economies

MDPI

Article

An Experiment on Autoregressive and Threshold Autoregressive Models with Non-Gaussian Error with Application to Realized Volatility

Ziyi Zhang †,‡ **and Wai Keung Li** *,‡

Department of Statistics and Actuarial Science, University of Hong Kong, Hong Kong, China; u3514050@connect.hku.hk or saas@hku.hk

* Correspondence: hrntlwk@hku.hk

† Current address: Room C, Front Portion, 2/F, No.425Z Queen's Road West, Hong Kong.

‡ These authors contributed equally to this work.

Received: 27 February 2019; Accepted: 11 June 2019; Published: 17 June 2019

Abstract: This article explores the fitting of Autoregressive (AR) and Threshold AR (TAR) models with a non-Gaussian error structure. This is motivated by the problem of finding a possible probabilistic model for the realized volatility. A Gamma random error is proposed to cater for the non-negativity of the realized volatility. With many good properties, such as consistency even for non-Gaussian errors, the maximum likelihood estimate is applied. Furthermore, a non-gradient numerical Nelder–Mead method for optimization and a penalty method, introduced for the non-negative constraint imposed by the Gamma distribution, are used. In the simulation experiments, the proposed fitting method found the true model with a rather insignificant bias and mean square error (MSE), given the true AR or TAR model. The AR and TAR models with Gamma random error are then tested on empirical realized volatility data of 30 stocks, where one third of the cases are fitted quite well, suggesting that the model may have potential as a supplement for current Gaussian random error models with proper adaptation.

Keywords: Autoregressive Model; non-Gaussian error; realized volatility; Threshold Autoregressive Model

1. Introduction

As the financial market and investment instruments grow more sophisticated, the need for the proper risk management of financial activities and the modeling of financial volatility has become more crucial. As there is no unique and unambiguous definition for volatility, observable quantities (such as daily high-lows or intra-day price changes) are used to approximate the quantity, thus dividing volatility modeling techniques into two sub-groups: Parametric and non-parametric (Anderson et al. 2002; Zheng et al. 2014). The first group are the traditional parametric latent volatility models, such as the Generalized Autoregressive Conditional Heteroscedastic (GARCH) model or the Stochastic Variance (SV) model. However, these parametric models have become increasingly restrictive in use, due to growing complexity. As mentioned in McAleer and Medeiros (2008), as the traditional standard latent volatility models cannot adequately describe the slowly decreasing auto-correlations of squared returns and as the usage of Gaussian standardized error has been criticized by many, the Realized Volatility (RV) model, as an alternative non-parametric method, has received increasing attention. In its simplest form, the RV model can be simply defined as

$$RV_t = \sum_{i=0}^{n_t} r_{t,i}^2, \tag{1}$$

where RV_t denotes the realized volatility at day t, $r_{t,i}$ denotes the ith intra-period return at day t, and n_t is the number of high frequency data observed. It has been shown that the RV model more accurately measures the 'true volatility' than daily squared returns (Anderson et al. 1999; Kambouroudis et al. 2016) and it is among the best for modeling the volatilities of the U.S. and E.U. stock indices (Kambouroudis et al. 2016). It is also a good measure for market risk, due to its ability to show clustering and fat-tail behavior for price fluctuations (Zheng et al. 2014).

A lot of work has been done towards constructing the realized volatility; see McAleer and Medeiros (2008) for a review. The focus of this article is to consider possible probabilistic models, given RV_t; particularly if it is possible to model it with a non-Gaussian random error structure. As models based on the Wishart distribution have been proposed for multi-variate realized volatility (Golosnoy et al. 2012) and multi-variate stochastic volatility (Gouriéroux et al. 2009), and as the Wishart distribution is the multi-variate analog of the chi-square distribution (which is a member of the Gamma distribution family), a Gamma random error structure in the univariate case has become of interest. Thus, traditional Autoregressive (AR) and Threshold-type non-linear AR (TAR) models with Gamma random error are explored. This article can be regarded as an extension of Li and McLeod (1988).

