**JTAM (Jurnal Teori dan Aplikasi Matematika)**
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 8, No. 4, October 2024, pp. 1096-1108

# Improvement of Real-GJR Model using Jump Variables on High Frequency Data

**Didit Budi Nugroho[1], Nadya Putri Wulandari[2], Yumita Cristin Alfagustina[2], Hanna Arini Parhusip[1], Faldy Tita[2], Bambang Susanto[1]**
[1]Master of Data Science, Universitas Kristen Satya Wacana, Indonesia
[2]Department of Mathematics, Universitas Kristen Satya Wacana, Indonesia
didit.budinugroho@uksw.edu

**Article History:**
Received : 06-06-2024
Revised : 16-09-2024
Accepted : 18-09-2024
Online : 01-10-2024

**Keywords:**
Adaptive Random Walk;
Continuous and Jump;
Realized GJR;
TOPIX.

**ABSTRACT**

Volatility is a key indicator in assessing risk when making investment decisions. In the world of financial markets, volatility reflects the degree to which the value of a financial asset fluctuates over a given period. The most common way to measure the future loss potential of an investment is through volatility. Focusing on the Realized GJR (RealGJR) volatility model, which consists of return, conditional volatility, and measurement equations, this study proposes the RealGJR-CJ model developed by decomposing the exogenous variable in the volatility equation of RealGJR into continuous C and discontinuous (jump) J variables. The decomposition of exogenous variables makes the RealGJR-CJ model follow realistic financial markets, where the asset volatility is a continuous process with some jump components. As an empirical illustration, the models are applied to an index in the Japanese stock market, namely Tokyo Stock Price Index, covering from January 2004 to December 2011. The observed exogenous variable in the volatility equation of RealGJR models is Realized Volatility (RV), which is calculated using intraday data with time intervals of 1 and 5 minutes. Adaptive Random Walk Metropolis method was employed in Markov Chain Monte Carlo algorithm to estimate the model parameters by updating the parameters during sampling based on previous samples from the chain. From the results of running the MCMC algorithm 20 times, the mean of the information criteria of competing models is significantly different based on standard deviation and the result suggests that the model with continuous and jump variables can improve the model without jump. The best fit model is provided by RealGJR-CJ with the adoption of 1-minute RV data.

Crossref
https://doi.org/10.31764/jtam.v8i4.24294


## A. INTRODUCTION

Volatility in the financial market world has become a major concern for market players and investment observers. It reflects the fluctuation level or change in financial asset values during a certain period and is a key indicator of risk assessment when making investment decisions (Nugroho et al., 2018; Wang et al., 2022). Volatility is statistically the standard deviation of return (Sheraz & Nasir, 2021). Risk involves the possibility of losing money, whereas volatility describes how much and how quickly prices change. A higher volatility indicates that the asset value can vary drastically within a short period of time. An increase in price changes that also raises the probability of loss carries an increase in risk. In other words, the higher the volatility, the riskier the asset.

One of the popular and widely used models to understand and model volatility is the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced by Bollerslev (1986). This model is the main basis for analyzing financial asset volatility and allows volatility to change over time, depending on past returns and volatility. With the availability of intra-daily data, the GARCH model has been developed into the GARCH-X model by Engle (2002), by incorporating realized measures (computed from high-frequency data) as exogenous variables. This model was further developed by Hansen et al. (2012) into the Realized GARCH (RealGARCH) model by by adding a measurement equation (for exogenous variable) that relates the observed realized measure to latent volatility.

Meanwhile, Glosten et al. (1993) introduced the Glosten-Jagannathan-Runkle (GJR) model, a development of the GARCH model. Unlike the GARCH model, the GJR model allows volatility variance to have different responses to previous returns (Chen et al., 2019). Due to this capability, the GJR model becomes widely applied in analyzing financial asset volatility. As an example, Chen et al. (2019), Mostafa et al. (2021), and Nugroho et al. (2019, 2022) proved that the GJR model is superior to the GARCH model. Most recently, the GJR model was developed by Nugroho et al. (2024) into the Realized GJR (RealGJR) model. The model's development follows the RealGARCH model from Hansen et al. (2012). Both models have been empirically proven to have better data matching than their constituent models. A different development for the GJR model was carried out by Zhang & Lan (2014) by decomposing the exogenous variable $X$ in the GJR-X model (similar to GARCH-X) into continuous component $C$ and jump component $J$, subsequently called GJR-CJ model. This model has improved the model's ability to measure and predict future volatility in the financial market compared to the GJR-X model.

In this study, a new RealGJR model is developed by decomposing the exogenous variables into continuous and jump components, which is then called the RealGJR-CJ model. This model is more flexible and can be adapted to various types of financial data and different market characteristics. Additionally, the model is also capable of providing a clear interpretation of the mechanisms behind volatility formation, particularly the roles of continuous components and jumps that accommodate realistic financial markets. Therefore, this study contributes to extend the model of Zhang & Lan (2014) in the context of adding measurement equation and to extend the model of Nugroho et al. (2024) in the context of decomposing the exogenous variable into continuous and jump components. By combining the advantages of its constituent models, this study provides a deeper understanding of market volatility, providing a more powerful tool for risk management in the investment world. By analyzing the comparison between the GJR, GJR-X, GJR-CJ, RealGJR, and RealGJR-CJ models, this study provides new insights about a more suitable model for practical applications in the real world.

