

2011

Alternate Models for Forecasting Hedge Fund Returns

Michael A. Holden

Michael Holden, MikeHolden23@aol.com

Creative Commons License



This work is licensed under a [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/).

Follow this and additional works at: <http://digitalcommons.uri.edu/srhonorsprog>

 Part of the [Finance and Financial Management Commons](#), [Management Sciences and Quantitative Methods Commons](#), and the [Technology and Innovation Commons](#)

Recommended Citation

Holden, Michael A., "Alternate Models for Forecasting Hedge Fund Returns" (2011). *Senior Honors Projects*. Paper 205.
<http://digitalcommons.uri.edu/srhonorsprog/205><http://digitalcommons.uri.edu/srhonorsprog/205>

This Article is brought to you for free and open access by the Honors Program at the University of Rhode Island at DigitalCommons@URI. It has been accepted for inclusion in Senior Honors Projects by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.

Michael Holden

Senior Honors Project

Professor Dash

May 5, 2011

Alternate Models for Forecasting Hedge Fund Returns

This study will make use of a radial basis function (RBF) artificial neural network (ANN) to calculate correlations between the dependent and independent variables. The neural network essentially acts as an artificial brain, first learning the data set to identify a function that fits the bond volatility, then testing its proposed equation against the remaining data. After the equation is calculated, we observe the accuracy of the solution by examining various measures of error. We can then identify what factors most influence foreign bond volatility, and how substantial these effects are.

Artificial neural networks attempt to replicate the structure and function of the human brain in order to allow a computer system to learn and capture knowledge of a data set for use in forecasting and other functions. An artificial neural network is a massively parallel distributed processor that has a natural propensity for storing experimental knowledge and making it available for use.

To understand the structure of an artificial neural network, it is first necessary to examine the human brain it is modeled after. The human brain, though still not well understood, contains a massively interconnected network of approximately ten billion neurons and sixty trillion synapses connections. The biological neuron is the simple “arithmetic computing” element; it consists of four parts: the soma, the axon, the dendrites, and the synapses. The soma, also known as the body cell, is a large, round central body in which almost all the logical functions of the neuron are realized. The axon is a nerve fiber attached to the soma which can serve as a final output channel of the neuron. An axon is usually highly branched. The dendrites serve as inputs for the neuron. The dendrites represent a highly branched tree of fibers. These long irregularly

shaped nerve fibers are attached to the soma. Synapses are specialized contacts on a neuron which are the termination points for the axons from other neurons. These interconnected neurons work together to perform the multitude of tasks required by our daily activities.

Three important principles essential to brain processing are massive parallelism, connectionism, and associative distributed memory. The processes of pattern recognition and manipulation, essential to both biological and artificial neural networks, are based upon these. **Massive parallelism** defines the brain computer as an information or signal processing system that is composed of a large number of simple processing elements called neurons. These neurons are interconnected by numerous direct links, called connections, and cooperate with each other to perform parallel distributed processing (PDP) in order to solve desired computation tasks. **Connectionism** suggests that the brain computer is a highly interconnected neurons system in such a way that the state of one neuron affects the potential of the large number of other neurons which are connected according to weights or strength. The key idea of such principle is the functional capacity of biological neural networks is determined not by the sum of single neurons but by its connections. **Associative distributed memory** suggests that storage of information in a brain is supposed to be concentrated in synaptic connections of the brain's neural network, or more precisely, in the pattern of these connections and strengths (weights) of the synaptic connections. Artificial neural networks rely upon these principles and knowledge of brain structure and function to produce mathematical models of human-brain computation.