2. Materials and Methods

This section aims to provide the specification for the proposed model and the fitting methodology. It will also briefly touch on the materials and methods for conducting the empirical data analysis.

2.1. Model Specification

Time-series models with non-Gaussian error were previously considered, in some detail, by Li and McLeod (1988), and earlier in Lawrance and Lewis (1980) and Ledolter (1979). In this article, specifically, the AR and TAR models are further explored. The AR(p) model is defined as follows:

$$RV_t = \sum_{i=1}^{p} \varphi_i * RV_{t-i} + \varepsilon_t \,, \tag{2}$$

where ε_t is the random error, assumed to follow a Gamma distribution; thus, $\varepsilon_t \sim \Gamma(\alpha, \beta)$, where the density function is defined as

$$f(x) = \frac{1}{\Gamma(\alpha) * \beta^{\alpha}} * x^{\alpha-1} * e^{-\frac{x}{\beta}} \,. \tag{3}$$

It should be noted that it is assumed that there is no drift term in the AR model; yet, the drift term could be easily incorporated into the model. The TAR(p) model, similar to that introduced in Tong (1978) but with a modification in the random error term, is defined as follows:

$$\begin{aligned} RV_t &= \sum_{i=1}^{p} \varphi_{1,i} * RV_{t-i} + \varepsilon_{1,t} \qquad if\ RV_{t-d} \leq T \,, \\ RV_t &= \sum_{i=1}^{p} \varphi_{2,i} * RV_{t-i} + \varepsilon_{2,t} \qquad if\ RV_{t-d} > T \,, \end{aligned} \tag{4}$$

where $d \geq 1$ is the lag of the model and T is the threshold, such that the model is divided into two regimes, according to the observations at d time periods earlier. The pivot element RV_{t-d} determines which regime RV_t falls into, with RV_t falling into the first regime if RV_{t-d} is less than or equal to the threshold and into the second regime, otherwise. Each regime follows an AR(p) model, as defined above, with different AR and Gamma parameters.

2.2. Model Estimation

The fitting of the AR(p) model is introduced in this part, followed by the extension of the procedure to the fitting of the TAR(p) model. Both procedures are fitted with the maximum likelihood procedure,

as it has been shown that the maximum likelihood estimators (MLE) are consistent for Gamma random error (Li and McLeod 1988).

The MLE for AR(p) model are derived by $l(\hat{\alpha}, \hat{\beta}, \hat{\varphi}) = argmin(-l)$, where l denotes the log-likelihood function, in the form of

$$l(\alpha, \beta, \varphi) = -n * ln(\Gamma(\alpha)) - n * \alpha * ln(\beta) + (\alpha - 1) * \sum_{t=1}^{n} ln(\varepsilon_t) - \frac{\sum_{t=1}^{n} \varepsilon_t}{\beta}, \quad (5)$$

where

$$\varepsilon_t = RV_t - \sum_{i=1}^{p} \varphi_i * RV_{t-i} \quad (6)$$

is the random error.

To further reduce the dimension of estimation, a profile likelihood method is used. The Gamma parameters α and β are replaced by the MLE of α and β, using the result of Wilk et al. (1962) and the approximation $\frac{dln(\gamma(\alpha))}{d\alpha} \approx ln(\alpha - \frac{1}{2})$. Thus, the final estimates of α and β are as follows:

$$\hat{\alpha} = \frac{A}{2 * (A - G)} \qquad and \qquad \hat{\beta} = \frac{A}{\hat{\alpha}}, \quad (7)$$

where A stands for the arithmetic mean of the random error and G is the geometric mean. Thus, the estimation of the model is achieved by estimating $\hat{\varphi} = argmin(-\hat{l})$, where

$$\hat{l}(\varphi) = -n * ln(\Gamma(\hat{\alpha})) - n * \hat{\alpha} * ln(\hat{\beta}) + (\hat{\alpha} - 1) * \sum_{t=1}^{n} ln(\varepsilon_t) - \frac{\sum_{t=1}^{n} \varepsilon_t}{\hat{\beta}}, \quad (8)$$

where $\hat{\alpha}$ and $\hat{\beta}$ are estimated by Equation (7) and, by Equation (6), thus depend on φ .

