Prediction of FX Volatility via an RBF Neural Network with Closed-Form Regularization

By

Gordon H. Dash, Jr. *
Associate Professor of Finance and Insurance

Nina Kajiji **
Assistant Professor of Research, Education

JEL CATEGORY

E40 MONEY DEMAND/INTEREST RATES; E47 Forecasting and Simulation
F30 INTERNATIONAL FINANCE; F31 Foreign Exchange
C10 ECONOMETRIC STATISTICS; C5 Econometric Modeling

For presentation at the 8th International Conference on Forecasting Financial Markets, London, England, 31 May -01 June, 2001.
Please address all correspondence to: *Gordon H. Dash, Jr., Department of Finance and Insurance, College of Business Administration, 7 Lippitt Road, University of Rhode Island, Kingston, RI 02881. Fax: (888) 298 7299 Email: GHDash@uri.edu .

**Department of Education, University of Rhode Island Kingston, Kingston, RI 02881.

The work reported in the paper was supported by a grant from The NKD Group, Inc. Wilmington, Delaware.

Abstract

Volatility modeling is a key to performance enhanced strategies in the derivative- and asset-pricing evaluation process. As such, it is understandable that a voluminous literature has evolved to discuss the temporal dependencies in financial market volatility. Much of this literature has been directed at daily and lower frequencies using ARCH and stochastic volatility type models. With access to high frequency and ultra high-frequency databases, more recent research has been able to explain about fifty percent of the interdaily forecasts of latent volatility. Relying upon hourly intervals, the GARCH(1,1) results presented here are consistent with prior studies. However, this paper adds to the tools available for conducting volatility exploration by introducing an adaptive radial basis function neural network that significantly lowers overall prediction error while maintaining a high explanatory ratio. The newly formulated RBF implements a closed-form regularization parameter with prior information. It is an algorithmic extension that will permit more accurate and insightful analyses to be performed on high frequency financial time series.

Over the past decade, research efforts increased significantly in the area of modeling volatility behavior in capital market high frequency data. Obtaining accurate volatility forecasts is a key part of the process by which derivative-pricing and asset-pricing models are evaluated. But, the modeling process is challenged by the complex patterns observed in many high-frequency financial time series. Conventional wisdom suggests that it is the continuous arrival of macro-economic news, return volatility, and overall market oscillations that shape these complex time-series. These factors are magnified by the fact that continuous-time data does not arrive at equally spaced time intervals. Faced with investment strategies that require modeling accuracy, the financial economist has an unyielding need to uncover prudent methods that best explain high-frequency (and now ultra high-frequency) data series. Ghysels, Harvey, and Renault (1996) provide recent evidence of this large-scale effort. One of the more popular strategic frameworks to surface in recent research on time-series volatility is the generalized autoregressive conditional heteroscedasticity framework (GARCH) of Bollerslev (1986). Another line of exploration that has proven fruitful has come via the utilization of artificial neural nets (ANN).

In this paper we focus our efforts on the explanatory power of a radial basis function (RBF) ANN when applied to high frequency volatility. We apply the RBF-ANN to the problem of measuring the information content in the one-day ahead volatility of foreign exchange futures-options. The RBF explored in this research is augmented by the inclusion of a closed-form solution for the estimate of the regularization parameter. Additionally, the RBF-ANN includes a Bayesian derived information set. The goal of these extensions is, of course, to produce a more accurate modeling framework.

The plan of the paper is as follows. Section 1 presents notation and introduces the sample data set. Section 2 focuses on modeling financial time-series under the GARCH framework. The results of the GARCH model applied to the sample data are presented in section 3. The RBF-ANN is introduced in Section 4. Section 5 concludes the paper by providing a comparison of the GARCH findings with those obtained from solving the RBF-ANN.

1. Notation and Data

The models presented in this paper are based on hourly returns obtained from closing quotes on the dollar exchange with the German deutsche mark (DM), Japanese Yen (JY), and the Swiss Franc (SF) as traded on the Chicago Mercantile Exchange (CME). Tick observations on currency futures options data are obtained from the Futures Industry Association while closing tick-quotes for futures contracts are obtained from Tick Data, Inc (Tick Data).[1] In addition to the currency-related data, high-frequency futures data on the U.S. Treasury Bill (TB), the dollar index (DX), and the U.S. Treasury Bond (US) are also obtained from Tick Data. The sample period for the DM extends from January 4, 1999 to August 06, 1999. For both the JY and the SF the sample period is from January 4, 1999 to December 31, 1999. The tick observations are aggregated into equally spaced intervals of one-hour beginning with the 9:00 a.m. closing quote. The last quote of the day is captured with the 1:59 p.m. trade. This results in 750 observations for the DM and 1,248 observations for all other contracts.

1.1 Hourly Risk-Free Rate

The continuously compounded risk-free rate is derived from the tick Futures contract (Fi) on the 90-day T-Bill (TB). First, we compute the price per $100 of par value (P100):

$$P_{100} = \$100 - \left(\$100 - F_i\right)\left(\frac{91}{360}\right). \tag{1}$$

Next the yield or the risk-free rate on the 91-day T-bill is computed as.

$$yield = \left(\frac{\$100}{P_{100}}\right)\left(\frac{365}{91}\right) - 1 \tag{2}$$

which is then subjected to a log transformation to produce the continuously compounded risk-free rate:

$$r_c = \ln(1 + yield) \tag{3}$$

[1] See http://www.fiafii.org and http://www.tickdata.com , respectively.

1.2 Hourly Futures Returns

The first one-hour period begins at 9:00 am, one-half hour after the market opens at 8:30 a.m. The last observation in the transformed data set ends at 1:59 p.m., one-hour before the market closing at 3:00 p.m. Both the first one-half hour and the last hour of trading are excluded in order to eliminate daily anomalies. Trade prices beginning with the trade closest to the 9:00 a.m. hour were used to compute rates of return on all futures contracts. Rates of return were derived as follows,

$$r_t = \ln(F_t) - \ln(F_{t-1}) \tag{4}$$

Where: r_i = i[th] period rate of return
F_i = Futures contract closing price for the i[th] period

To create an equally spaced hourly observation a geometric average of the within hour rates of returns was computed across the daily hours of trading. Stated differently, the target variable is a geometric average of the within hour trade returns which is derived by,

$$r_t = \left(\left(\prod_{i=1}^{n} (1 + r_i) \right)^{1/n} - 1 \right) 100 \tag{5}$$

Where:
r_t = t[th] hour geometric return in percentage format
r_i = i[th] period rate of return
n = number of trades in an hour

The option contract ISD is matched with the hourly sequence of continuously compounded futures contract returns until option expiration. The future volatility (or realized volatility) is obtained from the variability of these returns. Annualized volatility is measured by multiplying future volatility by 79.3725 (the square root of 252 multiplied by 5).

2. Modeling Futures Options

2.1 ISD and Observation Timing

Estimates of implied volatilities are calculated hourly for call options with 60 or fewer days to expiration. All implied volatilities are derived from the Black (1976) model for European options on futures,

$$C = e^{-r_c T}[FN(d_1) - EN(d_2)], \qquad (6)$$

Where:

$$d_1 = \frac{\ln(F/E) + \left(s^2 \big/ 2 \right) T}{s_X \sqrt{T}}, \text{ and}$$

$$d_2 = d_1 - s\sqrt{T}$$

In the above equation F is the futures rate , E is the exercise price, T is the time to option expiration, σ the volatility, and r is the risk-free rate.[2] Following Beckers (1981) we invert closest at-the-money calls.[3] When using tick-data there is always a possibility of non-simultaneous trades. As Jorion (1995) reports, measurement errors can substantially distort inferences on daily data. In this study we compute the hourly ISD based on the geometric average of calls traded within the time frame that begins on the hour and terminates 59 minutes past the hour.

