GARCH(1,1) model with the Yeo–Johnson transformed returns

D B Nugroho
Department of Mathematics, Study Center for Multidisciplinary Applied Research and Technology (SeMARTy), Satya Wacana Christian University, Jl. Diponegoro 52–60 Salatiga 50711, Jawa Tengah, Indonesia

Corresponding author’s e-mail: didit.budinugroho@staff.uksw.edu

Abstract. One of the most popular class in modeling financial volatility is GARCH-type model. Several extensions to the basic GARCH model have been developed to be more flexible in capturing various characteristics of financial time series data. In this study, a new class of GARCH models is proposed by applying Yeo–Johnson transformation to the return series. The proposed model is estimated by employing adaptive random walk Metropolis (ARWM) method in Markov chain Monte Carlo (MCMC) scheme. Our empirical results on the GARCH(1,1) models showed that the proposed model outperformed the initial model in fitting ten different international stock indices.

1. Introduction

A conditional variance of financial return series varying over time is important in risk management applications, such as asset pricing and hedging, for investing in risky assets. After the Engle's seminal paper [1] on autoregressive conditional heteroskedasticity (ARCH) model and its generalized version, so-called generalized ARCH (GARCH), by Bollerslev [2], a large number of variants of such models have been proposed (see, e.g. Poon and Granger [3] for a comprehensive review). The most popular one of such types in many applications is the GARCH(1,1) model, that is, tomorrow's variance regresses on today's squared return and today's variance. Furthermore, Hansen and Lunde [4] compared 330 ARCH-type models and found that the GARCH(1,1) model outperformed other models in the time series forecasting.

Various extensions and modifications of the ARCH model have been proposed and one of them is to take a non-linear form of the conditional variance (see, e.g. Higgins and Bera [5], Ding et al [6], Hentschel [7]) and time series variable (see, e.g. Sarkar [8], Utami and Subanar [9]). Their modifications were constructed by applying the Box–Cox [10] family of power transformations, where the basic transformation only worked with the positive variable. In the context of Stochastic Volatility (SV), those transformations were applied by Nugroho et al [11] to the volatility equation. Recently, Yeo–Johnson [12] introduced a new family of transformations which worked well on the whole real line and appropriate for reducing skewness and achieving normal approximation. These transformations are either monotonically convex or concave, which according to van Zwet (1964) [13] a convex, respectively concave, function reduces left-skewness, respectively right-skewness.
Therefore, this study extends the previous works of Sarkar [8] and Utami and Subunar [9] by applying the Yeo–Johnson (YJ) transformation to the return series data on the GARCH model.

This study compares the proposed model and basic model using ten different international stock indices. The main findings of our empirical analysis consisted three parts. First, we found that our proposed model yielded better in-sample fit than the basic model in terms of log-likelihood and deviance information criterion. Second, the long run average volatility of transformed-returns was more steady and less volatile over time. Third, the half-life of transformed-returns variance had a longer time period.

This paper proceeds as follows. Section 2 describes the proposed model and estimation using Markov chain Monte Carlo (MCMC) scheme. Section 3 presents the empirical results of the model performance, and summarizes some conclusions in Section 4.

2. Model and Estimation

GARCH model is one of econometric models which widely used for modelling and forecasting time-varying variance. One popular member of the GARCH family is the GARCH(1,1) model, which has an interpretation that tomorrow's variance is simply a function of the long-run average variance, today's squared return, and today's variance. The GARCH(1,1) form is given below,

\[ R_t = \sigma_t \varepsilon_t, \varepsilon_t \sim N(0,1), t = 1, 2, ..., T \]

\[ \sigma_t^2 = \omega + \alpha R_{t-1}^2 + \beta \sigma_{t-1}^2, t = 2, 3, ..., T \]

where \( R_t \) denotes the mean-corrected logarithmic return of an asset at time \( t \) and \( \omega, \alpha, \beta \) are unknown parameters. For a series of asset prices \( P_t \), the return \( R_t \) on a percentage scale is defined as

\[ R_t = 100 \times \left[ \ln \left( \frac{P_t}{P_{t-1}} \right) - \frac{1}{T} \sum_{t=1}^{T} \ln \left( \frac{P_t}{P_{t-1}} \right) \right]. \]

