VALUE-AT-RISK FORECAST USING COPULAS WITH REALIZED VOLATILITY: HAR, EWMA AND GARCH

Guilherme Ribeiro de Macêdo[a]
Igor Alexandre Clemente de Morais[b]
Pedro Lutz Ramos[c]
Marcelo Savino Portugal[d]

Tema : Gestão de Risco e Volatilidade.

[a] Professor do Programa de Pós-Graduação em Economia-UFRGS. Email: ribeiroguilherme@gmail.com.
[b] Professor do Programa de Pós-Graduação em Economia-Unisinos. Email: igor@vokin.com.br.
[c] Doutorando do Programa de Pós-Graduação em Economia-UFRGS. Email: plr2010@gmail.com.
[d] Professor do Programa de Pós-Graduação em Economia-UFRGS. Email: msp@ufrgs.br

Value-At-Risk Forecast Using Copulas With Realized Volatility: HAR, EWMA and GARCH

Abstract

Through empirical research is identified that the hypothesis of normal distribution of returns is no longer observed while verifying the existence of heavy tails and asymmetries in the distribution. Thus, the article has aimed empirically apply copula models using techniques of realized volatility (HAR) with high-frequency data and perform the calculation of Value at Risk for different periods. The results show that HAR models copula exhibited superior performance with respect to EWMA and GARCH in intervals of time greater than five minutes.

Keywords: Realized Volatility, Copula, Value at Risk

1. Introduction

The modeling of dependence between financial series over time is greatly important for asset allocation among portfolios and for risk management as well. Traditional asset allocation models take as proxy the linear correlation, which is a measure of intertemporal dependence between two random variables. Examples include CAPM (Capital Asset Pricing Model) and APT (Arbitrage Pricing Theory) (Campbell, Lo, e MacKinlay, 1997) which use correlation as a measure of dependence between different financial tools. They are derived under the hypothesis of a Gaussian multivariate distribution, i.e. a normal distribution of returns, to achieve an optimal portfolio.

The growing utilization of derivatives and related products in portfolios indicate the limitations of linear correlation in risk management modeling (Wang, 1997). That is so because the hypothesis on normal distribution of returns is no longer observed; on the contrary, heavy tails and distribution asymmetry are evidenced by empirical research. The increasing number of publications on economics that use copula modeling is explained by how effectively this statistical technique deals with the evidence of non-normality of financial asset return series. Non-normality appears as “volatility smile”, which is found in series of stock options near expiry date; also as “heavy tails” in portfolios of institutions, and therefore in institution risk management.

Value at Risk is a statistical technique aimed at calculating the maximum loss of a portfolio over a given time horizon considering an adopted level of significance. Originally created by J.P Morgan, it is today the main tool for risk mensuration, especially that of market and credit. Despite the demand for simplifier hypotheses, this statistical model is widely adopted for simplicity, both of implementation and interpretation. However, it is precisely some non-empirically verifiable hypotheses that cause a great variation in the results. The presence of heavy tails in the series poses a problem for determining the joint probability distribution, implying great difficulty to measure the degree of exposure to risk factors. That interferes with the correct and effective portfolio risk management, as the presence of non-normality makes it impossible to separate the effects of assets with different characteristics[2]. In cases of crises and bubbles, the portfolio may be riskier than desired or excessively conservative. In that sense, using copulas is interesting because it allows the separation of the marginal distribution of each asset from the dependence structure of variables. That implies discarding linear correlation.

In addition to theoretical contribution by applying the copula technique to VaR calculation, this article is of great empirical importance considering the recent developments in banking regulations. The article has two main purposes: to empirically apply copula models using realized volatility techniques (HAR) with high-frequency data; and to calculate the Value at Risk (VaR) for different periods. For such, the article is organized in five sections. The first section describes the main purposes of the paper. The second section presents: the main copula models used in finance to model the degree of dependence between two variables; the estimation of these structures; adequacy tests; and application to VaR calculation. The third section is dedicated to: the application of the microstructure correction technique to Brazilian intraday asset series to generate HAR; and copula modeling for market risk calculation using a theoretical portfolio with collected data. The fourth and fifth sections present the findings and point out possible future applications.

---

[2] For example, it is difficult to separate the effects of an asset backed by exchange from another backed by the SELIC Rate.

2. Methodology

The dependence between random variables X_1, X_2, X_3,..., X_n is formally described by the joint distribution function: $F(x_1,x_2, \ldots , x_n) = P[X_1 \leq x_1,\ldots,X_n \leq x_n]$. The concept of copulas emerges from the separation of the function F above into two parts: one describing the dependence structure and the other describing the marginal behavior of the variables.

The unicity of the copula-based representation for any multivariate continuous distribution is always guaranteed and the results are formally demonstrated in Schweizer and Sklar (1983, Chapter 6).

The guarantee of unicity allows one to obtain some families of copulas that can be applied to series of financial assets, which is the main object of analysis of this project.

