

Gumbel GARCH Model with Stock Application

Mehrnaz Mohammadpour¹ Fatemeh Ziaeenejad²

Abstract

The paper proposes a new GARCH model with Gumbel conditional distribution. Several statistical properties of the model are established, like autocorrelation function and stationarity. We consider two methods for estimating the unknown parameters of the model and investigate properties of the estimators. The performances of the estimators are checked by a simulation study. We investigate the application of the process using a real stock data.

1 Introduction

The generalized autoregressive conditionally heteroscedastic (GARCH) model has been found to be useful in many economic and financial studies which captures the tendency for volatility clustering (Bollerslev, 1986). In the classical GARCH model, normal distribution has been considered as a conditional distribution which is not quite logical consideration in the most financial studies.

Modeling of GARCH time series was first introduced by Bollerslev (1986) based on the normal conditional distribution. Among the GARCH models, we cite the standardized t-student model (Bollerslev, 1986), the normal poisson mixture model (Jorion, 1988), the power exponential model (Baillie and Bollerslev, 1989), the normal-log normal mixture model (Hsieh, 1988), the generalized exponential model (Nelson, 1990b), the normal model (Nelson, 1990a, 1992), the threshold GARCH model (Glosten et al., 1993) and the stable GARCH model (Liu and Brorsen, 1995 and Calzolari et al., 2014).

In this paper, a new GARCH model with Gumbel conditional distribution is introduced. The Gumbel GARCH model is justified by the need to model extreme observations more realistically than would be possible using the standard normal GARCH. The Gumbel model is the traditional model in extreme value analysis which has the same status as the normal model in other applications. The major advantage of the Gumbel model is that the distribution can be specified by location and scale parameters as in the Gaussian case. In the following, we briefly investigate Gumbel distribution and its features.

The Gumbel distribution is referred to the distribution corresponding to extremes. The Gumbel distribution with the location parameter α and the scale parameter γ (denoted by

¹Faculty of Mathematical Sciences, University of Mazandaran, Bbolsar, Iran; m.mohammadpour@umz.ac.ir

²Faculty of Mathematical Sciences, University of Mazandaran, Bbolsar, Iran; ziaeenejadsh@yahoo.com

Gumbel(α, γ) has the probability density function

$$f_X(x) = \frac{1}{\gamma} \exp\left(\frac{x - \alpha}{\gamma} - \exp\left(\frac{x - \alpha}{\gamma}\right)\right).$$

The mean and variance of the distribution are

$$\mu = \alpha + \nu\gamma,$$

and

$$\sigma^2 = \frac{1}{6}\pi^2\gamma^2,$$

where ν is Euler-Mascheroni constant. In this work, we consider $\alpha = 0$.

The rest of the paper is organized as follows. In Section 2, we construct the Gumbel GARCH model and investigate the stationarity condition and some properties of the model. Section 3 deals with the estimation of the model parameters by the Yule-Walker and maximum likelihood methods. The performances of the estimators are checked by a small Monte Carlo simulation. An application of the model for the stock data is given in Section 4.

2 Gumbel GARCH

Let $\{X_t\}_{t \in \mathbb{Z}}$, \mathbb{Z} the set of integers, be a discrete time second-order process and \mathcal{F}_{t-1} is a σ field generated by $\{X_s\}_{s < t}$. The process $\{X_t\}_{t \in \mathbb{Z}}$ defines as Gumbel GARCH model of orders p, q (Gumbel GARCH(p, q)) if

$$X_t | \mathcal{F}_{t-1} \sim \text{Gumbel}(0, \gamma_t), \quad (2.1)$$

and γ_t satisfies the model

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = \frac{1}{6}\pi^2\gamma_t^2 = \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (2.2)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $i = 1, \dots, p$, $j = 1, \dots, q$, $p \geq 1$, $q \geq 0$.

The conditional probability mass function of $\{X_t\}$ has the following form

$$f(x_t | \mathcal{F}_{t-1}) = \frac{1}{\gamma_t} \exp\left(\frac{x_t}{\gamma_t}\right) \exp\left(-\exp\left(-\frac{x_t}{\gamma_t}\right)\right),$$

where

$$\gamma_t = \left(\frac{6}{\pi^2}\sigma_t^2\right)^{\frac{1}{2}}.$$

The conditional mean is

$$E(X_t | \mathcal{F}_{t-1}) = \nu\gamma_t = \nu\left(\frac{6}{\pi^2}\sigma_t^2\right)^{\frac{1}{2}}.$$

In Proposition 1, we establish a necessary condition on the parameters of the model to ensure that the process is a second-order stationary process.

Proposition 1. For a second-order stationary process $\{X_t\}_{t \in \mathbb{Z}}$ to satisfy (2.1) and (2.2), it is necessary that $\sum_{i=1}^p \alpha_i(1 + \nu^2 \frac{6}{\pi^2}) + \sum_{j=1}^q \beta_j < 1$.

