Financial Volatility Forecasting by Nonlinear Support Vector Machine Heterogeneous Autoregressive Model:
Evidence from Nikkei 225 Stock Index

Md. Ashraful Islam Khan
Institute of Economic Research, Graduate School of Economics, Hitotsubashi University
Tokyo 186-8603, Japan

&
Department of Population Science and Human Resource Development, Rajshahi University
Rajshahi-6205, Bangladesh
E-mail: khan75ru@yahoo.com

Received: January 19, 2011        Accepted: February 18, 2011       doi:10.5539/ijef.v3n4p138

Abstract
Support vector machines (SVMs) are new semi-parametric tool for regression estimation. This paper introduced a new class of hybrid models, the nonlinear support vector machines heterogeneous autoregressive (SVM-HAR) models and aimed to compare the forecasting performance with the classical heterogeneous autoregressive (HAR) models to forecast financial volatilities. It was observed through empirical experiment that the newly proposed hybrid (SVM-HAR) models produced higher predicting ability than the classical HAR model.

Keywords: Realized volatility, Bi-power variation, Jump, Tri-power variation, Heterogeneous Autoregressive model, Nonlinear Support Vector Machine, High frequency Nikkei-225 data

1. Introduction
Volatility, the standard deviation of the continuously compounded returns of a financial instrument over a specific time horizon, is both the boon and bane of all traders, you can’t live with it and you can’t really trade without it. Most of the financial researchers are mainly concerned with modeling and forecasting volatility in asset returns to quantify the risk of financial instruments over a particular time period so that the risk manager and practitioners can realize whether their portfolio will decline in the future and they may want to sell it before it becomes too volatile. Therefore, volatility plays the key roles in the theory and applications of asset pricing, optimal portfolio allocation, and risk management.

Researches on time varying volatility using the time series models have been active ever since Engle (1982) introduced the ARCH model. The GARCH model, generalized by Bollerslev (1986), has been extended in various directions and these extensions recognize based on the various researcher’s empirical evidences that there may be important nonlinearity, asymmetry, and long memory properties in the volatility process. The popular extensions can be referred to Nelson’s (1991) EGARCH model, Glosten, Jaganathan, and Runkle’s (1993) GJR-GARCH which both account for the asymmetric relationship between stock returns and changes in variance (see, e.g., Black 1976, the beginning study of the asymmetric effect and Engle and Ng, 1993 for further discussion). Engle’s (1990) AGARCH; Ding, Granger and Engle’s (1993) APARCH; Zakoian’s (1994) TGARCH; and Sentana’s (1995) QGARCH models also have been developed for the flexibility of the models. The stochastic volatility (SV) modeling capitalized on and often contributed in turn to the concurrent development in the Bayesian statistical analysis using Markov chain Monte Carlo procedure (see, e.g., Shephard (2005)).

When GARCH type and SV latent volatilities models are used, a well established result in the financial time series literature is that the standardized returns do not have a Gaussian distribution. The excess kurtosis factor of time series motivates the use of heavy-tailed distributions. For example, Student’s t distribution has been used by Bollerslev (1987), GED by Nelson (1991), both Student’s t and GED by Hsieh (1989) as alternative distributional models for innovations. The researchers have found that returns usually exhibit empirical regularities including thick tails, volatility clustering, leverage effects (see, e.g., Bollerslev et al. 1994). Andersen et al. (2000a, b, 2001, 2003) showed that the distribution of the standardized exchange rate series was almost Gaussian when the realized volatility (RV) was used. Furthermore, the logarithm of the realized volatilities was also nearly Gaussian. It was also
corroborated for stock returns in Andersen et al. (2001a). Other literatures on realized volatility can be referred along with many researchers to Aït-Sahalia and Mancini (2006), Ghysels and Sinko (2006), Corradi et al. (2006), and recently Corsi et al. (2008).

In addition, there is significant evidence of long memory in the time series, which has been conventionally modeled as an ARFIMA \((p,d,q)\) process (see, e.g., Andersen et al. 2000a,b, 2001a, 2003). A large number of papers in the RV literature employ the ARFIMA model without a conditionally heteroskedastic error specification to fit daily RV series (see, e.g., Oomen 2001, Giot and Laurent 2004).

Corsi et al. (2001) and Corsi (2009) proposed the Heterogeneous Autoregressive Realized Volatility (HAR-RV) model as an alternative to the ARFIMA model and it has quickly become popular for modeling the dynamics of RV and other related volatility measures due to its ease estimation and extendability of the baseline model. The HAR-RV model employs a few predictor terms, the past daily RVs averaged over different horizons (typically a day, a week, and a month), and is capable to producing slow-decay patterns in autocorrelations exhibited by many RV series. Another recent development in the RV literature is the approach due to Barndorff-Nielsen and Shephard (2004, 2006), Andersen et al. (2003, 2007) of decomposing the RV into the contribution of continuous sample path variation and that of jumps. Extending the theory of quadratic variation of semimartingales, Barndorff-Nielsen et al. (2006) provided an asymptotic statistical foundation for this decomposition procedure under very general conditions.

However, all of the models do require specified distribution of innovations in order to estimate the model specification and to appropriately forecast future values. The semi-parametric approaches do not require any assumptions on data property (return distribution). These models have been successfully used for modeling and forecasting time series including volatility. One of such models is Support Vector Machine (SVM), introduced by Vepnik (1995), that guarantees to obtain globally optimal solution (see, e.g., Cristianini and Shawe-Taylor, 2000), which solves the problems of multiple local optima in which the neural network usually get trapped into. Pérez-Cruz et al. (2003) predicted GARCH \((1,1)\) based volatility by SVM and showed that the SVM-GARCH\((1,1)\) model yielded better predictive ability than the parametric GARCH\((1,1)\) model. Chen et al. (2008) proposed recurrent SVM as a dynamic process to model GARCH \((1,1)\) based Volatility and showed through simulated and real data that the model produced better performance than MLE based GARCH \((1,1)\) model. More recently, Ou and Wang (2010) proposed GARCH-LSSVM, EGARCH-LSSVM and GJR-LSSVM hybrid models based on modification of Suykens and Vandewalle (1999) to forecast the leverage effect volatilities of ASEAN stock markets. They showed that these models provided improved performances in forecasting the leverage effect volatilities especially during the recently global financial market crashes in 2008.

