Dependence of stock and commodity futures markets in India: implications for portfolio investment

1. Introduction-

In the last decade, India's commodity futures markets have experienced brisk growth. Thus, it is expected to have a large impact on the pricing of other assets and the investor portfolio decisions. This study endeavors to analyze how the fast pace of growth in India's commodity markets has been affected by the trends in its financial market. To this end, we examine the market co-movement between the commodity futures markets and stock markets in India. The information about the interdependence of commodity and stock markets is particularly relevant for investors because diversified portfolios may be composed of both commodities and stocks.

Recent empirical evidence shows evidence of increased co-movement between the commodity and stock markets since both markets are underpinned by some common factors. Furthermore, unlike stocks, commodities can serve as an inflation hedge. For this reason, investors are interested in adding commodity futures to their portfolios with the aim of diversifying and reducing the downside risk.

Empirically, the dependence relationship between stock and commodity markets has often been examined through assessing the correlation coefficient and using different multivariate models.Here, we plan to investigate the commodity-stock market dependence structure using copula functions.Many methods have been used before but this method is flexible since as it allows one to separately model the marginal behavior of the commodity and stock prices and the dependence structure. The study particularly analyzes the dependence structure between NIFTY50 and three commodity indices mentioned below.. The marginal distributions of asset returns for each index are modeled by an autoregressive moving average (ARMA) model with threshold GARCH (TGARCH) errors, whereas the market dependence is evaluated using different copula specifications.

COMDEX is an index which is a Composite commodity index. It has three subindices under it.The time span considered for the analysis is 17.09.2007 to 09.05.2016.Weekly data has been considered for the analysis.

MCX COMDEX	Commodity	Weight (New)	Group Adjusted Wts.
MCX METAL INDEX	Gold	15.16%	40.0%
	Silver	4.07%	
	Copper	7.56%	
	Aluminum	2.87%	
	Nickel	5.12%	
	Zinc	3.09%	
	Lead	2.13%	
MCX ENERGY INDEX	Crude Oil	35.22%	40.0%
	Natural Gas	4.78%	
MCX AGRI INDEX	Cardamom	2.01%	20.0%

	Mentha Oil	3.89%	
	Crude Palm Oil	6.32%	
	Cotton	7.78%	

2. Methodology

2.1 Copula functions

Let $R_{s,t}$ and $R_{c,t}$ (c = Metal,grain and agriculture) be random variables denoting India's stock and commodity sector futures returns, respectively, at time t. Moreover, let these assets' conditional continuous cumulative distribution functions (CDFs) be $F_s(R_{s,t}|y_{t-1})$ and $F_c(R_{c,t}|y_{t-1})$, respectively, where y_{t-1} denotes all past return information for the corresponding assets. Sklar's theorem states that the conditional joint distribution function G for $R_{s,t}$ and $R_{c,t}$ has a unique copula representation, C, such that:

$$G(R_{s,t},R_{c,t}\mid y_{t-1}) = C\left((F_{s,t}\mid y_{t-1}),F_{c,t}(R_{c,t}\mid y_{t-1})\right) \tag{1}$$

Assuming all CDFs are differentiable, the joint density can be obtained as:

$$g(R_{s,t},R_{c,t}\mid y_{t-1}) = \frac{\partial G(R_{s,t},R_{c,t}\mid y_{t-1})}{\partial R_{s,t}\partial R_{c,t}} = c(F_{s,t}(R_{s,t}\mid y_{t-1}),F_{c,t}(R_{c,t}\mid y_{t-1})\mid y_{t-1})\times f_{s,t}(R_{s,t}\mid y_{t-1})\times f_{c,t}(R_{c,t}\mid y_{t-1}) \tag{2}$$

where $c(u_t,v_t) = \partial^2 C(u_t,v_t\mid\psi_{t-1})/\partial u_t \partial v_t$, with $u_t = F_s(R_{s,t}|\psi_{t-1})$ and where $v_t = F_c(R_{c,t}|\psi_{t-1})$ is the conditional copula density. Thus, the conditional bivariate density function, $g(R_{s,t},R_{c,t}\mid y_{t-1})$, is represented by the product of the copula density and the two conditional marginal densities $f_{s,t}(R_{s,t}|\psi_{t-1})$ and $f_{c,t}(R_{c,t}|\psi_{t-1})$. Accordingly, the log-likelihood function can be written as:

$$\log\left[g(R_{s,t},R_{c,t}\mid y_{t-1})\right] = \log\left[c(F_{s,t}(R_{s,t}\mid y_{t-1})\right] + \log\left[f_{s,t}(R_{s,t}\mid y_{t-1})\right] + \log\left[f_{c,t}(R_{c,t}\mid y_{t-1})\right] \tag{3}$$