Proof. Let $\sigma^2 = E(\sigma_t^2)$. Since

$$\begin{aligned} \sigma^2 &= E(\sigma_t^2) = \alpha_0 + \sum_{i=1}^p \alpha_i E(X_{t-i}^2) + \sum_{j=1}^q \beta_j E(\sigma_{t-j}^2) \\ &= \alpha_0 + \sum_{i=1}^p \alpha_i (1 + \nu^2 \frac{6}{\pi^2}) \sigma^2 + \sum_{j=1}^q \beta_j \sigma^2, \end{aligned}$$

we obtain

$$\sigma^2 = \frac{\alpha_0}{\left(1 - \sum_{i=1}^p \alpha_i (1 + \nu^2 \frac{6}{\pi^2}) - \sum_{j=1}^q \beta_j\right)}. \quad (2.3)$$

In (2.3), the parameters must necessarily satisfy the condition

$$1 - \sum_{i=1}^p \alpha_i (1 + \nu^2 \frac{6}{\pi^2}) - \sum_{j=1}^q \beta_j > 0,$$

which completes the proof. \square

The following theorem presents a necessary and sufficient condition for the second order stationarity of the model. For the simplicity of notation, we assume that $p \geq q$.

Theorem 1. A necessary and sufficient condition for the process $\{X_t\}$ to be second-order stationary is that all roots of

$$1 - A\alpha_0 - \sum_{i=1}^q (A\alpha_i + \beta_i)z^i - \sum_{i=q+1}^p A\alpha_i z^i = 0, \quad (2.4)$$

lie inside the unit circle, where $A = (1 + \nu^2 6/\pi^2)$.

Proof. Let $\gamma_{i,t} = E(X_t X_{t-i})$, $i = 1, 2, \dots, p$. The conditional second moment is obtained as

$$\begin{aligned} E(X_t^2 | \mathcal{F}_{t-1}) &= \text{Var}(X_t | \mathcal{F}_{t-1}) + E^2(X_t | \mathcal{F}_{t-1}) \\ &= \sigma_t^2 + \nu^2 (6/\pi^2 \sigma_t^2) \\ &= A\sigma_t^2. \end{aligned}$$

Then

$$\begin{aligned} \gamma_{0,t} &= E(X_t^2) = E(E(X_t^2 | \mathcal{F}_{t-1})) = E(A\sigma_t^2) \\ &= A \left(\alpha_0 + \sum_{i=1}^p \alpha_i E(X_{t-i}^2) + \sum_{j=1}^q \beta_j E(\sigma_{t-j}^2) \right) \\ &= A\alpha_0 + \sum_{i=1}^p A\alpha_i \gamma_{0,t-i} + \sum_{j=1}^q \beta_j \gamma_{0,t-j} \\ &= A\alpha_0 + \sum_{i=1}^q (A\alpha_i + \beta_i) \gamma_{0,t-i} + \sum_{i=q+1}^p A\alpha_i \gamma_{0,t-i}. \end{aligned}$$

The necessary and sufficient condition for a non-homogeneous difference equation to have a stable solution is that all roots z_1, \dots, z_p of eqn (2.4) lie inside the unit circle, (Goldberg, 1958). \square

The following theorem gives the autocorrelation function (ACF) of $\{X_t^2\}$ which is used for Yule-Walker estimation.

Theorem 2. *Suppose that $\{X_t\}$ following the Gumbel GARCH(p, q) model is second-order stationary. Then $\gamma_x^2(k) := Cov(X_t^2, X_{t-k}^2)$ and $\gamma_{\sigma^2}^2(k) := Cov(\sigma_t^2, \sigma_{t-k}^2)$ satisfy the following equations*

$$\begin{aligned} \gamma_x^2(k) &= \sum_{i=1}^p (1 + \nu^2 6/\pi^2) \alpha_i \gamma_x^2(|i - k|) \\ &+ \sum_{j=1}^{\min(k-1, q)} \beta_j \gamma_x^2(k - j) + \sum_{j=1}^q (1 + \nu^2 6/\pi^2)^2 \beta_j \gamma_{\sigma^2}^2(|k - j|), k \geq 1, \end{aligned} \quad (2.5)$$

and

$$\begin{aligned} \gamma_{\sigma^2}^2(k) &= \sum_{i=1}^{\min(k, p)} (1 + \nu^2 6/\pi^2) \alpha_i \gamma_{\sigma^2}^2(|i - k|) \\ &+ \sum_{i=k+1}^p \alpha_i / (1 + \nu^2 6/\pi^2) \gamma_x^2(k - j) + \sum_{j=1}^q \beta_j \gamma_{\sigma^2}^2(|k - j|), k \geq 0. \end{aligned} \quad (2.6)$$

Proof. See Appendix A for details. \square

Corollary 1. *Suppose that $\{X_t\}$ following the Gumbel ARCH(p) model is second-order stationary. Then the autocovariance function $\gamma_x^2(\cdot)$ satisfies the following equation*

$$\gamma_x^2(k) = \sum_{i=1}^p (1 + \nu^2 6/\pi^2) \alpha_i \gamma_x^2(|k - i|), k \geq 1. \quad (2.7)$$

The equations of Corollary 1 are obviously nearly identical to the Yule Walker equations of the standard AR(p) model. As a consequence, the model of order p can be identified with the help of the partial autocorrelation function (PACF).

3 Estimation and Simulation Comparison

In this section, we will investigate two methods for parameter estimation of the Gumbel GARCH(p, q) model based on a realization X_1, \dots, X_n of the process. These estimators are compared via Monte Carlo simulations in terms of their means and standard deviations.

