Communications for Statistical Applications and Methods
2023, Vol. 30, No. 2, 163–178
https://doi.org/10.29220/CSAM.2023.30.2.163
Print ISSN 2287-7843 / Online ISSN 2383-4757

GARCH-X(1, 1) model allowing a non-linear function of the variance to follow an AR(1) process

Didit B Nugroho[1,ab], Bernadus AA Wicaksono[c], Lennox Larwuy[c]

[a]Master's Program in Data Science, Universitas Kristen Satya Wacana, Indonesia;
[b]Study Center for Multidisciplinary Applied Research and Technology, Universitas Kristen Satya Wacana, Indonesia;
[c]Mathematics Study Program, Universitas Kristen Satya Wacana, Indonesia

Abstract

GARCH-X(1, 1) model specifies that conditional variance follows an AR(1) process and includes a past exogenous variable. This study proposes a new class from that model by allowing a more general (non-linear) variance function to follow an AR(1) process. The functions applied to the variance equation include exponential, Tukey's ladder, and Yeo–Johnson transformations. In the framework of normal and student-t distributions for return errors, the empirical analysis focuses on two stock indices data in developed countries (FTSE100 and SP500) over the daily period from January 2000 to December 2020. This study uses 10-minute realized volatility as the exogenous component. The parameters of considered models are estimated using the adaptive random walk metropolis method in the Monte Carlo Markov chain algorithm and implemented in the Matlab program. The 95% highest posterior density intervals show that the three transformations are significant for the GARCH-X(1, 1) model. In general, based on the Akaike information criterion, the GARCH-X(1, 1) model that has return errors with student-t distribution and variance transformed by Tukey's ladder function provides the best data fit. In forecasting value-at-risk with the 95% confidence level, the Christoffersen's independence test suggest that non-linear models is the most suitable for modeling return data, especially model with the Tukey's ladder transformation.

Keywords: ARWM, GARCH-X, non-linear transformations, student-t, value-at-risk, volatility

1. Introduction

The volatility of financial asset values (such as values of currency rates, futures interest rates, stock indices, or commodities) is often understood as a risk indicator for financial market participants and observers. The higher the volatility, the higher the risk for investors. Because volatility comes from the word 'volatile', which refers to asset prices condition that tends to fluctuate sharply and regularly.

One of the volatility models that change consistently over time (called heteroscedasticity) to express the volatility of the financial asset returns (the differences in the financial asset values over time) is GARCH (generalized autoregressive conditional heteroscedasticity) model proposed by Bollerslev

This publication resulted (in part) from research supported by: (1)The Ministry of Research, Technology and Higher Education of the Republic of Indonesia for the 2022 fiscal year under Contract Number 158/E5/PG.02.00.PT/2022, 001/LL6/PB/AK.04/2022 and (2)Universitas Kristen Satya Wacana under Contract Number 171/SPK-PDKN/PR V/5/2022.
[1] Corresponding author: Master's Program in Data Science, Universitas Kristen Satya Wacana, Jl. Diponegoro 52–60 Salatiga 50711, Indonesia. E-mail: didit.budinugroho@uksw.edu

Published 31 March 2023 / journal homepage: http://csam.or.kr

(1986). Due to its effectiveness in capturing the regular fluctuations behavior of financial data volatility, the popularity and use of GARCH-type models in the financial sector's risk has increased in recent years (Mahajan, 2022).

The model was then extended by Engle and Patton (2001) and Engle (2002) into GARCH-X by adding high-frequency data, such as realized volatility (RV) measures, as exogenous components in the volatility dynamics equation. Han (2015) had proven empirically that the exogenous component's parameters are significant, causing the model to be able to capture the slow decaying autocorrelation of squared returns. In addition, the GARCH-X(1, 1) model provides a better data fit than the GARCH(1, 1) model.

Recently, Nugroho *et al.* (2021) generalized the GARCH(1, 1) volatility process by applying non-linear transformations (including Tukey, Box–Cox, exponential, modulus, and Yeo–Johnson) for the volatility of lag one. They showed empirically that the proposed model has the possibility that one might be able to fit the data better than the untransformed volatility model. Meanwhile, most empirical studies have shown that financial asset returns generally do not meet the characteristics of a normal distribution but typically leptokurtic and produce heavy tail behaviour. One of the distributions that are popular and should always be employed in the practical finance research to accommodate heavy tails is the student-t distribution. For example, Gunay (2015), Braione and Scholtes (2016), and Sampid *et al.* (2018) employed student-t distribution for return errors to real datasets and found that GARCH(1, 1)-type models with student-t distribution fit data better than normal distribution.

Motivated by the above studies, this study applies non-linear power transformations to the conditional variance process for the GARCH-X(1, 1) model. The application was carried out for all lags simultaneously. Therefore, the main contribution of this study is to propose a new class that generalizes the GARCH(1, 1) and GARCH-X(1, 1) models, namely the non-linear GARCH(1, 1) (simply N-GARCH(1, 1)) and non-linear GARCH-X(1, 1) (simply N-GARCH-X(1, 1)) model, respectively. Empirically, the performance of the proposed model is tested using real data of the FTSE100 and SP500 stock indices for a daily period from January 2000 to December 2020, where the model parameters are estimated using the ARWM (adaptive random walk metropolis) method from Atchade and Rosenthal (2005). To our knowledge, this empirical study on the new class of non-linear GARCH-X models provides the first results in the literature.

2. Methodology

2.1. Type of volatility model

The GARCH model introduced by Bollerslev (1986) is the simplest model of the GARCH-type class. Many empirical applications mainly use the GARCH model with one lag (i.e., GARCH(1, 1) model) to formulate conditional variance directly as a function of observables. The performance of the GARCH(1, 1) model has been proven by Hansen and Lunde (2005), who compared 330 volatility GARCH models, and by Namugaya *et al.* (2014), who compared four GARCH(p, q) models.

In the daily period, the GARCH(1, 1) model interprets that today's variance is a function that depends on yesterday's variance and squared returns. By denoting R_t as returns at time t, σ_t as returns' conditional volatility at time t, and $\mathcal{N}$ as normal distributions, the GARCH(1, 1) model with normal distribution has an expression as follows:

$$\begin{cases} R_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1), \\ \sigma_t^2 = \omega + \alpha R_{t-1}^2 + \beta \sigma_{t-1}^2, \end{cases} \tag{2.1}$$

where the necessary conditions to ensure that σ_t^2 remains positive are given by $\omega, \alpha, \beta > 0$ and a necessary condition for covariance stationarity is $\alpha + \beta < 1$.

Engle (2002) extended the GARCH model into GARCH-X by adding high-frequency data as exogenous components in the conditional variance equation. The high-frequency data illustrate the intraday volatility of each trading day. Specifically, the GARCH-X(1, 1) model with normal distribution has the form of:

$$\begin{cases} R_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0.1), \\ \sigma_t^2 = \omega + \alpha R_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma x_{t-1}, \end{cases} \tag{2.2}$$

where the following conditions hold:

$$\omega, \alpha, \beta, \gamma > 0 \quad \text{and} \quad 0 < \alpha + \beta < 1.$$

In the above model, variable x_t represents the exogenous component.