The Nelder–Mead method, which is a non-gradient optimization method, is proposed to optimize the negative log-likelihood function. Although such a procedure is heuristic and may converge to non-stationary points, its performance is much more stable than traditional gradient methods, such as the Hessian matrix method, which may not be easily calculated (even numerically) given the dependency of the log-likelihood function and as φ is quite complicated.

Additionally, before simply applying the method and carrying out the optimization, it should be noticed that, as the random error ε_t is assumed to be Gamma, it is required to be greater than zero, which is also evident from the term $ln(\varepsilon_t)$ in the expression of the log-likelihood function. To reflect this non-negativity constraint, a penalty method is applied and the log-likelihood function becomes:

$$\hat{l} = (-n * ln(\Gamma(\hat{\alpha})) - n * \hat{\alpha} * ln(\hat{\beta}) + (\hat{\alpha} - 1) * \sum_{t=1}^{n} ln(\varepsilon_t) - \frac{\sum_{t=1}^{n} \varepsilon_t}{\hat{\beta}}) * I_{all\ \varepsilon_t \geq 0} - M * I_{some\ \varepsilon_t < 0}, \quad (9)$$

where M is some large-enough number.

As the Nelder–Mead method is a heuristic search method, the choice of initial point may greatly affect the result and, thus, the estimation process takes various initial points and returns the result that yields a best fit, using the AIC or BIC . Furthermore, a candidate set of AR order p is given and the procedure searches for the best AR order within the set, again by AIC and BIC. Specifically, in the scope of the simulation study in this report, the initial points for φ are set uniformly within [0,1] and the initial points for T are set within $[\mu - n * \sigma, \mu + n * \sigma]$, where μ the sample mean of the RV, σ is the sample variance, and n is a pre-determined number to control the range, here set as 0.5. The step size of φ is set to be 0.25 and that of T to be 0.05σ. For empirical data analysis, values of φ in the ranges [0,0.5] and [0.5,1] are tested, with step size 0.125, and the results showed that the outcome from [0,0.5]

almost always dominated that from [0.5,1] and, thus, the range [0,0.5] and step size 0.125 were used for φ.

The fitting of the TAR(p) model is essentially the same, except that the random errors are classified into two different regimes. Thus, the log-likelihood function is expressed as:

$$\begin{aligned}\hat{l} = (&-n_1 * ln(\Gamma(\hat{\alpha_1})) - n_1 * \hat{\alpha_1} * ln(\hat{\beta_1}) + (\hat{\alpha_1} - 1) * \sum_{t=1}^{n_1} ln(\varepsilon_{1,t}) - \frac{\sum_{t=1}^{n_1} \varepsilon_{1,t}}{\hat{\beta_1}} \\ &-n_2 * ln(\Gamma(\hat{\alpha_2})) - n_2 * \hat{\alpha_2} * ln(\hat{\beta_2}) + (\hat{\alpha_2} - 1) * \sum_{t=1}^{n_2} ln(\varepsilon_{2,t}) - \frac{\sum_{t=1}^{n_2} \varepsilon_{2,t}}{\hat{\beta_2}}) \\ &* I_{all\ \varepsilon_{1,t},\varepsilon_{2,t} \geq 0} - M * I_{some\ \varepsilon_{1,t},\varepsilon_{2,t} < 0}\,,\end{aligned} \tag{10}$$

where $\varepsilon_{1,t}$ are the random errors corresponding to the observations in the first regime, n_1 is the number of observations in the first regime, and $\varepsilon_{2,t}$ and n_2 the corresponding counterparts in the second regime, respectively.