The parameters for the copula density and the marginal functions can be obtained by maximizing Eq. (3), using the two-step estimation procedure proposed by Joe (1997) called the inference for margins (IFM). This consists of first obtaining the marginal density parameters for both marginals via maximum likelihood and then using these estimates to obtain the copula parameters (q_c) by solving the following expression:

$$\hat{q}_c = \arg\max_{q_c} \sum_{t=1}^{T} \ln c(\hat{u}_t,\hat{v}_t;q_c) \tag{4}$$

where $\hat{u}_t = F_{s,t}(R_{s,t} \mid y_{t-1}; \hat{q}_s)$, $\hat{v}_t = F_{c,t}(R_{c,t} \mid y_{t-1}; \hat{q}_c)$, and $\hat{q}_s$ and $\hat{q}_c$ are the estimates of the marginal density parameters.

Modeling dependence using copulas is appealing since copulas offer flexibility in modeling separately the marginals and dependence structure, given by the copula function. Furthermore, copula functions are invariant to monotonic transformations of the variables because they relate the quantiles of the marginal distributions rather than the original variables. Copulas also provide a more complete description of dependence, offering information on both average dependence and tail dependence.

2.2 *The marginal distribution model*

In order to capture the main features of the stock and commodity futures returns described in Table 1, we employ an ARMA(*p*,*q*)-TGARCH(*r*,*m*) model with Student-t distribution errors for the marginal distributions. Thus, for the stock or commodity returns, denoted by R_t, the marginal model is given by:

$$R_t = \phi_0 + \sum_{j=1}^{p} \phi_j R_{t-1} + \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{7}$$

where ε_t is the product of the conditional volatility and the innovation z_t, $\varepsilon_t = \sigma_t z_t$, such that:

$$\sqrt{\frac{v}{v-2}} z_t \sim i.i.d.t_v \tag{8}$$

$$\sigma_t^2 = \omega + \sum_{j=1}^{r} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{m} \alpha_j \varepsilon_{t-i}^2 + \sum_{j=1}^{m} \gamma_j \varepsilon_{t-j} I_{t-j} \tag{9}$$

Eq. (7) decomposes the return at time *t* into a constant, an innovation (ε_t) and lags of R_t and ε_t, where ϕ and ϑ are the AR and MA parameters, respectively. Eq. (8) assumes that the standardized residuals follow the Student-t distribution, with *v* degrees of freedom. The leverage term in Eq. (9) is responsible for capturing the leverage effect, with $I_{t-j} = 1$ when ε_{t-j} is negative and $I_{t-j} = 0$ otherwise. In addition, the order of the ARMA terms and the lag orders of the TGARCH model are all specified according to the Akaike information criterion (AIC).

2.3 Bivariate Copulas

To model the dependence structure, we consider different types of copula functions with symmetric and asymmetric tail behavior. First, we consider elliptical Gaussian and Student-t copulas which are usual choices for the market dependence structure. They are defined, respectively, as:

$$C^{Gaussian}(u_t, v_t; \rho) = \Phi(\Phi^{-1}(u_t), \Phi^{-1}(v_t)) \quad (10)$$

$$C^{Student-t}(u_t, v_t; r, v) = T_v(t_v^{-1}(u_t), t_v^{-1}(v_t)) \quad (11)$$

where Φ is the bivariate standard normal CDF with correlation ρ ($-1 < \rho < 1$); $\Phi^{-1}(u_t)$ and $\Phi^{-1}(v_t)$ are standard normal quantile functions; T is the bivariate Student-t CDF with degree-of-freedom parameter v and correlation ρ ($-1 < \rho < 1$); and $t_v^{-1}(u_t)$ and $t_v^{-1}(v_t)$ are the quantile functions of the univariate Student-t distributions. Both copulas display symmetric dependence