---

[2] CME options are American style and those pose a small inconsistency with the Black-Scholes model. As shown in Jorion (1995) using a model based on European style options tends to overestimate the true volatility of the option by approximately 12 percent. As with Jorion, we consider this overestimation as inconsequential.

[3] It is well known that for out-of-the-money options as strike prices increase implied volatility increases. Conversely, it can be shown that in-the-money calls are less expensive than Black-Scholes theory predicts.

2.2 GARCH

The GARCH model was first developed to model data at the daily frequency level or greater. Because volatility persistence is known to exist in high frequency data as well, there is a keen interest in applying GARCH methods to short-term volatility questions. Research findings of how well the simple GARCH model is able to reproduce heteroscedastic behavior in high frequency data is mixed. Several studies are not supportive of the model when applied to high frequency data (see Andersen and Bollerslev (1994); Guillaume et al. (1994); Ghose and Kroner (1995); and, Dacorogna et al. (1998). Specifically, the consensus finding of these studies suggest that when high frequency data is modeled by GARCH, volatility memory is short-lived and weakly explained by ex-post squared returns. Conversely, daily (or lower) data displays a long-lived volatility memory. Andersen and Bollerslev (1997) address this apparent conflict. They show that standard GARCH models are capable of predicting close to fifty percent of the variance in the latent one-day ahead volatility factors. These results were achieved within a continuous-time stochastic volatility framework that allowed for the construction of a new ex-post volatility measurement that is based upon cumulative squared intra-day returns.[4]

The weak-form GARCH model of Bollerslev (1986) generalized the original autoregressive conditional heteroscedasticity (ARCH) model of Engle (1982). For a time series variable x_t, the model is expressed as:

$$x_t = s_t z_t \tag{7}$$

where: $$s_t^2 = a_0 + a_1 x_{t-1}^2 + b_1 s_{t-1}^2 \tag{8}$$

[4] For more comprehensive surveys on the subject of the GARCH framework see, for instance, Diebold and Lopez Diebold, F. X. and J. A. Lopez (1995). Modeling Volatility Dynamics. Macroeconometrics: Developments, Tensions, and Prospects. K. D. Hoover. Boston, Kluwer Academic Publishers., Bollerslev, Engle and Nelson Bollerslev, T., R. F. Engle, et al. (1994). ARCH Models. Handbook of Econometrics. R. F. a. M. Engle, D.L. Amsterdam, Elsevier Science., Bollerslev, Chou and Kroner Bollerslev, T., R. Y. Chou, et al. (1992). "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence." Journal of Econometrics **52**: 5-59., and Bond Bond, S. A. (2000). Asymmetry and Downside Risk in Foreign Exchange Markets. Department of Land Economy. U.K., Cambridge, University of Cambridge**:** 51..

and $z_t \sim NID(0,1)$, for $a_0, a_1 \geq 0$ and $t = 1...T$. The model implies that $x_t \mid \Omega_{t-1} \sim N(0, s^2_{t-1})$.[5] The model is particularly interesting in financial research as the model permits x_t to be leptokurtotic and can capture seasonality ('volatility clustering') that is known to characterize financial data.

3. Analysis

In this section we focus on describing real-time market dynamics within the context of the hourly returns. Our purpose is to provide a limited comparison of the weak-form GARCH framework applied to hourly data with the established results obtained from application to daily data. The analysis presented below focuses on modeling variation in return volatility. While it is of interest, this study does not attempt to develop findings regarding the long-term memory features of volatility. It is our belief that "memory" is best investigated with higher-frequency data. The modeling process is further challenged by factors that are symptomatic of high frequency data. These factors include, but are not limited to intraday seasonality, macroeconomic announcements, reaction to news releases, and errors that may arise in the information transmission and recording process. The effects of these low-frequency impacts are the subjects of continuing research. We proceed by generating descriptive statistics, followed by an analysis of returns and volatility. This section concludes by reporting the information content in ISD and conditional volatility as well as the prediction capability of these same measures.

3.1 Descriptive Statistics

Table I provides descriptive statistics of the three currency contracts. These statistics include the one-hour return, volatility, and ISD. The one-hour ISD is taken from an average hourly futures price as applied to the nearest at-the-money futures option within hour. At any hour the "realized" (future) volatility is the average of the contract's

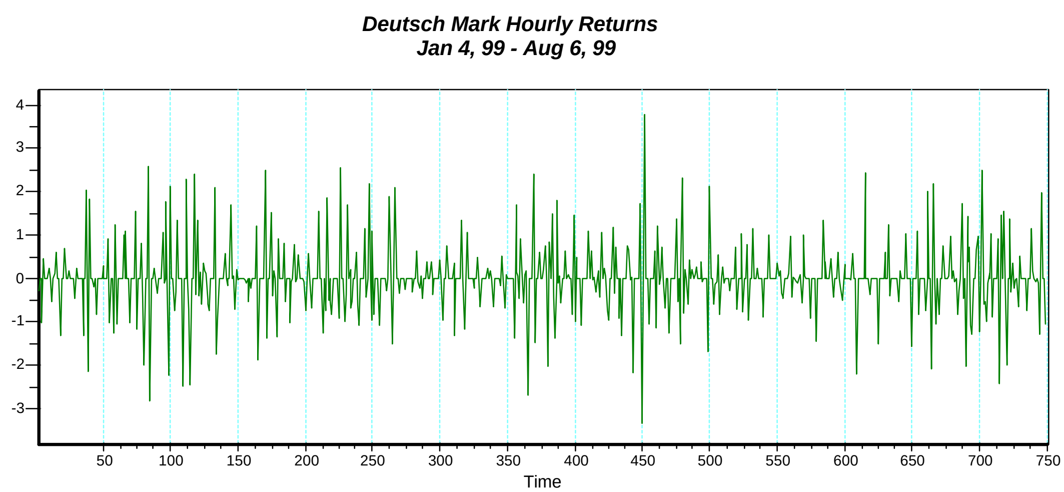
[5] Other non-normal conditional distributions have been used in the model specification. By way of example, see Nelson 1991.

volatility to expiration up to the hour of expiration.[6] Annualized volatility measures are obtained by multiplying the square root of 252 by 5 hours. Interest rates are specified in hourly terms. All contracts follow the March-June-September-December cycle. Rolling over expiring contracts into the nearest-at-the-money contract in the next expiration month creates a continuous contract of hourly returns, implied- and realized-volatilities.

3.1.1 Returns

The hourly returns for the three FX contracts follow a pattern that is well documented. The hourly standard deviations are 0.7375, 0.9763, and 0.8343 percent, respectively. These findings are similar to those reported by Jorion (1995) on daily data with one notable difference. For the hourly results, the JY has higher standard deviation than the other two contracts. But, as in the case of daily findings, the standard deviation for the DM and SF are about the same. In all three cases the hourly mean return is negative. Although the JY is more volatile, its mean return is no different than that of the SF. Both the SF and JY produce a return that is higher than the less volatile DM. Figures 1a, 1b, and 1c display the hourly returns for the DM, JY and SF, respectively.

Figure 1a



[6] Following Jorion 1995 and as shown by Baillie and Bollerslev 1989 we know that the weekend variance must be higher than during the weekday. However Jorion has shown that for daily observations the difference between the weekend and weekday variance is small. Hence, we do not distinguish between the two.