A final concern regarding the model estimation would be that, for the first few observations, the AR model may not be properly initiated, as there are no earlier observations. Therefore, the sample estimates are essentially estimated by a sample, with the first few observations serving only as the independent variable, but not the dependent variable; that is,

$$RV_{t+n} = \sum_{i=1}^{p} \varphi_i * RV_{t+n-i} + \varepsilon_{t+n}\,, \tag{11}$$

with n being the truncated size. Additionally, as the AIC and BIC are typically applied on the same sample with the same sample size, to allow for the comparison between models of different AR order and lag, a common truncation of size 10 is applied in the scope of this study, as the AR order and lag investigated did not exceed this reasonably.

As with the process of fitting the AR(p) model, the fitting for TAR(p) searches for the best model of AR order p and lag d, where p and d are given in the pre-determined candidate set and the threshold T.

2.3. Empirical Data Analysis Preparation

The data used in this paper were the consolidated realized volatility data from Shen et al. (2018), which are the realized volatilities for 30 stocks traded on the New York Stock Exchange (NYSE).

Graphs of PACF and the corresponding naive 95% confidence bound, proposed by Quenouille (1949), were first examined for the stock data, which showed that the PACF of the stocks were mostly significant within a lag of 5 and demonstrated a somewhat cut-off property; thus suggesting the fitting the AR model was potentially a good starting point. Non-linear threshold type AR models were also considered as a supplement to the AR model.

After considering the practical reasonableness of the model and the computational power available, an AR order up to 5 and lag order up to 3 were considered.

The final model for each stock was determined by both considering the AIC and BIC and the associated Ljung–Box test for each criterion. If the model selected by the two criteria differed with a similar goodness of fit, a simpler model was preferred. Otherwise, the model that gave a better goodness of fit result was preferred.

The data set and R code used for the study are available upon request, from either author.

3. Results

This section aims to briefly describe how the proposed AR and Threshold AR (TAR) models were fitted with a simulation study and some empirical data.

3.1. Simulation Study

A simulation study was conducted, by running the model-fitting process on batches of randomly generated AR or TAR models of observation length 500 and batch size 50 for all the results in this section (i.e., 50 simulated observations of length 500 were considered in each simulation experiment). The completion of each simulation took around half a day on a laptop. The following tables give the results for the bias and mean square error in the simulation study. Tables 1 and 2 give the results for the threshold models and Table 3 gives the result for AR models. The correct estimation of AR order and lag meant that the estimation of both the AR order p and the lag d were in line with the true parameters. The bias and MSE were calculated with the results in the simulations which gave the correct estimation of AR order and lag. The parameter for the true TAR model was selected such that the TAR structure was reasonably demonstrated (i.e., there were not too few observations in any regime).

Table 1. Simulation results for the threshold Autoregressive (TAR) (2) model with $d = 2$.

True Model	α_1	β_1	α_2	β_2	$\varphi_{1,1}$	$\varphi_{1,2}$	$\varphi_{2,1}$	$\varphi_{2,2}$	T
	5	2	5	2	0.5	0.3	0.3	0.2	30
AIC	Proportion of correct estimation of Autoregressive (AR) order and Lag: 44/50								
AIC Bias	0.032	0.023	0.341	−0.022	0.013	−0.007	0.003	−0.004	0.001
AIC MSE	1.245	0.083	2.759	0.123	0.002	0.003	0.002	0.002	0.001
BIC	Proportion of correct estimation of AR order and Lag: 50/50								
BIC Bias	0.015	0.022	0.384	−0.015	0.012	−0.006	0.004	−0.005	0.001
BIC MSE	1.198	0.08	3.887	0.143	0.002	0.003	0.002	0.002	0.001

Table 2. Simulation results for the TAR (1) model with $d = 1$.

