Thai Journal of Mathematics (2014) 11–23
Special Issue on : Copula Mathematics and Econometrics
http://thaijmath.in.cmu.ac.th
Online ISSN 1686-0209

Portfolio optimization of stock returns in high-dimensions: A copula-based approach

K. Autchariyapanitkul[†] [1] **S. Chanaim**[‡], **S. Sriboonchitta** [¶]

[†],[¶] Faculty of Economics, Chiang Mai University
e-mail : kittawit_autchariya@cmu.ac.th
[‡]Department of Mathematics, Faculty of Science, Chiang Mai University

Abstract : We used the multivariate t copula, which can capture the tail dependence to modeling the dependence structure of the risk in portfolio analysis. Multivariate t copula based on GARCH model was used to explain portfolio risk structure for high-dimensional asset allocation issue. With this method we used the Monte Carlo simulation and the results of multivariate t copula to estimate the expected shortfall of the portfolio. Finally, we obtained the optimal weighted for conditional Value-at-Risk (CVaR) model with the assumption of multivariate distribution to illustrate the potential model risk among portfolios returns.

Keywords : GARCH; Multivariate t Copula; CVaR; Expected Shortfall.
2010 Mathematics Subject Classification : 62P20; 91B84 (2010 MSC)

1 Introduction

The goal of portfolio optimization is to find the portfolio with highest returns. In this case the selection of the optimal portfolio depends on the underlying assumption on behavior of the assets and the choice on a measure of risk. In Markowitz (1952), the dependence between financial returns is totally explained by the linear correlation coefficient and efficient portfolios are the conventional mean variance optimization model. Generally, correlation is used to explain dependence between random variables in the linear regression, but it may be inappropriate for the financial analysis (see, Ang & Bekaert [3], Das & Uppal [15], Patton [16]

[1]Corresponding author.

and Hu [14]). They found that the performance of a portfolio based on dependence structure is better than a portfolio based on normal distribution model. The dependency among key factors in portfolios have to be considered. An incorrect model for dependence can lead to the loss of portfolios and missspecification to evaluate the liability. Several studies indicated the superiority of copula to model dependence. The reason they ignore to use correlation approach because of its failure to capture the tail dependency (see, Artzner et al., [4] and Szego [12]) and extreme events. Copulas can be easily used to obtain multivariate distributions and offer much more flexibility than the conventional one (see, Embrechts et al. [19] and Lee and Long [20].

Harvey and Siddique [21], who considered multivariate GARCH model with skewness, same as in Sriboonchitta et al. [7] who applied the time-varying copula based on GARCH model to predict the agriculture price. More studies from Chang et al. [6], who constructed dynamic portfolio of crude oil, soybean and corn by GARCH and ARJI models to estimate value at risk (VaR). This model allows for time-varying conditional correlation, but they cannot exhibit asymmetries in asymptotic tail dependence. To fix this loop hole, we introduce an optional approach to modeling the dependence structure of multivariate data by using an appropriate Student's t based on copula theory.

In this paper we are using a multivariate t copula, which is applied to portfolio optimization in financial risk management. Multivariate t copula have been used extensively in the context of modeling multivariate financial return data, and have been shown the superior to the normal copula (see, Chan and Kroese [13]). Similarly, the works from Kole et al. [11] provided the test of fit to select the right copula for a portfolio consisting of stocks, bonds and real estate, the result clearly showed that Student's t copula passes the tests with success and dominated Gaussian and Gumbel copulas.

To determine the portfolio risk management, the conventional portfolio Value-at-Risk (VaR) model with the assumption of normal joint distribution, which is wildly used in empirical studies, shows considerable biased due to model specification error (see, Miller and Liu, [18]). VaR is has been criticized for not being diversified risk measure. From Pflug [17], CVaR has been proved to be a coherent risk measure. For more application about VaR and CVaR, we refer the reader to the studies from Rockafellar and Uryasev [9], Acerbi and Tasche [1, 2].

