

# A rolling analysis on the prediction of value at risk with multivariate GARCH and copula

Yang Bai<sup>a</sup>, Yibo Dang<sup>a</sup>, Cheolwoo Park<sup>a</sup>, Taewook Lee<sup>1,b</sup>

<sup>a</sup>Department of Statistics, University of Georgia, USA;

<sup>b</sup>Department of Statistics, Hankuk University of Foreign Studies, Korea

---

## Abstract

Risk management has been a crucial part of the daily operations of the financial industry over the past two decades. Value at Risk (VaR), a quantitative measure introduced by JP Morgan in 1995, is the most popular and simplest quantitative measure of risk. VaR has been widely applied to the risk evaluation over all types of financial activities, including portfolio management and asset allocation. This paper uses the implementations of multivariate GARCH models and copula methods to illustrate the performance of a one-day-ahead VaR prediction modeling process for high-dimensional portfolios. Many factors, such as the interaction among included assets, are included in the modeling process. Additionally, empirical data analyses and backtesting results are demonstrated through a rolling analysis, which help capture the instability of parameter estimates. We find that our way of modeling is relatively robust and flexible.

**Keywords:** backtesting, copula, dependency, multivariate GARCH, GOGARCH, multidimensional VaR, portfolio risk management, quantitative finance

---

## 1. Introduction

Risk management is an old subject for banking industry, regulators and academia. After the “Black Monday”, financial professionals with quantitative background worried about the firm-wide risk management. Thus, big banks started evaluating risks with quantitative tools. In the early 90s, most banks’ methods were already close to the formal concept of Value at Risk (VaR). In 1994, JP Morgan published its extensively developed quantitative methodology on evaluating the VaR and made a free access, the *RiskMetrics*, to the public. Thereafter, US Securities and Exchange Commission (SEC) required all big banks on Wall Street to disclose their quantitative risk evaluation of trading activities associated with derivatives. VaR evaluation became a more formal and common requirement after the release of Basel Accord II, which included detailed requirements on both VaR estimation and backtesting. The release of Basel III, together with the release of regional regulation frameworks such as Comprehensive Capital Analysis and Review (CCAR, the regulation framework released and used in USA), has made VaR a popular risk measure in the practice of risk evaluation.

This paper presents a hybrid method on the one-day-ahead prediction of high-dimensional portfolios VaR. In time series analysis, a popular way to validate the predictive power of a methodology is to use historical data sets sequentially with moving windows of a certain size. An analysis conducted

---

Both authors contributed equally to this work.

<sup>1</sup> Corresponding author: Department of Statistics, Hankuk University of Foreign Studies, 81, Oedae-ro, Mohyeon-eup, Cheoin-gu, Yongin-si, Gyeonggi-do 17035, Korea. E-mail: [twlee@hufs.ac.kr](mailto:twlee@hufs.ac.kr)

Published 30 November 2018 / journal homepage: <http://csam.or.kr>

© 2018 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.

with moving window is often referred as a rolling analysis. This paper also presents real data rolling analyses with sample portfolios and the corresponding backtesting results, utilizing publicly available historical data sets and different moving windows.

**Definition 1.** *To make it clear, in this paper, the VaR is defined and calculated with the general mathematical form:*

$$\text{VaR}_\alpha(L) = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\},$$

where  $F$  is the cumulative probability distribution of the loss function, and the confidence level of the VaR is  $1 - \alpha$ . VaR is therefore defined as the potential loss in the worst case at  $1 - \alpha$  confidence level. This paper demonstrates a modeling methodology at the 95% confidence level.

Generally, parametric, nonparametric, and hybrid are three specific approaches to define and calculate VaR. Each approach of VaR calculation requires the forward prediction of the variance, because the variance is the source of fluctuation. To estimate the variance in a sound accuracy level or to even actually estimate it at all, it is inevitable to take the time horizon into account. For instance, with a parametric distribution assumption, a simple percentage based VaR can be calculated with the parametric approach:

$$\text{VaR}_\alpha(L) = z_\alpha \times \sigma,$$

where  $\sigma$  is the standard deviation in percentage from a provided time horizon, and  $z_\alpha$  denotes the critical value with respect to a standard normal distribution at level  $\alpha$ . An applicable VaR prediction has to consider the mean effect of the target portfolio, or assets. The consideration on this mean effect also requires the input of a time horizon. Therefore, the time horizon is a crucial part of the VaR prediction, and the moving window approach is applied to establish the historical time horizon for the demonstrated prediction methodology.

The rest of the paper is organized as follows. Section 2 introduces the background knowledge of the methodology. Section 3 illustrates simple multi-dimensional examples. Section 4 demonstrates 3-, 5-, and 9-dimensional cases. Section 5 concludes the study and gives remarks.

## 2. Introduction to the modeling methodology

### 2.1. ARMA and multivariate GARCH models

The Autoregressive-Moving-Average (ARMA) model is a combination of the Autoregressive model and the Moving-Average model. The general form of the ARMA( $p, q$ ) model follows:

$$X_t = \theta + \epsilon_t + \sum_{i=1}^p \phi_i X_{t-i} + \sum_{j=1}^q \psi_j \epsilon_{t-j},$$

where  $X_t$ , a current value from a time series at time point  $t$ , can be expressed as the result of some general effect  $\theta$ , the historical values of the series  $X_1, \dots, X_{t-1}$  and the error terms (also referred to innovation terms or residual terms)  $\epsilon_1, \dots, \epsilon_t$ .

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model developed by Bollerslev (1986), is the extension of the Autoregressive Conditional Heteroskedasticity (ARCH) model (Engle, 1982). The multivariate version of GARCH has many different specifications mainly to resolve the dependency problem. In this paper, Constant Conditional Correlation (CCC) specification (Bollerslev, 1990), its enhanced version, Dynamic Conditional Correlation (DCC) specification

(Engle, 2002), and Generalized Orthogonal GARCH (GO-GARCH) model (Van der Weide, 2002) are used in modeling procedures.