True Model	α_1	β_1	α_2	β_2	$\varphi_{1,1}$	$\varphi_{2,1}$	T
	4	2	4	2	0.7	0.3	15
AIC	Proportion of correct estimation of AR order and Lag: 36/50						
AIC Bias	0.34	−0.074	0.068	−0.006	−0.008	0.007	0.019
AIC MSE	0.929	0.086	0.722	0.078	0.003	0.001	0.002
BIC	Proportion of correct estimation of AR order and Lag: 50/50						
BIC Bias	0.199	−0.035	0.016	0.024	−0.002	0.007	0.019
BIC MSE	0.825	0.079	0.711	0.082	0.003	0.001	0.002

Table 3. Simulation results for the AR (2) model.

True Model	α	β	φ_1	φ_2
	5	2	0.6	0.2
AIC	Proportion of correct estimation of AR order and Lag: 7/50			
AIC Bias	0.627275	−0.08103	0.020833	−0.02968
AIC MSE	2.004022	0.087312	0.001747	0.002992
BIC	Proportion of correct estimation of AR order and Lag: 46/50			
BIC Bias	0.137581	−0.02197	0.006732	−0.00646
BIC MSE	0.605966	0.037434	0.001058	0.001104

The estimates of the AR order and lag were generally good, except the AIC criterion for the AR model, as the AIC tends to pick a more complicated model. In fact, the AIC estimated the AR order correctly in 36 out of 50 cases; yet, in most of these cases, it preferred a threshold structure.

The simulation results show that the model could identify the correct AR order p and the lag d with good accuracy in general, the estimate for the threshold T was very consistent; and the results for the AR parameters φ were rather accurate when p and d were estimated correctly. It should be

noted that the accuracy here is defined as the probability of identifying the correct AR model order and correct threshold, given that the underlying model was indeed an AR/TAR model.

3.2. Empirical Data Analysis

The realized volatilities of 30 stocks traded on the New York Stock Exchange (NYSE) were tested by the proposed models. Please kindly refer to Appendix A–C for the best model selected by the AIC, BIC, and the final model.

From the results, the AR/TAR model seemed to be a good fit for around 33% of the cases, with almost all of the final models having a threshold structure and a marginally good fit for another 10% of the cases, where the Ljung–Box test was marginally significant. This demonstrates that, overall, the proposed model has the potential to explain a little less than half of the empirical data, in this case, and further investigation, through other data sets or improved fitting algorithms, is worthwhile.

4. Discussion

This section aims to provide a brief discussion as a supplement to the results found above. It is divided into discussions regarding the simulation study and the empirical data, respectively.

4.1. Simulation Study

While, as mentioned before, the bias and MSE were acceptable overall, with consistent estimates for the AR parameter and threshold, it can be noticed that the estimates for Gamma parameters were more volatile. This is possibly due to the profile likelihood methodology adopted for estimation, which increases the complexity in estimating the Gamma parameters.

Additionally, as the simulation study was constructed in such a way that the true model was within the set of candidate models, the BIC would select the true model with probability tending to one and, thus, outperformed the AIC. However, in practice, the true model may not reside within the candidate set, and the AIC may give a better result, yet may also choose a more complicated model (as mentioned above), while the BIC would prefer a simpler model. Therefore, in terms of forecasting MSE, both criteria are considered, in practice, for model selection.

4.2. Empirical Data Analysis

A residual analysis was conducted by looking at the PACF plots for the models with significant goodness of fit test results. It was observed that, in some cases, the PACF still demonstrated a rather clear cut-off at a higher order, suggesting that the AR order of the model could be further increased. Thus, it is suggested that, in this case, it is possible that the model was not a good enough fit, as it did not select a high enough order. This was possible, as the model fitting limited the highest AR order to be less than five, for practical concerns, and as the optimization process was sensitive to the selection of initial points and the initial points were evaluated in a sparser set at higher AR orders, thus resulting in a less-than-ideal fit.