The approach for minimizing CVaR and optimization problems with CVaR constraints can be found in Sriboonchitta et al. [8], Rockafellar and Uryasev [10], Chekhlov et al. (2000), Pflug [17]. They found that optimization with CVaR is much more efficient in the empirical studies.

In this study, we are also considers whether a more accurate CVaR or expected shortfall estimation under the t copulas based joint distributions could be illustrated. t copulas allowed the researcher to construct flexible multivariate distributions showing various patterns of tail behavior, expanding the characters of tails independence to dependence. Thus, multivariate t copula may be considered for measuring the risk of portfolio investment.

This study focused on returns of securities in the Stock Exchange of Thailand

(SET). With this method, we measure the risk of a high-dimensional stock returns portfolio. Thus, the main contributes of this paper can be summarize in two folds. First, we emphasize that the multivariate t copula can illustrate the asymmetric dependence structure and evaluate the complex nonlinear relations among financial portfolio management. Second, we use stock returns in high-dimensions with the minimum lost to show the weight of assets in portfolios.

The remainder of this paper is organized as follows: Section 2 provides the theoretical background of GARCH model and multivariate t-copula, while Section 3 shows the empirical application to stock market. Section 4 reports the empirical results, and final Section gives conclusions.

2 Theoretical Background

2.1 GARCH

GARCH model was proposed by Bollerslev (1986), which can relaxed the assumption that volatility is constant overtime, because GARCH can capture the characteristics of financial time series data (heteroscedasticity and volatility). If the data indicate a skewness or heavy tail, we can choose an innovation that support these information. Thus, ARMA(p,q) and GARCH(k,l) are defined by

$$\begin{aligned}
r_t &= \mu + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \psi_i \varepsilon_{t-i} + \varepsilon_t, \\
\varepsilon_t &= \sigma_t \cdot \nu_t, \\
\sigma_t^2 &= \omega + \sum_{i=1}^{k} \alpha_i \varepsilon_{i-t}^2 + \sum_{i=1}^{l} \beta_i \sigma_{t-i}^2,
\end{aligned}$$

where $\sum_{i=1}^{n} \phi_i < 1,\ \omega > 0,\ \alpha_i, \omega_i \geq 0$ and $\sum_{i=1}^{k} \alpha_i + \sum_{i=1}^{l} \beta_i \leq 1$, ν_t is an standardized residual of a chosen innovation. In this case, we used t distribution because the data was considered as a heavy tail distribution, which well defined for the financial time series data.

2.2 Multivariate t Copula

In contrast to Gaussian copulas, copulas extracted from multivariate t-distribution (called t-copulas) exhibit tail dependence. Copula is a way to construct a joint distribution function. The joint distribution function can define by

$$H(x_1, x_2, \cdots, x_n) = C(u_1, u_2, \cdots, u_n), \tag{2.1}$$

where $u_i = F_{X_i}(x_i),\ i = 1, 2, \cdots, n$, where $F_{X_i}(\cdot)$ are distribution functions. By Sklar's Theorem, if $F_{X_i}(\cdot)$, for all $i = 1, \cdots, n$ are continuous, the n copula function $C(u_1, u_2, \cdots, u_n)$ is unique. The high-dimensional copula is a high-dimensional distribution function with uniform marginals on space $[0, 1]^n$. The multivariate t distribution with degrees of freedom $\nu = n-1$, μ is the mean vector and Σ as a positive definite dispersion matrix, t distributed as $t \sim t_n(\nu, \mu, \Sigma)$, has density written as

$$f(x) = \frac{\Gamma[(\nu+n)/2]\left(1 + \frac{1}{\nu}(x-\mu)\sum^{-1}(x-\mu)^T\right)^{-(\nu+n)/2}}{\Gamma(\nu/2)\sqrt{(\nu\pi)^n|\sum|}}, \tag{2.2}$$

and the correlation matrix define by

$$\Sigma = \begin{pmatrix} 1 & \rho_{u_1,u_2} & \cdots & \rho_{u_1,u_n} \\ \rho_{u_2,u_1} & 1 & \cdots & \rho_{u_2,u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{u_n,1} & a_{u_n,2} & \cdots & 1 \end{pmatrix} \text{ where } \rho_{ij} \in [-1,1] \text{ and } i,j = \{1,\cdots,n\}, \tag{2.3}$$