The standard univariate GARCH ( $p, q$ ) model has the form as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_q \epsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2,$$

where  $\alpha$ 's and  $\beta$ 's are coefficients, and  $\epsilon$ 's are the error terms from the ARCH model. GARCH models explicitly account for the effects from error patterns and the historical variances. The formula of univariate GARCH (1, 1) is given as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

The multivariate version of GARCH pushes the edge even further to account for the dependency relationship, or covariance, of the multiple volatility that is included in the multivariate time series model. Assume that we have a time-varying  $N$  by 1 vector  $x_t$  and that we also have a mean vector  $\mu_t$ . With given  $\mathcal{F}_{t-1}$  which includes all available information on the time horizon up to the time point  $t - 1$ , we can model the vector  $x_t$ :

$$x_t | \mathcal{F}_{t-1} = \mu_t + \epsilon_t.$$

Note that

$$\epsilon_t = H_t^{\frac{1}{2}} \Omega_t,$$

where  $H_t$  is the conditional covariance matrix of the vector  $x_t$  and  $\Omega_t$  is an i.i.d. random vector with  $E(\Omega_t) = 0$ , and  $\text{Var}(\Omega_t) = I_N$ . The conditional variance matrix  $H_t$  can also be defined as (Bauwens *et al.*, 2006):

$$\text{Var}(x_t | \mathcal{F}_{t-1}) = \text{Var}_{t-1}(x_t) = \text{Var}_{t-1}(\epsilon_t) = H_t^{\frac{1}{2}} \text{Var}_{t-1}(\Omega_t) (H_t^{\frac{1}{2}})^{\top}.$$

For the bivariate case, the correlation between variables, or two time series, is presumed to be static over the time. The stochastic relationship does not change between two variables over time. Thus, a simple CCC-GARCH(1, 1) is expressed as:

$$\begin{aligned} h_{ii,t} &= \omega_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i h_{ii,t-1}, \\ h_{ij,t} &= \rho_{ij} \sqrt{h_{ii,t} h_{jj,t}}, \end{aligned}$$

where  $h_{ii,t}$  indicates the variance of the  $i^{\text{th}}$  entry at time  $t$ ,  $h_{ij,t}$  indicates the covariance between  $i^{\text{th}}$  entry and  $j^{\text{th}}$  entry at time  $t$ ,  $\rho_{ij}$  is the static correlation coefficient for  $i^{\text{th}}$  entry and  $j^{\text{th}}$  entry,  $\omega$ 's,  $\alpha$ 's,  $\beta$ 's are coefficients, and  $\epsilon$ 's stand for error terms.

According to CCC, the covariance entirely depends on the prediction of the variance terms, once a fixed correlation estimation is provided with a certain historical time horizon. Therefore, CCC accounts for only a partial effect of the interaction between variables in the system. But a major advantage of this specification is the unrestricted applicability for large systems of time series (Franke *et al.*, 2008).

As a successor of CCC, DCC specification is proposed, which bridges the gap of time varying interaction between variables (Engle, 2002). The idea is similar to the CCC specification. The DCC

specification estimates the covariance and variance separately, but allows the existence of the time-varying correlations instead of assuming that the correlation remains the same. The correlation matrix  $R_t$  can be expressed as:

$$R_t = \text{diag}(Q_t)^{-\frac{1}{2}} Q_t \text{diag}(Q_t)^{-\frac{1}{2}}$$

with

$$Q_t = \Omega + \gamma v_{t-1} v_{t-1}^\top + \delta Q_{t-1},$$

where  $v_t = (\epsilon_{1,t}/\sqrt{h_{11,t}}, \dots, \epsilon_{d,t}/\sqrt{h_{dd,t}})^\top$ ,  $\Omega$  is a symmetric positive definite parameter matrix,  $\gamma$  and  $\delta$  are scalar parameters, and  $Q_t$  is positive definite. Once the changing correlation matrix is found at a certain time point of the multivariate time series, the corresponding variance-covariance matrix at the time point can be defined by the same logic of CCC specification.

The generalized orthogonal GARCH or GO-GARCH (Van der Weide, 2002) is a special case of the early BEKK specification (named after Baba, Engle, Kraft, and Kroner), which is one of several ground-breaking specifications of multivariate GARCH models. With the assumption that the errors in the standard GARCH model are dependent on some uncorrelated component and can be modeled by that component through a linear mapping. The logic of GO-GARCH can be expressed as:

$$\epsilon_t = Z y_t,$$

where  $Z$  is a linear mapping factor, and  $y_t$  is a unobservable factor. Notice that the  $Z$  is assumed to be invertible. If  $y_t$  follows a GARCH(1, 1), then

$$H_t = Z \Omega Z^\top + \sum_{i=1}^d \alpha_i \lambda_i \omega_i^\top H_{t-1} \omega_i \lambda_i^\top + \sum_{i=1}^d \beta_i \lambda_i \omega_i^\top \mu_{t-1} \mu_{t-1}^\top \omega_i \lambda_i^\top,$$

where  $\Omega = \sum_{i=1}^d (1 - \alpha_i - \beta_i) e_i e_i^\top$ ,  $\lambda_i = Z e_i$ ,  $\omega_i = (Z^{-1})^\top e_i$ . Here,  $e_i$  is a vector that has only the  $i^{\text{th}}$  entry equal to 1, and the rest are all zero.

With the above assumption and definition,  $Z$  can have the following decomposition:

$$Z = H_t^{\frac{1}{2}} R,$$

where  $R$  is a rotation matrix. The unconditional covariance matrix is then

$$Z Z^\top = H_t^{\frac{1}{2}} R R^\top H_t^{\frac{1}{2}} = H_t.$$

This specification balances the computational cost and the complexity of covariances over time (Broda and Paoletta, 2009).