This paper, closer to Andersen et al. (2003, 2007), aims to apply the SVM approach on HAR-RV models to forecast empirically the daily RV of the Nikkei 225 index. Watanabe and Yamaguchi (2007), Ishida and Watanabe (2009) among other researchers studied the RV of the Nikkei 225 index and reported empirical findings. But, to the author’s knowledge, this paper is the first to apply the SVM-HAR-RV model to RV literature.

The plan for the rest of the paper is as follows. In section 2, we briefly discuss the realized volatility, realized bi-power variation, and jump component extraction. Section 3 describes the data and summary statistics. Section 4 describes the SVM volatility model. Section 5 describes different HAR-RV models. Section 6 reports the forecasting results of the RV and Section 7 concludes with suggestions for further research.

2. Realized volatility, realized bi-power variation and jump component extraction

If we consider a simple diffusion process

\[
dp(s) = \mu(s)dt + \sigma(s)dW(s)
\]

where \(p(s)dt\) is the instantaneous log-price, \(W(s)\) is a standard Brownian process and \(\sigma(s)\) is the standard deviation of \(dp(s)\), which may be time-varying but is assumed to be independent of \(dW(s)\). Then the volatility for day 1 is defined as the integral of \(\sigma^2(s)\) over the interval \((t, t+1)\) i.e., \(\int_t^{t+1} \sigma^2(s)ds\), which is known as integrated volatility and it is unobserved. Let the discretely sampled \(\Delta\)-period returns be denoted by, \(r_t^{\Delta} = p(t) - p(t - \Delta)\). If the process (in our case the log of Nikkei 225 index level process) is a continuous semimartingale then under mild regularity conditions,

\[
RV_t = \sum_{j=1}^{\left\lfloor \frac{t}{\Delta} \right\rfloor} |r_t^{\Delta}|^{1/2} \rightarrow \int_t^{t+1} \sigma^2(s)ds \quad \text{as} \quad \Delta \downarrow 0
\]

\(RV_t\) is the \(t\)-th day realized variance since \(t\) has the daily unit and \(\left\lfloor \frac{t}{\Delta} \right\rfloor\) is integer. We will hereafter use the terms “realized volatility” or “realized variance” interchangeably, or their common abbreviation “RV.”
Again, if the process is semimartingale with finite-activity jumps, i.e., only a finite number of jumps occurring in any finite time interval, such as Poisson jumps, then the realized variance converges to the quadratic variation, which can be decomposed as,

\[ RV_t \overset{p}{\to} \int_t^{t+1} \sigma^2(s)ds + \sum_{s < t + \Delta} k^2(s) \text{ as } \Delta \downarrow 0 \]

where \( k(s) \) refers to the size of the jump occurring at time \( s \). Barndorff-Nielsen and Shephard (2004, 2006) showed that even in the presence of jumps the bipower variation

\[ BV_t \equiv \mu_t^2 \sum_{j=3}^{\Delta} \prod_{i=1}^{[s-\Delta_i]} |r_{t+i\Delta}| \prod_{j=1}^{[s-\Delta_j]} |r_{t+j\Delta}| \]

where \( \mu_t \equiv \sqrt{2/\pi} \), holds under mild conditions and proposed to use

\[ RV_t - BV_t \overset{p}{\to} \sum_{s < t + \Delta} k^2(s) \]

or

\[ J_t \equiv \max((RV_t - BV_t), 0) \]

as an estimator for \( \sum_{s < t + \Delta} k^2(s) \). \( J_t \) is known to take non-zero, small values very frequently due to measurement and possibly due to the presence of jumps infinite-activity types.

Andersen et al. (2007) introduced shrinkage estimator for the jump contribution based on the asymptotic distribution theory developed by Barndorff-Nielsen and Shephard (2004, 2006) and Barndorff-Nielsen et al. (2006) as

\[ SJ_t \equiv I(Z_t > \Phi_0). (RV_t - BV_t) \]

where \( I \) is an indicator function, \( Z_t \equiv \Delta^{-1/2} \frac{(RV_t - BV_t)\Phi_0^{-1}}{\sqrt{[\mu_t^4 + 2\mu_t^2 - 5]_{\max}}[1, T Q_t, BV_t^{-2}]} \) is asymptotically standard normally distributed under the null hypothesis of no jumps, \( \mu_t \equiv \sqrt{2/\pi} \), \( \Phi_0 = \Phi(\alpha) \), the standard normal distribution function where \( \alpha \) is usually set to the values such as .999 so that \( J_t \) can picks up only “significance jumps” and the realized tripower variation

\[ TQ_t \equiv \Delta^{-1} \mu_t^{-2} \sum_{j=3}^{\Delta} |r_{t+j\Delta}|^{4/3} |r_{t+(j-1)\Delta}|^{4/3} |r_{t+(j-2)\Delta}|^{4/3} \int_t^{t+1} \delta^4(s)ds \text{ as } \Delta \downarrow 0 \]

where \( \mu_t^{-2} \equiv 2^{2/3} \Gamma^4(7/6) \Gamma^4(1/2)^{-1} \). The convergence result holds even in the presence of jumps.

Andersen et al. (2007) introduced the shrinkage estimator for the continuous sample path variation as

\[ C_t \equiv I[Z_t < \Phi_0] RV_t + I[Z_t > \Phi_0] BV_t \]

Andersen et al. (2007) also proposed microstructure-noise-robust versions of \( BV_t \) and \( TQ_t \) as

\[ BV_t \equiv \mu_t^{-2} (1 - 2\Delta)^{-1} \sum_{j=3}^{\Delta} |r_{t+j\Delta}| \prod_{j=1}^{[s-\Delta_j]} |r_{t+j\Delta}| \]

\[ TQ_t \equiv \Delta^{-1} \mu_t^{-2} (1 - 4\Delta)^{-1} \sum_{j=3}^{\Delta} |r_{t+j\Delta}|^{4/3} |r_{t+(j-1)\Delta}|^{4/3} |r_{t+(j-2)\Delta}|^{4/3} \]

The definitions of \( J_t \) and \( C_t \) will be modified as well.