3.1 Yule-Walker Estimator

Let $\hat{\gamma}_x^2(k) = \frac{1}{n} \sum_{t=1}^{n-k} (X_t^2 - \overline{X^2}) (X_{t+k}^2 - \overline{X^2})$, $0 \leq k < n$, be the sample autocovariance function of $\{X_t^2\}$, where $\overline{X^2} = \frac{1}{n} \sum_{t=1}^n X_t^2$ is the sample mean. The Yule-Walker (YW) estimators of the model are obtained by substituting the sample $\hat{\gamma}_x^2(k)$ in eqn (2.7) and solving them.

Example 1. Consider the Gumbel ARCH (2) as

$$X_t | \mathcal{F}_{t-1} \sim \text{Gumbel}(0, \gamma_t),$$

$$\text{Var}(X_t | \mathcal{F}_{t-1}) = (\pi^2/6)\gamma_t^2 = \sigma_t^2 = \alpha_0 + \sum_{i=1}^2 \alpha_i X_{t-i}^2.$$

The explicit YW estimators of α_1 and α_2 are

$$\hat{\alpha}_1 = \frac{\rho_x^2(1)(1 - \rho_x^2(2))}{(1 - \rho_x^2(1))(1 + \nu^2 6/\pi^2)},$$

$$\hat{\alpha}_2 = \frac{\rho_x^2(2) - (\rho_x^2(1))^2}{(1 - (\rho_x^2(1))^2)^2 (1 + \nu^2 6/\pi^2)}$$

where $\rho_x^2(k) = \frac{\hat{\gamma}_x^2(k)}{\hat{\gamma}_x^2(0)}$. Also note that α_0 can be estimated from eqn (2.3).

3.2 Maximum Likelihood

Here we derive the maximum likelihood estimator (MLE) of the unknown parameter θ where $\theta = [\alpha_0, \alpha_1, \dots, \alpha_p, \beta_1, \beta_2, \dots, \beta_q]$. The MLE of the parameter is obtained by maximization of the conditional log-likelihood function

$$\ell(\theta) = \sum_{t=1}^n \ell_t(\theta) = \sum_{t=1}^n \left[-\log \gamma_t + \frac{X_t}{\gamma_t} - \exp\left(-\frac{X_t}{\gamma_t}\right) \right],$$

where

$$\gamma_t = \left(\frac{6}{\pi^2} \sigma_t^2 \right)^{\frac{1}{2}},$$

and

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2.$$

Solving the system of equations $\frac{\partial \ell(\theta)}{\partial \theta} = 0$, the MLE of θ is obtained. This can be done by using standard nonlinear maximization procedures which may be found in most of the statistical and data analysis packages.

3.3 Simulation

Here we have carried out two simulation studies. In the first study, we compare the two method of estimation. To examine the performances of the YW and ML estimators, a Monte Carlo simulation is conducted for different sample sizes $n = 100, 300, 500$ with $m = 200$ replications for models Gumbel ARCH(1), Gumbel ARCH(2) and Gumbel GARCH(1,1) for true parameter values:

- Gumbel ARCH(1): $(\alpha_0, \alpha_1) = A1(3, 0.5); A2(4, 0.3);$
- Gumbel ARCH(2): $(\alpha_0, \alpha_1, \alpha_2) = B1(3, 0.4, 0.2); B2(4, 0.3, 0.3);$
- Gumbel GARCH(1,1): $(\alpha_0, \alpha_1, \beta_1) = C1(2, 0.4, 0.3); C2(4, 0.3, 0.3).$

For the maximization of the log-likelihood function, the YW estimates were used as the initial values. Table 1 provides the mean and mean absolute deviation error (MADE) of the estimators for different values of the parameters and different sample sizes. As for the stationarity discussed, $\sum_{i=1}^p \alpha_i(1 + \nu^2 \frac{6}{\pi^2}) + \sum_{j=1}^q \beta_j$ in the models are all less than one and the stationary condition holds for the fitted models. It can be seen that as the sample size increases, the estimates seem to converge to the true parameter values. Two estimation methods seem to perform reasonably well but the MLEs provide better performance, which was expected.

We compare the two models with respect to their general properties and coherent forecasting ability. To achieve this, we have simulated 200 series, each of size 100 from Gumbel GARCH(1,1) process with three sets of parameter values viz.; (a) $\alpha_0 = 2, \alpha_1 = 0.2, \beta_1 = 0.3$, (b) $\alpha_0 = 3, \alpha_1 = 0.3, \beta_1 = 0.4$ and (c) $\alpha_0 = 3, \alpha_1 = 0.4, \beta_1 = 0.6$. From Table 2 it can be observed that the values of the performance measures mean square prediction error (MSPE) for the Gumbel GARCH(1,1) model are relatively lower, supporting the fact that if the actual process is Gumbel GARCH(1,1), then it gives a better fit (less prediction error) than classic GARCH(1,1). It can be also seen that, as k in k -step ahead forecasts increases, the values of the measures increases, indicating that the error in forecast increases as lag increases.

