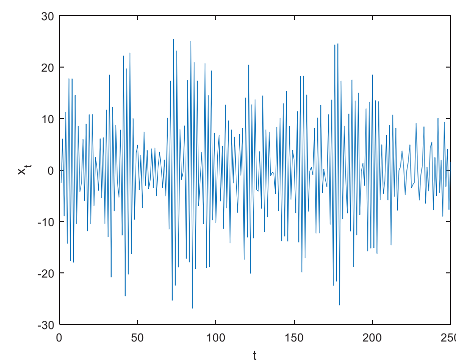

Figure 4 Synthetic time series with ARMA (4, 2) model.

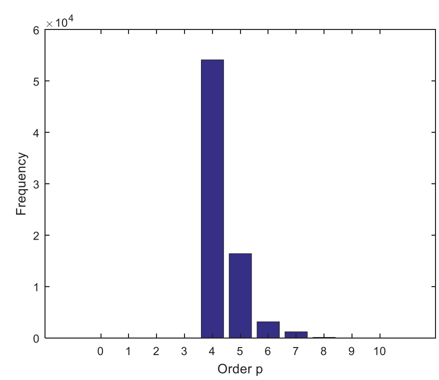

Figure 5 Histogram for order p of the synthetic ARMA (4, 2) model.

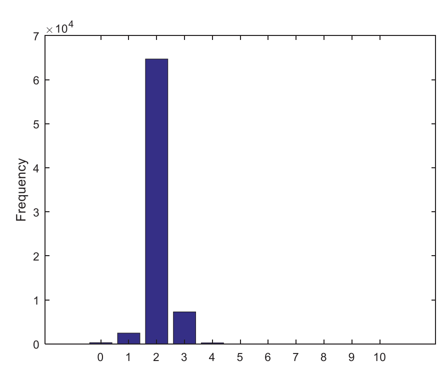

Figure 6 Histogram for order q of the synthetic ARMA (4, 2) model.

Table 4 Estimation of the parameter for the ARMA (4, 2) model.

$(\hat{p},\hat{q})$	$\hat{\phi}^{(4)}$	$\hat{\theta}^{(2)}$	$\hat{\beta}$
(4, 2)	$\begin{pmatrix} 1.4393 \\ 1.1033 \\ 1.4193 \\ 0.9323 \end{pmatrix}$	$\begin{pmatrix} -0.5146 \\ 0.2518 \end{pmatrix}$	1.1015

If the underlying model is wrong will produce a biased estimator. So that the ARMA model is not suitable for modeling data. There is a way to identify whether the wrong model is chosen in the following way. The model is used to predict nth data based on previous $n - 1$ data. Then, the difference between the nth data forecast value and the nth data value is calculated. If the difference is relatively small, the model chosen is correct. Conversely, if the difference is large, the chosen model is wrong.

5. CONCLUSION

This paper is an effort to develop a stationary and invertible ARMA model by assuming that noise has a Laplacian distribution. The ARMA model can be used to describe future behavior only if the ARMA model is stationary. The ARMA can be used to forecast the future values of the dependent variable only if the ARMA model is invertible. Identification of ARMA model orders, estimation of ARMA model coefficients, and estimation of noise variance carried out simultaneously in the Bayesian framework. The Bayes estimator is determined using the MCMC reversible jump algorithm. The performance of reversible jump MCMC is validated in the simulations. The simulation shows that the reversible jump MCMC algorithm can estimate the parameters of the ARMA model correctly.

ACKNOWLEDGMENTS

The author would like to thank to the Ahmad Dahlan University (Indonesia) in providing grant to publish this paper. Also, the author would like to thank the reviewers who commented on improving the quality of this paper. Not to forget, the author would like to thank Professor A. Salhi, Essex University (UK).

REFERENCES

[1] S.-J. Huang, K.-R. Shih, Short-term load forecasting via ARMA model identification including non-Gaussian process considerations, IEEE Trans. Power Syst. 18 (2003), 673–679.

[2] H. Li, R. Li, Y. Zhao, Wind speed forecasting based on autoregressive moving average-exponential generalized autoregressive conditional heteroscedasticity-generalized error distribution (ARMA-EGARCH-GED) model, Int. J. Phy. Sci. 6 (2011), 6867–6871.

[3] M.R. Makhwiting, M. Lesaoana, C. Sigauke, Modelling volatility and financial market risk of shares on the Johannesburg stock exchange, Afr. J. Bus. Manag. 6 (2012), 8065–8070.

[4] M.S. Aminzadeh, Sequential and non-sequential acceptance sampling plans for autocorrelated processes using ARMA (p, q) models, Comput. Stat. 24 (2009), 95–111.

[5] I. Turki, B. Laignel, N. Kakeh, L. Chevalier, S. Costa, A new hybrid model for filling gaps and forecast in sea level: application to the eastern English channel and north Atlantic sea (Western France), Ocean Dyn. 65 (2015), 509–521.

[6] K. Benmouiza, A. Chekname, Small-scale solar radiation forecasting using ARMA and nonlinear autoregressive neural network models, Theor. Appl. Climatol. 124 (2016), 945–958.

[7] Suparman, M. Doisy, Hierarchical Bayesian of ARMA models using simulated annealing algorithm, Telkomnika. 12 (2014), 87–96.

[8] B. Friedlander, B. Porat, Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from high-order moments, IEEE Trans. Automatic Control. 35 (1990), 27–35.

[9] G.B. Giannakis, J.M. Mendel, Cumulant-based order determination of non-Gaussian ARMA models, IEEE Trans. Acoustics. Speech. Signal Process. 38 (1990), 1411–1423.

[10] A. Minchole, L. Sornmo, P. Laguna, Detection of body position changes from the ECG using a Laplacian noise model, Biomed. Signal Process. Control. 14 (2014), 189–196.

[11] F.I. Miertoiu, B. Dumitrescu, Feasibility pump algorithm for sparse representation under Laplacian noise, Math. Probl. Eng. 2019 (2019), 1–9.

[12] P. Neri, The statistical distribution of noisy transmission in human sensors, J. Neural Eng. 10 (2013), 016014.

[13] P.J. Green, Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, Biometrika. 82 (1995), 711–732.

[14] J. Vermaak, C. Andrieu, A. Doucet, S.J. Godsill, Reversible jump Markov chain Monte Carlo strategies for Bayesian model selection in autoregressive processes, J. Time Ser. Anal. 25 (2004), 785–809.

[15] Suparman, Reversible jump MCMC method for hierarchical Bayesian model selection in moving average model, Int. J. Geomate. 16 (2019), 9–15.

[16] O. Barndorff-Nielsen, G. Schou, On the parametrization of autoregressive models by partial autocorrelation, J. Multivar. Anal. 3 (1973), 408–419.

[17] R.J. Bhansali, The inverse partial correlation function of a time series and its applications, J. Multivar. Anal. 13 (1983), 310–327.

[18] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. 6 (1984), 721–741.

[19] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equations of state calculations by fast computing machines, J. Chem. Phy. 21 (1953), 1087–1092.

[20] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika. 57 (1970), 97–109.