Volatility and Value-at-Risk Forecasting with Realized Volatility and HAR: A comparative Approach

1. Introduction

The development of a new technique for measuring and estimating asset volatility is a field of great interest in finance, especially concerning the asset pricing and risk management theory which is based on that statistical measure. To market makers and Banking Institutions in general, these new techniques should be more comprehensible and easier to measure, estimate and forecast. The current techniques, like conditional variance or stochastic volatility, are not simple to understand and estimation processes such as almost-maximum likelihood are complex and difficult to converge. To academics, it is necessary to develop a direct way of estimating volatility, which is a non-observed variable, by implementing models to capture all stylized facts of financial series.

In that sense, a new great wave in finance has been growing - the Realized Volatility (RV), which seeks to meet those needs. This technique is usually accompanied by econometric models denominated Heterogeneous Autoregressive (HAR), developed by Corsi (2004, 2009), aiming to model and forecast volatility. All this literature originates from the seminal article of Merton (1980), according to which it is possible to estimate an asset's latent or non-observed volatility over a given period using the sum of *n* intradaily squared returns, when *n* tends to infinity. However, this technique took many years to be employed due to a problem with the availability of ultra-high frequency data, i.e., 1-minute or more frequent samples. Further, these models could only be tested in reality by using supercomputers. Even nowadays, handling databases with many assets at this frequency requires a large data processing. Finally, and more importantly, Black (1976) developed a theory about the problem caused by high-frequency sampling, which was denominated as microstructure noise.

This problem partly prevented the theme to be further explored because the theory sustained that an asset's observable price consisted of the efficient, or real, price plus a random error. However, when the asset's squared return is calculated, the error no longer disappears when it is summed, systematically generating a bias in this estimate. Moreover, the higher the intraday return frequency, the larger the microstructure noise. Hence, the utilization of models for measuring the realized volatility depended on a way of solving those problems.

Therefore, in an attempt to solve such problems, some authors like Andersen and Bollerslev (1998) started to employ a sampling frequency that could statistically converge to the continuous function of latent volatility, but was not high enough to cause relevant bias due to the microstructure noise. However, some authors like Harris (1990), Zhou (1996) and Andersen and Bollserslev (1998), and later, employing simulation techniques, Zhang, Mykland and Aït-Sahalia (2005), demonstrate that ignoring this problem could lead to serious mensuration errors. Hence, it was necessary to develop techniques for estimating latent volatility without or with a minimum microstructure error.

In order to solve this problem, several techniques were found that successfully dealt with the microstructure noise - through the optimal choice of sampling frequency (Bandi and Russel (2006a) and Zhang, Mykland and Aït-Sahalia (2005); through filters based on the estimation of an AR(p) or MA(q) of the intraday data (Ebens (1999), Andersen, Bollerslev, Diebold and Ebens (2001) and Hansen, Large and Lunde (2006)); through the so-called Realized Kernel, which develops a HAC-type technique with Bartlett Kernel and covariance matrix of Newey and West (1987) to correct the microstructure noise (Barndorff-Nielsen, Hansen, Lunde, Shephard (2006a, 2006b, 2008a, 2008b)); and finally, through a technique that seeks to combine two different frequencies, with an aim to take advantage of each of them (Zhang, Mykland and Aït-Sahalia (2005) and Aït-Sahalia, Mykland, and Zhang (2006, 2009)).

As the microstructure noise started to be solved, a literature emerged with an aim at modeling and forecasting the Realized Volatility without the use of conditional variance, as well as capturing stylized facts of financial series that were not modeled before. This literature started with the work of Andersen, Bollerslev, Diebold and Labys (2003) and Corsi (2004, 2009), who presented a way of modeling and forecasting realized volatility from models estimated by Ordinary Least Squares (OLS). Moreover, these techniques demonstrated some properties that were not always present in models like GARCH *(*Generalized Autoregressive Conditional Heterocedastik), ARFIMA (Autoregressive Fractal Integrated Moving Avarage), FIGARCH *(*Fractal Integrated Generalized Autoregressive Conditional Heteroscedasticity*)* and Stochastic Volatility (SV), such as: long memory, low computational cost, multivariated extensions, quick responses to shocks in the short term and economic explanation for model designing.

Andersen, Bollerslev, Diebold and Labys (2003) and Bollerslev, Chou and Kroner (1992) find that GARCH and Stochastic Volatility models do not satisfy multivariate models, because the estimation is made by Almost-Maximum Likelihood or Kalman Filter, which is complex and of difficult convergence with many assets.

Moreover, Corsi (2004, 2009) points out that conditional variance models do not capture all the characteristics of financial series, such as quick responses to short-term shocks and long memory. He argues that since GARCH models (p,q) present high persistence, as identified by Bollerslev, Chou and Kroner (1992) and are often modeled by Integred GARCH (IGARCH), they have a p close to zero and a q close to 1, which causes the model to give a slow response to sudden changes. This happens because the parameter p, which models the conditional variance of the shock over the t-1 period, is small, and the parameter q, which models ARCH (p) of infinite order and has a longer dependence, is much more important in the determination of volatility. Thus, the short-term impacts take long to be assimilated by the models. If it were the opposite, i.e., a high p and a low q, the long-term effect would be once more neglected, as this technique imposes a strong trade-off on the short and long-term relationship and fails to adequately model these relations.

In an attempt to model the thusly neglected long memory relationship, Corsi (2004, 2009) utilizes traditional long memory models, such as the ARFIMA (p,d,q) and the FIGARCH (p,d,q); however those are also unable to fully perform their task. According to the author, the mathematical operation of the fractional difference operator may result in a loss of information and may not be able to capture fast changes in the long-term dynamics, which is often observed in financial data. Furthermore, he points out that operator precedence of the fractional difference operator (parameter d) together with the other parameters (p and q) is not trivial to estimate, making the estimation of these models often impossible. Additionally, this difficulty once more causes trouble in computational terms to multivariate extensions. Finally, estimating the parameter d separately in order to facilitate the convergence process may incur in bias and inefficiency of the estimators.

Thus, Corsi (2004, 2009) demonstrated that the Heterogeneous Autoregressive (HAR) model could correct those problems. Further, it could provide superior fitting and forecasting performance in relation to traditional models, because it would be easily implemented, would capture long memory and its parameters would adequately respond to short-term shocks. Later, a number of works found the same result for several assets - Chang and McAller (2010) for exchange rate; Scharth and Medeiros (2009) for stocks; Allen, McAleer, and Scharth (2009) and Jou, Wang, and Chiu (2010) for derivatives. Others extended the model of Corsi (2004, 2009) - Markovian regime switching in Bordignon and Raggi (2010); jumps modeling and leverage effects in Corsi, Pirino and Reno (2009), and in Chung, Huang and Tseng (2008); Multiple-Regime Smooth Transition HAR model, in Medeiros and McAleer (2008); and multivariate extensions in Audrino and Corsi (2008), confirming the best performance of these models.

With respect to Brazilian data, there are no works to properly test these techniques and no use of the HAR model. Some of the works are Andrade and Tabak (2001) for exchange rate; Carvalho, Freire, Medeiros and Souza (2005) for IBOVESPA and Sá Mota and Fernandes (2004) for IBOVESPA stocks. However, they either apply the realized volatility without correcting the microstructure noise or utilize old econometric techniques, such as EWMA (exponentially weighted moving average) and GARCH. No one utilizes the HAR models, failing to take advantage of what is best in using realized volatility in practical terms.

Given the circumstances above, the paper seeks to meet two goals. First, to analyze whether the HAR models are superior to traditional models in forecasting ability, at the same time observing which is the best method and sampling frequency to minimize the microstructure noise in the Brazilian data, as we have a more volatile market, with less liquidity and more restrict data availability than the North-American market. Further, our second goal is to analyze whether the HAR models succeed in the empirical application for Value-at-Risk (VaR) and whether they are superior to the GARCH and EWMA models.

Results indicate that the HAR is superior to the GARCH and EWMA models in forecasting ability, especially at 2, 5 and 10 steps ahead. Besides, the correction method of Hansen, Large and Lunde (2008) did not fit the Brazilian data. Curiously, the 1-minute frequency which was the highest employed, produced the best models of forecasting and fitting to VaR.

In the Value-at-Risk application, the HAR models did not demonstrate superiority to the GARCH model. The realized volatility-based model performed well in three of the four stocks, as did the GARCH, especially at short forecasting horizons (1 and 2 steps ahead), being fit at all maximum loss levels for the next day. Furthermore, the model had an excellent performance for GGBR4, passing all tests and at all forecasting horizons. However, it failed to model PETR4 and was outperformed by the GARCH model for VALE5. We emphasize that the GARCH model did not fit the GGBR4 series and had a draw in performance when compared to the USIM5 series; also, it was barely approved in VaR configurations with a maximum loss of 10% and 5%. Hence, we believe that the models are complementary to each other, with none demonstrating a significant superiority. The EWMA showed problems with the criterion de independence of violations, being rejected in a large part of the models estimated.

For our purpose, this paper is organized in four more parts, as follows. In the methodology, we introduce the microstructure noise and the options for correction, as well as the HAR model and the traditional estimated models. The third section is dedicated to the treatment of the database and the fourth provides the findings, where we detail the results of forecasting and application to VaR. Finally, we proceed to the conclusion of the paper.