Alternative models with non-Gaussian error provide another perspective of improvement. A Buffered Threshold Autoregressive (BAR) model, as described in Li et al. (2015), has been examined, using a fitting methodology similar to that of the TAR model. However, as the goodness of fit did not improve much, and as it is natural to choose a simpler model given similar goodness of fit, the results of BAR model have not yet been reported. However, other models (such as the Autoregressive Moving-Average model (ARMA)) could still be considered.

5. Conclusions

In this article, the model fitting of a non-Gaussian model on the realized volatility is explored. As the definition of realized volatility requires it to be positive, previous works established a Wishart model (a multi-variate analog of the chi-square distribution) that belongs to the Gamma family;

considering this selection, a univariate Gamma random error is proposed and the AR and TAR models are explored. MLE estimation, based on the AIC and BIC, and with some adjustment, is proposed. A profile likelihood method, which replaces the Gamma parameters with their MLE counterparts, is used to reduce the dimension of the estimation and a non-gradient numerical optimization method is employed, as the calculation of gradient may not be feasible. A penalty method is introduced into the likelihood function, to enforce the non-negative constraint imposed by Gamma random error. The proposed process manages to find the true model with a rather insignificant bias and MSE, when the true model is AR or TAR. Finally, the model is tested on the empirical realized volatility data of 30 stocks and managed to fit one third of the cases quite well, suggesting that the model may have the potential to be further generalized, in order to act as a good supplement for current Gaussian random error models. The lack of fit may be improved by considering higher AR orders or a denser initial point selection for the Nelder–Mead method, which requires more computational time. Other possible directions of improvement include using a better method (instead of AIC or BIC) to reduce the ambiguity in choosing the model and possibly using other AR structures, such as the Heterogeneous Auto-Regressive (HAR) model. Other time-series models with non-Gaussian error may also be considered and the model fitting methodology proposed in this article could possibly be extended to these models without difficulty.

Author Contributions: Conceptualization, Z.Z. and W.K.L.; methodology, Z.Z. and W.K.L.; software, Z.Z.; validation, Z.Z. and W.K.L.; formal analysis, Z.Z. and W.K.L.; investigation, Z.Z. and W.K.L.; resources, W.K.L.; data curation, W.K.L.; writing—original draft preparation, Z.Z.; writing—review and editing, W.K.L.; visualization, Z.Z.; supervision, W.K.L.; project administration, W.K.L.; funding acquisition, W.K.L.

Funding: This research was funded by HK SAR GRF grant number 17304417 and 17303315.

Acknowledgments: Special thanks to the referees of the journal for the invaluable advice they provide to better this report.

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

Abbreviations

The following abbreviations are used in this manuscript:

AR	Autoregressive
AIC	Akaike Information Criterion
BAR	Buffered Threshold Autoregressive
BIC	Bayesian Information Criterion
EU	European Union
GARCH	Generalized Autoregressive Conditional Heteroscedastic
MSE	Mean Square Error
NYSE	New York Stock Exchange
PACF	Partial Auto-Correlation Function
RV	Realized Variance
SV	Stochastic Variance
TAR	Threshold Autoregressive
US	United States

Appendix A

Table A1. Best Model Selected by the AIC.

