Review on Three New Value at Risk (VaR) Models

Heying Liu[1,a,*]

[1]College of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao,Hebei, China, 066000
a. lyee_alice@163.com
**corresponding author*

Abstract: The emergence of financial derivatives complicates traditional financial products and increases financial market volatility. Individuals and financial institutions are both exposed to more complex and uncontrollable risks in this environment. Because of the risk's uncertainty, we must use reasonable methods to predict and estimate it in order to achieve the goal of risk control. This paper discusses three new VaR (Value at Risk) models that have emerged in recent years based on the ARCH family model using a method of literature review. The ARMA-EGARCH model, for example, combines the ARMA model to describe constant variance time series and the EGARCH model to describe heteroscedasticity phenomena, and theoretically can better describe the fluctuations of financial time series and obtain an independent time series with the same distribution. The sequence is processed using extreme value theory, which is the ARMA-EGARCH-GPPD model, in conjunction with the GPD model. We used the ARMA-EGARCH-semi-parametric method in conjunction with the historical simulation method and the parameter method to avoid cumbersome quantile calculation because the model algorithm is more complex. The generalized EWMA risk value prediction model has more advantages for financial data with large peaks.

Keywords: VaR, ARCH series model, ARMA-EGARCH-GPPD model, generalized-EWMA model

1. Introduction

In the last century, there have been unprecedented developments in international financial markets, which are now beginning to take place. The emergence of financial derivatives is one of the most significant changes in comparison to previous financial markets. Derivatives have increased the complexity of traditional financial products, resulting in increased volatility in financial markets and an unstoppable trend toward globalization. Under such conditions, both individuals and financial institutions face more complex and uncontrollable risks, and the financial market's whims and whims will have an impact on human economic life worldwide. We must proactively manage uncontrollable and uncertain risks rather than simply avoiding them or doing nothing at all. Only through the use of proper methods and active control can investors and financial institutions reduce losses and grow more steadily. Value at risk is a risk index, statistical data, and measurement method [1]. VAR enables people to intuitively understand their maximum loss value and probability in the future, allowing them to address and manage risks. Risk is quantified by the value at risk. It is simply the proportion of assets that are at risk when market volatility is normal. This is the risky value.

VaR is a relatively new approach of gauging financial market risk, yet it has become a research hotspot. G-30, an international private research company, originally proposed regulating risk via VaR in 1993 [2]. VaR prediction methods include historical simulation (HS), moving average (MA), and exponentially weighted moving average (EWMA) (ARCH model). Because the ARCH model depicts the fluctuation clustering of financial time series, people have promoted it, forming a general ARCH model and ARCH series model. These models have different advantages in characterizing financial time series, so scholars will choose different models according to the different needs of time series when estimating VaR, so a large number of VaR models with different manifestations. Ning Tang proposed the ARMA-EGARCH-GPPD model in 2016 [2]. Aerambamoorthy Thavaneswaran, Alex Paseka, and Julieta Frank suggested a data-driven generalized EWMA model to reduce the asymptotic variance of the volatility estimator [3]. Hemant Kumar Badaye and Jason Narsoo proposed the MC-GARCH-Copula model, which can predict VaR and ES for the next 1 min [4]. Tianyi Wang et al. introduced the GARCH-RSRK model, which determines return, realized volatility, realized skewness, and realized kurtosis [5].

Starting from the theoretical foundation of value-at-risk, this paper introduces several value-at-risk models in recent years, and compares the model's shortcomings and advantages using the method of comparison, in order to look into the future development direction of value-at-risk theory.

2. Value at Risk Theory

VaR is a statistical estimate of an asset's loss in a given risk range caused by normal market fluctuations during a holding period. VaR can be defined by financial institutions as the maximum loss of financial assets with a given probability over a given time period. In the case of normal market fluctuations, VaR is defined as the maximum loss that is expected to occur in the value of a specific financial asset or security portfolio within a specific period in the future under a given probability level (confidence level). For example, if a company has a $100,000 value at risk and a 95% confidence level, there is a 5% chance that each of the company's $1 million assets will suffer a maximum loss of $100,000 in the next 24 hours. People can make decisions based on predicted data as well as their own risk preference or aversion.