2. Methodology

In this section, we will formally introduce the methodology employed in the paper, dividing it into five parts. In the first subsection, we will present the theoretical construction of realized volatility and its correction methods. In the second, we will introduce the econometric models HAR, GARCH and EWMA. In the third, we will analyze the criteria developed to evaluate the models' forecasting performance. In the fourth, we will show the use of Value-at-Risk and the empirical validation tests of the technique. Finally, in the last subsection we will show the empirical method employed by us that makes the HAR model's performance much superior to traditional methods.

2.1 Realized Volatility

Merton (1980) showed that it would be possible to create a proxy of latent volatility using the sum of N intraday squared returns over a given time period t. It would be possible, as when N tends to infinite there is a convergence in probability to the continuous function of integrated volatility. In other words, by collecting an asset's price on one day N+1 times, with N being frequent enough, applying the logarithmic return, squaring and then summing all returns, we would reach the latent or non-observed volatility for that day. Formally, considering that an asset's price follows a diffusion process:

$$dp_{t+\tau} = \mu_{t+\tau} d\tau + \sigma_{t+\tau} dW_{t+\tau}, \qquad 0 < \tau < 1\,,\ t = 1,2,\dots,T \qquad (1)$$

where p_t is the logarithm of instantaneous price on time $t+\tau$, $\mu_{t+\tau}$ is the *drift* component (equal to zero in this case), $dW_{t+\tau}$ is a standard Brownian motion and $\sigma_{t+\tau}$ is a standard deviation. Thus, it is demonstrated that the instantaneous volatility from t-1 to t is the integral of the standard deviation of the Brownian motion, as follows:

$$\sigma_t = IV = \left(\int_{t-1}^{t} \sigma^2(\omega)d\omega\right)^{\frac{1}{2}} \tag{2}$$

However, this variable is not directly observable and the data collection is discreet. Thus, Merton (1980) and Andersen, Bollerslev, Diebold and Labys (2001) point out that:

$$RV_t = \sqrt{\sum_{n=1}^{N} r_{t+n}^2} \tag{3}$$

will be an approximate measure of integrated volatility when:

$$plim_{N\to\infty} \to RV_t = \sigma_t \tag{4}$$

2.1.1 Microstructure Noise

As presented in the introduction, the implementation of the Realized Volatility aims at correcting the old microstructure noise, found by Black (1986). In this theory there is a price $P_{t,n}^0$, which is the observable price on day t and on the n-th division, comprised of $p_{t,n+1}^0 = p_{t,n+1}^L + \varepsilon_{t,n+1}$, where $p_{t,n}^L$ is the latent price and $\varepsilon_{t,n}$ is an IID disturbance with $E[\varepsilon_t] = 0$, $E|\varepsilon_t|^4 < \infty$ and not correlated with the latent (or efficient) price. If we follow Merton's (1980) technique and sum the N returns to the square derived from the N+1 partitions of the day, with N being sufficiently large, our estimate will be biased because the error term will accumulate. In other words, when we take the first difference of the logarithm of $P_{t,n}^0$ and square it, we will accumulate the N $\varepsilon_{t,n}$ in the realized volatility. Therefore, the problem is to create a sample that is sufficiently frequent to converge to continuous function, but not so frequent as to incur in a large bias of the microstructure noise. It should be observed that the higher the partition of day t, the larger the microstructure bias. Formally, the logarithm of the observable price is

$$p_{t,n}^0 = p_{t,n}^L + \varepsilon_{t,n} \tag{5}$$

Taking the first difference of (5) and defining $r_{t,n+1}$ as the return, we have:

$$p_{t,n+1}^o - p_{t,n}^0 = p_{t,n+1}^L + \varepsilon_{t,n+1} - p_{t,n}^L - \varepsilon_{t,n} \to r_{t,n+1}^o = r_{t,n+1}^L + \varepsilon_{t,n+1} - \varepsilon_{t,n} \tag{6}$$

But

$$v_{t,n+1} = \varepsilon_{t,n+1} - \varepsilon_{t,n}, \quad r_{t,n+1}^o = r_{t,n+1}^L + v_{t,n+1} \tag{7}$$

By squaring it, we have: $\left(r_{t,n+1}^o\right)^2 = \left(r_{t,n+1}^L\right)^2 + \left(v_{t,n+1}\right)^2 + 2\left(r_{t,n+1}^L v_{t,n+1}\right)$ (8)

Summing the N returns, using definition of (3) and $\sum_{n=1}^{N}\left(v_{t,n+1}\right)^2 = \epsilon_t^2$

$$(RV_t^o)^2 = (RV_t^L)^2 + \left(\epsilon_{t,}\right)^2 + 2\sum_{n=1}^{N}\left(r_{t,n+1}^L v_{t,n+1}\right) \tag{9}$$

Assuming that the microstructure problem is IID, with $\mathrm{E}v_{t,n} = 0$, as $p_{t,n}^0$ is not stochastically correlated with $v_{t,n}$ and the estimator variance is non-infinite, $E\left|v_{t,n}\right|^4 < \infty$ and taking into consideration that Var $(v_{t,n})$=$E\left(v_{t,n}\right)^2$, we have:

$$E[(RV_t^o)^2|(RV_t^L)^2] = (RV_t^L)^2 + 2n(\epsilon_t)^2 \tag{10}$$

demonstrating to be clearly a biased sampling process. Following the approach of Zhang, Mykland and Aït-Sahalia (2005), which confirms the derivations below, the conditional variance becomes:

$$var[(RV_t^o)^2|(RV_t^L)^2] = 4nE(\epsilon_t)^4 + (8RV_t^L E(\epsilon_t)^2 - 2var\,(\epsilon_t)^2) + O_p(n^{-\frac{1}{2}}) \quad (11)$$

Therefore, putting N to infinite:

$$n^{-\frac{1}{2}}((RV_t^o)^2 - 2n((\epsilon_t)^2) \xrightarrow{\mathcal{L}} 2(E(\epsilon_t)^4)^{-\frac{1}{2}} Z_{Noise} \quad (12)$$

Z_{Noise} has a normal distribution and originates from the disturbance in RV_t^o. Moreover, the authors find that in addition to the mean and the variance being affected by the microstructure noise, there is also the discretization problem, i.e., the process is not effectively continuous in practice. Hence, they indicate the existence of convergence in distribution to Observed Realized Volatility, as:

$$RV_t^o \overset{d}{\approx} \underbrace{IV + 2N_t E(\epsilon_t^2)}_{\text{Bias due to noise}} + \underbrace{\left[\underbrace{4N_t E(\epsilon_t^4)}_{\text{due to noise}} + \underbrace{\frac{2}{N_t}\int_0^1 \sigma_t{}^4 dt}_{\text{due to discretization}}\right]^{\frac{1}{2}}}_{\text{Total Variance}} N(0,1) \quad (13)$$

2.1.2 Correction methods

In this article two correction methods are applied to ensure that the microstructure noise is not relevant to our forecasts with the HAR model and to our empirical value-at-risk applications. Hence, it is used the procedure of finding the optimal frequency for each stock, which will be presented in the next subsection, and also a correction system for the microstructure noise, developed by Hansen, Large and Lunde (2008).

2.1.2.1 Optimal sampling frequency

The most used correction method is the optimal frequency, whose methodology follows the article of Hansen and Lunde (2006), and Zhang, Mykland and Aït-Sahalia (2005). Both articles demonstrate that this correction method is highly efficient in dealing with the microstructure noise. In this paper, we derive the optimal sampling system of Bandi and Russel (2005a, 2006), which works an approximation of the formula by arbitrating the optimal frequency of the estimator's variance.

Bandi and Russel (2005a, 2006) derive and minimize the function of error caused by microstructure noise, in order to ensure the convergence to a continuous function of integrated volatility. Hence, the mean quadratic error function is given by:

$$MSE_t = 2\left(\frac{1}{N}\right)\left(IQ_t + o_{a.s.}(1)\right) + N\left(2E(v_t^4) - 3\left(E(v_t^2)\right)^2\right) + N^2\left(E(v_t^2)\right)^2 + 4E(v_t^2)\left(\int_{t-1}^t \sigma_\tau^4 d\tau\right) - E(v_t^4) + 2\left(E(v^2)\right)^2 \quad (14)$$

with T being the total number of days and IQ_tbeing called the *Integrated Quarticity*, theoretically defined by a diffusion process presented in (1) [1].Minimizing the MSE_t of equation (14) we have:

$$N^* = 2N^3\left(E(v^2)\right)^2 + N^2(2E(v^4) - 3\left(E(v^2)\right)^2 - 2IQ_t \quad (15)$$

The authors derive that $E(v^2)$ is equal to $\frac{(\sum_{t=1}^T \sum_{n=1}^N r_{t+n}^2)}{TN}$ and $E(v^4)$ is equal to $\frac{(\sum_{t=1}^T \sum_{n=1}^N r_{t+n}^4)}{TN}$. Thus, they define the approximation of optimal N:

$$N^* \sim \left(\frac{IQ_t}{\left(E(v^2)\right)^2}\right)^{\frac{1}{3}} = \left(\frac{IQ_t}{\left(\frac{(\sum_{t=1}^T \sum_{n=1}^N r_{t+n}^2)}{TN}\right)^2}\right)^{\frac{1}{3}} \quad (16)$$

[1] $IQ_t = \int_0^1 \sigma^4(t+\tau-1)d\tau$

However, IQ_t must be estimated and Bandi and Russel define equation (16) for a 15-minute sampling frequency that would be sufficiently fast to approximate, given the empirical experiments, without incurring in microstructure error. In statistical tests, the IQ is derived with 15 and 30-minute frequency, being that the result obtained for all series investigated is about 7 minutes.

$$IQ_t = \frac{N_t}{3} \sum_{n=0}^{N_t} r_t^4 \qquad (17)$$

Using this methodology, we find a way to significantly minimize the microstructure noise, like Zhang, Mykland and Aït-Sahalia (2005) point out by simulation. Moreover, Hansen and Lunde (2006) demonstrate that the microstructure error for DJIA stocks is small in sampling frequencies lower than 20 minutes, which indicates that this technique would be sufficient to implement the models like Andersen (2007) addresses in his article.