4 Real Example

In this section, we discuss some possible applications of the Gumbel GARCH model for a real time series of weekly data of Tehran Price Index (TEPIX). The data consist of 152 observation from 11 Nov. 2012 to 9 Sep. 2015. The data are obtained from the website of <http://www.tse.ir>. The sample paths, difference of log data, autocorrelation functions (ACFs) and partial autocorrelation functions (PACFs) of the difference log series are displayed in Figure 1. For selecting the model for the data series, we compare the classical ARCH(1), ARCH(2), ARCH(3), GARCH(1,1), GARCH(1,2), GARCH(1,3), Gumbel ARCH(1), Gumbel ARCH(2), Gumbel ARCH(3), Gumbel GARCH(1,1), Gumbel GARCH(1,2), Gumbel GARCH(1,3) models. For each model, based on MLEs we provide some well-known measures of goodness-of-fit statistics to check the adequacy of a time series model among a finite set of models. These statistics are the Akaike information criterion (AIC) and Bayesian information criterion (BIC). The obtained results, for

the data series, are shown in Table 3. As it can be seen from this table, the values of the goodness-of-fit statistics are the smallest for the Gumbel GARCH(1,3) model. The Pearson residual for Gumbel GARCH(1,3) is defined by

$$\frac{X_n - E(X_n|\mathcal{F}_{n-1})}{\sqrt{Var(X_n|\mathcal{F}_{n-1})}}$$

where

$$E(X_n|\mathcal{F}_{n-1}) = \nu\hat{\gamma}_n; \quad Var(X_n|\mathcal{F}_{n-1}) = \frac{1}{6}\pi^2\hat{\gamma}_n^2.$$

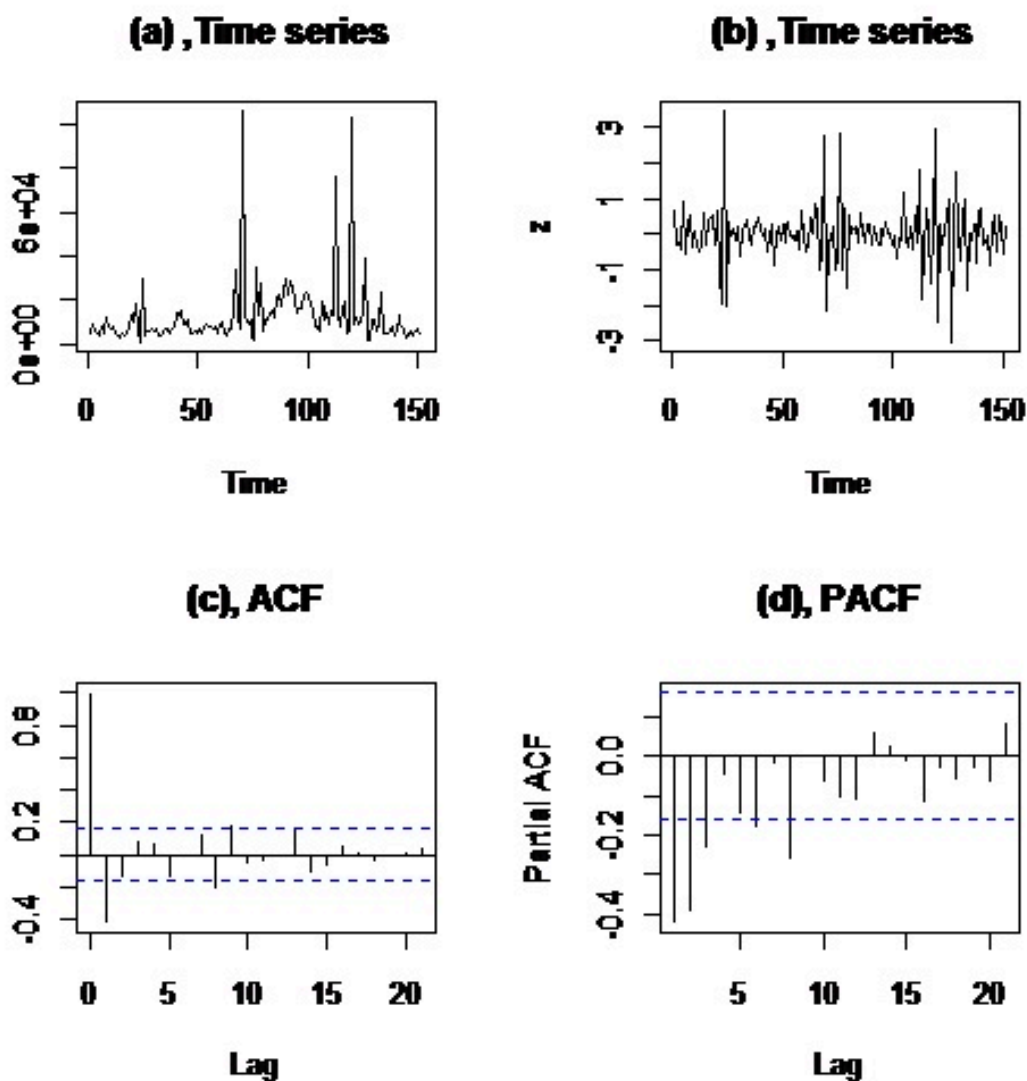


Figure 1: (a) Sample path of the time series, (b) Difference log time series, (c) ACF, (d) PACF

The residual analysis is shown in Figure 2. Figure 2 shows density and normal Q-Q plot for Pearson residuals which appear to be almost normally distributed. The Ljung-Box

statistic is 15.0825 by 15 lags ($\chi_{0.05}^2(14) = 23.6848$). The results show that the residuals are independent.