For the purpose of analysis, this study tests the models during the financial crisis period, causing the asset prices to jump sharply. Analyzing the stock market during a crisis period is a significant challenge for any open economy. Japan, as the country with the third-largest economy based on nominal GDP, has experienced significant impacts from the global financial crisis of 2008–2009 and the earthquake and tsunami on March 11, 2011, which caused the stock market to fluctuate more. Therefore, as an illustration, the observed data is the Tokyo Stock Price Index (known as TOPIX) of Japan for the period from January 2004 to December 2011 during trading days. TOPIX is a stock index that measures the performance of large companies listed on the Tokyo Stock Exchange. Since the aim of this study is to get an overview of volatility dynamics, the empirical study considers a simple Realized Volatiliy (RV) measure estimator, which is the square root of the sum of squared returns, with a reasonable choice of sampling frequency, i.e. 5-minute or 10-minute RV, following recent literature (see Floros et al. (2020); Gkillas et al. (2020)).

Based on the model application on real data, the model parameters are estimated using the Adaptive Random Walk Metropolis (ARWM) method in the Markov Chain Monte Carlo (MCMC) algorithm. The practical implemetation of this method can be very simple, efficient, and computationally fast. The method can be said to be simple because it is able to generate samples of model parameters from complex posterior probability distributions such as in the context of the Realized GJR model. Nugroho et al. (2024) proved the efficiency of the method in estimating RealGJR models in terms of autocorrelation time. Meanwhile, Nugroho (2018) showed that the method is much faster than the other two MCMC methods.

## B. METHODOLOGY

This study aims to describe the volatility characteristics of the financial time series by accounting for jump effects in intra-day data. A quantitative approach is taken through an empirical study by examining the volatility of stock indices using the GJR and RealGJR models. In addition, all calculations is implemented by an own Matlab code.

### 1. Volatility Model

A common thing to study in finance is the return on asset prices. Assuming that an asset is modeled as Geometric Brownian Motion (GBM), the return (defined as the profit or loss on an investment) is estimated using the difference in natural logarithms of the prices within successive periods. Under these assumptions, suppose $P_t$ and $P_{t-1}$ are the asset prices within successive periods, then the return at time period $t$ is (Sinha, 2021):

$$R_t = \ln P_t - \ln P_{t-1}. \tag{1}$$

The terms "log return" and "return" are often used interchangeably in the financial literature to refer to the same quantity. This study simply uses the term "return". The return distribution for GBM can be expressed as:

$$R_t \sim N(0, \sigma_t^2), \tag{2}$$

where $\sigma_t$ is the return volatility. The GJR volatility model is one of the asymmetric model types of GARCH. The asymmetric nature indicates that positive or negative values of past returns (with the same absolute value) have different effects on current volatility (Caporin & Costola, 2019). Meanwhile, the GJR-X(1,1) model was developed from the GJR(1,1) model by adding exogenous variable $X$ to the volatility dynamics equation:

$$\sigma_t^2 = \omega + \left(\alpha_1 + \alpha_2 I_{[R_{t-1}<0]}\right) R_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma X_{t-1}, \tag{3}$$

where $I_{[y]}$ is an indicator function that has the value of 1 when $y$ is true and 0 otherwise. Note that if $\alpha_2 = 0$, it means that positive and negative returns, with the same absolute value, have the same effect on volatility, and the model is being reduced back to the GARCH model. By decomposing the exogenous variable $X$ into continuous component $C$ and jump component $J$, a GJR-CJ(1,1) model was obtained, expressed as follows:

$$\sigma_t^2 = \omega + \left(\alpha_1 + \alpha_2 I_{[R_{t-1}<0]}\right) R_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma_1 C_{t-1} + \gamma_2 J_{t-1}\,, \tag{4}$$

The continuous component $C_t$ and jump component $J_t$ were calculated using the following formulas (Degiannakis et al., 2022):

$$C_t = I_{[Z_t > \emptyset_a]}(X_t - M_t), \tag{5}$$

$$J_t = I_{[Z_t \leq \emptyset_a]} X_t + I_{[Z_t > \emptyset_a]} M_t, \tag{6}$$

where $\emptyset_\alpha$ represents the quantile of the standard normal distribution function with confidence level $\alpha$, and $M_t$ represents the Median Realized Volatility. The $Z_t$ statistic is expressed by:

$$Z_t = \sqrt{n}\frac{1 - M_t X_t^{-1}}{\sqrt{\left(\left(\frac{\pi}{2}\right)^2 + \pi - 5\right)\text{maks}\{1, Q_t M_t^{-2}\}}}, \tag{7}$$

where $n$ represents the number of intra-daily observations and $Q_t$ represents the Median Realized Tri-power Quarticity. For the exogenous variable $X$ as a measure of Realized Volatility, the realized measures were calculated using the formula:

$$X_t^2 = \sum_{i=1}^{m} R_{t,i}^2, \tag{8}$$

$$MedRV_t = \frac{\pi}{6 - 4\sqrt{3} + \pi}\left(\frac{n}{n-2}\right) \times \sum_{i=2}^{m-1} Med\left(|R_{t,i-1}||R_{t,i}||R_{t,i+1}|\right)^2, \tag{9}$$

$$MedRTQ_t = \frac{3\pi n}{9\pi + 72 - 52\sqrt{3}}\left(\frac{n}{n-2}\right) \times \sum_{i=2}^{m-1} Med\left(|R_{t,i-1}||R_{t,i}||R_{t,i+1}|\right)^4, \tag{10}$$

