Investment Modelling Using Value at Risk Bayesian Mixture Modelling Approach and Backtesting to Assess Stock Risk

Brina Miftahurrohmah[1] *, **Catur Wulandari[2]**, **Yogantara Setya Dharmawan[3]**

[1)2)3)]*Universitas Internasional Semen Indonesia, Indonesia*

Kompleks PT. Semen Indonesia (Persero) Tbk. Jl. Veteran, Gresik

[1)]brina.miftahurrohmah@uisi.ac.id, [2)]catur.wulandari@uisi.ac.id, [3)]yogantara.dharmawan@uisi.ac.id

Abstract

Background: Stock investment has been gaining momentum in the past years due to the development of technology. During the pandemic lockdown, people have invested more. One the one hand, stock investment has high potential profitability, but on the other, it is equally risky. Therefore, a value at risk (VaR) analysis is needed. One approach to calculate VaR is by using the Bayesian mixture model, which has been proven to be able to overcome heavy-tailed cases. Then, the VaR's accuracy needs to be tested, and one of the ways is by using backtesting, such as the Kupiec test.
Objective: This study aims to determine the VaR model of PT NFC Indonesia Tbk (NFCX) return data using Bayesian mixture modelling and backtesting. On a practical level, this study can provide information about the potential risks of investing that is grounded in empirical evidence.
Methods: The data used was NFCX data retrieved from Yahoo Finance, which was then modelled with a mixture model based on the normal and Laplace distributions. After that, the VaR accuracy was calculated and then tested by using backtesting.
Results: The test results showed that the VaR with the mixture Laplace autoregressive (MLAR) approach (2;[2],[4]) was accurate at 5% and 1% quantiles while mixture normal autoregressive MNAR (2;[2],[2,4]) was only accurate at 5% quantiles.
Conclusion: The better performing NFCX VaR model for this study based on backtesting using Kupiec test is MLAR(2;[2],[4]).

Keywords: Backtesting, Bayesian, Kupiec Test, Mixture Model, Value at Risk

Article history: Received 23 October 2020, first decision 11 December 2020, accepted 2 February 2021, available online 28 April 2021

I. INTRODUCTION

Economic uncertainty in the current pandemic has changed some aspects of consumer behaviour in Indonesia. Stock investment is on the rise among the Indonesians people today. This is shown by a DBS bank survey, which states that people will choose to save and invest their money rather than spending it on unnecessary consumer products or services [1]. In doing investment, people certainly hope to get a return after a certain period of time, but they also face a risk of loss. Risk assessment is needed to minimize such risk of loss.

As most people were in lockdown and shopping was considered more convenient online, sales on e-commerce soared to 66% during the COVID-19 pandemic [1]. This makes risk analysis on e-commerce stocks interesting to research. In 2020, three e-commerce companies took the floor on the Indonesia Stock Exchange (IDX), one of which is PT NFC Indonesia Tbk (NFCX) [2]. This company is the newest e-commerce company joining the IDX and there has been no research on the company related to the risk analysis [2]. Therefore, this study chose NFCX as the subject.

Value at Risk (VaR) is a tool commonly used to measure financial risks when the distribution of potential loss positions is not generally known [3]. VaR is generally estimated by assuming returns that have a normal distribution. However, this may not be appropriate because, in reality, return distribution is not always normal [4]. Some results of stock return visualization indicate that the data distribution is leptokurtic, which tends to resemble Laplace distribution. That being said, this does not mean that Laplace distribution is the one suitable for determining VaR. It is partially suitable because of the long-tail and asymmetric conditions that may occur in the stock return data. This kind of financial data could be handled by the asymmetrical distribution of Laplace [5]. Research shows that a parametric method based on skewed and fat-tails is the best method for determining VaR, especially when time variations are considered and independent and identical return distributions are ignored [6].

A method with a mixture model approach has been used in several studies to solve problems that occur in unimodal data [7] [8]. By considering the skewed and the heterogeneity, the mixture model method can improve the accuracy

---

* Corresponding author

of predictive cases [7] [8]. It has been proven to be better than the separate one [9] [10] and suitable to be applied to the autoregressive (AR) model, a time series model. Meanwhile, a mixture normal autoregressive (MNAR) has been developed both to consider the mixture of normal distribution and to analyse the time series that shows regime-switching behaviour. The method considers the probability and weight of mixing based on past values, so that a) the stationarity and ergodicity of the underlying stochastic process could be easily established; and b) the explicit expression of the low-order of stationary marginal distribution is known. This is in contrast with other majority nonlinear autoregressive models [11]. The ability of MNAR to overcome the heavy-tailed problems that occur in unimodal data makes it applicable for VaR modelling [12] [13] [14] [15].

As for the pattern of actual conditions, Laplace distribution is able to capture it in mixture autoregressive models [16]. It has been proven to be more robust than the normal distribution in linear mixture models [17]. Since the distribution of returns tends to be peak similar to that of the asymmetric Laplace distribution, the current research uses MLAR for the VaR modelling [18]. Research [18] has shown how Bayesian MLAR approaches model the VaR of Islamic stock; the performance was then compared with the Bayesian MNAR analysis. The results showed that the Bayesian MLAR model performed better than the Bayesian MNAR model.