Two parameters must be determined to calculate value at risk: asset holding period and result confidence. Confidence level is loss probability interval. The likelihood increases the risk value. Higher confidence means the loss is less likely to surpass the projected value at risk. There is no hard and fast rule for determining whether a higher or lower confidence level is preferable. Because a high confidence level may overestimate the risk, a low confidence level may underestimate the risk. In general, a confidence level of 90% to 99% is appropriate. The holding period can be set to be a day, a week, a month, or a year, among other options. If the holding period is excessively long, an assumption should be made about the data, namely that the sample data are independent and identically distributed. If the data is relatively large, the holding period could be as short as one hour or as long as one minute. As a result, the holding period should be chosen based on the characteristics of the data and the needs of the users.

If x represents the return of the user's venture capital, x is a random variable and its probability density function is $f(x)$, then the value at risk with confidence level α is defined as[3]:

$$P(x \leq VaR) = \alpha \tag{2.1}$$

$$\alpha = \int_{-\infty}^{VaR} f(x)\, dx \tag{2.2}$$

3. Introduction to Several Types of VaR Estimation Models

There are three types of VAR estimation methods: non-parametric, semi-parametric, and parametric

DOI: 10.54254/2754-1169/17/20231068

[6]. More computational models are derived from this, which play a role in various financial markets. This paper begins with ARCH family models, then discusses three models based on these models that have emerged in recent years, as well as their benefits and drawbacks.

3.1. ARCH Series Model

To calculate VaR, the variance and standard deviation of the return rate, i.e. its volatility, must be estimated. Because the default volatility of the return rate did not change in the early studies, the volatility of the return rate was set to a constant. However, according to additional research and reflection on practical problems, volatility is not a constant, but rather changes with time and has heap and aggregation characteristics.

3.1.1. ARCH Model

The ARCH model was first proposed by Engle [7], whose autoregressive conditional abnormal difference model can provide a better explanation and analysis of volatility. He defined $ARCH(q)$ model as:

$$y_t = u_t + \varepsilon_t (t = 1,2,\cdots,n) \tag{3.1}$$

$$\varepsilon_t = \sigma_t e_t, E(e_t) = 0, D(e_t) = 1 \tag{3.2}$$

$$\sigma_t{}^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}{}^2 + \alpha_2 \varepsilon_{t-2}{}^2 + \cdots + \alpha_q \varepsilon_{t-q}{}^2 \tag{3.3}$$

The formula (3.1) is called the mean equation of $ARCH(q)$ model. ε_t is the random perturbation term of the sequence at time t. e_t is a sequence of independent same-distribution residuals. The formula (3.3) is called the conditional equation of variance.

However, for this model, if you want to get a better fitting effect, the lag order q will be quite large, which increases the difficulty and accuracy of the model fitting. And in practice, σ_t^2 does not necessarily satisfy the linear function conditions in the equations of variance.

3.1.2. GARCH Model

To improve on both shortcomings, Bollerslav generalized the ARCH model to the generalized ARCH model (GARCH model). The definition of the $GARCH(p,q)$ model is based on the ARCH model definition, and the conditional equation of variance is changed as:

$$\sigma_t{}^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}{}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}{}^2 \tag{3.4}$$

where the coefficient $\alpha_0 > 0, \alpha_i \geq 0 (i = 1,2,\ldots,q), \beta_j \geq 0 (j = 1,2,\ldots,p)$. This avoids the defect of large lagging order and linear relationship on the basis of the original $ARCH(q)$.