where $\Gamma : \alpha > 0 \to \Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$. In the same line as Gaussian random vectors, general random vectors whose multivariate distributions are t-distributions have the stochastic representation as

$$X = \mu + \frac{\sqrt{\nu}}{S}Z, \tag{2.4}$$

where S distributed as $\chi^2(\nu)$, $Z \sim N(0,\Sigma)$ and S and Z are independent. Thus, the t-copulas with distribution function defined as

$$C^n_{\nu,\Sigma} = \int_{-\infty}^{T_\nu^{-1}(u_1)} \cdots \int_{-\infty}^{T_\nu^{-1}(u_n)} f(x)dx, \tag{2.5}$$

where T_v^{-1} is the quantile function of univariate distribution $T_1(\nu,0,1)$. Thus, the high-dimensional copula density is (see, Demarta and McNeil [22]).

$$c^n_{\nu,\Sigma}(t_\nu(x_1),\cdots,t_\nu(x_n)) = |\Sigma|^{-1/2}\frac{\Gamma(\dfrac{\nu+2}{2})}{\Gamma(\dfrac{\nu}{2})}\left[\frac{\Gamma(\dfrac{\nu}{2})}{\Gamma(\dfrac{\nu+1}{2})}\right]^n \frac{\left(1+\frac{\zeta'\Sigma^{-1}\zeta}{\nu}\right)^{-\frac{\nu+n}{2}}}{\prod_{i=1}^n\left(1+\frac{\zeta_i^2}{2}\right)^{-\frac{\nu+1}{2}}}, \tag{2.6}$$

where $\zeta = (T_\nu^{-1}(u_1),\cdots,T_\nu^{-1}(u_n))$ is the t-student univariate vector inverse distribution functions.

3 Simulated data for risk management

3.1 Equally weighted portfolio for Var and CVar

Using the t copulas, we can simulate returns for time series data in high dimensions for the next period to describe the correlation structure. Suppose, we would like to calculate the empirical VaR and CVaR of and equally weighted portfolio with n assets. Then, the equations given by

$$\begin{aligned} \text{Min } ES &= E[r|r \leq r_\alpha], && (3.1a)\\ \text{subject to } \; r_i &= w[r_{(1,t+1)} + r_{(2,t+1)} + \cdots + r_{n,t+1}], && (3.1b)\\ w_1 = w_2 = \cdots &= w_n = \frac{1}{n}, && (3.1c)\\ 0 \leq w_i \leq 1,\; i &= 1, 2, \cdots, n, \end{aligned}$$

where r_α is the lower $\alpha - quantile$, and $r_{i,t+1}$ is the return on individual asset at time $t+1$.

3.2 Optimal portfolio with minimum risk via t copula

To make multivariate t copula useful, we use the Monte Carlo simulation to estimate the expected shortfall of an optimal weighted portfolio. After that, the optimal portfolio weights of the selected assets are constructed under minimize expected shortfalls with respect to maximize returns. The method for calculating the expected shortfall can be summarized into four steps. First, we use t copula to simulate events which length is sample size N. Second, we plug the random number into inverse functions of the probability distributions often random variables, such as the skewed generalized error distribution in this study, and employ the mean and variance equations of the ARMA-GARCH model to get the N values of each variable at period $t+1$. Third, at the beginning we set the weights to each variable equally. Finally, the investor need to minimize her portfolio (P) with respect to her expected returns given by:

$$\text{Min } ES = E[r|r \leq r_\alpha], \tag{3.2}$$

subject to

$$\begin{aligned} r_i = w_1 r_{(1,t+1)} + w_2 r_{(2,t+1)} + \cdots + w_n r_{n,t+1}, && (3.3a)\\ w_1 + w_2 + \cdots + w_n = 1, && (3.3b)\\ 0 \leq w_i \leq 1,\; where\; i = 1, 2, \cdots, n, \end{aligned}$$

where r_α is the lower $\alpha - quantile$, and $r_{i,t+1}$ is the return on individual asset at time $t+1$.