To independent random variables, the copula is as follows:

$$C(x_1, x_2, \ldots , x_n) = x_1.x_2. \ldots .x_n$$

For the Gaussian copula, with $F_1 = \Phi$ and correlation ρ and considering that $Y \mid X = x \sim N(\rho x, 1\text{-}\rho^2)$, we have:

$$\lambda = 2 \lim_{x\to\infty} \Phi(\frac{x\sqrt{1-\rho}}{\sqrt{1+\rho}})$$

The interpretation of this result indicates that this family of copulas has asymptotic independence, for $\rho < 1$. In other words, extreme events seem to occur independently of each other in each tail. (Sibuya(1961) and Resnick (1987)).

However, for the bivariate distribution t with ν degrees of freedom and correlation ρ, the value of the tail dependence coefficient is:

$$\lambda = 2t_{\nu+1}(\frac{\sqrt{(\nu+1)(1-\rho)}}{\sqrt{1+\rho}})$$

A critical point for the methodology is to determine the dependence structure. Suppose that there are two vectors of random variables and that the marginal distributions are normal with null mean and variance σ_a^2 e σ_b^2, respectively. Using the Sklar Theorem[3], we have:

$$G.(\epsilon_a,\epsilon_b) = C.(N(\epsilon_a/\sigma_{a,.}), N(\epsilon_b/\sigma_{b,.}))$$

In the above equation, the notation "." means the choice of any copula function. The goal is to obtain the maximum likelihood function to find the parameters ($\omega.$, $\sigma_{a,.}$, $\sigma_{b,.}$) of the joint distribution function of G.

The maximum likelihood function is then written as:

$$\ln L.(\omega.,\sigma_{a,.}, \sigma_{b,.} ; \epsilon_{a,t}, \epsilon_{b,t}) = \sum_{t=1}^{T} \ln \frac{1}{\sigma_{a,.}\sigma_{b,.}} c.\,(N\left(\frac{\epsilon_{a,t}}{\sigma_{a,.}}\right), N(\frac{\epsilon_{b,t}}{\sigma_{b,.}})N'(\frac{\epsilon_{a,t}}{\sigma_{a,.}})N'(\frac{\epsilon_{b,t}}{\sigma_{b,.}}))$$

In the above function, N' is the density of the standard normal distribution and c. is the density of some known copula. As the estimates of parameters of the copula $\widehat{\omega}_T$ have maximum likelihood properties[4], the estimates of the tail dependence parameters $\hat{\lambda}_T = \lambda(\widehat{\omega}_T)$ will be consistent and asymptotically distributed according to a normal distribution with:

$$\sqrt{T}\left(\hat{\lambda}_T - \lambda\right) \to N(0,\sigma_\lambda^2)$$

The variance σ_λ^2 can be approximated by using the first order of the Taylor polynomial [5] with:

$$\sigma_{\hat{\lambda}}^2 = (\frac{\partial\lambda(\widehat{\omega}_T)}{\partial\omega_T})^T \sum\nolimits_{\widehat{\omega}_T} \frac{\partial\lambda(\widehat{\omega}_T)}{\partial\omega_T}$$

---

[3] As explained in the beginning of this section, the Sklar theorem guarantees the unicity so we have the following result: $H(z_1,z_2) = C(F_{Z_1}(z_1), F_{Z_2}(z_2))$. The interpretation is immediate: the dependence structure and the univariate marginals can be separated.

[4] That is, they allow the maximization of the function L.

[5] A function can be approximated by the Taylor polynomial through the formula $f(x) = \sum_{|\alpha|=0}^{n} \frac{1}{\alpha!} \frac{\partial^\alpha f(a)}{\partial x^\alpha}(x-a)^\alpha + \sum_{|\alpha|=n+1} R_\alpha(x)(x-a)^\alpha$

Another estimation possibility is to consider some non-parametric model based on empirical distribution. According to Deheuvels (1979), the result consists of copulas with discontinuities and without geometrical interpretation[6].

Fermanian (2005) discusses the difficulty of conducting tests given that the cumulative distribution functions of the marginals are unknown. Therefore, this author states that the adequacy test problem is still a field for great development in statistics. In that article, he proposes some alternatives for that issue.

To Fermanian (2005), the most usual method is to conduct the test separately on each marginal distribution and, once it has been accepted, to proceed with it on the multivariate distribution as a whole. Although it is simple from a methodological viewpoint, there are some computational restrictions for the testing part on the complete multivariate distribution. Moreover, it is more important to study the dependence structure, independently of the specifications of the marginals.

In order to correct the microstructure problem, two steps are required. The first is to determine the optimal sampling frequency, and the second is to apply the filter according to the methodology proposed by Hansen, Large and Lunde (2006). The optimal sampling system of Bandi and Russel (2005a, 2006) approximates the formula, arbitrating the optimal frequency of the estimator variance.

Bandi and Russel (2005a, 2006) derive and minimize the function of error caused by the microstructure problem, in order to ensure the convergence to a continuous function of integrated volatility.

By using this methodology, it is possible to significantly minimize the microstructure problem, as Zhang, Mykland and Aït-Sahalia (2005) point out through simulation. Moreover, Hansen and Lunde (2006) show that the microstructure error for the DJIA stocks is small in sampling frequencies lower than 20 minutes. This indicates that this technique would be sufficient to implement the models, as approached by Andersen (2007) in his article.