3.1.3. TARCH Model

After a deeper examination of the GARCH model, it was discovered that the condition that the coefficient in this model was not negative was difficult to guarantee, but many financial time series had obvious leverage effects, which the model was unable to characterize, so Zakoian proposed the TARCH model in the study. The TARCH model retains the GARCH model's mean equation but replaces the variance equation to:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i \varepsilon_{t-i}{}^2 + \varphi_i \varepsilon_{t-1}^2 I_{t-1}) + \sum_{j-1}^{p} \beta_j \sigma_{t-j}^2 \tag{3.5}$$

DOI: 10.54254/2754-1169/17/20231068

Where when ε_{t-1} is not negative, $I_{t-1}=1$; when ε_{t-1} is negative, $I_{t-1}=0$. This reflects the effect of the ε_{t-1} symbol on σ_t^2.

3.1.4. EGARCH Model

By 1991, Nelson had proposed the EGARCH model, which had similar advantages to the TARCH model in terms of characterizing the volatility of financial time series. The $EGARCH(p,q)$ model is defined as:

$$y_t = u_t + \varepsilon_t (t = 1,2,\cdots,n) \tag{3.6}$$

$$\varepsilon_t = \sigma_t e_t, E(e_t) = 0, D(e_t) = 1 \tag{3.7}$$

$$log\, \sigma_t{}^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i \left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| + \varphi_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}}) + \sum_{j=1}^{p} \beta_j\, log\, \sigma_{t-j}{}^2 \tag{3.8}$$

There is no requirement for positive or negative coefficients. Also, $\alpha_i \left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right|$ and $\varphi_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}}$ reflect the magnitude of ε_{t-i} and the effect of positive and negative changes on the variance of conditions

3.1.5. ARMA-EGARCH Model

Financial time series show heteroscedasticity as a whole, but they show constant variance from the perspective of individual time periods. In practical applications we can use the ARMA model (Autoregreesive-Moving Average Model) to describe time series with constant variance. The general form of the $ARMA(p,q)$ model [8] is:

$$R_t = c_0 + \sum_{i=1}^{p} \phi_i R_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{3.9}$$

Where $\varepsilon_t \sim WN(0,\sigma^2)$, and satisfies $\forall s < t, E(R \bullet \varepsilon_t) = 0$.

In fact, the EGARCH model can accurately describe the variance time variability and leverage effect of the majority of financial time series. As a result, the two models can be combined, first with the ARMA model to characterize the mean equation, and then with the EGARCH model to depict the residual after the ARMA model, to theoretically better describe the fluctuations of the financial time series. Combined with the general expression of the ARMA and EGARCH models, the $ARMA(p,q)-EGARCH(m,n)$ model can be expressed as:

$$R_t = u_t + \varepsilon_t, u_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{i-1} + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} \tag{3.10}$$

$$\varepsilon_t = \sigma_t e_t, E(e_t) = 0, D(e_t) = 1 \quad (3.11) \tag{3.11}$$

$$log\, \sigma_t{}^2 = \phi_0 + \sum_{\kappa=1}^{m} (\phi_\kappa \left|\frac{\varepsilon_{t-\kappa}}{\sigma_{t-\kappa}}\right| + \varphi_\kappa \frac{\varepsilon_{t-\kappa}}{\sigma_{t-\kappa}}) + \sum_{\xi=1}^{n} \theta_\xi\, log\, \sigma_{t-\xi}{}^2 \tag{3.12}$$

where e_t is an independent residuals sequence with the same distribution; σ_t is the standard deviation of the sequence R_t; ε_t is a random perturbation term that enters the system at t moment, and satisfies:

DOI: 10.54254/2754-1169/17/20231068

3.2. ARMA-EGARCH-GPD Model

3.2.1. Generalized Pareto Distribution

Generalized Pareto distribution (GPD) [9], is a right-skewed distribution that is generally expressed as [9]:

$$G_{\tau,\gamma}(y) = \begin{cases} 1-(1+\frac{\tau y}{\gamma})^{-\frac{1}{\tau}} & \tau \neq 0 \\ 1-e^{-\frac{y}{\gamma}} & \tau=0 \end{cases} \tag{3.14}$$

Where $y \in \begin{cases} [0, x_F - u] & \tau \geq 0 \\ [0, -\frac{y}{\tau}] & \tau<0 \end{cases}$, x_F is the right endpoint of the distribution function F.