4 Application to the stock market

4.1 Data and Statistical test

In this paper, we used the stock returns in SET50 index. We applied this method to several companies which have big market capitalization, high volatility and high market value. There are Airports Of Thailand Public Company Limited (AOT), Bankok Bank Public Company Limited (BBL), The Siam Commercial Bank Public Company Limited (SCB), and The Siam Cement Public Company Limited (SCC). All the weekly data are extracted from Datastream from March 2009 until Jan 2014 with a total of 260 observations for each selected companies. Thus, innovation for GARCH model was used by t distribution. we checked all growth rate values are stationary by using Augmented Dickey-Fuller (ADF) and Phillip-Perron (PP) tests shown in table 1.

Table 1: Summary statistics

	AOT	BBL	SCB	SCC
Mean	0.0097	0.0033	0.0039	0.0055
Median	0.0048	0.0000	0.0000	0.0045
Maximum	0.2863	0.1476	0.1457	0.1650
Minimum	-0.1235	-0.1039	-0.1322	-0.1304
SD.	0.0511	0.0374	0.0417	0.0401
Skewness	0.9340	0.3276	0.2089	0.2502
Kurtosis	6.2145	3.5271	3.9875	4.4801
ADF-test	-15.7446	-16.7535	-17.6415	-16.9495
PP-test	-15.7446	-16.7535	-17.6415	-16.9419
JB	149.7406	7.66217	12.4560	26.4453
Obs.	260	260	260	260

All values are the log return.

4.2 ARMA-GARCH process

For each data series, we use the ARMA-GARCH process to estimate the marginals and we have shown that all the marginals are follow t distributions. We select the optimal lag for ARMA(p,q) by using Akaike information criterion (AIC) and found that the returns on AOT, BBL, SCB and SCC satisfied $ARMA(3,2)$, $ARMA(1,1)$, $ARMA(1,1)$, and $ARMA(5,4)$ with $GARCH(1,1)$ respectively. Table 2 gives the solutions of the estimated parameter

We used the Kolmogorov-Smirnov test (KS-test) to ensure the marginals are uniform distribution in $(0,1)$ and Box-Ljung test to confirm residuals are independent and identically distributed random variables (i.i.d). The results show that

Table 2: Estimates of ARMA-GARCH parameters for raw returns.

	AOT	BBL	SCB	SCC
C	0.0310	0.0004	0.0008	0.0022
	(0.0034)	(0.0006)	(0.0007)	(0.0017)
AR(1)	-1.8987	0.8326	0.7735	-0.0759
	(0.0678)	(0.2328)	(0.1793)	(0.0692)
AR(2)	-0.8755	-	-	1.1640
	(0.1322)	-	-	(0.0324)
AR(3)	0.0267	-	-	0.0795
	(0.06633)	-	-	(0.0856)
AR(4)	-	-	-	-0.9194
	-	-	-	(0.0342)
AR(5)	-	-	-	-0.1094
	-	-	-	(0.0645)
MA(1)	1.9942	-0.8744	-0.8453	0.0091
	(0.0187)	(0.2015)	(0.1495)	(0.0179)
MA(2)	1.0000	-	-	-1.1668
	(0.0190)	-	-	(0.0197)
MA(3)	-	-	-	0.0278
	-	-	-	(0.0172)
MA(4)	-	-	-	0.9784
	-	-	-	(0.0190)
K	0.0001	0.0012	0.0011	0.0001
	(0.0001)	(0.0008)	(0.0004)	(0.0001)
GARCH(1)	0.8633	0.0001	0.0395	0.7934
	(0.0947)	(0.6024)	(0.2061)	(0.1030)
ARCH(1)	0.0730	0.1569	0.3199	0.1426
	(0.0523)	(0.1234)	(0.1384)	(0.0708)
DoF	8.7515	12.857	10.2490	5.5719
	(3.4256)	(8.7887)	(6.8386)	(2.1589)
LogL	430.6423	490.3632	469.6514	492.3380