Nevertheless, the discussion on how to optimally calculate the Integrated Quarticity induces many authors, like Bandi and Russel (2005a, 2006b), to calculate it with 15-minute sampling frequency, thus avoiding a complex method demonstrated in Zhang, Mykland and Aït-Sahalia (2005). Hence, given the limitations of optimal choice the sampling is performed in various frequencies (1, 2, 5, 15 and 30 minutes) in order to solve this problem. However, the variance of the estimators is calculated as demonstrated above, indicating the optimal choice for each stock.

The filter-based estimator was introduced by Ebens (1999) and Andersen, Bollerslev, Diebold and Ebens (2001). Its general idea was to estimate an autoregressive or moving average model from intraday return, as that autocorrelation and partial autocorrelation process is exclusively derived from a process generated by the microstructure problem. Hence, when estimating an AR or MA model the intraday return part that is microstructure bias can be identified; it is possible to filter it through the estimated parameters. However, Bandi and Russel (2005) criticize the model and show that the technique is not sufficiently effective to stop the tendency. Hence, Hansen, Large and Lunde (2006) demonstrate that a larger time lag in relation to the MA is necessary to ensure consistency.

Therefore, considering the error as being correlated to the latent price and under the assumption of serial independence, the price return is defined as following the MA(q) process:

$$r_{t,N} = u_{t,i} - \theta_1 u_{t,i-1} - \theta_2 u_{t,i-2} - \cdots - \theta_q u_{t,i-q}, \qquad i = 0, \ldots, N$$

where the sequence $\{u_{t,m}\}$ é $IID(0,\sigma^2_{u,t})$. Hence, the filter becomes:

$$RV^{MA} = \left(\frac{(1 - \theta_1 - \theta_2 - \cdots - \theta_q)^2}{1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2}\right) RV$$

[6] It is not possible to obtain the graph due to discontinuities.

However, as availability of data with frequency higher than 1 minute is extremely low for the series of assets traded at Bovespa, the authors show that the estimator consistency problem decreases by using an equidistant interval and considering that the volatility in the interval must be constant.

3. Data

This article analyzes four stocks traded at the São Paulo Stock Exchange, whose trading codes are PETR4 (Petrobras), USIM5 (Usiminas), GGBR4 (Gerdau) and VALE5 (Vale do Rio Doce). The analysis period was March 3, 2006 to April 30, 2010, from the database provided by BM&FBovespa Educational Institute. The assets were adopted in terms of data availability and liquidity, as this measure is important for Realized Volatility models given the microstructure problem. Daily closing prices were used for EWMA and GARCH data, whereas for Realized Volatility data the time considered was 10:00am to 5:00pm for trading sessions conducted in standard time, and 11:00am to 6:00pm when in daylight savings time[7]. It is worth noting that the extraction method is important for error-correction mechanisms, therefore this method was time related. Hence, the algorithm searches for data based on predetermined date and time, considering as valid the price closest to that time. This is important, as the provided database is not regular with regard to trading, leading to significant changes in the number of samples. Therefore, we do not have a calendar-based sampling model, but a hybrid system between trading and calendar. Finally, Stock Split adjustments were performed.

The following table presents the number of observations per asset:

Asset	Daily	30 minutes	15 minutes	5 minutes	2 minutes	1 minute
GGBR4	1026	14364	28728	86184	215461	430921
PETR4	1026	14364	28728	86184	215461	430921
USIM5	1026	14364	28728	86184	215461	430921
VALE5	1026	14364	28728	86184	215461	430921

Table 1 – Number of observations per fixed time interval

To compose the theoretical portfolio for Value at Risk calculation, each asset is assumed to be equally composed. The transaction costs necessary to rebalance that portfolio are ignored.

4. Results

The series of financial assets have as characteristic the volatility clustering, which is distinguished by the presence of heteroscedasticity. Therefore, serial autocorrelation tests on the return series are requered to prove that hypothesis. In order to do so, the Engle test is conducted on the four series of assets for 10, 15 and 20 lags.

The results are presented in the following table:

Asset	Hypothesis	Statistics	Critical Value	Lags
USIM5	1	202.8864	18.307	10
	1	240.405	24.9958	15
	1	264.3578	31.4104	20

[7] Bovespa changes trading hours according to daylight saving time, although the start and end of this period not necessarily coincide with the time change in trading sessions.

PETR4	1	237.1285	18.307	10
	1	272.4248	24.9958	15
	1	304.6029	31.4104	20
GGBR4	1	280.4628	18.307	10
	1	292.6219	24.9958	15
	1	302.5517	31.4104	20
VALE5	1	200.317	18.307	10
	1	205.2666	24.9958	15
	1	222.4825	31.4104	20

Table 2 – ARCH Test for heteroscedasticity detection

4.1 Estimation of Copulas with GARCH (1,1) models

By using the four GARCH (1,1) models estimated for each series of assets, standardized residuals are obtained that are utilized to estimate the cumulative probability function, with a Gaussian Kernel function as smoother[8]. This can be problematic concerning extreme values, as non-parametric smoothing functions only perform well within the distribution with the highest data frequency. However, it tends to perform poorly in the tails, which are precisely where the interest of VaR lies. In order to deal with this fact, the generalized Pareto distribution[9] is used in approximately 10% of the data, involving lower and upper tails, given a cut value. This procedure allows for the modeling of heavier tails, which is extremely advantageous for the purpose of this article.

