International Journal on Advanced Science Engineering Information Technology

Vol.14 (2024) No. 4
ISSN: 2088-5334

Analyzing Composite Stock Price Index Volatility in Response to Changes in Data Structure Using Bayesian Markov-Switching GARCH

Ermanely [a], Dodi Devianto [a,*], Ferra Yanuar [a], Rahmat Hidayat [b]

[a] *Department of Mathematics and Data Science, Universitas Andalas, Padang, Indonesia*
[b] *Department of Information Technology, Politeknik Negeri Padang, Padang, Indonesia*
*Corresponding author: *ddevianto@sci.unand.ac.id*

Abstract—**Rapid economic growth was observed to have motivated economic actors and investors to maximize business performance using different methods. An example of these are stocks or shares securities showing that a person or entity has a stake in a company. Investors are primarily drawn to stocks due to their potential for substantial profits closely tied to the volatility in the stock market. The prevalence of this volatility has long been addressed using the ARCH/GARCH model, but it is not ideal for datasets experiencing structural changes. Therefore, a model was developed using the Bayesian Markov-switching GARCH approach to effectively capture the heteroscedastic component and structural changes in data and mitigate certain limitations, especially those associated with a small sample size. This study adopted the composite stock price index (CSI) data from March 2020 to April 2021 to model and understand the volatility. The results showed that the Bayesian Markov-switching GARCH model with the most negligible variance provided the best fit. It was also discovered that a more minor error variance corresponded to lower data volatility. Moreover, the concept of value at risk was used to assess the investment risk based on the criterion that a decrease in the CSI investments led to a reduction in the level of risk faced by the investors.**

Keywords— **Volatility; Bayesian Markov-switching GARCH; small sample size; CSI.**

Manuscript received 10 Dec. 2023; revised 25 Mar. 2024; accepted 10 May 2024. Date of publication 31 Aug. 2024.

I. INTRODUCTION

The composite stock price index (CSI) is typically used as a benchmark in the Indonesia Stock Exchange (IDX) to measure the profitability of market movements. The weekly CSI data usually presents the fluctuations and changes in the price of stock, and this concept is known as volatility. Therefore, volatility modeling has been discovered to be essential to risk management, portfolio management, and price equality, and this is necessary because high volatility values represent high asset risk.

The standard model usually used in monitoring stock price data, specifically when the data has constant variance or homoscedasticity, is the autoregressive integrated moving average (ARIMA). However, when the variance of the data is not continuous, the suitable choice is the generalized autoregressive conditional heteroscedasticity (GARCH). This model also becomes less effective when dealing with structural changes in the data, leading to adopting the Markov-switching model [1]. Therefore, a more comprehensive solution is proposed in the Markov-switching generalized autoregressive conditional heteroscedasticity (MS-GARCH) model to tackle structural changes and heteroscedasticity. The model can also explain the high persistence and poor forecasting performance issues associated with the single-regime GARCH [1]-[4].

The development of the probabilistic and estimation properties of the MS-GARCH model was observed to have proved to be a better fit than the conventional system. A previous study showed that GARCH was the better technique to study volatility even with the soft computing approach [5]-[6]. Moreover, the MS-GARCH was reported to be a better fit for implied volatility market [7]-[8], inflation determination model [9], oil and natural gas trading [10], volatility intraday between exchange rate, gold and crude oil [11], and dynamics volatility of cryptocurrencies and bitcoin [12]-[15]. Besides, the time-varying probability transition of volatility is also applicable using the MS-GARCH model approach [16]-[17].

Furthermore, the GARCH-MIDAS model with a regime transition was utilized to investigate the connection between oil price volatility and macroeconomic fundamentals [18], forecast renewable energy stock [19], and stock market volatility [20].

The volatility fluctuation and the changes in data structure or the effect of asymmetric trends make the fit of the variance model less accurate, so it is necessary to adjust the model to volatility movements. The exponential GARCH model was also more precise in predicting volatility [21]. Moreover, the MS-GARCH models, by using the Bayesian approach, have been used effectively in non-parametric panels [22] and in modeling agricultural commodities [23]. Moreover, the two-regime MS GARCH model was found to have outperformed the single-regime GARCH and also the best model of the Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) in analyzing housing returns and asymmetric volatility [24]. Several studies also concluded that the two-regime MS GARCH model delivered more accurate volatility estimates than the single-regime GARCH model [12]. Furthermore, a time series model was proposed to capture long memory [25], where the results showed the Markov-switching long memory model performed better than the other GARCH models [26].

The study on Markov-switching GARCH also relies on stability analysis of semi-Markov-switching stochastic delay systems. [27], Markov-switching Poisson [28], and the zero-drift GARCH model [29]. This research has also been adapted in many fields since the MS-GARCH package was introduced for the Markov-switching model [30], but a special approach was required to form a model with a Bayesian approach to overcome the change of data structure, asymmetry problem, and small data sample size. Besides, the non-Gaussian assumption residual model has to be considered by using the infinitely divisible distribution, such as exponential or gamma distribution with Bayesian approach [31]-[32]. These infinitely divisible distributions have been applied in integer-valued GARCH models and option pricing [33]-[34]. Therefore, this critical research focused on using the MS-GARCH to CSI data in order to provide information on short-time volatility based on the Bayesian approach.

