Forecasting Software Using Laplacian AR Model based on Bootstrap-Reversible Jump MCMC: Application on Stock Price Data

Suparman
Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Ahmad Dahlan, Yogyakarta, Indonesia.

Yahya Hairun*
Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Khairun, Ternate, Indonesia. E-mail: hairunyahya@gmail.com

Idrus Alhaddad
Faculty of Teacher Training and Education, Universitas Khairun, Ternate, Indonesia.

Tedy Machmud
Faculty of Mathematics and Sciences, Universitas Negeri Gorontalo, Gorontalo, Indonesia.

Hery Suharna
Department of Mathematics Education, Faculty of Teacher Training and Education, Universitas Khairun, Ternate, Indonesia.

Mohd Saifullah Rusiman
Department of Mathematics and Statistics, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia.

Received April 03, 2021; Accepted July 10, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V18SI04/WEB18180

Abstract

The application of the Bootstrap-Metropolis-Hastings algorithm is limited to fixed dimension models. In various fields, data often has a variable dimension model. The Laplacian autoregressive (AR) model includes a variable dimension model so that the Bootstrap-Metropolis-Hasting algorithm cannot be applied. This article aims to develop a Bootstrap reversible jump Markov Chain Monte Carlo (MCMC) algorithm to estimate the Laplacian AR model. The parameters of the Laplacian AR model were estimated using a Bayesian approach. The posterior distribution has a complex structure so that the Bayesian estimator cannot be calculated analytically. The Bootstrap-reversible jump MCMC algorithm was applied to calculate the Bayes estimator. This study provides a procedure for estimating the parameters of the Laplacian AR model. Algorithm performance was tested using simulation studies. Furthermore, the algorithm is applied to the finance sector to predict stock price on the stock market. In general, this study can be useful for decision makers in predicting future events. The novelty of this study is the triangulation between the bootstrap algorithm and the reversible jump MCMC algorithm. The Bootstrap-reversible jump MCMC algorithm is useful especially when the data is large and the data has a variable dimension model. The study can be extended to the Laplacian Autoregressive Moving Average (ARMA) model.

Keywords

Autoregressive, Bootstrap, Laplace Noise, Reversible Jump MCMC.

Introduction

Forecasting is an activity to make predictions of future events. Forecasting methods can be grouped into 2 categories, namely: qualitative methods and quantitative methods (Makridakis et al., 1983). In qualitative method, forecasting is based on expert judgment. While in quantitative methods, forecasting is based on the relationship between variables. Forecasting is applied in various fields. Forecasting applications can be found in various literatures, for example (Guizzardi et al., 2021) and (Kwas, 2021). From time to time, the need for forecasting is increasing along with the increasing fields that require forecasting. Therefore, the development of accurate forecasting methods is an important topic for research.

Time series is one of the quantitative forecasting methods. One of the stochastic models in the time series is the autoregressive (AR) model (Box et al., 2015). Several authors have conducted research on the topic of AR models, for example (Brouste et al., 2014) and (Suparman & Rusiman, 2018). In these studies, noise is assumed to be normally distributed. In some applications, the data is often found to be not normally distributed. To solve this problem, several studies assume that noise is not normally distributed, for example (Larbi & Fellag, 2016) and (Suparman & Diponegoro, 2020). In (Larbi & Fellag, 2016), the AR model assumes that noise has an exponential distribution. While in (Suparman & Diponegoro, 2020), Laplace noise is used in the AR Model. Research related to AR with Laplace noise has not been done much. Therefore, an estimation method for the Laplace AR noise model needs to be developed.

In (Suparman & Diponegoro, 2020), the parameters of the AR model with Laplace noise are estimated using a Bayesian approach. The order of the AR model is assumed to be unknown and the order of the AR model is estimated based on the data. Thus, the posterior distribution has a complex structure so that the Bayesian estimator cannot be calculated analytically. In (Suparman & Diponegoro, 2020), the Bayes estimator is estimated using the reversible jump Markov chain Monte Carlo (MCMC) method.