After simulating standard errors via non-parametric modeling, a dependence structure between the marginals must be chosen. This article utilizes the maximum likelihood estimation algorithm, which is discussed in the methodology section. As the goal is to determine which copula fits best the VaR calculation, this step consists of four estimates for the same volatility model estimated: t, Clayton, Frank and Gumbel copulas

The theoretical portfolio required for dependence simulation will be composed of equal weights for the four assets. The transaction costs of rebalancing are assumed to be very low in comparison to the value of the portfolio, so this restriction can be relaxed.

The results show that Frank family copulas tend to forecast a larger loss for more extreme confidence intervals, indicating that they have a fatter tail. All the forecasts show an asymmetrical characteristic of distribution, indicating that, on average, losses are greater than gains in modulus. The Kupiec test for checking the null hypothesis that the adopted model is adequate can be found in the table below:

GARCH(1,1) Volatility Model					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	**78**	115	**80**	105
	% of the total number	**7.43%**	10.95%	**7.62%**	10.00%
	Kupiec (Acceptance region)	84 < N < 123			
5%	Number of violations	**34**	50	**36**	55
	% of the total number	**3.71%**	4.76%	**4.00%**	5.24%
	Kupiec (Acceptance region)	38 < N < 67			

[8] Gaussian kernel smoothing function is a non-parametric model to adjust empirical data to a known function.

[9] The probability density function for a generalized Pareto distribution is written as $y = f(x|k,\sigma,\theta) = (\frac{1}{\sigma})(1 + k\frac{(x-\theta)}{\sigma})^{-1-\frac{1}{k}}$

	Number of violations	**2**	10	**4**	10
1%	% of the total number	**1.14%**	0.95%	**0.67%**	0.95%
	Kupiec (Acceptance region)	5 < N < 18			

Table 3 – Kupiec Test for the GARCH(1,1) models estimated using Back Testing

It is observed that the VaR calculation caused non-acceptance of the null hypothesis for t and Frank copula models, indicating that the models can be excessively allocating. At other levels of confidence, all the models are accepted.

4.2. Estimation of copulas with EWMA model

Through the same procedure as described in the previous section, residuals from the series modeled by EWMA will be used to generate the marginal distribution and then to simulate the copula chosen for VaR calculation. The Kupiec test results are presented in the table below:

EWMA Volatility Model					
	Copulas	t	Clayton	Frank	Gumbel
	Number of violations	**130**	**75**	**81**	91
10%	% of the total number	**10.95%**	**7.14%**	**10.48%**	8.67%
	Kupiec (Acceptance region)	84 < N< 123			
	Number of violations	**71**	**34**	**28**	42
5%	% of the total number	**2.86%**	**4.76%**	**2.67%**	4.00%
	Kupiec (Acceptance region)	38 < N < 67			
	Number of violations	**23**	**3**	**4**	14
1%	% of the total number	**2.19%**	**2.67%**	**0.57%**	1.33%
	Kupiec (Acceptance region)	5 < N < 18			

Table 4 - Result of the Kupiec Test *for* EWMA models

As noticed from the above table, the Back Testing showed rejection of the VaR model for Clayton copulas, rejection of t copulas, and rejection of Frank copulas for all levels of significance. When compared to the GARCH (1,1) model previously estimated, the results suggest that EWMA is less effective concerning Value at Risk calculation. This is probably because the model is more rigid than GARCH as a value of 94% is fixed for the volatility decay, according to recommendation of Basileia III.

4.3. Estimation of copulas with HAR model

The section on the HAR model estimation presented the method in a three-group division: volatility without microstructure correction; volatility with microstructure correction using estimated optimal frequency; and volatility with microstructure correction using fixed time intervals. Therefore, this section is subdivided according to each estimated group.

4.3.1. Estimation of copulas with HAR models without correction of the data microstructure problem

The procedure for copula generation and Value at Risk calculation is the same presented for the EWMA and GARCH(1,1) models. The entry vector for tail dependence simulation is the standardized residual of the volatility model estimated for each series, denoted by: $ALL^1, ALL^2, ALL^5, ALL^{15}$ and ALL^{30}.

The Kupiec test for these models is presented in the tables below:

Volatility model HAR ALL[1]					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	121	86	**82**	**133**
	% of the total number	11.52%	8.19%	**8.10%**	**12.48%**
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	42	40	**29**	**74**
	% of the total number	4.00%	3.81%	**4.29%**	**4.10%**
	Kupiec (Acceptance region)	38 < N < 67			
1%	Number of violations	12	9	**2**	**22**
	% of the total number	1.14%	0.86%	**0.19%**	**0.67%**
	Kupiec (Acceptance region)	5 < N < 18			
Volatility model HAR ALL[2]					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	100	85	**81**	**91**
	% of the total number	9.52%	8.10%	**8.00%**	**8.67%**
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	46	42	**35**	**71**
	% of the total number	4.38%	4.00%	**4.86%**	**4.57%**
	Kupiec (Acceptance region)	38 < N < 67			
1%	Number of violations	9	7	**0**	**21**
	% of the total number	0.86%	0.67%	**0.00%**	**0.29%**
	Kupiec (Acceptance region)	5 < N < 18			