Like a human brain, the ANN first analyzes a given data set and trains itself. Knowledge is acquired by the network through a learning (training) process; the learning process is a procedure of adapting weights with a learning algorithm in order to capture the knowledge of a data set. The aim of the learning process is to map a given relation between inputs and output (outputs) of the network. Once the knowledge is captured, the artificial neural network can utilize this knowledge to perform numerous tasks. A key principle of artificial neural networks is the distinction between programming and training. The neural network is not programmed to solve tasks; rather, the ability to learn is programmed, thus allowing the solution of an endless possibility of tasks. This allows neural networks to adapt themselves to special environmental conditions brought up by a given task. Another significant principle of neural networks is their non-linear functionality. Every new state of a neuron is a nonlinear function of the input pattern

created by the firing of nonlinear activity of the other neurons. In addition to its use in financial analysis and forecasting, artificial neural networks are used in advanced robotics, intelligent control devices, technical diagnostics, signal processing, intelligent security systems, image and pattern recognition, machine vision technology, and various other cutting-edge technologies.

The artificial neural network used for these studies in finance, the Kajiji-4 network, is a radial basis function network. Though there are many different types of artificial neural networks, each with their own advantages, RBF neural networks are able to learn data sets very well in an extremely quick manner. Thus, these neural networks are very well suited for real-time data analysis, such as that required of high-frequency securities trading. The Kajiji-4 artificial neural network is also used in this respect in addition to its use as a research tool.

The first step of this project after becoming familiarized with the complexities of the radial-basis function artificial neural network (RBF-ANN) was to collect the necessary data. For the independent variables, I use two types of interest rates and the S&P 500 equity volatility index. Each of these variables is called a node, and each observation made by the RBF-ANN explains something about that node. The first interest rate I use was the 30-Day Treasury bill rate to represent the short-term interest rates throughout the economy. This rate is also commonly referred to as the risk-free rate. I also use the 20-year Treasury bond rate in order to represent the longer term interest rates throughout the economy. This variable is also significant because it corresponds to the overall sentiment in the market in regard to whether interest rates will be going up or down in the short-term. The final variable I use is the S&P Volatility Index (VIX), which is representative of the amount of volatility within the broad US equity markets. The independent variables are composed of these three variables.

The dependent variables consisted of various types of hedge fund indices which I gathered from the Dow Jones Credit Suisse Hedge Fund Indexes. Of the 11 index strategies available, I use four as my dependent variables:

- Equity Market Neutral
- Event Driven
- Global Macro
- Long/Short Equity

The managers of these four various index strategies seek to profit based on different economic conditions. Managers of equity market neutral funds seek to reduce the systematic risk of the market by taking both long and short positions in stocks in an attempt to exploit investment opportunities within a select group of stocks while maintaining neutral exposure in the broad stock market. Managers of event driven funds “typically invest in various asset classes and seek to profit from potential mispricing of securities related to a specific corporate or market event.” Managers of global macro funds attempt to profit by exploiting opportunities in asset classes such as currencies and interest rates which would be largely affected by macroeconomic and/or political trends across the globe. Finally, managers of long/short equity funds attempt to profit by holding both long and short positions in the equity markets in order to hedge their overall exposure to certain sectors of the market through diversification¹. These four various strategies represent the dependent variables which I predict through use of the previously discussed independent variables.