This study extends the GARCH(1, 1) and GARCH-X(1, 1) models by transforming the conditional variance using a non-linear power function to make the models more general and allow non-linear variance to follow the AR(1) process. The GARCH(1, 1) and GARCH-X(1, 1) models, where the variances are transformed for all lags, are constructed consecutively as follows:

$$\text{N-GARCH}(1,1): h\left(\sigma_t^2, \lambda\right) = \omega + \alpha R_{t-1}^2 + \beta h\left(\sigma_{t-1}^2, \lambda\right). \tag{2.3}$$

$$\text{N-GARCH-X}(1,1): h\left(\sigma_t^2, \lambda\right) = \omega + \alpha R_{t-1}^2 + \beta h\left(\sigma_{t-1}^2, \lambda\right) + \gamma x_{t-1}, \tag{2.4}$$

where $h(\sigma_t^2, \lambda)$ represents the power transformation for variance σ_t^2 with transformation parameter λ. By taking $h(\sigma_t^2, \lambda)$ as a new variable h_t, then the model can be rewritten as:

$$R_t = \sqrt{g_t} \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, 1), \tag{2.5}$$

where $g_t = \sigma_t^2(h_t, \lambda)$, with the variance function is

$$h_t = \omega + \alpha R_{t-1}^2 + \beta h_{t-1} + \gamma x_{t-1}, \tag{2.6}$$

in which $\gamma = 0$ for N-GARCH(1, 1)-type models and $\gamma > 0$ for N-GARCH-X(1, 1)-type models.

2.2. Exogenous variable

In fact, empirical studies take realized measures of volatility as an exogenous variable as it is much more informative on the volatility level than the squared returns rate (Hansen *et al.*, 2012). Realized measures of volatility are theoretically high-frequency and are non-parametric estimators of variations in asset value changes when traded frequently (Floros *et al.*, 2020).

The first class of realized measures of volatility includes realized volatility (RV) measures introduced by Andersen *et al.* (2001), which are constructed as the root of the sum of intraday squared returns. Suppose that on day t, there is a set of returns $R_{t,1}, R_{t,2}, \ldots, R_{t,m}$ for a certain time range, then RV is formulated by:

$$\text{RV}_t = \sqrt{\sum_{j=1}^{m} R_{t,j}^2}\,. \tag{2.7}$$

In this study, the exogenous component x_t in the GARCH-X(1, 1) model takes RV 10 minutes. According to Liu *et al.* (2015) and Floros *et al.* (2020), RV 10 is a "good" estimator as it has high accuracy and a significant positive impact on daily returns and future volatility.

Table 1: The three transformation families and their inverses

Transf.	$h_t = h(\sigma_t^2, \lambda)$	$g_t = \sigma_t^2(h_t, \lambda)$
TL	$= \begin{cases} \sigma_t^{2\lambda}, & \lambda > 0 \\ \log\left(\sigma_t^2\right), & \lambda = 0 \end{cases}$	$= \begin{cases} h_t^{\lambda^{-1}}, & \lambda > 0 \\ \exp\{h_t\}, & \lambda = 0 \end{cases}$
Exp.	$= \begin{cases} \dfrac{\exp\{\lambda\sigma_t^2\} - 1}{\lambda}, & \lambda \neq 0 \\ \sigma_t^2, & \lambda = 0 \end{cases}$	$= \begin{cases} \dfrac{\ln(\lambda h_t + 1)}{\lambda}, & \lambda \neq 0 \\ h_t, & \lambda = 0 \end{cases}$
YJ	$= \begin{cases} \dfrac{(\sigma_t^2 + 1)^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log\left(\sigma_t^2 + 1\right), & \lambda = 0 \end{cases}$	$= \begin{cases} (1 + \lambda h_t)^{\lambda^{-1}}, & \lambda \neq 0 \\ \exp\{h_t\} - 1, & \lambda = 0 \end{cases}$

2.3. Power transformation family

Because this study transforms the variance parameter (σ_t^2), which value is always positive, there are three non-linear power transformations applied, namely Tukey's ladder, exponential, and Yeo–Johnson transformations. Table 1 gives the non-linear transformations for the variable σ_t^2 and their inverses.

Manly (1976) proposed an exponential (simply Exp) transformation that allows both positive and negative response values. The idea is to take the natural exponential of the response variable and then apply the transformation of Box and Cox (1964). The exponential transformation is useful for changing a unimodal skew distribution into a normal distribution.

Tukey's ladder (simply TL) of power transformation introduced by Tukey (1977) is the simplest power transformation and aims to simplify data analysis. This way changes the asymmetric distribution so that it becomes normal or nearly-normal and can also help to reduce error variability.

Yeo and Johnson (2000) proposed power transformation that can be used without limitation of the response variable and is suitable to reduce the distribution skew. For the positive variable, the Yeo–Johnson (simply YJ) transformation is identical to the Modulus transformation introduced by John and Draper (1980), thus also equal to Box–Cox transformation if the response variable is susbtituted with the variable minus one.

2.4. Distributional assumptions

The most widely and commonly used distribution as an initial study is the normal distribution. Therefore, as a framework for easy initial work, the normal distribution is also considered here. A type of probability distribution that is similar to the normal distribution is student-t distribution, where the curve is symmetrically bell-shaped but has thicker tails. The student-t distribution tends to be close to the normal distribution when the degree of freedom heads towards infinity (in this case ν greater than or equal to 30), see Ramachandran and Tsokos (2021).

The likelihood function for a normal random variable Y with mean μ and variance ϕ, denoted as $Y \sim \mathcal{N}(\mu, \phi)$, is defined as follows:

$$L(y \mid \mu, \phi) = \frac{1}{\sqrt{2\pi\phi}} \exp\left\{-\frac{(y-\mu)^2}{2\phi}\right\}. \tag{2.8}$$

for $-\infty < y < \infty$, with the log-likelihood function $\mathscr{L}(y|\mu, \phi)$ expressed by:

$$\mathscr{L}(y \mid \mu, \phi) = -\frac{1}{2}\log(2\pi\phi) - \frac{(y-\mu)^2}{2\phi}. \tag{2.9}$$

Meanwhile, the random variable Y that follows a student-t distribution with mean μ, variance ϕ, and degree of freedom ν (not required to be an integer), denoted as $Y \sim \mathcal{T}(\mu, \phi, \nu)$, has a likelihood function expressed by (see, Choi and Yoon (2020)):

$$L(y \mid \mu, \phi, \nu) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi\phi(\nu-2)}\Gamma(\nu/2)}\left(1 + \frac{(y-\mu)^2}{\phi(\nu-2)}\right)^{-\frac{1+\nu}{2}} \tag{2.10}$$

for $-\infty < y, \mu < \infty$, $\phi > 0$, and $\nu > 2$. In the above function, $\Gamma(\cdot)$ represents the gamma function. Therefore, the log-likelihood function $\mathscr{L}(y|\mu, \phi, \nu)$ is expressed by:

$$\mathscr{L}(y \mid \mu, \phi, \nu) = \log\Gamma\left(\frac{\nu+1}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log(\pi\phi(\nu-2)) - \frac{1+\nu}{2}\log\left(1 + \frac{(y-\mu)^2}{\phi(\nu-2)}\right). \tag{2.11}$$