Let the random variable X of the independent same distribution satisfy the unknown distribution as $F(X)$, and when given a determined threshold of u, if you make $y=x-u$, there is a conditional super-threshold distribution function $F_u(y)=P(x-u \leq y | x>u)$. If the threshold value u is large enough, the conditional over-threshold distribution function F_u is similar to GPD, that is $F_u(y) \approx G_{\tau,\gamma}(y)$.

From this equation (3.14) can be deformed to:

$$F_u(y) = \frac{F(x)-F(y)}{1-F(u)} \tag{3.15}$$

In the equation(3.15), with $(1-\frac{n_u}{n})$ as the estimation of $F(u)$ (where n is the total number of samples and n_u is the number of numbers in the sample whose values are greater than the threshold of u, then the tail estimation function of the distribution function $F(X)$ can be obtained:

$$\hat{F}(x) = \begin{cases} 1-\frac{n_u}{n}[1+\frac{\hat{\tau}}{\hat{\gamma}}(x-u)]^{-\frac{1}{\hat{\tau}}} & \hat{\tau} \neq 0 \\ 1-\frac{n_u}{n}e^{-\frac{x-u}{\hat{\gamma}}} & \hat{\tau}=0 \end{cases} \tag{3.16}$$

where $\hat{\tau}, \hat{\gamma}$ are the estimation of parameter τ, γ in the GPD by the maximal likelihood method respectively. For sample $y_i(y_i = x_i - u, x_i > u)$ beyond the threshold u, the likelihood function of the generalized Pareto function is:

$$l = \begin{cases} -n_u \log\gamma + (\frac{1}{\tau}-1)\sum_{i=1}^{n_u} \log(1-\frac{\tau y_i}{\gamma}) & \tau \neq 0 \\ -n_u \log\gamma - \frac{1}{\gamma}\sum_{i=1}^{n_u} y_i & \tau=0 \end{cases} \tag{3.17}$$

If the two formulas of the tail estimation function are equal to $1-\alpha$ (α is the significant level) , the expression of the upper tail α quantile function VaR_α of the Pareto distribution can be regurgitated:

$$VaR_{\alpha} = \begin{cases} u + \frac{\hat{\gamma}}{\hat{\tau}}[(\frac{n}{n_u}\alpha)^{-\hat{\tau}} - 1] & \hat{\tau} \neq 0 \\ u - \hat{\gamma}\, ln^{\frac{n}{n_u}\alpha} & \hat{\tau}=0 \end{cases} \tag{3.18}$$

3.2.2. Introduction of ARMA-EGARCH-GPD Model

For independent time series with the same distribution, the extreme value theory can be applied to the unprocessed time series to calculate VaR. The general time series model is not independent of the same distribution, at this time it needs to be filtered by the ARCH series model, and the resulting residual sequence can meet the conditions of independent same distribution. Therefore, it is theoretically possible to combine extreme value theory and ARCH family models to study VaR. At the same time, with a large number of empirical analysis studies by scholars in recent years, the accuracy of this method in calculating VaR has also been confirmed. In the ARCH series model, the ARMA-EGARCH model has the dual advantages of ARMA model characterizing the constant variance and the EGARCH model characterizing heteroscedasticity, leverage effect, etc.

Therefore, when estimating VaR in this paper, the mean u_t , standard deviation σ_t and corresponding residual sequence $\{e_t\}$ of the t-time sequence $\{R_t\}$ are estimated by establishing an ARMA-EGARCH model for sequence $\{R_t\}$, and then use the GPD model to characterize the tail cloth of the $\{e_t\}$ sequence to obtain VaR_e , and finally get the estimate of VaR from $VaR_t = u_t + \sigma_t VaR_e$. Because this estimation model is a combination of the ARMA-EGARCH model and the GPD model, it is called the ARMA-EGARCH-GPD model [2].