Table I: Descriptive Statistics

	Descriptive Statistics		Autocorrelations								
	Mean	Std. Dev	Lag1	Lag2	Lag3	Lag4	Lag5	Lag10	Lag20	Lag100	Lag250
DM											
1-hour return	-0.0011	0.7375	-0.216	-0.092	-0.156	-0.047	-0.064	0.111	0.035	-0.082	0.083
1-hour volatility (%)	0.1726	0.4090	0.179	0.031	0.031	0.017	-0.016	-0.023	0.079	-0.023	-0.010
ISD Volatility (%)	0.7704	0.7306	0.571	0.370	0.219	0.143	0.181	0.087	-0.008	0.019	0.035
Conditional Volatility	0.5491	0.2321	0.872	0.759	0.657	0.555	0.476	0.224	0.128	-0.100	0.072
JY											
1-hour return	-0.0008	0.9763	-0.342	-0.124	-0.011	-0.058	0.094	0.058	0.056	0.089	0.043
1-hour volatility (%)	1.4429	8.7089	0.002	0.007	0.107	0.005	0.008	0.007	0.019	-0.007	-0.004
ISD Volatility (%)	1.5485	1.3246	0.425	0.396	0.373	0.369	0.408	0.377	0.285	0.047	-0.026
Conditional Volatility	0.9707	0.4712	0.991	0.978	0.963	0.944	0.923	0.813	0.612	0.019	0.032
SF											
1-hour return	-0.0008	0.8343	-0.324	-0.092	-0.083	0.012	0.004	0.017	0.043	-0.030	-0.010
1-hour volatility (%)	0.4653	1.3982	0.193	-0.003	0.060	0.036	0.062	0.102	0.026	0.007	-0.016
ISD Volatility (%)	0.8453	0.7105	0.557	0.393	0.291	0.284	0.231	0.264	0.179	-0.057	0.001
Conditional Volatility	0.7088	0.5115	0.990	0.976	0.960	0.940	0.920	0.832	0.635	0.027	-0.133

Figure 1b

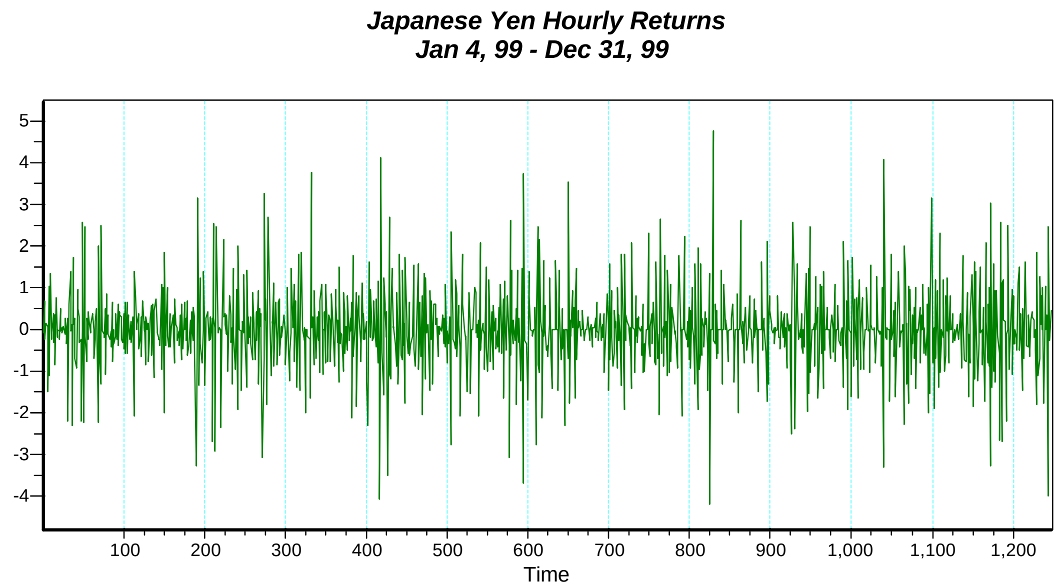
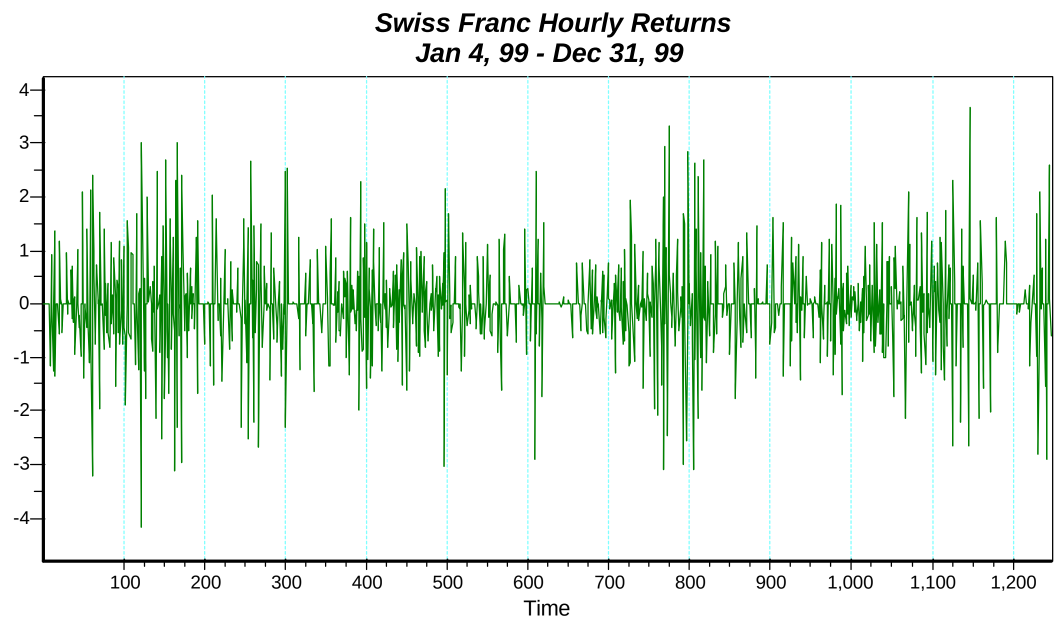


Figure 1c



3.1.2 Volatility

Volatility findings are consistent with those provided on daily data. The observed autocorrelation in the hourly realized (future) volatility provides strong evidence of volatility persistence. Autocorrelation is significantly positive in lag 1 for all contracts and remains positive up through lag 4. For the JY and the SF autocorrelation remains significant up to lag 20. Although the lags are small, the longer-term persistence in positive lags for both the JY and SF are quite evident.

The implied volatility findings provide further insight into the autocorrelation structure. Clearly, the JY was the most volatile of the three currencies over the test period. The standard deviations of the implied volatility for the both the DM and the SF are nearly identical. The analysis is more striking upon the comparison of the mean values for one-hour volatility and ISD volatility. The JY mean values are nearly the same. Stated differently, for all currencies the mean ISD volatility is higher than the mean of the one-hour volatility. The results reported here do not immediately lead to any conclusions about the robustness of the ISD estimate. That is, there is no evidence of consistent over- or under-representation of the realized future volatility.

Figures 1, 2 and 3 display the volatility patterns for the DM, JY, and SF, respectively. Each figure shows time variation in ISD, realized future volatility, and conditional volatility measured in percent per hour. Time variation is obvious in both the ISD and future volatility across all country denominated series. In the next section we present GARCH estimates for each time series.

3.2 GARCH Estimate

Using hourly returns data the GARCH(1,1) is estimated. Parameters α_1 and β were subjected to the typical stationarity constraint. This constraint is necessary and sufficient to examine a finite, time-independent variance of the innovations process. The reported χ^2 statistic of the GARCH estimate confirms that a GARCH process is describing a statistically significant amount of the conditional variance in returns.[7] The results presented in Table II of this study are consistent with GARCH(1,1) results applied to daily data as reported in all prior research.[8] Realized hourly volatility is stationary, but it does change over time.