2.5. Estimation method

The observed models are estimated using an adaptive method called ARWM and performed in the MCMC algorithm. According to Van Ravenzwaaij *et al.* (2018), MCMC is an increasingly popular method in Bayesian inference when the posterior distribution is difficult to perform via analytic examination. The MCMC algorithm combines two properties: Markov chain and Monte Carlo. The idea is to generate a series of random samples that have the "Markov" property so that the stationary distribution is the same as the estimated posterior distribution and to estimate the distribution properties by testing random samples using the Monte Carlo approach. Therefore, the MCMC has two stages, namely generating the Markov chain and estimating the distribution properties (such as mean, standard deviation, and Bayesian interval) based on the Monte Carlo approach.

The ARWM method proposed by Atchade and Rosenthal (2005) is one of the Markov chain generation methods, which is an improvement from the random walk metropolis (RWM) method, the most often used method because it is the simplest. Nugroho (2018) showed empirically that the ARWM method has high computational and convergence speeds. In summary, the steps of the ARWM method can be explained as follows.

At iteration n in the MCMC algorithm, the candidate of random sample θ is generated by the equation:

$$\theta^{(n)} = \theta^{(n-1)} + \sqrt{\Delta^{(n)}}z^{(n)}, \quad z^{(n)} \sim N(0, 1), \tag{2.12}$$

where $\Delta^{(n)}$ is the step width. Bayesian inference uses information from the observation data D given θ, formally called likelihood, to update the parameter prior state into a posterior one. Formally, the natural logarithm of the posterior distribution or probability of θ given D, is calculated using Bayes' rule (see, Yang (2019)):

$$\log p(\theta \mid D) \propto \mathscr{L}(D \mid \theta) + \log p(\theta), \tag{2.13}$$

where $\mathscr{L}(D|\theta)$ indicates the log-likelihood or log-probability of D given θ, and $p(\theta)$ indicates the prior probability of θ. The symbol $\propto$ means "proportional to". The candidate sample $\theta^{(n)}$ is accepted if $(p(\theta^{(n)}|D))/p(\theta^{(n-1)}|D) > u$, for a standard uniformly distributed random variable u. Next, the step width $\Delta^{(n)}$ is going to change adaptively based on the formula (Atchade and Rosenthal, 2005; Andrieu and Thom, 2008):

$$\Delta^{(n)} = \Delta^{(n-1)} + \eta_n \left(\frac{N}{n} - r \right) \tag{2.14}$$

with the desire optimal acceptance rate r is fixed at 0.234 or 0.44 (Rosenthal, 2011). Here, N represents the number of accepted candidates after n iterations of the MCMC chain. In other words, N/n represents the acceptance probability of last candidate. Assuming monotocity on N/n, if $N/n - r > 0$ then Δ_{n-1} is probably too small and should be increased while if $N/n - r < 0$ then Δ_{n-1} should be decreased. To control how quickly the impact of the tuning mechanism decays such that the variations of $\{\Delta_n\}$ vanish, the standard approach chooses a positive sequence of real number $\{\eta_n\}$ deterministic and non-increasing. A sequence that satisfies these conditions is $\eta_n = n^{-\delta}$ for some constant $0.5 < \delta \leq 1$ (called averaging or damping parameter).

After the Markov chain of parameter θ is constructed, the next step is to calculate descriptive statistics from random samples as the output of the MCMC. The descriptive statistics include mean, standard deviation, and Bayesian confidence interval. This study chose the HPD (highest posterior density) interval as the Bayesian confidence interval following the Chen & Shao approach in Le *et al.* (2020) and Nugroho *et al.* (2021). The 95% HPD interval from the Markov chain with length M is constructed with the following steps:

Step 1 : Calculate $M_{\text{cut}} = [0.05 \times M]$ and $M_{\text{span}} = M - M_{\text{cut}}$, where $[x]$ represents the standard rounding function of x.

Step 2 : Sort the estimated values from the smallest to the largest, i.e. $\{\theta_j\}_{j=1}^{M}$, where $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_M$.

Step 3 : Find the index j^* so that $\theta_{j^*+M_{\text{span}}} - \theta_{j*} = \min\limits_{1 \leq j \leq M_{\text{cut}}} (\theta_{j+M_{\text{span}}} - \theta_j)$.

Step 4 : Determine the 95% HPD interval:

$$\left(\theta_{j^*}, \theta_{j^*+M_{\text{span}}} \right). \tag{2.15}$$

2.6. Model selection

The goodness-of-fit of the studied models to real data is evaluated through the Akaike information criterion (AIC). This criterion is adequate for both nested and non-nested models. It should be noted that the models considered in this study are non-nested, meaning that a model is not a special case from another model for a certain parameter value. For the statistical model with dimension k (the number of parameter), the AIC is defined by (see, Portet (2020)):

$$\text{AIC} = 2\,(k - \mathscr{L})\,, \tag{2.16}$$

where $\mathscr{L}$ represents the maximum value of the log-likelihood function. In principle, given a set of competing models for the same data, the model with the minimum AIC values indicates the best fit model.

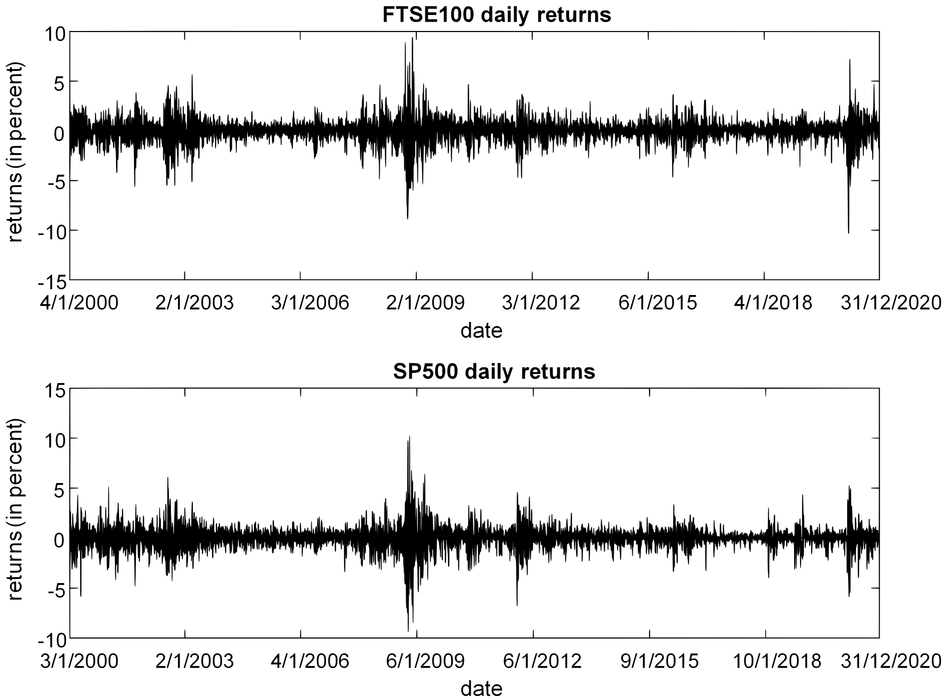

Figure 1: *Time plots of daily return of all the indices.*

3. Applications on real data

This section applies the models and estimation method to real data to empirically analyze the performance of sampling method, parameter estimation, and goodness-of-fit of the models.