In practice, the results of VaR modelling such as this should be evaluated using a backtesting method in order to determine the best model. Kupiec Test developed by [19] is a backtesting procedure commonly used for VaR modelling studies [6] [20] [21]. All things considered, this study aims to determine the NFCX VaR model using Bayesian mixture modelling and backtesting.

II. METHODS

The analysis was carried out using stock closing data retrieved from https://finance.yahoo.com/quote/NFCX.JK/history?p=NFCX.JK from 12 June 2018 to 4 August 2020. Generally, there are four steps in doing the risk-return analysis: component identification of the mixture autoregressive models; analysis of the Bayesian mixture autoregressive models; VaR modelling; and evaluation of VaR model using Kupiec test. Before conducting the risk-return analysis, the return of NFCX stock must be calculated by (1).

$$y_t = \frac{P_t + P_{t-1}}{P_{t-1}} \qquad (1)$$

where y_t is the stock price return on the tth day, P_t is the stock price on the tth day, and P_{t-1} is the stock price on the $(t-1)^{th}$ day [22].

The first step of conducting a return-risk analysis is the component identification of the mixture autoregressive models. At this stage, the numbers of autoregressive mixture components and the autoregressive AR(p) components are determined by analysing the histogram of the return and autoregressive modelling respectively. The following is the AR(p) model [23] as seen on (2).

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t. \qquad (2)$$

The y_{t-p} is the return on the $(t-p)^{th}$ time, p is the autoregressive order, ϕ is the parameter of autoregressive, and ε_t is white noise. The pth order in AR(p) model is determined by ensuring that the return data is a qualified stationery in mean and variance. The parameters of the AR(p) model that have been specified are then estimated and tested to find out its significance. Only the significant ones are qualified to be mixed.

The AR(p) models that will be mixed in MNAR and MLAR modelling are the same. The difference is that the MNAR model uses normal distribution and MLAR model uses Laplace distribution. In MNAR($K; AR(p)$) model, the conditional distribution of $Y_t|y_{t-1}; \Phi_K$ is a normal mixture with conditional density function of (3).

$$f(Y_t|y_{t-1}; \Phi_K) = \sum_{k=1}^{K} \pi_k \lambda(y_t; \mu_{t,k}, \sigma_k^2) \qquad (3)$$

where π_k is the proportion of kth mixed components, and $\lambda(y_t; \mu_{t,k}, \sigma_k^2)$ is the density function of $N(\mu_{t,k}, \sigma_k^2)$. Let $\Phi_K = (\pi_1, \pi_2, \ldots, \pi_K, \phi_1^T, \phi_2^T, \ldots, \phi_K^T, \sigma_1, \sigma_2, \ldots, \sigma_K)^T$ denotes the vector of autoregressive parameters [15]. While in MLAR($K; AR(p)$) model, the conditional distribution of $Y_t|y_{t-1}; \vartheta$ is a Laplace mixture with conditional density function of (4).

$$f(Y_t|y_{t-1}; \vartheta) = \sum_{k=1}^{K} \pi_k \lambda \left(y_t; \phi_{k,i0} + \sum_{i=1}^{p} \phi_{ki,i} y_{t-i}, \sigma_k \right) \tag{4}$$

where

$$\lambda(y; \mu, \sigma) = \left(\sigma\sqrt{2}\right)^{-1} exp\left(-\sqrt{2}|x - \mu|/\sigma\right) \tag{5}$$

Equation (5) is Laplace density function with mean μ, variance σ^2, and (6) is the parameter vector of mixed models. The residual of the MLAR model is also considered to have a Laplace distribution [17].

$$\vartheta = (\pi_1, \pi_2, \dots, \pi_K, \phi_1^T, \phi_2^T, \dots, \phi_K^T, \sigma_1, \sigma_2, \dots, \sigma_K)^T \tag{6}$$

The estimation parameter determined by Bayesian inference was developed based on the Bayes theorem. Posterior distribution of Bayesian inference is shown on (7).

$$p(\theta|y) \propto f(y|\theta)p(\theta). \tag{7}$$

Markov Chain Monte Carlo (MCMC) is an algorithm of the Bayesian inference commonly used to estimate parameter θ by generating samples from a given distribution. The subsequent sample is chosen based on the sample taken previously. This is done by determining the initialization θ at the start of the sampling. As a result, the sample taken forms the MCMC $\theta^1, \theta^2, \dots, \theta^b$ chain. The distribution of the given θ^b depends only on all the preceding θ at the most recent value, which is θ^{b-1}. The generated samples are not independent, but still identically distributed if the Markov chain is stationary [24]. This parameter is then corrected to obtain a θ value that is closer to the target of the posterior distribution of $p(\theta|y)$. Gibbs sampler is one of the MCMC methods that can solve multidimensional problems. In normal and Laplace case, $\theta = (\mu, \sigma, w)$ and posterior is $p(\mu, \sigma, w|y)$. The Gibbs sampler will help estimate μ, σ, and w iteratively following the sampling scheme.