() standard error is in parenthesis, C and K are constant terms

Table 3: KS Test and p-value of Box-Ljung Test (Q-Test)

	AOT	BBL	SCB	SCC
KS	0.3956	0.4439	0.5724	0.0537
Q(5)	0.9292	0.9988	0.9956	0.7984
Q(10)	0.8179	0.9991	0.9939	0.8167
Q(15)	0.8562	0.9923	0.9796	0.9165
Q(20)	0.8366	0.9841	0.9822	0.9381

at given lag with significant level 0.05, these stocks satisfied all the requirements. Table 3 exhibits the results of the test

We can compare the residuals and the corresponding conditional standard deviations of four stocks extracted from their raw returns. The Fig. 1 clearly illustrates heteroskedasticity present in the filtered residuals. Having the model residuals from each return series, standardize the residuals by the corresponding conditional standard deviation. The returns reveals that the standardized residuals are now approximately i.i.d.

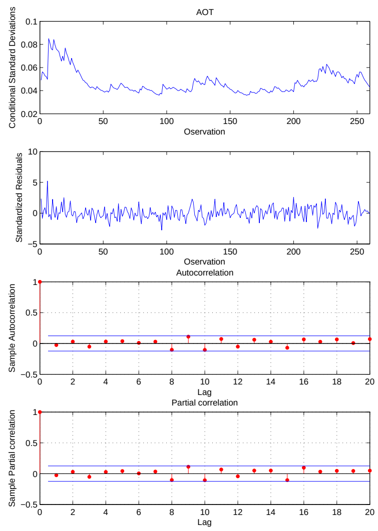
4.3 *t* copulas parameter estimation

Table 4 shows that the solutions of multivariate t copula parameters. We can use these values to construct efficient portfolio and find optimal plans for best expected returns with minimum loss.

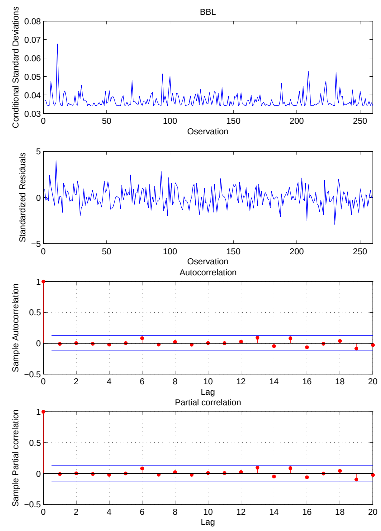
Table 4: Empirical t copulas parameters ($\hat{\rho}$)

	AOT	BBL	SCB	SCC
AOT	1.0000	0.4608	0.5252	0.4700
BBL	0.4608	1.0000	0.8053	0.6330
SCB	0.5252	0.8053	1.0000	0.6398
SCC	0.4700	0.6330	0.6398	1.0000