Second, we consider two other copulas with symmetric tail dependence, namely, the Plackett and the Frank copulas, specified, respectively, as:

$$C^{Plackett}(u_t, v_t; \pi) = \frac{1}{2(\pi-1)}(1+(\pi-1)(u_t+v_t) - \sqrt{(1+(\pi-1)(u_t+v_t))^2 - 4\pi(\pi-1)u_t v_t}) \quad (12)$$

$$C^{Frank}(u_t, v_t; \lambda) = \frac{-1}{\lambda}\log\left(\frac{(1-e^{-\lambda}) - (1-e^{-\lambda u_t})(1-e^{-\lambda v_t})}{(1-e^{-\lambda})}\right) \quad (13)$$

where $\pi \in [0,\infty)\setminus\{1\}$ and $\lambda \in (-\infty,\infty)\setminus\{0\}$. Both copulas display tail independence.

Given that dependence may change under different market circumstances — in booms or bursts, for instance — we consider copula functions with asymmetric tail dependence structures. The Gumbel copula reflects upper tail dependence, whereas its rotation reflects lower tail dependence, given, respectively, by:

$$C^{Gumbel}(u_t, v_t; \delta) = \exp\left(-\left((-\log u_t)^{\delta} + (-\log v_t)^{\delta}\right)^{1/\delta}\right) \quad (14)$$

$$C^{Rotated\_Gumbel}(u_t, v_t; \delta) = u_t + v_t - 1 + C^{Gumbel}(1-u_t, 1-v_t; \delta) \quad (15)$$

where $\delta \in (1,\infty)$. We also consider the symmetrized Joe-Clayton (SJC) copula. It is given by:

$$C^{SJC}(u_t, v_t; \lambda_U^{SJC}, \lambda_L^{SJC}) = 0.5(C^{JC}(u_t, v_t; \lambda_U^{JC}, \lambda_L^{JC}) + C^{JC}(1-u_t, 1-v_t; \lambda_U^{JC}, \lambda_L^{JC}) + u_t + v_t - 1) \quad (16)$$

where $C^{JC}(u_t, v_t; \lambda_U^{JC}, \lambda_L^{JC}) = 1-(1-\{[1-(1-u_t)^{\kappa}]^{-\gamma} + [1-(1-v_t)^{\kappa}]^{-\gamma} - 1\}^{-1/\gamma})^{1/\kappa}$, $\kappa = 1/\log_2(2-\lambda_U^{JC})$ and $\gamma = -1/\log_2(\lambda_L^{JC})$. Moreover, $\lambda_U^{SJC}(v) \in (0,1)$ and $\lambda_L^{SJC}(v) \in (0,1)$. For this copula function, the tail dependence coefficients are themselves the parameters of the copula. If $\lambda_U^{SJC} = \lambda_L^{SJC}$, then the market structure is symmetric, otherwise it is asymmetric.

Results-

NOTE: All the results are tabulated at 5 percent level of significance.