However, the discussion about how to optimally calculate the *Integrated Quarticity* leads many authors such as Bandi and Russel (2005a, 2006b) to calculate it with a 15-minute sampling frequency, thus avoiding the use of a complex method demonstrated in Zhang, Mykland and Aït-Sahalia (2005). Hence, given the optimal choice limitations, the option is to sample in several frequencies (1, 2, 5, 15 and 30 minutes) in order to solve this problem, although we calculate the estimators' variance as demonstrated above pointing out the optimal choice for each stock.

2.1.2.2 Filter-based estimator

The filter-based estimator was introduced by Ebens (1999) and Andersen, Bollerslev, Diebold and Ebens (2001). The general idea was to estimate an auto-regressive model or moving average from the intraday return, because that process of autocorrelation and partial autocorrelation derived exclusively from a process generated by the microstructure noise. Hence, when estimating an AR or MA model it is possible to identify the part of the intraday return that is a microstructure bias, filtering it through the estimated parameters. However, Bandi and Russel (2005) criticize the model and demonstrate that the technique is not sufficient to stop the tendency. Hence, Hansen, Large and Lunde (2006) demonstrate that it is necessary a larger time lag relative to MA to ensure consistency.

Therefore, considering that the error is correlated to the latent price and assuming serial independence, the price return follows the MA(q) process: $r_{t,N} = u_{t,i} - \theta_1 u_{t,i-1} - \theta_2 u_{t,i-2} - \cdots - \theta_q u_{t,i-q},\ for\ any \quad i = 0, \ldots, N$. Where the sequence $\{u_{t,m}\}$ is *IID*$(0, \sigma^2_{u,t})$. Thus, the filter becomes:

$$RV^{MA} = \left(\frac{(1-\theta_1-\theta_2-\cdots-\theta_q)^2}{1+\theta_1^2+\theta_2^2+\cdots+\theta_q^2} \right) RV \qquad (18)$$

However, as the availability of data with frequency higher than 1 minute is too low for assets series traded at Bovespa, the authors demonstrate that the estimator's consistency problem diminishes with an equidistant interval and considering the volatility to be constant in this interval. In this article, the 20/20 second interval is used, which ensures a variance approximately constant in the subinterval. On the other side, less frequent intervals are allowed in a market not so liquid like NYSE and S&P500, enabling the utilization of the data provided.

2.2. Heterogeneous Autoregressive (HAR)

The model of Corsi (2004, 2009), denominated Heterogeneous Autoregressive (HAR), is based on the assumption that markets are heterogeneous as presented by Muller, Dacorogna, Dav, Pictet, Olsen, and Ward (1993). This theory argues that different agents operate in the market, with goals, institutional restrictions, performance horizons, information, knowledge, and other variables of their own, which causes each change in the market to have a different response from these agents. This theory attempts to explain why the assets most searched for have the higher volatility, since if the markets were homogeneous the oscillations out of the efficient (or real) price would be more quickly corrected and less oscillations would occur.

In practical terms, this theory proves that while changes in the long term affect the short-term strategies, changes in the latter do not affect the former. This occurs because, according to Müller, Dacorogna, Dav and Pictet (1993), the change in the type of agents is associated with three performance horizons: short, medium and long term. The short-term group would include the brokers aiming to profit with day or intraday operations; the medium-term group would include the operators of high-risk funds; and the long-term group would include the central banks, commercial banks and pension funds. Hence, when brokers modify their operations by changes in the short term, it does not necessarily follow that the others will change their strategies. However, when the Central Banks and the pension funds modify their positions, they end up affecting the dynamics of the short-term agents. Thus, the estimated models is

$$\log\left(RV_t^d\right) = C + \theta_\mathrm{d}\log\left(RV_{t-1}^d\right) + \theta_\mathrm{w}\log(RV_{t-1}^w) + \theta_\mathrm{m}\log(RV_{t-1}^m) + \widetilde{\omega_{t+1}^d} \qquad (19)$$

The expression above is estimated by OLS and uses the Newey-West covariance correction for serial correlation. To forecast from (23), we estimated a model for each forecasting horizon h, as follows:

$$\log(RV_t^d) = C + \theta_\mathrm{d}\log\left(RV_{t-h}^d\right) + \theta_\mathrm{w}\log(RV_{t-h}^w) + \theta_\mathrm{m}\log(RV_{t-h}^m) + \widetilde{\omega_t^d} \qquad (20)$$

The models used for comparison to the HAR's performance are the Exponential Weighted Moving Average (EWMA), with $\lambda = 0{,}94$, which is the model most used in the market for value-at-risk forecasting, and the traditional GARCH model (1,1)

2.2.1 Theoretical VaR approach

The standard methodology we select a standard normal distribution and the quantile to a given probability "α" that we wish. As the distribution has a deviation equal to 1, we only have to multiply the estimated standard deviation (assuming that the mean is equal to ZERO) by the value of the cumulative probability distribution to find the maximum variation for the period. Formally, the maximum loss for h-th days ahead at a probability α is:

$$VaR(\alpha, h) = Asset \times (x_\alpha \times \sigma_{t+h}), \qquad (21)$$

where x_α is the p-th quantile of a standard normal cumulative distribution and σ_{t+h} is the standard deviation estimated by the GARCH, EWMA or HAR models. We emphasize that for backtesting purposes our asset is equal to 1 (monetary value of the portfolio) and our volatility is always multiplied by -1, so as the $VaR(\alpha, h)$ is violated when $r_t < VaR(\alpha, h)$. For evaluations fo the VaR used Kupiec Test (traditional) and Christoffersen Test.

2.2.1.2 Christoffersen Test

A problem proposed by the Kupiec Test is that we might be correlating the violations of the model such that the percentage found, although similar to the desired, may incur in more errors over a given evaluation period. Therefore, based on Markov chains, Christoffersen (1998) verifies whether the period t-1 is correlated with t.

Additionally, we define $n_{i,j}$ as the number of observations in state j after having been in state i the previous day.

	$I_{t-1} = 0$	$I_t = 1$
$I_{t-1} = 0$	n_{00}	n_{01}
$I_t = 1$	n_{10}	n_{11}

Now, defining φ_i as the probability of a violation occurring given that i occurred in the previous day, we have: $\varphi_0 = \frac{n_{01}}{n_{00}+n_{01}}$, $\varphi_1 = \frac{n_{11}}{n_{10}+n_{11}}$ e $\varphi = \frac{n_{01}+n_{11}}{n_{00}+n_{01}n_{10}+n_{11}}$

With these probabilities, the test checks whether statistically $\varphi_0 = \varphi_1$, then verifying whether the probability of failure is equal to that of non failure. In other words, failing today does not increase the probability of failing tomorrow again. Thus, Christoffersen (1998) uses the LM test to check whether there is difference:

$$TC = -2\ln\left(\frac{(1-\varphi)^{n_{00}+n_{01}}\varphi^{n_{10}+n_{11}}}{(1-\varphi_0)^{n_{00}}\varphi_0^{n_{01}}(1-\varphi_1)^{n_{10}}\varphi_1^{n_{11}}}\right) \sim \chi^2(1) \quad (22)$$

2.4 Forecasting methods

The forecasting performance of the models is tested from the traditional evaluation on the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE), the Mean Percentage Absolute Error (MPAE) and by the Mincer-Zarnowits test. However, these criteria are constructed on the basis of the h-step-ahead forecast error of volatility[2]. Also, to generate the EWMA and GARCH models we do not have an endogenous variable for comparison. Thus, we use as volatility in our main evaluation the Realized Volatility (RV), as commonly used in the literature; see Hansen and Lunde (2001), Medeiros and McAleer (2008) and Corsi (2004). However, as this variable is the dependent variable of the HAR models we believe that this represents a natural advantage to these models. Thus, the h-step-ahead forecast error of the GARCH models is shown against a proxy of volatility, normally using the root of the squared return.

3. Data

This paper used four stocks traded at BOVESPA, which have the trading codes PETR4 (Petrobrás), USIM5 (Usiminas), GGBR4 (Gerdau) and VALE5 (Vale do Rio Doce), covering the period from 03/03/2006 to 30/04/2010 from the database provided by the Instituto Educacional BM&FBovespa. The assets were chosen based on data availability and liquidity, since this is an important measure for Realized Volatility models due to the microstructure noise. For EWMA and GARCH, we used the closing daily prices, while for Realized Volatility we used data from 10am to 5pm for trading sessions conducted in normal time, and from 11am to 6pm for those in daylight-saving time[3]. We emphasize that the extraction method is important to error correction mechanisms, and for that reason we explain that this is a time-based method. For such, our algorithm searched data on a predetermined day and time, considering as valid the price closest to that time. This is important, as the database provided is not regular with respect to trading, which causes the number of daily samples to significantly change; thus there is not a calendar-based sample model, but a hybrid system between trading and calendar. Finally, we call attention to the fact that Stock Split adjustments were made. The table below presents the number of observations by asset (All table are annexes section):

4. Results

[2] The HAR model generates a result to the standard deviation σ and the GARCH models generate the variance σ^2. In order to compare, we had to take the square root of the squared return and the 1-step ahead predictions of the GARCH models. We did the opposite too and the results did not change.