To help understand the modeling process with the combination of the GARCH and ARMA models, let us take DCC-GARCH(1, 1)-ARMA(1, 1) as an example. With DCC-GARCH(1, 1), the variance and covariance of the multivariate time series can be modeled as:

$$\begin{aligned} h_{ii,t} &= \omega_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i h_{ii,t-1}, \\ h_{ij,t} &= \rho_{ij,t} \sqrt{h_{ii,t} h_{jj,t}}, \end{aligned}$$

where  $h_{ii,t}$  is the variance of  $i^{\text{th}}$  time series at time  $t$ ,  $h_{ij,t}$  is the covariance of the  $i^{\text{th}}$  and  $j^{\text{th}}$  time series at time  $t$ , and  $\rho_{ij,t}$  is the changing correlation for the covariance. At the mean time, the multivariate mean of the time series can be modeled with the ARMA(1, 1) model. More details are included in Section 3.

## 2.2. Introduction to copulas

Modeling returns' dependency of different assets is important, especially for portfolios related to a volatile capital market such as future market and derivatives market. The conventional techniques for modeling the dependence among different variables are multivariate distributions, usually the multivariate normal distribution. However, when modeling the dependency among returns from different assets in a portfolio, the dependency is restricted by a specific theoretical distribution assumption, and the multivariate distribution does not provide a flexible fitting. Furthermore, for higher dimensional portfolios, the joint distribution is difficult to estimate. In addition, if the multivariate normal distribution is used to fit the returns, the tail part is automatically assumed to be not fat and the returns are assumed to be symmetrically distributed. Unfortunately, these assumptions are not verified by the market. A popular alternative for dependency quantification is to utilize copulas, which provides different structures to simplify the estimation of a multivariate joint distribution and also allows different marginal distributions to appear in the structures. For random variables  $x_i$ 's, an informal explanation of copula technique can be symbolized as:

$$F(x_1, \dots, x_n) = C_\theta(F(x_1), \dots, F(x_n)),$$

where  $C_\theta(F(x_1), \dots, F(x_n))$  is a copula expression of the multivariate joint distribution  $F(x_1, \dots, x_n)$ ,  $\theta$  is a parameter of the copula structure, and  $F(x_i)$ 's are marginal distributions. For time-varying scenarios, copula allow models to capture the information of time-varying dependency.

Copula can be defined in different ways for different situations. Commonly used copulas include Gaussian copula,  $t$  copula, Gumbel copula, Clayton copula, Ali-Mikhail-Haq copula, and those extensions of Archimedean copulas in hierarchical structures and nested structures. For completeness, definitions of some popular copulas are listed below. The general definition of copula is:

**Definition 2.** Define  $C(\cdot)$  as a function mapping from  $[0, 1]^d$  to  $[0, 1]$  with the following properties:

1.  $C(\mu_1, \dots, \mu_d)$  is an increasing function, if all  $\mu_i$ 's  $\in [0, 1]$ .
2.  $C(1, \dots, 1, \mu_i, 1, \dots, 1) = \mu_i$  for  $\forall \mu_i \in [0, 1]$ .
3. For all  $(\mu_1, \dots, \mu_d), (\mu'_1, \dots, \mu'_d) \in [0, 1]^d$  with  $\mu_i < \mu'_i$ .

Then,  $\sum_{i=1}^2 \dots \sum_{i_d=1}^2 (-1)^{(i_1 + \dots + i_d)} C(v_{j_1}, \dots, v_{j_d}) > 0$ , where  $v_{j_1} = \mu_j$  and  $v_{j_2} = \mu'_j$  for all  $j = 1, \dots, d$ .

The function  $C(\cdot)$  is called a  $d$  dimensional copula. This setting makes the copula  $C(\cdot)$  equivalent to a  $d$  dimensional uniform distribution. For a multivariate distribution with all marginal distributions being continuous, there exists a unique copula that explains the joint distribution together with those marginal distributions (Sklar, 1959). Using this uniqueness, many inverse operation methods of copula can be developed.

**Definition 3.** Assume that a random variate vector  $X \sim N_d(\mu, \Sigma)$  with the entry  $x_i \sim N(\mu_i, \sigma_i^2)$  for  $i = 1, \dots, d$ . Then, a copula exists as:

$$F_X(x_1, \dots, x_d) = C(F(x_1), \dots, F(x_d)),$$

where  $F_X$  is the multivariate joint distribution for  $X$  and  $F(x_i)$  is the marginal distribution for the associated  $x_i$ . Let  $Y$  be a vector containing an one on one transformation of the standardized entries from  $X$  with  $y_i \sim N(0, 1)$ . Define the standardization process as a function  $S$ . Then,  $y_i = S(x_i) =$

$(x_i - \mu_i)/\sigma_i$  and  $Y = (y_1, \dots, y_d)^\top \sim N_d(0, \Psi)$ , where  $\Psi$  is the correlation matrix corresponding to covariance matrix  $\Sigma$ . A copula  $C_{Gauss}$  can be defined as:

$$F_Y(y_1, \dots, y_d) = C_{Gauss}(\Phi(y_1), \dots, \Phi(y_d)),$$

where  $\mu_i = \Phi(y_i)$ . A further expression can be derived:

$$\begin{aligned} C_{Gauss}(\mu_1, \dots, \mu_d) &= C_{Gauss}(\Phi(y_1), \dots, \Phi(y_d)) \\ &= F_Y(y_1, \dots, y_d) = F_Y(\Phi^{-1}(\mu_1), \dots, \Phi^{-1}(\mu_d)) \\ &= \int_{-\infty}^{\Phi^{-1}(\mu_1)} \cdots \int_{-\infty}^{\Phi^{-1}(\mu_d)} (2\pi)^{-\frac{d}{2}} |\Psi|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} r^\top \Psi^{-1} r\right) dr_1 \cdots dr_d, \end{aligned}$$

where  $r = (r_1, \dots, r_d)$ . Since the given transformation function  $S$  is increasing, it follows that  $C = C_{Gauss}$  (Franke et al., 2008), and Gaussian copula is therefore obtained. Gaussian copula explains the dependence structure of the multivariate normal distribution together with the corresponding normal marginal distributions. Gaussian copula is also part of the ‘‘Recipe of Disaster’’, which was often used in financial asset pricing prior to 2008, especially in the area of Collateralized Debt Obligation (CDO) pricing.