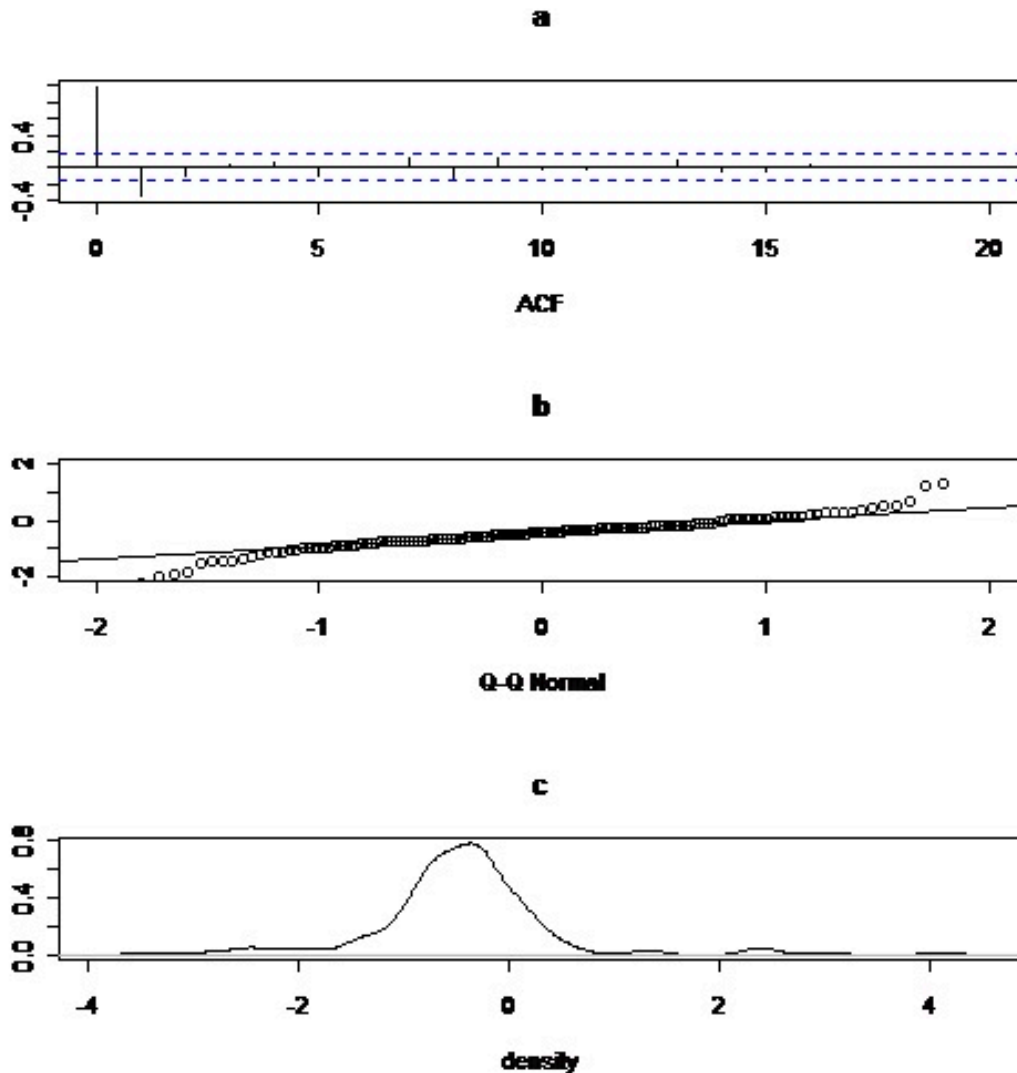


Figure 2: (a) ACF, (b) Q-Q plot, (c) density of Residuals

5 Conclusion

This article discusses the financial time series modeling with potential extreme observations. The Gumbel GARCH model, a generalization of the classic GARCH model is proposed to modeling. Stationarity conditions are given as well as the autocorrelation function. For estimation, we present two approaches with the focus on the maximum likelihood method. Simulation results show that two estimation methods are sufficiently accurate but the MLEs provide better performance. Results on the real stock data indicate that the proposed method performs better than the classic GARCH model.

Table 1: Mean (\pm MADE) of the estimators for different values of the parameters