[3] The Bovespa changes the trading time according to the daylight-saving time, but not necessarily the beginning and end of this time will coincide with the time change in the trading session.

In this paper, the realized volatility was estimated at several sampling frequencies considered high and capable of reducing the microstructure noise, as presented in the methodology section. For such, our sample consists of frequency series of 1, 2, 5, 15 and 30 minutes without correction of the microstructure noise; they are denoted $ALL^1, ALL^2, ALL^5, ALL^{15}$ and ALL^{30}, respectively.

Moreover, according to equation (18), we derive the optimal frequency for our data series in agreement with Bandi and Russel (2005), who estimate the *Integrated Quarticity* (IQ) every 15 minutes. The authors point out that using this frequency even with higher or lower frequencies, but still incurring in little microstructure noise, would be sufficient and would not significantly modify the sampling choice. Hence, since we have a different and less liquid market, we also estimate the IQ every 30 minutes, just for checking. We conclude that the optimal frequency is similar for all assets and the different optimal frequencies are about 7 minutes, with the 30-minute IQ decreasing approximately 1 minute at the optimal frequency. The Realized Volatility estimated by this method was denominated OPT. The table below presents the results.

Finally, we estimated the realized volatility with the microstructure correction method of Hansen, Large and Lunde (2006), as presented in the methodology section, at three sampling frequencies: 1, 2 and 5 minutes, denoted HLL^1, HLL^2 and HLL^5 respectively.

4.2 GARCH Estimation

The GARCH models (1.1) were estimated via maximum likelihood, using the Eviews 5 software, with Marquardt optimization algorithms configured with 500 maximum iterations and convergence 0.001. Moreover, the series presented ARCH effect at least until lag 10, and the estimation residuals became white noises after the estimation of the models. This produced coefficients of significant parameters. Finally, we took the natural logarithm of volatility estimated by the GARCH and EWMA for comparisons with HAR. At this point we observe, the GARCH models estimates are very similar to the Realized Volatility, only with a little more Kurtosis and Asymmetry.

4.3. HAR estimation and comparison with the literature

The HAR estimated with Brazilian data demonstrated to be as fit or better than that found by the literature for North-American data. Our estimated parameters were significant in all models, proving that the structure proposed by the HAR which is using the mean of the realized volatility for the last 5 and 22 days is coherent in our data. Another important point is that the adjusted R^2 found slightly drops as the lags of the variables increase. For instance, in lag 10 the models still have a considerable explanatory level, about 0.35. (Table with authors)

4.4 Prediction of the GARCH and EWMA models

The volatility forecasting models GARCH and EWMA, constructed with conditional variance, do not have a reference like the OLS estimation models which have their own endogenous variable to measure their performance. Therefore, it has always been very complicated to accurately know if these models are fit. In the article of Hansen (2001), it is applied a realized volatility model which would be the theoretically correct measure for forecast comparison of conditional variance models, becoming a reference to evaluate these models. In this article we used two measures of realized volatility - correction filter of Hansen, Large and Lunde (2006) with 1-minute-frequency; and the optimal frequency choice of Bandi and Russel (2005). Further, it was used the squared return reference and the GARCH and EWMA's one-step-ahead forecast to demonstrate the loss of precision as the number of steps ahead (PAF) increase.

As we can see in Tables 15, 16 and 17, the GARCH and EWMA models have a very similar performance, with the EWMA showing slightly lower RMSE, MAE and MPAE. Varying according to the asset and using the realized volatility as reference, the GARCH model has a one-step-ahead RMSE between 0.09 and 0.17, while the EWMA has it between 0.06 and 0.09. Moreover, the MPAE of the GARCH model is between 5.8% and 7.4%, with the EWMA a little below, between 4.53% and 5.43%.

Further, when we extend the forecast to between 2 and 10 steps, the EWMA model's error increases less than the GARCH model's error, as in the 10-step-ahead forecast compared to its own 1-step-ahead forecast the EWMA always has a percentage result lower or equal to that of the GARCH. With respect to the stocks, no one presented a very different behavior; only the VALE5 series showed the RMSE, MAE and MPAE a little higher at all horizons and references and in both models compared.

The Mincer-Zarnovitz tests against the realized volatility of HLL demonstrate that the deterministic volatility models had an explanatory level close to 50% of the realized volatility value with the 1-step-ahead forecast. However, neither model's forecast is able to explain the endogenous variable, which indicates that even showing relatively little error they fail to explain the variations that actually occur at long horizons. In practical terms, this means that the model gets very close to what actually happens because it presents few errors, however without being accurate enough to explain all the volatility variations in each time t. In this criterion the GARCH model showed the best results.

In general, the models performed well with low mean absolute percentage error at long horizons and presented good adaptation in the Mincer-Zarnowitz tests. Naturally, the best fit models showed the best forecasting ability - the HAR models with RV HLL, with correction filter; and with RV ALL, without microstructure error correction, both with 1-minute frequency. However, it can be observed that the fit and forecasting ability of the models drop as the sampling frequency decrease, independently from any correction methods. That is perhaps the reason why some authors that work with the methodology of *Jumps* jointly tend to ignore the microstructure error and directly apply the realized volatility. It is interesting to note that the microstructure error should increase as the frequency increases, causing the performance of these models to fall since there would be a "non-modeled" disturbance component inside the estimates. This could indicate that the error in Brazilian models would be small and irrelevant to the estimation of models like the HAR or that the correction methods used have non-feasible hypotheses with regard to local data.

The 1-minute frequency models reached the 1-step-ahead forecast: 60% in the Mincer-Zarnowitz test with mean squared error of forecast of 0.25 and percentage mean absolute error of 4.6%. However, the advantages over the others decrease as the horizon increases, reaching approximately 37% in the Mincer-Zarnowitz test at the 10-step-ahead horizon. It is also interesting that the mean percentage error increases only 1.2 percentage points for all models as the horizon increases, indicating high forecasting ability of this technique.

The performance comparison between GARCH, EWMA and HAR is based on the realized volatility measure, which is the method used by Hansen and Lunde (2001) and adopted by the literature as the main way of comparing the forecasting performance of these models. Otherwise, the only aspect we can evaluate is that both models provide consistent forecasts at larger forecasting horizons such as 5 and 10 steps. The HAR increases less the forecast error relative to RV than the EWMA and GARCH do in relation to their own 1-step-ahead forecast. Hence, using the realized volatility as reference, it can be observed that the mean squared error of the 1-step-ahead forecasts of the EWMA and the GARCH is higher than that of the HAR: 0.31 to 0.42 and 0.30 to 0.36 against 0.25 to 0.39, respectively, varying according to the chosen asset. Furthermore, the MPAE also presents higher values: 4.32% to 5.97% of the HAR against 6.42% to 7.46% and 5.8% to 8.2% of the EWMA and GARCH, respectively. With respect to the Mincer-Zarnowitz test, the HAR obtained 0.65 to 0.44, while the EWMA obtained 0.55 to 0.35 and 0.59 to 0.40, demonstrating a higher ability of the HAR models.

When the criterion for comparison is the forecast more than one step ahead, the EWMA and GARCH increase the root mean squared error between 0.13 and 0.16 in relation to their own 1-step-ahead forecast, while the HAR increases 0.05 and 0.08, demonstrating that the loss of precision occurs more often in the conditional variance models. In addition to that aspect, the other tests show that the RMSE, MAE, MPSE and the Mincer-Zarnowitz test for the HAR are superior to those of the EWMA and GARCH. Thus, since the differences were not significant, we believe that the HAR models' ability is slightly superior to that of the conditional variance models. The table below summarizes the results:

4.7 Value-at-Risk

Overall, the HAR model performed well in the value-at-risk for GGBR4, USIM5 and VALE5 only, since the result for PETR4 was poor. In the case of GGBR4 and USIM5, the HAR model constructed without microstructure noise correction with 1- minute frequency was approved in the Kupiec Test at all forecasting horizons and fitted the maximum loss of 10%, 5%, 2,5% and 1%[4]. In the Christoffersen Test, we perceived that the errors are independent, i.e., the fact that an error occurred on time t does not mean that the probability of another violation increases on t+1. However, we emphasize that in the independence criterion, the USIM5 series provided forecasts adequate to the VaR at short horizons, such as 1 and 2 steps ahead. Table 24 presents the data:

In the case of the VALE5 series, the low volatility presented by the model caused the number of violations of the Value-at-Risk to slightly increase, as we can see in Table 25. In this case, since the model is estimated by OLS, the forecast directly generated by the equation is the most likely source of error, but it presents standard deviations that end up constructing an interval of possibilities, thus the most distant from the center, the less likely it becomes. Considering that, we added a standard deviation to the HAR model's forecast for the VALE5 series. Thus, the model performed satisfactorily in the Kupiec Test for 5%, 2.5% and 1% of maximum loss and at all horizons. The model did not fit in 10%. In the Christoffersen Test, the model also presented good results and passed all the tests, being rejected at the 10-step-ahead horizon only. In other words, the HAR model fully performed the activity for more accurate VaRs (5%, 2.5% and 1%) at 1- to 5-step-ahead horizons at least.

