Modeling Heterogeneous Risk Behavior of the South African Rand via A Closed-form Radial Basis Function Neural Network

A Preliminary Analysis

By

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

Nina Kajiji **
Assistant Professor of Research, Education

Initial Printing: 31-JUL-02

JEL CATEGORY
C22 ECONOMETRIC METHODS: Time Series Models
C45 ECONOMETRIC AND STATISTICAL METHODS; Neural Networks
C53 ECONOMETRIC MODELING; Forecasting


**Please address all correspondence to: Gordon H. Dash, Jr., 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

© 2002 By Gordon H. Dash, Jr. and Nina Kajiji. All rights reserved. Explicit permission to quote is required for text quotes beyond two paragraphs. In all cases, full credit, including © notice is required.

The authors wish to acknowledge The NKD Group, Inc. of Wilmington, Delaware for providing a grant for the acquisition of data, computational algorithms and travel offsets.
Abstract

Stylized facts are uncovered for a domestic (U.S.) examination of the South African Rand futures contract (ZAR). In this preliminary study, we model complex volatility patterns by a nonparametric artificial neural network (ANN) that incorporates a performance enhancing closed-form regularization technique. The modeling characteristics revealed by the Kajiji-4 radial basis function (RBF) ANN provides significant new information about the behavior and prediction of the ZAR futures contract. The impact of U.S. based news proves to be an important determinant of volatility prediction. However, the proxy for the U.S. trade weighted dollar index does not prove to be particularly important. Additionally, there is evidence that future studies should continue a focus on the role of one-day lagged behavior.
I. Introduction

It is widely known that high frequency time series from the FX markets are best modeled by complex nonlinear functions. Modeling the volatility of FX futures contracts is one of the complex nonlinear functions that have been a subject of numerous academic studies proposing alternative methods to best account for the factors underlying volatility patterns over time. One of the more promising nonparametric methods used to replicate FX futures contract volatility falls under the domain of artificial neural networks (ANN). Hornik, Stinchcombe, and White (1989) provide significant evidence about the ability of ANN models to approximate any Borel measurable function to any degree of accuracy. Not surprisingly, the literature is rich with studies that offer various enhancements to the ANN topology to better exploit this method as a financial time-series modeling tool. By way of example, Malliaris and Salchenberger (1996); Niranjan (1997); Hutchinson et al., (1996) and Kajiji (2001) have each focused on incorporating enhancements to dampen noise-related error inflation in a technique that is employed in this study – the radial basis function neural network (RBF).

The findings presented in this paper are preliminary. The principal focus of this initial effort is to determine the applicability of the Kajiji-4 RBF-ANN when applied to the task of modeling the volatility of the ZAR futures contract. A subsequent study will focus on a more definitive approach to the sample period. For now, we limit our objective to the application of the Kajiji-4 RBF-ANN to the problem of measuring the information content in the one-day ahead volatility of the ZAR futures contract as traded on the U.S. Chicago Board of Trade. Additionally, we report the results of the GARCH(1,1) volatility model when applied to the same contract.

2. The RBF ANN with Prior Information and Regularization

The objective of the RBF ANN is to combine a hidden layer of radial units that model a Gaussian response surface into a network of outputs. Since these functions are non-linear, it is not actually necessary to have more than one hidden layer to model any
shape of function. With sufficient radial units the RBF design will always produce an efficient modeling of any function. But, the method does not exist without limitations. Stylized facts have emerged from RBF implementations that report extrapolation inaccuracies when data points are far from the training set. Because it is sufficient to use a linear combination of these outputs (i.e. a weighted sum of the Gaussians), RBF networks are generally constructed as a supervised least squares based method that is first applied to a training set as a means of deriving the optimal weighting values. The supervised learning function may be stated as,

\[ y = f(x) \]  

(1)

where, \( y \) the output vector is a function of \( x \) the input vector with \( n \) number of inputs.

Alternatively, the supervised learning function can be restated as the following linear model,

\[ f(x_i) = \sum_{j=1}^{m} w_j h_j(x) \]  

(2)

where, \( m \) is the number of basis functions (centers), \( h \) is the hidden units, \( w \) is the weight vectors, and \( i = 1..K \) output vectors (target variables). The flexibility of \( f \) and its ability to model many different functions is inherited from the freedom to choose different values for the weights.