3. Data Description and summary statistics

3.1. Calculation of intraday returns and related realized volatility measures from minute-by-minute Nikkei 225 data

This paper measures the realized volatility of the Nikkei 225 index for the sample of the period 11 March 1996 to 30 September 2009. First, construct a “five- minute (percentage) returns” series by taking the five-minute log differences multiplied by hundred from the minute-by-minute data. This choice is made to mitigate the effect of microstructure related noise and increase the precision of volatility measures. (see, e.g., Ishida and Watanabe, 2009; Watanabe and Yamaguchi, 2007).

The Tokyo Stock Exchange is open only for 9:00-11:00 (Morning Session) and 12:30-15:00 (Afternoon Session). Our database includes every minute prices of the Nikkei 225 stock index for both sessions. This paper first extracts prices for 9:01, 9:05, 9:10,……, 11:00 in the morning session and for 12:31, 12:35, 12:40,……, 15:00 in the afternoon session. Sometimes, the last transaction price for morning (and/or afternoon) session is observed slightly after 11:00 (and/or 15:00). In such cases, the last prices instead of prices at 11:00 (and/or 15:00) are used. Next using these prices the five-minute returns as mentioned in section 2 are calculated. There are 54 five-minute returns for a typical trading day in total, 24 from the morning session and 30 from the afternoon session.
Given the recent literature on the market microstructure noise effect on realized volatility estimation, the optimal choice of sampling frequency as studied by Bandi and Russell (2003, 2008) has been considered here. The sampling frequency $M^{opt}$ (the number of observations per day) is calculated as (see also Zhang et al., 2005 and Clements et al. 2008)

$$M^{opt} = \left( \frac{\hat{a}}{\hat{b}} \right)^{\frac{1}{2}}$$

where $\hat{a} = \left( \frac{1}{T} \sum_{t=1}^{T} r_{t} \right)^{2}$ and $\hat{b} = \frac{1}{T} \sum_{t=1}^{T} \hat{Q}_{t}$, $Q_{t} = \frac{M_{15}}{3} \sum_{j=1}^{M_{15}} r_{t}^{4}$. $M_{15}$ is the 15-minute returns and $M$ is the highest frequency at which data are available. In our case, it is 1-minute returns. The 15-minute intraday returns are being considered to calculate realized volatility as well.

We cannot calculate the 5-minute, 15-minute and optimally-sampled returns for the non-trading hours including lunch time and overnight period though we can calculate the lunch time and overnight returns by considering the last price of the morning session and the first prices of the afternoon session, and the last price of the afternoon session and the first price of the next morning session but following Hansen and Lunde (2005), we drop this idea and scale the realized volatility as follow,

$$RV_{t} \equiv C^{*} RV_{t}^{(0)}$$

where $C^{*} = \frac{\sum_{t=1}^{T} (R_{t} - \bar{R})^{2}}{\sum_{t=1}^{T} RV_{t}^{(0)}}$, where $\bar{R} = \frac{\sum_{t=1}^{T} R_{t}}{T}$, and $T$ is the number of complete trading days.

In my sample period, the first trading in the second session from January 1, 2006 to April 21, 2006 observed at 13:01. Therefore, I remove these trading days along with the sessions from half trading days including the first and the last trading days of each year. The remaining number of complete trading days, $T$ is 3279. We calculate $RV_{t}$ and $J_{t}$ by using this 3279 days data for the four series.

3.2. Properties of the realized volatility and related measures

Summary statistics of daily returns, the daily RV, its standard deviation form, i.e., $RV^{1/2}$, the logarithmic form i.e., $ln(RV)$, the daily jump, microstructure noise robust version of daily jump (MSNR-Jump) (where $BV_{t}$ has been calculated according to Andersen et al., 2007) series and their standard deviation and logarithmic) are presented on Table 1a. The summary statistics of continuous path component, significant jump series, the microstructure noise robust version of the continuous path component (MSNR-C) and significant jump (MSNR-SJ) series due to Andersen et al. (2007), and their standard deviation and logarithmic form are presented on Table 1b. In addition to the sample skewness and kurtosis, the Jarque-Bera (JB) statistic for testing normality and the Ljung-Box statistics of order 5, 10 and 22 (corresponding to roughly one week, two weeks and a month) for testing serial correlations up to their respective order are also presented on the Tables.

From Table 1a, we observed that the unconditional distribution of the daily return series is negatively skewed but highly significantly nonnormal with high positive kurtosis. The LB statistics also indicate that the series is significantly serially correlated. We also observed that the daily RV, Jump and MNSR-Jump series are highly unconditionally nonnormally distributed with large positive values of skewness and kurtosis and highly significantly serially correlated. The average value of Jump and MSNR-Jump are 1.471 and 1.504 respectively with positive minimum values, that implies more than one jump occurred in every single days.

The square-root transformation reduces the deviation from normality but still huge. The log transformation brings down the sample skewness and kurtosis values for the RV series but still significantly nonnormal. All the transformed series remain highly significantly serially correlated.

Looking at the summary statistics from Table 1b, where $\alpha = 0.999$ in (7) and (9), we observed that the average of the significant jump and MSNR-SJ are slightly reduced but still greater than one while the minimum values reduced to zero. The Jauque-Bera statistic shows the strong evidence of highly significant nonlinearity for all series and the LB statistic shows the strong evidence of highly significant serial correlation.

Figure 1. shows the daily RV, Jump, MSNR-Jump, Significant Jump and MSNR-SJ. We visually observe few big jumps in the initial stage and the biggest jumps in the ending part of our sample period, the period of global financial market crashes.