| Model | n | Method | α_0 | α_1 | α_2 or β_1 | $\sum_{i=1}^6 (1 + \nu^2 \frac{6}{\pi^2}) \alpha_i + \beta_1$ |
|-------|-----|--------|---------------------|---------------------|-------------------------|---|
| A1 | 100 | YW | 3.2348 ± 0.3991 | 0.4431 ± 0.1284 | | |
| | | ML | 3.1525 ± 0.3580 | 0.4773 ± 0.1209 | | 0.2324 |
| | 300 | YW | 3.1866 ± 0.3272 | 0.4593 ± 0.1075 | | |
| | | ML | 3.1540 ± 0.2660 | 0.4834 ± 0.0858 | | 0.2368 |
| | 500 | YW | 3.1475 ± 0.2621 | 0.4741 ± 0.0863 | | |
| | | ML | 3.0663 ± 0.2171 | 0.5012 ± 0.0512 | | 0.2455 |
| A2 | 100 | YW | 5.1435 ± 1.1690 | 0.2524 ± 0.0799 | | |
| | | ML | 4.6023 ± 0.8525 | 0.2660 ± 0.0587 | | 0.1303 |
| | 300 | YW | 4.7894 ± 0.8179 | 0.2721 ± 0.0612 | | |
| | | ML | 4.3769 ± 0.6936 | 0.2811 ± 0.0551 | | 0.1377 |
| | 500 | YW | 4.4469 ± 0.4858 | 0.2920 ± 0.0468 | | |
| | | ML | 4.1923 ± 0.4759 | 0.2964 ± 0.0436 | | 0.1452 |
| B1 | 100 | YW | 2.5347 ± 0.6879 | 0.3324 ± 0.1212 | 0.1438 ± 0.1233 | |
| | | ML | 2.8489 ± 0.5592 | 0.3756 ± 0.1225 | 0.1772 ± 0.1020 | 0.3612 |
| | 300 | YW | 2.6893 ± 0.5626 | 0.3466 ± 0.1028 | 0.1621 ± 0.0961 | |
| | | ML | 2.9035 ± 0.4659 | 0.3723 ± 0.0838 | 0.1893 ± 0.0768 | 0.3717 |
| | 500 | YW | 2.8767 ± 0.4542 | 0.3595 ± 0.0865 | 0.1794 ± 0.0740 | |
| | | ML | 3.0288 ± 0.3821 | 0.3902 ± 0.0753 | 0.1988 ± 0.0598 | 0.3899 |
| B2 | 100 | YW | 5.6798 ± 2.1657 | 0.2731 ± 0.1643 | 0.2432 ± 0.1905 | |
| | | ML | 4.8586 ± 1.6984 | 0.2762 ± 0.1593 | 0.2463 ± 0.1732 | 0.3816 |
| | 300 | YW | 5.0924 ± 1.6661 | 0.2794 ± 0.1528 | 0.2653 ± 0.1716 | |
| | | ML | 4.9401 ± 1.5831 | 0.2804 ± 0.1468 | 0.2695 ± 0.1690 | 0.4068 |
| | 500 | YW | 4.5712 ± 1.1649 | 0.2812 ± 0.1417 | 0.2851 ± 0.1563 | |
| | | ML | 4.4499 ± 1.1028 | 0.2827 ± 0.1377 | 0.2876 ± 0.1539 | 0.4261 |
| C1 | 100 | YW | 3.5619 ± 1.6097 | 0.2761 ± 0.0634 | 0.1579 ± 0.2611 | |

continued ...

...continued

| Model | n | Method | α_0 | α_1 | α_2 or β_1 | $\sum_{i=1} (1 + \nu^2 \frac{6}{\pi^2}) \alpha_i + \beta_1$ | |
|-------|-----|--------|---------------------|---------------------|-------------------------|---|--------|
| C2 | 300 | ML | 2.4378 ± 0.3580 | 0.3673 ± 0.0437 | 0.2379 ± 0.1842 | 0.4178 | |
| | | YW | 2.8406 ± 0.9385 | 0.3022 ± 0.0420 | 0.2029 ± 0.2138 | | |
| | 500 | ML | 2.2283 ± 0.4981 | 0.3767 ± 0.0208 | 0.2652 ± 0.1860 | 0.4498 | |
| | | YW | 2.2464 ± 0.3405 | 0.3243 ± 0.0299 | 0.2571 ± 0.1746 | | |
| | 100 | ML | 2.1563 ± 0.1184 | 0.3854 ± 0.0184 | 0.2983 ± 0.1545 | 0.4887 | |
| | | YW | 5.4571 ± 2.7082 | 0.2785 ± 0.1701 | 0.1376 ± 0.2537 | | |
| | 300 | ML | 4.4554 ± 2.2111 | 0.2898 ± 0.1711 | 0.2617 ± 0.1796 | 0.4037 | |
| | | YW | 4.9820 ± 2.2210 | 0.2876 ± 0.1518 | 0.1907 ± 0.1911 | | |
| | 500 | ML | 4.4697 ± 2.0628 | 0.2967 ± 0.1546 | 0.2778 ± 0.1859 | 0.4231 | |
| | | YW | 4.5606 ± 1.7099 | 0.2898 ± 0.1346 | 0.2327 ± 0.1339 | | |
| | | | ML | 4.0485 ± 1.5657 | 0.2904 ± 0.1365 | 0.2978 ± 0.1298 | 0.4400 |

Table 2: MSPE of the data generated from Gumbel GARCH(1,1) model for k -step ahead forecasts

| k | GARCH(1,1) | Gumbel GARCH(1,1) | GARCH(1,1) | Gumbel GARCH(1,1) | GARCH(1,1) | Gumbel GARCH(1,1) |
|-----|---|-------------------|---|-------------------|---|-------------------|
| | $(\alpha_0, \alpha_1, \beta_1) = (2, 0.2, 0.3)$ | | $(\alpha_0, \alpha_1, \beta_1) = (3, 0.3, 0.4)$ | | $(\alpha_0, \alpha_1, \beta_1) = (3, 0.4, 0.6)$ | |
| 1 | 0.84 | 0.69 | 0.76 | 0.67 | 0.72 | 0.67 |
| 2 | 0.93 | 0.82 | 0.90 | 0.83 | 0.81 | 0.77 |
| 3 | 1.04 | 0.94 | 1.01 | 0.90 | 0.93 | 0.86 |
| 4 | 1.10 | 0.97 | 1.12 | 1.03 | 1.07 | 1.01 |
| 5 | 1.13 | 1.02 | 1.18 | 1.08 | 1.20 | 1.10 |