3.1. The descriptions of observation data

The data used in our investigation are the financial times stock exchange 100 stock indices, also called the FTSE100, and the standard and poor's 500, or simply the SP500. Both stock indices were chosen in this study because according to Chaudhary *et al.* (2020), both of them are part of the top 10 stock market indices that accounted for 66% of gross domestic product (GDP) as the world economy reference. Therefore, both stock indices can represent the world economy. Specifically, this study used the daily stock indices values from January 2000 to December 2020. All data were obtained from the Oxford-Man Institute's "realized library" at https://realized.oxford-man.ox.ac.uk/data/.

Figure 1 presents time plots of daily returns of FTSE100 (top) and SP500 (bottom), where both returns are observed to fluctuate around the mean (red line). This visually interprets that the returns are stationary, or means that there is no up or down trend during the observation period. This is the reason why financial studies are more interested in returns than asset prices, because stationary data make statistical analysis easier.

Meanwhile, Table 2 depicts descriptive statistics for daily returns from each observation dataset. The average value of each stock index return is close to zero so the formulation of the returns equation which uses a mean of zero is appropriate. The kurtosis value clearly indicates that both returns are leptokurtic (heavy tails), a finding found in many empirical studies in the literature. The probabil-

Table 2: Descriptive statistics

Particulars	FTSE100	SP500
Mean	–0.0060	0.0065
Standard deviation	1.166	1.128
Kurtosis	10.22	11.33
JB statistics	1.16×10^4	1.52×10^4
Probability of JB	0	0
Observations	5291	5258

ity value of the computed Jarque–Bera (JB) is less than 0.05 in both returns series. This normality test confirms that the null hypothesis of normality is rejected at the 5% significance level, meaning that both return series are not normally distributed. For this reason, the considered volatility models are estimated with a student-t distribution framework, which is a type of heavy-tailed distribution. Since this study mainly focused on the extension of the GARCH(1, 1) and GARCH-X(1, 1) models to capture nonlinearity in volatility process, the models can be extensively enhanced in future work by joining the stylized facts of heavy-tails and skewness.

Construction of log-likelihood functions for models

The log-likelihood functions in Equations (2.9) and (2.11) need to be rewritten for each considered model. By applying non-linear power transformation for variances on all lags, the NL-GARCH(1, 1) and NL-GARCH-X(1, 1) models, where return errors follow a normal distribution with a mean of zero and variance of g_t, have the following log-likelihood function:

$$\mathscr{L}\left(R_t \mid g_t\right) = -\frac{1}{2}\log\left(2\pi g_t\right) - \frac{R_t^2}{2g_t}, \tag{3.1}$$

where $g_t = \sigma_t^2(h_t, \lambda)$ is given as in Table 1 with the process of h_t follows Equation (2.6).

Furthermore, when the return error ε_t follows student-t distribution with a mean of zero, variance of g_t, and degrees of freedom $\nu > 2$, the log-likelihood functions of the NL-GARCH(1, 1) and NL-GARCH-X(1, 1) models are given as follows:

$$\mathscr{L}\left(R_t \mid g_t, \nu\right) = \log\Gamma\left(\frac{\nu+1}{2}\right) - \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log\left(\pi g_t\left(\nu-2\right)\right) - \frac{1+\nu}{2}\log\left(1 + \frac{R_t^2}{g_t\left(\nu-2\right)}\right). \tag{3.2}$$

3.2. MCMC implementation

Following the Bayes' formula in Equation (2.13), a prior distribution needs to be set for each model parameter. This study specifies a truncated normal distribution with a mean of zero and a variance of 10 for the prior of $\omega, \alpha, \beta, \lambda, \gamma$ and an exponential distribution with hyper-parameter 0.01 for the prior of ν according to the practical standards (see, Nugroho *et al.* (2021)). Then, the prior distribution is combined with the likelihood function to form a posterior distribution.

To run the MCMC algorithm, the initial values for the parameters were set as follows:

$$\omega^{(0)} = 0.01, \quad \alpha^{(0)} = 0.2, \quad \beta^{(0)} = 0.7, \quad \gamma^{(0)} = 0.1, \quad \lambda^{(0)} = 0.5, \quad \nu^{(0)} = 10.$$

The number of MCMC iterations was 6000, where the first 1000 iterations were discarded as burn-in to eliminate the non-stationary caused by the arbitrary selection of the initial values. In the ARWM method, the initial step width was chosen to be 0.005. By setting the optimal acceptance rate $r = 0.44$

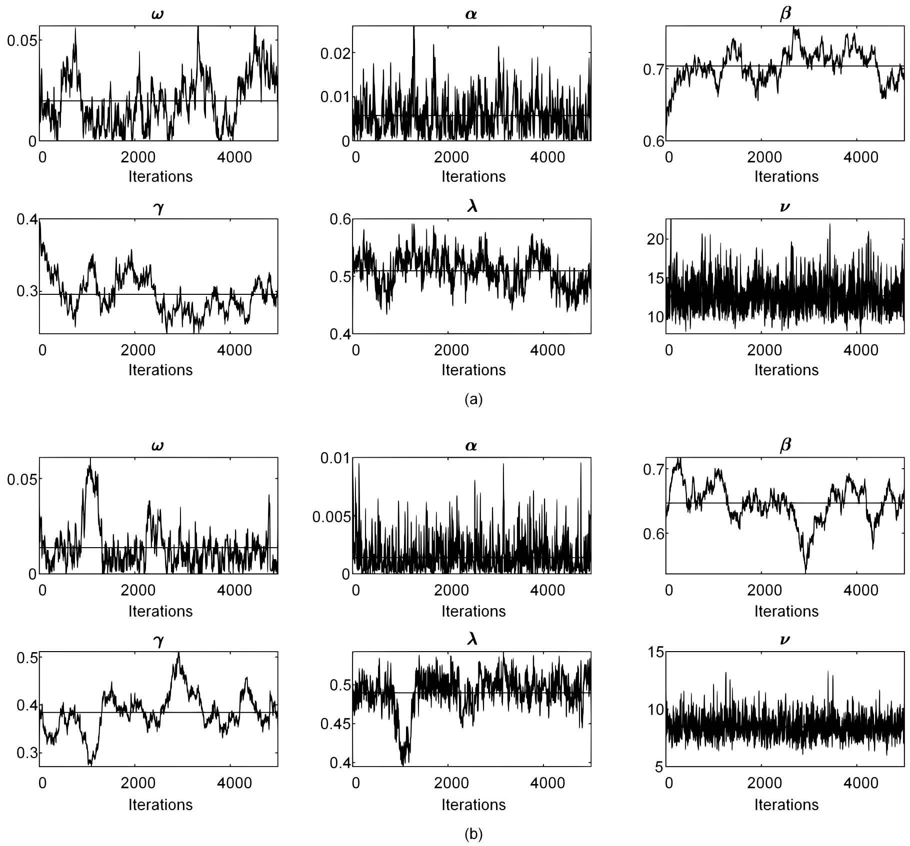

Figure 2: *Trace plots of the MCMC chain for the parameters of ST-GARCHt-X(1, 1) model generated by the ARWM method: (a) FTSE100 and (b) SP500.*