II. Material and Method

This study was conducted using the weekly CSI data obtained through the website http://finance.yahoo.com. A total of 60 data from March 2, 2020, to April 19, 2021, were retrieved for this purpose. Moreover, the GARCH, MS-GARCH, hybrid MS-GARCH and Bayesian models were discussed and applied to form the combined Bayesian MS-GARCH model in the following methods of model building.

A. GARCH Model

The GARCH model is normally used to simulate the volatility of data on stock movement with the parameters calculated through the maximal likelihood estimation (MLE) approach. The conditional variance equation of the model is stated as follows:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \qquad (1)$$

This is based on the conditions $\varepsilon_t^2 \sim N(0, \sigma_t^2)$ and $\varepsilon_t = \sigma_t e_t$. Some parameters estimated in the GARCH model are $\omega, \alpha_i,$ and β_j for $i = 1, 2, \ldots, p$ and $j = 1, 2, \ldots, q$. The parameter σ_t^2 is the conditional variance, and ε_t is an error of the ARIMA model.

Some of the steps to follow in forming a GARCH model:

- Check whether the data used is stationary to the mean using the augmented Dickey-Fuller (ADF) test or to the variance of Box-Cox transformation.
- Identify the ARIMA models based on ACF and PACF plots.
- Estimate significant parameters in the ARIMA model with a significant level of $\alpha = 0.05$.
- Select the ideal ARIMA model using the AIC criteria.

The next step was to conduct a diagnostic check using the following tests:

1) The Autocorrelation Test: The correlogram of residuals can be used to determine the autocorrelation based on the criterion that the residuals have autocorrelation when the correlogram displays a substantial ACF and PACF plot at the early lags and vice versa. It can also be assessed using the Ljung-Box test with the following hypothesis:

H_0: $\rho_1 = \rho_2 = \cdots = \rho_m = 0$ there is no autocorrelation in the residuals).

H_1: $\exists \quad \rho_i \neq 0$, untuk $\quad i = 1, 2, \ldots, m$ (there is an autocorrelation in the residuals).

The equation for the Ljung-Box test statistic can also be written as follows:

$$Q = T(T+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{T-k}, \qquad (2)$$

where, T is the number of log return data, k is the number of lags, $\hat{\rho}_k^2$ is the autocorrelation value up to lag-k, and m is the maximum lag tested.

2) Homoscedasticity Test: A model is normally classified as good when it is homoscedastic, or there is no heteroscedasticity problem. This variable can be determined through the correlogram of squared residuals based on the criterion that the residual variance is not constant when the ACF and PACF plots are significant at the first lags, and vice versa. It can also be evaluated using the Lagrange Multiplier (LM) test by regressing the squared residuals through the following model:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1} + \cdots + \alpha_k \varepsilon_{t-k}. \qquad (3)$$

The hypothesis is:

H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ (there is no heteroscedasticity).

H_1: $\exists \alpha_i \neq 0$, for $i = 1, 2, \ldots, m$ (there is a heteroscedasticity).

The statistic for this test is

$$LM = nR^2, \qquad (4)$$

where the number of observations is n and the coefficient of determination is R^2.

3) Residual Normality Test: The residual normality test is normally used to determine the distribution status of the residuals collected. This can be achieved through the Shapiro-Wilk test based on the following hypotheses:

H_0: The data has a normal distribution.

H_1: The data is not normal.

A p-value of $\geq \alpha$=0.05 shows that H_0 is accepted, and this means the data are normally distributed while the p-value <

α=0.05 indicates H_0 is rejected, and the data are not normally distributed. The equation for the *Shapiro-Wilk* test statistics can be written as follows:

$$T_3 = \frac{1}{D}\left(\sum_{i=1}^{n} \alpha_i (x_{n-i+1} - x_i)^2\right), \quad (5)$$

where $D = \sum_{i=1}^{n}(X_i - \bar{X})^2$, α_i is the Shapiro-Wilk test coefficient, X_{n-i+1} is data to $n-i+1$, X_i is data to *i*, and $\bar{X}$ is the mean of the data.

4) Structural Change Test: Structural changes are usually evaluated to determine any variation in the pattern of the data. The existence of any structural changes within a certain period normally indicates the data is less stable. This can be achieved using the following linear regression model:

$$y_i = x_i\beta_i + u_i. \text{ for i=1,2,...,n}, \quad (6)$$

The hypothesis for the *F* test statistic is as follows:
H_0: $\beta_0 = \beta_i$, i = 1,2, …, *n* (no structural change).
H_1: $\beta_0 \neq \beta_i$= i, i = 1, 2,…, *n* (there is a change in structure).