where $R_{t,i}$ is the return on day $t$ for the $i$-th observation. Similar to the RealGARCH model proposed by Hansen et al. (2012), which expresses the exogenous variable $X$ as a normal process, the RealGJR(1,1) volatility equation is then obtained as follows (Nugroho et al., 2024):

$$\sigma_t^2 = \omega + \left(\alpha_1 + \alpha_2 I_{[R_{t-1}<0]}\right) R_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma X_{t-1}, \tag{11}$$

$$X_t = \xi + \varphi\sigma_t^2 + \tau_1 \frac{R_t}{\sigma_t} + \tau_2\left(\frac{R_t^2}{\sigma_t^2} - 1\right) + u_t, \quad u_t \sim N(0, s_u^2). \tag{12}$$

To ensure that $\sigma_t^2$ is positive, the model requires sufficient conditions that $\omega$, $\alpha_1$, $\alpha_1+\alpha_2$, $\beta$, and $\gamma$ are positive. Meanwhile, to ensure that the variance is stationary, the conditions are as follows (see Gerlach & Wang (2016)):

$$0<\alpha_1+0.5\alpha_2+\beta+\gamma\varphi<1 \text{ dan } \omega+\gamma\xi>0. \tag{13}$$

By decomposing the exogenous variable $X$ into continuous component $C$ and jump component $J$, this study proposes the RealGJR-CJ(1,1) model, expressed as follows:

$$\sigma_t^2=\omega+\left(\alpha_1+\alpha_2 I_{[R_{t-1}<0]}\right)R_{t-1}^2+\beta\sigma_{t-1}^2+\gamma_1 C_{t-1}+\gamma_2 J_{t-1}\,, \tag{14}$$

$$C_t=\xi+\varphi\sigma_t^2+\tau_1\frac{R_t}{\sigma_t}+\tau_2\left(\frac{R_t^2}{\sigma_t^2}-1\right)+u_t,\ \ u_t\sim N(0,s_u^2). \tag{15}$$

## 2. Estimation Method

The MCMC algorithm is a popular algorithm for dealing with simulations of varying distributions with high levels of complexity (Robert et al., 2018). The algorithm consists of two steps, namely the Markov chain generation and the Monte Carlo-based distribution properties approach. The Markov chain generation for parameter $\theta$ is based on the Bayes' rule:

$$f(\theta\,|\,\text{data})=L(\text{data}\,|\,\theta)\times p(\theta), \tag{16}$$

where $f$ represents the posterior distribution, $L$ represents the likelihood function, and $p$ represents the prior distribution. One of the Markov chain generation methods for a model parameter is ARWM, introduced by Atchade & Rosenthal (2005). The algorithm of this method can be seen in detail in (Nugroho et al., 2023, 2024). Consider the RealGJR-CJ(1,1) model again. The total log-likelihood function for the data of return $\{R_t\}_{t=1}^T$ with conditional parameters $\theta_1=(\omega,\alpha_1,\alpha_2,\beta,\gamma_1,\gamma_2)$ can be expressed in the form of:

$$\mathcal{L}(R_1,\dots,R_T\,|\theta_1)=-\frac{1}{2}\sum_{t=1}^T\left[\ln(2\pi\sigma_t^2)+\frac{R_t^2}{\sigma_t^2}\right]. \tag{17}$$

Meanwhile, the exogenous data $\{C_t\}_{t=1}^T$ with conditional parameters $\theta_2=(\xi,\varphi,\tau_1,\tau_2,\eta)$ has a total log-likelihood function as follows:

$$\mathcal{L}(C_1,\dots,C_T\,|\theta_2)=-\frac{1}{2}\sum_{t=1}^T\left[\ln(2\pi\eta^2)+\frac{\left(X_t-\xi-\varphi\sigma_t^2-\tau_1\frac{R_t}{\sigma_t}-\tau_2\left(\frac{R_t^2}{\sigma_t^2}-1\right)\right)^2}{\eta}\right]. \tag{18}$$

Thus, the total log-likelihood function of the RealGJR-CJ(1,1) model with parameter $\theta=(\theta_1,\theta_2)$ is expressed as follows:

$$\mathcal{L}(\text{data}|\theta)=\mathcal{L}(R_1,\dots,R_T\,|\theta_1)+\mathcal{L}(C_1,\dots,C_T\,|\theta_2). \tag{19}$$

To measure the efficiency of the estimation method, this study used Integrated Autocorrelation Time (IACT), defined as the number of MCMC iterations of the Markov chain required to obtain one independent sample. Smaller IACT values indicate that the method is more efficient and the convergence is faster, so the estimation is more accurate. See Nugroho et al. (2021) for the IACT value estimation procedure. Furthermore, to describe and summarize the uncertainty associated with the estimated parameters, it is important to find a confidence interval. This study used the Highest Posterior Density (HPD) interval, and the estimation procedure (Le et al. (2020)(Nugroho et al. (2023, 2024).