Table II: Estimation of GARCH(1,1) Process

Currency	**Model**	μ	α_0	α_1	β	**Long-Run Volatility (% pa)**	**Log-Likelihood**	**Pr > χ^2**
DM	GARCH	0.0046	0.0568	0.0850	0.8128	26.41	-814.63	0.0001
	(pr)	(0.8376)	(0.0001)	(0.0001)	(0.0001)			
JY	GARCH	-0.0192	0.2371	0.1603	0.5986	35.17	-1704.52	0.0001
	(pr)	(0.4592)	(0.0001)	(0.0001)	(0.0001)			
SF	GARCH	-0.0108	0.0356	0.1233	0.8296	30.75	-1444.49	0.0001
	(pr)	(0.5528)	(0.0001)	(0.0001)	(0.0001)			

(pr) = p-value

3.3 Information Content

We turn our attention to role of the hourly ISD in forecasting next period volatility. The following regression models examine the information content of ISD:

$$\sqrt{R_{t+1}^2} = a + bs^{ISD} + e_{t+1} \tag{9}$$

$$\sqrt{R_{t+1}^2} = a + bs^{GARCH} + e_{t+1} \tag{10}$$

$$\sqrt{R_{t+1}^2} = a + bs^{ISD} + cs^{GARCH} + e_{t+1} \tag{11}$$

[7] It is well known that daily rates are correlated. The pairwise correlations in the hourly futures prices changes do not show a significant correlation profile. The correlations are: DM/JY, 0.0022; DM/SF, 0.0540; and, JY/SF, 0.0206. The pairwise correlation p-values are: DM/JY, 0.9522; DM/SF, 0.1397; and, JY/SF, 0.4671. Despite these findings, we caution the reader to exercise caution when comparing regression-based results.

[8] OLS, Hansen-White, and GARCH estimates are obtained from version 8e of the Statistical Analysis System (SAS), North Carolina.

Our expectation for each model is the slope parameter for each equation. Our objective is to identify whether s^{ISD} and s^{GARCH} explain meaningful information in the future volatility. We expect the parameters to be statistically significant and non-zero in all cases.

Table III: Information Content Regression

Currency	Slopes on			R^2	MSE
	Intercept	ISD	GARCH		
DM - (9a)	0.4323 (0.0001)	-0.0548 (0.0806)		0.0041	0.3912
(9b)	0.2754 (0.0001)		0.2088 (0.0343)	0.0060	0.3905
(9c)	0.3146 (0.0001)	-0.0591 (0.0592)	0.2204 (0.0255)	0.0107	0.3892
JY - (10-a)	0.5975 (0.0001)	0.0341 (0.0285)		0.0038	0.5290
(10b)	0.5689 (0.0001)		0.0840 (0.0553)	0.0029	0.5295
(10c)	0.5131 (0.0001)	0.0348 (0.0256)	0.0860 (0.0494)	0.0069	0.5278
SF – (11a)	0.5621 (0.0001)	-0.0367 (0.0158)		0.0047	0.4363
(11b)	0.3301 (0.0001)		0.2517 (0.0001)	0.0378	0.4218
(11c)	0.3849 (0.0001)	-0.0656 (0.0113)	0.2526 (0.0001)	0.0427	0.4199

The information content findings are presented in Table III.[9] Based on the reported R^2, none of the models offer much explanatory power in the determination of future one-hour volatility. R^2 Statistics range from a low of 0.0060 to a high of 0.0427. Focusing on the parameters, the role of ISD is mixed. The role of ISD is significant for all currencies. Except for the JY, and in contrast to reported findings on daily data, the ISD coefficient is negative. The inverse relationship between future volatility and ISD is not an intuitive one. However, in prior studies the role of α was statistically insignificant. In this study α is positive and significant at the 5-percent level for all currencies. Given the relatively low explanatory power of the hourly models, this could

[9] As with prior studies, the methodology employed here does not produce an overlap in the dependent variable. Hence, OLS derived p-values are presented here.

be an indication that the models are missing an important explanatory variable. We would suspect that the information content regressions would benefit by the inclusion of a variable to account for economic announcements. The results presented here suggest that at the hourly level, macroeconomic events may not be captured by ISD or GARCH volatility.[10]

Of particularly importance is the presentation of MSE for each model. The reported MSE is used later to compare the efficiency of the information content models to the RBF-ANN counterparts. The MSE results presented in Table III follow the other statistical results in this section. On average, the MSE is lowest for the DM and highest for the JY. This reflects the volatility analysis from above.

When the ISD and GARCH time-series are treated together, the explanatory power of the future returns is as expected. Except for a small discrepancy with the DM, all slopes are significant at the 5-percent level. R^2 values increase for each model within currency model group. The reported MSE value is the lowest when compared to the other two models within currency group. Again, for each model α is statistically significant; this is a finding that continues to suggest the effect of an omitted variable.

Figure 2a, 2b, and 2c display the time-variation in s^{ISD}, s^{GARCH} and the corresponding realized volatility measured in percent per hour. The average annualized percentage volatility for the DM, JY, and SF is about 23, 114, and 37 percent, respectively. The performance of the GARCH(1,1) model is quite satisfactory when applied to the squared returns. Except for areas demonstrating intraday seasonality, the coherence between the volatility forecasts and the ex post volatility measure is quite good.

[10] The DM parameter does report a positive sign. However, as mentioned this variable is not significant at the 10-percent level. The JY is clearly insignificant and is reported with a negative sign.