1. Given state: $\theta^b = (\mu, \sigma, w)^b$ in iteration $b = 0$
2. Generating component parameters of each mixture
   a. $\mu^{(b+1)}$ is generated from $p\left(\mu|y, \sigma^{(b)}, w^{(b)}\right)$
   b. $\sigma^{(b+1)}$ is generated from $p\left(\sigma|y, \mu^{(b+1)}, w^{(b)}\right)$
   c. $w^{(b+1)}$ is generated from $p\left(w|y, \mu^{(b)}, \sigma^{(b+1)}\right)$
3. Repeat step two T times, $\rightarrow \infty$

In estimating the mixture parameters, step 2 must estimate as much as K of the mixture components of both μ, σ, and w. The samples generated using the above algorithm will have a convergent and stationary data pattern and will be proportional to their respective distributions [18] [25]. The confirmation of an ergodic Markov chain must be made to identify the existence of limiting distribution in this chain. It can be divided into three sections, namely irreducible, periodicity, and recurrent and transient states [26]. The parameter significance test is used to select the suitable ones for the model. Testing the parameters resulting from the estimation with Bayesian MCMC assumes that the null hypothesis of $\theta = 0$ and the alternative hypothesis of $\theta \neq 0$. The null hypothesis is rejected if in confidence interval $(1 - \alpha)$ of posterior, the credible interval does not contain null [27]. After obtaining the MNAR and MLAR model, deviance information criterion (DIC) for each MLAR model is calculated and the model with the smallest DIC is selected. The DIC formula is shown on (8).

$$DIC = \underline{D} + 2p_D \tag{8}$$

$\underline{D}$ is the posterior mean of the deviance that is defined as $-2 \, log \, log \, p(y|\theta)$. p_D is the effective number of parameters and is given by $p_D = \underline{D} - \widehat{D}$ where $\widehat{D} = -2 \, log \, log \, p(y|\underline{\theta})$ [28] [29]. The best model of MNAR and MLAR that has been obtained is used to determine the VaR of each mixture model. It is calculated on (9).

$$VaR_{mix} = \pi_1\sqrt{h}VaR_{1,\alpha} + \pi_2\sqrt{h}VaR_{1,\alpha} + \cdots + \pi_K\sqrt{h}VaR_{1,\alpha} \tag{9}$$

where for MNAR best model is shown on (10),

$$VaR_{1,\alpha} = \Phi^{-1}(1-\alpha) \tag{10}$$

and for MLAR best model is shown on (11).

$$VaR_{1,\alpha} = -\left(\sigma\left(ln\,ln\left(4(-\alpha^2 + \alpha)\right)\right)\right). \tag{11}$$

The last stage to evaluate the VaR model is by using backtesting, which is a statistical procedure to systematically compare the actual gains and losses with the estimated VaR. The most widely used backtesting, the Kupiec test, which is also known as the POF (proportion of failure) test, measures whether the number of exceptions is consistent with the αth quantile [30], which follows the binomial distribution. In other words, the information needed to perform the Kupiec test is the number of observations (T), the number of exceptions (x), and the αth quantile [2]. The null hypothesis of this test is $p = \hat{p}$ and the alternative hypothesis is $p \neq \hat{p}$. The test statistic used is the likelihood ratio (LR) [19] equated on (12).

$$LR = -2\,ln\,ln\left(\frac{(1-p)^{T-x}p^x}{\left(1-\frac{x}{T}\right)^{T-x}\left(\frac{x}{T}\right)^x}\right) \tag{12}$$

where p is the probability of failure in the αth quantile. LR is asymptotic with a chi-square distribution (χ^2) and a degree of freedom is 1. The null hypothesis is rejected if LR is greater than χ^2. Accordingly, the VaR model is declared valid if the null hypothesis is accepted. Finally, the last result of this analysis will show the VaR that the investors of PT. NFC Indonesia Tbk (NFCX) will face.

III. RESULTS

This section presents the results of the VaR modelling process, which consists of four steps-the identification, the analysis, the modelling and the backtesting. The best mixture model determined based on a normal and Laplace distribution is derived from the comparison of several mixture VaR models.

A. Component Identification of Mixture Autoregressive Model

The identification determines the number of components and which components of autoregressive and AR(p) to be mixed. The number of components of mixture autoregressive is detected from the histogram, whereas the mixable autoregressive components AR(p) are those with significant parameters.

The histogram of the NCFX return data, as presented in Fig. 1, shows that the data has outliers. The outliers of the data are indicated in the histogram as right-skewed or positively-skewed patterns. Besides, Fig. 1 also identifies the shapes of its frequency distribution, namely leptokurtic, platykurtic, and mesokurtic. If the peak of the curve is higher than the normal distribution, it is considered leptokurtic; if it is lower, it is platykurtic; and if it is the same, it is mesokurtic [31]. The histogram shows that the data have a higher peak than the normal distribution, so they are leptokurtic, which tends to be more similar to the Laplace distribution (blue line) than the normal distribution (red line). However, the outliers and the high variance result in a mismatched Laplace distribution. Variability in a histogram is higher when the taller bars are spread away from the mean; and lower when they are closer to the mean. The solution to these cases is to form a combination for each distribution. This results in two components for each distribution. One component is leptokurtic and the other is platykurtic. The two components in the Laplace distribution are shown by the black and yellow dash line, while the normal distribution is presented by the green and magenta dash line. Platykurtic conditions are expected to overcome the high variation.
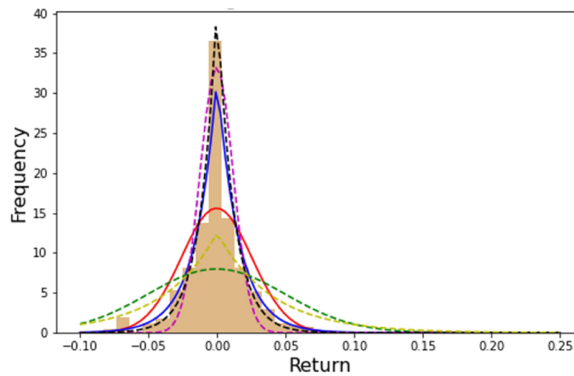
Fig. 1 The Histogram of NFCX Return