| | A | B | C | D | E | F | G | H |
|----|--------|-----------------------|--------------|--------------|-------------------|-----------------------|-----------------------|------------------------|
| | Date | Equity Market Neutral | Event Driven | Global Macro | Long/Short Equity | 3-Month Treasury Bill | 20-Year Treasury Bond | Volatility Index (VIX) |
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| 4 | Jan-94 | -0.520 | 3.680 | 0.170 | 1.200 | 0.087 | 0.028 | -0.182 |
| 5 | Feb-94 | 0.240 | -0.150 | -5.650 | -2.470 | 0.074 | 0.063 | 0.405 |
| 6 | Mar-94 | -0.240 | -1.290 | -4.270 | -3.900 | 0.050 | 0.056 | 0.427 |
| 7 | Apr-94 | 0.250 | -0.660 | -1.590 | -1.560 | 0.118 | 0.019 | -0.496 |
| 8 | May-94 | -0.110 | -0.150 | 3.800 | 0.570 | 0.000 | -0.004 | -0.154 |
| 9 | Jun-94 | 0.700 | -0.130 | -0.920 | -2.310 | 0.045 | 0.021 | 0.154 |
| 10 | Jul-94 | -1.000 | 0.540 | -0.350 | 1.290 | 0.034 | -0.007 | -0.241 |
| 11 | Aug-94 | -0.990 | 0.820 | 2.750 | 1.640 | 0.031 | 0.032 | 0.000 |
| 12 | Sep-94 | -0.940 | 0.750 | 0.250 | 0.250 | 0.069 | 0.026 | 0.310 |
| 13 | Oct-94 | -0.370 | -0.730 | -2.220 | 0.270 | 0.066 | 0.015 | -0.069 |
| 14 | Nov-94 | -0.300 | -1.060 | 2.090 | -2.850 | 0.057 | -0.026 | 0.134 |
| 15 | Dec-94 | 1.270 | -0.780 | 0.500 | -0.350 | 0.019 | -0.003 | -0.134 |
| 16 | Jan-95 | 0.470 | 0.510 | -2.470 | -0.740 | 0.010 | -0.031 | -0.241 |
| 17 | Feb-95 | 0.940 | 2.410 | 0.070 | 3.040 | -0.007 | -0.021 | 0.087 |

¹ Hedgeindex.com

After gathering historical data for all of the variables dating back to January 1994, I plugged it into the RBF-ANN to predict the returns.

Inputting the data into the Windows Operating Resource System with e-data and artificial intelligence (WinORS) was relatively straightforward. Upon inputting the data and marking independent variables with the character “I” and dependent variables with the character “D,” I opened the RBF dialog box. This screen is quite intuitive, as it clearly establishes the purpose of all the input areas:

Radial Basis Model Setup

*Range: AIB2:V208

Forecast:

Table: BIB4

Model: General

Radius: 1.00

Patterns: Number Percentage

Training: 69 33.3%

Validation: 138 66.7%

Simulate

67

Data Transformation

Include Target

Standardized Method 1

RBF Method

Iterative Global Ridge

Kajiji-1

Kajiji-2

Kajiji-3

Kajiji-4

Error Minimization Rule

UEV

FPE

GCV

BIC

Transfer Functions

Gaussian

Cauchy

Multiquadric

Inverse Multiquadric

Softmax

* Required Input

OK Cancel Help Expert Mode

The numbers I input for the training and validation are very important, as the neural network will not produce any meaningful solutions if it trains with too much or too little data. Training is essentially the act of learning for the neural network. By training with too much data, the neural network will become over-trained and model both true variability as well as error in the variability of the data. By training with too little data, on the other hand, the neural network may not be able to learn a sufficient amount. I chose to train using 69 of the data points, or 33.3% of the data, so that the neural network could determine its fitness by validating what it learned using the other 66.7% of the data. A model that did a good job learning the data is said

to be fit, whereas a model that did not learn the data as well is said to be unfit. The fitness of the data can be measured by the residuals, which are calculated by taking the difference between the actual returns and the predicted returns in a given time period.

For the data transformation section, I included the target variable and chose to scale the data using the Standardized Method 1. This method is a well-known approach for creating a standardized value using the following formula: $Z = (X - \mu) / \sigma$. For the algorithm I opted to use the Kajiji-4 method. This “method is an RBF ANN method constructed upon the unbiased ridge estimation (URE) of Crouse, Jin, and Hanumara for parametric models.” It is augmented by a prior information matrix which serves to supplement the Tikhonov’s regularization method.² For the transfer functions, I used the Gaussian method which is unique because it monotonically decreases with distance from the center. This method also involves transferring information between the nodes. Finally, I chose to stick with the WinORS default error minimization method of generalized cross validation (GCV). A summation of the algorithmic settings I opted to use for this project can be seen in the table below.

| Algorithmic Settings | User Choice |
|----------------------|-----------------|
| Method | <i>Kajiji-4</i> |
| Error Min. Rule | <i>GCV</i> |
| Transfer Function | <i>Gaussian</i> |

After setting all of the functions, I was able to solve the RBF-ANN and I obtained three new tabs in my WinORS workbook: RBF Parameters, RBF Weights, and RBF Predicted.