### 3. Model Selection Criteria

In selecting models involving data matching, this study used four criteria based on the log-likelihood, namely Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Adjusted BIC (ABIC), and Consistent AIC (CAIC). These four criteria are formulated as follows (Dziak et al., 2020):

$$AIC = -2\mathcal{L} + 2k, \tag{20}$$

$$CAIC = -2\mathcal{L} + k(1 + \ln T), \tag{21}$$

$$BIC = -2\mathcal{L} + k \ln T, \tag{22}$$

$$ABIC = -2\mathcal{L} + k \ln\left(\frac{T+2}{24}\right), \tag{23}$$

where $k$ represents the number of parameters. In a set of models being compared, the best model is indicated by the criteria that have the smallest values compared to the other models. In the case of comparing two models, when more criteria of a model have lower values, it generally concludes that the model provides better fit. Meanwhile, when the number of criteria favoring each model is the same, both models are said to be competitive.

## C. RESULTS AND DISCUSSION

### 1. Data Description

The observational data used in this study are secondary data, namely the intra-daily data of the Japanese Tokyo Stock Price Index (TOPIX) from January 2004 to December 2011. The data consists of daily returns and daily exogenous data for 1-minute and 5-minute time intervals. Specifically, Figure 1 displays the plot of absolute returns as well as continuous and jump components for 1-minute time intervals. The plot shows high (extreme) fluctuation movements in certain periods, where the continuous and jump components follow the return movements. Even though the Ljung–Box normality test indicates rejection of the normal distribution for the data of returns and continuous components, this study assumed a normal distribution for both data as a simple framework. In addition, the main objective of this study is not to propose the best distribution, but rather to investigate whether decomposing the realized measure into continuous and jump components is going to improve the existing GJR-X and RealGJR models in terms of data fit, as shown in Figure 1.

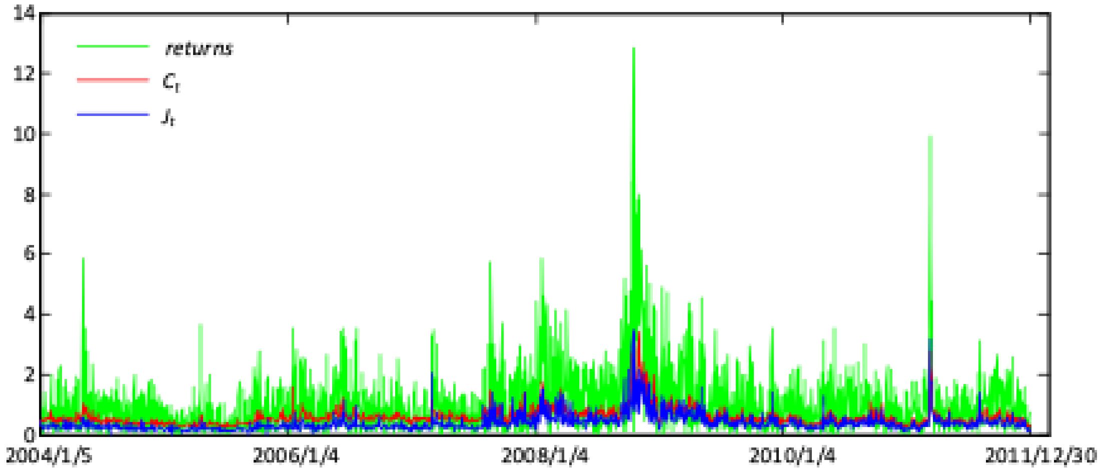


**Figure 1.** Plot of Return $|R_t|$, Continuous Component $C_t$, and Jump Component $J_t$ in Daily TOPIX Data from 2004 to 2011

## 2. Estimation Results

First, the efficiency of the ARWM method was assessed, as indicated by the convergence of the Markov chain. The simplest way to examine the convergence is visually through a trace plot. A trace plot is a time series plot that shows the results of parameter estimation (Markov chain) at each iteration. According to Roy (2020), if the Markov chain is stuck at several consecutive iterations, meaning that too many estimation proposals are rejected sequentially, then the trace plot shows very slow convergence.

This study ran the MCMC for 6000 iterations, of which the first 1000 iterations were not used in the Monte Carlo calculations to reduce non-stationarity caused by initial values. The initial values of the parameters were $\omega = 0.005$, $\alpha_1 = 0.05$, $\alpha_2 = 0.1$, $\beta = 0.45$, $\gamma_1 = 1.5$, $\gamma_2 = 0.05$, $\xi = 0.01$, $\varphi = 0.2$, $\tau_1 = 0.15$, $\tau_2 = 0.5$, and $\eta^2 = 0.2$. To complement the Bayesian method, the common prior distribution for the model parameters was the normal distribution. The mean 0 and variance 1000 were used to provide estimated values over a wide range.

Figure 2 presents the trace plot of specific estimated values for key parameters of the RealGJR-CJ model with the intra-daily data application with 1-minute time intervals. Visually, the trace plot shows that the estimated values of each parameter can be considered to be convergent or stationary. The estimated values fluctuate around the mean (red line). In other words, it indicates that the ARWM method is efficient in estimating the model. Fast convergence occurs in the estimation of all parameters, where the trace plot shows a dense visual during the iteration. The only exception is for the parameter $\gamma_1$, where there is a movement that is not dense around the mean during the iteration. This result is similar to Nugroho et al. (2024) that found a slow convergence for the parameters of the exogenous components of RV. This slow convergence is caused by the very low acceptance rate of sample proposals $\gamma_1^*$ (generated by ARWM) in the MCMC algorithm. The parameter $\gamma_1$ may be highly correlated, so the ARWM method slow to explore the entire posterior distribution, as shown in Figure 2.