**Definition 4.** With  $X \sim t_d(v, \mu, \Sigma)$ , the  $t$  copula is given by:

$$C_v(\mu_1, \dots, \mu_d) = |\Psi|^{-\frac{1}{2}} \frac{\Gamma\left(\frac{v+d}{2}\right) \left(\Gamma\left(\frac{v}{2}\right)\right)^{d-1} \left(1 + \frac{1}{v} \zeta^\top \Psi^{-1} \zeta\right)^{-\frac{v+d}{2}}}{\left(\Gamma\left(\frac{v+1}{2}\right)\right)^d \prod_{j=1}^d \left(1 + \frac{1}{v} \zeta_j^2\right)^{-\frac{v+1}{2}}},$$

where  $\Psi$  is the correlation matrix,  $v$  is the degree of freedom, and  $\zeta_j = t_v^{-1}(\mu_j)$  (Franke et al., 2008). The derivation of  $t$  copula is similar to Gaussian copula. Note that Gaussian copula and  $t$  copula belong to the elliptical copula family.

**Definition 5.** Gumbel copula is specified as:

$$C_\theta(\mu_1, \dots, \mu_d) = -\exp\left\{\left[\sum_{j=1}^d (-\log \mu_j)^\theta\right]^{\frac{1}{\theta}}\right\}.$$

For  $x \in [0, \infty]$ , Gumbel copula generator is defined as:  $\phi(x, \theta) = \exp(-x^\theta)$ , where  $1 \leq \theta < \infty$ .

As opposed to the previous two copulas, Gumbel copula belongs to the Archimedean copula family. It is also a frequently implemented copula in the financial area. One of the major differences between the Archimedean copula and the elliptical copula is that Archimedean copula allows an asymmetric fit of the data with only one parameter. If we visualize a bivariate  $t$  copula, its shape looks close to the letter ‘‘X’’. If we visualize a bivariate Gaussian copula, the plot usually looks like a circle. While the bivariate visualization of a Gumbel copula often has an unbalanced concentration on the diagram. In fact, Archimedean copulas are mostly manually designed to explain the unbalanced dependence among variables. This differs from both  $t$  copula and Gaussian copula that are made out of known distributions and only explain balanced dependence.

**Definition 6.** Clayton copula is defined as

$$C_{\theta}(\mu_1, \dots, \mu_d) = \left[ \left( \sum_{j=1}^d \mu_j^{-\theta} \right) - d + 1 \right]^{\theta}.$$

For  $x \in [0, \infty]$ , the generator is defined as:  $\phi(x, \theta) = (\theta x + 1)^{-1/\theta}$ , where  $\theta \in [-1, \infty] - \{0\}$ . Compared to Gumbel copula, Clayton copula puts more weight on the lower tail rather than the upper tail.

**Definition 7.** Ali-Mikhail-Haq (AMH) copula is defined as:

$$C_{\theta}(\mu_1, \dots, \mu_d) = \frac{\prod_{j=1}^d \mu_j}{1 - \theta \left( \prod_{j=1}^d 1 - \mu_j \right)},$$

where  $\theta \in [-1, 1]$ . If  $\theta = 1$ , there exists significant lower tail dependence. Other Archimedean copulas, such as Joe (1997) copula, Frank (1979) are not reviewed and implemented in this work.

**Definition 8.** Hierarchical Archimedean Copula (HAC) defines a dependence structure with different Archimedean copulas in a nested form. Its structure resembles a hierarchical classification tree. In detail,

$$C(\mu_1, \dots, \mu_d) = \Phi_{d-1} \left[ \Phi_{d-1}^{-1} \circ C((\Phi_1, \dots, \Phi_{d-2}))(\mu_1, \dots, \mu_d) + \Phi_{d-1}^{-1}(\mu_d) \right],$$

where  $\Phi$ 's are copula generators at different levels. Different copulas can be used for different levels if needed and this is the advantage of the HAC. One simplification of HAC is the Nested Archimedean Copula (NAC), which applies the idea of HAC in a completely nesting way with only one type of Archimedean copula at all levels. In this paper, we use mainly NAC to demonstrate the modeling process.

### 3. Introduction to the modeling methodology and model validation

#### 3.1. Modeling methodology

Statistically,  $\text{VaR}_{1-\alpha}$ , or sometimes  $\text{VaR}_{\alpha}$ , which indicates the VaR at  $1 - \alpha$  confidence level, can be calculated using the lowest  $100 \times (1 - \alpha)$  percentile of the returns or log returns. For example, the simplest nonparametric method to calculate VaR of an asset is the historical simulation method. Basically, it takes the value at the lowest  $100 \times \alpha$  percentile of historical returns, or historical log returns, directly as  $\text{VaR}_{\alpha}$ . Moreover, the log returns  $X_t$ , which is the return time series, can be modeled with the formula:

$$X_t = \mu_t + \sigma_t \epsilon_t,$$

where  $\mu_t$  is the mean of log returns by the time point  $t - 1$  or the estimated log returns at time point  $t$ ,  $\sigma_t$  is the square root of the associated variances predicted from the GARCH model at time point  $t$ , and  $\epsilon_t$  is the error terms from the ARMA-GARCH model with distribution assumption for the included noises at time point  $t$ . Note that the distribution assumptions can be flexible, since we only care about the fit of the tails.

For example, if the multivariate GARCH(1, 1) and the ARMA(1, 1) are applied to model the returns, the variance-covariance matrix and the error vector on the right hand side of the above equation

can be produced from the multivariate GARCH(1, 1) model with distribution assumption for noises included in the error terms. At the mean time, ARMA(1, 1) can produce the modeling results for mean return vector on the right hand side of the equation.