RBF Parameters

The first section of the RBF Parameters tab, the computed measures segment, is displayed as follows:

| Target | Equity Market | Event | Global | Long/Short |
|--------------------------|----------------------|---------------|---------------|-------------------|
| Computed Measures | Neutral | Driven | Macro | Equity |
| Actual Error | 1.33E-01 | 1.33E+00 | 2.13E+00 | 1.10E+00 |
| Training Error | 1.66E-03 | 1.10E-01 | 3.55E-02 | 6.54E-03 |
| Validation Error | 1.73E-03 | 4.22E-02 | 1.13E-02 | 6.36E-03 |
| Fitness Error | 1.71E-03 | 6.45E-02 | 1.92E-02 | 6.42E-03 |

² WinORS software documentation

All of the computed measures are the mean squared errors (MSE) of the given hedge fund indices. The actual error is the standard deviation of the data prior to RBF extensions. The training error is the MSE of the training data set. The validation error is the MSE of the validation set. Finally, the fitness error is the MSE of the training and validation sets. It is a positive sign for the RBF-ANN that 75% of the dependent variables learned better than they trained, as the validation error was less than the training error for all strategies except for the Equity Market Neutral strategy. This is because the Equity Market Neutral index has the lowest actual error, causing the training, validation, and fitness errors to be fractions of the other strategies' error sizes.

The second section of the RBF Parameters tab, the performance measures segment, is displayed as follows:

| Performance Measures | Equity Market Neutral | Event Driven | Global Macro | Long/Short Equity |
|-----------------------------|------------------------------|---------------------|---------------------|--------------------------|
| Direction | 0.981 | 0.932 | 0.951 | 0.990 |
| Modified Direction | 0.994 | 0.963 | 0.961 | 1.000 |
| TDPM | 0.000 | 0.007 | 0.002 | 0.001 |
| R-Square | 99.99% | 99.45% | 99.89% | 99.98% |
| AIC | -1299.784 | -555.89 | -803.838 | -1028.749 |
| Schwarz | -1289.815 | -545.921 | -793.869 | -1018.78 |
| MAPE | 10.17 | 29.71 | 14.67 | 8.23 |

This first performance measure, direction, calculates how many times the RBF-ANN correctly predicts the increasing or decreasing movement of the respective underlying index. This means that a direction of 1 would be ideal, as the RBF-ANN will have perfectly predicted the up and down movements of the underlying indexes during each time period. The RBF-ANN still did learn very well, however, as the lowest direction is only 93.2% for the Event Driven index.

Next, modified direction is very similar to direction, as it measures the same movements in a slightly different manner:

Modified Direction = ((# of correct up predictions / # of times index up) + (# of correct down predictions / # of times index down)) – 1.

The measures for modified direction are better across the board, with the Event Driven modified direction at 96.3% and the Long/Short Equity modified direction at a perfect 100%.

The Total Downtime Performance Measure (TDPM) is a correction weight that compensates for incorrect directional forecasts by overall magnitude of the movement. Smaller weights are indicative of a more accurate training phase, and in my study the largest weight is only 0.007 for the Event Driven index which also had the lowest direction and modified direction performance measures. Overall, this shows that the RBF-ANN is very efficient in predicting the up and down movements of the underlying indices.

The R-Square measure is the traditional coefficient of determination. This measure is high for all of the indexes; however it is only meaningful when the data is linear. Since this project involves working with nonlinear information, there is not much use for this statistic.

The Akaike Information Criterion (AIC) measures the overall goodness of fit within the RBF-ANN. The AIC scores are always negative numbers, with the smaller likelihoods being more negative, or further from zero.³ The Schwarz measure is another non-parametric measure that shares a similar calculation to the AIC measure, which is why the numbers are so similar. These performance measures are important for comparing different models (i.e. Standardized 1 vs. Standardized 2) but since this project does not compare the various models, the measures are not of particular significance.