Figure 2a

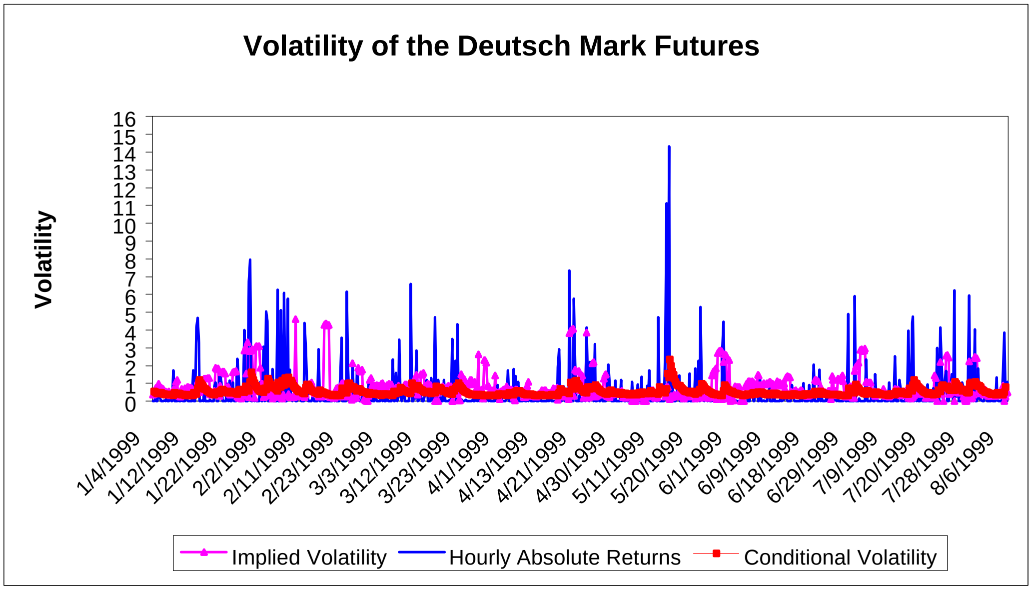


Figure 2b

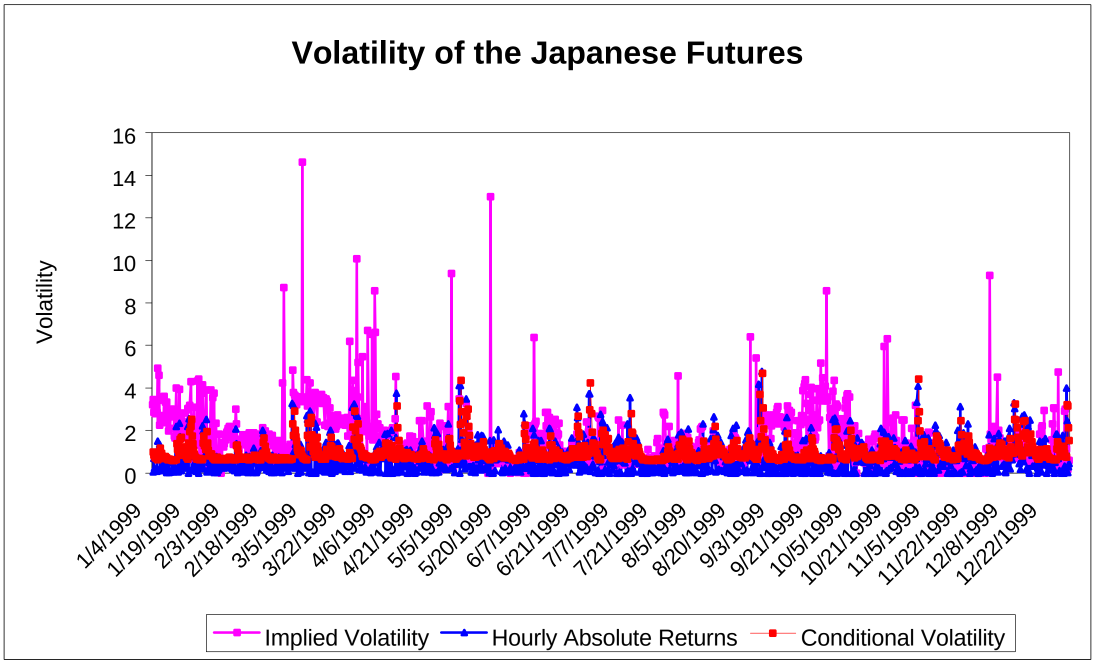
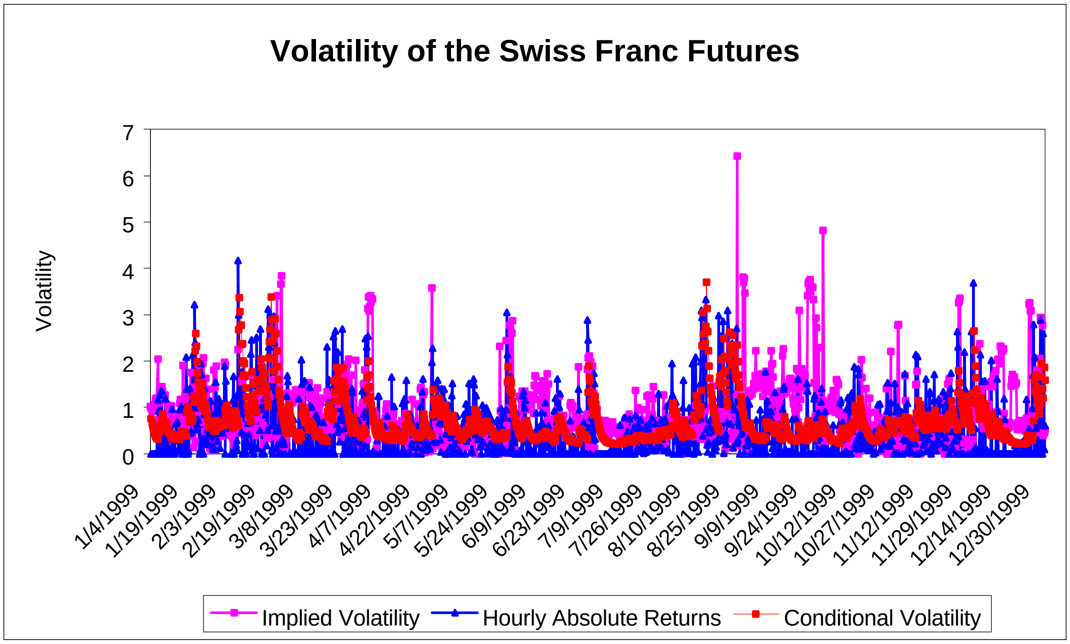


Figure 2c



4. Futures Options Predictability via a RBF ANN

Artificial neural networks and nonparametric methods have become increasingly important in the study of high-frequency financial data. Impressive results utilizing daily frequencies have already been reported. Malliaris and Salchenberger (1996) and Niranjan (1997) each report superior results from the application of the neural network methods to the problem of modeling options volatility when compared to other statistical methodologies. Hutchinson et al., (1996) propose the use of 'smoothing' based neural net topologies as a method to price options. Other nonparametric or mixed applications are prominent in the literature as well. For example, foreign exchange rates have been successfully modeled in a mixed Kalman filter neural network architecture (see Bolland et al. (1998)); prediction of financial distress (Coats, et al. (1992)), and the modeling stochastic systems by Elanyar, et al. (1994) offer contrasts between traditional methodologies and the incorporation of nonparametric methods. The prediction of stock market performance by Refenes and Bolland (1996), forecasting model selection by Sohl and Venakatachalam (1995), and forecasting futures trading volume by Kaastra and Boyd

(1995) provide additional useful examples. ANN have also been collapsed with other artificial learning methods as well as multivariate statistical methods to derive improved model accuracy. Olaf (1997), for example, blended a genetic algorithm with an ANN to produce improved prediction of stock index returns.

Because of the adherence to contemporary financial theory, the Hutchinson, Lo, and Poggio study deserves additional mention. In this study a RBF ANN was trained on artificially induced prices from solving the Black-Scholes option pricing formula (BSOP). With as little as six months of daily data, the ANN pricing results compared reasonably with the outcomes produced by the BSOP. The authors note that an attractive feature of the RBF ANN is its reliance on regularization techniques. The regularization approach to RBF application was shown to perform well in its approximation of both the BSOP equation and its derivatives. This latter point is of particular interest as optimal hedging strategies require methods that are consistent with accurate derivative estimation.[11]

4.1 Regularization Theory and RBF Neural Networks

Poggio and Giorsi (1990a; 1990b) introduced the theory of regularization in capital market neural net applications. They suggested the use of regularization theory in neural networks as a way of controlling the smoothness properties of a mapping function. The supervised learning function is stated as,

$$y = f(x) \tag{12}$$

Where:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

[11] For a more complete discussion on the approximation accuracy of the RBF ANN when applied to a function and its derivatives see, Poggio and Girosi 1990. For a discussion on the general arguments of the approximation issue see Gallant and White 1992 and Hornick, Stinchcombe, and White 1990.

and,

y = output vector
x = input vector
n = number of inputs

The supervised learning function can be stated as a linear model,

$$f(x_i) = \sum_{j=1}^{m} w_j h_j(x) \qquad (13)$$

where,

m	= number of basis functions (centers)
h	= hidden units via Broomhead & Lowe
w	= weight vectors
i	= 1..K output vectors (target variables)

The flexibility of *f* and its ability to fit many different functions is inherited from the freedom to choose different values for the weights. The RBF-ANN is constructed as a supervised ANN least-squares method that is capable of solving for the optimal weight values from a designated data set known as the "training set." Further, we note that the RBF expression may also be viewed as a restatement of the Tikhonov's (1977) regularization equation. Under Tikhonov regularization a weight decay parameter is added to the error function to penalize mappings that are not smooth. The minimum error function value is found by optimizing the set of smoothing weights.