The equation for the *F* test statistic is written as follows:

$$F_i = \frac{\hat{u}^T\hat{u} - \hat{u}(i)^T\hat{u}(i)}{\hat{u}(i)^T\hat{u}(i)/(n-2k)}, \quad (7)$$

where, k represents several parameters, $\hat{u}(i)$ is the residuals of each fault, and *i* is the fault/breakpoint. The decision-making criteria is that the hypothesis be rejected when *p*-value <0.05, and this means there is a change in the data structure.

B. Markov-Switching GARCH Model

The Markov Regime Switching GARCH model was proposed by Gray (1996) to explain the structural changes and shifts in the volatility of time series data using simpler parameters. This model is presented as follows:

$$r_t = \mu_{s_t} + \varepsilon_t, \text{ for } \varepsilon_t = u_t\sqrt{\sigma_t^2}, \quad (8)$$

$$\sigma_t^2 = \omega_{s_t} + \alpha_{s_t}\varepsilon_{t-1}^2 + \beta_{s_t}\sigma_{t-1}^2, \quad (9)$$

where, s_t is the state variable with value 0 or 1, μ_{s_t} and σ_t^2 are the mean and conditional variance, α_{s_t} and β_{s_t} are the parameters to be estimated for each regime in the Markov-switching model, ε_{t-1}^2 is the squared residual, and σ_{t-1}^2 is the conditional variance of the previous period.

The maximum likelihood function is normally used to estimate the parameters of the MS GARCH model through the algorithm. Moreover, the conditional variance of the GARCH model is dependent on the past values of the state variables and this presents a new challenge in the calculation process. For example, *K* states and *T* samples are required to be considered as many as K^T cases to determine the likelihood function but this can be overcome easily through the Bayesian technique.

C. The Proposed Bayesian MS-GARCH Model

The Bayesian MS-GARCH model is a hybrid between the MS-GARCH normally used to overcome heteroscedasticity problems while capturing the changes in volatility structure and the parameter estimation through a Bayesian approach. It is important to state that the Bayesian estimation approach introduces a concept known as prior distribution which requires obtaining information about the estimated parameters. Moreover, this concept can be combined with the likelihood function to determine the posterior distribution in estimating the parameters of the MS-GARCH model.

The model used for the two regimes is presented in Equations (8) and (9), suppose the vectors $r_t = (r_1, r_2, \dots, r_t)$, $s_t = (s_1, s_2, \dots, s_t)$, and the model parameters consist of $\eta = (\eta_{11}, \eta_{21}, \eta_{12}, \eta_{22})'$, $\mu = (\mu_1, \mu_2)'$ and $\theta = (\theta'_1, \theta'_2)'$, where $\theta_k = (\omega_k, \alpha_k, \beta_k)'$ for *k* = 1, 2.

The joint probability density function of r_t and s_t is presented as follows:

$$f(r_t, s_t|\varphi_{t-1}) = f(s_t = j, \varphi_{t-1})P(\varphi_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} exp\left(-\frac{(y_t - \mu_{s_t})^2}{2\sigma_t^2}\right).\eta_{jt} \quad (10)$$

The likelihood function is:

$$L(\theta) \propto \prod_{t=1}^{T} \sigma_t^{-1} exp\left(-\frac{(y_t-\mu_{s_t})^2}{2\sigma_t^2}\right)\eta_{jt}, \quad (11)$$

The determination of the marginal posterior distribution for this model was not easy due to the need for an integral process with a reasonably high dimension. Therefore, the *Gibbs sampler algorithm* was introduced to overcome the problem.

D. The Best Model selection

Some of the criteria used in selecting the best model are stated as follows:

1) Akaike Information Criterion (AIC): AIC was first introduced by Akaike to identify a model from a dataset using the following equation:

$$AIC = \log\hat{\sigma}^2 + \frac{2k}{n}, \quad (12)$$

where, $\log\hat{\sigma}^2$ is the likelihood measure, *k* is a number of parameters, and *n* represents several observations.

2) Bayesian Information Criterion (BIC): BIC is a model selection method with the Penalized Maximum Likelihood approach. It was first introduced by Schwartz to be used in selecting a model using the following equation:

$$BIC = \log\hat{\sigma}^2 + \frac{k\log(n)}{n}, \quad (13)$$

The model was considered to be good for selection when the AIC and BIC values were smaller.

E. Value at Risk by Using Volatility Model

Value at Risk (VaR) is a statistical risk evaluation technique normally used to calculate the potential maximum loss in a portfolio at a specific level of confidence. This is achieved through the adoption of some statistical methods in its application as a risk measurement tool. The following formula can be used to calculate VaR for a period of *T* days:

$$VaR = -V_0 * Z_\alpha * \hat{\sigma}\sqrt{T}, \quad (14)$$

VaR is used in risk management to confidently determine the greatest possible loss in an investment or portfolio and recommend the right steps to lower the risk or safeguard the capital.

III. Results and Discussion

This research has developed a new model of MS-GARCH model by using Bayesian inference approach. This part was also generally used to analyze the CSI volatility modeling using the time series model of GARCH, MS-GARCH, and Bayesian MS-GARCH methods.