After obtaining the number of components, the AR(p) was selected using parameter significance. The order of AR(p) was determined by ensuring that the data was stationary in mean and variance. Stationarity of mean was detected from time series plot and augmented dickey fuller (ADF) test. The data in the time series plot, as presented in Fig.2, fluctuate around the mean. The plot indicates that the data are stationary in mean. This result was confirmed by the ADF test. To recall, the null hypothesis is when the data are not stationary in mean, whereas the alternative hypothesis is when the data are stationary in mean. Because the P-value (0.01) is less than the significant level (0.05), the null hypothesis is rejected. Furthermore, the data must also be stationary in variance. Stationarity in variance was detected by rounded value (λ) where the rounded value of the data equals 1. However, the NFCX return data did not fulfill the assumption since $\lambda = -1$; thus, the data must be Box-Cox transformed until $\lambda = 1$. After $\lambda = 1$, the order of AR(p) can be detected.
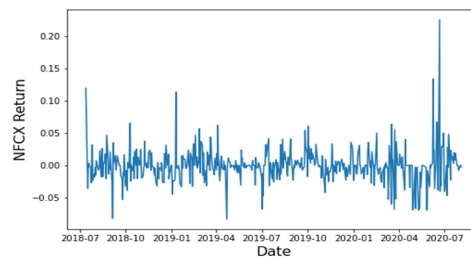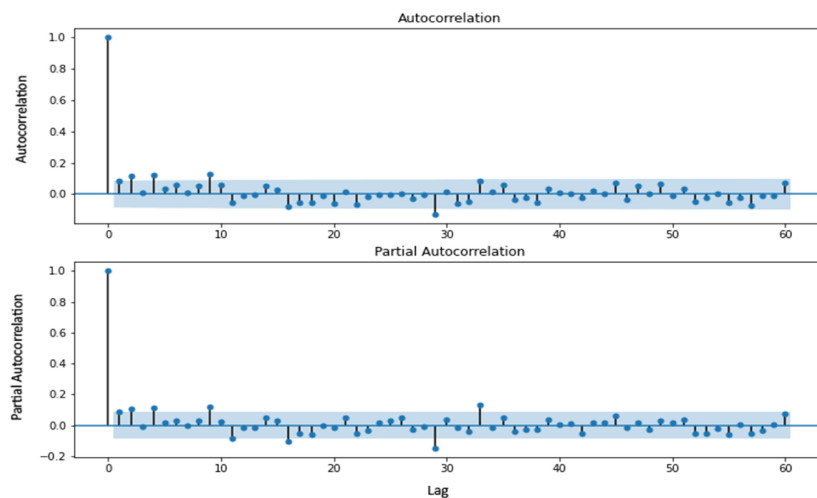

Fig. 2 Time Series Plot of NFCX Return

Fig. 3 Autocorrelation (ACF) and Partial Autocorrelation (PACF) of NFCX Return

15

AR(p) order is determined by autocorrelation (ACF) and partial autocorrelation (PACF) plot, namely those which had the same cut off p-lag. The cut off is decided based on the lag that exits the blue interval limit. ACF and PACF cut off lag 1, 2, 4, and 9, respectively. The order of AR(p) models can be seen in Table 1.

Parameters were estimated using the Bayesian method with the function of $y_t \sim N(\phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p}, \sigma)$. The prior distribution for autoregressive parameters (ϕ_p) was conjugate prior and σ was noninformative prior. This is defined by limiting the model to a relatively simple likelihood function with a suitable formula for the previous distribution [27]. The noninformative prior is a distribution that has a greater range of uncertainty than the reasonable parameter value [32]. The prior autoregressive parameter is normal distribution, while the standard deviation (σ) is inverse Gamma. The estimation and significance of the parameters results can be seen in Table 1. A 95% credible interval indicates the 2.5[th] percentile and the 97.5[th] percentile since this has been used by some authors and software [27]. Table 1 shows that all parameters of autoregressive models that contain $p = 1$ are not significant because the credible interval contains 0. However, all autoregressive models that do not contain $p = 1$ are significant.