4. The Support Vector Machines (SVMs)

The Support Vector Machines (SVMs) were introduced by Vapnik (1995) based on the statistical learning theory, which had been developed over the last three decades by Vapnik, Chervonenkis and others (see, e.g., Vapnik 1982,
1995) from a nonlinear generalization of the Generalized Portrait algorithm. SVMs were developed to solve the classification problem, but recently they have been extended to the domain of regression problems (e.g., Vapnik et al. 1997). The SVMs usually map data to a high-dimensional feature space and apply a simple linear method to the data in that high-dimensional space nonlinearly related to the input space. Moreover, even though we can think of SVMs as a linear algorithm in a high-dimensional space, in practice, it does not involve any computations in that high-dimensional space (see, e.g., Karatzoglou and Meyer 2006). The terminology for SVMs can be slightly confusing in the literature. In few literatures, SVM refers to both classification and regression with support vector methods. In this paper, the term SVM will be used for the Nonlinear Support Vector Regression (NL-SVR). The mathematical formulation of SVM is as follows,

In the $\epsilon$-insensitive support vector regression of Vapnik (1995), our goal is to find a function $f(x)$ that has an $\epsilon$ deviation from the actually obtained targets $y_t$ for all training data, and at the same time, is as flat as possible. Suppose $f(x)$ takes the following form

$$f(x) = k(\omega, x) + b \quad \text{with} \quad \omega \in X, b \in \mathbb{R}$$

(14)

where $X$ is the space of the input patterns and $k(\cdot, \cdot)$ denotes the kernel function. Flatness of the above model means need to find the small $\omega$. One way to ensure this is to minimize the Euclidean norm, i.e., $\|\omega\|^2$ (see, e.g., Smola 1998). By applying the soft margin formulation of Cortes and Vapnik, (1995), and the Karush-Kuhn-Tucker (KKT) conditions (Karush, 1939, Kuhn and Tucker 1951) one can estimate the above model as

$$f(x) = \sum_{t=1}^{T} (\alpha_t - \alpha_t^*) k(x_t, x) + b$$

(15)

where $b$ can be computed as

$$\tilde{b} = y_t - k(\omega, x_t) - \epsilon \quad \text{for} \quad \alpha_t \in (0, C)$$

$$\tilde{b} = y_t - k(\omega, x_t) + \epsilon \quad \text{for} \quad \alpha_t^* \in (0, C)$$

where, $C > 0$ determines the trade-off between the flatness of the $f(x)$ and the amount up to which derivations larger than $\epsilon$ are tolerated and $\alpha_t, \alpha_t^* \geq 0$. See, e.g., Smola and Schölkopf (1998) for further discussion. A several numbers (see, e.g., Kernlab in R, MATLAB, etc) of statistical software are available to handle SVM method.

According to Cortes and Vepnik(1995), any symmetric positive semi-definite function that satisfies the Mercer’s conditions can be used as a kernel function in the SVMs context. The Mercer’s conditions are

$$\int \int K(x, y) g(x)g(y)dx dy > 0 \quad \text{and} \quad \int g^2(x)dx < \alpha,$$

where

$$K(x, y) \equiv \sum_{t=1}^{\infty} \alpha_t \psi(x) \psi(y), \quad \alpha_t \geq 0$$

This paper used the Polynomial kernel function (used for out-of-sample forecast) and Laplacian kernel function (used for in-sample forecast) for SVMs. The general form of the Polynomial kernel function is

$$K(x, y) \equiv (\text{scale}.(x, y) + \text{offset})^{\text{degree}}$$

and the Laplacian kernel function is

$$K(x, y) \equiv \exp \left(-\frac{|x-y|}{\sigma}\right)$$

See, e.g., Smola and Schölkopf (1998) for further discussion.

5. HAR-RV models

The HAR-RV class volatility models proposed by Corsi (2003) on the basis of a straightforward extension of the so-called Heterogeneous ARCH (HARCH) class of models analyzed by Müller et al. (1997). To sketch the HAR-RV model, define the multi-period realized volatilities by the normalized sum of the one-period volatilities,

$$RV_{t+h} = h^{-1}(RV_{t+1} + RV_{t+2} + \cdots + RV_{t+h})$$

(16)

Note that, by definition of the daily volatilities, $RV_{t+h} \equiv RV_{t+1}$. Also, provided the expectations exist, $E(RV_{t+h}) \equiv E(RV_{t+1})$ for all $h$. (see, e.g., Andersen et al. 2003, 2007). Also $h=5$ and $h=22$ will produce the weekly and monthly volatilities, respectively. The daily HAR-RV model of Corsi (2003) may then be expressed as

$$RV_{t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + e_{t+1}$$

(17)

where $t = 1, 2, \ldots, \ldots, T$.

Andersen et al. (2003, 2007) included the jump component, which has been explained in the Section 2, as an explanatory variable to the above model and introduced the new model as

$$RV_{t+h} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5,t} + \beta_M RV_{t-22,t} + \beta J_t + e_{t+h}$$

(18)
The standard deviation and logarithmic form of the above model respectively are

\[
(RV_{t+h})^{1/2} = \beta_0 + \beta_D (RV_t)^{1/2} + \beta_w (RV_{t-s_1})^{1/2} + \beta_M (RV_{t-s_2})^{1/2} + \beta_J (J_t)^{1/2} + \epsilon_{t+h}
\]  
(19)

and

\[
\log(RV_{t+h}) = \beta_0 + \beta_D \log(RV_t) + \beta_w \log(RV_{t-s_1}) + \beta_M \log(RV_{t-s_2}) + \beta_J \log(1 + J_t) + \epsilon_{t+h}
\]  
(20)

After introducing the so-called shrinkage and microstructure-noise-robust estimator for the significance jump and continuous sample path variation, those have been discussed in the Section 2, Andersen et al. (2007) represented the HAR-RV-CJ model as

\[
RV_{t+h} = \beta_0 + \beta_{CD} C_t + \beta_{CW} C_{t-5,t} + \beta_{CM} C_{t-22,t} + \beta_{DJ} S_J t + \beta_{W} S_{J-t-5,t} + \beta_{JM} S_{J-t-22,t} + \epsilon_{t+h}
\]  
(21)

where,

\[
C_{t+h} = h^{-1}(C_{t+1} + C_{t+2} + \cdots + C_{t+h})\]

and

\[
S_{J,t+h} = h^{-1}(S_{J,t+1} + S_{J,t+2} + \cdots + S_{J,t+h})
\]