A. Data Identification and ARIMA Model

The weekly CSI data were plotted and visualized to demonstrate the trend and seasonality of the pattern and stationarity. The focus was on the period from March 2, 2020, to April 19, 2021 as presented in the following Fig. 1.

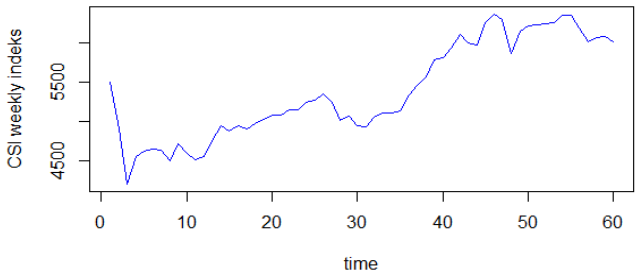


Fig. 1 Plot of CSI weekly data from March 2, 2020, to April 19, 2021.

Fig. 1 shows that the time series plot of CSI weekly data has both upward and downward trends at any given time. The data was observed not to change over time at a fixed variance or around a fixed mean during the observation period. This showed that the CSI weekly data were not stationary in terms of mean and variance.

The next step was to convert the data into a return format using the natural logarithm value of the simple net return. This was achieved using the following formula:

$$r_t = ln\left(\frac{P_t}{P_{t-1}}\right) \qquad (15)$$

The data plot returned is presented in the following Fig. 2.

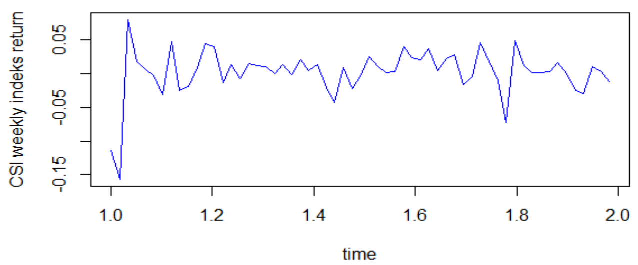
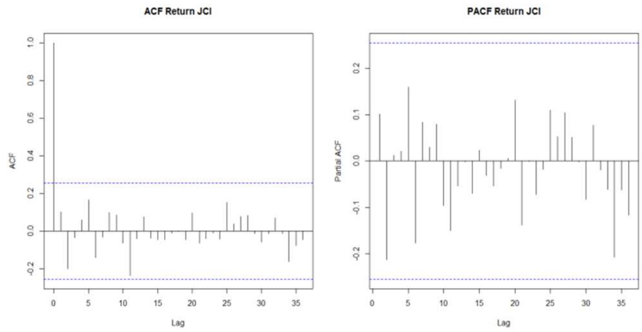


Fig. 2 Plot of weekly return data of CSI from March 2, 2020 to April 19, 2021.

Fig. 2 shows that the CSI weekly data plot has upward and downward trends at any given time. The plot also shows that the highest CSI share price occurred on January 11, 2021, with a price index of 6373.41, and the lowest CSI share price was on March 16, 2020 with a price index of 4194.94. The data plot also showed that the data did not fluctuate around the mean and the variance was not fixed over time. Even though, this indicated that the weekly CSI data regarding mean and variance were not clearly stationary.

The conversion of the data into return format was followed by the stationarity and ADF tests, and the results showed that the p-value, 0.0116, was smaller than $\alpha = 0.05$. This led to the conclusion that the H_0 was rejected, thereby indicating the CSI weekly data were stationary.

The ACF and PACF values obtained from the R-studio and presented in Fig. 3 were used to identify the starting steps to create the ARIMA Model.



a) b)

Fig. 3 a). ACF plot, and b). PACF plot.

The ACF output indicated that the coefficient was significant at lag-1 while the PACF did not have any significant coefficient value at any point. This showed that the potential models to be formed were ARIMA (1,0,1), ARIMA(1,0,0), and ARIMA(0,0,1). Moreover, the parameter estimation process was conducted using the R-studio software, and the results are presented in the following TABLE I.

TABLE I
Parameter Estimation Results and AIC Value for the ARIMA Model

Model	ϕ_1	θ_1	AIC
ARIMA(1,0,1)	-0.6736283	0.8741824	-223.71
ARIMA(1,0,0)	0.1246359	-	-222.49
ARIMA(0,0,1)	-	0.183416	-222.94

Table I shows that the best ARIMA model was the ARIMA(1,0,1) because it had the smallest AIC value. This was followed by the determination of the GARCH effect using the LM, autocorrelation, and normality tests.

B. Diagnostic Check of Heteroscedasticity, Autocorrelation and Normality

The determination of the best model was followed by the evaluation of the heteroscedasticity to understand the effect of the remaining ARCH in the ARIMA model estimation results on the residuals. This was achieved using the Lagrange Multiplier test based on the following hypothesis:

H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ (no heteroscedasticity effect)

H_1: There is at least one $\alpha_i \neq 0$, for i=1,2,...,m (there is a heteroscedasticity).