Note that the RBF mapping function may be restated as a Tikhonov’s (1977) regularization equation. Tikhonov regularization adds a weight decay parameter to the error function to penalize mappings that are not smooth. Traditionally, 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. Computational experience suggests that when iterative methods are employed, it is not uncommon to experience local minimum, or to produce inflated residual sums of squares when the weight decay parameter goes to infinity.
Kajiji (2001) reasoned that an optimally derived regularization estimate would reduce the “curse” of dimensionality and thereby assist in achieving a reduction in noise-induced inflation in the residual sum of squares. Through the use of a closed-form solution method for the estimation of an optimal ridge regression parameter when enhanced by Bayesian prior information set, Kajiji implemented a parallel extension to the RBF topology as a means by which to derive the initial estimate of the regularization parameter. Hemmerle (1975) proposed an alternative closed-form solution to the estimation of the ridge-regression parameter by offering a modification to the original Hoerl and Kennard (1970) iterative method. Further, Hemmerle’s method produced a vector of optimized ridge parameters unlike a single non-optimized ridge parameter of Hoerl and Kennard. By contrast, Crouse et al. (1995) produced an algorithm that further enhanced the prediction ability of the Hoerl and Kennard single ridge parameter by adding a prior information matrix. In tests of algorithmic efficiency Kajiji found her extension to the regularization parameter modeled complex nonlinear financial time series with a consistently smaller fitness function MSE than did the non-optimized iterative-based RBF with regularization. Although Kajiji presented four alternative RBF algorithms (Kajiji-1 to Kajiji-4), we postulate that the Kajiji-4 algorithm is a sufficiently reasonable method to model high frequency financial time series.

3. Conditional Volatility and the ZAR Futures Contract

3.1 Data

The models reported in this section are based on daily returns obtained from closing quotes on the futures contract against the dollar exchange with the South African Rand. In addition to the currency-related data, high-frequency futures data on the dollar index (DX), and the U.S. Treasury Bond (TB) were obtained. Daily observations on the ZAR futures contract are obtained from Global Financial Data. All other futures contract
data are obtained from Tick Data, Inc. In this preliminary study, we limit the sample period to January 4, 1999 to December 31, 1999 inclusive. Where tick observations are provided, we employ a harmonic mean to aggregate the data into equally spaced daily observations beginning with the Monday 9:00 a.m. closing quote. The last quote of the day is captured with the closest tick to the 1:59 p.m. stamp. This process is repeated for all available trading days of the business week within the sample period. This results in 250 daily observations. For modeling purposes are training set size is 186 observations. The validation set is 63 observations.

3.2 Modeling Conditional Volatility

The GARCH model was first developed to model data at the daily frequency level or greater. The stylized facts report volatility persistence in high frequency financial data. Hence, we are not surprised to observe the keen attention of 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 volatility 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.³

² CME upper floor trading hours are 7:20 a.m. to 2:00 p.m. However, the data supplied from Tick Data begins at 8:30 a.m.

Modeling Heterogenous Risk…. -6- Dash & Kajiji
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:

$$
\begin{align*}
    x_t &= \sigma_t z_t \\
    \sigma_t^2 &= \alpha_0 + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2
\end{align*}
$$

where:

$z_t \sim NID(0,1)$, for $\alpha_0, \alpha_1 \geq 0$ and $t=1...T$. The model implies that $x_t | \Omega_{t-1} \sim N(0, \sigma_{t-1}^2)$.

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.

4. Estimation Results

4.1 Descriptive Statistics

Table 1 provides descriptive statistics computed from daily observations on the ZAR FX futures contract. Presented in the table are results on the daily observations: one-day return, volatility (%), and conditional volatility. The returns pattern for the ZAR shows negative lags at daily frequencies of 1, 3, 4, and 5. The relatively small value of the lags beyond one-day suggests that a one-day lag may be useful in explaining the returns generating process of the ZAR futures contract. The focus here is on volatility. We note that one-day volatility shows a positive lag of approximately the same magnitude for up to 3 days. Beginning with day 4 the positive lag begins a steady erosion. No negative lags are observed for ZAR daily volatility.