and damping parameter $\delta = 0.66$, our experimental results demonstrate that the optimal acceptance rate has been achieved by all parameters about 0.44. The only exception is the parameter ω which achieves about 0.234.

For example, the MCMC stages to estimate the NL-GARCH-X(1, 1) model with student-t distribution, or simply the N-GARCHt-X(1, 1), which is applied to data returns $\mathbf{R} = \{R_t\}_{t=1}^T$ and $\mathbf{X} = \{\mathrm{RV}_t\}_{t=1}^T$, are summarized as follows.

Stage 1 : Generate Markov chains from the posterior probability by using the ARWM method.

1. Draw $\omega|\alpha, \beta, \gamma, \lambda, \nu, \mathbf{R}, \mathbf{X}$.
2. Draw $\alpha|\omega, \beta, \gamma, \lambda, \nu, \mathbf{R}, \mathbf{X}$.
3. Draw $\beta|\omega, \alpha, \gamma, \lambda, \nu, \mathbf{R}, \mathbf{X}$.
4. Draw $\gamma|\omega, \alpha, \beta, \lambda, \nu, \mathbf{R}, \mathbf{X}$.
5. Draw $\lambda|\omega, \alpha, \beta, \gamma, \nu, \mathbf{R}, \mathbf{X}$.
6. Draw $\nu|\omega, \alpha, \beta, \gamma, \lambda, \mathbf{R}, \mathbf{X}$.

Stage 2 : Calculate statistics.

Table 3: Mean and standar deviation (in bracket) of the parameter estimates on GARCH(1, 1)-type models construction process for FTSE100 index

Parameter	Model			
	GARCH	Exp-GARCH	TL-GARCH	YJ-GARCH
ω	0.0191 (0.0028)	0.0175 (0.0035)	0.0205 (0.0059)	0.0179 (0.0034)
α	0.1152 (0.0082)	0.1205 (0.0121)	0.1146 (0.0102)	0.1194 (0.0114)
β	0.8708 (0.0082)	0.8713 (0.0126)	0.8713 (0.0087)	0.8720 (0.0117)
λ		0.0207 (0.0214)	0.9788 (0.0563)	1.0436 (0.0624)
	GARCH*t*	Exp-GARCH*t*	TL-GARCH*t*	YJ-GARCH*t*
ω	0.0165 (0.0031)	0.0158 (0.0034)	0.0187 (0.0051)	0.0163 (0.0032)
α	0.1094 (0.0092)	0.1062 (0.0120)	0.1032 (0.0098)	0.1029 (0.0100)
β	0.8791 (0.0097)	0.8829 (0.0119)	0.8819 (0.0094)	0.8832 (0.0095)
λ		0.0060 (0.0236)	0.9632 (0.0528)	0.9626 (0.0837)
ν	9.46 (1.21)	9.53 (1.22)	9.51 (1.23)	9.47 (1.18)
	GARCH-X	Exp-GARCH-X	TL-GARCH-X	YJ-GARCH-X
ω	0.0011 (0.0011)	0.0009 (0.0009)	0.0221 (0.0135)	0.0008 (0.0008)
α	0.1169 (0.0106)	0.1081 (0.0120)	0.0074 (0.0048)	0.0286 (0.0092)
β	0.8057 (0.0204)	0.7966 (0.0225)	0.6948 (0.0198)	0.7161 (0.0283)
γ	0.0905 (0.0167)	0.1028 (0.0175)	0.5140 (0.0286)	0.1843 (0.0229)
λ		−0.0387 (0.0105)	0.3006 (0.0202)	0.1390 (0.0808)
	GARCH*t*-X	Exp-GARCH*t*-X	TL-GARCH*t*-X	YJ-GARCH*t*-X
ω	0.0012 (0.0012)	0.0012 (0.0013)	0.0253 (0.0146)	0.0009 (0.0009)
α	0.1077 (0.0103)	0.0976 (0.0103)	0.0052 (0.0038)	0.0294 (0.0106)
β	0.8275 (0.0132)	0.8257 (0.0152)	0.7039 (0.0222)	0.7535 (0.0199)
γ	0.0758 (0.0113)	0.0821 (0.0137)	0.2899 (0.0289)	0.1580 (0.0175)
λ		−0.0365 (0.0241)	0.4978 (0.0315)	0.1802 (0.0978)
ν	10.43 (1.44)	10.66 (1.51)	12.55 (2.02)	11.31 (1.77)

3.3. Parameter estimates

Based on the results of running the MCMC algorithm, the remaining 5000 iterations were used to calculate the posterior mean, standard deviation, and 95% HPD interval. For example, Figure 2 displays the plot of the estimated values (Markov chain) for parameters in the GARCH-X(1, 1) model with student-t distribution, where the variance transformed by the TL function follows the AR(1) process. Visually, the results show that the ARWM method is very efficient for estimating parameters α and ν, but less efficient for other parameters. Even so, the stationarity for the estimated values of ω, β, γ, and λ are considered good enough. Moreover, the convergence of algorithm were checked by using the integrated autocorrelation time (IACT). This IACT value, denoted by τ, is estimated by using Sokal's adaptive truncated periodogram estimator (see, Lampart and Sbalzarini (2012)) with the results as follows:

$$\begin{aligned}\text{FTSE100}: &\ \tau(\omega) = 148.7,\ \ \tau(\alpha) = 34.3,\ \ \tau(\beta) = 115.5,\ \ \tau(\gamma) = 125.4,\\ &\ \tau(\lambda) = 136.8,\ \ \tau(\nu) = 6.6,\\ \text{SP500}: &\ \tau(\omega) = 89.0,\ \ \tau(\alpha) = 9.3,\ \ \tau(\beta) = 137.9,\ \ \tau(\gamma) = 141.3,\\ &\ \tau(\lambda) = 72.6,\ \ \tau(\nu) = 7.3.\end{aligned}$$

The results indicate that the ARWM sampling method in MCMC provides quite efficiency and reasonably good mixing performance, which yield quite fast convergence. Therefore, it recommends to make reliable inferences on the interest parameters for the next analysis.