Traditionally, in the RBF framework iterative techniques are used to compute the weight decay parameter (see Orr (1996; 1997)). But, iterative techniques have known drawbacks. In addition to being computationally burdensome, iterative methods lack specificity, as they require an initial estimate for the regularization parameter. It is not uncommon for iterative methods to terminate at a local minimum, or produce inflated residual sums of squares when the ridge parameter goes to infinity.

4.2 Regularization Parameter: A Closed Form Solution

Crouse, Jin, and Hanumara (1995) offer a closed-form method for the estimation of an optimal ridge-parameter in ridge-regression. These authors go on to enhance the derivation by including a Bayesian enveloped prior information matrix. The Kajiji (2001) RBF-ANN implements a similar closed form solution (with a prior information matrix) to derive the initial estimate of the regularization parameter used in the solution of the RBF-ANN. Kajiji reasoned that an optimally derived estimate would lead to a reduction of the noise-induced inflation in the residual sum of squares (MSE inflation). Hemmerle (1975) proposed an alternative closed-form solution to the estimation of the ridge-regression parameter by offering a modification to the original Hoerl & Kennard (1970) iterative method. In contrast to the Crouse et al. method that produces a globally optimized parameter, Hemmerle's method produces a vector of optimized ridge parameters. Kajiji (2001) adapted both closed-form proofs to produce a new RBF-ANN.

In the next section, we apply the Kajiji-RBF framework to the volatility prediction problem. The results of solving the Kajiji-RBF algorithms produce an MSE. This MSE extraction leads to a direct comparison of the MSE results obtained from the information content of the GARCH analysis.

5. Comparative Analysis

The Kajiji-RBF algorithm is capable of solving several alternative RBF topologies. For conciseness and brevity we limit our detailed comments to the solutions produced from the algorithmic specification of greatest interest. Specifically, we refer to this algorithm as Kajiji-4; it is a RBF algorithm that relies upon a closed form derivation for estimation of the regularization parameter with a prior information set. The Kajiji-4 model topology produced the smallest MSE statistics across all versions of the Kajiji RBF algorithmic framework.

5.1 Macroeconomic Factors

There is increasing evidence that intraday return volatility is linked to the release of regularly scheduled macroeconomic announcements and other news events. Bollerslev, Cai, and Song (2000) find strong evidence of announcement effects in the 5-minute returns of U.S. Treasury bonds. Utilizing these findings we introduce two new variables to the RBF simulations. Added are the hourly intraday returns for U.S. Treasury bond futures (TB) and the U.S. Dollar Index (DX). Both variables are treated as proxies for macroeconomic- and news-announcements. The TB futures contract is well known and documented in the literature. The DX is the futures contract on the Federal Reserve Bank of Atlanta's trade-weighted dollar index. The latest revision of the DX produced an index that captures the effects of the United States' 15 largest trading partners. We expect the TB to capture the U.S. based announcement effects and the DX to capture the combined effects of announcements by the represented trading partners.[12]

5.2 The RBF-ANN Model Specifications

The Kajij-4 algorithm was applied to four statistical models. Model I is an RBF application of equation (11). The Model II specification adds the hourly returns for the DX to the Model I specification. Model III is Model II with the addition of the hourly returns for U.S. Treasury bond futures (TB). Model IV reduces Model III by the dropping the ISD component. Model V includes two predictor variables: GARCH and DX. Finally, Model VI examines the predictability of GARCH and TB together.

5.3 RBF Results

The results of solving the Kajiji-4 algorithm on the four economic models are presented in Table IV. The table displays the optimal weights obtained from the training set of data; the training set MSE and the test set MSE. The test set is the data held in reserve to examine the ex-post predictive power of the trained ANN. For this

[12] The pairwise correlation between TB and DX is –0.6643 with a p-value of 0.0001. Although the two proxies for the macroeconomic effect are significantly correlated this should not pose a problem to the RBF ANN as the underlying methodology of ridge regression is well suited to handle collinear predictor variables.

experiment, the test set was kept to a minimum number of observations. Since our objective is to compare the ability of the RBF-ANN to explain information content rather than to predict future occurrences of volatility, our efforts are extended on efficient training of the RBF.

Table IV: Information Content of Returns: RBF-ANN

	Weights				**Training Set MSE**	**Test Set MSE**
	ISD	**GARCH**	**DX**	**TB**		
DM						
Model I	-0.211825	0.954984			0.152787	0.437599
Model II	0.499772	0.258728	0.336745		0.174179	0.457883
Model III	0.396330	0.352969	0.308326	0.327398	0.175335	0.507098
Model IV		0.954023	1.121414	-0.624527	0.175334	0.494009
Model V		-0.171523	0.940443		0.174314	0.459409
Model VI		-0.198441		1.023938	0.175329	0.505138
JY						
Model I	-1.066420	1.597764			0.467589	0.951658
Model II	1.202824	-0.370107	1.197315		0.467589	0.769282
Model III	3.457789	-1.526313	-1.343548	2.033734	0.467589	0.790742
Model IV		1.922160	-3.375070	4.202657	0.467589	0.781442
Model V		0.157938	0.660742		0.467589	0.952444
Model VI		-2.469492		3.113836	0.467589	0.953019
SF						
Model I	1.093531	0.737385			0.024500	0.536022
Model II	1.143419	1.610866	1.146307		0.000001	0.613796
Model III	-1.616530	3.050662	2.187535	1.116123	0.000001	0.651792
Model IV		-3.004039	3.482911	2.200567	0.000001	0.647218
Model V		1.144421	0.620053		0.000001	0.617666
Model VI		1.269823		0.891271	0.000001	0.623689

As with the earlier GARCH framework, the results obtained for the JY stand apart from the DM and the SF. For the SF all models except Model I are over-trained (over-fitted).[13] When compared to the test set, the training set MSE values indicate an ANN that is not capable of sustaining any reasonable level of prediction. Hence, for the analysis of information content, we conclude that Model I presents the more accurate analysis when applied to both the DM and SF.

[13] In an over-fitted network the output variable fits the data forming the underlying function so closely that it also models the noise in the data set.

The Model I test set results may be compared directly to findings presented in the GARCH framework. The MSE findings for the GARCH framework were reported as 0.3892 for the DM and 0.4199 for the SF, respectively. The RBF reduced the MSE by a factor of 2.5 for the DM and by a factor greater than 17 for the SF. The results for the JY are strikingly different. Each model formulation produced approximately identical MSE values for the training set MSE. When compared to the MSE from the GARCH framework, the improvement is only marginal with an improvement factor of 1.13. It is only the lower value of the test set that suggests which model may be preferred. The lowest MSE value on the test set simulation is recorded for Model II, or the model that includes DX as a proxy for the macroeconomic news effect.

These results are complemented by the graphical representation of the findings. Figures 3a, 3b, and 3c present the modeling characteristics for the three currencies.