Figure 1 provides a descriptive view of the daily returns for ZAR. Except for January 1999 the volatility patterns in the daily returns present a pattern that is well documented for high frequency data. Volatility is persistent but is clearly changing and diminishing over the sample period. This observation is confirmed by the diminishing

---

1 For more on the subject of the GARCH framework see the review in Dash and Kajiji (2001).
2 Other non-normal conditional distributions have been used in the model specification. By way of example, see Nelson 1991.

---

Modeling Heterogenous Risk.... -7- Dash & Kajiji
value of the positive lag reported for one-day volatility as reported in Table 1. With a value of 0.980 for day-one, the effect of the positive lag diminishes to just slightly more than half of that value by day ten (0.564). The same trend is evident in the contract’s conditional volatility finding.

Table I: Descriptive Statistics

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>Autocorrelations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Rand</td>
<td>0.000191</td>
</tr>
<tr>
<td>1-day return</td>
<td>0.003812</td>
</tr>
<tr>
<td>1-day volatility (%)</td>
<td>0.000056</td>
</tr>
</tbody>
</table>

Figure 1: ZAR/USD Daily Returns

South African Rand Returns on Futures
Jan 4, 1999 - Dec 31, 1999
4.2 GARCH Estimate

Conditional volatility is estimated by applying GARCH(1,1) to the daily returns. Parameters $\alpha_1$ and $\beta$ 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 $\chi^2$ statistic of the GARCH estimate confirms that a GARCH process is describing a statistically significant amount of the conditional variance in returns. 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.\(^5\) Stated differently, realized daily volatility is stationary, but it does change over time.

<table>
<thead>
<tr>
<th>Currency</th>
<th>Model</th>
<th>$\mu$</th>
<th>$\alpha_0$</th>
<th>$\alpha_1$</th>
<th>$\beta$</th>
<th>Long-Run Volatility (% pa)</th>
<th>Log-Likelihood</th>
<th>Pr &gt; $\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rand</td>
<td>GARCH (pr)</td>
<td>$3.22 \times 10^{-5}$</td>
<td>$3.8534 \times 10^{-7}$</td>
<td>$0.1757$</td>
<td>$0.8234$</td>
<td>$0.0573$</td>
<td>$986.22$</td>
<td>$0.0193$</td>
</tr>
</tbody>
</table>

4.3 RBF Volatility Modeling Efficiency

In order to fully examine the RBF-ANN in its appointed task of volatility modeling, it will be necessary to examine alternative RBF model specifications. The results of solving alternative RBF models is examined by comparing computed MSE statistics for each training, validation and fitness execution. Model errors for RBF model applications are computed as,

\[
MSE_{training} = \frac{1}{T_1} \sum_{i=1}^{T_1} (y_i - \hat{y}_i)^2
\]

\(^5\) OLS, Hansen-White, and GARCH estimates are obtained from version 8e of the Statistical Analysis System (SAS), North Carolina.
\[ MSE_{\text{validation}} = \frac{1}{T_2} \sum_{i=1}^{T_2} (y_i - \hat{y}_i)^2 \] (6)

\[ MSE_{\text{fitness}} = \frac{1}{T} \sum_{i=1}^{T} (y_i - \hat{y}_i)^2 \] (7)

where the training error MSE is restricted to the training set sample. The validation MSE is the error reported over the out-of-sample validation set and the fitness MSE is computed over all \( T \) observations.

All RBF model applications use the Kajiji-4 default settings for all parameters. This restriction only permits comparison of the “default” state of the RBF. Performance characteristics that differ from those reported below may be possible with further refinement(s) to the optional parameter settings.

4.4 RBF Model Efficiency

Table 3 presents the results of solving ten alternative models. For each model the MSE for the training, validation, and resultant fitness function is presented for comparative review. Model III produced the lowest fitness MSE at 0.0410. All other models produced a fitness MSE values in excess of 0.500. Model III confirms the importance of conditional volatility and the U.S. news effect in explaining ZAR futures volatility. Models with 1, 2, 7, 5, and 10 day lags did achieve the same level of fitness response as Model III. However, the nature of the ZAR returns volatility as displayed in figure 1 can be observed by a comparison of the training set MSE statistics. For example, in addition to the conditional volatility parameter and DX, Model V includes both a 2-day and 7-day lag. This model produces the low value for training set MSE at 0.1405 (compared to Model III with a training set MSE of 0.5034). This fact reflects the influence in the sub-sample of the unique January, 1999 volatility pattern. However, as time passes, the model’s explanatory power dissipates. This is further explained by a known shortcoming of the RBF algorithmic design: its inability to extrapolate data that falls far from the training set.
Table 3: ZAR Model MSE by Model