The reversible jump MCMC method is a very powerful tool for analyzing data of complex structures. The Markov Monte Carlo chain method requires a large number of iterations when applied to big data (Liang et al., 2016). In (Liang et al., 2016), the Metropolis-Hastings method is combined with the Bootstrap method to analyze big data. Bootstrap is a data-based simulation method for statistical inference, which can be used to generate inferences (Efron & Tibshirani, 1993). However, the Bootstrap-Metropolis-Hastings method assumes that the dimensions of the model parameters are constants. For model parameters with variable dimensions, the Bootstrap-Metropolis-Hastings method cannot be used.

This article aims to develop a forecasting method on an AR model with Laplace noise based on Bootstrap-reversible jump MCMC. The focus of this research is on quantitative forecasting method using AR model with Laplacian noise. As an application example, forecasting methods are applied to the finance sector to predict stock prices on the Indonesian stock exchange.

Literature Review

This section briefly describes the basic ideas about the topic, Bootstrap, reversible jump MCMC, and the Laplacian AR model.

Bootstrap

The basic idea in bootstrap is resampling using the actual sample to simulate the distribution of the relevant test statistic. Let $x_1, \ldots, x_n$ be n actual data. Let $\hat{F}$ be an empirical distribution, assigning a probability of $\frac{1}{n}$ to each observed value x_i $(i = 1, \ldots, n)$. A bootstrap sample is defined as a random sample of size n taken from $\hat{F}$, say $x_1^*, \ldots, x_n^*$. This Bootstrap sample $x_1^*, \ldots, x_n^*$ is a sample of size n taken with replacement from the population of n objects $(x_1, \ldots, x_n)$ (Efron & Tibshirani, 1993).

Reversible Jump MCMC

Reversible jump MCMC (Green, 1995) is an extension of the standard MCMC. In reversible jump MCMC, the Markov chain can jump between parameter subspaces of different dimensions. Therefore, the reversible jump MCMC can be applied in the model determination problem. In this article, reversible jump MCMC is applied to estimate the parameters of the Laplacian AR model.

Laplacian AR Model

Let $x_1, \dots, x_n$ be time series data where n is the number of observations. The time series $x_1, \dots, x_n$ has an AR model if this time series satisfies (Box et al., 2015):

$$x_t = \sum_{i=}^{p} \psi_t x_{t-i} + \epsilon_t \qquad (1)$$

where p is the order of the model, $\psi_1, \dots, \psi_p$ is the model coefficient, and $\epsilon_1, \dots, \epsilon_n$ is the noise. In this article, time series data $x_1, \dots, x_n$ are modeled as Laplacian AR. The Laplacian AR model means that the autoregressive model contains noise that has a Laplace distribution. Research related to Laplace noise can be found in various literatures, for example (Miertoiu & Dumitrescu, 2019), (Minchole et al., 2014) and (Suparman & Diponegoro, 2020).

The random variable ϵ is said to have a Laplace distribution with parameters μ and β, written as $\epsilon \sim Laplace(\mu, \beta)$, if the probability density function of ϵ has the following form:

$$f(\epsilon|\mu, \beta) = \frac{1}{2\beta} exp - \frac{|\epsilon - \mu|}{\beta}.$$

In this article, $\epsilon \sim Laplace(0, \beta)$.

Methodology

This study uses data that has an AR model with Laplace noise. Let $x_1, \dots, x_n$ be the data that satisfies the AR model with Laplace noise where n is the number of data. Next, this data is resampled with replacement to get Bootstrap samples. Let B be the number of resamplings. Then, the j-th bootstrap sample $j = 1, 2, \dots, B)$ was used to estimate the AR model parameters with Laplace noise. Parameter estimation using Bayesian method. The steps in the Bayesian method include determining the likelihood function, selecting the prior distribution, determining the posterior distribution, and calculating the Bayesian estimator using the reversible jump MCMC algorithm. After that, the j-th Bayesian estimator $j = 1, 2, \dots, B)$ is used to calculate the Bootstrap estimator of the parameters. Finally, the AR model with Laplace noise is used to predict the value of x_{n+1}. The procedure for calculating the Bootstrap-reversible jump MCMC estimator is presented in Figure 1.