Figure 3a

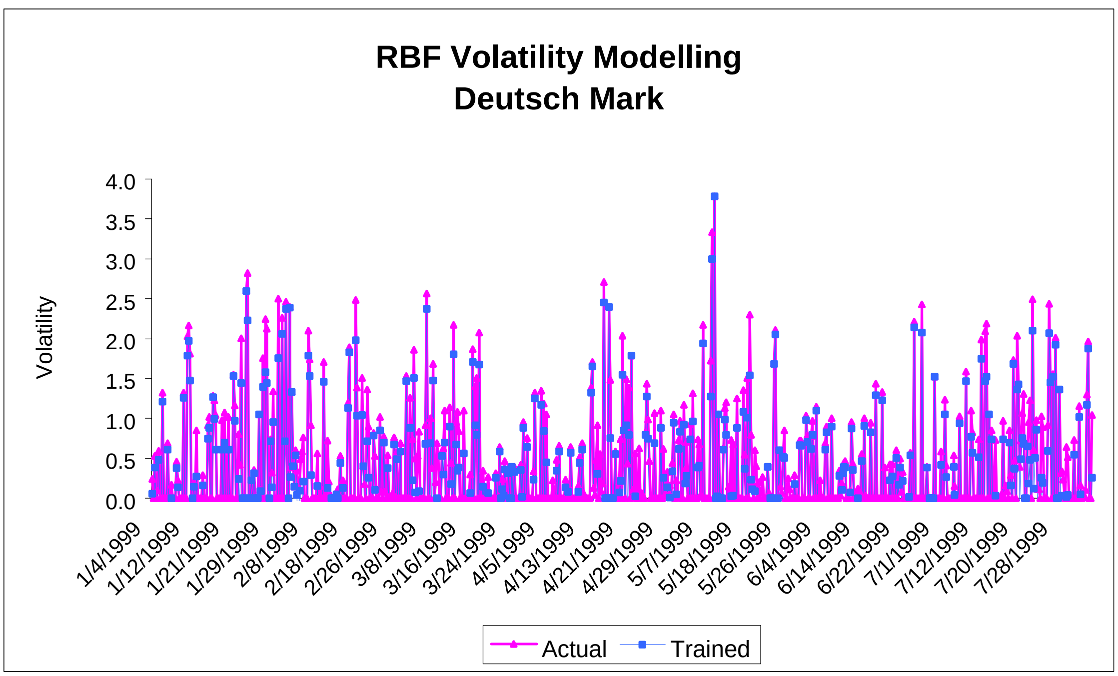


Figure 3b

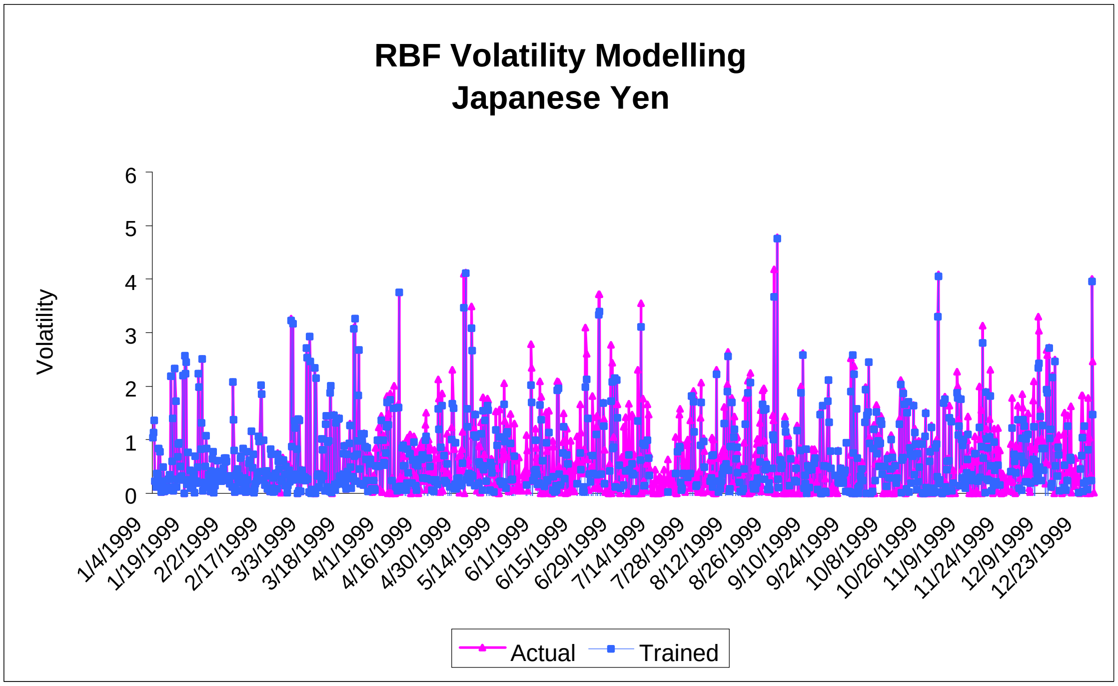
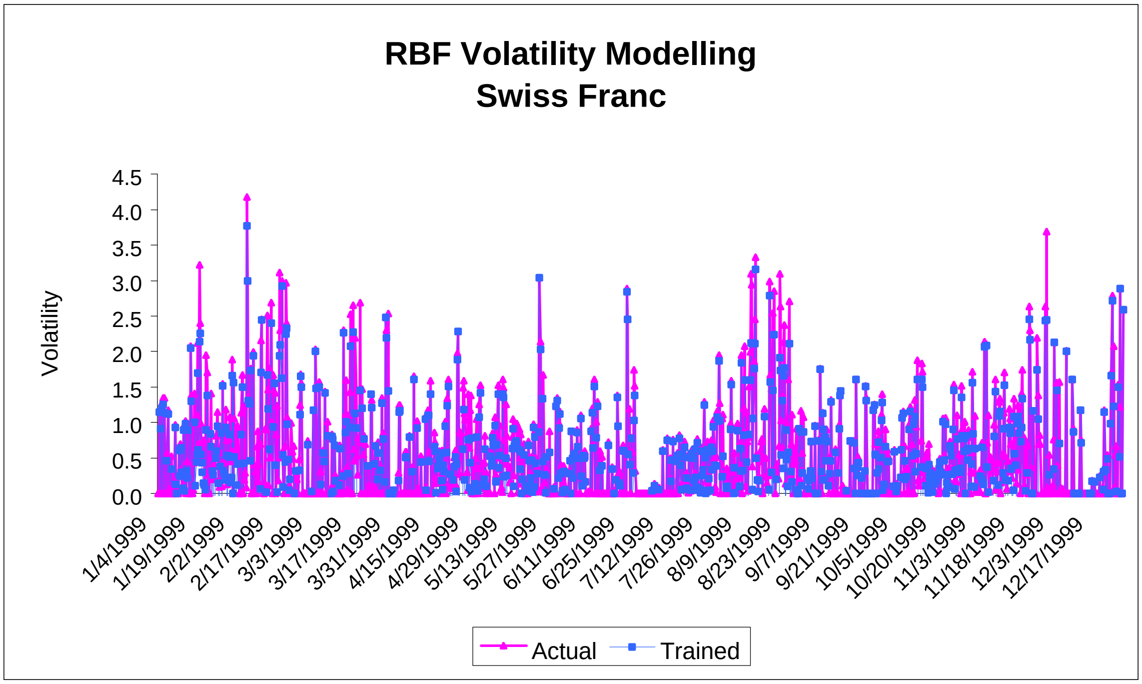


Figure 3c



5.4 Summary and Conclusions

Within the context of the general theory of continuous time arbitrage-free price processes, this study presents a comparison of volatility modeling. The study explicitly compares the predictive power of a GARCH framework to that of an RBF ANN framework. In this study we rely upon futures options volatility measures inverted from hourly intervals. The results of our implementation of a GARCH discrete model produced results that supported earlier findings on daily intervals. Additionally, the GARCH based results presented in this paper support more recent findings that incorporate 5-minute interval high frequency data.