The HAR models did not perform well for the PETR4 series estimation. Only the HAR model based on the 1-minute RV HLL performed reasonably in the Kupiec Test, and even so, only at 1- to 5-step-ahead horizons. In addition, it did not fit all maximum loss percentages of the test; further, the model performed poorly in the Christoffersen independence test, demonstrating that it presents errors in a correlated way. Ultimately, it can be said that the model presented errors in a correlated way in the past, which means that the maximum loss for more than one day ahead may be higher than suggested by the percentage of violations estimated. Attached are all the tables with the value-at-risk estimates for all the HAR models estimated, with the forecast center and a standard deviation.

The results of the GARCH and EWMA models indicate that none of the models works properly for all the stocks used in the paper. Moreover, the main problem is that these models fail the independence test of Christoffersen, demonstrating that they violate the value-at-risk in a non-random way.

The EWMA model was approved in the Kupiec Test in 10%, 5% and 2.5% at all horizons for the GGBR4, USIM5 and VALE5 stock series. Further, in 1% it was only approved for USIM5. However, the Christoffersen test accepts GGBR4 and VALE5 only, rejecting almost all the USIM5 forecasts. In short, the EWMA model worked at all horizons in 10%, 5% and 2.5% for GGBR4 and VALE5.

On the other hand, the GARCH models were approved in the Kupiec Test at all forecasting horizons with 10%, 5%, 2.5% and 1% of maximum loss for the next period for PETR4, USIM5 and VALE5. However, when the maximum loss for the next period is 10% for the PETR4 price series, and 10% and 5% for the USIM5 price series, the Christoffersen test rejects the independence of the model's violation, putting into doubt the model's reliability with this configuration. In other words, the GARCH models showed superiority to the EWMA models, since they can provide safe forecasts for PETR4, USIM5 and VALE5 in maximum losses for the next period in the more accurate 2.5% and 1% systems.

In our assessment, we concluded that none of the models showed superiority to the others, with each of them presenting advantages and disadvantages. The GARCH models succeeded with three stocks when the loss estimated for the next period was 1% and 2.5% for all forecasting horizons. On the other hand, for three stocks the HAR models performed better at short forecasting horizons - 1 and 2 steps ahead - for all the maximum loss percentages for the next day. For comparison purposes, we can only point out that none of the models was superior; however, the HAR could not model the value-at-risk of PETR4 and was inferior to the GARCH for VALE5. In turn, the GARCH could not model the value-at-risk of GGBR4, had an equivalent performance for USIM5 and was superior for VALE5. In that case, we may conclude that the models are equivalent, with a small advantage of the GARCH.

[4] We want to remind that the reference value of the Kupiec Test and the Christoffersen Test is 3.87 to 5% of confidence.

Comparing our results with the international literature, it is possible to highlight that the HAR models actually provide better forecasts, as pointed out by almost all the literature on HAR. However, as noted by Giot and Laurent (2001, 2003), still using realized volatility modeled with ARFIMA, they do not outperform better the RV models. Further, Kruse (2006) tests the realized volatility and does not prove that these models are statistically superior to the GARCHs. In that sense, we believe that further studies are required to analyze these models, as we know that comparisons can always be criticized due to the reference issue. However they should be investigated, as the RV models have all the theory on their side, and the GARCHs have all their history and successful empirical applications.

The difference between the group of stocks that succeeded and the one that did not may have two explanations: (i) the HAR specification should be improved to find models with better fit levels, like the Thresold-HAR of Medeiros and McAller (2008); and (ii) it may be due to the microstructure noise that the hypotheses of the model of Hansen, Large and Lunde (2006) did not fit the VaR in Brazilian data. Since the market liquidity is higher for PETR4 and VALE5 series, perhaps they require correction methods more accurate if compared to the GGBR4 and USIM5 series, which showed good results. Models with Jumps - Andersen (2007), Zhang, Mykland and Sahalia (2005), or Barndorff-Nielsen, Hansen, Lunde, Shephard (2006a, 2006b, 2008a, 2008b) - have different and more restrictive hypotheses that can provide better results. We still highlight the contribution of this article from the perspective of the realized volatility. Unlike the USA market, the best models had the highest frequencies, suggesting that the microstructure noise is very small and does not significantly grows as the sampling frequency increases, at least for USIM5 and GGBR4.

5. Conclusion

This article applies the recent realized volatility techniques and the Heterogeneous Auto-Regressive (HAR) models to data of four stocks traded at the São Paulo Stock Exchange with an effective participation at the Ibovespa index. The goal is to find out whether these models have a superior forecasting ability to the traditional GARCH (1,1) and the Exponential Weighted Moving-Average (EWMA). Furthermore, we performed an empirical application to the Value-at-Risk with the different volatility estimates generated from the models above.

For that purpose, we had to deal with the microstructure noise in connection with observation of the intraday price to create a realized volatility measure that was really consistent and not biased. Hence, we used the optimal frequency of Bandi and Russel (2005), the microstructure correction filter of Hansen, Large and Lunde (2006), and a variety of frequencies between 30 and 15 minutes. In the next stage, we estimated the HAR models for all the different estimates and compared the results to the forecasts of the GARCH and EWMA models. Then we concluded that the HAR model fits Brazilian data as well or better than in the international literature, being that the higher the sampling frequency of intraday data, the better the model fits. Moreover, in the forecasting comparison with 1- to 10-step-ahead horizons against the GARCH and the EWMA, we found out that the HAR model's forecast was slightly superior, especially at 10-day horizons.

Finally, we used the realized volatility prediction to forecast the maximum loss for the next period at a 1- to 10-day horizon, with different expected loss percentages (10%, 5%, 2.5% and 1%). Our results, based on the Kupiec Test and the Christoffersen Test, point out that the techniques complement each other by offering qualities and disadvantages. The HAR model performed well in three stocks at 1- and 2-step-ahead horizons, independently from the maximum loss percentage attributed in the value-at-risk. Furthermore, it outperformed the GARCH for the GGBR4 series. On the other hand, the GARCH models were also fit to 3 stocks, at all forecasting horizons, but only when the value-at-risk expected losses were small: 2.5% and 1%. The GARCH was not fit to GGBR4 and the HAR was not sufficient to properly model the PETR4 series. Hence, we found out that the models do not offer any advantages over the others and therefore there is not a clear evidence of superiority relative to the utilization for Value-at-Risk.

Thus, the contribution of this article is basically to present evidence that the microstructure noise is small or the filter-based correction method hypotheses do not fit Brazilian data, since the models with high sampling frequency and with no correction method were more successful. Moreover, we demonstrated that the HAR models are fit to Brazilian data, showing superiority with data of S&P500 and DJIA. Further, we found out that this technique's forecasting ability is slightly superior to that of traditional techniques. With respect to the VaR, the application of the technique does not show advantages in any of the econometric models.

References:

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

Aït-Sahalia, Y., Mykland, P. A. and Zhang, L. (2006), Comments on "Realized variance and market microstructure noise." *J. Bus. Econom. Statist.* **24**162–167.

Allen, D.; McAleer,; e Scharth, M. (2009), Princing Options by simulation Using Realized Volatility 18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 http://www.mssanz.org.au/modsim09/D10/allen_D10b.pdf (20/09/2010)

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.

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, Novembere

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., 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° 694*2010, http://www2.dse.unibo.it/wp/694.pdf(20/09/2010)

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 Aproach. 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 Conterbury. Working paper. 2010 http://www.econ.canterbury.ac.nz/RePEc/cbt/econwp/1002.pdf (20/09/2010)

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

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

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

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)

Jou,Y-J.; Wang, C-W; and Chiu, W-C. (2010), Pricing American Options by Using HAR-RV Model. http://www.campus-for-finance.com/fileadmin/pagecontent/Papers_2010/Chiu__Wan-Chien_-Pricing_American_Options_by_Using_HAR-RV_Model.pdf (20/09/2010)

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.

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)

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

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.

Mota, B. de Sá, and Fernandes, M. (2004), Desempenho de estimadores de volatilidade na bolsa de valores de São Paulo. *Rev. Bras. Econ.* [online]. vol.58, n.3, pp. 429-448.
Muller, U., Dacorogna, M., Dav, R., Pictet, O., Olsen, R. and Ward, J.: (1993), Fractals and intrinsic time - a challenge to econometricians. XXXIXth International AEA Conference on Real Time Econometrics, 14-15 Oct 1993, Luxembourg.
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.
Zhang, L., Mykland, P.A., and Aït-Sahalia Y., (2005), A tale of two time scales: Determining integrated volatility with noisy high frequency data, *Journal of the American Statistical Association*, 100, 1394 – 1411.
Zhou B., (1996), High-frequency data and volatility in foreign-exchange rates. *J. Bus. & Econom.Statist.* 14 45–52.