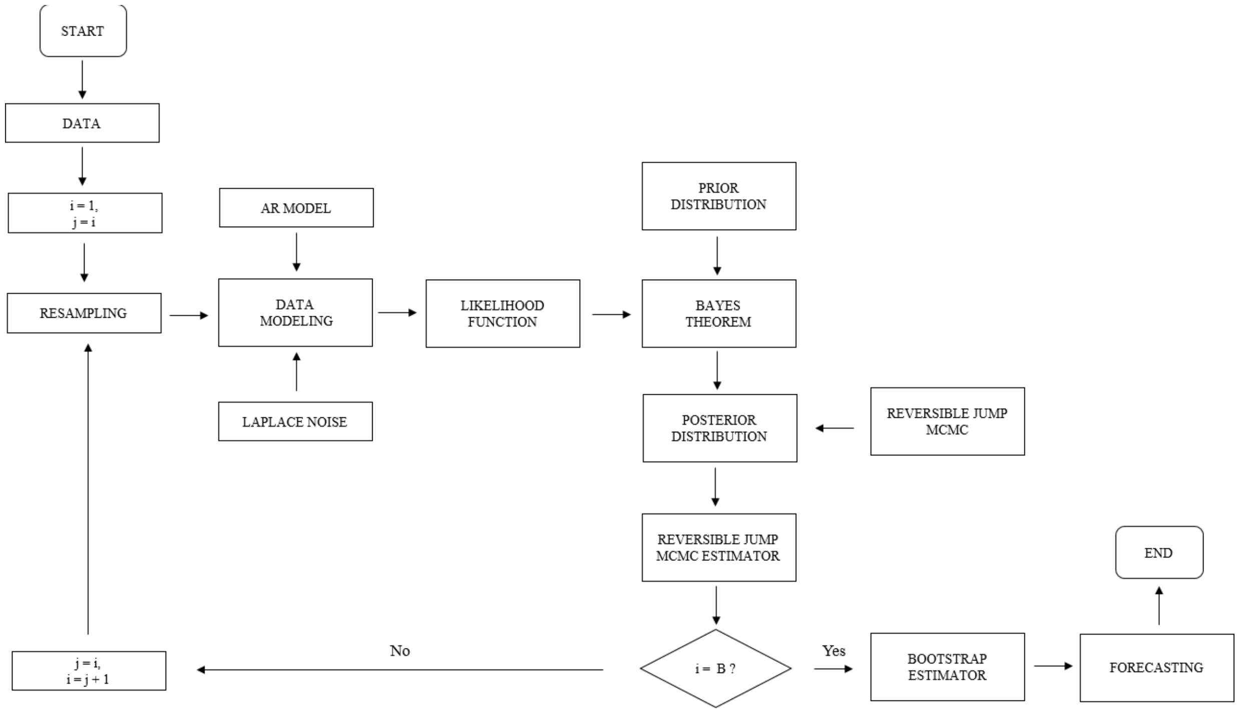

Figure 1 Flowchart

Results and Discussion

Likelihood Function of Bootstrap Samples

Let $x_1, \dots, x_n$ be the data of size n and $x_1^{*j}, \dots, x_n^{*j}$ is the j-th Bootstrap sample $(j = 1, \dots\dots, B)$. Bootstrap sample $x_1^{*j}, \dots, x_n^{*j}$ is obtained by taking n points with replacement from $\{x_1, \dots, x_n\}$ (Efron & Tibshirani, 1993).