Table 5 - Kupiec tests for the HAR ALL[1] e HAR ALL[2] models

Volatility model HAR ALL[5]					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	111	91	87	105
	% of the total number	10.57%	8.67%	8.29%	10.00%
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	55	51	60	54
	% of the total number	5.24%	4.86%	5.71%	5.14%
	Kupiec (Acceptance region)	38 < N < 67			
1%	Number of violations	17	14	5	7
	% of the total number	1.62%	1.33%	0.48%	0.67%
	Kupiec (Acceptance region)	5 < N < 18			
Volatility model HAR ALL[15]					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	101	97	93	117
	% of the total number	9.62%	9.24%	8.86%	11.14%
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	47	56	40	43
	% of the total number	4.48%	5.33%	3.81%	4.10%
	Kupiec (Acceptance region)	38 < N < 67			

1%	Number of violations	15	12	7	10
	% of the total number	1.43%	1.14%	0.67%	0.95%
	Kupiec (Acceptance region)	5 < N < 18			
	Volatility model HAR ALL[30]				
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	97	101	110	120
	% of the total number	9.24%	9.62%	10.48%	11.43%
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	49	54	41	47
	% of the total number	4.67%	5.14%	3.90%	4.48%
	Kupiec (Acceptance region)	38 < N < 67			
1%	Number of violations	13	14	9	8
	% of the total number	1.24%	1.33%	0.86%	0.76%
	Kupiec (Acceptance region)	5 < N < 18			

Table 6 - Kupiec tests for the HAR ALL^5, HAR ALL^{15} and HAR ALL^{30} models

The Kupiec test tables indicate that higher frequencies without microstructure error correction tend to overestimate the VaR calculation, indicating a high allocation bias. From the 5-minute frequency onwards, it is observed that all models are accepted at the adopted levels of significance.

4.3.2. Estimation of copulas with HAR models using estimated optimal frequency

The estimation of optimal frequency for the data series indicated 7 minutes for all assets. With the estimated model, residuals were used to generate the marginal distributions and to obtain the VaR.

Volatility model OPT					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	87	94	105	116
	% of the total number	8.29%	8.95%	10.00%	11.05%
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	41	52	43	44
	% of the total number	3.90%	4.95%	4.10%	4.19%
	Kupiec (Acceptance region)	38 < N < 67			
1%	Number of violations	15	14	10	7
	% of the total number	1.43%	1.33%	0.95%	0.67%
	Kupiec (Acceptance region)	5 < N < 18			

Table 7 – Kupiec test for the OPT model

It is noticed from the above table that all models were accepted at the levels of significance investigated, indicating good fit of the optimal frequency model to data. Moreover, the Gumbel copula indicates good fit to extreme tails, as the number of violations at 1% was lower among the estimated families than verified at 10%. In the t copulas it is the inverse, as the number of violations proportionally increases in the interval as the confidence interval increases.

4.3.3. Estimation of copulas with HAR models using microstructure error correction with fixed frequencies

The series estimated using HAR with microstructure error correction are denoted by HLL^1, HLL^2 and HLL^5, for frequencies of 1, 2, and 5 minutes respectively. The procedure for copula estimation is the same as described in the previous sections, where what changes in the algorithm is the data entry matrix. In this case, it corresponds to the residuals of these estimated models.

The estimated copulas for HLL^1, HLL^2 and HLL^5 models exhibit the asymmetrical character of all the distributions: the value of the loss in modulus is larger than the gain value. Moreover, the HLL^2models indicate a greater tendency to more extreme values for Value at Risk. However, no statements can be made before the Kupiec test has been conducted, which is presented in table 8.

Volatility model HAR HLL^1					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	**117**	**134**	109	115
	% of the total number	**11.14%**	**12.76%**	10.38%	10.95%
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	**60**	**69**	62	57
	% of the total number	**5.71%**	**6.57%**	5.90%	5.43%
	Kupiec (Acceptance region)	38 < N < 67			
1%	Number of violations	**19**	**13**	15	9
	% of the total number	**1.81%**	**1.24%**	1.43%	0.86%
	Kupiec (Acceptance region)	5 < N < 18			
Volatility model HAR HLL^2					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	105	119	89	93
	% of the total number	10.00%	11.33%	8.48%	8.86%
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	52	61	48	53
	% of the total number	4.,95%	5.,81%	4.,57%	5.05%
	Kupiec (Acceptance region)	38 < N < 67			
1%	Number of violations	14	12	8	7
	% of the total number	1.,33%	1.14%	0.,76%	0.,67%
	Kupiec (Acceptance region)	5 < N < 18			
Volatility model HAR HLL^5					
	Copulas	t	Clayton	Frank	Gumbel
10%	Number of violations	97	112	85	89
	% of the total number	9.24%	10.,67%	8.10%	8.48%
	Kupiec (Acceptance region)	84 < N< 123			
5%	Number of violations	49	44	52	58
	% of the total number	4.67%	4.19%	4.95%	5.,52%
	Kupiec (Acceptance region)	38 < N < 67			
1%	Number of violation	15	14	11	9
	% of the total number	1.43%	1.33%	1.05%	0.86%

Kupiec (Acceptance region)	5 < N < 18

Table 8 – Kupiec test for the HLL^1, HLL^2 and HLL^5 models

The Kupiec test for microstructure correction models using fixed time intervals indicated rejection of the Value at Risk models for the Clayton and t copulas at all adopted levels of significance. For the remaining intervals, all the models are statistically significant. The result corroborates the test of the 1-minute HAR model without microstructure correction, which also presented non-acceptance of the null hypothesis.