The standard deviation and logarithmic version of this model respectively are

\[
(RV_{t+h})^{1/2} = \beta_0 + \beta_{CD} C_t^{1/2} + \beta_{CW} (C_{t-5,t})^{1/2} + \beta_{CM} (C_{t-22,t})^{1/2} + \beta_{DJ} S_J t^{1/2} + \beta_{W} S_{J-t-5,t}^{1/2} + \beta_{JM} S_{J-t-22,t}^{1/2} + \epsilon_{t+h}
\]  
(22)

and

\[
\log(RV_{t+h}) = \beta_0 + \beta_{CD} \log(C_t) + \beta_{CW} \log(C_{t-5,t}) + \beta_{CM} \log(C_{t-22,t}) + \beta_{DJ} \log(1 + S_J t) + \beta_{W} \log(1 + S_{J-t-5,t}) + \beta_{JM} \log(1 + S_{J-t-22,t}) + \epsilon_{t+h}
\]  
(23)

See, Andersen et al. (2003, 2007) for further discussion.

6. Modeling and forecasting RV with HAR-RV and SVM-HAR-RV models

This paper compared the forecasting performance of the SVM-HAR-RV class models with the classical HAR-RV class model. For this comparison, the in-sample period considered from March 11, 1996 to December 29, 2004 and out-of-sample period from January 5, 2005 to September 30, 2009, the period including global financial market crashes. First, estimated model (17) (using RV, standard deviation and logarithm of RV series), (18), (19), (20), (21), (22) and (23) by ordinary least squares (OLS) method and next, by SVM setting the values \( C = 1 \) and \( \epsilon = 0.1 \) to these models and named SVM-HAR-RV models. The R 2.12.0-win32 and R 2.12.0-win32’s Kernlab package were used for both HAR-RV and SVM-HAR-RV class models.

Both class of models for horizons \( h = 1, 5, \) and 22 days were estimated. To compare the forecasting performance, the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Root Mean Square Percentage Error (RMSPE) and Mean Absolute Percentage Error (MAPE) were computed, which defined as follow:

\[
RMSE \equiv \frac{1}{N} \sum_{t=1}^{T} (RV_t - \bar{RV}_{t-[t-1]}),
\]

\[
MAE \equiv \frac{1}{N} \sum_{t=1}^{T} |RV_t - \bar{RV}_{t-[t-1]}|,
\]

\[
RMSPE \equiv \frac{1}{N} \sum_{t=1}^{T} \left( \frac{1 - \bar{RV}_{t-[t-1]}}{RV_t} \right)^2, \quad \text{and} \quad MAPE \equiv \frac{1}{N} \sum_{t=1}^{T} \left| \frac{1 - \bar{RV}_{t-[t-1]}}{RV_t} \right|,
\]

where \( \bar{RV}_{t-[t-1]} \) denotes one-day ahead realized volatility forecast. We evaluate these errors for 5 days ahead and 22-days ahead volatility forecast as well.

To save space, this paper did not include the estimation results of all models. The values of \( R^2 \) for different models are presented in Table 2a and Table 2b while the forecasting errors are presented in Table 3a and 3b.

6.1 Empirical Results

Let us first compare the \( R^2 \) results. It is observed from Table 2a, presents values of \( R^2 \) for different models, that the value of \( R^2 \) successively increases for the standard deviation of RV series than RV series and for the log RV series than standard deviation series for all the models and all different horizons but successively decreases for higher horizons in each and every series. In each series and horizon, the in-sample forecasting performance of SVM-HAR-RV models is remarkably better than HAR-RV models for each and every series and horizon. The out-of-sample forecasting of SVM-HAR-RV models is also higher that the HAR-RV models for standard deviation and logarithmic series. Only the values of out-of-sample \( R^2 \) of HAR-RV model for the RV series are slightly higher than the SVM-HAR-RV. Almost similar results (differ in values) have been observed to compare the forecasting performance of the HAR-RV-J (and MSNR-J) and SVM-HAR-J (and MSNR-J) models. It is also observed for both classes of models that the model performances improved after adding the jump (and/or MSNR-J) component as explanatory variable. The MSNR-Jump remarkably improves the predictive ability for the SVM-HAR-RV-MSNR-J class models but not for the HAR-RV-J class models.
Table 2b presents values of $R^2$ for the HAR-RV-RV-CJ, HAR-RV-RV-MSNR-CJ, SVM-HAR-RV-CJ and SVM-HAR-RV-MSNR-CJ models. This table also produced the similar results as Table 2a. The value of $R^2$ successively increases for the standard deviation of RV series than RV series and for the log RV series than standard deviation series for all the models and all different horizons but successively decreases for higher horizons in each and every series. In the in-sample case, The SVM-HAR-RV-CJ (and/or SVM-HAR-RV-MSNR-CJ) class of models performed well that the HAR-RV-CJ (and/or HAR-RV-MSNR-CJ) class of models. The out-of-sample performances of SVM-HAR-RV class models are also satisfactory.

The logarithmic transformed series produced better performances compared to RV and standard deviation of RV series for both classes of models. For both class of models, the best performances observed when 5-minute intraday returns are used to estimate the realized volatility.

Next, the different errors are calculated for the logarithmic transformed series for both classes of models.

Let us now compare the results based on different above defined error squares. Table 3a represents the forecasting errors for HAR-RV, SVM-HAR-RV, HAR-RV-J, SVM-HAR-RV-J, HAR-RV-MSNR-J and SVM-HAR-RV-MSNR-J models while and 3b presents the forecasting errors for HAR-RV-CJ, SVM-HAR-RV-CJ, HAR-RV-MSNR-CJ and SVM-HAR-RV-MSNR-CJ models. It is observed that in the in-sample case, the SVM-HAR class models completely defeat the HAR-RV class models for every series, horizon and intraday returns series. For the Out-of-sample case, the performance of SVM-HAR class models is also satisfactory compared to HAR-RV class models. Figure 2 presents the out-of-sample forecasting performances of the above models when 5-minute intraday returns are used.