TABLE 1
THE SIGNIFICANCE OF THE PARAMETER IN THE AUTOREGRESSIVE MODEL

Model	Parameter	Estimation	Standard Deviation	Credible Interval		Results
				2.50%	97.50%	
AR([1])	ϕ_1	0.068	0.043	-0.014	0.153	Not Significant
	σ	0.025	0.001	0.024	0.027	
AR([2])	ϕ_2	0.105	0.043	0.023	0.190	Significant
	σ	0.025	0.001	0.024	0.027	
AR([4])	ϕ_4	0.125	0.043	0.043	0.209	Significant
	σ	0.025	0.001	0.024	0.027	
AR([9])	ϕ_9	0.140	0.043	0.058	0.224	Significant
	σ	0.025	0.001	0.024	0.027	
AR([1,2])	ϕ_1	0.064	0.044	-0.020	0.151	Not Significant
	ϕ_2	0.101	0.050	0.016	0.185	
	σ	0.025	0.001	0.024	0.027	
AR([1,4])	ϕ_1	0.070	0.044	-0.015	0.156	Not Significant
	ϕ_4	0.124	0.050	0.040	0.209	
	σ	0.025	0.001	0.024	0.027	
AR([1,9])	ϕ_1	0.064	0.044	-0.021	0.150	Not Significant
	ϕ_9	0.137	0.050	0.053	0.222	
	σ	0.025	0.001	0.024	0.027	
AR([2,4])	ϕ_2	0.109	0.044	0.025	0.195	Significant
	ϕ_4	0.113	0.050	0.029	0.198	
	σ	0.025	0.001	0.023	0.027	
AR([2,9])	ϕ_2	0.122	0.044	0.038	0.208	Significant
	ϕ_9	0.139	0.050	0.055	0.223	
	σ	0.025	0.001	0.023	0.027	
AR([4,9])	ϕ_4	0.123	0.044	0.039	0.209	Significant
	ϕ_9	0.135	0.050	0.051	0.219	
	σ	0.025	0.001	0.023	0.027	
AR([1,2,4])	ϕ_1	0.063	0.044	-0.021	0.147	Not Significant
	ϕ_2	0.105	0.052	0.019	0.192	
	ϕ_4	0.113	0.045	0.030	0.197	
	σ	0.025	0.001	0.023	0.027	
AR([2,4,9])	ϕ_2	0.109	0.043	0.024	0.194	Significant
	ϕ_4	0.110	0.046	0.025	0.196	
	ϕ_9	0.135	0.050	0.051	0.219	
	σ	0.025	0.001	0.023	0.026	

B. Analysis of Bayesian Mixture Autoregressive Model

The mixture models are created by mixing two components of significant AR(p) as presented in Table 1. The number of mixed models is 42 consisting of 21 MNAR and 21 MLAR. In this study, only six models—three MNAR,

and three MLAR—are shown. Parameter estimation began by creating a directed acyclic graph (DAG) for each model. DAG for MNAR and MLAR can be seen in Fig.4. The models are assumed as follows.

$$y_t \sim N(\mu_t, \tau_t) \text{ for MNAR and } y_t \sim Laplace(\mu_t, \tau_t) \text{ for MLAR}$$
$$\phi \sim N(a, b) \text{ for MNAR and } \phi \sim Laplace(a, b) \text{ for MLAR}$$
$$\tau 1 \sim \Gamma(0.001, 0.001)$$
$$\tau 2 \sim \Gamma(0.001, 0.001)$$
$$T_t \sim Cat(\pi)$$
$$\pi \sim Dir(\alpha).$$

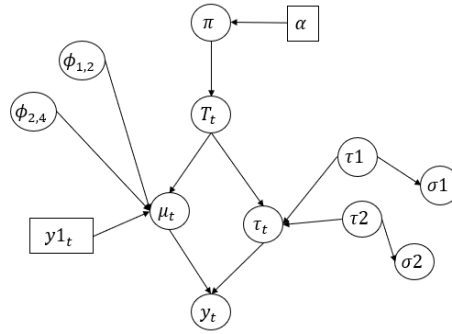

Fig. 4 The DAG of Mixture Autoregressive Models (example: MLAR (2;[2],[4]) models)

$\mu_t = \phi_{T_\pi, lag-p} y1_t$ and $\tau_t = \tau_{T_\pi}$ where T_π represents the category of $\pi = 1,2$ in level 1 and α is constant in level 2. The number of ϕ nodes depends on the number of lags and K components of the mixture. The number of independent variables in Fig. 4 is constant nodes ($y1_t$) and depends on the maximum parameter of each component. If component one has two parameters and component two has three parameters, the number of independent variables is three ($y1_t, y2_t$, and $y3_t$). The number of σ depends on the number of mixed components, and in this case, the number of mixed components is $\sigma 1$ and $\sigma 2$, where $\sigma = \frac{1}{\sqrt{\tau}}$ for MNAR models and $\sigma = \frac{\sqrt{2}}{\sqrt{\tau}}$ for MNAR models. The prior distribution for the autoregressive parameters (ϕ_p) is the conjugate prior for MNAR models; whereas the prior distribution for MLAR model is the pseudo priors, whose value is determined based on the frequentist estimation [33]. The pseudo prior of the MLAR models is based on the parameter estimation of the AR(p) models. The prior distribution for other parameters (σ, T, π) is noninformative prior.

The results of parameter estimation using DAG structure—as presented in Fig. 4—for each mixture autoregressive models are shown in Table 2. It shows that all MLAR and MNAR models are significant because the 95% credible interval does not contain 0. The selection of MNAR and MLAR models was measured using DIC with those being the smallest DIC is considered the best. Table 3 presents the best model of MNAR, namely MNAR (2;[2],[2,4]); and MLAR, namely MLAR(2;[2],[4]). The DIC of those models are -3293.10 and -3698.88 respectively. The DIC of the MLAR(2;[2],[4]) model is smaller than the MNAR(2;[2],[2,4]); however, the isolated model cannot be said to be the best model for estimating VaR; instead, both models can be used to calculate VaR.