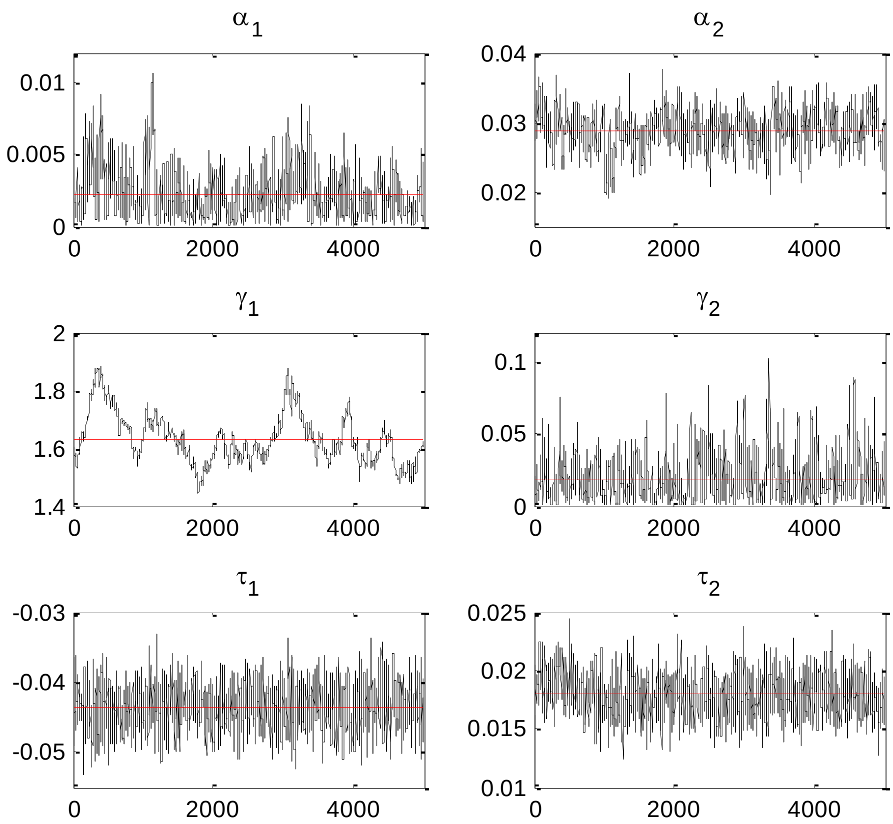


**Figure 2.** Trace Plot of Estimations for Key Parameters from the RealGJR-CJ Model that Applies High-frequency Data with 1-minute Time Intervals.

Next, Figure 2 presents several summary statistics for key parameters, such as mean, Standard Deviation (SD), Lower Bound (LB), and Upper Bound (UB) values of the 95% HPD intervals, of the GJR-CJ and RealGJR-CJ models. In all cases, the LB and UB values for the parameter $\alpha_2$ are positive, meaning that 95% of the HPD intervals do not contain the value of 0. It indicates that the asymmetric effect between returns and volatility is significantly positive, meaning that negative returns cause greater volatility compared to positive returns (of the same value). Similarly, 95% of the HPD intervals for parameters $\tau_1$ do not contain the value of 0. It indicates that there is an asymmetric effect between returns and continuous components. In the case of intra-daily data application with 1-minute intervals, the estimated values of $\tau_1$ are negative, meaning that negative returns cause a larger continuous variable size. However, the estimated values of $\tau_1$ in the intra-daily data application with 5-minute time intervals are positive. This means that the larger effect for the continuous variable size is caused by positive returns. Specifically, the effects caused by negative returns are called leverage (according to Black (1976) in Caporin & Costola (2019)).

One important economic issue in daily stock return data is the degree to which conditional volatility is persistent or permanent. In essence, volatility persistence is the propensity for extended periods of high or low volatility to continue occurring. This implies that a market is likely to be volatile in the future if it goes through a period of high volatility, and vice versa. The persistence phenomenon in the GJR-CJ and RealGJR-CJ model is measured by the coefficient, respectively,

$$\alpha_1 + 0.5\alpha_2 + \beta \tag{24}$$

and

$$\alpha_1 + 0.5\alpha_2 + \beta + \gamma_1\varphi. \tag{25}$$

In the case of TOPIX data, the GJR-CJ estimates reveal persistence coefficients of 0.8572 and 0.8325 for 1-minute and 5-minute, respectively. Meanwhile, the estimated persistence coefficient in the RealGJR-CJ models is 0.8633 dan 0.8187 for 1-minute and 5-minute, respectively. The values close to 1 indicate a quite strong volatility persistence of the TOPIX index, although this is lower than the finding of Nugroho & Morimoto (2019) in the context of stochastic volatility. The implication is that investors must be prepared to face the possibility of high volatility periods, especially if the market has recently experienced such times. Trading strategies that rely on volatility, such as options, may be highly sensitive to the persistence of volatility in the TOPIX market. High persistence suggests that historical volatility data may serve as a good predictor for future volatility. In particular, 1-minute TOPIX data has higher persistence compared to 5-minute TOPIX data. This means that volatility shocks on 1-minute data tend to have a longer impact. This suggests investors and traders who use 1-minute data to be more cautious about volatility risk.