3.2.3. Advantages and Disadvantages

The ARMA-EGARCH-GPD model is theoretically accurate, and no special restrictions are required for general time series to obtain more accurate VaR values. It overcomes the hypothetical linear function situation required by the ARCH model, the parameter non-negative condition required by the GARCH model, and can accurately and widely characterize the fluctuation situation and leverage effect of the sequence. The GPD model's depiction of the end of the sequence also makes the calculation of VaR more reasonable in comparison to the extreme value theory.

This method, however, is more time-consuming, and the thresholds in the GPD model cannot be too large or too small. Quantile estimation is time-consuming.

3.3. ARMA-EGARCH-semi-parametric Model

3.3.1. Introduction of ARMA-EGARCH-semi-parametric Model

ARMA-EGARCH-GPD model is more accurate in theory, but the method is more cumbersome, ARMA-EGARCH-semi-parameter methodcombines parameter method and historical simulation method on the basis of extreme value theory, avoiding cumbersome quantile calculation [2].

Similar to the previous section, we can estimate u_t and σ_t by building an ARMA-EGARCH model using parametric methods. According to VaR's general calculation formula

$$VaR_t = u_t + \sigma_t X_{1-\alpha} \tag{3.19}$$

only the $1-\alpha$ quantile $X_{1-\alpha}$ of the normalized distribution function of sequence $\{R_t\}$ is required. Inspired by the historical simulation method, we can first obtain VaR through the historical simulation method, and then substitute the general formula to find $X_{1-\alpha}$. The specific method is: the sequence $\{R_t\}$ is sorted from the largest to the smallest, take the sample point value corresponding to the total number of samples of $1-\alpha$ times as VaR_0, and set the standard deviation of the series to be σ_0 and the mean to be u, then:

$$X_{1-\alpha} = \frac{VaR_0 - u}{\sigma_0} \tag{3.20}$$

The estimate of $X_{1-\alpha}$ is obtained by historical simulation, and the VaR estimate is obtained by bringing u_t and σ_t obtained by the parameter method into the formula (3.19).

Because this method is to estimate VaR by using the ARMA-EGARCH model to estimate the parameter and nonparametric method to obtain quantiles, this model is called ARMA-EGARCH - semiparametric model.

3.3.2. Advantages and Disadvantages

ARMA-EGARCH-semi-parametricmodel combines the characteristics of the ARMA model to fit the constant variance and the EGARCH model to fit the heteroscedasticity effectively [10]. And instead of directly using the historical simulation method to find the quantile, but combining the parameter method on the basis of the historical simulation method, which avoids any assumptions about the distribution of the return, and can make the quantile more in line with the actual distribution. However, the method of estimating $X_{1-\alpha}$ using historical simulation is not necessarily optimal, so this model is not necessarily the optimal model for a particular sequence, and then you can study the estimation method of $X_{1-\alpha}$ to obtain the optimal model.

3.4. Generalized-EWMA Model

3.4.1. Introduction of Generalized-EWMA Model

Exponentially weighted moving average (EWMA) decreases the weighted coefficient of each value exponentially over time. The closer to the current moment, the bigger the numerical weighting coefficient. This approach does not need to save all prior data, unlike standard averaging methods. Second, the quantity of calculation has been minimized. In VaR estimation, for the square of the continuous compound return, EWMA can be used to get a prediction of conditional variance σ^2, and then take its square root to get an estimate of conditional volatility σ_p. But the square root of the variance is an inefficient estimate of volatility, so we can directly estimate volatility and get the optimal VaR prediction through a data-driven generalized EWMA model.