Inspired by the properties of ARMA models, GARCH models and the concept of copulas, the error terms from ARMA-GARCH models for multivariate cases can be modeled and simulated with copulas and marginal distributions, and a sequential modeling process can be implemented to model and predict VaR:

Step 1: Model the log return time series with multivariate ARMA-GARCH.

Step 2: Fit marginal distributions and a copula using the errors from the multivariate ARMA-GARCH.

Step 3: Get the prediction of mean and variance from the multivariate ARMA-GARCH.

Step 4: Simulate the error terms from the results of Step 2.

Step 5: Acquire a set of predicted log return values with results from Step 3 and Step 4, and determine the value at  $100 \times \alpha$  percentile of the log return values, which is the VaR at the corresponding time point.

To avoid any confusion, the error terms from GARCH, or ARMA-GARCH refer to the error terms of the mean process. GARCH models inherit the error terms from the setup of the ARCH model, and the error terms applied to model the variance in the ARCH model ultimately refer back to the error terms in the mean process, which is the ARMA model in our modeling process.

### 3.2. Copula fit

The copula fit of the error terms' dependency structure is the key to the modeling process. We assume the error distribution to be multivariate normal; however, the choice of copula is flexible depending on the emphasis on the tail distribution of the shocks. The estimation of a copula based multivariate distribution involves both the estimation of the copula parameters  $\theta$  and the estimation of margins  $F(x_i)$ . Suppose  $\delta_i$  is the parameter of  $F(x_i)$ , the parameter vector is  $\Omega = (\delta_1, \delta_2, \dots, \delta_d, \theta)^T$ . Then the log-likelihood function is:

$$l(\Omega | x_1, x_2, \dots, x_T) = \sum_{t=1}^T \log(cF_1(x_{1,t}; \delta_1), \dots, F_d(x_{d,t}; \delta_d); \theta) + \sum_{t=1}^T \sum_{i=1}^d \log f_i(x_{i,t}; \delta_i),$$

where  $c(u_1, \dots, u_d) = \partial^d C_\theta(u_1, \dots, u_d) / (\partial u_1 \dots \partial u_d)$ , and  $x_t = (x_{1,t}, x_{2,t}, \dots, x_{d,t})$ . In order to obtain the MLE, we firstly need to estimate the parameter vector  $\Omega$ . By Full Maximum Likelihood (FML) Estimation, estimating  $\Omega$  is equal to solve the equation:

$$\frac{\partial l(\Omega|X)}{\partial \delta_i} = \frac{\partial l(\Omega|X)}{\partial \theta} = 0.$$

However, it costs too much time in practice if the dimension is not moderate. Another popular estimating method is called the Inference for Margins (IFM), unlike the former one step FML method, this method estimates the parameter from the marginal distributions  $\delta_i$  in the first step; second, it estimates the dependence parameter  $\theta$  through the pseudo log likelihood function. The main advantage of IFM method is that it highly reduces the computational and notational complexity (Franke *et al.*, 2008).



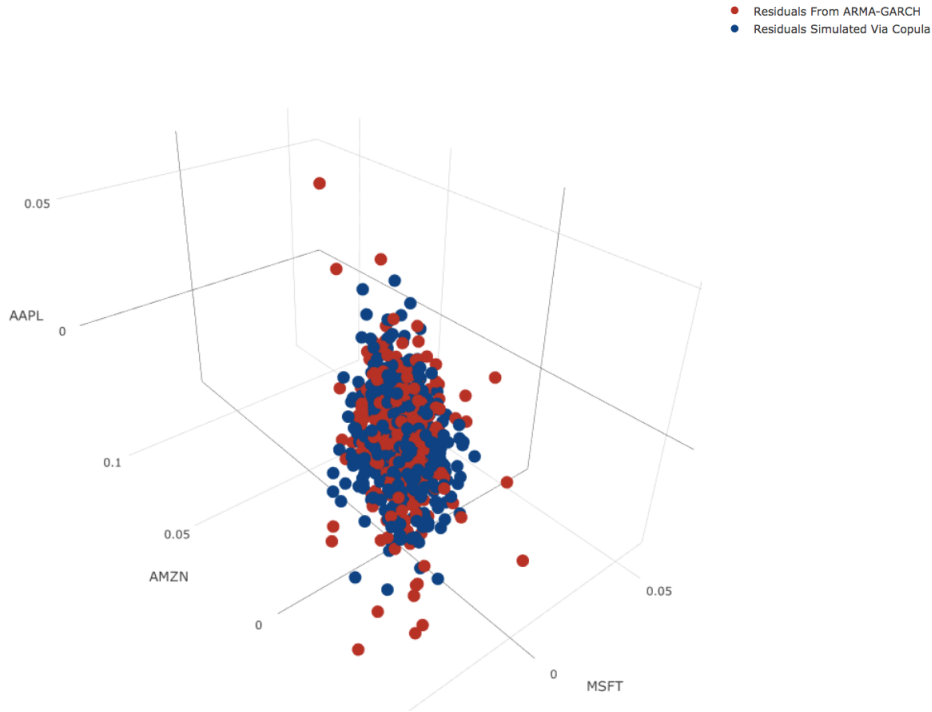


Figure 1: Copula fit with error terms of MSFT, AMZN, and AAPL from ARMA-GARCH from 2016/06/01 to 2017/12/01 with the simulation size equal to the time window.

This work is not limited with the gaussian copula or  $t$  copula. As a visual demonstration of the goodness of fit, we show an example of a 3-dimensional 250-day dependency structure with returns of MSFT, AMZN, and AAPL in Figure 1. Note that the simulated multivariate distribution of the error terms reasonably covers the entire area of the realized errors.