Considering the continuous and jump coefficients, for instance, for the 1-minute data estimated by the RealGJR-CJ model, a unit increase in the RV comes from the continuous component implies an average increase in volatility on the following day of 1.6838 for days when there was no jump on the previous day. Meanwhile, for days in which part of the RV comes from the jump component, the increase in volatility on the following day is only 0.0206 times the jump component. In other words, if volatility is entirely caused by jumps, then this only leads to a slight increase for the following day's volatility. This indicates that although there are events that cause jumps in volatility, the impact is relatively short-lived and quickly subsides within a 1-minute timeframe. Similary result is found for the 5-minute-based GJR-CJ. In this case, market participants for TOPIX should be more aware of the rapid and frequent changes in volatility, but the impact tends to be temporary. In contrast, the 1-minute-based GJR-CJ and 5-minute-based RealGJR-CJ models result the large relative increase about 0.4. In this case, market participants for TOPIX need to increase their awareness of events that could potentially trigger jumps in volatility, as shown in Table 1.

**Table 1.** Estimation Results for Key Parameters

| **Statistics** | **Parameter** | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\alpha_1$ | $\alpha_2$ | $\beta$ | $\gamma_1$ | $\gamma_2$ | $\tau_1$ | $\tau_2$ |
| GJR-CJ with data is sampled at a 1-minute frequency | | | | | | | |
| Mean | 0.0169 | 0.2369 | 0.7219 | 0.1278 | 0.4564 | - | - |
| SD | 0.0132 | 0.0386 | 0.0376 | 0.0814 | 0.1686 | - | - |
| LB | 0.0000 | 0.1650 | 0.6440 | 0.000 | 0.1151 | - | - |
| UB | 0.0411 | 0.3097 | 0.7969 | 0.2821 | 0.7992 | - | - |
| GJR-CJ with data is sampled at a 5-minute frequency | | | | | | | |
| Mean | 0.0127 | 0.2460 | 0.6968 | 0.3733 | 0.0838 | - | - |
| SD | 0.0108 | 0.0392 | 0.0470 | 0.0942 | 0.0756 | - | - |
| LB | 0.0000 | 0.1721 | 0.6043 | 0.1966 | 0.0001 | - | - |
| UB | 0.0358 | 0.3185 | 0.8010 | 0.5512 | 0.2362 | - | - |
| RealGJR-CJ with data is sampled at a 1-minute frequency | | | | | | | |

| Statistics | Parameter | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\alpha_1$ | $\alpha_2$ | $\beta$ | $\gamma_1$ | $\gamma_2$ | $\tau_1$ | $\tau_2$ |
| Mean | 0.0027 | 0.0291 | 0.4087 | 1.6838 | 0.0206 | −0.0438 | 0.0184 |
| SD | 0.0020 | 0.0032 | 0.0257 | 0.0807 | 0.0180 | 0.0030 | 0.0018 |
| LB | 0.0000 | 0.0235 | 0.3548 | 1.5073 | 0.0000 | −0.0498 | 0.0150 |
| UB | 0.0064 | 0.0358 | 0.4544 | 1.8210 | 0.0564 | −0.0378 | 0.0218 |
| RealGJR-CJ with data is sampled at a 5-minute frequency | | | | | | | |
| Rata-rata | 0.0007 | 0.0346 | 0.4580 | 1.0976 | 0.3836 | −0.0583 | 0.0533 |
| SB | 0.0007 | 0.0034 | 0.0240 | 0.0631 | 0.0780 | 0.0048 | 0.0032 |
| BB | 0.0000 | 0.0278 | 0.4124 | 0.9775 | 0.2432 | −0.0683 | 0.0477 |
| BA | 0.0019 | 0.0415 | 0.5037 | 1.2154 | 0.5442 | −0.0498 | 0.0601 |

## 3. Model Selection

Table 2 presents the log-likelihood estimates (means and standard deviations in brackets) along with the four criteria for the GJR-X, GJR-CJ, RealGJR, and RealGJR-CJ models for three intra-daily cases. For each case the value of the log-likelihood is evaluated by averaging 20 runs of the MCMC. The standard deviations of the log-likelihood and information criteria estimates are smaller than 0.9 in all cases. The relatively small standard errors in all models indicate that the estimation results from 20 runs of the MCMC are quite stable. This means that the models are quite consistent in providing similar results. It can be seen that, in each data case, the four models have significantly different information criteria. This is convincing because there is no overlap of the confidence intervals between the criterion values of the different models.

The results of the four criteria in all cases show that the models with continuous and jump components outperform the models without decomposition, as indicated by the smaller criteria values. These results are consistent with the results of by Zhang & Lan (2014) in the context of the GJR and EGARCH models. In addition, in all data cases, the models that treat the exogenous component as a dynamic process have a better fit. These results demonstrate the ability of models with more complex structures to capture more complex data features. In the case of intra-daily data application, all four criteria support the intra-daily data with 1-minute time intervals to provide a better model. Therefore, overall, the RealGJR-CJ model that applies 1-minute RV data provides the best fit model for the TOPIX data. As an implication, the application of RealGJR-CJ model with 1-minute intra-day data is more reasonable in measuring volatility in financial practices such as financial risk measures, financial contract pricing, and asset allocation, as shown in Table 2.