ANNEXIES

Table 01 – Number of observations in each frequency used

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 02 – Estimated optimal frequency

Asset	Sampling Freq. of IQ	Optimal Freq. (min)	Asset	Sampling Freq. of IQ	Optimal Freq. (min)
GGBR4	15 min	7.7553	USIM5	15 min	8.2067
	30 min	6.5562		30 min	6.9453
PETR4	15 min	7.7040	VALE5	15 min	7.5756
	30 min	6.6477		30 min	6.4190

Table 17 – Mean absolute error of forecast of GARCHs e EWMAs

		EWMA				GARCH			
ASSET	PAF	GARCH	RV HLL¹	RV OPT	EWMA	GARCH	RV HLL¹	RV OPT	EWMA
	1	0,0615	0,1950	0,1662	0,0000	0,0000	0,2416	0,2749	0,0615
	2	0,0736	0,2025	0,1746	0,0307	0,0341	0,2454	0,2812	0,0723
	5	0,1015	0,2150	0,1914	0,0686	0,0730	0,2514	0,2955	0,0962
GGBR4	10	0,1366	0,2311	0,2152	0,1100	0,1095	0,2582	0,3109	0,1266
	1	0,0842	0,1524	0,1879	0,0000	0,0000	0,2339	0,2801	0,0776
	2	0,0852	0,1564	0,1933	0,0319	0,0380	0,2365	0,2883	0,0860
	5	0,1083	0,1632	0,1784	0,0529	0,0799	0,2429	0,3051	0,1075
USIM5	10	0,1484	0,2054	0,2499	0,1179	0,1156	0,2483	0,3174	0,1330
	1	0,0776	0,1719	0,1677	0,0000	0,0000	0,2217	0,3000	0,0842
	2	0,0711	0,1745	0,1733	0,0318	0,0356	0,2290	0,3097	0,0925
	5	0,1036	0,1906	0,1925	0,0720	0,0794	0,2511	0,3323	0,1164
PETR4	10	0,1431	0,2076	0,2173	0,1144	0,1185	0,2698	0,3534	0,1492
	1	0,0823	0,2520	0,3282	0,0000	0,0000	0,3456	0,4316	0,1006
	2	0,0928	0,2559	0,3307	0,0317	0,0378	0,3499	0,4357	0,1038
	5	0,1208	0,2685	0,3413	0,0717	0,0802	0,3648	0,4519	0,1177
VALE5	10	0,1557	0,2898	0,3595	0,1137	0,1158	0,3830	0,4681	0,1469

Table 18 – Mincer-Zarnowitz Test of GARCHs and EWMAs

Assets	PETR4		GGBR4		USIM5		VALE5	
PAF\Model	GARCH	EWMA	GARCH	EWMA	GARCH	EWMA	GARCH	EWMA
1	0,5954	0,5534	0,4760	0,4258	0,4897	0,4777	0,4020	0,3574
2	0,5633	0,5208	0,4531	0,4025	0,4618	0,4282	0,3913	0,3430
5	0,4886	0,5135	0,4015	0,3607	0,4080	0,3698	0,3348	0,3003
10	0,4305	0,3920	0,3460	0,3073	0,3654	0,3210	0,2659	0,2382

Table 19 – Forecasting performance of the HARs for GGBR4

	1-step-ahead model				2-step-ahead Model			
Asset	RMSE	MAE	MAPE	Mincer-Z.	RMSE	MAE	MAPE	Mincer-Z.
ALL 1	0,2569	0,1990	4,64%	0,5954	0,2730	0,2099	4,91%	0,5433
ALL 2	0,2702	0,2102	4,85%	0,5787	0,2869	0,2217	5,13%	0,5255
ALL 5	0,2966	0,2303	5,27%	0,5462	0,3125	0,2427	5,55%	0,4965
ALL 15	0,3335	0,2607	5,87%	0,4928	0,3446	0,2688	6,06%	0,4622
ALL 30	0,3786	0,3012	6,71%	0,4375	0,3867	0,3059	6,82%	0,4136
HLL 1	0,2569	0,1990	4,73%	0,5954	0,2730	0,2099	5,00%	0,5433
HLL 2	0,2702	0,2102	4,93%	0,5787	0,2869	0,2217	5,21%	0,5255
HLL 5	0,2966	0,2303	5,33%	0,5462	0,3125	0,2427	5,62%	0,4965
OPT	0,3051	0,2365	5,39%	0,5424	0,3228	0,2502	5,71%	0,4878
	5-step-ahead Model				10-step-ahead Model			
	RMSE	MAE	MAPE	Mincer-Z.	RMSE	MAE	MAPE	Mincer-Z.
ALL 1	0,2992	0,2310	5,41%	0,4527	0,3222	0,2501	5,87%	0,3676
ALL 2	0,3126	0,2433	5,65%	0,4380	0,3341	0,2591	6,03%	0,3608
ALL 5	0,3371	0,2625	6,02%	0,4154	0,3549	0,2765	6,36%	0,3546
ALL 15	0,3659	0,2852	6,45%	0,3916	0,3817	0,2967	6,73%	0,3406
ALL 30	0,4025	0,3181	7,12%	0,3664	0,4171	0,3289	7,38%	0,3222
HLL 1	0,2992	0,2310	5,51%	0,4527	0,3222	0,2501	5,98%	0,3676
HLL 2	0,3126	0,2433	5,74%	0,4380	0,3341	0,2591	6,13%	0,3608
HLL 5	0,3371	0,2625	6,10%	0,4154	0,3549	0,2765	6,43%	0,3546
OPT	0,3453	0,2692	6,16%	0,4152	0,3630	0,2818	6,47%	0,3560

Table 20 – Forecasting performance of the HARs for PETR4

	1-step-ahead Model				2-step-ahead Model			
Asset	RMSE	MAE	MAPE	Mincer-Z.	RMSE	MAE	MAPE	Mincer-Z.
ALL 1	0,2590	0,1950	4,32%	0,6549	0,2724	0,2065	4,59%	0,6187
ALL 2	0,2874	0,2172	4,78%	0,6098	0,2904	0,2197	4,83%	0,6012
ALL 5	0,3026	0,2302	5,02%	0,6087	0,3171	0,2407	5,26%	0,5708
ALL 15	0,3410	0,2669	5,72%	0,5529	0,3540	0,2739	5,89%	0,5186
ALL 30	0,3834	0,3005	6,38%	0,5081	0,3971	0,3097	6,59%	0,4728
HLL 1	0,2590	0,1950	4,37%	0,6549	0,2724	0,2065	4,63%	0,6187
HLL 2	0,2759	0,2082	4,60%	0,6397	0,2904	0,2197	4,87%	0,6012
HLL 5	0,3026	0,2302	5,02%	0,6087	0,3171	0,2407	5,26%	0,5708
OPT	0,3090	0,2358	5,10%	0,6006	0,3212	0,2445	5,30%	0,5690
	5-step-ahead Model				10-step-ahead Model			
	RMSE	MAE	MAPE	Mincer-Z.	RMSE	MAE	MAPE	Mincer-Z.
ALL 1	0,3080	0,2338	5,20%	0,5133	0,3321	0,2570	5,72%	0,4360
ALL 2	0,3258	0,2489	5,48%	0,4992	0,3508	0,2732	0,60%	0,4211
ALL 5	0,3490	0,2675	5,85%	0,4809	0,3721	0,2891	6,33%	0,4115
ALL 15	0,3809	0,2965	6,39%	0,4433	0,3997	0,3111	6,72%	0,3878
ALL 30	0,4312	0,3307	7,05%	0,4075	0,4387	0,3437	7,36%	0,3572
HLL 1	0,3080	0,2338	5,26%	0,5133	0,3321	0,2570	5,78%	0,4360
HLL 2	0,3258	0,2489	5,52%	0,4992	0,3507	0,2732	6,07%	0,4211
HLL 5	0,3490	0,2675	5,85%	0,4809	0,3721	0,2891	6,33%	0,4115
OPT	0,3542	0,2721	5,91%	0,4764	0,3783	0,2947	6,41%	0,4038

Table 21 – Forecasting performance of the HARs for USIM5

	1-step-ahead Model				2-step-ahead Model			
Asset	RMSE	MAE	MAPE	Mincer-Z.	RMSE	MAE	MAPE	Mincer-Z.
ALL 1	0,2589	0,1999	4,62%	0,5707	0,2726	0,2113	4,89%	0,5243
ALL 2	0,2707	0,2090	4,79%	0,5521	0,2843	0,2216	5,08%	0,5064
ALL 5	0,2973	0,2316	5,25%	0,5165	0,3107	0,2419	5,50%	0,4723
ALL 15	0,3332	0,2658	5,93%	0,4536	0,3420	0,2732	6,11%	0,4246
ALL 30	0,3728	0,2993	6,63%	0,3999	0,3803	0,3054	6,77%	0,3761
HLL 1	0,2589	0,1999	4,73%	0,5707	0,2726	0,2113	5,01%	0,5243
HLL 2	0,2707	0,2090	4,90%	0,5521	0,2843	0,2216	5,20%	0,5064
HLL 5	0,2973	0,2316	5,32%	0,5165	0,3107	0,2419	5,57%	0,4723
OPT	0,3047	0,2401	5,42%	0,5026	0,3190	0,2523	5,70%	0,4555
	5 step-ahead Model				10-step-ahead Model			
	RMSE	MAE	MAPE	Mincer-Z.	RMSE	MAE	MAPE	Mincer-Z.
ALL 1	0,2943	0,2294	5,33%	0,4473	0,3154	0,2455	5,71%	0,3670
ALL 2	0,3053	0,2386	5,49%	0,4321	0,3256	0,2533	5,85%	0,3557
ALL 5	0,3315	0,2583	5,89%	0,4011	0,3488	0,2684	6,13%	0,3392
ALL 15	0,3573	0,2852	6,39%	0,3739	0,3735	0,2943	6,61%	0,3187
ALL 30	0,3943	0,3256	7,01%	0,3306	0,4083	0,3218	7,17%	0,2852
HLL 1	0,2943	0,2294	5,46%	0,4473	0,3154	0,2455	5,86%	0,3670
HLL 2	0,3053	0,2386	5,62%	0,4321	0,3256	0,2533	5,99%	0,3557
HLL 5	0,3315	0,2583	5,97%	0,4011	0,3488	0,2684	6,22%	0,3392
OPT	0,3369	0,2646	6,00%	0,3941	0,3509	0,2727	6,20%	0,3451