C.   Value at Risk Modelling and Backtesting Using Kupiec Test

Value at Risk (VaR) calculated using MNAR (2;[2],[2,4]) and MLAR(2;[2],[4]) approach can be seen in Table 4. The best VaR was obtained by performing the POF Kupiec test as shown in Table 5. The Kupiec test used the null hypothesis where the $p = \hat{p}$ and the alternative hypothesis of $p \neq \hat{p}$. The result shows that VaR MLAR (2;[2],[4]) approach is accurate at 5% and 1% quantiles. To be precise, $p = \hat{p}$ because the LR value is less than $\chi^2_{0.05}$, thus the null hypothesis is accepted. Meanwhile, VaR MNAR (2;[2],[2,4]) approach is only accurate at 5% quantile. Thus, the better result is the VaR MLAR (2;[2],[4]) approach because in addition to being accurate at 5% and 1% quantiles, MLAR (2;[2],[4]) VaR values are more accurate than MNAR (2;[2],[2,4]). It is considered more accurate because VaR MLAR (2;[2],[4]) can be trusted with a higher level of confidence of 99% (100% − 1%), while VaR MNAR (2;[2],[2,4]) only has up to a 95% confidence level (100% − 5%).

TABLE 2
THE SIGNIFICANCE OF THE PARAMETER IN THE MNAR AND MLAR MODEL

Model	Parameter	Estimation	Standard Deviation	Credible Interval 2.50%	Credible Interval 97.50%	Results
MNAR(2;[2],[2,4])	π_1	0.501	0.015	0.470	0.531	Significant
	π_2	0.499	0.015	0.469	0.530	Significant
	$\phi_{1,2}$	0.104	0.043	0.020	0.186	Significant
	$\phi_{2,2}$	0.110	0.044	0.024	0.195	Significant
	$\phi_{2,9}$	0.113	0.043	0.030	0.198	Significant
	σ_1	0.025	0.001	0.024	0.027	Significant
MNAR(2;[4],[2,4])	π_1	0.500	0.015	0.469	0.530	Significant
	π_2	0.500	0.015	0.470	0.531	Significant
	$\phi_{1,4}$	0.124	0.043	0.040	0.206	Significant
	$\phi_{2,2}$	0.110	0.044	0.024	0.195	Significant
	$\phi_{2,4}$	0.113	0.043	0.030	0.198	Significant
	σ_1	0.025	0.001	0.024	0.027	Significant
	σ_2	0.025	0.001	0.023	0.026	Significant
MNAR(2;[2],[4])	π_1	0.501	0.015	0.471	0.532	Significant
	π_2	0.499	0.015	0.469	0.530	Significant
	$\phi_{1,2}$	0.105	0.043	0.022	0.189	Significant
	$\phi_{2,4}$	0.124	0.044	0.041	0.208	Significant
	σ_1	0.025	0.001	0.024	0.027	Significant
	σ_2	0.025	0.001	0.024	0.027	Significant
MLAR(2;[2],[4])	π_1	0.501	0.015	0.470	0.531	Significant
	π_2	0.499	0.015	0.469	0.530	Significant
	$\phi_{1,2}$	0.105	0.003	0.099	0.110	Significant
	$\phi_{2,4}$	0.124	0.003	0.117	0.129	Significant
	σ_1	0.022	0.001	0.020	0.024	Significant
	σ_2	0.022	0.001	0.020	0.024	Significant
MLAR(2;[2],[2,9])	π_1	0.504	0.015	0.473	0.534	Significant
	π_2	0.497	0.015	0.466	0.527	Significant
	$\phi_{1,2}$	0.105	0.003	0.098	0.110	Significant
	$\phi_{2,2}$	0.121	0.003	0.115	0.127	Significant
	$\phi_{2,9}$	0.138	0.004	0.127	0.144	Significant
	σ_1	0.022	0.001	0.020	0.024	Significant
MLAR(2;[4,9],[2,4,9])	π_1	0.500	0.016	0.470	0.531	Significant
	π_2	0.500	0.016	0.469	0.530	Significant
	$\phi_{1,4}$	0.122	0.003	0.115	0.127	Significant
	$\phi_{2,2}$	0.109	0.003	0.102	0.114	Significant
	$\phi_{1,9}$	0.134	0.004	0.125	0.141	Significant
	$\phi_{2,4}$	0.109	0.003	0.101	0.114	Significant

TABLE 3
THE DIC OF THE MIXTURE MODEL

Model	DIC
MNAR(2;[2],[2,4])	-3293.10
MNAR(2;[4],[2,4])	-3289.41
MNAR(2;[2],[4])	-3288.68
MLAR(2;[2],[4])	-3698.88
MLAR(2;[2],[2,9]))	-3667.33
MLAR(2;[4,9],[2,4,9])	-3616.79

Therefore, investors of PT NFC Indonesia Tbk. (NFCX) with US\$57.00 at 5%-quantile during one day have a 5% chance of losing more than US\$2.34 calculated from $0.041 \times US\$57.00 = US\$2.34$ of investment over the next one day. In other words, the maximum risk of loss is US\$2.34. In general, the other interpretation of VaR, as presented in Table 4, investors of PT NFC Indonesia Tbk. (NFCX) have US\$57.00 at α %-quantile, and for x-days have a α% chance of losing more than \$ y of this investment over the next x days; or in other words, the maximum risk of loss

for the investor is $ y. This interpretation is only for the accepted models as shown in Table 5 where α is % -quantile, x is the length of investment, and y is the multiplication of VaR and investment value.