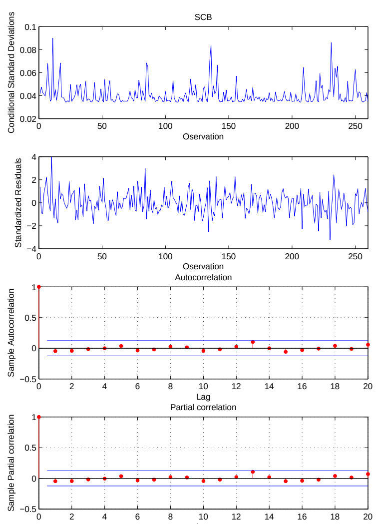
$\hat{\nu} = 9.4558$



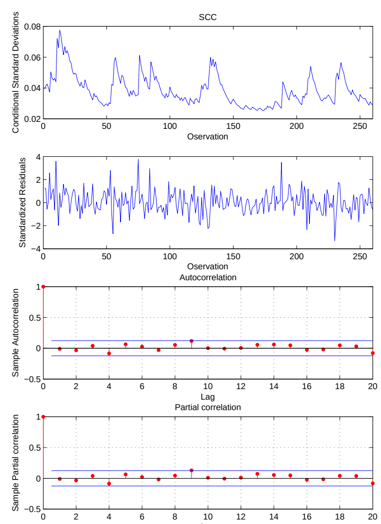
(a) AOT



(b) BBL



(c) SCB



(d) SCC

Figure 1: : Variation in volatility and auto-correlation plots

4.4 Experimental results

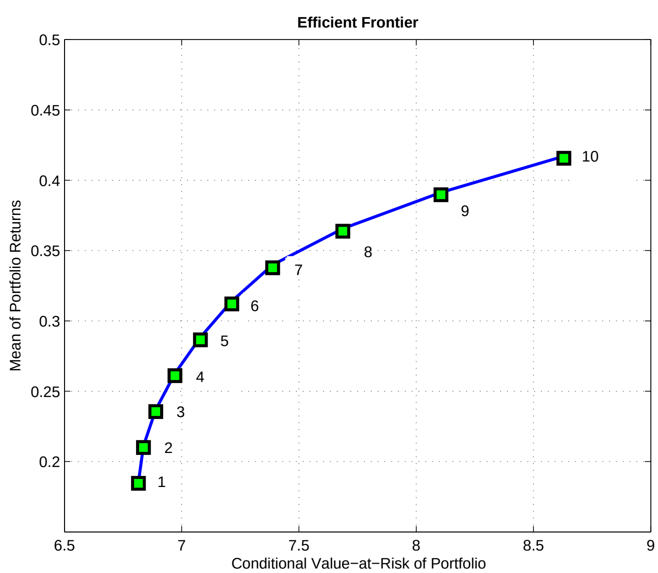
Table 4 exhibits the expected returns, VaR and CVaR at levels of 1%, 5% and 10% with equally weighted. We notice that the estimated CVaR converges to -0.0573, -0.0704 and -0.1019 at 10%, 5% and 1% levels in period $t+1$, respectively.

Table 5: Expected shortfall of equally weighted portfolios

	Expected Returns	VaR	CVaR
10%	0.0024	-0.0383	-0.0573
5 %	0.0024	-0.0515	-0.0704
1 %	0.0024	-0.0810	-0.1019

We used the Monte Carlo simulation to generate a set of 1,000,000 samples described in section 3. Then, given significant level of 5%, we optimize the portfolio by using the mean-CVaR model and obtained the efficient frontier of the portfolio under various expected returns, as shown in Fig. 2.

Finally, we also obtained the optimal weight of the portfolios varies to the ES. Table 6 shown some of the results of optimal weight with the expected returns in the frontier.



(a) ES5

Figure 2: : The efficient frontiers of CVaR under mean

Table 6: Optimal weighted portfolios for ES 5 %

Portfolios	w_1	w_2	w_3	w_4	Returns
1	0.1685	0.4541	0.0000	0.3773	0.0018
2	0.1754	0.4923	0.0320	0.3004	0.0021
3	0.1767	0.5010	0.0885	0.2338	0.0024
4	0.1795	0.5075	0.1458	0.1672	0.0026
5	0.1815	0.5163	0.2019	0.1003	0.0029
6	0.1839	0.5249	0.2579	0.0333	0.0031
7	0.1531	0.4616	0.3853	0.0000	0.0034
8	0.0915	0.3242	0.5843	0.0000	0.0036
9	0.0293	0.1875	0.7832	0.0000	0.0039
10	0.0000	0.0000	1.0000	0.0000	0.0042

5 conclusions

In this paper, we estimates the risk in portfolio management by using CVaR and applied mean-CVaR model to optimize portfolio. We used the t copulas to described dependence structure between individual stock returns affects the returns of portfolio. We conducted our analysis in two steps. First, we examined the dependence structure of stock returns obtained from ARMA-GARCH process. Second, we studied how the dependence structure of the stock returns affects portfolio optimization. We used an optimization technique to allocate risk in the portfolio. It is reasonable to conclude that t copulas can described dependency structure of the asset in the portfolio management.