First, the effect of the non-linear power transformation on the parameters in basic model is analyzed. This is seen from the mean and standard deviation of the estimates over the MCMC algorithm

Table 4: Mean and standar deviation (in bracket) of the parameter estimates on GARCH(1, 1)-type models construction process for SP500 index

Parameter	Model			
	GARCH	Exp-GARCH	TL-GARCH	YJ-GARCH
ω	0.0152 (0.0023)	0.0152 (0.0021)	0.0135 (0.0027)	0.0145 (0.0018)
α	0.1188 (0.0101)	0.1274 (0.0094)	0.1269 (0.0087)	0.1240 (0.0078)
β	0.8701 (0.0102)	0.8658 (0.0090)	0.8666 (0.0079)	0.8701 (0.0074)
λ		0.0212 (0.0206)	1.0306 (0.0397)	1.0581 (0.0552)
	GARCH*t*	Exp-GARCH*t*	TL-GARCH*t*	YJ-GARCH*t*
ω	0.0090 (0.0019)	0.0093 (0.0018)	0.0097 (0.0025)	0.0093 (0.0020)
α	0.1161 (0.0097)	0.1174 (0.0088)	0.1149 (0.0096)	0.1157 (0.0111)
β	0.8808 (0.0097)	0.8794 (0.0086)	0.8814 (0.0093)	0.8797 (0.0103)
λ		−0.0012 (0.0167)	0.9861 (0.0395)	0.9579 (0.0622)
ν	6.68 (0.55)	6.65 (0.58)	6.64 (0.59)	6.58 (0.57)
	GARCH-X	Exp-GARCH-X	TL-GARCH-X	YJ-GARCH-X
ω	0.0005 (0.0005)	0.0005 (0.0006)	0.0168 (0.0117)	0.0005 (0.0006)
α	0.1171 (0.0099)	0.1082 (0.0098)	0.0016 (0.0015)	0.0090 (0.0039)
β	0.7906 (0.0178)	0.7863 (0.0165)	0.6130 (0.0337)	0.5864 (0.0420)
γ	0.1042 (0.0145)	0.1116 (0.0138)	0.4154 (0.0449)	0.2929 (0.0333)
λ		−0.0359 (0.0082)	0.5115 (0.0237)	0.0121 (0.0115)
	GARCH*t*-X	Exp-GARCH*t*-X	TL-GARCH*t*-X	YJ-GARCH*t*-X
ω	0.0007 (0.0007)	0.0006 (0.0006)	0.0173 (0.0150)	0.0006 (0.0006)
α	0.1217 (0.0128)	0.1168 (0.0114)	0.0015 (0.0015)	0.0093 (0.0043)
β	0.8252 (0.0192)	0.8207 (0.0162)	0.6299 (0.0405)	0.5790 (0.0396)
γ	0.0665 (0.0146)	0.0724 (0.0132)	0.3988 (0.0388)	0.3029 (0.0322)
λ		−0.0260 (0.0112)	0.4874 (0.0234)	0.0133 (0.0153)
ν	6.97 (0.65)	7.07 (0.69)	8.49 (0.96)	7.50 (0.87)

presented in Tables 3 and 4. From the estimated parameter values, it is necessary to see the difference between the results in the basic model and the transformation models. The difference is calculated based on the relative difference with the formula $|(\theta^{\text{basic}} - \theta^{\text{transf}})/\theta^{\text{basic}}| \times 100\%$ for the parameter θ. Intuitively, it is determined that if the relative difference is less than 25%, then the difference is small and the transformation is considered not to affect parameter θ. Due to space limitation, the results for the relative difference are not presented.

The comparison between the basic GARCH(1, 1) model and the Exp-transformed models generally shows that the estimated values have small differences, which are less than 9% for the Exp-GARCH(1, 1) model and less than 19% for the Exp-GARCH-X(1, 1) model on all parameters. For the TL-transformed models, the estimation differences are small (less than 14%) on all parameters for the TL-GARCH(1, 1) model but large (greater than 25%) for the TL-GARCH-X(1, 1) model, except for parameters β and ν. Finally, compared to the YJ-transformed models, the estimation differences are small (less than 7%) on all parameters for the YJ-GARCH(1, 1) model but large (greater than 25%) for the YJ-GARCH-X(1, 1) model, except for parameters β (FTSE100 only) and ν. The results from Tables 3 and 4 conclude that the non-linear power transformation fairly affects the parameters of the GARCH-X(1, 1) model, except for parameters β and ν.

The second analysis is conducted on the significance of the transformation parameter λ. It is investigated whether the HPD interval contained the basic value as described in the previous section, namely $\lambda = 0$ for the Exp transformation and $\lambda = 1$ for the TL and YJ transformations. Table 5 presents the 95% HPD interval for parameter λ of all transformed models.

Table 5 shows that the three transformations are not significant in the GARCH(1, 1)-type models. This is indicated by the 95% HPD intervals containing the value of 0 for the exp-transformed model

Table 5: HPD intervals at 95% confidence level for the transformation parameter λ in the non-linear models

Model	FTSE100		SP500	
	Normal	Student-t	Normal	Student-t
Exp-GARCH	(−0.0207,0.0636)	(−0.0373,0.0500)	(−0.0163,0.0613)	(−0.0283,0.0327)
TL-GARCH	(0.8627,1.0734)	(0.8434,1.0636)	(0.9518,1.1032)	(0.9097,1.0631)
YJ-GARCH	(0.9311,1.1674)	(0.7956,1.1171)	(0.9463,1.1603)	(0.8407,1.0931)
Exp-GARCH-X	(–0.0554,–0.0174)	(–0.0591,–0.0087)	(–0.0496,–0.0200)	(–0.0432,–0.0039)
TL-GARCH-X	(0.4565,0.5667)	(0.4414,0.5574)	(0.3353,0.5008)	(0.4402,0.5301)
YJ-GARCH-X	(0.0001,0.2884)	(0.0037,0.3658)	(0.2428,0.3649)	(0.0001,0.0390)

Table 6: Comparison of alternative models

Model	FTSE100				SP500			
	Normal		Student-t		Normal		Student-t	
	AIC	Rank	AIC	Rank	AIC	Rank	AIC	Rank
GARCH	14408.77	5	14308.11	5	13561.66	8	13309.78	5
Exp-GARCH	14410.20	7	14311.02	8	13561.62	7	13312.54	6
TL-GARCH	14410.92	8	14310.05	6	13561.39	6	13312.66	7
YJ-GARCH	14409.96	6	14310.13	7	13560.44	5	13312.82	8
GARCH-X	14343.64	4	14263.59	4	13446.41	4	13261.66	4
Exp-GARCH-X	14338.31	3	14260.55	3	13438.07	3	13258.47	3
TL-GARCH-X	14190.14	1	14135.01	1	13132.21	1	12997.34	1
YJ-GARCH-X	14280.30	2	14218.01	2	13299.23	2	13165.28	2

and the value of 1 for the TL- and YJ-transformed models. However, when the non-linear power transformation is applied to the GARCH-X(1, 1) model, the result is significant. This is indicated by intervals that exclude basic values. These results show that the presence of RV component greatly affects the significance of non-linear power transformation. Furthermore, these results indicate that the conditional variance process in the GARCH-X(1, 1) model needs to be non-linearly transformed.