5. Conclusion

The study of variance to determine the risk level is essential in the literature on finance, as it is an important variable for several theories in the fields of asset pricing and Value at Risk. In that sense, a significant number of studies can be observed nowadays whose goal is to provide a more effective estimate for the non-observable parameter. From the 1980s onwards, many volatility models have been developed with an aim of properly forecasting the risk level for the next period. This discussion includes the bank regulations promoted by the *Bank for International Settlements* with respect to allocation of minimum capital by banks, especially the adoption of internal Risk Management models.

In a scenario which combines research on effective volatility forecasting models with rigid bank supervision, this article presents the results of using realized volatility models, known as HAR, to generate tail dependence through the use of copulas. The sample of this study consisted of utilizing the four most net assets of the Ibovespa index from March 3, 2006 to April 30, 2010, thus minimizing the liquidity risk problem. The data generated was 1,026 daily data and 430,921 data for the highest 1-minute frequency.

In the first stage, the research consisted of estimating the volatility models, with the optimal frequency technique of Bandi e Russel (2005) applied to HAR. The result was an average 7-minute interval for the four analyzed assets. Another measure for correcting the microstructure problem is using the filter of Hansen, Large and Lunde (2006), with intervals of 1, 2 and 5 minutes. This number of combinations generated 44 volatility models for each data series, which were organized in accordance with the generated model for the four assets. Eleven possible combinations were generated for a portfolio composed of equally weighted assets. In a second stage, the copulas were estimated using these 11 combinations, each with the generation of four different dependence structures: t, Frank, Gumbel and Clayton copulas.

Once the copulas were generated, the next stage consisted of calculating the Value at Risk at the levels of 90%, 95% and 99%, and using the Kupiec test to validate the null hypothesis that the models are adequate. The results found are:

i) The 44 estimated volatility models exhibited good fit to data, as the residuals did not present heteroscedasticity;

ii) The comparison between empirical and estimated copulas for the 11 groups of models was satisfactory and presented white noises as residuals, indicating that the estimation was successful;

iii) The choice of generalized Pareto distribution for tail modeling proved to be satisfactory for the same reason above;

iv) The model for Value at Risk calculation for each group and estimated copula demonstrated that: the GARCH(1,1) model is excessively allocating for t and Frank copulas*;* the EWMA model was statistically significantly for the Gumbel copula only*;* the HAR model without microstructure error correction presented inadequate models at 1 and 2-minute frequencies; the HAR model with optimal frequency was significant for all estimated copulas; and finally, the HAR model with microstructure correction error for fixed intervals did not have good fit at the 1-minute frequency only.

The findings of this article point out that HAR models with copulas outperformed GARCH and EWMA models with time intervals longer than 5 minutes, with no influence from the microstructure error correction. Furthermore, the optimal frequency using copulas presented results that fit the null hypothesis. This possible contradiction between the results can be explained by assuming that the portfolio in this study is diversified, whereas the portfolio in the article was entirely composed of a single asset. This study utilizes temporal structures that allow correlation change overtime through copulas, whereas the first utilized linear covariance matrix.

The utilization of different tail dependence structures did not indicate superiority of one family over others. However, it was evidenced the acceptance of the null hypothesis that the model for Value at Risk calculation is acceptable for the great majority of the studied models.

As future research, this study can be extended to other assets with high liquidity, tail dependence structure analysis using different markets and time windows, and also models that use volatility dynamics other than those analyzed in this article.

References

Aït-Sahalia, Y., Mykland, P. A. and Zhang, L. (2009), Ultra high frequency volatility estimation with dependent microstructure noise. Journal of Econometrics, forthcoming.

Alexander, Carol (2001) "Modelos de Mercados". Editora Saraiva.

Andersen, T. and T. Bollerslev (1998), Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. International Economic Review, 39(4): 885- 906.

Andersen, T.G., T. Bollerslev, Diebold, F.X. and Eben, H. (2001), The distribution of realized stock return volatility, Journal of Financial Economics, 61, 43 – 76.

Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2003), Modeling and forecasting realized volatility, Econometrica, 71, 529 – 626.

Andrade, S.C.; Tabak, B. M. (2001), Is it Worth Tracking Dollar/Real Implied Volatility? Banco Central do Brasil, Working Papers Series, n. 15, p. 1-25.

Ané, T. and C. Kharoubi, 2003. Dependence structure and risk measure. *Journal of Business* 76, 411-438.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999): "Coherent measures of risk," Mathematical Finance, 9(3), 203–228.

Bandi, F.M. and Russell, J.R. (2006a), Volatility, in J.R. Birge and V. Linetsky (eds.), Handbook of Financial Engineering, Elsevier, to appear.

Bandi, F.M. and Russell, J.R. (2006b), Separating market microstructure noise from volatility, Journal of Financial Economics, 79, 655 – 692.

Bandi, F.M. and Russell, J.R. (2006c), Comment on Hansen and Lunde, Journal of Business and Economic Statistics, 24, 167 – 173.