Table 22 – Forecasting performance of the HARs for VALE5

	1-step-ahead Model				2-step-ahead Model			
Asset	RMSE	MAE	MAPE	Mincer-Z.	RMSE	MAE	MAPE	Mincer-Z.
ALL 1	0,3025	0,2400	5,30%	0,5774	0,3147	0,2474	5,47%	0,5430
ALL 2	0,3130	0,2490	5,46%	0,5772	0,3265	0,2583	5,68%	0,5404
ALL 5	0,3393	0,2706	5,88%	0,5488	0,3557	0,2819	6,14%	0,5045
ALL 15	0,3749	0,3000	6,44%	0,4901	0,3911	0,3098	6,66%	0,4456
ALL 30	0,4127	0,3321	7,07%	0,4475	0,4260	0,3412	7,27%	0,4115
HLL 1	0,3025	0,2400	5,35%	0,5774	0,3147	0,2474	5,53%	0,5430
HLL 2	0,3130	0,2490	5,53%	0,5772	0,3265	0,2583	5,74%	0,5404
HLL 5	0,3393	0,2706	5,95%	0,5488	0,3557	0,2819	6,21%	0,5045
OPT	0,3450	0,2766	6,00%	0,5426	0,3611	0,2866	6,23%	0,4993
	5 step-ahead Model				10-step-ahead-model			
	RMSE	MAE	MAPE	Mincer-Z.	RMSE	MAE	MAPE	Mincer-Z.
ALL 1	0,3378	0,2694	5,97%	0,4739	0,3634	0,2845	6,34%	0,3921
ALL 2	0,3507	0,2802	6,17%	0,4696	0,3766	0,2965	6,56%	0,3893
ALL 5	0,3797	0,3029	6,62%	0,4355	0,4035	0,3181	6,98%	0,3631
ALL 15	0,4132	0,3285	7,08%	0,3810	0,4328	0,3422	7,41%	0,3218
ALL 30	0,4462	0,3581	7,65%	0,3548	0,4621	0,3689	7,91%	0,3088
HLL 1	0,3378	0,2694	6,04%	0,4739	0,3634	0,2845	6,41%	0,3921
HLL 2	0,3507	0,2802	6,24%	0,4696	0,3766	0,2965	6,64%	0,3893
HLL 5	0,3797	0,3029	6,69%	0,4355	0,4035	0,3181	7,06%	0,3631
OPT	0,3852	0,3077	6,71%	0,4289	0,4088	0,3223	7,06%	0,3581

Table 23 – HAR HLL, GARCH and EWMA – Forecasting ability

		RMSE			MAE			MPSE			Micer-Zarnowitz		
PAF	Asset	GARCH	EWMA	HAR HLL[1]	GARCH	EWMA	HAR HLL[1]	GARCH	EWMA	HAR HLL[1]	GARCH	EWMA	HAR HLL[1]
	GGBR4	0,3092	0,3497	0,2569	0,2416	0,2729	0,1990	0,0580	6,52%	4,73%	47,60%	0,4258	0,5954
	PETR4	0,3149	0,3089	0,2590	0,2454	0,2355	0,1950	0,0590	5,24%	4,37%	59,54%	0,5534	0,6549
	USIM5	0,3260	0,3186	0,2589	0,2514	0,2521	0,1999	0,0605	6,02%	4,73%	48,97%	0,4777	0,5707
1	VALE5	0,3365	0,4236	0,3025	0,2582	0,3456	0,2400	0,0623	7,46%	5,35%	40,20%	0,3574	0,5774
	GGBR4	0,2939	0,3583	0,2730	0,2339	0,2778	0,2099	0,0561	6,64%	5,00%	45,31%	0,4025	0,5433
	PETR4	0,3008	0,3220	0,2724	0,2365	0,2453	0,2065	0,0568	5,47%	4,63%	56,33%	0,5208	0,6187
	USIM5	0,3127	0,3345	0,2726	0,2429	0,2627	0,2113	0,0584	6,28%	5,01%	46,18%	0,4282	0,5243
2	VALE5	0,3202	0,4286	0,3147	0,2483	0,3501	0,2474	0,0597	7,56%	5,53%	39,13%	0,3430	0,5430
	GGBR4	0,2834	0,3721	0,2992	0,2217	0,2872	0,2310	0,0492	6,88%	5,51%	40,15%	0,3607	0,4527
	PETR4	0,2949	0,3275	0,3080	0,2290	0,2502	0,2338	0,0509	5,61%	5,26%	48,86%	0,5135	0,5133
	USIM5	0,3203	0,3533	0,2943	0,2511	0,2734	0,2294	0,0558	6,55%	5,46%	40,80%	0,3698	0,4473
5	VALE5	0,3394	0,4433	0,3378	0,2698	0,3611	0,2694	0,0599	7,82%	6,04%	33,48%	0,3003	0,4739
	GGBR4	0,4233	0,3901	0,3222	0,3456	0,3005	0,2501	0,0739	7,22%	5,98%	34,60%	0,3073	0,3676
	PETR4	0,4278	0,3735	0,3321	0,3499	0,2861	0,2570	0,0747	6,41%	5,78%	43,05%	0,3920	0,4360
	USIM5	0,4460	0,3693	0,3154	0,3648	0,2838	0,2455	0,0781	6,81%	5,86%	36,54%	0,3210	0,3670
10	VALE5	0,4677	0,4650	0,3634	0,3830	0,3793	0,2845	0,0821	8,24%	6,41%	26,59%	0,2382	0,3921

Table 24 – Performance of the HAR ALL[1] model in the Value-at-Risk for USIM5 and GGBR4

		USIM5				GGBR4			
	Test\PAF	1	2	5	10	1	2	5	10
10%	Violations	**94**	**99**	**92**	**100**	**96**	**96**	**100**	**98**
	% Violations	**9,40%**	**9,90%**	**9,20%**	**10,00%**	**9,60%**	**9,60%**	**10,00%**	**9,80%**
	Kupiec	**0,4073**	**0,0111**	**0,7287**	**-**	**0,1799**	**0,1799**	**0,00**	**0,0447**
	Christoffersen	**3,6248**	5,5785	12,2844	11,2971	**0,2579**	**0,5783**	**1,7920**	**0,7304**
5%	Violations	**53**	**58**	**52**	**61**	**47**	**40**	**49**	**57**
	% Violations	**5,30%**	**5,80%**	**5,20%**	**6,10%**	**4,70%**	**4,00%**	**4,90%**	**5,70%**
	Kupiec	**0,1860**	**1,2843**	**0,0832**	**2,3877**	**0,1932**	**2,2534**	**0,0212**	**0,9889**
	Christoffersen	**3,3300**	**3,2163**	11,6095	12,3945	**0,2161**	**2,3537**	**1,0168**	**1,9127**
2,5%	Violations	**29**	**29**	**33**	**38**	**18**	15	**25**	**31**
	% Violations	**2,90%**	**2,90%**	**3,30%**	**3,80%**	**1,80%**	1,50%	**2,50%**	**3,10%**
	Kupiec	**0,6248**	**0,6248**	**2,3895**	**5,9961**	**2,2240**	4,7774	**0,0000**	**1,3739**
	Christoffersen	**0,65**	**0,65**	9,90	21,94	**0,00**	0,00	**0,20**	**2,30**
1%	Violations	**15**	**14**	**15**	**22**	**11**	**10**	**14**	**15**
	% Violations	**1,50%**	**1,40%**	**1,50%**	**2,20%**	**1,10%**	**1,00%**	**1,40%**	**1,50%**
	Kupiec	**2,1892**	**1,4374**	**2,1892**	**10,8382**	**0,0978**	**-**	**1,4374**	**2,1892**
	Christoffersen	**0,0000**	**0,0000**	**0,0000**	29,9335	**0,0000**	**0,0000**	**0,0000**	**0,0000**

Table 25 – Performance of the HAR ALL[1] model in the Value-at-Risk for VALE5

	PAF	1		2		5		10	
	Test	With 1 SD	Without SD	With 1 SD	Without SD	With 1 SD	Without SD	With 1 DP	Without SD
10%	Violations	**84**	144	**91**	141	79	139	76	151
	% Violations	**8,40%**	14,40%	**9,10%**	14,10%	7,90%	13,90%	7,60%	15,10%
	Kupiec	**2,9914046**	19,2042659	**0,9251578**	16,7891012	5,24210404	15,2613959	6,91998522	25,4033048
	Christoffersen	5,04080362	22,1143405	5,55190822	19,6714595	0,91295951	15,5415761	8,38326141	25,8900969
5%	Violations	**44**	99	**42**	101	**41**	98	**40**	95
	% Violations	**4,40%**	9,90%	**4,20%**	10,10%	**4,10%**	9,80%	**4,00%**	9,50%
	Kupiec	**0,7884785**	39,8251532	**1,4214956**	42,813945	**1,8120182**	38,3642907	**2,2534116**	34,1182944
	Christoffersen	**0,0021541**	6,50568575	**0,7925603**	0,64482277	**2,582174**	3,29243511	5,2275551	2,9610076
2,5%	Violations	**22**	69	**21**	66	**22**	68	**23**	71
	% Violations	**2,20%**	6,90%	**2,10%**	6,60%	**2,20%**	6,80%	**2,30%**	7,10%
	Kupiec	**0,384553**	54,1180371	**0,6935456**	47,8916079	**0,384553**	52,0108564	**0,1685458**	58,425391
	Christoffersen	**0,8280432**	54,3477052	**1,2428039**	45,5946819	**0,8280432**	49,2711348	6,20457364	58,003175
1%	Violations	**7**	35	**10**	32	**9**	40	**15**	44
	% Violations	**0,70%**	3,50%	**1,00%**	3,20%	**0,90%**	4,00%	**1,50%**	4,40%
	Kupiec	**1,0156325**	38,3301031	**0**	30,9342029	**0,1045205**	51,8219642	**2,1892484**	63,5624781
	Christoffersem	**0**	38,7757514	**0**	31,7210201	**0**	46,5951126	13,2976301	61,3656862