### 3.3. Model validation

To examine the adequacy of the VaR measures produced by our copula-based models, we turn to a conventional “backtesting” validation method, which is used through our entire modeling process. The backtesting procedure evaluates the quality of the forecast of a risk model by comparing the actual results with those generated by the VaR models. For example, consider an event that the loss of a portfolio exceeds its reported VaR,  $\text{VaR}(\alpha)$  at quantile  $\alpha$ . Denoting the profit or loss of the portfolio between time  $t$  to  $t + 1$  as  $r_{t,t+1}$ , the “hit” function  $I_{t+1}(\alpha)$  can be defined as (Christoffersen, 1998):

$$I_{t+1}(\alpha) = \begin{cases} 1, & \text{if } r_{t,t+1} \leq -\text{VaR}(\alpha), \\ 0, & \text{otherwise.} \end{cases}$$

Thereupon, the hit function sequence, e.g., (1, 0, 1, 0, 0, 0, 1), tells the times that the loss of the portfolio has exceeded the labeled  $\text{VaR}(\alpha)$  during the observed period.

Backtests can often be classified by if they examine an unconditional coverage property or independence property of a VaR measure. However, the test is called an Unconditional Coverage (UC) test

Table 1: Setup of example portfolios

Dimension (portfolio no.) <sup>a</sup>	Assets <sup>b</sup>	Time horizon (moving window)	Prediction coverage <sup>c</sup>
3 (A)	AMZN, F, RDS.A	1000 (250)	01/01/05–01/25/16
3 (B)	BA, SPY, GS		
3 (C)	DDD, SPY, GS		
3 (D)	BAC, DAL, DG		
5 (A)	APPL, NFLX, DDD, BA, TUC	1750 (500)	01/01/05–01/25/16
5 (B)	AMZN, F, GS, RDS.A, SPY		
5 (C)	DDD, TUC, F, BAC, MSFT		
5 (D)	WMT, C, WMAR, MS, ACET		
5 (E)	RDS.A, C, BAC, MS, ACUR		
5 (F)	APPL, DAL, ENZN, SPY, NFLX		
5 (G)	AMZN, F, GS, RDS.A, SPY		
9 (A)	APPL, WMT, NFLX, SPY, C, F, AMZN, MSFT, GS	2550 (250)	01/01/05–01/25/16
9 (B)	ACET, C, ACUR, BAC, MS, DDD, TUC, SH, SPY		
9 (C)	WSTL, C, ACUR, BA, MS, DDD, TUC, WMAR, ENZN		
9 (D)	AAPL, GOOG, NFLX, AMZN, F, MSFT, GS, RDS.A, SPY, DEXJUP		
			11/01/05–02/27/15
			12/07/05–02/27/15

a: The combination of dimension and portfolio number of associated dimension is used as the portfolio name in latter tables.

b: Assets included in this table are all from open markets such as NYSE and NASDAQ. Missing values are omitted. Portfolios contain only one share of each asset.

c: Date format is mm/dd/yy.

Table 2: Modeling and validation of 3-dimensional portfolios

Portfolio	GARCH	Marginal distributions	Copula (simulation size)	Unconditional coverage	Conditional coverage	Exceedance (95%)
3 (A)			Gumbel (250)	0.1386943	0.2460887	3.87% (29 out of 750)
			Clayton (250)	0.1027796	0.3124458	4.0% (30 out of 750)
			Gaussian (250)	0.4333781	0.2670521	5.06% (38 out of 750)
			<i>t</i> (250)	0.3452114	0.1640517	3.47% (26 out of 750)
			NAC (250)	0.6221392	0.2290134	5.06% (38 out of 750)
3 (B)			Gumbel (250)	0.0026605	0.0035204	2.8% (21 out of 750)
			Clayton (250)	0.0037568	0.0058276	7.47% (56 out of 750)
			Gaussian (250)	0.2673953	0.4725123	5.46% (41 out of 750)
			<i>t</i> (250)	0.0315854	0.0701514	6.80% (51 out of 750)
			NAC (250)	0.0213867	0.0691763	6.93% (52 out of 750)
3 (C)	DCC-GARCH (1, 1) - ARMA(1, 1)	Skewed- <i>t</i>	Gumbel (250)	0.0962584	0.2504802	3.73% (28 out of 750)
			Clayton (250)	NA	NA	NA
			Gaussian (250)	0.1386943	0.3316343	3.87% (29 out of 750)
			<i>t</i> (250)	0.1386943	0.3316343	3.87% (29 out of 750)
			NAC (250)	0.1033458	0.2134489	4.40% (33 out of 750)
3 (D)			Gumbel (250)	0.1935932	0.3367489	4.93% (37 out of 750)
			Clayton (250)	0.00011298	0.0000013	2.26% (17 out of 750)
			Gaussian (250)	0.4760852	0.4405354	5.60% (42 out of 750)
			<i>t</i> (250)	NA	NA	NA
			NAC (250)	0.5632115	0.5253472	5.47% (41 out of 750)

if we are only interested in the times of the reported VaRs being violated. According to Christoffersen and Pelletier (2004), the interim between two VaR violations should exhibit no duration dependency, which is specified as the independence property. Therefore, the Conditional Coverage (CC) test is a joint test that checks both the violated times and the independence property.

#### 4. Empirical examples with real data sets

To examine the stability of the modeling methodology and its accuracy, different example portfolios were established with assets from global financial markets including stocks, index futures, and commodities. A key assumption is that the parameter estimates should be constant over time when