In this article, the AR model is assumed to have a Laplace noise with a mean of 0 and a variance of $2\beta^2$. Thus, the likelihood function of the jth Bootstrap sample can be written as

$$f\left(x^{*j} \middle| p, \psi^{(p)}, \beta\right) = \left(\frac{1}{2\beta}\right)^{n-p} exp - \frac{1}{\beta} \sum\nolimits_{t=p+1}^{n} \left| \sum\nolimits_{i=1}^{p} \psi_i x_{t-i}^{*j} + x_t^{*j} \right|$$

where $x^{*j} = \left(x_1^{*j}, \dots, x_n^{*j}\right)$ and $\psi^{(p)} = \left(\psi_1, \dots, \psi_p\right)$. The stationary AR model is very useful in forecasting. For small model orders, the stationary region of the AR model is easy to identify. However, for a large order model, the stationarity region of the AR model is very difficult to identify. To solve this problem, the coefficients of the AR model are transformed into a partial autocorrelation function (Barndorff-Nielsen & Schou, 1973). Let Ω_p be the stationary region of the AR model and F is a transformation from $\psi^{(p)} = \left(\psi_1, \dots, \psi_p\right) \in$

Ω_p to $r^{(p)} = (r_1, \dots, r_p) \in (-1, 1)^p$ where $r_1, \dots, r_p$ are the partial autocorrection functions of the AR model. Through this transformation F, an AR model with order p is said to be stationary if $|r_i| < 0$ for $i = 1, \dots, p$. With the new parameters, the likelihood function of the j-th Bootstrap sample becomes

$$f\left(x^{*j} \middle| p, r^{(p)}, \beta\right) = \left(\frac{1}{2\beta}\right)^{n-p} exp - \frac{1}{\beta} \sum_{t=p+1}^{n} \left| \sum_{i=1}^{p} F^{-1}(r_i) x_{t-i}^{*j} + x_t^{*j} \right|$$

where F^{-1} is the inverse transformation of F.

Prior Dan Posterior Distributions

AR model parameters include p, $r^{(p)}$, λ, β, and v. As in (Suparman & Diponegoro, 2020), the prior distribution of $\left(p, r^{(p)}, \lambda, \beta, v\right)$ is

$$\pi\left(p, r^{(p)}, \lambda, \beta, v\right) = C_p^{p_{max}} \lambda^p (1-\lambda)^{p_{max}-p} \frac{1}{2^p} \frac{v^u}{\Gamma(u)} \beta^{-(u+1)} exp - \frac{v}{\beta} \frac{1}{v}.$$

According to Bayes' Theorem, the posterior distribution of $\left(p, r^{(p)}, \lambda, \beta, v\right)$ is

$$\pi\left(p, r^{(p)}, \lambda, \beta, v \middle| x^{*j}\right)$$
$$= \left(\frac{1}{2}\right)^{n-p} \left(\frac{1}{\beta}\right)^{n-p-1} exp - \frac{1}{\beta} \sum_{t=p+1}^{n} \sum_{t=p+1}^{n} \left| \sum_{i=1}^{p} F^{-1}(r_i) x_{t-i}^{*j} + x_t^{*j} \right|$$
$$C_p^{p_{max}} \lambda^p (1-\lambda)^{p_{max}-p} \frac{1}{2^p} \frac{v^{u-1}}{\Gamma(u)} \beta^{-(u+1)} exp - \frac{v}{\beta}.$$

The posterior distribution has a complex structure so that the Bayesian estimator is not calculated analytically. Therefore, the parameters $\left(p, r^{(p)}, \lambda, \beta, v\right)$ are estimated using the reversible jump MCMC algorithm.

Bootstrap-Reversible Jump MCMC

The posterior distribution is simulated in 2 stages, namely: simulation of conditional distribution of (λ, β, v) if given $\left(p, r^{(p)}\right)$ and simulation of conditional distribution of $\left(p, r^{(p)}\right)$ if given (λ, β, v). Furthermore, the conditional distribution of (λ, β, v) given $\left(p, r^{(p)}\right)$ is simulated in 3 stages, namely: $\beta \sim IG\left(n-p, v + \sum_{t=p+1}^{n} \left| \sum_{i=1}^{p} F^{-1}(r_i) x_{t-i}^{*j} + x_t^{*j} \right|\right)$, $\lambda \sim B(p+1, p_{max} - p + 1)$, and $v \sim G(u, \frac{1}{\beta})$.

However, the conditional distribution of $\left(p, r^{(p)}\right)$ given (λ, β, v) has a complex structure. Therefore, the conditional distribution of $\left(p, r^{(p)}\right)$ if given (λ, β, v) is simulated using the reversible jump MCMC algorithm. Furthermore, the reversible jump MCMC algorithm uses 3 types of transformations, namely: change in coefficient, birth of order, and death of order (Suparman & Diponegoro, 2020).