7. Concluding remarks

This paper combined the Support Vector Machine (SVM) regression with Heterogeneous Autoregressive (HAR) model as a hybrid model (SMV-HAR model) to improve the volatility forecasting ability. It is examined the realized volatility forecasting ability of the models for Nikkei 225 stock returns. The empirical results presented here are suggestive for several interesting extensions. First, the values $C = 1$ and $\epsilon = 0.1$ for the SVM-HAR-RV class models were set and observed better forecasting ability. The appropriate choice of the value $C$ and $\epsilon$ could be helpful to improve the forecasting ability.

Second, the Polynomial and Laplaceian kernel were considered for the SVMs and observed better performances. The appropriate choice of other existing kernels in SVM literature or an appropriate new kernel could improve the forecasting ability.

Third, the optimally sampled sampling frequencies are considered to mitigate the market microstructure noise. This choice failed to improve the forecasting performances. It would be interesting to consider the other market microstructure noise mitigation techniques.

Those topics are left for further research.

References


Table 1a. Summary Statistics for Nikkei 225 Daily Returns, Realized Volatility and Jumps

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RV</td>
<td>-0.024</td>
<td>1.621</td>
<td>-0.163</td>
<td>0.831</td>
<td>-11.923</td>
<td>939885.37</td>
<td>13.589</td>
<td>24.205</td>
<td>41.57</td>
<td></td>
</tr>
<tr>
<td>ln (RV)</td>
<td>0.496</td>
<td>0.872</td>
<td>0.117</td>
<td>1.365</td>
<td>-2.614</td>
<td>4.246</td>
<td>58.760</td>
<td>71.102</td>
<td>203.17</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>1.471</td>
<td>1.700</td>
<td>6.542</td>
<td>37.139</td>
<td>0.066</td>
<td>26.163</td>
<td>667585.43</td>
<td>5247.666</td>
<td>8861.280</td>
<td>14596.52</td>
</tr>
<tr>
<td>ln (J)</td>
<td>1.111</td>
<td>0.486</td>
<td>2.082</td>
<td>12.789</td>
<td>0.263</td>
<td>5.115</td>
<td>15350.645</td>
<td>6597.039</td>
<td>114877.356</td>
<td>20901.50</td>
</tr>
<tr>
<td>J^2</td>
<td>0.792</td>
<td>0.426</td>
<td>1.262</td>
<td>6.220</td>
<td>0.262</td>
<td>3.302</td>
<td>22705.338</td>
<td>66996.961</td>
<td>116743.345</td>
<td>20728.59</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>1.504</td>
<td>1.760</td>
<td>6.149</td>
<td>36.888</td>
<td>0.063</td>
<td>28.759</td>
<td>503835.151</td>
<td>5387.071</td>
<td>92048.210</td>
<td>15255.27</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>1.120</td>
<td>0.499</td>
<td>0.263</td>
<td>12.121</td>
<td>0.252</td>
<td>5.173</td>
<td>13595.248</td>
<td>6458.825</td>
<td>11335.073</td>
<td>19988.77</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>0.800</td>
<td>0.436</td>
<td>1.273</td>
<td>6.090</td>
<td>0.232</td>
<td>3.214</td>
<td>6747.022</td>
<td>11425.650</td>
<td>20373.58</td>
<td></td>
</tr>
</tbody>
</table>

Optimally Sampled

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RV</td>
<td>2.621</td>
<td>4.014</td>
<td>7.868</td>
<td>96.357</td>
<td>0.063</td>
<td>69.809</td>
<td>1216009.157</td>
<td>4754.751</td>
<td>79584.077</td>
<td>12526.44</td>
</tr>
<tr>
<td>ln (RV)</td>
<td>1.431</td>
<td>0.757</td>
<td>2.600</td>
<td>16.160</td>
<td>0.200</td>
<td>8.355</td>
<td>271646.726</td>
<td>5830.113</td>
<td>1011552.102</td>
<td>17369.59</td>
</tr>
<tr>
<td>J</td>
<td>1.434</td>
<td>2.008</td>
<td>6.658</td>
<td>70.959</td>
<td>0.036</td>
<td>36.743</td>
<td>650619.489</td>
<td>3818.668</td>
<td>6281.099</td>
<td>10176.56</td>
</tr>
<tr>
<td>ln (J)</td>
<td>1.045</td>
<td>0.585</td>
<td>2.367</td>
<td>13.050</td>
<td>0.189</td>
<td>6.062</td>
<td>167417.888</td>
<td>4596.683</td>
<td>7995.894</td>
<td>13779.03</td>
</tr>
<tr>
<td>J^2</td>
<td>0.724</td>
<td>0.496</td>
<td>1.558</td>
<td>6.599</td>
<td>0.035</td>
<td>3.631</td>
<td>2988.122</td>
<td>4549.871</td>
<td>7996.372</td>
<td>13995.17</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>1.448</td>
<td>2.562</td>
<td>8.836</td>
<td>122.242</td>
<td>0.025</td>
<td>54.221</td>
<td>1971380.206</td>
<td>3640.938</td>
<td>67074.458</td>
<td>10963.37</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>1.036</td>
<td>0.611</td>
<td>2.776</td>
<td>17.646</td>
<td>0.123</td>
<td>7.240</td>
<td>332529.970</td>
<td>4459.291</td>
<td>7767.385</td>
<td>13235.28</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>0.714</td>
<td>0.507</td>
<td>1.678</td>
<td>7.266</td>
<td>0.035</td>
<td>3.978</td>
<td>3979.658</td>
<td>4730.037</td>
<td>7697.203</td>
<td>13737.17</td>
</tr>
</tbody>
</table>

Key: The sample of the period 11 March 1996 to 30 September 2009, there are total 3281 Daily observations. The 5% critical values for Jarque-Bera (i.e., χ² (k)) and LB (k) are 5.991 (k=2), 11.070 (k=5), 18.924 (k=10) and 33.924 (k=22) respectively.