The ADF test showed that the p-value 0.00479 was smaller than α = 0.05 and this led to the conclusion that the hypothesis H_0 was rejected, and the CSI weekly data was stationary. Therefore, the ARCH/GARCH affected the residuals of the ARIMA (1,0,1) model.

The autocorrelation test was also conducted using the Ljung Box based on the following hypothesis:

H_0: $\rho_1 = \rho_2 = \cdots = \rho_m = 0$ (there is no autocorrelation in the residuals).

H_1: There is at least one $\rho_i \neq 0$, for i=1, 2, ..., *m* (there is autocorrelation in the residuals).

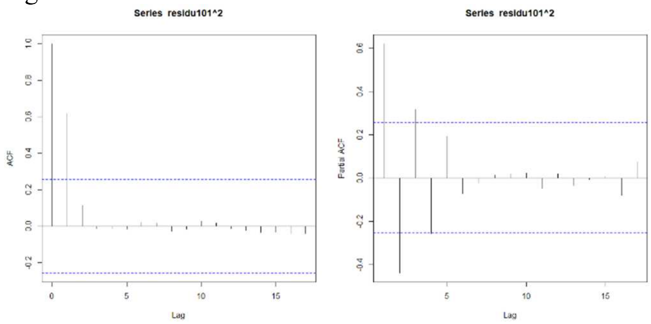
The results showed that the Q_{LB} value 4.7593 obtained was smaller than the $\chi^2_{0.05}(8) = 15.5073$ and the probability p-value of the Ljung Box was also 0.783 more significant than the α value $= 0.05$. This led to the conclusion that there was no autocorrelation in the residuals of the ARIMA (1,0,1) model.

The normality test was also applied to ascertain the distribution status of the residuals from the ARIMA model using the Shapiro-Wilk technique based on the following hypotheses:

H_0: The data has a normal distribution.

H_1: The data does not.

The results showed that the probability value obtained from the Shapiro-Wilk test was 2.001×10^{-05} and this was smaller than the α value of 0.05. This indicated that the residuals were clearly not distributed normally. The normality test status can be ignored because fluctuating data changes with a heavy tailed tendency do not provide changes to the ARIMA model for CSI financial data. The residual assumption test conducted was observed to have indicated the existence of heteroscedasticity in the ARIMA (1,0,1) and this could be solved using the ARCH/GARCH model. This led to the plotting of the ACF and PACF as presented in the following Fig. 4.



a) b)

Fig. 4 a). ACF and b). Plot PACF residuals of ARIMA (1,1).

The results from the ACF plot showed that the coefficient value was only significant at lag-1 while the PACF Plot indicated a lack of significant coefficient value at any lag until lag-3. Therefore, it was possible to form GARCH (3,2), GARCH (3,1), GARCH(2,2), GARCH(1,2), GARCH(1,1), and GARCH(1,0). This was followed by the presentation of the parameter estimation in TABLE II.

TABLE II
PARAMETER ESTIMATION OF GARCH MODEL AND AIC VALUE

Model	ω	α_1	α_2	α_3	β_1	β_1	AIC
GARCH (3,2)	-1.598×10^{-04}	3.518×10^{-02}	-1.789×10^{-02}	-2.396×10^{-02}	$1.276 \times 10^{+00}$	$5.025 \times 10^{-0-3}$	-3.947921
GARCH (3,1)	-1.456×10^{-04}	2.153×10^{-02}	-9.918×10^{-03}	-2.653×10^{-02}	$1.263 \times 10^{+00}$	-	-3.992398
GARCH (2,2)	-9.451×10^{-05}	2.874×10^{-02}	-5.128×10^{-02}	-	$-1.184 \times 10^{+00}$	1.376×10^{-02}	-4.026991
GARCH (1,2)	5.664×10^{-05}	1.085×10^{-08}	-	-	8.963×10^{-01}	9.464×10^{-09}	-3.881724
GARCH (1,1)	7.725×10^{-04}	7.075×10^{-01}	-	-	-3.334×10^{-01}	-	-4.216274
GARCH (1,0)	4.876×10^{-04}	5.817×10^{-01}	-	-	-	-	-4.198221

The best GARCH model with small value of AIC is determined as GARCH (1,1), and it is obtained from the parameter estimation results and presented as follows:

$$\sigma_t^2 = 0.0007725 + 0.7074639\varepsilon_{t-1}^2 - 0.3333629\sigma_{t-1}^2 \quad (15)$$

This model identified volatility of CSI depending on only one lag. This short-term lag of the model indicates that volatility changes fast for the next time.

C. MS-GARCH Modeling for Volatility of CSI

The CSI weekly data were observed to have the tendency to increase and decrease rapidly, thereby indicating the existence of a structural change. This was determined using the F-test statistics based on the following hypotheses:

H_0: $\beta_{i,j} = \beta_i$, i = 1,2, …, n (no structural change).

H_1: $\beta_{i,j} \neq \beta_i$= i, i = 1, 2,…, n (there is a change in structure).