The use of nonparametric ANN methods in high frequency data modeling was the primary focus of this study. We provide evidence in this paper that RBF-ANN topologies constructed on a closed-form regularization parameter is capable of modeling volatility information that is not fully captured in the traditional GARCH framework. In all model formulations, the RBF solution dominated the error minimization functions recorded within the comparable GARCH formulation. These results bode well for future model examinations that rely upon this family of RBF design. Nevertheless, these encouraging research findings are tempered by a number of unanswered research questions. For example, the use of MSE as a benchmark error measure is not without issue. Additional error reporting statistics must be reported in future examinations. The question of data frequency also requires further investigation. Whether the augmented RBF will perform consistently on higher (lower) frequencies must be examined.

The Kajiji-RBF ANN needs to address the derivation of "optimal-value" parameter settings. For example, the inclusion of dynamic parameter settings will, in all likelihood, lead to a genetically optimized version of the Kajiji-RBF ANN. As applied in this research it is possible that the subjectively determined parameter settings produced an optimal but non-dominant solution; albeit, a solution that dominates the comparative GARCH findings. Finally, there is the question of macroeconomic proxies. At least one set of results presented here suggests that this issue deserves continued investigation when posited in an RBF formulation.

References

Andersen, T. G. and T. Bollerslev (1994). Intraday Seasonality and Volatility Persistence in Foreign Exchange and Equity Markets. Kellogg Graduate School of Management, Northwestern University.

Andersen, T. G. and T. Bollerslev (1997). Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. Time Series Analysis of High Frequency Financial Data, San Diego, CA.

Beckers, S. (1981). "Standard deviations implied in option prices as predictors of future stock price variability." Journal of Banking and Finance: 363-381.

Black, F. (1976). "The Pricing of Commodity Contracts." Journal of Financial Economics **3**(Jan-Feb): 167-179.

Bolland, P. J., J. T. Connor, et al. (1998). Application of Neural Networks to Forecast High Frequency Data: Foreign Exchange. Nonlinear Modelling of High Frequency Financial Time Series. C. Dunis and B. Zhou. Chichester, John Wiley & sons**:** 225-246.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity." Journal of Econometrics **31**: 307-327.

Bollerslev, T., J. Cai, et al. (2000). "Intraday Periodicity, Long Memory Volatility, and Macroeconomic Announcement Effects in the US Treasury Bond Market." Journal of Empirical Finance **7**: 37-55.

Bollerslev, T., R. Y. Chou, et al. (1992). "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence." Journal of Econometrics **52**: 5-59.

Bollerslev, T., R. F. Engle, et al. (1994). ARCH Models. Handbook of Econometrics. R. F. a. M. Engle, D.L. Amsterdam, Elsevier Science.

Bond, S. A. (2000). Asymmetry and Downside Risk in Foreign Exchange Markets. Department of Land Economy. U.K., Cambridge, University of Cambridge**:** 51.

Coats, P. and L. Fant (1992). "A Neural Network Approach to Forecasting Financial Distress." Journal of Business Forecasting **10**(4): 9-12.

Crouse, R. H., C. Jin, et al. (1995). "Unbiased Ridge Estimation with Prior Information and Ridge Trace." Communication in Statistics **24**(9): 2341-2354.

Dacorogna, M. M., U. A. Muller, et al. (1998). Modelling Short-term Volatility with GARCH and HARCH Models. Nonlinear Modelling of High Frequency Financial Time Series. C. Dunis and B. Zhou. Chichester, England, John Wiley & Sons**:** 161-176.

Diebold, F. X. and J. A. Lopez (1995). Modeling Volatility Dynamics. Macroeconometrics: Developments, Tensions, and Prospects. K. D. Hoover. Boston, Kluwer Academic Publishers.

Elanyar, V. T. and Y. C. Shin (1994). "Radial Basis Function Neural Network for Approximation and Estimation of Non linear Stochastic Dynamic Systems." IEEE Transactions on Neural Networks **5**(4): 594-603.

Engle, R. F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation." Econometrica **50**: 987-1008.

Ghose, D. and K. F. Kroner (1995). "Temporal Aggregation of High Frequency Financial Data." Proceedings of the HFDF-I Conference **2**: 1-31.

Ghysels, E. A., A. Harvey, et al. (1996). Stochastic Volatility. Handbook of Statistics. G. S. Maddala. Amsterdam, North Holland. **Vol. 14**.

Guillaume, D. M., O. V. Pictet, et al. (1994). On the Intra-day Performance of GARCH Processes'. Internal Document dMG.1994-07-30. Zurich, Switzerland, Olsen & Associates.

Hemmerle, W. J. (1975). "An Explicit Solution for Generalized Ridge Regression." Technometrics **17**(3): 309-314.

Hoerl, A. E. and R. W. Kennard (1970). "Ridge Regression: Biased Estimation for Nonorthogonal Problems." Technometrics **12**(3): 55-67.

Hutchinson, J. M., Lo, A.W., and Poggio, T. (1996). A Nonparametric Approach to Pricing and Hedging Derivative Securities via Learning Networks. Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real-World Performance. R. R. Trippi and E. Turban. New York, McGraw Hill**:** Chapter 33.

Jorion, P. (1995). "Predicting Volatility in the Foreign Exchange Market." The Journal of Finance **Vol. L.**(No. 2): 507-528.

Kaastra, I. and M. S. Boyd (1995). "Forecasting futures trading volume using neural networks." Journal of Futures Markets **15**(8): 953-970.

Kajiji, N. (2001). Adaptation of Alternative Closed Form Regularization Parameters with Prior Information to the Radial Basis Function Neural Network for High Frequency Financial Time Series. Applied Mathematics. Kingston, University of Rhode Island.

Malliaris, M. and L. Salchengerger (1996). Neural Networks for Predicting Options Volatility. Neural Networks in Finance and Investing. R. R. Trippi and E. Turban. New York, McGraw-Hill. **2:** 613-622.

Niranjan, M. (1997). Sequential Tracking in Pricing Financial Options using Model Based and Neural Network Approaches. Advances in Neural Information Processing Systems. M. C. Mozer, Jordan, Michael I., and Petsche, Thomas. Boston, The MIT Press. **9:** 960-972.

Olaf, W. (1997). Predicting Stock Index Returns by Means of Genetically Engineered Neural Networks. Department of Management. Los Angeles, University of California.

Orr, M. J. L. (1996). Introduction to Radial Basis Function Networks, Center for Cognitive Science, Scotland, UK.

Orr, M. J. L. (1997). MATLAB Routines for Subset Selection and Ridge Regression in Linear Neural Networks., Center for Cognitive Science, Scotland, UK.

Poggio, T. and F. Girosi (1990a). "Networks for Approximation and Learning." Proceedings of the IEEE **78**(9): 1481-1497.

Poggio, T. and F. Girosi (1990b). "Regularization Algorithms for Learning that are Equivalent to Multilayer Networks." Science **247**: 978-982.

Refenes, A. N. and P. Bolland (1996). Modeling Quarterly Returns on the FTSE: A Comparitive Study with Regression and Neural Networks. Fuzzy Logic and Neural Network Handbook. C. H. Chen. New York, McGraw-Hill**:** 19.1-19.28.

Sohl, J. E. and A. R. Venkatachalam (1995). "A Neural Network Approach to Forecasting Model Selection." Information Management **29**(6): 297-303.

Tikhonov, A., and Arsenin, V. (1977). Solutions of Ill-Posed Problems. New York, Wiley.