Let $\left(p^{*j}, r^{*j(p^{*j})}, \lambda^{*j}, \beta^{*j}, \nu^{*j}\right)$ be the Bayesian estimator obtained by the MCMC reversible jump algorithm based on the j-th Bootstrap sample for $j = 1, \ldots, B$. The Bootstrap-reversible jump MCMC estimator of the parameters $\left(p, \psi^{(p)}, \lambda, \beta, v\right)$ is calculated using the following formulas: $\hat{p} = \frac{1}{B}\sum_{j=1}^{B} p^{*j}$, $\hat{\psi}^{\widehat{(p)}} = \frac{1}{B}\sum_{j=1}^{B} F^{-1}(r^{*j(p^{*j})})$, $\hat{\lambda} = \frac{1}{B}\sum_{j=1}^{B} \lambda^{*j}$, $\hat{\beta} = \frac{1}{B}\sum_{j=1}^{B} \beta^{*j}$, and $\hat{\nu} = \frac{1}{B}\sum_{j=1}^{B} \nu^{*j}$.

Simulated Data

The accuracy of the Bootstrap-reversible jump MCMC method was demonstrated through a simulation study. Simulated data is made according to equation (1). The number of simulated data is 250. The order model is $p = 2$ and the coefficient of the AR model is $\psi^{(2)} = (-0.44, 0.43)$. The noise is assumed to have a Laplace distribution with $\mu = 0$ and $\beta = 2$. Simulated data is given in Figure 2.

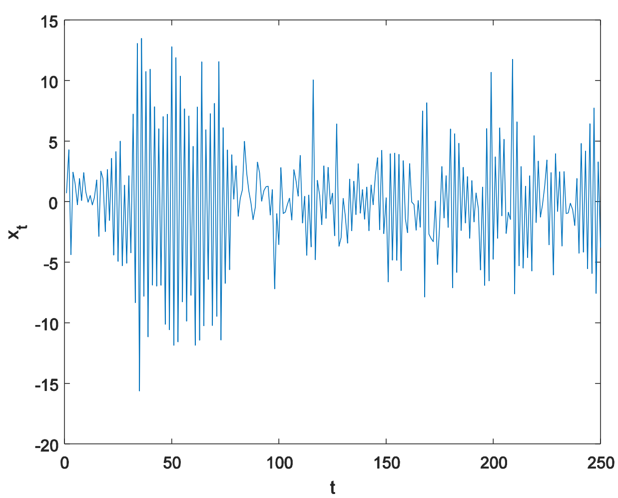

Figure 2 Simulated data

Furthermore, this simulated data is resampled as much as $B = 11$. The reversible jump algorithm is implemented on the j-th Bootstrap sample ($j = 1, \ldots.., 11$). The estimates of the order and coefficients of the Laplacian AR model for each Bootstrap sample are presented in Table 1.

Table 1 Order and coefficient of Laplacian AR model for each Bootstrap sample using simulated data

j-th Bootstrap Sample	p^{*j}	ψ_1^{*j}	ψ_2^{*j}
1	2	-0.3609	0.4551
2	2	-0.3898	0.4996
3	2	-0.3866	0.5182
4	2	-0.4039	0.4254
5	2	-0.4836	0.3799
6	2	-0.3900	0.4756
7	2	-0.4653	0.3966
8	2	-0.4652	0.4503
9	2	-0.3819	0.4742
10	2	-0.4066	0.4772
11	2	-0.3883	0.4743

Based on Table 1, the Bootstrap-reversible jump MCMC estimator of the order is $\hat{p} = 2$. While the Bootstrap-reversible jump MCMC estimator of the coefficients is $\hat{\psi}^{(2)} = (-0.4411, 0.4337)$.