The analysis conducted in the R-studio showed a *p*-value of 2.2×10^{-16} which was smaller than the α = 0.05, thereby indicating the rejection of the hypothesis and this meant there was a structural change in the CSI weekly data.

The MS GARCH model was required to determine the volatility conditions that led to structural changes. This was observed from the difference between regime 1 for high volatility and regime 2 for low volatility. Moreover, the order selected to estimate the MS-GARCH model was based on the previous GARCH, and this led to the preference for MS-GARCH (1,1). This was followed by the estimation process and the equation of two regimes GARCH model is presented as follows:

$$\sigma_t^2 = \begin{cases} 0.0005 + 0.3175\varepsilon_{t-1}^2 + 0.0001\sigma_{t-1}^2; \text{ regime1} \\ 0.0005 + 0.3178\varepsilon_{t-1}^2 + 0.0001\sigma_{t-1}^2; \text{ regime2} \end{cases} \quad (16)$$

The MS GARCH model of transition probability matrix is also obtained as follows:

$$P = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} = \begin{bmatrix} 1.0000 & 0.0000 \\ 0.3148 & 0.6852 \end{bmatrix} \quad (17)$$

The results of the transition probability matrix were used to conclude that the probability of low volatility remaining constant was 1.0000 while the high changing to low was 0.0000. Moreover, the chance of high volatility remaining constant was 0.6852 while the probability of low volatility turning to high was 0.3148. The information of probability the change of volatility is very useful for decision in stock investmens.

D. Bayesian MS-GARCH Modeling for Volatility of CSI

The Bayesian MS-GARCH) model was combined with the Bayesian methods for the estimation process. The purpose was to determine the volatility conditions experiencing structural changes and to overcome the problems associated with a limited amount of data. Therefore, the MS-GARCH (1,1) obtained from the previous estimation results was further used to determine the Bayesian MS-GARCH (1,1) model through the R-studio software, and the equation obtained is presented as follows:

$$\sigma_t^2 = \begin{cases} 0.0004 + 0.4853\varepsilon_{t-1}^2 + 0.1111\sigma_{t-1}^2; \text{regime1} \\ 1.1661 + 0.3643\varepsilon_{t-1}^2 + 0.3361\sigma_{t-1}^2; \text{regime2} \end{cases} \quad (18)$$

The transition probability matrix of the Bayesian MS-GARCH model were also obtained as follows:

$$P = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix} = \begin{bmatrix} 0.8938 & 0.1062 \\ 0.6425 & 0.3575 \end{bmatrix} \quad (19)$$

These results showed that the probability of the low volatility remaining constant was 0,8938 while the high changing to the low was 0.1062. Moreover, the chance of high volatility remaining constant was 0.3575 while the low turning to high was 0.6425. The results showed that the model Bayesian MS-GARCH could be used more naturally for data with the cases not only changes in data structure but also to overcome the effect heteroscedasticity and asymmetric effects on its volatility.

E. The Best Model Selection

The best model was selected using the smallest values of AIC and BIC as indicated in the following Table III.

TABLE III
COMPARISON OF AIC AND BIC OF THE MODEL

Model	**Best Selection Model**	
	AIC	**BIC**
GARCH	-4.2163	-4.1106
MS GARCH	-243.1777	-226.5574
Bayesian MS-GARCH	-388.9084	-372.2881

The results showed that the best model for CSI volatility was the Bayesian MS-GARCH model due to its low AIC and BIC values compared to the others. The following gives the application of Bayesian MS-GARCH model in investment. The VaR approach was used to determine the level of risk investors were expected to experience when investing in CSI. The value was calculated at an error rate of $\alpha = 0.05$ and the investment time T was varied at 1, 5, 10, 20, and 30 weeks with the initial capital to be invested assumed to be IDR 100,000,000. The results showed that the VaR values for each model were IDR 3,722,966 for a holding period of 1 week, for 5 weeks IDR 8,324,805, for 10 weeks IDR 11,773,052, for 20 weeks IDR 16,649,610, and for 30 weeks IDR 20,391,525. The longer holding periods than higher the VaR value as the investor's consideration.

This study modeled the CSI financial data volatility using the Bayesian MS-GARCH model with due consideration for the stochastic changes in the data structure. Moreover, the model was used to determine the VaR at a confidence level of $\alpha = 0.05$ and this provided information on the potential losses that might be experienced when investing over a specific period. These results were expected to assist the investors in managing investment risk and making more informed decisions.

IV. CONCLUSION

This study was conducted to determine the performance of the GARCH and MS-GARCH models. It was observed that the MS GARCH estimation process focused on the maximum likelihood algorithm and the Bayesian approach to model the volatility of CSI stock data. The results showed that the MS-GARCH model with a Bayesian approach had the smallest AIC and BIC values, thereby indicating the existence of the lowest variance and volatility. This further indicated that the potential risk of investing in CSI data was at the smallest using this model. The results further showed that the model could be used for data with the cases of heteroscedasticity, changes in data structure, and asymmetric effects. The Bayesian model of MS-GARCH has described volatility of CSI very well, this result can be used to assess the investment risk based on the Value at Risk as the effort to minimize the risk faced by the investors.