Acknowledgement(s) : The authors thank Prof. Dr. Hung T. Nguyen for his helpful comments and suggestions.

References

[1] C. Acerbi and D. Tasche. (2002). Expected Shortfall: A Natural Coherent Alternative to Value at Risk. *Economic Notes*, Vol. 31 (2), 379-338.

[2] C. Acerbi and D. Tasche. (2002) On the coherence of expected shortfall. *Journal of Banking & Finance*, Vol. 26 (7), 1487-1503.

[3] A. Ang and G. Bekaert. (2002). International Asset Allocation With Regime Shifts. *The review of financial studies*, Vol. 15 (4), 1137-1187.

[4] P. Artzner, F. Delbaen, J.M. Erber and D. Heath. (1998). Coherent measures of risk. *Mathematical Finance*, Vol. 9, 203-228.

[5] C.L. Chang, M. McAleer, and R. Tansuchat. (2011). Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Economics*, Vol. 33 (5), 912-923.

[6] C.L. Chang, M. McAleer, and R. Tansuchat. (2011). Crude oil hedging strategies using dynamic multivariate GARCH. *Energy Economics*, Vol. 33 (5), 912-923.

[7] S., Sriboonchitta, H.T., Nguyen, A., Wiboonpongse and J., Liu. (2013). Modeling volatility and dependency of agricultural price and production indices of Thailand: Static versus time varying copulas. *International Journal of Approximate Reasoning*, Vol. 54, 793-808.

[8] S. Sriboonchitta, J. Liu and A. Wiboonpongse. (2014). Vine copula-cross entropy evaluation of dependence structure and financial risk in agricultural commodity index returns. *Advances in Intelligent Systems and Computing*, Vol. 251, 275-287.

[9] R.T. Rockafellar and S. Uryasev. (2002). Conditional Value-at-risk for general loss distribution. *Journal of Banking & Finance*, Vol. 26, 1443-1471.

[10] R.T. Rockafellar and S. Uryasev. (2000). Optimization of conditional Value-at-risk. *Journal of Risk*, Vol. 2 , 21-41.

[11] E. Kole K. Koedijk and M. Verbeek. (2007). Selecting copulas for risk management. *Journal of Banking & Finance*, Vol. 31 (8), 2405-2423.

[12] G. Szego. Measure of risk. (2005). *European Journal of Operational Research*, Vol. 163, 5-19.

[13] J.C.C. Chan and D.P. Kroese. (2010). Efficient estimation of large portfolio loss probabilities in t-copula models. *European Journal of Operational Research*, Vol. 205(2), 361-367.

[14] L. Hu. (2006). Dependence patterns across financial markets: A mixed copula approach. *Applied Financial Economics*, Vol. 16, 717-729.

[15] S.R. Das and R., Uppal. (2004). International Portfolio Choice. *The Journal of Finance*, Vol. 59 (6), 2809-2834.

[16] AJ. Patton. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, Vol. 2 (1), 130-168.

[17] G. Ch. Pflug. (2000). Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk. *Probabilistic Constrained Optimization*, Vol. 49, 272-281.

[18] D. J. Miller and W. Liu. (2006). Improved estimation of portfolio value-at-risk under copula models with mixed marginals. *Journal of Futures Markets*, 26, 997-1018.

[19] P. Embrechts, F. Lindskog and A. McNeil. (2003). Modeling Dependence with Copulas and Applications to Risk Management. *Handbook of heavy tailed distributions in finance*, 8(1), 329-384.

[20] T.H. Lee and X. Long. (2009). Copula-based multivariate GARCH model with uncorrelated dependent errors. *Journal of Econometrics*, Vol. 150 (2), 207-218.

[21] C.R. Harvey and A. Siddique. (2000). Conditional skewness in asset pricing texts. *The Journal of Finance*, 55(3), 1263-1295.

[22] S. Demarta and A.J. McNeil. (2005). The t copula and related copulas. *International statistical review*, 73(1), 111-129.

(Received 30 May 2014)
(Accepted 10 September 2014)