Table 26 – Performance in the Value-at-Risk of the GARCHs and EWMAs for GGBR4

Model		EWMA				GARCH			
	PAF	Violations	% Viol.	Kupiec	Christoffersen	Violations	% Viol.	Kupiec Test	Christ. Test
	1	**106**	**10,60%**	**0,3931**	**1,4512**	**99**	**9,91%**	**0,0090**	**0,5737**
	2	**109**	**10,90%**	**0,8770**	**1,6522**	**106**	**10,61%**	**0,4066**	11,0774
	5	**109**	**10,90%**	**0,8770**	4,7189	**108**	**10,81%**	**0,7129**	8,6710
10%	10	**110**	**11,00%**	**1,0798**	**2,2771**	**111**	**11,11%**	**1,3276**	11,1719
	1	**55**	**5,50%**	**0,5105**	**0,3208**	**49**	**4,90%**	**0,0191**	**1,1612**
	2	**55**	**5,50%**	**0,5105**	**2,6187**	**62**	**6,21%**	**2,8515**	13,5336
	5	**61**	**6,10%**	**2,3877**	**2,6764**	**61**	**6,11%**	**2,4110**	11,0710
5%	10	**64**	**6,40%**	**3,8054**	**3,4317**	**60**	**6,01%**	**2,0054**	13,8398
	1	**29**	**2,90%**	**0,6248**	**0,0298**	**25**	**2,50%**	**0,0000**	**0,2002**
	2	**28**	**2,80%**	**0,3556**	**0,0577**	40	4,00%	7,8633	15,9755
	5	**32**	**3,20%**	**1,8494**	**0,7868**	39	3,90%	6,9165	15,5454
2,5%	10	**35**	**3,50%**	**3,6560**	**0,0476**	38	3,80%	6,0230	15,1918
	1	18	1,80%	5,2251	**0,0000**	11	1,10%	0,0999	**0,0000**
	2	16	1,60%	3,0766	**0,0000**	19	1,90%	6,4908	26,0279
	5	20	2,00%	7,8272	**0,0000**	19	1,90%	6,4908	26,0279
1%	10	20	2,00%	7,8272	**0,0000**	19	1,90%	6,4908	26,0279

Table 27 – Performance in the Value-at-Risk of the GARCHs and EWMAs for PETR4

Model		EWMA				GARCH			
	PAF	Violations	% Viol.	Kupiec	Christoffersen	Violations	% Viol.	Kupiec Test	Christ. Test
	1	**112**	**11,20%**	**1,5463**	4,9460	**99**	**9,91%**	**0,0090**	5,5411
	2	**112**	**11,20%**	**1,5463**	11,1145	**101**	**10,11%**	**0,0134**	9,5868
	5	**113**	**11,30%**	**1,8099**	10,5754	**103**	**10,31%**	**0,1059**	7,0654
10%	10	**118**	**11,80%**	**3,4238**	13,2259	**103**	**10,31%**	**0,1059**	7,0654
	1	**65**	**6,50%**	**4,3455**	**1,7783**	**55**	**5,51%**	**0,5211**	**0,8375**
	2	**64**	**6,40%**	**3,8054**	7,2738	**55**	**5,51%**	**0,5211**	**1,7395**
	5	71	7,10%	8,2609	6,3864	**59**	**5,91%**	**1,6353**	**3,3529**
5%	10	72	7,20%	9,0221	2,7501	**58**	**5,81%**	**1,3013**	**2,0833**
	1	**32**	**3,20%**	**1,8494**	**0,0007**	**27**	**2,70%**	**0,1641**	**0,2583**
	2	**32**	**3,20%**	**1,8494**	**0,0007**	**29**	**2,90%**	**0,6330**	**0,0000**
	5	39	3,90%	6,8875	1,2349	**28**	**2,80%**	**0,3618**	**0,4187**
2,5%	10	40	4,00%	7,8323	5,2202	**31**	**3,10%**	**1,3863**	**2,3072**
	1	17	1,70%	4,0910	**0,0000**	**14**	**1,40%**	**1,4455**	**0,0000**
	2	18	1,80%	5,2251	**0,0000**	**12**	**1,20%**	**0,3838**	**0,0000**
	5	20	2,00%	7,8272	**0,0000**	**15**	**1,50%**	**2,1994**	**0,0000**
1%	10	22	2,20%	10,8382	0,4435	17	1,70%	4,1052	5,2233

Table 28 – Performance in the Value-at-Risk of the GARCHs and EWMAs for USIM5

		EWMA				GARCH			
	PAF	Violations	% Viol.	Kupiec	Christoffersen	Violations	% Viol.	Kupiec Test	Christ. Test
	1	**98**	**9,80%**	**0,0447**	6,7810	**92**	**9,21%**	**0,7111**	4,4237
	2	**100**	**10,00%**	**0,0000**	11,2971	**89**	**8,91%**	**1,3667**	**3,5471**
	5	**102**	**10,20%**	**0,0442**	9,0407	**88**	**8,81%**	**1,6342**	5,2401
10%	10	**94**	**9,40%**	**0,4073**	9,2025	**86**	**8,61%**	**2,2443**	8,8878
	1	**58**	**5,80%**	**1,2843**	6,4468	**45**	**4,50%**	**0,5334**	4,0078
	2	**56**	**5,60%**	**0,7308**	12,0687	**44**	**4,40%**	**0,7759**	**1,3151**
	5	**60**	**6,00%**	**1,9842**	8,5864	**49**	**4,90%**	**0,0191**	**0,9936**
5%	10	**61**	**6,10%**	**2,3877**	10,0068	**53**	**5,31%**	**0,1924**	11,2251
	1	**26**	**2,60%**	**0,0405**	**1,8308**	**22**	**2,20%**	**0,3784**	**0,8154**
	2	**28**	**2,80%**	**0,3556**	**1,4233**	**23**	**2,30%**	**0,1645**	**0,5092**
	5	**34**	**3,40%**	**2,9923**	11,6661	**24**	**2,40%**	**0,0395**	**0,0000**
2,5%	10	37	3,70%	5,1594	10,4229	**33**	**3,30%**	**2,4060**	11,7758
	1	**15**	**1,50%**	**2,1892**	**1,5151**	**11**	**1,10%**	**0,0999**	**0,0000**
	2	**14**	**1,40%**	**1,4374**	**1,7458**	**12**	**1,20%**	**0,3838**	**0,0000**
	5	**15**	**1,50%**	**2,1892**	10,0658	**12**	**1,20%**	**0,3838**	**0,0000**
1%	10	22	2,20%	10,8382	14,4233	**16**	**1,60%**	**3,0887**	16,7392

Table 29 – Performance in the Value-at-Risk of the GARCHs and EWMAs for VALE5

Model		EWMA				GARCH			
	PAF	Violations	% Viol.	Kupiec	Christoffersen	Violations	% Viol.	Kupiec Test	Christ. Test
	1	**106**	**10,60%**	**0,3931**	**1,4512**	**85**	**8,50%**	**2,6204**	**3,7676**
	2	**109**	**10,90%**	**0,8770**	**1,6522**	**88**	**8,80%**	**1,6606**	**2,3853**
	5	**109**	**10,90%**	**0,8770**	4,7189	**90**	**9,00%**	**1,1458**	**3,1559**
10%	10	**110**	**11,00%**	**1,0798**	**2,2771**	**87**	**8,70%**	**1,9553**	**3,6161**
	1	**55**	**5,50%**	**0,5105**	**0,3208**	35	3,50%	5,2684	5,7123
	2	**55**	**5,50%**	**0,5105**	2,6187	**46**	**4,60%**	**0,3457**	**0,3533**
	5	**61**	**6,10%**	**2,3877**	2,6764	**44**	**4,40%**	**0,7885**	**1,3393**
5%	10	**64**	**6,40%**	**3,8054**	3,4317	**51**	**5,10%**	**0,0209**	8,4600
	1	**29**	**2,90%**	**0,6248**	**0,0298**	**15**	**1,50%**	**4,7774**	**0,0000**
	2	**28**	**2,80%**	**0,3556**	**0,0577**	**25**	**2,50%**	**0,0000**	**2,0569**
	5	**32**	**3,20%**	**1,8494**	**0,7868**	**26**	**2,60%**	**0,0405**	**0,0000**
2,5%	10	**35**	**3,50%**	**3,6560**	**0,0476**	**26**	**2,60%**	**0,0405**	**1,8685**
	1	18	1,80%	5,2251	**0,0000**	2	0,20%	9,6267	**0,0000**
	2	**16**	**1,60%**	**3,0766**	**0,0000**	**12**	**1,20%**	**0,3798**	**0,0000**
	5	20	2,00%	7,8272	**0,0000**	**13**	**1,30%**	**0,8306**	**0,0000**
1%	10	20	2,00%	7,8272	**0,0000**	**14**	**1,40%**	**1,4374**	7,6100