The GARCH(1,1) model imposes some restrictions such as \( \omega > 0, \alpha, \beta \geq 0 \) for positive variance, and \( \alpha + \beta < 1 \) for the covariance stationarity. Under the stationarity condition, the unconditional variance or long-run average variance per unit of time implied by the model can be calculated as

\[ V_L = \frac{\omega}{1 - \alpha - \beta}. \]

So the magnitude of \( \alpha + \beta \) controls the speed to which the conditional variance reverts to its long-run average variance. This magnitude is the so-called persistence of variance. A variance persistence of 1 does not imply a reversion to its mean. A high persistence (close to 1) indicates slow reversion while low persistence indicates rapid reversion. A measure of the persistence is a 'half-life' of variance . Engle and Patton [14] defined the variance half-life as the time taken by variance to move a halfway back to its unconditional variance. In the GARCH(1,1) context, the variance half-life is given by \( \ln(0.5)/\ln(\alpha + \beta) \) (see Ahmed et al [15] for detail).

This study proposes a generalized GARCH model by applying the YJ transformation to return series and called it “YJR-GARCH”. The return equation in the GARCH(1,1) model at time \( t \) is now represented by

\[ R(t, \delta) = \sigma_t \varepsilon_t, \]

where \( R(t, \delta) \) denotes the YJ function of the variable \( R_t \) given by

\[ R(t, \delta) = \begin{cases} 
\frac{(R_t + 1)^\delta - 1}{\delta}, & \text{if } \delta \neq 0, R_t \geq 0; \\
\ln(R_t + 1), & \text{if } \delta = 0, R_t \geq 0; \\
\frac{(1 - R_t)^{2-\delta} - 1}{\delta - 2}, & \text{if } \delta \neq 2, R_t < 0; \\
-\ln(-R_t + 1), & \text{if } \delta = 2, R_t < 0.
\end{cases} \]

The value of \( \delta = 1 \) implies no transformation. This new class has several attractive features: (i) on the positive line, it is equivalent to the shifted Box–Cox transformation, \( (R_t + 1)^\delta - 1/\delta \), for \( R_t > -1 \); (ii) it has a greater flexibility in functional form; and (iii) it displays a measure of the degree of departure from the basic GARCH model.
In order to estimate the GARCH parameters, this study employs the Adaptive Random Walk Metropolis (ARWM) method in the MCMC algorithm. The method was proposed by Atchade and Rosenthal [16] and used by Nugroho et al [17] and Nugroho [18] in the context of GARCH-type models. Alternatively, the HMC-based methods as in Takaishi [19] and Nugroho and Morimoto [20] can be used. The Bayesian MCMC algorithm is run for 6,000 iterations, but the first 1,000 iterations are discarded. The MCMC samples are drawn from the joint posterior distribution of parameters vector \( \theta = (\omega, \alpha, \beta, \delta) \):

\[
p(\theta|R(t, \delta)) = \text{prior}(\theta) \times L(R(t, \delta)|\theta) \times |J(R(t, \delta))|
\]

where \( R(t, \delta) \) denotes the YJ transformed returns vector, i.e. \( R(t, \delta) = \{R(t, \delta)\} \), \( L(R(t, \delta)|\theta) \) is the likelihood function of transformed data \( R(t, \delta) \) given the parameters vector \( \theta \), and \( |J(R(t, \delta))| \) is the determinant of the Jacobian matrix of YJ transformation. The same prior distribution is used for the all parameters, i.e. \( N(0,1000) \). All calculations were performed on an Intel Core i7 CPU, 2.10GHz, 16 GiB of memory on a 64-bit Windows 10 operating system, running the Matlab 2016a version by writing own code.