Application on Stock Price Data

The Bootstrap-reversible jump MCMC algorithm is implemented in the finance sector, especially in forecasting stock prices on the stock market. One of the stock price data on the Indonesia Stock Exchange (IDX), say stock C, is presented in Figure 3. The stock price data is observed from January 4, 2021 to April 20, 2021. The number of stock price data is 82. This stock price data has also been used in (Suparman et al., 2021).

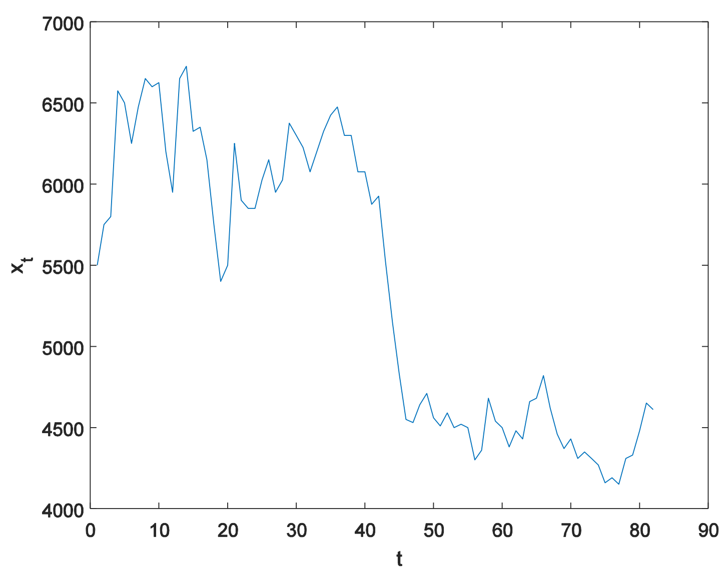

Figure 3 Stock C

As for the simulated data, the price data for stock C is also resampled as much as $B = 11$. The reversible jump algorithm is run on the j-th Bootstrap sample ($j = 1, \dots ,11$). Table 2 presents the estimators of the order and coefficients of the Laplacian AR model for each Bootstrap sample.

Table 2 Order and coefficient of Laplacian AR model for each Bootstrap sample using the price data for stock C

j-th Bootstrap Sample	p^{*j}	ψ_1^{*j}
1	1	0.9522
2	1	0.9517
3	1	0.9518
4	1	0.9517
5	1	0.9525
6	1	0.9525
7	1	0.9516
8	1	0.9525
9	1	0.9514
10	1	0.9516
11	1	0.9526

Based on Table 2, the Bootstrap-reversible jump MCMC estimator of the order is $\hat{p} = 1$. While the Bootstrap-reversible jump MCMC estimator of the coefficients is $\hat{\psi}^{(1)} = (0.9520)$.

Discussion

In the simulation study, the values of the Bootstrap-reversible jump MCMC estimator of the order and coefficients of the Laplacian AR model are $\hat{p} = 2$ and $\hat{\psi}^{(2)} = (-0.4411, 0.4337)$. This estimator value is close to the parameter values used in the Laplacian AR model to create simulated data, namely: $p = 2$ and $\psi^{(2)} = (-0.44, 0.43)$. The results of this simulation show that the Bootstrap-reversible jump MCMC algorithm can estimate the order and coefficients of the Laplacian AR model very accurately.

In an application to forecast prices for stock C, the Bootstrap-reversible jump MCMC estimator of the order is $\hat{p} = 1$. While the Bootstrap-reversible jump MCMC estimator of the coefficients is $\hat{\psi}^{(1)} = (0.9520)$. Suppose $\hat{x}_{83}$ is a stock price forecast one day ahead. Then, $\hat{x}_{83} = 0.9520\ (4610) = 4389$. The price of Stock C is predicted to decline in the next one day. The estimation results are not much different from the results in (Suparman et al., 2021). Estimator for order and coefficient in (Suparman et al., 2021) is $\hat{\psi}^{(1)} = (0.95)$.

This study can be applied to forecasting in various fields in order to make decisions, for example in the field of finance. To expand the application, further research can be carried out on the development of the Bootstrap-reversible jump MCMC algorithm for the Laplacian Autoregressive Moving Average (ARMA) model. The AR model is a special form of the ARMA model when the order q is equal to zero.