First, we plot the data.The plot is as follows-

Figure1-

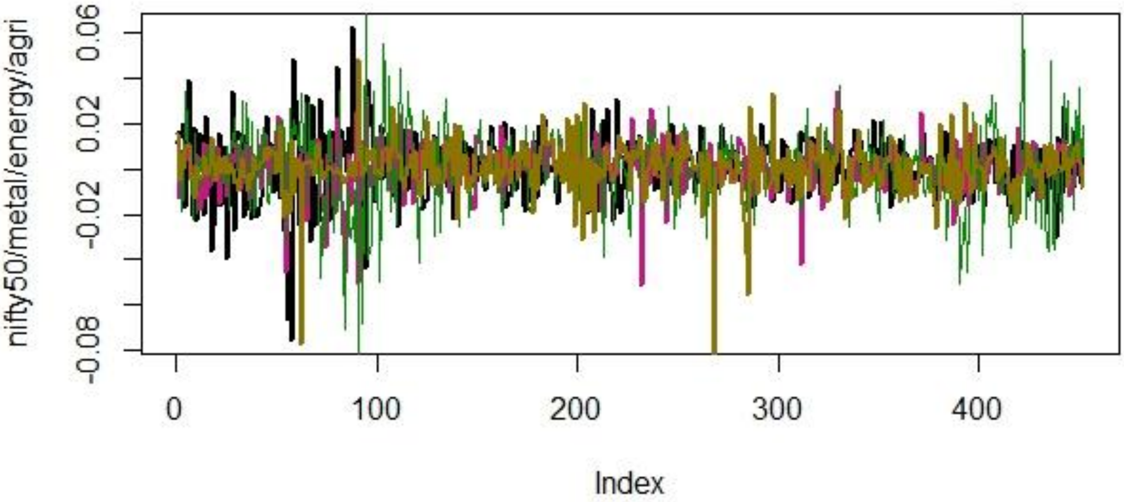


Figure 1 depicts the stock and commodity futures price dynamics throughout the sampling period.It can be seen that they exhibit similar behaviour. The black indicates the return for NIFTY50, violet for METAL, green for ENERGY and golden for AGRICULTURE. After this, the descriptive statistics are studied from the data. The reported statistics (Table 1) show that the return series are all skewed and exhibit excess kurtosis, indicating that those returns are not normally distributed. The highly significant Jarque-Bera statistics confirm the evidence of a non-normal distribution for returns. Meanwhile, the Ljung-Box statistics and the ARCH-Lagrange multiplier (LM) test suggest the presence of serial correlation and ARCH effects in returns.

Table1-Descriptive statistics

	NIFTY50	METAL	ENERGY	AGRICULTURE
Mean	0.0004852	0.0004579	2.647e-05	0.0002986
Std dev	0.01406738	0.01078335	0.01885222	0.01157467
Max	0.0623500	0.0337457	7.941e-02	0.0477852
Min	-0.0754600	-0.0510021	-9.276e-02	-0.0912027
Skewness	-0.3845517	-0.8574728	-0.2775518	-1.818419
Kurtosis	3.442983	3.607287	2.577093	13.88361

Jarque-bera statistic	238.19(+)	304.83 (+)	133.34 (+)	3920.3 (+)
Ljung-box	0.085277(+)	2.0081(+)	0.1954(+)	3.2845(+)
ARCH-LM	91.49(+)	32.448(+)	89.848(+)	1.6441(+)

NOTE: (+) denotes that H0 is rejected

Table 2 shows the Pearson linear correlation between the Indian stock market and the three commodity futures markets. The positive value indicates that the stock and commodity futures markets move together and in the same direction while negative value indicated opposite direction.

Table2-Pearson's Correlation coefficient

	NIFTY50	METAL	ENERGY	AGRICULTURE
NIFTY50	1	-0.004093025	-0.09536905	0.07387603
METAL	-0.004093025	1	0.46835420	0.11688302
ENERGY	-0.095369048	0.468354199	1	0.08382863
AGRICULTURE	0.073876030	0.116883016	0.08382863	1

Marginal model estimates-

Different values of p,q,r and m are considered.The AIC for each of these models is very close. For NIFTY, ARMA(4,0)+GARCH(1,1) is considered, for METAL and ENERGY, ARMA(2,0)+GARCH(1,1) and for agriculture, ARMA(6,0)+GARCH(1,1) is considered.

Model parameters of the marginal distribution-

1.NIFTY50-

```
Estimate  Std. Error  t value Pr(>|t|)
mu      7.714e-04   5.197e-04    1.484  0.13769
ar1     2.323e-02   4.849e-02    0.479  0.63189
ar2     3.523e-02   4.897e-02    0.719  0.47185
ar3    -5.379e-02   4.800e-02   -1.121  0.26242
ar4     2.411e-02   4.846e-02    0.498  0.61873
omega   2.924e-06   1.841e-06    1.588  0.11221
alpha1  8.149e-02   2.511e-02    3.245  0.00117 **
beta1   9.007e-01   2.902e-02   31.036  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:
 1346.8    normalized:  2.979647
```

2.METAL

```
Estimate           Std. Error   t value  Pr(>|t|)

mu      4.237e-04   4.698e-04    0.902   0.3671
```

```
ar1     6.210e-02   5.326e-02    1.166   0.2437
ar2    -1.065e-03   5.118e-02   -0.021   0.9834
omega   1.063e-05   6.769e-06    1.570   0.1163
alpha1  9.036e-02   4.599e-02    1.965   0.0494 *
beta1   8.178e-01   9.547e-02    8.565   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log Likelihood:
 1425.737    normalized:  3.154286
```

3.ENERGY

```
Estimate  Std. Error  t value Pr(>|t|)

mu      2.647e-04   6.996e-04    0.378  0.70522
ar1     2.942e-02   4.868e-02    0.604  0.54560
ar2    -5.607e-02   4.926e-02   -1.138  0.25499
omega   8.068e-06   4.251e-06    1.898  0.05774 .
alpha1  1.138e-01   2.943e-02    3.866  0.00011 ***
beta1   8.662e-01   3.193e-02   27.125  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log Likelihood:
 1205.881    normalized:  2.667878
```