3. Real data application

This section examines the empirical performance of the YJR-GARCH(1,1) model on empirical examples based on real financial time series data that demonstrate the statistical significance of the transformation parameter estimates.

3.1. The observed data series

We address the application by adopting ten different international stock indices, namely the DJIA, FTSE 100, IPC Mexico, KOSPI Composite Index, Nasdaq 100, Nikkei 225, Russel 2000, S&P 500, S&P CNX Nifty, and Swiss Market Index. These data are publicly available by Oxford-Man Institute's Realised Library (https://realized.oxford-man.ox.ac.uk/data/download). The sample data of the study are spanning daily data period from January 2000 to December 2017.

Table 1 presents descriptive statistics of the daily return series, such as the mean, standard deviation (SD), minimum, maximum, skewness, and Jarque–Bera (JB) normality test. All the index return series display a negative skewness (except the Nasdaq 100). Moreover, they clearly exhibit asymmetrical distribution (except the DJIA and IPC Mexico), with the absolute values for skewness between -0.105 and -1.007. Thus, the use of YJ transformation seems appropriate as occurring skewness in the returns series. Furthermore, the JB test strongly rejects the hypothesis of normal distribution for all returns series at the 5% significance level.

| Stock          | Mean   | SD    | Skewness | Min.   | Max.   | JB stats. |
|----------------|--------|-------|----------|--------|--------|-----------|
| DJIA           | 0.000  | 1.120 | -0.009   | -8.428 | 10.731 | 13992.9 (>5.98) |
| FTSE 100       | -0.000 | 0.936 | -0.143   | -5.724 | 7.080  | 3629.5 (>5.98)  |
| IPC Mexico     | -0.000 | 1.284 | -0.001   | -8.299 | 9.16   | 5060.1 (>5.98)  |
| KOSPI          | -0.000 | 1.182 | -0.342   | -11.735 | 8.802  | 6682.1 (>5.98)  |
| Nasdaq 100     | 0.000  | 1.359 | 0.105    | -8.028 | 14.926 | 10721.2 (>5.98) |
| Nikkei 225     | -0.000 | 1.164 | -0.553   | -10.530 | 11.692 | 21361.5 (>5.98) |
| Russel 2000    | 0.000  | 1.404 | -0.263   | -11.063 | 7.767  | 3660.2 (>5.98)  |
| S&P 500        | 0.000  | 1.164 | -0.174   | -9.361 | 10.210 | 12051.2 (>5.98) |
| S&P CNX Nifty  | -0.000 | 1.199 | -1.007   | -13.405 | 7.107  | 19789.3 (>5.98) |
| Swiss Market Index | -0.000 | 0.976 | -0.305   | -9.712 | 8.700  | 13445.8 (>5.98) |

3.2. Results

First, we perform a model comparison of the GARCH(1,1) and YJR-GARCH(1,1) on the basis of the log-likelihood (LL) and the DIC (Deviance Information Criterion) estimators. The DIC was
introduced by Spiegelhalter et al [21] to compare the performance of Bayesian hierarchical models. The DIC can be written as a function of the log-likelihood, i.e.
\[
DIC = 2 \ln L(R(t, \delta)|\bar{\theta}) - 4E_{\theta|R} [\ln L(R(t, \delta)|\theta)],
\]
where \(\ln L(R(t, \delta)|\bar{\theta})\) is the log-likelihood function of \(R(t, \delta)\) given the posterior mean of the parameters and \(E_{\theta|R}[\ln L(R(t, \delta)|\theta)]\) is the posterior expectation of the log-likelihood function of \(R(t, \delta)\) given \(\theta\). A small value of DIC indicates a better model fit to data. We repeated the MCMC computation 10 times to obtain estimates of the performance in terms of means and standard deviations of both log-likelihood and DIC estimates, see Table 2. The YJR-GARCH(1,1) model provides smaller values DIC than GARCH(1,1), indicating that the proposed model fits the sample data better than the GARCH model. This suggests that the YJ transformation can be considered as an alternative for transforming the return series.