Conclusion

This article has studied the parameter estimation procedure of the Laplacian AR model using Bootstrap-reversible jump MCMC. Bootstrap-reversible jump MCMC algorithm is used to overcome the Laplacian AR model where the order of the model is assumed to be unknown. The focus of this research is on the Laplacian AR model. The AR model is a special form of the ARMA model so that research can be extended to the development of the parameter estimation procedure of the Laplacian ARMA model.

The performance of the Bootstrap-reversible jump MCMC algorithm was evaluated using simulated data. The simulation study results show that this algorithm can estimate the parameters of the Laplacian AR model accurately.

The advantage of this research is that both the order model and the model coefficients are estimated simultaneously using the data. Research has been well applied to the field of finance, especially to forecast stock prices. The method proposed in this study is recommended to be used for forecasting in other fields.

Acknowledgements

The authors would like to thank Universitas Ahmad Dahlan (Indonesia) for providing a research grant with a basic research scheme so that this research can be completed properly. The authors would also like to thank the reviewers for their comments and suggestions. Comments and suggestions from reviewers are very useful for improving the quality of articles so that this article becomes better.

References

Barndorff-Nielsen, O., & Schou, G. (1973). On the parametrization of autoregressive models by partial autocorrelations. *Journal of multivariate Analysis, 3*(4), 408-419.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., & Ljung, G.M. (2015). *Time Series Analysis: Forecasting and Control* (5th Editio). John Wiley & Sons.

Brouste, A., Cai, C., & Kleptsyna, M. (2014). Asymptotic properties of the MLE for the autoregressive process coefficients under stationary Gaussian noise. *Mathematical Methods of Statistics*, *23*(2), 103–115.

Efron, B., & Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. Chamman & Hall.

Green, P.J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, *82*(4), 711–743.

Guizzardi, A., Pons, F.M.E., Angelini, G., & Ranieri, E. (2021). Big data from dynamic pricing: A smart approach to tourism demand forecasting. *International Journal of Forecasting*, *37*(3), 1049–1060.

Kwas, M. (2021). *Forecasting Commodity Prices : Looking for a Benchmark*, 447–459.

Larbi, L., & Fellag, H. (2016). Robust bayesian analysis of an autoregressive model with exponential innovations. *Afrika Statistika*, *11*(1), 955-964.

Liang, F., Kim, J., & Song, Q. (2016). A Bootstrap Metropolis-Hastings Algorithm for Bayesian Analysis of Big Data. *Technometrics*, *58*(3), 304–318.

Makridakis, S., Wheelwright, S.C., & McGee, V.E. (1983). *Forecasting Methods and Applications*. Wiley & Sons.

Miertoiu, F.I., & Dumitrescu, B. (2019). Feasibility pump algorithm for sparse representation under Laplacian noise. *Mathematical Problem in Engineering*, 1–9.

Minchole, A., Sornmo, L., & Laguna, P. (2014). Detection of body position changes from the ECG using a Laplacian noise model. *Biomedical Signal Processing and Control*, *14*, 189–196.

Suparman, & Diponegoro, A.M. (2020). *Mathematical Model with Laplace Autoregressive Process for Heart Rate Signals*, *62*(3), 1001-1011.

Suparman, S., Diponegoro, A.M., Ritonga, M., Hairun, Y., Machmud, T., & Price, S. (2021). *Triangulation between Bernoulli Distribution and Laplacian Autoregressive Model to Predict Probability of Increase in Stock Price*. *9*(4), 588–593. https://doi.org/10.13189/ujaf.2021.090406

Suparman, S., & Rusiman, M.S. (2018). Hierarchical Bayesian Estimation for Stationary Autoregressive Models Using Reversible Jump MCMC Algorithm. *International Journal of Engineering & Technology*, *7*(4.30), 64–67. www.sciencepubco.com/index.php/IJET

Keshavarz, H., Givi, M.E., & Vafaeian, A. (2016). Students' sense of self-efficacy in searching information from the Web: A PLS approach. *Webology*, *13*(2), 16-31.