Bandi, F.M. and Russell, J.R. (2006d), Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations, unpublished paper, Graduate School of Business, University of Chicago.

Barndorff-Nielsen, O. E., Hansen, P. R. Lunde, A. and Shephard, N. (2006a), Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise, unpublished manuscript, Stanford University.

Barndorff-Nielsen, O. E., Hansen, P. R. Lunde, A. and Shephard, N. (2006b), Subsampling realised kernels, unpublished manuscript, Stanford University.

Barndorff-Nielsen, O. E., Hansen, P. R. Lunde, A. and Shephard, N. (2008a), Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise, Econometrica, Econometric Society, vol. 76(6), pages 1481-1536, November.

Barndorff-Nielsen, O. E., Hansen, P. R. Lunde, A. and Shephard, N. (2008b), Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and nonsynchronous trading. Unpublished paper: Oxford-Man Institute, University of Oxford

Barndorff-Nielsen, O. and N. Shephard (2002), Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models. Journal of the Royal Statistical Society B, 64(2): 253-280.

Black, F. (1976), Studies in stock price volatility changes, Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, American Statistical Association, 177-181.

Bollerslev, T. (1986) "Generalized Autoregressive Conditional Heteroscedasticity", Journal of Econometrics, 31.

Bollerslev, T., Chou R. Y., and Kroner K. F. (1992), ARCH modeling in Finance, Journal of Econometrics, 52, 5-59.

Bordignon, S e Raggi, D. (2010), Long memory and nonlinearities in realized volatility: a Markov switching approach. Universita' Degli Studi di Bologna, Dipartimento di Scienze Economiche. Quaderni - Working Papers DSE N° 6942010, http://www2.dse.unibo.it/wp/694.pdf. (20/09/2010).

Campbell, J. Y., A. W. Lo, and A. MacKinlay (1997): The Econometrics of Financial Markets. Princeton University Press, Princeton.

Carvalho, M. R. C., Freire, M. A, Medeiros, M. C. and Souza, L. R. (2006), Modeling and Forecast the Volatility of Brazilian Asset Returns: A Realized Variance Approach. Textos para discussão 530, Department of Economics PUC-Rio (Brazil).

Chang, C-L., and McAleer, M. (2010), Aggregation, Heterogeneous Autoregression and Volatility of Daily International Tourist Arrivals and Exchange Rates. Department of Economics and Finance University Canterbury. Working paper. 2010 http://www.econ.canterbury.ac.nz/RePEc/cbt/econwp/1002.pdf (20/09/2010)

Cherubini, U., Luciano, E., 2001. Value-at-risk trade-off and capital allocation with copulas. *Economic Notes* 30(2), 2001.

Christoffersen, P. (1998), Evaluating Interval Forecasts. International Economic Review, Issue 39, pp. 841-862.

Corsi, F., (2004), A Simple Long Memory Model of Realized Volatility. Universidade de Lugano e Swiss Finance Institute. Working paper. 2004 http://papers.ssrn.com/sol3/papers.cfm?abstract_id=626064 (20 de setembro de 2010)

Corsi, F., (2009), A Simple approximate Long Memory Model of Realized Volatility ". Journal of Finance Econometrics, vol. 7, Issue 2, pp. 174-196, 2009.

Corsi, F., Pirino, D. and Reno, R. (2009), Volatility Forecasting: The Jumps Do Matter, Global COE Hi-Stat Discussion Paper Series gd08-036, Institute of Economic Research, Hitotsubashi University

Corsi, F. and Audrino, F. (2008), Realized Covariance Tick-by-Tick in Presence of Rounded Time Stamps and General Microstructure Effects, University of St. Gallen Department of Economics working paper series 2008 2008-04, Department of Economics, University of St. Gallen

Chung, H.; Huang, C-T.; and Tseng. C.-H. (2008), Modeling and Forecasting of Realized Volatility Based on High-Frequency Data: Evidence from Taiwan. International Research Journal of Finance and Economics. International Research Journal of Finance and Economics. ISSN 1450-2887. Issue 22 http://www.eurojournals.com/finance.htm

Deheuvels P., La fonction de dependance empirique et ses proprietes. Un test non parametrique d'independance, Acad. R. Belg. Bull. Cl. Sci., 5 Ser. 65 (1979) 274–292.

Deheuvels P., A Kolmogorov–Smirnov type test for independence and multivariate samples, Rev. Roum. Math. Pures et Appl. Tome XXVI, (2) (1981a) 213–226.

Deheuvels P., A nonparametric test of independence, Publications de l'ISUP 26, 1981b, pp. 29–50.

Dias, A., and Embrechts, P. (2003). Dynamic copula models for multivariate high frequency data in finance. Working Paper, ETH Zurich: Department of Mathematics.

Dhaene, J., and M. J. Goovaerts (1996): "Dependency of risks and stop-loss order," ASTIN Bulletin, 26(2), 201–212.

Diebold, Francis X (2007) "Elements of Forecasting", 4ª edição, Thomson South Western.

Ebens, H. (1999), Realized stock index volatility, Working Paper No. 420. Department of Economics, Johns Hopkins University, Baltimore.

Embrechts, P., M. Alexander, and D. Straumann (1999): Correlation and Dependence in Risk Management: Properties and Pitfalls. Working paper.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997): Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

Fama, Eugene (1970) "Efficient Capital Markets: A Review of Theory and Empirical Work", Journal of Finance, Maio de 1970, pág. 383-417.