ACKNOWLEDGMENT

This study was supported by the grant scheme of master thesis research from the Ministry of Education, Culture, Research and Technology of Indonesia with contract number 115/E5/PG.02.00.PL//2023.

REFERENCES

[1] J.D. Hamilton and R. Susmel, "Autoregressive conditional heteroscedasticity and changes in regime," *Journal of Econometrics*, vol. 64, no. 1-2, pp. 307-333, 1994, doi: 10.1016/0304-4076(94)90067-1.

[2] S.F. Gray, "Modeling the conditional distribution of interest rates as a regime-switching process," *Journal of Financial Economics,* vol. 42, no. 1, pp. 27-62, 1996, doi: 10.1016/0304-405X(96)00875-6.

[3] M. Cavicchioli, "Spectral analysis of Markov switching GARCH models with statistical inference Scandinavian," *Journal of Statistics*. vol. 50, no. 1, pp. 102-119, 2023, doi: 10.1111/sjos.12571.

[4] D.C.H. Wee, F. Chen, and W.T.M. Dunsmuir, "Likelihood inference for Markov-switching GARCH(1,1) models using sequential Monte Carlo," *Econometrics and Statistics*, vol. 21, pp. 50-68, 2022, doi:10.1016/j.ecosta.2020.03.004.

[5] D. Devianto, M. Maiyastri, and D.R. Fadhilla, "Time series modeling for risk of stock price with value at risk computation," *Applied Mathematical Sciences*, vol. 9, no. 56, pp. 2779-2787, 2015, doi:10.12988/ams.2015.52144.

[6] D. Devianto, M. Yolanda, M. Maiyastri, and F. Yanuar, "The soft computing FFNN method for adjusting heteroscedasticity on the time series model of currency exchange rate," *Applied Mathematical Sciences*, vol. 9, pp.1–18, 2023, doi: 10.3389/fams.2023.1045218.

[7] Y. Runfang, D. Jiangze, and L. Xiaotao, "Improved forecast ability of oil market volatility based on combined Markov switching and GARCH-class model," *Procedia Computer Science*, vol. 122, pp. 415–422, 2017, doi: 10.1016/j.procs.2017.11.388.

[8] L.D. Persio and S. Vettori, "Markov switching model analysis of implied volatility for market indexes with applications to S&P 500 and DAX," *Journal of Mathematics*, vol. 2014, pp. 1-17, 2014, doi:10.1155/2014/753852.

[9] R.C.K. Burdekin and R Tao, "Chinese influences on inflation determination in Australia and the ASEAN group: A Markov-switching analysis," *Asia and the Global Economy,* vol. 2, no. 2, pp. 1-17, 2022, doi: 10.1016/j.aglobe.2022.100037

[10] O.V.D. Torre-Torres, E. Galeana-Figueroa, and J. Alvarez-Garcia, "A Test of Using Markov-Switching GARCH Models in Oil and Natural Gas Trading," *Energies,* vol. 13, no. 129, pp. 1-24, 2020, doi:10.3390/en13010129.

[11] M. Shakeel, M.R. Rabbani, I.T. Hawaldar, V. Chhabra, and F.K. Zaidi, "Is there an intraday volatility spillover between exchange rate, gold and crude oil?," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 9, no. 3, pp. 100094, 2023, doi:10.1016/j.joitmc.2023.100094.

[12] D. Ardia, K. Bluteau, and M. Ruede, "Regime changes in Bitcoin GARCH volatility dynamics," *Finance Research Letters*, vol. 29, pp. 266–271, 2019, doi: 10.1016/j.frl.2018.08.009.

[13] C-Y. Tan, Y-B. Koh, K-Haur Ng, and K-Huat Ng, "Dynamic volatility modelling of Bitcoin using time-varying transition probability Markov-switching GARCH model," *The North American Journal of*

Economics and Finance, vol. 56, pp. 101377, 2021, doi:10.1016/j.najef.2021.101377.

[14] G.M. Caporale and T. Zekokh, "Modeling volatility of cryptocurrencies using Markov-switching GARCH models," *Research in International Business and Finance*, vol. 48, pp. 143–155, 2019, doi:10.1016/j.ribaf.2018.12.009.

[15] C. Wu, "Window effect with Markov-switching GARCH model in the cryptocurrency market," *Chaos, Solitons and Fractals*, vol. 146, pp. 110902, 2021, doi: 10.1016/j.chaos.2021.110902.

[16] I. Barjasic and N. Antulov-Fantulin, "Time-Varying Volatility in Bitcoin Market and Information Flow at Minute-Level Frequency," *Frontiers in Physics*, vol. 9, pp. 1-12, 2021, doi:10.3389/fphy.2021.644102.

[17] J. Wang, F. Ma, C. Liang, and Z. Chen, "Volatility forecasting revisited using Markov-switching with time-varying probability transition," *International Journal of Finance and Economics*, vol. 27, no.1, pp. 1387-1400, 2019, doi: 10.1002/ijfe.2221.