Table 2. Estimated likelihood on a logarithmic scale (Log-Like) and DIC. Standard deviations are in parentheses.

| Stock               | GARCH(1,1)         | YJR-GARCH(1,1)       |
|---------------------|--------------------|----------------------|
|                     | LL     | DIC    | LL     | DIC    |
| DJIA                | -5727.6 (0.038) | 11457.9 (0.149) | -5724.2 (0.051) | 11452.2 (0.235) |
| FTSE 100            | -5107.1 (0.075)  | 10217.0 (0.297) | -5104.9 (0.086) | 10213.4 (0.334) |
| IPC Mexico          | -6707.8 (0.051)  | 13418.5 (0.203) | -6705.3 (0.055) | 13414.6 (0.222) |
| KOSPI               | -6028.9 (0.078)  | 12060.6 (0.312) | -6019.2 (0.080) | 12042.3 (0.326) |
| Nasdaq 100          | -6490.6 (0.046)  | 12984.1 (0.181) | -6488.6 (0.054) | 12981.0 (0.218) |
| Nikkei 225          | -6093.9 (0.094)  | 12190.8 (0.374) | -6065.8 (0.088) | 12135.6 (0.360) |
| Russel 2000         | -7093.4 (0.066)  | 14189.8 (0.265) | -7090.7 (0.068) | 14185.5 (0.269) |
| S&P 500             | -5866.6 (0.091)  | 11736.0 (0.359) | -5861.5 (0.067) | 11726.9 (0.264) |
| S&P CNX Nifty       | -5422.3 (0.068)  | 10847.5 (0.267) | -5403.4 (0.224) | 10811.1 (0.907) |
| Swiss Market Index  | -5393.1 (0.111)  | 10789.1 (0.442) | -5381.7 (0.115) | 10767.4 (0.469) |

Table 3. MCMC simulation results in the GARCH(1,1) and GARCH-YJR(1,1) models adopting ten international stock market indices.

| Stock            | Model         | Mean   |
|------------------|---------------|--------|
|                  | \(\omega\) | \(\alpha\) | \(\beta\) | \(\delta\) | \(\alpha + \beta\) | \(\sqrt{V_L}\) |
| DJIA             | GARCH(1,1)   | 0.0115 | 0.1068 | 0.8852 | 0.9920 | 1.1990 |
|                  | YJR- GARCH(1,1) | 0.0113 | 0.1089 | 0.8837 | 1.0435 | 0.9927 | 1.2442 |
| FTSE 100         | GARCH(1,1)   | 0.0066 | 0.0990 | 0.8954 | 0.9944 | 1.0856 |
|                  | YJR- GARCH(1,1) | 0.0061 | 0.0975 | 0.8973 | 1.0407 | 0.9948 | 1.0831 |
| IPC Mexico       | GARCH(1,1)   | 0.0142 | 0.0791 | 0.9130 | 0.9921 | 1.3407 |
|                  | YJR- GARCH(1,1) | 0.0132 | 0.0771 | 0.9157 | 1.0331 | 0.9928 | 1.3540 |
| KOSPI            | GARCH(1,1)   | 0.0084 | 0.0824 | 0.9122 | 0.9946 | 1.2472 |
|                  | YJR- GARCH(1,1) | 0.0075 | 0.0775 | 0.9174 | 1.0687 | 0.9949 | 1.2159 |
| Nasdaq 100       | GARCH(1,1)   | 0.0124 | 0.089  | 0.9027 | 0.9920 | 1.2450 |
|                  | YJR- GARCH(1,1) | 0.0115 | 0.0866 | 0.9059 | 1.0317 | 0.9925 | 1.2383 |
| Nikkei 225       | GARCH(1,1)   | 0.0328 | 0.1238 | 0.8562 | 0.9800 | 1.2806 |
|                  | YJR- GARCH(1,1) | 0.0316 | 0.1282 | 0.8522 | 1.1046 | 0.9804 | 1.2570 |
| Russel 2000      | GARCH(1,1)   | 0.0277 | 0.0855 | 0.8979 | 0.9834 | 1.2918 |
|                  | YJR- GARCH(1,1) | 0.0262 | 0.0860 | 0.8985 | 1.0365 | 0.9840 | 1.2796 |
| S&P 500          | GARCH(1,1)   | 0.0110 | 0.0975 | 0.8945 | 0.9920 | 1.1726 |
|                  | YJR- GARCH(1,1) | 0.0105 | 0.0977 | 0.8950 | 1.0514 | 0.9927 | 1.1993 |
| S&P CNX Nifty    | GARCH(1,1)   | 0.0230 | 0.1149 | 0.8722 | 0.9871 | 1.3353 |
|                  | YJR- GARCH(1,1) | 0.0194 | 0.1005 | 0.8872 | 1.0956 | 0.9877 | 1.2559 |
| Swiss Market Index | GARCH(1,1) | 0.0175 | 0.1198 | 0.8636 | 1.0763 | 0.9834 | 1.0268 |
Second, the MCMC simulation results for all observed data series derived from the competing models are summarised in Table 3. In all cases, we found that the 95% Highest Posterior Density (HPD) interval for $\delta$ excluded 1, so there was significant evidence to transform the return series using the YJ transformation. This result confirms the previous findings that it is important to apply a YJ transformation to return series. The HPD interval was estimated by using the Chen–Shao algorithm (Chen and Shao [22]; Chen et al [23]), which was basically the shortest interval on a Bayesian credible intervals set.