4.AGRICULTURE

```
Estimate  Std. Error  t value Pr(>|t|)

mu      1.153e-03   5.802e-04    1.988  0.04686 *
ar1     1.463e-01   6.219e-02    2.353  0.01865 *
ar2     7.588e-02   6.308e-02    1.203  0.22900
ar3    -3.167e-02   5.792e-02   -0.547  0.58454
ar4    -7.834e-02   5.342e-02   -1.467  0.14247
ar5    -1.526e-02   5.723e-02   -0.267  0.78969
ar6     9.383e-03   4.814e-02    0.195  0.84547
omega   2.531e-05   8.574e-06    2.952  0.00316 **
alpha1  2.326e-01   7.624e-02    3.051  0.00228 **
beta1   6.284e-01   7.518e-02    8.359  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log Likelihood:
 1390.709    normalized:  3.07679
```

After this,it is important to check for the goodness of fit. The KS test is ued to test whether the marginal models are not misspecified and the copula model accurately captures the co-movement of Indian commodity futures and stock markets. KS test is performed to examine whether the probability integral transforms are uniform (0,1).

KS test-

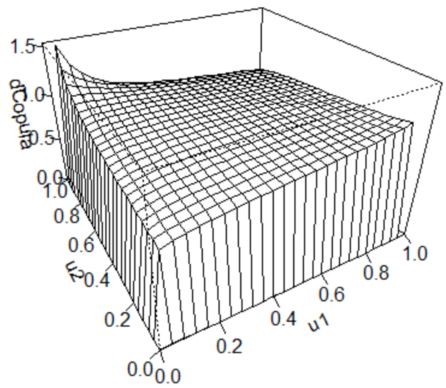
```
p-value-for NIFTY50-  0.9541
p-value-for Metal- 0.7321
p-values-for Energy- 0.94784
p-values-for agriculture- 0.96321
```

```
Thus,H0 is rejected for all the above and we can conclude that the the pro
bability integral transform of the marginal distribution are uniform(0,1).
```

After modelling the marginal distributions, the copula estimates can be calculated by the Maximum likelihood method. We need to choose the copula that most adequately represents the dependence structure between markets. This can be done by calculating the AIC value.

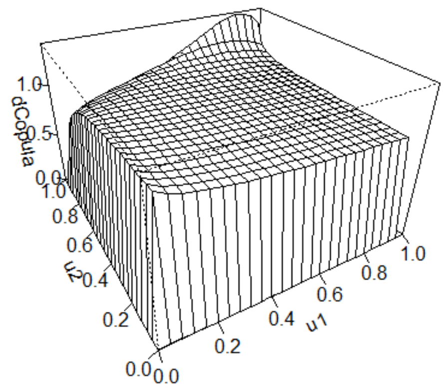
- NSE,Energy

```
Rotated Gumbel 90 degrees (par = -1.05, tau = -0.05)
```



- NSE and Metal are found to be independent
- NSE , Agriculture

  ```
  Rotated Tawn type 1 180 degrees (par = 1.42, par2 = 0.09, tau =0.05)
  ```