3.4. Model evaluation

To choose a model that provides the best data fit, the AIC values from the competing models are observed in Table 6. It is observed that models with smaller AIC values indicate better models.

When the transformation is applied to the GARCH(1, 1)-type models, the basic GARCH(1, 1) model provides a better fit. There is an exception to the SP500 data fit using the normal distribution, where the transformation models are superior with extremely small AIC differences. In contrast, when the transformation is applied to the GARCH-X(1, 1)-type models, all transformation models provide a better fit than the basic model in all data cases and distribution specifications. These results are consistent with the significance of the transformation parameters obtained in the previous section. In this case, the TL-transformed model is the best, followed by the YJ- and exp-transformed model.

Furthermore, when the results of the two distributions are compared, the model with the student-t distribution is superior to the normal distribution in all models. Overall, the AIC chooses the TL-GARCH-X(1, 1) model with student-t distribution as the best fit model.

3.5. Value-at-risk estimation

The most popular market risk measure is value-at-risk (VaR), which quantities the possible financial losses for a portfolio, investment, or entity over a specified time period and for a given confidence level. Since volatility estimation is a key input to VaR (value-at-risk) models, the selection of the appropriate volatility model is one of the most important factors in determining the accuracy of VaR.

Table 7: Estimation and testing windows in dates

t	Estimation window			Testing window		
	Start	$t + W_E - 1$	End	VaR(Obs.)	$t + W_E$	VaR forecast
1	4/1/2000	5039	31/12/2019	VaR(1)	5040	VaR(2/1/2020)
2	5/1/2000	5040	2/1/2020	VaR(2)	5041	VaR(3/1/2020)
⋮	⋮	⋮	⋮	⋮	⋮	⋮
252	29/12/2000	5290	30/12/2020	VaR(252)	5291	VaR(31/12/2020)

Table 8: Results of backtesting VaR for the CCI test at the 5% risk level using FTSE100 and SP500

Model	FTSE100				SP500			
	Normal		Student-t		Normal		Student-t	
	LR	p-value	LR	p-value	LR	p-value	LR	p-value
GARCH	0.61	0.44	0.61	0.44	0.009	0.93	0.03	0.86
Exp-GARCH	0.61	0.44	0.61	0.44	0.002	0.97	0.03	0.86
TL-GARCH	0.61	0.44	0.61	0.44	0.002	0.97	0.002	0.97
YJ-GARCH	0.61	0.44	0.61	0.44	0.009	0.93	0.002	0.97
GARCH-X	0.23	0.63	0.23	0.63	1.95	0.16	0.03	0.86
Exp-GARCH-X	0.40	0.53	0.40	0.53	1.95	0.16	0.03	0.86
TL-GARCH-X	0.12	0.74	0.04	0.84	1.95	0.16	2.23	0.14
YJ-GARCH-X	0.12	0.74	0.04	0.84	2.23	0.14	0.10	0.76

According to the distribution of errors, when the mean return is zero, the 1-day horizon VaR for an individual asset with probability $p\%$ can be calculated as follows (Xu, 2017):

$$\begin{aligned} &\text{under the normal distribution: } \mathrm{VaR}_p = Z_p \hat{\sigma}_{t+1}, \\ &\text{under the student-}t\text{ distribution: } \mathrm{VaR}_{p,\nu} = \sqrt{\nu^{-1}(\nu-2)}\mathcal{T}^{-1}_{\nu,p}\hat{\sigma}_{t+1}, \end{aligned} \tag{3.3}$$

where Z_p denotes the p-quantile of a standard normal distribution and $\mathcal{T}^{-1}_{\nu,p}$ represents the p-quantile of a student-t distribution. When implementing VaR predictions from a GARCH-X type model, one needs to calculate $\hat{\sigma}_{t+1}$ given the last volatility estimate $\hat{\sigma}_t$ and the parameter estimate vector.

In order to measure the accuracy of VaR estimates, the backtesting is applied to the VaR forecast for to the year 2020. We estimate each model of VaR using the estimation window size, W_E, of 5039 trading days starting from January 4^{th}, 2000 onwards with the beginning of the estimation period is on day t. The testing window for the year 2020 starts from its first day January 2^{nd}, 2020 to its last day December 31^{st}, 2020, having a total of 252 observations. Every observation in the testing window corresponds to a day for which the VaR is calculated using the previous 5039 trading days. The calculation procedure of estimation and testing windows is shown in Table 7.

In this study, the conditional coverage independence (CCI) test of Christoffersen (1998) is used to evaluate the accuracy of VaR models at the daily horizon. This test, which seeks to determine whether the VaR breaches are independent from one period to the next or not, is one of the most well-known backtesting tests. The Christoffersen's CCI test has the advantage that it respects the conditionality in the volatility forecasts. Table 8 presents the outputs for the Christoffersen's CCI test at the significance level of 5%. As a result, in any case, all VaR models passes the Christoffersen's CCI test, indicated by likelihood ratio (LR) less than < 5.99 (p-value > 0.5), and therefore the accuracy of all VaR models is accepted. In case of adopting FTSE100 index data, the best performing models with the lowest LR value are the TL-GARCH-X and YJ-GARCH-X models in each distribution specification. Differents results are given by the adoption of SP500 index, which show that the best VaR estimate is provided by Exp-GARCH and TL-GARCH in the normal distribution case and by TL-GARCH and YJ-GARCH in the student-t distribution case.

4. Conclusions and future works

This study proposed an extension of the GARCH(1, 1) and GARCH-X(1, 1) models by applying non-linear power transformation for the variance equation, namely non-linear transformation of variance following the AR(1) process. These non-linear power transformations include exponential, Tukey's ladder, and Yeo–Johnson transformations. In this study, the return errors in the model were assumed to follow normal and student-t distributions. The data used for empirical analysis were FTSE100 and SP500 stock indices data in a daily period from 2000 to 2020. The method used to estimate the model parameters was the ARWM method in the MCMC algorithm.

The empirical results showed that the non-linear power transformation greatly affects the parameter estimate of the ARCH components as well as the exogenous variable component. Significantly, the non-linear power transformation needs to be applied to the GARCH-X(1, 1) model. In particular, the AIC provides evidence that the GARCH-X(1, 1) model with student-t distribution, where the variance equation is transformed by Tukey's ladder, provides the best fit. Therefore, this study found a more general GARCH-X(1, 1) model which potentially provides a better data fit than the basic model.

Furthermore, this study evaluates VaR estimates produced by eight GARCH-type models, made under symmetric (normal dan student-t) error distributions. In the Christoffersen's independence test, the results for VaR estimation with the 95% confidence interval support all models and confirm that the univariate GARCH VaR models allowing a non-linear function of variance to follow an AR(1) process are among the best performing models.