[18] Z. Pan, Y. Wang, C. Wu, and L. Yin, "Oil price volatility and macroeconomic fundamentals: A regime-switching GARCH-MIDAS model," Journal of Emperical Finance, vol. 43, pp. 130-142, 2017, doi:10.1016/j.jempfin.2017.06.005.

[19] L. Wang, J. Wu, Y. Cao, and Y. Hong, "Forecasting renewable energy stock volatility using short and long-term Markov switching GARCH-MIDAS models: Either, neither or both?," *Energy Economics*, vol. 111, pp. 106056, 2022, doi: 10.1016/j.eneco.2022.106056.

[20] M. Segnon, R. Gupta, and B. Wilfling, "Forecasting stock market volatility with regime-switching GARCH-MIDAS: The role of geopolitical risks, *International Journal of Forecasting*, vol. 40, no. 1, pp. 29-43, 2024, doi: 10.1016/j.ijforecast.2022.11.007.

[21] C.B. Kim, "Leverage effect of HRCI volatility and the volatility impact on Korean export container volume before and after the global financial crisis: Application of ARIMA-EGARCH and GIRF," *.The Asian Journal of Shipping and Logistics*, vol. 34, no. 3, pp. 231–237, 2018, doi: 10.1016/j.ajsl.2018.09.006.

[22] R. Casarin, M. Costantini, and A. Osuntuyi, "Bayesian nonparametric panel Markov-switching GARCH Models," *Journal of Business and Economic Statistics*, vol. 14 Feb 2023, pp. 1-12, 2023, doi:10.1080/07350015.2023.2166049.

[23] Y.A. Shiferaw, "The Bayesian MS-GARCH Model and Value at Risk in South African Agricultural Commodity Price Markets. *International Association of Agricultural Economists,* vol. 275991, pp. 1-15, 2018, doi: 10.22004/ag.econ.275991.

[24] W. Nyakabawo, R. Gupta, and H.A. Marfatia, " High frequency impact of monetary policy and macroeconomic surprises on us MSAs, aggregate US housing returns and asymmetric volatility, *Advances in Decision Sciences*, vol. 22(A), pp. 204-229, 2018, doi:10.47654/v22y2018i1p204-229.

[25] D. Devianto, K. Ramadani, M. Maiyastri, Y. Asdi, and M. Yolanda, "The hybrid model of autoregressive integrated moving average and fuzzy time series Markov chain on long-memory data," *Applied Mathematical Sciences*, vol. 8, pp.1–15, 2022, doi:10.3389/fams.2022.1045241.

[26] S.D. Sanzo, "A Markov switching long memory model of crude oil price return volatility," *Energy Economics*, vol. 74, pp. 351-359, 2018, doi: 10.1016/j.eneco.2018.06.015.

[27] H. Xiao, Q. Zhu, and H.R. Karimi, "Stability analysis of semi-Markov switching stochastic mode-dependent delay systems with unstable subsystems," *Chaos, Solitons and Fractals*, vol. 165, pp. 112791, 2022, doi: 10.1016/j.chaos.2022.112791.

[28] J. Liu, Y. Pan, J. Pan, and A. Almarashi. Markov-switching Poisson generalized autoregressive conditional heteroscedastic models," *Statistics and Its Interface*, vol. 16, no. 4, pp. 531 – 544, 2023, doi:10.4310/22-SII741

[29] L. Feng, T. Fu,Y. Shi, and Z. Wang, "Discussions on the zero-drift GARCH model: Evidence from an Markov regime-switching extension," *Finance Research Letters*, vol. 40, pp. 101713, 2021, doi:10.1016/j.frl.2020.101713.

[30] D. Ardia, K.B. Bluteau, K. Boudt, L. Catania, and D-A. Trottier, "Markov-Switching GARCH Models in R: The MSGARCH Package," *Journal of Statistical Software*, vol. 91, no. 4, pp. 1-38, 2019, doi:10.18637/jss.v091.i04.

[31] D. Devianto, "On the class of infinitely divisible exponential distribution," *Journal of Physics: Conference Series*, vol. 1097. IOP Publishing, pp. 012084, 2018, doi: 10.1088/1742-6596/1097/1/012084.

[32] D. Devianto, Maiyastri, and K. Azmi, "On the Levy measure property of gamma distribution," *AIP* Conference *Proceedings,* vol. 2242, pp. 030003, 2020, doi: 10.1063/5.0007842.

[33] F. Goncalves, N. Mendes-Lopes, and F. Silva, "Infinitely divisible distributions in integer-valued GARCH models," *Journal of Time Series Analysis*, vol. 36, no. 64, pp. 503-527, 2015, doi:10.1111/jtsa.12112.

[34] F. Zhu, M.L. Bianchi, Y.S.A. Kim, F. Fabozzi, and H. Wu "Learning for infinitely divisible GARCH models in option pricing," *Studies in Nonlinear Dynamics and Econometrics*,vol. 25, no. 3, pp. 35-62, 2021, doi: 10.1515/snde-2019-0088.