Appendix

R codes

```
library(readxl)
library(readr)
nif=read_csv("C:/Users/sony/Desktop/project/nifty.csv")
MCXMETAL=read_excel("C:/Users/sony/Desktop/project/MCXMETAL.xlsx")
MCXENERGY=read_excel("C:/Users/sony/Desktop/project/MCXENERGY.xlsx")
MCXAGRI=read_excel("C:/Users/sony/Desktop/project/MCXAGRI.xlsx")
MCXMETAL=as.matrix(MCXMETAL)
MCXENERGY=as.matrix(MCXENERGY)
MCXAGRI=as.matrix(MCXAGRI)
metal=as.numeric(MCXMETAL[,8])
energy=as.numeric(MCXENERGY[,8])
agri=as.numeric(MCXAGRI[,8])
nift=as.matrix(nif)
nift
nifty=as.numeric(nift[,9])
n=length(nifty)
n
nifty
m=length(metal)
m
#nifty data is in reverse order,thus correction-
```

```
nse=numeric(n)
j=1
for(i in 1:n)
{
  nse[i]=nifty[n-j+1]
  j=j+1
}
nse



###data is daily,we do time aggregation to make it weekly
m=length(metal)
me=numeric(n)
en=numeric(n)
ag=numeric(n)
k=1
for(i in 1:n)
{
  j=1
  sum1=0
  sum2=0
  sum3=0
  while(j<7)
  {
    sum1=sum1+metal[k]
    sum2=sum2+energy[k]
    sum3=sum3+agri[k]
    j=j+1
    k=k+1
  }
  me[i]=sum1
  en[i]=sum2
  ag[i]=sum3
```

```
}
dat=cbind(nse,me,en,ag)
nrow(dat)
ncol(dat)
#plot of the data-figure 1
plot.new()
plot(nse,type="l",lwd=2,ylab="nifty50/metal/energy/agri")
lines(me,col="mediumvioletred",lwd=2.5)
lines(en,col="green4",lwd=1.5)
lines(ag,col="gold4",lwd=2)
#legend(350,-0.02,legend=c("violet-metal","green-energy","golden-agriculture"))
library(timeDate)
summary(dat)
sd(nse)
skewness(nse)
kurtosis(nse)
sd(me)
skewness(me)
kurtosis(me)
sd(en)
skewness(en)
kurtosis(en)
sd(ag)
skewness(ag)
kurtosis(ag)

#jarque bera test for testing normality
library(tseries)
jarque.bera.test(nse)
jarque.bera.test(me)
jarque.bera.test(en)
jarque.bera.test(ag)

#L-jung box test for correlation
```

```
Box.test(nse, lag = 1, type = "Ljung-Box")
Box.test(me, lag = 1, type = "Ljung-Box")
Box.test(en, lag = 1, type = "Ljung-Box")
Box.test(ag, lag = 1, type = "Ljung-Box")

#ARCH LM test for autoregressive conditional heteroschedasticity
library(FinTS)
ArchTest(nse)
ArchTest(me)
ArchTest(en)
ArchTest(ag)

cor(dat,method="pearson")

##finding p,q,r and m for the ARMA-GARCH model
library(fGarch)
ar=matrix(0,5,5)
gar=matrix(0,5,5)

##to find min aic,below formula is executed for different p,q,r and m and the values with
min aic are chosen
ar[3,3]=garchFit(formula=~arma(4,2)+garch(1,1),data=nse)@fit$ics[1]
##aic values
garchFit(formula=~arma(4,0)+garch(1,1),data=nse)@fit$ics[1]
garchFit(formula=~arma(2,0)+garch(1,1),data=me)@fit$ics[1]
garchFit(formula=~arma(2,0)+garch(1,1),data=en)@fit$ics[1]
garchFit(formula=~arma(6,0)+garch(1,1),data=ag)@fit$ics[1]


a=garchFit(formula=~arma(4,0)+garch(1,1),data=nse)
b=garchFit(formula=~arma(2,0)+garch(1,1),data=me)
c=garchFit(formula=~arma(2,0)+garch(1,1),data=en)
d=garchFit(formula=~arma(6,0)+garch(1,1),data=ag)
x=quantile(ecdf(nse),probs=seq(0,1,0.01))
```

```
ks.test(x,punif)
y=quantile(ecdf(me),probs=seq(0,1,0.01))
ks.test(y,punif)
z=quantile(ecdf(en),probs=seq(0,1,0.01))
ks.test(z,punif)

uv=quantile(ecdf(ag),probs=seq(0,1,0.01))
ks.test(uv,punif)
library(QRM)
sig=cor(cbind(nse,me))
summary(fit.tcopula(cbind(nse,me),2, method ="Spearman"))
summary(fit.tcopula(cbind(nse,me), method ="Spearman"))

##modelling the copulae
library(copula)
library(VineCopula)
u <-pobs(nse)
v <- pobs(en)
cop.en.nse<- BiCopSelect(u,v,familyset=NA)
cop.en.nse
persp(r90GumbelCopula(par = -1.05),dCopula)


u <-pobs(nse)
v <- pobs(ag)
cop.ag.nse<- BiCopSelect(u,v,familyset=NA)
cop.ag.nse
persp( tawnT1Copula(param = c(1.42, 0.09)),dCopula)

u <-pobs(nse)
v <- pobs(me)
cop.me.nse<- BiCopSelect(u,v,familyset=NA)
cop.me.nse
```

Data : Attached