Finally, the Mean Absolute Percentage Error (MAPE) is the sum of the deviations based on absolute values expressed as a percentage. On average, there is a 10% error in the forecast produced by this model, and observably smaller values are superior. Although the Equity Market Neutral and Long/Short Equity indices have average MAPEs, the Global Macro index MAPE is slightly higher and the Event Driven index MAPE is significantly higher. This does not come as a surprise, as the RBF-ANN has had the most difficulty in predicting the returns of the Event Driven index based on the majority of the afore mentioned performance measures.

³ AIC Scores as Evidence – a Bayesian Interpretation

RBF Weights

The second tab I obtained upon solving the RBF-ANN is the Weights tab. This worksheet gives the relativity of each of the independent variables to the respective dependent variables. The table is displayed as follows:

| | A | B | C | D | E |
|---|-------------------------|------------------------------|---------------------|---------------------|--------------------------|
| 1 | Variables | Weights | | | |
| 2 | | <i>Equity Market Neutral</i> | <i>Event Driven</i> | <i>Global Macro</i> | <i>Long/Short Equity</i> |
| 3 | <i>3-Month T-Bill</i> | -0.0290 | 0.1350 | -0.1840 | -0.0570 |
| 4 | <i>20-Year T-Bond</i> | 0.0130 | 0.3890 | -0.0640 | -0.0450 |
| 5 | <i>Volatility Index</i> | 0.0490 | 0.0410 | 0.0910 | -0.0510 |

It is important to note that the signs of the weights are not as significant as the absolute values. Higher absolute values indicate more predictive ability of the independent variables. The signs show whether the returns of the dependent variables move directly or inversely with the independent variable returns. For instance, increases in the T-Bill yields will explain an approximate 0.029 of the inverse movement of the Equity Market Neutral index. Overall, the independent variables are the least effective in predicting the returns of the Equity Market Neutral Index, with the most significant weight (Volatility Index) explaining a mere 0.049 of the movements. This weight of 0.049 is, however, approximately 4 times larger than the weight of the 20-Year T-Bond which carries a weight of 0.013. This shows that the VIX is approximately 4 times more important than the T-Bond in predicting the returns of Equity Market Neutral funds. The weights are only slightly better at predicting the returns of the Long/Short Equity index, with the most significant weight (3-Month T-Bill) explaining 0.057 of the movements. For the Global Macro and Event Driven indexes, however, the weights have a higher predictive ability. Increases in the interest rates of both short and long-term bonds predict significant movements in both the Global Macro and Event Driven indexes, however they move in opposite directions.

Increases in the 3-Month T-Bill yield explain 0.184 of the inverse movement of the Global Macro index. Likewise, increases in the 20-Year T-Bond explain 0.064 of the inverse movement of the same index. The inverse movement can be explained by the fact that yields and prices move in opposite directions, so increases in the interest rates cause the prices of Treasuries

to decrease along with their returns. Also, high interest rates can cause people to invest their money in risk-free assets such as Treasury securities which will drive down demand for other investments such as commodities and currencies – two of the major asset classes held by Global Macro managers. Finally, increased volatility explains 0.091 of the movement of the same index. This is because, as a whole, increased volatility allows hedge fund managers to profit through the use of derivative securities such as options and futures across all asset classes.

The most effective use of the independent variables, specifically the interest rates, was their predictive ability within the Event Driven index. Increases in both the 3-Month T-Bill and 20-Year T-Bond explained 0.135 and 0.389 respectively, of the movements in the Event Driven index. I believe this is due to the managers of these funds exploiting investment opportunities that arise due to major events surrounding monetary policy decisions. In the Global Macro section, I explained how increases in interest rates cause the returns of Treasuries to decrease. Although this still holds true, there are other implications for hedge fund managers that are not as directly exposed to asset classes such as Treasuries and commodities.

A major reason for the Fed to increase interest rates through the use of monetary policy is to attempt to slow down an overheating economy. An overheated economy occurs when people are spending their money much more than they are saving which drives up the prices of goods, the profits of businesses, and the prices of their stocks. Managers of Event Driven indexes will exploit these opportunities to invest in stocks, which will allow them to profit greatly during times of interest rate hikes.