Table 1b. Summary Statistics for Nikkei 225 Continuous Path Components and Significant Jumps

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RV</td>
<td>1.231</td>
<td>2.209</td>
<td>12.430</td>
<td>237.994</td>
<td>0.051</td>
<td>59.537</td>
<td>7576674.613</td>
<td>3346.519</td>
<td>5690.879</td>
<td>9141.11</td>
</tr>
<tr>
<td>ln (RV)</td>
<td>0.396</td>
<td>0.927</td>
<td>0.117</td>
<td>1.365</td>
<td>-2.614</td>
<td>4.246</td>
<td>58.760</td>
<td>71.102</td>
<td>203.17</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>1.434</td>
<td>2.008</td>
<td>6.658</td>
<td>70.959</td>
<td>0.036</td>
<td>36.743</td>
<td>650619.489</td>
<td>3818.668</td>
<td>6281.099</td>
<td>10176.56</td>
</tr>
<tr>
<td>ln (J)</td>
<td>1.045</td>
<td>0.585</td>
<td>2.367</td>
<td>13.050</td>
<td>0.189</td>
<td>6.062</td>
<td>167417.888</td>
<td>4596.683</td>
<td>7995.894</td>
<td>13779.03</td>
</tr>
<tr>
<td>J^2</td>
<td>0.724</td>
<td>0.496</td>
<td>1.558</td>
<td>6.599</td>
<td>0.035</td>
<td>3.631</td>
<td>2988.122</td>
<td>4549.871</td>
<td>7996.372</td>
<td>13995.17</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>1.448</td>
<td>2.562</td>
<td>8.836</td>
<td>122.242</td>
<td>0.025</td>
<td>54.221</td>
<td>1971380.206</td>
<td>3640.938</td>
<td>67074.458</td>
<td>10963.37</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>1.036</td>
<td>0.611</td>
<td>2.776</td>
<td>17.646</td>
<td>0.123</td>
<td>7.240</td>
<td>332529.970</td>
<td>4459.291</td>
<td>7767.385</td>
<td>13235.28</td>
</tr>
<tr>
<td>ln (J^2)</td>
<td>0.714</td>
<td>0.507</td>
<td>1.678</td>
<td>7.266</td>
<td>0.035</td>
<td>3.978</td>
<td>3979.658</td>
<td>4730.037</td>
<td>7697.203</td>
<td>13737.17</td>
</tr>
</tbody>
</table>

Key: The sample of the period 11 March 1996 to 30 September 2009, there are total 3281 Daily observations. For continuous sample path variation and significant jump measures, we set ϵ = 0.999. The 5% critical values for Jarque-Bera (i.e., χ² (k)) and LB (k) are 5.991 (k=2), 11.070 (k=5), 18.924 (k=10) and 33.924 (k=22) respectively.

Published by Canadian Center of Science and Education
Table 2a. The $R^2$-Values for HAR, SVM-HAR, HAR-J, SVM-HAR-J, HAR-MSNR-J and SVM-HAR-MSNR-J models

<table>
<thead>
<tr>
<th>Horizon Days</th>
<th>HAR</th>
<th>SVM-HAR</th>
<th>HAR-J</th>
<th>SVM-HAR-J</th>
<th>HAR-MSNR-J</th>
<th>SVM-HAR-MSNR-J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>HAR</td>
<td>SVM-HAR</td>
<td>HAR-J</td>
<td>SVM-HAR-J</td>
<td>HAR-MSNR-J</td>
<td>SVM-HAR-MSNR-J</td>
</tr>
<tr>
<td>RV</td>
<td>0.59</td>
<td>0.52</td>
<td>0.57</td>
<td>0.59</td>
<td>0.53</td>
<td>0.58</td>
</tr>
<tr>
<td>lnRV</td>
<td>0.48</td>
<td>0.44</td>
<td>0.45</td>
<td>0.47</td>
<td>0.4</td>
<td>0.47</td>
</tr>
<tr>
<td>In-Sample</td>
<td>0.39</td>
<td>0.36</td>
<td>0.39</td>
<td>0.40</td>
<td>0.38</td>
<td>0.39</td>
</tr>
<tr>
<td>Out-of-sample</td>
<td>0.38</td>
<td>0.35</td>
<td>0.37</td>
<td>0.40</td>
<td>0.38</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Key: The sample of the period 11 March 1996 to 30 September 2009, there are total 3279 Daily Observations.

Table 2b. The $R^2$-Values for HAR-CJ, SVM-HAR-CJ, HAR-MNR-CJ and SVM-HAR-MNR-CJ models for different horizons

<table>
<thead>
<tr>
<th>Horizon Days</th>
<th>HAR</th>
<th>SVM-HAR</th>
<th>HAR-MNR</th>
<th>SVM-HAR-MNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>HAR</td>
<td>SVM-HAR</td>
<td>HAR-MNR</td>
<td>SVM-HAR-MNR</td>
</tr>
<tr>
<td>RV</td>
<td>0.59</td>
<td>0.52</td>
<td>0.59</td>
<td>0.58</td>
</tr>
<tr>
<td>lnRV</td>
<td>0.48</td>
<td>0.44</td>
<td>0.47</td>
<td>0.46</td>
</tr>
<tr>
<td>In-Sample</td>
<td>0.39</td>
<td>0.36</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>Out-of-sample</td>
<td>0.38</td>
<td>0.35</td>
<td>0.37</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Key: The sample of the period 11 March 1996 to 30 September 2009, there are total 3279 Daily Observations.