**Tabel 2.** Average log-likelihood and criteria values with standard deviations (in brackets)

| Model | LL | AIC | ABIC | BIC | CAIC |
|---|---|---|---|---|---|
| Data is sampled at a 1-minute frequency | | | | | |
| GJR-X | −3147.7<br>(0.23) | 6305.5<br>(0.46) | 6322.4<br>(0.46) | 6333.2<br>(0.46) | 6338.2<br>(0.46) |
| GJR-CJ | −3136.9<br>(0.22) | 6285.8<br>(0.44) | 6306.3<br>(0.44) | 6319.3<br>(0.44) | 6325.3<br>(0.44) |
| RealGJR | −2499.2<br>(0.22) | 5018,5<br>(0.43) | 5052.5<br>(0.43) | 5074.3<br>(0.43) | 5084.3<br>(0.43) |
| RealGJR-CJ | −1908.3<br>(0.29) | 3838.7<br>(0.58) | 3876.1<br>(0.58) | 3900.1<br>(0.58) | 3911.1<br>(0.58) |
| Data is sampled at a 5-minute frequency | | | | | |

| Model | LL | AIC | ABIC | BIC | CAIC |
|---|---|---|---|---|---|
| GJR-X | −3144.0<br>(0.21) | 6298.1<br>(0.41) | 6315.1<br>(0.41) | 6326.0<br>(0.41) | 6331.0<br>(0.41) |
| GJR-CJ | −3129.4<br>(0.14) | 6270.9<br>(0.27) | 6291.3<br>(0.27) | 6304.3<br>(0.27) | 6310.3<br>(0.27) |
| RealGJR | −2904.3<br>(0.42) | 5828.6<br>(0.85) | 5862.7<br>(0.85) | 5884.4<br>(0.85) | 5894.4<br>(0.85) |
| RealGJR-CJ | −2792.2<br>(0.26) | 5606.4<br>(0.51) | 5643.9<br>(0.51) | 5667.8<br>(0.51) | 5678.8<br>(0.51) |

Finally, notice that since the objective of this study is not to propose the best volatility model, but rather to investigate whether decomposing the exogenous variable RV in the conditional volatility process into continuous and jump components will improve on the existing GJR models in terms of data fit, inclusion of a non-Normal distribution is not necessary to answer the question. However, in addition, the assumption of non-normality in models can certainly have a significant impact on the estimation results and validity of the model. A misspecified distribution can lead to a poor fit, even if the volatility structure is correct. If the true distribution of the return errors is known or can be estimated, an alternative distribution can be used in the model to improve the accuracy of the fit and forecast. However, this study will not provide any empirical results of any non-Normal assumption because it is believed that the purpose of this study can be carried out without the empirical results of non-Normal assumption for return errors.

## D. CONCLUSION AND SUGGESTION

This study proposes the RealGJR-CJ model constructed from the Realized GJR model by decomposing the exogenous components into continuous and jump components. As an empirical study, the model was applied to the TOPIX stock index data from 2004 to 2011 with 1 and 5-minute time intervals. The ARWM method was applied in the MCMC algorithm to estimate the model. The results show the efficiency of the ARWM method visually through a trace plot. Based on the empirical results from the comparison between the GJR-X, GJR-CJ, and RealGJR models, the four selection criteria indicate the superiority of models with continuous and jump components. In addition, in some cases, the jump component coefficient show a significant effect on volatility due to jumps in the previous time period. Therefore, this study recommends the decomposition of the exogenous data into continuous and jump components. In particular, the use of intra-daily data with 1-minute time intervals is more recommended than intra-daily data with 5-minute time intervals.

It would be interesting to investigate whether the empirical performance of the proposed model can be improved by incorporating recent (return) information in the volatility process, such as the Real-Time GARCH model which has better fitting and forecasting than GARCH models. In addition, decomposition in recent models, such as the GARCH@CARR model which has better forecasting than the Realized GARCH model, would affirm the power of continuous and jump components. These suggestions can be used for further research.

**ACKNOWLEDGEMENT**

This publication resulted (in part) from research supported by Universitas Kristen Satya Wacana under the Fundamental Research Grant Scheme with Contract Number 032/SPK-PF/RIK/07/2024.

**REFERENCES**

Atchade, Y. F., & Rosenthal, J. S. (2005). On adaptive Markov chain Monte Carlo algorithms. *Bernoulli*, *11*(5), 815–828. https://doi.org/10.3150/bj/1130077595

Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the 1976 Meetings of the Business and Economic Statistics*, 177–181. https://www.scirp.org/reference/referencespapers?referenceid=2030459

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Caporin, M., & Costola, M. (2019). Asymmetry and leverage in GARCH models: a News Impact Curve perspective. *Applied Economics*, *51*(3), 3345–3364. https://doi.org/10.1080/00036846.2019.1578853

Chen, H., Zhang, J., Tao, Y., & Tan, F. (2019). Asymmetric GARCH type models for asymmetric volatility characteristics analysis and wind power forecasting. *Protection and Control of Modern Power Systems*, *4*(1), 29. https://doi.org/10.1186/s41601-019-0146-0

Degiannakis, S., Filis, G., Klein, T., & Walther, T. (2022). Forecasting realized volatility of agricultural commodities. *International Journal of Forecasting*, *38*(1), 74–96. https://doi.org/10.1016/J.IJFORECAST.2019.08.011

Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in Bioinformatics*, *21*(2), 553–565. https://doi.org/10.1093/bib/bbz016

Engle, R. (2002). New Frontiers for ARCH Models. *Journal of Applied Econometrics*, *17*(5), 425–446. https://doi.org/10.1002/jae.683