StockNum	p	d	AIC	Ljung–Box Test	*p*-Value
1	2	2	275.7274	Significant	0.0073
2	2	2	2.2509	Insignificant	0.2034
3	5	3	3.8177	Insignificant	0.0220
4	2	2	246.9670	Insignificant	0.3230
5	5	1	143.1260	Significant	0.0155
6	4	2	−28.6272	Significant	0.0218
7	4	1	−9.54336	Insignificant	0.0716
8	5	1	−21.6088	Insignificant	0.1159
9	5	1	5.7717	Significant	0.0000
10	2	1	144.2935	Insignificant	0.6415
11	2	2	−17.8815	Somewhat Significant	0.0093
12	5	2	4.7316	Significant	0.0012
13	4	1	−81.0311	Somewhat Significant	0.0064
14	2	1	−241.8272	Insignificant	0.4850
15	2	2	157.0407	Significant	0.0000
16	3	1	−180.8073	Significant	0.0000
17	2	1	−127.2748	Somewhat Significant	0.0194
18	1	1	−90.0935	Significant	0.0000
19	5	1	−117.8152	Significant	0.0000
20	3	1	30.9568	Significant	0.0136
21	4	1	−60.7726	Significant	0.0005
22	1	3	−192.5901	Significant	0.0000
23	2	1	−99.0953	Significant	0.0000
24	3	2	−92.6265	Significant	0.0000
25	2	2	68.0884	Insignificant	0.8520
26	1	1	16.6098	Significant	0.0005
27	1	1	44.5472	Insignificant	0.2261
28	2	2	−121.6981	Significant	0.0000
29	2	1	−120.5036	Significant	0.0000
30	1	1	−93.9557	Significant	0.0002

Appendix B

Table A2. Best Model Selected by the BIC.

StockNum	p	d	BIC	Ljung-Box Test	*p*-Value
1	2	2	582.7805	Significant	0.0073
2	2	2	35.8275	Insignificant	0.2304
3	1	1	36.9940	Significant	0.0000
4	2	2	525.2598	Insignificant	0.3203
5	5	0	318.3741	Significant	0.0000
6	1	1	−28.8146	Significant	0.0000
7	1	3	17.7029	Insignificant	0.8858
8	2	0	−11.8445	Significant	0.0060
9	1	1	38.3802	Significant	0.0000
10	2	1	319.9127	Insignificant	0.6415
11	2	2	−4.4373	Somewhat Significant	0.0093
12	1	1	43.3797	Significant	0.0000
13	1	1	−134.143	Significant	0.0000
14	2	1	−452.3287	Insignificant	0.4850
15	2	2	345.4071	Significant	0.0000
16	1	1	−331.1404	Significant	0.0000
17	2	1	−223.2238	Somewhat Significant	0.0194
18	1	1	−155.8226	Significant	0.0000
19	1	3	−192.3672	Significant	0.0000
20	1	1	93.1204	Significant	0.0000
21	1	1	−86.9778	Significant	0.0183
22	1	3	−360.8156	Significant	0.0000
23	2	1	−166.8648	Significant	0.0000
24	1	2	−159.1344	Significant	0.0000
25	2	2	167.5026	Insignificant	0.8520
26	1	1	57.5842	Significant	0.0002
27	1	1	113.4588	Insignificant	0.2261
28	2	2	−212.0704	Significant	0.0000
29	2	1	−209.6814	Significant	0.0000
30	1	1	−163.5470	Significant	0.0002

Appendix C

Table A3. Best Model Selected by considering both AIC and BIC and goodness of fit.

StockNum	p	d	Info Cri *	Ljung-Box Test	*p*-Value
1	2	2	BIC	Significant	0.0073
2	2	2	AIC	Insignificant	0.2304
3	5	3	AIC	Insignificant	0.0220
4	2	2	AIC	Insignificant	0.3203
5	5	0	BIC	Significant	0.0000
6	1	1	BIC	Significant	0.0000
7	1	3	BIC	Insignificant	0.8858
8	5	1	AIC	Insignificant	0.1159
9	1	1	BIC	Significant	0.0000
10	2	1	AIC/BIC	Insignificant	0.6415
11	2	2	AIC/BIC	Somewhat Significant	0.0093
12	1	1	BIC	Significant	0.0000
13	4	1	AIC	Somewhat Significant	0.0064
14	2	1	AIC/BIC	Insignificant	0.4850
15	2	2	AIC/BIC	Significant	0.0000
16	1	1	BIC	Significant	0.0000
17	2	1	AIC/BIC	Somewhat Significant	0.0194
18	1	1	AIC/BIC	Significant	0.0000
19	1	3	BIC	Significant	0.0000
20	1	1	BIC	Significant	0.0000
21	1	1	BIC	Significant	0.0000
22	1	3	AIC/BIC	Significant	0.0000
23	2	1	AIC/BIC	Significant	0.0000
24	1	2	BIC	Significant	0.0000
25	2	2	AIC/BIC	Insignificant	0.8520
26	1	1	AIC/BIC	Significant	0.0005
27	1	1	AIC/BIC	Insignificant	0.2261
28	2	2	AIC/BIC	Significant	0.0000
29	2	1	AIC/BIC	Significant	0.0000
30	1	1	AIC/BIC	Significant	0.0002