<table>
<thead>
<tr>
<th>Futures Contract</th>
<th>Inputs</th>
<th>MSEs for Kajiji-4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GARCH DX TB Lag1 Lag2 Lag7 Lag5 Lag10</td>
<td>Training Validation Fitness</td>
</tr>
<tr>
<td>RAND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model I</td>
<td>I I I I</td>
<td>0.4999 0.0470 0.0507</td>
</tr>
<tr>
<td>Model II</td>
<td>I I I I</td>
<td>0.4669 0.0484 0.0676</td>
</tr>
<tr>
<td>Model III</td>
<td>I I I I</td>
<td><strong>0.5034</strong> 0.0402 0.0410</td>
</tr>
<tr>
<td>Model IV</td>
<td>I I I I</td>
<td>0.1599 0.0481 0.0601</td>
</tr>
<tr>
<td>Model V</td>
<td>I I I I</td>
<td>0.1405 0.0461 0.0780</td>
</tr>
<tr>
<td>Model VI</td>
<td>I I I I</td>
<td>0.1665 0.0487 0.0573</td>
</tr>
<tr>
<td>Model VII</td>
<td>I I I I</td>
<td>0.4822 0.0465 0.0516</td>
</tr>
<tr>
<td>Model VIII</td>
<td>I I I I</td>
<td>0.4863 0.0412 0.0513</td>
</tr>
<tr>
<td>Model IX</td>
<td>I I I I</td>
<td>0.4695 0.0430 0.0545</td>
</tr>
<tr>
<td>Model X</td>
<td>I I I I</td>
<td>0.4715 0.0418 0.0594</td>
</tr>
</tbody>
</table>

Figure 2 provides visual support for the results presented in table 3. Clearly evident is a predicted function that performs well over the training set. However, two observations may be gleaned from these comparative time series. First, as time passes and volatility declines by showing less dispersion among the peaks and troughs, it is clear that the Kajiji-4 algorithm is able to capture this trend with extreme precision. However, the algorithm also shows an under-prediction bias in its replication of the volatility pattern. That is, the predicted values are below the actual values despite the ability of the algorithm to match peaks and troughs on a daily basis.
Figure 2: Actual and Predicted Values of Rand

Actual and Predicted Values of the South African Rand

5. Summary and Conclusions

As a preliminary study on the volatility of the ZAR futures contract, we focused on the role of the RBF-ANN to provide an efficient prediction platform for financial decision-makers. Before exploring the applicability of the Kajiji-4 RBF-ANN, the GARCH(1,1) model was applied to the daily returns data. As with other studies from the FX markets performed on daily returns data, we found volatility persistence with a time effect. This finding is important for two reasons. First, it suggests that a successful application of the RBF-ANN to the ZAR may warrant the application of this nonparametric method to other high frequency financial data obtained from the global FX markets. Second, the GARCH findings presented in this paper further corroborate the conclusion that volatility for this contract, like other FX contracts, is appropriately – albeit, not completely – modeled by the ARCH style models.
The results of solving thirteen alternative RBF-ANN model specifications clearly point to the role of both conditional volatility and U.S. news as volatility components. The analysis did not find significant support for the inclusion of the U.S. dollar exchange weighted by its trading partners (DX). Nor did the analysis find increased significance with the inclusion of a one-day lag. This evidence alone is not sufficient to conclude that traders around the world can predict the amount and direction of returns volatility on a daily basis. However, it is clear that if a trading rule exists it is one that reacts to both the daily release of FX news and the conditional volatility of the contract. These findings beg the further interrogation of the role news (both domestic and South African) plays in the daily volatility pattern of this contract. The volatility of the ZAR contract seems to be somewhat isolated from U.S. trade related information. This finding only serves to increase the curiosity of how domestic South African information news may shape the volatility patterns that appear efficiently modeled by the Kajiji-4 RBF-ANN. This latter point is left for future study and investigation.
References