Floros, C., Gkillas, K., Konstantatos, C., & Tsagkanos, A. (2020). Realized measures to explain volatility changes over time. *Journal of Risk and Financial Management*, *13*(6), 125. https://doi.org/10.3390/jrfm13060125

Gerlach, R., & Wang, C. (2016). Forecasting risk via realized GARCH, incorporating the realized range. *Quantitative Finance*. https://doi.org/10.1080/14697688.2015.1079641

Gkillas, K., Gupta, R., & Pierdzioch, C. (2020). Forecasting realized gold volatility: Is there a role of geopolitical risks? *Finance Research Letters*, *35*, 1–6. https://doi.org/10.1016/j.frl.2019.08.028

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

Hansen, P. R., Huang, Z., & Shek, H. H. (2012). Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility. *Journal of Applied Econometrics*, *27*(6), 877–906. https://doi.org/10.1002/jae.1234

Le, H., Pham, U., Nguyen, P., & Pham, T. B. (2020). Improvement on Monte Carlo estimation of HPD intervals. *Communications in Statistics - Simulation and Computation*, *49*(9), 2164–2180. https://doi.org/10.1080/03610918.2018.1513141

Mostafa, F., Saha, P., Islam, M. R., & Nguyen, N. (2021). GJR-GARCH Volatility Modeling under NIG and ANN for Predicting Top Cryptocurrencies. *Journal of Risk and Financial Management*, *14*(9), 421. https://doi.org/10.3390/jrfm14090421

Nugroho, D. B. (2018). Comparative analysis of three MCMC methods for estimating GARCH models. *IOP Conference Series: Materials Science and Engineering*, *403*(1), 012061. https://doi.org/10.1088/1757-899X/403/1/012061

Nugroho, D. B., Kurniawati, D., Panjaitan, L. P., Kholil, Z., Susanto, B., & Sasongko, L. R. (2019). Empirical Performance of GARCH, GARCH-M, GJR-GARCH and log-GARCH Models for Returns Volatility. *Journal of Physics: Conference Series*, *1307*(1), 012003. https://doi.org/10.1088/1742-6596/1307/1/012003

Nugroho, D. B., Mahatma, T., & Pratomo, Y. (2021). Applying the non-linear transformation families to the lagged-variance of EGARCH and GJR models. *IAENG International Journal of Applied Mathematics*, *51*(4), 908–919. https://www.iaeng.org/IJAM/issues_v51/issue_4/IJAM_51_4_12.pdf

Nugroho, D. B., & Morimoto, T. (2019). Incorporating realized quarticity into a realized stochastic volatility model. *Asia-Pacific Financial Markets*, *26*(4), 495–528. https://doi.org/10.1007/s10690-019-09276-2

Nugroho, D. B., Panjaitan, L. P., Kurniawati, D., Kholil, Z., Susanto, B., & Sasongko, L. R. (2022). GRG non-linear and ARWM methods for estimating the GARCH-M, GJR, and log-GARCH models. *Jurnal Teori Dan Aplikasi Matematika*, *6*(2), 448–460. https://doi.org/10.31764/jtam.v6i2.7694

Nugroho, D. B., Susanto, B., & Rosely, M. M. M. (2018). Penggunaan MS Excel untuk estimasi model GARCH(1,1). *Jurnal Matematika Integratif*, *14*(2), 71–83. https://doi.org/10.24198/jmi.v14.n2.17680.71-82

Nugroho, D. B., Wibowo, H., & Saragih, A. (2024). Modeling daily return volatility through GJR(1, 1) model and realized volatility measure. *Thailand Statistician*, *22*(1), 50–62. https://ph02.tci-thaijo.org/index.php/thaistat/article/view/252221

Nugroho, D. B., Wicaksono, B. A. A., & Larwuy, L. (2023). GARCH-X(1,1) model allowing a non-linear function of the variance to follow an AR(1) process. *Communications for Statistical Applications and Methods*, *30*(2), 163–178. https://doi.org/10.29220/CSAM.2023.30.2.163

Robert, C. P., Elvira, V., Tawn, N., & Wu, C. (2018). Accelerating MCMC algorithms. *Wiley Interdisciplinary Reviews: Computational Statistics*, *10*(5), 1–14. https://doi.org/10.1002/wics.1435

Roy, V. (2020). Convergence diagnostics for Markov chain Monte Carlo. *Annual Review of Statistics and Its Application*, *7*, 387–412. https://doi.org/10.1146/annurev-statistics-031219-041300

Sheraz, M., & Nasir, I. (2021). Information-theoretic measures and modeling stock market volatility: A comparative approach. *Risks*, *9*(5), 89. https://doi.org/https://doi.org/10.3390/risks9050089

Sinha, A. K. (2021). The reliability of geometric Brownian motion forecasts of S&P500 index values. *Journal of Forecasting*, *40*(8), 1444–1462. https://doi.org/10.1002/for.2775

Wang, Y., Xiang, Y., Lei, X., & Zhou, Y. (2022). Volatility analysis based on GARCH-type models: Evidence from the Chinese stock market. *Economic Research-Ekonomska Istrazivanja* , *35*(1), 2530–2554. https://doi.org/10.1080/1331677X.2021.1967771

Zhang, H., & Lan, Q. (2014). GARCH-type model with continuous and jump variation for stock volatility and its empirical study in China. *Mathematical Problems in Engineering*, *2014*. https://doi.org/10.1155/2014/386721