Fang, K.-T., S. Kotz, and K.-W. Ng (1987): Symmetric Multivariate and Related Distributions. Chapman & Hall, London.

Fermanian, J. D. (2005): "Goodness-of-fit tests for copulas". Journal of Multivariate Analysis, 95, 119-152.

Fréchet, M. (1957): "Les tableaux de corrélation dont les marges sont données," Annales del'Université de Lyon, Sciences Mathématiques et Astronomie, Série A, 4, 13–31.

Galambos, J. (1987): The Asymptotic Theory of Extreme Order Statistics. Kreiger Publishing Co., Melbourne, FL.

Genest C., Ghoudi K., Rivest C., A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, Biometrika 82 (1995) 543–552.

Giot, P. and Laurent, S. 2004. Modelling daily Value-at-Risk using realized volatility and ARCH type models, Journal of Empirical Finance, Elsevier, vol. 11(3), pages 379-398, June.

Giot, P. and Laurent, S., 2001. Value-at-risk for long and short trading positions," CORE Discussion Papers 2001022, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE)

Hamerle, A., Rosch, D., Misspecified Copulas in Credit Risk Models: How Good is Gaussian?, Journal of Risk 8(1), 2005, 41-58.

Hansen, P.R., Large, J., and Lunde, A., (2006), Moving average-based estimators of integrated variance, Econometric Reviews, to appear.

Hansen, P.R. and Lunde, A., (2006), Realized Variance and Market Microstructure Noise. (with discussion). Journal of Business and Economic Statistics, 24(2): 127-218.

Hansen, P.R. and Lunde, A., (2001). "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?," Working Papers 2001-04, Brown University, Department of Economics.

Harris, L. (1990), Estimation of stock variance and serial covariance from discrete observations, Journal of Financial and Quantitative Analysis, 25, 291 – 306.

Höffding, W. (1940): “Massstabinvariante Korrelationstheorie,” Schriften des Mathematischen Seminars und des Instituts f¨ur Angewandte Mathematikder Universit¨at Berlin, 5, 181–233.

Hull, John (2008) “Options, Futures, and Other Derivatives”, 7a. Ed., Prentice Hall.

Hutchinson, T. P., and C. D. Lai (1990): Continuous Bivariate Distributions, Emphasizing Applications. Rumsby Scientific Publishing, Adelaide.

Joag-dev, K. (1984): “Measures of Dependence,” in Handbook of Statistics, ed. by P. R. Krishnaiah, vol. 4, pp. 79–88. North–Holland/Elsevier, New York.

Jorion, Philippe. *Value at Risk*- The New Benchmark for Managing Financial Risk 3a. Edition. McGraw-Hill. New York: 2007.

J.P. Morgan (1996) “RiskMetrics – Technical Document”, 4a. Ed., NY.

Kendall, M., and A. Stuart (1979): Handbookof Statistics. Griffin & Company, London.

Kruse, R. (2006), Can Realized Volatility Improve the Accuracy of Value-at-Risk Forecasts?.http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.172.8176&rep=rep1&type=pdf (10/04/2011)

McAleer, M. and Medeiros, M.C. (2008), A multiple regime smooth transition Heterogeneous Autoregressive model for long memory and asymmetries, Journal of Econometrics, Elsevier, vol. 147(1), pages 104-119, November.

Merton, R. C. (1980), On estimating the expected return on the market: An exploratory investigation, Journal of Financial Economics 8, 323-61.

Miller, D. J., Liu, W. H., 2006. Improved estimation of portfolio value-at-risk under copula models with mixed marginals. *Journal of Futures Markets* 26(10), 997-1018.

Newey, W.K. and West, K.D., (1987), A simple positive semi-definite, heteroskedasticityand autocorrelation consistent covariance matrix, Econometrica, 55, 703 – 708.

Owen, J., and Rabinovitch, R. 1983. "On the class of elliptical distributions and their applications to the theory of portfolio choice", Journal of Finance 38, 745-752.

Patton, A. (2003a). Modelling asymmetric exchange rate dependence. Working paper, University of California, San Diego.

Resnick, S. I. (1987): Extreme Values, Regular Variation and Point Processes. Springer, New York.

Ruppert, David (2006) “Statistics and Finance: An Introduction”, Springer.

Scharth, M. and Medeiros, M. C., (2009), Asymmetric effects and long memory in the volatility of Dow Jones stocks, International Journal of Forecasting, Elsevier, vol. 25(2), pages 304-327.

Schweizer, B., and A. Sklar (1983): Probabilistic Metric Spaces. North–Holland/Elsevier, New York.

Shi J., Louis T., Inferences on the association parameter in copula models for bivariate survival data, Biometrics 51 (1995) 1384–1399.

Sibuya, M. (1961): “Bivariate extreme statistics,” Annals of Mathematical Statistics, 11, 195–210.

Spearman, C. "The proof and measurement of association between two things" American Journal of Psychology, 15 (1904) pp. 72–101.

Wang, S. (1997): “Aggregation of correlated Risk Portfolios: Models and Algorithms,” Preprint Casualty Actuarial Society (CAS).