Because the conditional distribution of financial returns tends to be heavy tail with large peaks, and in VaR estimation, the mean of the return sequence μ is usually set to 0. By analogy with the GARCH model of conditional variance, based on the recursive form of volatility estimation

$$\hat{\sigma}_{n+1} = \frac{n}{n+1}\hat{\sigma}_n + \frac{1}{n+1}\frac{|r_{n+1}-\mu|}{\rho} \tag{3.21}$$

a data-driven generalized EWMA model of time-varying volatility is proposed [5]:

$$\hat{\sigma}_{t+1} = \alpha\hat{\sigma}_t + (1-\alpha)\frac{|r_t|}{\rho}, \alpha \in (0,1) \tag{3.22}$$

Therefore, we can predict the VaR value at the moment of t+1 with the following formula:

$$VaR_{t+1}(p) = -\sigma_{t+1}F_R^{-1}(p) \tag{3.23}$$

3.4.2. Advantages and Disadvantages

One of the obvious advantages of generalized EWMA models is that the estimation of ρ can be used to identify t-distributions with appropriate degrees of freedom, and this method avoids false specifications of the model. When the t distribution's degrees of freedom are fewer than 4, the TGARCH model's projected variance of the conditional variance becomes infinite, hence it can't be employed. Generalized EWMA can compensate for this issue, and its estimated asymptotic variance is substantially smaller than the standard estimate, making it ideal for financial data.

4. Conclusion

In this paper, the ARMA-EGARCH-GPD model is based on the relationship between the risk value of the original sequence and the risk value of the corresponding residual sequence: $VaR_t = u_t + \sigma_t VaR_e$, the ARMA-EGARCH model is first used to estimate the σ_t and u_t of the original time series, and then the GPD model in the extreme value theory is used to fit the residual sequence and estimate the VaR_e, so as to estimate the VaR of the original sequence. The ARMA-EGARCH-semi-parametric model is based on the general calculation formula of VaR (3.19), and similarly, the ARMA-EGARCH model is used to fit the time series, estimating to σ_t and u_t, and then using the historical simulation method to obtain $X_{1-\alpha}$ inversely, and finally obtaining the estimate of VaR. For models with infinite kurtosis, VaR can be estimated using a generalized EWMA volatility model in the recursive form of volatility estimation.

As for the limitations, the model proposal in this paper is only theoretical and does not include regression testing or practical application. Because VaR is an estimate, the model that estimates VaR must be tested, with the main idea being to count the total number of times lost in the entire sample exceeds the VaR value. In future studies, authors should link specific examples to more intuitively study the advantages and disadvantages of these VaR models.

References

[1] *Liu Dingci. The comparison of different approaches to estimate value at risk and optimization [D]. Jilin university, 2020. The DOI: 10.27162 /, dc nki. Gjlin. 2020.005143.*
[2] *Tang Ning. Statistical research and Empirical analysis based on Value at Risk (VaR) model and backtest test [D]. China West Normal University,2016.*
[3] *Jing Yongqiang. A study on the Value at risk model of portfolio investment [J]. North University of China,2017.*
[4] *He Ying. Research on Financial risk Assessment of Small and medium-sized listed Companies based on VAR [D]. Shenyang Institute of Technology,2016.*
[5] *Thavaneswaran Aerambamoorthy,Paseka Alex,Frank Julieta. Generalized value at risk forecasting[J]. Communications in Statistics - Theory and Methods,2020,49(20).*
[6] *Wang Tianyi,Liang Fang,Huang Zhuo,Yan Hong. Do realized higher moments have information content? - VaR forecasting based on the realized GARCH-RSRK model[J]. Economic Modelling,2022,109.*
[7] *Badaye Hemant Kumar,Narsoo Jason. Forecasting multivariate VaR and ES using MC-GARCH-Copula model[J]. The Journal of Risk Finance,2020,ahead-of-print(ahead-of-print).*
[8] *Tian Yahao. Using Convolutional neural network to determine ARCH time series model [D]. Xiamen university, 2020. DOI: 10.27424 /, dc nki. Gxmdu. 2020.001647.*
[9] *YU Huiqin. Comparative study on industry risk of listed companies based on VaR model of extreme value Theory [D]. Zhejiang University,2013.*
[10] *Wang Liming, Wang Lian, et al. Applied Time Series Analysis [M]. Shanghai: Fudan University Press,2012.*