TABLE 4
THE BEST VALUE AT RISK MODEL OF MLAR AND MNAR

Model	Long Investment	5% quantiles	1% quantiles
MNAR(2;[2],[2,4])	1 day	-0.041	-0.058
	5 days	-0.092	-0.130
	20 days	-0.184	-0.260
MLAR(2;[2],[4])	1 day	-0.037	-0.071
	5 days	-0.082	-0.160
	20 days	-0.164	-0.319

TABLE 5
THE KUPIEC POF TEST

Model	% quantiles	LR	$\chi^2_{0.05}$	Results
MNAR(2;[2],[2,4]	5%	1.70	3.84	Accepted
	1%	8.19	3.84	Rejected
MLAR(2;[2],[4])	5%	0.06	3.84	Accepted
	1%	2.66	3.84	Accepted

IV. DISCUSSION

This research has a limitation on the accuracy of the VaR because it is tested only as a one-day investment. This is based on several references that have been referred to [34] [21]. Although it is known that the longer the investment the greater the risks, this information is not enough. The amount of risk that will be faced during the investment period that has been determined in this study—the five-day and twenty-day periods—also needs to be tested for its accuracy so that investors will be more confident in investing. In addition to knowing the accuracy of the risks faced during a specific investment period, a reanalysis is needed to corroborate the finding in this study.

This study result denotes that the VaR accuracy done using the Bayesian MLAR approach is more accurate than the Bayesian MNAR in terms of backtesting. This is in line with the comparison results of the mixture models based on DIC, where the best mixture model produces an accurate VaR model. In reality, this does not mean that the best mixture model produces accurate VaR, which can be seen in the research done by [2]. Hence, the backtesting test should be done on all already-formed mixture models. Also, the use of one backtesting method also limits the interpretation of the study because the Kupiec method is not always correct. Thus, backtesting methods other than Kupiec can also be added to increase the level of accuracy. Lastly, the analysis is only based on stock data in the past period. The research should be carried out by considering the factors that influence changes in share prices. Several factors are believed to affect stock prices, including oil and gold market prices and their volatilities [35].

V. CONCLUSIONS

This research was conducted to obtain an NFCX stock investment VaR model (NFCX VaR model) obtained with the Bayesian mixture model approach. The best VaR model is accurate at 5% and 1% quantiles based on the backtesting results using the Kupiec test. The research shows that the best VaR model results from MLAR(2;[2],[4]) approach, which consist of two-component Laplace mixed distribution with one component being leptokurtic, and the other platykurtic. This model's goodness shows that the model is accurate at 5% and 1% quantiles with accuracy of testing limited to one day.

Based on the discussion results, several further studies can be developed, one of which is to test VaR's accuracy not only during one day but during a predetermined period horizon. The backtesting method used should not only be the Kupiec test in order to improve the accuracy of the model. Modelling should be done with historical data from stock prices and data from factors that affect stock prices.

Author Contributions: *Brina Miftahurrohmah*: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Supervision. *Catur Wulandari*: Visualisation, Data Curation, Validation, Writing - Original Draft. *Yogantara Setya Dharmawan*: Software, Investigation, Writing – Review & Editing.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

[1] C. C. Widjaja and A. Sim, COVID-19's Impact on Indonesian Consumers Accelerating shifts in consumer behaviour, Asian Insights Office • DBS Group Research, 2020.

[2] B. Miftahurrohmah, Y. S. Dharmawan and C. Wulandari, "Bayesian Mixture Laplace Autoregressive Modeling To Estimate Value-At-Risk In E-Commerce Stocks".

[3] K. C. Cheung and F. L. Yuen, "On the uncertainty of VaR of individual risk," *Journal of Computational and Applied Mathematics,* vol. 367, 2020.

[4] S.-Y. Choi and J.-H. Yoon, "Modeling and Risk Analysis Using Parametric Distributions with an Application in Equity-Linked Securities," *Mathematical Problems in Engineering,* vol. 2020, no. Special Issue, pp. 1-20, 2020.

[5] C. Chen, "Tests of fit for the asymmetric Laplace distribution," *Statistics and Interface,* vol. 7, no. 3, pp. 405-414, 2014.

[6] P. Abad, S. Benito and C. López, "A comprehensive review of Value at Risk methodologies," *The Spanish Review of Financial Economics,* vol. 12, no. 1, pp. 15-32, 2013.

[7] Q. Li, M. Guindani, B. J. Reich, H. D. Bondell and M. Vannucci, "A Bayesian mixture model for clustering and selection of feature occurrence rates under mean constraints," *Statistical Data and Data Mining, The ASA Data Science Journal,* vol. 10, no. 6, pp. 393-409, 2017.

[8] F. Noor, A. Sajid, S. B. H. Shah, M. Zaman, M. Gheisari and V. Mariappan, "Bayesian estimation and prediction for Burr-Rayleigh mixture model using censored data," *International Journal of Communication System,* vol. 32, no. 15, pp. 1-13, 2019.

[9] M. Huang, W. Yao, S. Wang and Y. Chen, "Statistical Inference and Applications of Mixture of Varying Coefficient Models," *Scandinavian Journal of Statistics,* vol. 45, no. 3, pp. 618-643, 2018.