* denotes the information criteria by which the fitted model is selected. The goodness of fit is regarded as somewhat significant if, out of the different lags considered in the Ljung–Box Test, which is five in this case, around half (two or three) are insignificant, while the others are only marginally significant.

References

Andersen, Torben G., Tim Bollerslev, and Steve Lange. 1999. Forecasting financial market volatility: Sample frequency vis-a-vis forecast horizon. *Journal of Empirical Finance* 6: 457–77. [CrossRef]

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, Yacine Aït-Sahalia, Lars Hansen, and J. A. Scheinkman. 2002. *Handbook of Financial Econometrics*. Amsterdam: Elsevier.

Golosnoy, Vasyl, Bastian Gribisch, and Roman Liesenfeld. 2012. The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics* 167: 211–23. [CrossRef]

Gouriéroux, Christian, Joann Jasiak, and Razvan Sufana. 2009. The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics* 150: 167–81. [CrossRef]

Kambouroudis, Dimos S., David G. McMillan, and Katerina Tsakou. 2016. Forecasting stock return volatility: A comparison of GARCH, Implied volatility, and realized volatility models. *Journal of Futures Markets* 36: 1127–63. [CrossRef]

Lawrance, A. J., and P. A. W. Lewis. 1980. The exponential autoregressive-moving average EARMA (p, q) process. *Journal of the Royal Statistical Society. Series B (Methodological)* 42: 150–61. [CrossRef]

Ledolter, Johannes. 1979. Inference robustness of ARIMA models under non-normality—Special application to stock price data. *Metrika* 26: 43–56. [CrossRef]

Li, Guodong, Bo Guan, Wai Keung Li, and Philip L. H. Yu. 2015. Hysteretic autoregressive time series models. *Biometrika* 102: 717–23. [CrossRef]

Li, Wai Keung, and A. Ian McLeod. 1988. ARMA modelling with non-Gaussian innovations. *Journal of Time Series Analysis* 9: 155–68. [CrossRef]

McAleer, Michael, and Marcelo C. Medeiros. 2008. Realized volatility: A review. *Econometric Reviews* 27: 10–45. [CrossRef]

Quenouille, M. H. 1949. Approximate tests of correlation in time-series 3. In *Mathematical Proceedings of the Cambridge Philosophical Society*. Cambridge: Cambridge University Press, vol. 45, pp. 483–84.

Shen, Keren, Jianfeng Yao, and Wai Keung Li. 2018. Forecasting high-dimensional realized volatility matrices using a factor model. *Quantitative Finance*. [CrossRef]

Tong, Howell. 1978. On a Threshold Model. In *Pattern Recognition and Signal Processing*. Edited by C. H. Chen. Amsterdam: Sijthoff and Noordhoff, pp. 575–86.

Wilk, M. B., R. Gnanadesikan, and M. M. Huyett. 1962. Probability plots for the gamma distribution. *Technometrics* 4: 1–20. [CrossRef]

Zheng, Zeyu, Zhi Qiao, Tetsuya Takaishi, H. Eugene Stanley, and Baowen Li. 2014. Realized volatility and absolute return volatility: A comparison indicating market risk. *PLoS ONE* 9: e102940. [CrossRef] [PubMed]