Considering the GARCH parameters, both models provided similar estimates. The unconditional volatility $\sqrt{V_L}$ were very close to the sample standard deviation of returns reported in Table 1. In particular, the unconditional volatility of returns per day implied by the proposed model was lower than the basic model in most of the sample data. This means that the long run average volatility of YJ transformed returns are more steady and less volatile over time than the untransformed. Regarding the persistence level of conditional variance, measured by $\alpha + \beta$, all sample data demonstrated a very high persistence of variance for both models. In particular, applying the YJ transformation to the return series displayed a slightly high persistence. Hence, the forecasting variance for the YJ transformed stock's return series approached the long run average variance $V_L$ slower than for the untransformed stock's return series. This result implied a variance half-life of the YJ transformed returns had a longer time period than the untransformed returns. For example, the variance half-life of DJIA returns would take about $\ln(0.5)/\ln(0.992)$ or 95 days for the GARCH(1,1) and about $\ln(0.5)/\ln(0.9927)$ or 95 days for the YJ-GARCH(1,1).

4. Conclusions
In this study, the YJ transformation was applied to the return series in which its variance was captured by GARCH(1,1). In adopting ten different international stock indices covering the period January 2000 to December 2017, the DIC criterion and HPD interval provided evidence to support the YJ transformed returns against the untransformed returns. By assuming a stationary process of conditional variance, we observed that the YJ transformed returns fitted to the GARCH(1,1) model produced a weaker reaction to a shock return than the untransformed return.

There are some possibilities to extend our study. It is possible to extend this to return errors from a non-normal distribution such as Student-$t$ as in Nugroho and Susanto [24], or non-central Student-$t$ and skew Student-$t$ distributions as in Nugroho and Morimoto [25,26] and Nugroho et al [27]. Other interesting extensions would be to consider other transformations such as exponential and modulus transformations proposed by Manly [28] and John and Draper [29]. Finally, it would be interesting to evaluate forecasting performances of the YJR-GARCH model relative to the basic GARCH model using out-of-sample predictions.

Acknowledgments
Sincere thanks to Satya Wacana Christian University for funding. Because of their financial supports, we were able to finish this research work and final publication.