Table 3: Modeling and validation of 5-dimensional portfolios

Portfolio	GARCH	Marginal distributions	Copula (simulation size)	Unconditional coverage $p$ -value	Conditional coverage $p$ -value	Exceedance (95%)
5 (A)			Gumbel (250)	0.3208206	0.5854642	4.40% (55 out of 1250)
			Clayton (250)	NA	NA	NA
			Gaussian (250)	0.7440254	0.8036172	4.80% (60 out of 1250)
			$t$ (250)	0.9481972	0.7883348	4.96% (62 out of 1250)
			NAC-Gumbel (250)	0.7893542	0.8231094	4.40% (55 out of 1250)
5 (B)			Gumbel (250)	0.4063832	0.6386581	5.52% (69 out of 1250)
			Clayton (250)	0.9481972	0.7883348	4.96% (62 out of 1250)
			Gaussian (250)	0.8450712	0.9801981	4.88% (61 out of 1250)
			$t$ (250)	0.4063832	0.6386581	5.52% (69 out of 1250)
			NAC-Gumbel (250)	0.8964532	0.9354267	4.88% (61 out of 1250)
5 (C)			Gumbel (250)	0.1152822	0.2105959	6.00% (75 out of 1250)
			Clayton (250)	0.8462265	0.7121321	5.12% (64 out of 1250)
			Gaussian (250)	0.1386943	0.3316343	5.60% (70 out of 1250)
			$t$ (250)	0.1386943	0.3316343	5.68% (71 out of 1250)
			NAC-Gumbel (250)	0.4418941	0.6722767	6.00% (75 out of 1250)
5 (D)	GO-GARCH (1, 1)	Skewed- $t$	Gumbel (250)	0.0523736	0.1324256	6.24% (78 out of 1250)
	AR (1)		Clayton (250)	NA	NA	NA
			Gaussian (250)	0.2592072	0.1463091	4.32% (54 out of 1250)
			$t$ (250)	0.0394249	0.1073051	6.32% (79 out of 1250)
			NAC-Gumbel (250)	0.9481972	0.8661561	4.96% (62 out of 1250)
5 (E)			Gumbel (250)	0.0523736	0.1520078	6.24% (78 out of 1250)
			Clayton (250)	0.3208206	0.0733361	4.40% (55 out of 1250)
			Gaussian (250)	0.32082062	0.5692731	4.40% (55 out of 1250)
			$t$ (250)	0.9483278	0.0510023	5.04% (63 out of 1250)
			NAC-Gumbel (250)	0.4690168	0.0421323	4.56% (57 out of 1250)
5 (F)			Gumbel (250)	0.0223748	0.0420299	8.00% (100 out of 1250)
			Clayton (250)	NA	NA	NA
			Gaussian (250)	0.0026155	0.0051958	6.96% (87 out of 1250)
			$t$ (250)	0.1152820	0.2235687	6.00% (75 out of 1250)
			NAC-Gumbel (250)	0.6524971	0.8679374	5.28% (66 out of 1250)
5 (G)		$t$	NAC-Clayton (500)	0.4727282	0.03804881	5.33% (120 out of 2250)

analyzing financial time series data using the Multivariate GARCH and Copula models. However, in a real financial market, considerable changes make it unreasonable to make this assumption. Therefore, we compute the parameter estimates over a rolling window of fixed sizes, e.g., 250, 500, through the samples. If the parameters are truly constant over the entire time horizon, then the estimates over the rolling windows should not be too different. If the parameters oscillate too much at some point during the period, then the rolling estimates should capture this volatility. Details of the example portfolios can be found in Table 1. We implemented mainly DCC-GARCH for 3-dimensional cases and GO-GARCH for 5-dimensional and 9-dimensional cases. Different copulas, simulation sizes and moving windows were tested. Details of the modeling specification and corresponding backtesting results can be found in Tables 2–4. Note that the key observation should be on the backtesting results, which include the  $p$ -values for both unconditional coverage and conditional coverage tests as well as the exceedance.

With our 3- and 5-dimensional experimental data sets, the modeling process provides a good and stable performance as indicated with two backtesting results, as well as the exceedance. In addition, it is important to pay attention to the balance of the results in the evaluation of such process. The exceedance should not be too far away from 95%, because capital reserve requirements for financial

Table 4: Modeling and validation of 9-dimensional portfolios

Portfolio	GARCH	Marginal distributions	Copula (simulation size)	Unconditional coverage $p$ -value	Conditional coverage $p$ -value	Exceedance (95%)
9 (A)			NAC-Gumbel (250)	0.2428564	0.4969938	4.48% (103 out of 2300)
			Clayton (250)	NA	NA	NA
			Gaussian (250)	0.0	0.0	1.65% (38 out of 2300)
			$t$ (250)	0.0018221	0.0031106	1.74% (40 out of 2300)
			NAC-Clayton (250)	0.0228109	0.0031106	1.74% (40 out of 2300)
9 (B)	GO-GARCH (1,1)	Skewed- $t$	NAC-Gumbel (250)	0.003	0.01	3.74% (86 out of 2300)
			Clayton (250)	NA	NA	NA
			Gaussian (250)	0.0	0.0	1.48% (34 out of 2300)
			$t$ (250)	0.0	0.0	1.52% (35 out of 2300)
			NAC-Clayton (250)	0.0073	0.025	3.48% (80 out of 2300)
9 (C)	AR (1)		NAC-Gumbel (250)	0.02	0.005	3.74% (86 out of 2300)
			Clayton (250)	NA	NA	NA
			Gaussian (250)	0.0	0.0	1.69% (39 out of 2300)
			$t$ (250)	0.0	0.0	1.83% (42 out of 2300)
			NAC-Clayton (250)	0.002	0.007	3.70% (85 out of 2300)
9 (D)		$t$	NAC-Clayton (500)	0.2428564	0.4969938	4.48% (103 out of 2300)
			Gumbel (500)	0.7731826	0.5368803	4.87% (112 out of 2300)

institutions will be evaluated according to the VaR models and VaR Stress Testing performance. If a model is claimed to be at a 95% confidence level but always delivers results close to 99%, the efficiency of capital allocation will be reduced due to the 4% difference. From this perspective, the 3- and 5-dimensional modeling examples also present a good performance. The exceedance of the modeling results are reasonably close to the 95% target level.

For 9-dimensional examples, especially those examples with skewed- $t$  and the provided NAC copula structure, the modeling process demonstrates a very conservative projection on VaR. The modeling process highly depends on copula structures; therefore, it is clear that copula's capability of explaining high-dimensional dependency decreases along with the increase of dimensionality. The modeling process tends to overestimate the risk in 9-dimensional cases. One potential way to improve this part is to compare the performance among different copulas in order to select the best copula for the corresponding dependency structure. Another simpler way would be to increase the simulation size.