For the next study, the model can be developed by applying non-linear power transformation to a realized GARCH model as in Hansen *et al.* (2012) and Hansen and Huang (2016). It would be a good idea to increase the number of models and evaluate whether an asymmetric GARCH model and a skewed student-t distribution would produce better results than the models and distributions explored in this study.

References

Andrieu C and Thoms J (2008). A tutorial on adaptive MCMC, *Statistics and Computing*, **18**, 343–373.

Andersen TG, Bollerslev T, Diebold FX, and Labys P (2001). The distribution of realized exchange rate volatility, *Journal of the American Statistical Association*, **96**, 42–55, Available from: https://doi.org/ 10.1198/016214501750332965

Atchade YF and Rosenthal JS (2005). On adaptive Markov chain Monte Carlo algorithms, *Bernoulli*, **11**, 815–828, Available from: https://doi.org/10.3150/bj/1130077595

Box GEP and Cox DR (1964). An analysis of transformations, *Journal of the Royal Statistical Society. Series B (Methodological)*, **26**, 211–252, Available from: https://doi.org/10.1111/j.2517-6161.1964. tb00553.x

Braione M and Scholtes N (2016). Forecasting value-at-risk under different distributional assumptions, *Econometrics*, **4**, 1–27, Available from: https://doi.org/10.3390/econometrics4010003

Chaudhary R, Bakhshi P, and Gupta H (2020). Volatility in international stock markets: An empirical study during COVID-19, *Journal of Risk and Financial Management*, **13**, 1–17, Available from: https://doi.org/ 10.3390/jrfm13090208

Choi SY and Yoon JH. (2020). Modeling and risk analysis using parametric distributions with an application in equity-linked securities, *Mathematical Problems in Engineering*, **2020**, 1–20, Avail-

able from: https://doi.org/10.1155/2020/9763065
Christoffersen PF (1998). Evaluating interval forecasts, *International Economic Review*, **39**, 841–862, Available from: https://doi.org/10.2307/2527341
Engle R (2002). New frontiers for ARCH models, *Journal of Applied Econometrics*, **17**, 425–446, Available from: https://doi.org/10.1002/jae.683
Engle RF and Patton AJ (2001). What good is a volatility model?, *Quantitative Finance*, **1**, 237–245, Available from: https://doi.org/10.1088/1469-7688/1/2/305
Floros C, Gkillas K, Konstantatos C, and Tsagkanos A (2020). Realized measures to explain volatility changes over time, *Journal of Risk and Financial Management*, **13**, 1–19, Available from: https://doi.org/ 10.3390/jrfm13060125
Gunay S (2015). Markov regime switching GARCH model and volatility modeling for oil returns, *International Journal of Energy Economics and Policy*, **5**, 979–985.
Han H (2015). Asymptotic properties of GARCH-X processes, *Journal of Financial Econometrics*, **13**, 188–221, Available from: https://doi.org/10.1093/jjfinec/nbt023
Hansen PR, Huang Z, and Shek HH (2012). Realized GARCH: A joint model for returns and realized measures of volatility, *Journal of Applied Econometrics*, **27**, 877–906, Available from: https://doi.org/ 10.1002/jae.1234
Hansen PR and Huang Z (2016). Exponential GARCH modeling with realized measures of volatility, *Journal of Business and Economic Statistics*, **34**, 269–287, Available from: https://doi.org/10.1080/07350015.2015.1038543
Hansen PR and Lunde A (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1, 1)?, *Journal of Applied Econometrics*, **20**, 873–889, Available from: https://doi.org/10.1002/jae.800
John JA and Draper NR (1980). An alternative family of transformations, *Applied Statistics*, **29**, 190–197.
Lampart T and Sbalzarini I (2012). Implementation and performance comparison of an ensemble sampler with affine invariance, *Technical report of the MOSAIC group*, Institute of Theoretical Computer Science.
Le H, Pham U, Nguyen P, and Pham TB (2020). Improvement on Monte Carlo estimation of HPD intervals, *Communications in Statistics - Simulation and Computation*, **49**, 2164–2180, Available from: https://doi.org/10.1080/03610918.2018.1513141
Liu LY, Patton AJ, and Sheppard K (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes, *Journal of Econometrics*, **187**, 293–311, Available from: https://doi.org/10.1016/j.jeconom.2015.02.008
Mahajan V, Thakan S, and Malik A (2022). Modeling and forecasting the volatility of NIFTY 50 using GARCH and RNN models, *Economies*, **10**, 1–20, Available from: https://doi.org/10.3390/economies 10050102
Manly BFJ (1976). Exponential data transformations, *The Statistician*, **25**, 37–42.
Namugaya J, Weke PGO, and Charles WM (2014). Modelling volatility of stock returns: Is GARCH (1, 1) enough?, *International Journal of Sciences: Basic and Applied Research*, **16**, 216–223.
Nugroho DB, Mahatma T, and 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**, 908–919.
Nugroho DB (2018). Comparative analysis of three MCMC methods for estimating GARCH models, *IOP Conference Series: Materials Science and Engineering*, **403**, 1–7, Available from: https://doi.org/ 10.1088/1757-899X/403/1/012061

Portet S (2020). A primer on model selection using the Akaike information criterion, *Infectious Disease Modelling*, **5**, 111–128, Available from: https://doi.org/10.1016/J.IDM.2019.12.010

Ramachandran KM and Tsokos CP (2021). *Mathematical Statistics with Applications in R* (3rd ed), Academic Press, Available from: https://doi.org/10.1016/C2012-0-07341-3

Rosenthal JS (2011). Optimal proposal distributions and adaptive MCMC. In S Brooks, A Gelman, GL Jones, and XL Meng (Eds), *Handbook of Markov Chain Monte Carlo* (pp. 93–112), Chapman & Hall/CRC, Boca Raton, FL.

Sampid MG, Hasim HM, and Dai H (2018). Refining value-at-risk estimates using a Bayesian Markov-switching GJR-GARCH copula-EVT model, *PLoS ONE*, **13**, e0198753, Available from: https://doi.org/ 10.1371/journal.pone.0198753

Tukey JW (1977). *Exploratory Data Analysis*, Addison-Wesley, Reading, MA.

van Ravenzwaaij D, Cassey P, and Brown SD (2018). A simple introduction to Markov chain Monte-Carlo sampling, *Psychonomic Bulletin & Review*, **25**, 143–154, Available from: https://doi.org/10 .3758/ s13423-016-1015-8

Xu S (2017). A VaR assuming student-*t* distribution not significantly different from a VaR assuming normal distribution, *Risk Management*, **19**, 189–201.

Yang XS (2019). *Introduction to Algorithms for Data Mining and Machine Learning*, Academic Press, London.

Yeo IK and Johnson R (2000). A new family of power transformations to improve normality or symmetry, *Biometrika*, **87**, 954–959.

Received September 16, 2022; Revised December 27, 2022; Accepted January 24, 2023