[10] J. Lu, C. Wang, J. Zhang and J. Tao, "A mixture model for responses and response times with a higher-order ability structure to detect rapid guessing behaviour," *British Journal of Mathematical and Statistical Psychology,* vol. 73, no. 2, pp. 261-288, 2019.

[11] L. Kalliovirta, M. Meitza and P. Saikkonen, "Gaussian mixture vector autoregression," *Journal of Econometrics,* vol. 192, no. 2, pp. 485-498, 2016.

[12] J. Valecký, "Mixture normal Value at Risk models of some European market portfolios," in *6th International Scientific Conference Managing and Modelling of Financial Risks*, Ostrava, 2012.

[13] R. Chen and L. Yu, "A novel nonlinear value-at-risk method for modeling risk of option portfolio with multivariate mixture of normal distributions," *Economic Modelling,* vol. 35, p. 796–804, 2013.

[14] K. Dowd, "Value-at-Risk," *Wiley StatsRef:Statistics Reference Online,* pp. 1-11, 2014.

[15] A. Khalili, J. Chen and D. A. Stephens, "Regularization and selection in Gaussian mixture of autoregressive models," *The Canadian Journal of Statistics,* vol. 45, no. 4, pp. 356-374, 2017.

[16] H. D. Nguyen, M. J. Geoffrey, J. F. Ullmann and A. L. Janke, "Laplace mixture autoregressive models," *Statistics and Probability Letters,* vol. 110, pp. 18-24, 2016.

[17] H. D. Nguyen and G. J. McLachlan, "Laplace mixture of linear experts," *Computational Statistics and Data Analysis 93,* vol. 93, pp. 177-191, 2016.

[18] B. Miftahurrohmah, N. Iriawan and K. Fithriasari, "On The Value at Risk Using Bayesian Mixture Laplace Autoregressive Approach for Modelling the Islamic Stock Risk Investment," *Journal of Physics: Conference Series,* vol. 855, no. 1, p. 1, 2017.

[19] P. H. Kupiec, "Techniques for verifying the accuracy of risk measurement models," *The Journal of Derivatives,* vol. 3, pp. 73-84, 1995.

[20] F. Iorgulescu, "Backtesting value-at-risk: Case study on the Romanian capital market," *Procedia - Social and Behavioral Sciences,* vol. 62, p. 796 – 800, 2012.

[21] R. Summinga-Sonagadu and J. Narsoo, "Risk Model Validation: An Intraday VaR and ES Approach Using the Multiplicative," *Risk,* vol. 7, no. 10, pp. 1-23, 2019.

[22] P. Jorion, Value at Risk The New Benchmark for Managing Financial Risk, 3 ed., New York: McGraw-Hill, 2011.

[23] A. S. Ahmar and E. B. d. Val, "SutteARIMA: Short-term forecasting method, a case: Covid-19 and stock," *Science of the Total Environment,* vol. 729, no. 138883, pp. 1-6, 2020.

[24] C. Proppe, "Markov Chain Monte Carlo Simulation Methods for Structural Reliability Analysis," *Procedia Engineering,* vol. 199, pp. 1122-1127, 2017.

[25] S. H. Cheung and S. Bansal, "A new Gibbs sampling based algorithm for Bayesian model updating with incomplete complex modal data," *Mechanical Systems and Signal Processing,* vol. 92, pp. 156-172, 2017.

[26] N. N. Zakaria, M. Othman, R. Sokkalingam, H. Daud, L. Abdullah and E. A. Kadir, "Markov Chain Model Development for Forecasting Air Pollution Index of Miri, Sarawak," *Sustainability,* vol. 11, no. 19, pp. 1-11, 2019.

[27] J. K. Kruschke, Doing Bayesian Data Analysis, Bloomington: Academic Press, 2015.

[28] D. J. Spiegelhalter, N. G. Best, B. P. Carlin and A. van der Linde, "Bayesian Measures of Model Complexity and Fit" (with discussion)," *Journal of the Royal Statistical Society,* vol. 64, pp. 583-639, 2002.

[29] J. C. Chan and A. L. Grant, "Fast computation of the deviance information criterion for latent variable models," *Computational Statistics and Data Analysis,* vol. 100, pp. 847-859, 2016.

[30] K. Dowd, "Retrospective Assessment of Value-at-Risk," *Risk Management: A Modern Perspective,* pp. 183-202, 2006.

[31] O. C. Ibe, Fundamentals of Applied Probability and Random Processes, 2nd ed., Lowell: Elsevier, 2014.

[32] N. P. Lemoine, "Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses," *Oikos,* vol. 128, p. 912–928, 2019.

[33] B. P. Charlin and S. Chip, "Bayesian Model Choice via Markov Chain Monte Carlo Methods," *Journal of the Royal Statistical Society, Series B (Methodological),* vol. 57, no. 3, pp. 473-484, 1995.

[34] H. Sanel and V. Mia, "Backtesting Value at Risk Forecast: The Case of Kupiec Pof-Test," *European Journal of Economic Studies,* vol. 17, no. 3, pp. 393-404, 2016.

[35] K. K. Gokmenoglu and N. Fazlollahi, "The Interactions among Gold, Oil, and Stock Market: Evidence from S&P 500," *Procedia Economics and Finance,* vol. 25, no. Special issue, pp. 478-488, 2015.