## 5. Discussion on model selection

### 5.1. Marginal distributions in multi-dimensional copulas

Copula methods are introduced to decompose the continuous joint distribution into individual marginal distributions so that the dependence structure can be easily studied. This raises concerns on the selection of marginal distribution and fitting of the distributions using the copula structures. Commonly used dispersion structures include: autoregressive of order 1, exchangeable, Toeplitz, and unstructured. Some researchers use one of the correlation matrices (usually autoregressive of order one and simple exchangeable Archimedean (Yan, 2007)) as a direct proxy for multivariate dependency. However, it is widely acknowledged that volatility parameters and variables are not always normally distributed. Many financial related variables often have fat tails and exhibit "tail dependence". In our study, the skewed- $t$  distribution fits the real log-return density function well most of the time. However, the skewed- $t$  tends to underestimate peak areas.

According to the definition of Normal Mean-Variance Mixture Representation of Skewed- $t$  Dis-

tribution, if we fix the degree of freedom, it is possible to fit financial data well with elliptical distributions. In practice, the normal mean-variance mixture formatted skewed- $t$  distribution presents good flexibility to fit different types of portfolios. In addition, skewed- $t$  distribution has non-zero tail dependence since the covariance matrix  $\Sigma$  is positive semi-definite. The skewed  $t$  distribution also gives heavier tails at either upper or lower part of the margins. Therefore, we prefer skewed- $t$  distribution to skewed normal as a margin to fit copulas.

## 5.2. GO-GARCH Model

A DCC-GARCH model is designed to overcome the drawbacks on the covariance estimation that early specifications such as BEKK and CCC models present (Van der Weide, 2002); however, the DCC specification is still not perfect due to the limitation on the possible failure of finding an invertible matrix. DCC assumes the existence of at least one orthogonal matrix that links the observed variables linearly to a set of components similar to the concepts of latent variables. Since the corresponding latent components are assumed to be independent, it is now possible to get dynamically time-varying conditional covariance for each of the multivariate time series systems presenting heteroskedasticity. Nevertheless, it is not guaranteed that a linkage or an invertible matrix will always exist. For example, if the diagonal elements of  $H$ , the covariance matrix, are not distinct, the transformed orthogonal matrix will no longer coincide with the linkage  $Z$ . Orthogonal GARCH (O-GARCH) and Generalized Orthogonal GARCH (GO-GARCH) were proposed successively in-between 1994 to 2002. GO-GARCH allows  $Z$  to be given by any possible invertible matrix. Instead of using all of the historical information, the estimation of  $Z$  depends on the recent time point. Thus, GO-GARCH introduces a much easier and balanced computation method on dynamic conditional multivariate time series systems, especially high dimensional time series systems. For our exhibitions, we prefer GO-GARCH in high dimensional cases. It is important to mention that  $Z$  is chosen to be constant over time for the prediction of covariance. Since the direction of  $Z$  does not change during the process, this constant assumption holds.

## 6. Conclusion

Through the investigation on applicability of the proposed modeling methodology, we conclude that the proposed modeling process is a good alternative approach to study and predict VaR. With multivariate time series and copula, the proposed modeling methodology keeps a good balance between the complexity and flexibility. At the meantime, it helps reduce the computational cost. Furthermore, a methodology combining different techniques provides an opportunity to conduct the model selection over a system of models which can be easily constructed. The copula technique plays an important role in the description of the dependency relationship. More importantly, it enables us to directly model the VaR with a regard to the interactions among assets at a high dimensional level. During the study and experimental modeling investigation, we also noticed that the effect of the mean time series on VaR was weak when comparing to the variance effect. The fit of the marginal distributions is also not as important as the fit of GARCH and copula. Additionally, there was no universal solution. Therefore, when modeling with the methodology presented in this paper, we should use a greater amount of caution and patience to adjust different components such as moving window size and copula selection. With the careful adjustment, a model selection process based on backtesting results should be conducted. For future studies on the high-dimension dependency structures, a careful evaluation is highly recommended for the difference between the simulated error terms and the realized error terms. As the dimensionality increases, the difference for fitted multivariate dependency struc-

tures can be very vulnerable in terms of precision and accuracy, compared to the realized dependency structure.

### Acknowledgement

The fourth author's work was supported by Hankuk University of Foreign Studies Research Fund of 2018.

### References

- Bauwens L, Laurent S, and Rombouts JVK (2006). Multivariate GARCH models: a survey, *Journal of Applied Econometrics*, **21**, 79–109.
- Bollerslev T (1986). Generalized autoregressive conditional Heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.
- Bollerslev T (1990). Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model, *The Review of Economics and Statistics*, **72**, 498–505.
- Broda S and Paoletta M (2009). CHICAGO: a fast and accurate method for portfolio risk calculation, *Journal of Financial Econometrics*, **7**, 412–436.
- Christoffersen P (1998). Evaluating interval forecasts, *International Economic Review*, **39**, 841–862.
- Christoffersen P and Pelletier D (2004). Backtesting value-at-risk: a duration-based approach, *Journal of Empirical Finance*, **2**, 84–108.
- Engle RF (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1008.
- Engle RF (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroscedasticity models, *Journal of Business & Economic Statistics*, **20**, 339–350.
- Frank MJ (1979). On the simultaneous associativity of  $F(x, y)$  and  $x + y - F(x, y)$ , *Aequationes Mathematicae*, **19**, 194–226.
- Franke J, Hafner CM, and Härdle WK (2008). *Statistics of Financial Markets An Introduction*, Springer, Verlag Berlin Heidelberg.
- Joe H (1997). *Multivariate Models and Multivariate Dependence Concepts*, Chapman and Hall/CRC, London.
- Sklar A (1959). Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231.
- Van der Weide R (2002). GO-GARCH: a multivariate generalized orthogonal GARCH model, *Journal of Applied Econometrics*, **17**, 549–564.
- Yan J (2007). Enjoy the joy of copulas: with a package copula, *Journal of Statistical Software*, **21**, 1–21.

Received April 21, 2018; Revised October 10, 2018; Accepted November 9, 2018