References
[1] Engle R F 1982 *Econometrica* 50 987–1007 DOI: 10.2307/1912773
[2] Bollerslev T 1986 *J. Econom.* 31 307–27 DOI: 10.1016/0304-4076(86)90063-1
[3] Poon S-H and Granger C W J 2003 *J. Econ. Lit.* 41 478–539 DOI: 10.1257/002205103765762743
[4] Hansen P R and Lunde A 2005 *J. Appl. Econom.* 20 873–89 DOI: 10.1002/jae.800
[5] Higgins M L and Bera A K 1992 *Int. Econ. Rev. (Philadelphia).* 33 137–58 DOI: 10.2307/2526988
[6] Ding Z, Granger C and Engle R 1993 *J. Empir. Financ.* 1 83–106 DOI: 10.1016/0927-5398(93)90006-D
[7] Hentschel L 1995 *J. Financ. Econ.* 39 71–104 DOI: 10.1016/0304-405X(94)00821-H
[8] Sarkar N 2000 Stat. Probab. Lett. **50** 365–74 DOI: 10.1016/S0167-7152(00)00117-6
[9] Utami H and Subanar 2013 *J. Indones. Math. Soc.* **19** 99–110 DOI: 10.22342/jims.19.2.166.99-110
[10] Box G E P and Cox D R 1964 *J. R. Stat. Soc. Ser. B* **26** 211–52 DOI:
[11] Nugroho D B, Mahatma T and Pratomo Y 2017 *Proceedings of The ISI Regional Statistics Conference* (Statistics Department – Bank Indonesia) pp 560–6
[12] Yeo I-K and Johnson R 2000 *Biometrika* **87** 954–9
[13] van Zwet W R 1964 *Convex transformations of random variables* (Amsterdam: Mathematisch Centrum)
[14] Engle R F and Patton A J 2001 *Quant. Financ.* **1** 237–45 DOI: 10.1088/1469-7688/1/2/305
[15] Ahmed R R, Vveinhardt J, Streimikiene D and Channar Z A 2018 *Econ. Res. Istraživanja* **31** 1198–217 DOI: 10.1080/1331677X.2018.1456358
[16] Atchade Y F and Rosenthal J S 2005 *Bernoulli* **11** 815–28
[17] Nugroho D B, Susanto B and Pratama S R 2017 *J. Econ. dan Ekon. Stud. Pembang.* **9** 65–75 DOI: 10.17977/um002v9i12017p065
[18] Nugroho D B 2018 *IOP Conference Series: Materials Science and Engineering* (IOP Publishing) DOI: 10.1088/1757-899X/403/1/012061
[19] Takaishi T 2007 *Proceedings of the 9th Joint Conference on Information Science* p 159
[20] Nugroho D B and Morimoto T 2015 *Comput. Stat.* **30** 491–516 DOI: 10.1007/s00180-014-0546-6
[21] Spiegelhalter D J, Best N G, Carlin B P and van der Linde A 2002 *J. R. Stat. Soc. B* **64** 583–639 DOI: 10.1111/1467-9868.00353
[22] Chen M-H and Shao Q-M 1999 *J. Comput. Graph. Stat.* **8** 69–92
[23] Chen M-H, Shao Q-M and Ibrahim J G 2000 *Monte Carlo Methods in Bayesian Computation* (New York: Springer)
[24] Nugroho D B and Susanto B 2017 *AIP Conference Proceedings* vol 1868 (AIP Publishing LLC) p 040005 DOI: 10.1063/1.4995120
[25] Nugroho D B and Morimoto T 2014 **44** 83–118 DOI: 10.14490/jjss.44.83
[26] Nugroho D B and Morimoto T 2016 *J. Appl. Stat.* **43** 1906–27 DOI: 10.1080/02664763.2015.1125862
[27] Nugroho D B, Mahatma T and Pratomo Y 2018 *Gadjah Mada Int. J. Bus.* **20** 165–85 DOI: 10.22146/gamaijb.26006
[28] Manly B F J 1976 *Stat.* **25** 37–42
[29] John J A and Draper N R 1980 *Appl. Stat.* **29** 190–7