Table 3: Estimated parameters, AIC and BIC for the TEPIX time series

| Model | α_0 | α_1 | α_2 | α_3 | β_1 | β_2 | β_3 | AIC | BIC |
|-------------------|------------|------------|------------|------------|-----------|-----------|-----------|-------|-------|
| ARCH(1) | 0.4362 | 0.3748 | | | | | | 350.6 | 356.6 |
| ARCH(2) | 0.4213 | 0.3106 | 0.2219 | | | | | 339.4 | 345.5 |
| ARCH(3) | 0.4147 | 0.3216 | 0.3105 | 0.2221 | | | | 335.6 | 340.7 |
| GARCH(1,1) | 0.3927 | 0.2520 | | | 0.1984 | | | 350.9 | 361.3 |
| GARCH(1,2) | 0.3524 | 0.2212 | | | 0.1874 | 0.1523 | | 351.5 | 363.9 |
| GARCH(1,3) | 0.3021 | 0.1980 | | | 0.1722 | 0.1281 | 0.1125 | 347.7 | 352.2 |
| Gumbel ARCH(1) | 0.4993 | 0.3216 | | | | | | 318.2 | 322.7 |
| Gumbel ARCH(2) | 0.4325 | 0.2992 | 0.2579 | | | | | 315.3 | 319.7 |
| Gumbel ARCH(3) | 0.4091 | 0.2537 | 0.2280 | 0.1852 | | | | 312.4 | 317.1 |
| Gumbel GARCH(1,1) | 0.3423 | 0.2758 | | | 0.1980 | | | 306.4 | 314.9 |
| Gumbel GARCH(1,2) | 0.3219 | 0.2227 | | | 0.1825 | 0.1653 | | 304.7 | 311.5 |
| Gumbel GARCH(1,3) | 0.3051 | 0.2113 | | | 0.1812 | 0.1521 | 0.1210 | 303.2 | 307.4 |

References

- [1] Baillie, R. T. and Bollerslev, T. (1989): Common stochastic trends in a system of exchange rates. *The Journal of Finance*, **44**, 167–181.
- [2] Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.
- [3] Calzolari, G., Halbleib, R. and Parrini, A. (2014): Estimating GARCH-type models with symmetric stable innovations: Indirect inference versus maximum likelihood. *Computational Statistics and Data Analysis*, **76**, 158–171.
- [4] Goldberg, S. (1958): *Introduction to Difference Equations, with Illustrative Examples from Economics, Psychology, and Sociology*. New York: Wiley.
- [5] Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993): On the relation between the expected value and the volatility of the nominal excess return of stocks. *Journal of Finance*, **48**, 1779–1801.
- [6] Hsieh, D. A. (1988): The statistical properties of daily foreign exchange rates, 1974–1983. *Journal of International Economics*, **24**, 129–145.
- [7] Jorion, P. (1988): On jump processes in the foreign exchange and stock markets. *Review of Financial Studies*, **1**(4), 427–445.
- [8] Liu, S. and Brorsen, B. (1995): Maximum likelihood estimation of a GARCH-stable model. *Journal of Applied Econometrics*, **10**(3), 273–285.
- [9] Nelson, D. B. (1990a): ARCH models as diffusion approximation, *Journal of Econometrics*, **45**, 7–38.
- [10] Nelson, D. B. (1990b): Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, **59**, 347–370.
- [11] Nelson, D. B. and Cao, C. Q. (1992): Inequality constraints in the univariate GARCH model. *Journal of Business and Economic Statistics*, **10**(2), 229–235.

A Appendix

Let \mathcal{I}_t be the σ -field generated by $\{\sigma_t^2, \sigma_{t-1}^2, \dots\}$, then we have

$$E(X_t^2 | \mathcal{F}_{t-1}, \mathcal{I}_t) = E(X_t^2 | \mathcal{F}_{t-1}) = \sigma_t^2(1 + \nu^2 6/\pi^2),$$

and

$$\sigma^2 := E(X_t^2) = E(E(X_t^2 | \mathcal{F}_{t-1})) = E(\sigma_t^2(1 + \nu^2 6/\pi^2)).$$

For $k \geq 0$, since

$$\begin{aligned}
& Cov(X_t^2 - \sigma_t^2(1 + \nu^2 6/\pi^2), \sigma_{t-k}^2(1 + \nu^2 6/\pi^2)) \\
&= E[(X_t^2 - \sigma_t^2(1 + \nu^2 6/\pi^2))(\sigma_{t-k}^2(1 + \nu^2 6/\pi^2) - \sigma^2)] \\
&= E[(\sigma_{t-k}^2(1 + \nu^2 6/\pi^2) - \sigma^2)E(X_t^2 - \sigma_t^2(1 + \nu^2 6/\pi^2)|\mathcal{I}_t)] \\
&= E[(\sigma_{t-k}^2(1 + \nu^2 6/\pi^2) - \sigma^2)E(E(X_t^2|\mathcal{F}_{t-1}, \mathcal{I}_t)|\mathcal{I}_t - \sigma_t^2(1 + \nu^2 6/\pi^2))] \\
&= E(\sigma_{t-k}^2(1 + \nu^2 6/\pi^2) - \sigma^2)E(\sigma_t^2(1 + \nu^2 6/\pi^2)|\mathcal{I}_t - \sigma_t^2(1 + \nu^2 6/\pi^2)) \\
&= 0,
\end{aligned}$$

we obtain

$$Cov(X_t^2, \sigma_{t-k}^2(1 + \nu^2 6/\pi^2)) = Cov(\sigma_t^2(1 + \nu^2 6/\pi^2), \sigma_{t-k}^2(1 + \nu^2 6/\pi^2)).$$