Table 3a. Forecasting Errors of HAR, SVM-HAR, HAR-J, SVM-HAR-J, HAR-MSNR-J and SVM-HAR-MSNR-J models for the logarithmic series

<table>
<thead>
<tr>
<th>Horizon Days</th>
<th>HAR</th>
<th>SVM-HAR</th>
<th>HAR-J</th>
<th>SVM-HAR-J</th>
<th>HAR-MSNR-J</th>
<th>SVM-HAR-MSNR-J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>HAR</td>
<td>SVM-HAR</td>
<td>HAR-J</td>
<td>SVM-HAR-J</td>
<td>HAR-MSNR-J</td>
<td>SVM-HAR-MSNR-J</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.89</td>
<td>0.80</td>
<td>0.82</td>
<td>0.83</td>
<td>0.84</td>
<td>0.87</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.66</td>
<td>0.61</td>
<td>0.64</td>
<td>0.65</td>
<td>0.66</td>
<td>0.68</td>
</tr>
<tr>
<td>MAPE%</td>
<td>0.54</td>
<td>0.50</td>
<td>0.53</td>
<td>0.54</td>
<td>0.55</td>
<td>0.57</td>
</tr>
<tr>
<td>MAPE%</td>
<td>0.47</td>
<td>0.43</td>
<td>0.45</td>
<td>0.46</td>
<td>0.47</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Key: The sample of the period 11 March 1996 to 30 September 2009, there are total 3279 Daily Observations.

Table 3b. The $R^2$-Values those have been calculated for daily (h=1), weekly (h=5) and monthly (h=22) horizons.

<table>
<thead>
<tr>
<th>Horizon Days</th>
<th>HAR</th>
<th>SVM-HAR</th>
<th>HAR-MNR</th>
<th>SVM-HAR-MNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>HAR</td>
<td>SVM-HAR</td>
<td>HAR-MNR</td>
<td>SVM-HAR-MNR</td>
</tr>
<tr>
<td>RV</td>
<td>0.59</td>
<td>0.52</td>
<td>0.59</td>
<td>0.58</td>
</tr>
<tr>
<td>lnRV</td>
<td>0.48</td>
<td>0.44</td>
<td>0.47</td>
<td>0.46</td>
</tr>
<tr>
<td>In-Sample</td>
<td>0.39</td>
<td>0.36</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>Out-of-sample</td>
<td>0.38</td>
<td>0.35</td>
<td>0.37</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Key: The sample of the period 11 March 1996 to 30 September 2009, there are total 3279 Daily Observations.

Table 4. The $R^2$-Values those have been calculated for daily (h=1), weekly (h=5) and monthly (h=22) horizons.

<table>
<thead>
<tr>
<th>Horizon Days</th>
<th>HAR</th>
<th>SVM-HAR</th>
<th>HAR-MNR</th>
<th>SVM-HAR-MNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>HAR</td>
<td>SVM-HAR</td>
<td>HAR-MNR</td>
<td>SVM-HAR-MNR</td>
</tr>
<tr>
<td>RV</td>
<td>0.59</td>
<td>0.52</td>
<td>0.59</td>
<td>0.58</td>
</tr>
<tr>
<td>lnRV</td>
<td>0.48</td>
<td>0.44</td>
<td>0.47</td>
<td>0.46</td>
</tr>
<tr>
<td>In-Sample</td>
<td>0.39</td>
<td>0.36</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>Out-of-sample</td>
<td>0.38</td>
<td>0.35</td>
<td>0.37</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Key: The sample of the period 11 March 1996 to 30 September 2009, there are total 3279 Daily Observations.
Table 3b. Forecasting Errors of HAR-CJ, SVM-HAR-CJ, HAR-MNR-CJ and SVM-HAR-MNR-CJ models for different horizons for the logarithmic series

<table>
<thead>
<tr>
<th>Horizon Days(s)</th>
<th>Model</th>
<th>In-Sample</th>
<th>Opt</th>
<th>Out-of-sample</th>
<th>Opt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HAR</td>
<td>SVM-HAR</td>
<td></td>
<td>HAR</td>
<td>SVM-HAR</td>
</tr>
<tr>
<td></td>
<td>C3</td>
<td>MNR-CJ</td>
<td>C3</td>
<td>MNR-CJ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.401</td>
<td>0.401</td>
<td>0.451</td>
<td>0.454</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>0.548</td>
<td>0.548</td>
<td>0.646</td>
<td>0.645</td>
</tr>
<tr>
<td></td>
<td>RMSPE</td>
<td>0.527</td>
<td>0.527</td>
<td>0.554</td>
<td>0.556</td>
</tr>
<tr>
<td></td>
<td>RMSPE</td>
<td>0.590</td>
<td>0.590</td>
<td>0.552</td>
<td>0.556</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>1.276</td>
<td>1.292</td>
<td>1.356</td>
<td>1.377</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.421</td>
<td>0.422</td>
<td>0.317</td>
<td>0.321</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>0.590</td>
<td>0.590</td>
<td>0.599</td>
<td>0.603</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>0.590</td>
<td>0.590</td>
<td>0.552</td>
<td>0.556</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>1.276</td>
<td>1.292</td>
<td>1.356</td>
<td>1.377</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.421</td>
<td>0.422</td>
<td>0.317</td>
<td>0.321</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>0.590</td>
<td>0.590</td>
<td>0.599</td>
<td>0.603</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>0.590</td>
<td>0.590</td>
<td>0.552</td>
<td>0.556</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>1.276</td>
<td>1.292</td>
<td>1.356</td>
<td>1.377</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>0.421</td>
<td>0.422</td>
<td>0.317</td>
<td>0.321</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>0.590</td>
<td>0.590</td>
<td>0.599</td>
<td>0.603</td>
</tr>
</tbody>
</table>

Key: The sample of the period 11 March 1996 to 30 September 2009, there are total 3279 Daily observations. The Table reports the $R^2$-Values those have been calculated for daily (h=1), weekly (h=5) and monthly (h=22) horizons.

Figure 1. Realized Volatility, Log-Realized Volatility, Jumps and Significant Jumps series.
The first panels show the HAR-RV/SVM-HAR-RV, the second is HAR-RV-J/SVM-HAR-RV-J, the third is HAR-RV-MSNR-J/SVM-HAR-RV-MSNR-J, the fourth is HAR-RV-CJ/SVM-HAR-RV-CJ and the fifth is HAR-RV-MSNR-CJ/SVM-HAR-RV-MSNR-CJ model’s out-of-sample forecasts for the logarithmic transformed series.

Figure 2. Daily, Weekly and Monthly out-of-sample realized volatility forecasts from HAR-RV, SMV-HAR-RV-R, HAR-RV-J and SVM-HAR-RV-J models