Similarly for $k < 0$, since

$$\begin{aligned}
& Cov(X_t^2, X_{t-k}^2 - \sigma_{t-k}^2(1 + \nu^2 6/\pi^2)) \\
&= E[(X_t^2 - \sigma^2)(X_{t-k}^2 - \sigma_{t-k}^2(1 + \nu^2 6/\pi^2))] \\
&= E[(X_t^2 - \sigma^2)E((X_{t-k}^2 - \sigma_{t-k}^2(1 + \nu^2 6/\pi^2))|\mathcal{F}_{t-k-1})] \\
&= E[(X_t^2 - \sigma^2)(\sigma_{t-k}^2(1 + \nu^2 6/\pi^2) - E(\sigma_{t-k}^2(1 + \nu^2 6/\pi^2)|\mathcal{F}_{t-k-1}))] \\
&= 0,
\end{aligned}$$

we have

$$Cov(X_t^2, \sigma_{t-k}^2(1 + \nu^2 6/\pi^2)) = Cov(X_t^2, X_{t-k}^2).$$

Therefore

$$Cov(X_t^2, \sigma_{t-k}^2(1 + \nu^2 6/\pi^2)) = \begin{cases} (1 + \nu^2 6/\pi^2)^2 Cov(\sigma_t^2, \sigma_{t-k}^2), & k \geq 0 \\ Cov(X_t^2, X_{t-k}^2), & k < 0 \end{cases}.$$

Let us now derive $\gamma_{\sigma^2}^2(k)$. For $k \geq 0$, since

$$\begin{aligned}
\gamma_{\sigma^2}^2(k) &= Cov(\sigma_t^2, \sigma_{t-k}^2) = \sum_{i=1}^p Cov(X_{t-i}^2, \sigma_{t-k}^2) + \sum_{j=1}^q \beta_j Cov(\sigma_{t-j}^2, \sigma_{t-k}^2) \\
&= \sum_{i=1}^p (1 + \nu^2 6/\pi^2)/(1 + \nu^2 6/\pi^2) \alpha_i Cov(X_{t-i}^2, \sigma_{t-k}^2) + \sum_{j=1}^q \beta_j Cov(\sigma_{t-j}^2, \sigma_{t-k}^2) \\
&= \sum_{i=1}^p \alpha_i/(1 + \nu^2 6/\pi^2) Cov(X_{t-i}^2, (1 + \nu^2 6/\pi^2)\sigma_{t-k}^2) + \sum_{i=1}^p \beta_j Cov(\sigma_{t-j}^2, \sigma_{t-k}^2),
\end{aligned}$$

we have

$$\begin{aligned}
\gamma_{\sigma^2}^2(k) &= \sum_{i=1}^{\min(k,p)} (1 + \nu^2 6/\pi^2) \alpha_i Cov(\sigma_{t-i}^2, \sigma_{t-k}^2) + \\
&\quad \sum_{i=1}^p \alpha_i/(1 + \nu^2 6/\pi^2) Cov(X_{t-i}^2, X_{t-k}^2) + \sum_{j=1}^q \beta_j Cov(\sigma_{t-j}^2, \sigma_{t-k}^2).
\end{aligned}$$

For the case, when $k \geq 1$, we obtain $\gamma_x^2(k)$ as

$$\begin{aligned}
\gamma_x^2(k) &= Cov(X_t^2, X_{t-k}^2) = E(X_t^2, X_{t-k}^2) = E(E(X_t^2, X_{t-k}^2) | \mathcal{F}_{t-1}) \\
&= E(E(X_t^2 | \mathcal{F}_{t-1}, X_{t-k}^2)) = E((1 + \nu^2 6 / \pi^2) \sigma_t^2, X_{t-k}^2) \\
&= (1 + \nu^2 6 / \pi^2) E(\alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, X_{t-k}^2) \\
&= (1 + \nu^2 6 / \pi^2) (\sum_{i=1}^p \alpha_i E(X_{t-i}^2, X_{t-k}^2) + \sum_{j=1}^q \beta_j E(\sigma_{t-j}^2, X_{t-k}^2)) \\
&= \sum_{i=1}^p (1 + \nu^2 6 / \pi^2 \alpha_i Cov(X_{t-i}^2, X_{t-k}^2)) + \sum_{j=1}^{\min(k-1, q)} \beta_j Cov(X_{t-j}^2, X_{t-k}^2) \\
&\quad + \sum_{j=1}^q (1 + \nu^2 6 / \pi^2)^2 \beta_j Cov(\sigma_{t-j}^2, \sigma_{t-k}^2). \quad \square
\end{aligned}$$