RBF-Predicted

The third and final tab I obtained upon solving the RBF-ANN is the Predicted tab which I used to calculate the residuals. The predicted tab is displayed as follows:

| | A | B | C | D | E | F | G | H | I |
|----|--------|-----------------------|-----------|--------------|-----------|--------------|-----------|-------------------|-----------|
| 1 | | Equity Market Neutral | | Event Driven | | Global Macro | | Long/Short Equity | |
| 2 | Date | Actual | Predicted | Actual | Predicted | Actual | Predicted | Actual | Predicted |
| 3 | Jan-94 | 1.14 | 1.001 | -1.6 | -3.142 | -0.52 | 4.457 | 0.57 | 0.934 |
| 4 | Feb-94 | -4.09 | -3.878 | 2 | -0.143 | 0.24 | 4.46 | -0.44 | 0.012 |
| 5 | Mar-94 | -3.57 | -3.105 | 7.19 | 3.442 | -0.24 | 7.471 | 1.86 | 2.707 |
| 6 | Apr-94 | -1.74 | -1.906 | 1.28 | 0.244 | 0.25 | 4.033 | -0.96 | -0.655 |
| 7 | May-94 | 2.23 | 1.938 | 2.25 | 1.949 | -0.11 | 1.524 | 0.25 | 0.269 |
| 8 | Jun-94 | -0.81 | -0.83 | 6.35 | 5.125 | 0.7 | 3.556 | 0.18 | 0.425 |
| 9 | Jul-94 | 0.35 | -0.059 | -1.18 | -1.194 | -1 | 1.194 | 0.57 | 0.601 |
| 10 | Aug-94 | 2.77 | 2.695 | -5.66 | -6.207 | -0.99 | -0.015 | 1.69 | 1.781 |
| 11 | Sep-94 | 0.67 | 0.673 | 1.58 | 0.459 | -0.94 | 1.466 | 0.38 | 0.595 |
| 12 | Oct-94 | -1.35 | -1.335 | -0.41 | -1.736 | -0.37 | 2.512 | 0.2 | 0.466 |
| 13 | Nov-94 | 0.41 | 0.654 | 3.03 | 0.689 | -0.3 | 4.131 | 0.61 | 1.077 |
| 14 | Dec-94 | -0.21 | -0.555 | -0.19 | -0.862 | 1.27 | 5.202 | 0.26 | 0.514 |
| 15 | Jan-95 | -1.97 | -2.3 | 1.31 | 0.815 | 0.47 | 3.229 | 1.61 | 1.695 |
| 16 | Feb-95 | 0.45 | 0.006 | -2.24 | -2.103 | 0.94 | 2.855 | 1.88 | 1.877 |
| 17 | Mar-95 | 3.48 | 3.255 | 0.38 | 0.222 | 2.92 | 4.125 | 0.63 | 0.666 |

Based on the data I obtained in this tab, I was able to calculate the residuals by taking the difference of the actual returns and the predicted returns. I added an additional column and performed the calculations which gave me this screen:

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|----|--------|-----------------------|-----------|----------|--------------|-----------|----------|--------------|-----------|----------|-------------------|-----------|----------|
| 1 | | Equity Market Neutral | | | Event Driven | | | Global Macro | | | Long/Short Equity | | |
| 2 | Date | Actual | Predicted | Residual | Actual | Predicted | Residual | Actual | Predicted | Residual | Actual | Predicted | Residual |
| 3 | Jan-94 | -0.52 | -0.512 | -0.008 | 3.68 | 3.172 | 0.508 | 0.17 | 0.234 | -0.064 | 1.2 | 1.366 | -0.166 |
| 4 | Feb-94 | 0.24 | 0.24 | 0 | -0.15 | -0.153 | 0.003 | -5.65 | -5.649 | -0.001 | -2.47 | -2.469 | -0.001 |
| 5 | Mar-94 | -0.24 | -0.24 | 0 | -1.29 | -1.292 | 0.002 | -4.27 | -4.27 | 0 | -3.9 | -3.899 | -0.001 |
| 6 | Apr-94 | 0.25 | 0.252 | -0.002 | -0.66 | -0.693 | 0.033 | -1.59 | -1.577 | -0.013 | -1.56 | -1.548 | -0.012 |
| 7 | May-94 | -0.11 | -0.122 | 0.012 | -0.15 | -0.481 | 0.331 | 3.8 | 3.753 | 0.047 | 0.57 | 0.666 | -0.096 |
| 8 | Jun-94 | 0.7 | 0.665 | 0.035 | -0.13 | -0.335 | 0.205 | -0.92 | -1.085 | 0.165 | -2.31 | -2.269 | -0.041 |
| 9 | Jul-94 | -1 | -0.992 | -0.008 | 0.54 | 0.299 | 0.241 | -0.35 | -0.301 | -0.049 | 1.29 | 1.371 | -0.081 |
| 10 | Aug-94 | -0.99 | -1.063 | 0.073 | 0.82 | 0.223 | 0.597 | 2.75 | 2.423 | 0.327 | 1.64 | 1.782 | -0.142 |
| 11 | Sep-94 | -0.94 | -0.943 | 0.003 | 0.75 | 0.729 | 0.021 | 0.25 | 0.234 | 0.016 | 0.25 | 0.254 | -0.004 |
| 12 | Oct-94 | -0.37 | -0.405 | 0.035 | -0.73 | -1.202 | 0.472 | -2.22 | -2.376 | 0.156 | 0.27 | 0.395 | -0.125 |
| 13 | Nov-94 | -0.3 | -0.307 | 0.007 | -1.06 | -1.096 | 0.036 | 2.09 | 2.058 | 0.032 | -2.85 | -2.843 | -0.007 |
| 14 | Dec-94 | 1.27 | 1.255 | 0.015 | -0.78 | -1.124 | 0.344 | 0.5 | 0.44 | 0.06 | -0.35 | -0.252 | -0.098 |
| 15 | Jan-95 | 0.47 | 0.472 | -0.002 | 0.51 | 0.44 | 0.07 | -2.47 | -2.456 | -0.014 | -0.74 | -0.717 | -0.023 |
| 16 | Feb-95 | 0.94 | 0.922 | 0.018 | 2.41 | 2.31 | 0.1 | 0.07 | -0.019 | 0.089 | 3.04 | 3.059 | -0.019 |
| 17 | Mar-95 | 2.92 | 2.894 | 0.026 | 0.85 | 0.707 | 0.143 | 5.96 | 5.834 | 0.126 | 0.81 | 0.837 | -0.027 |

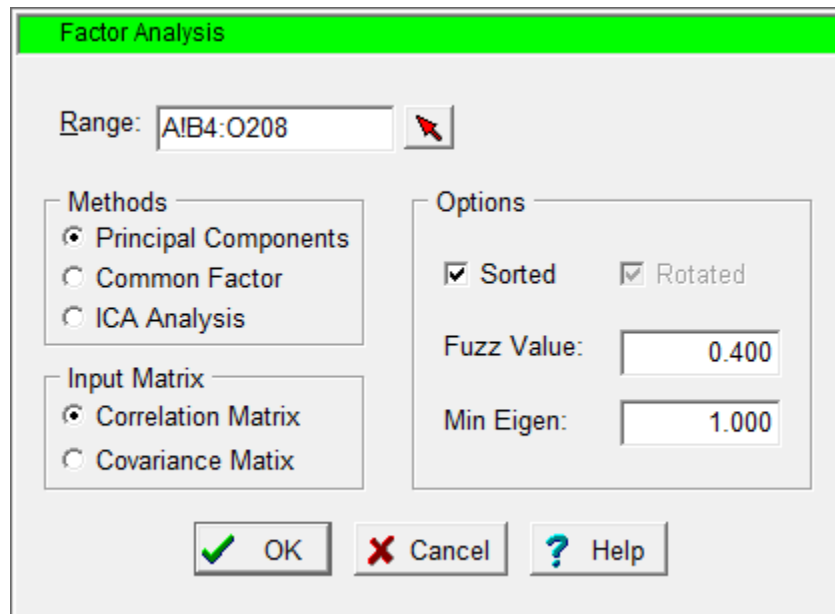
I proceeded to gather data in order to more thoroughly analyze the residuals. I calculated the maximum value, minimum value, mean, and median for each category. My outcomes are as follows:

| | Equity Market Neutral | Event Driven | Global Macro | Long/Short Equity |
|------------------|-----------------------|--------------|--------------|-------------------|
| Max Value | 0.144 | 0.765 | 0.524 | 0 |
| Min Value | -0.07 | 0 | -0.505 | -0.185 |
| Mean | 0.025 | 0.171 | 0.022 | -0.058 |
| Median | 0.014 | 0.1 | 0.005 | -0.042 |

It is interesting that the largest residual of the 205 months included in this study is only .765. More than anything else this shows what a good job the RBF-ANN did in learning the patterns prevalent in this data. Also, the average and median residual values for all hedge fund returns with the exception of the Event Driven index are less than 1 basis point – another testament to the fitness of the RBF-ANN. Although the Event Driven index did have slightly

higher residual values, they were still low overall and the RBF-ANN did a good job predicting the data across all of the indexes.

Finally, after computing all of the residuals, I was able to analyze the factors using the component analysis capabilities of WinORS:



I opted to use a Principle Component (PCA) method to examine latent correlation in the structure of the residual returns. The result of applying PCA is to locate factors which are known as principle components. That also explains why I opted to use the Correlation Matrix for the input matrix. I kept the default Fuzz Value at 0.4 (minimum display value) and I kept the Minimum Eigen Value of 1. The Eigen Value is the sum of the squared factor loadings, and it serves to explain a percentage of the variance. After setting all of these options, the Varimax Rotated Loadings tab was produced. The Varimax rotation is designed to “clean-up” factor loadings by producing either very high or very low loadings. This helps to “label” the latent dimension. The Varimax rotation displayed as follows:

| | A | B | C | D |
|---|-----------------------|----------|----------|-------|
| 1 | | Factor 1 | Factor 2 | Row % |
| 2 | | | | |
| 4 | Long Short Equity | -0.979 | . | 0.958 |
| 5 | Event Driven | 0.912 | . | 0.875 |
| 6 | Equity Market Neutral | 0.558 | -0.729 | 0.843 |
| 7 | Global Macro | . | -0.969 | 0.941 |
| 8 | Eigen Val | 2.103 | 1.515 | |
| 9 | Cum%VarExp | 0.526 | 0.904 | |

The RBF-ANN calculated a total of 2 factors which act as correlation coefficients to explain variance correlation between the variables and factors. Factor loadings are explained as if they are correlation coefficients. Correlation coefficients essentially explain statistical relationships between random variables. The correlation coefficient variables must be a number between -1 and 1. A correlation coefficient of -1 represents a perfect negative correlation, whereas a correlation coefficient of 1 represents a perfect positive correlation. Perfect negative correlation occurs when one variable moves in perfect inverse tandem with another, whereas perfect positive correlation occurs when one variable moves in perfect tandem with another.

Factor 1 was able to account for -.979 of the variance in Long/Short Equity, .912 of the variance in Event Driven, and .558 of the variance in Equity Market Neutral. Factor 2 was able to explain -.729 of the variance in Equity Market Neutral and -.969 of the variance in Global Macro. Overall, Factor 1 and Factor 2 explained a cumulative amount of 90.4% of the variance present in the four various hedge indexes. The factors loadings most effectively explained variance in the Long/Short Equity and Global Macro indices; however they still did well within all indexes. In conclusion, after taking into consideration all of the information calculated within the course of this project, I can state that the RBF-ANN truly has predictive abilities in forecasting hedge fund returns.

Works Cited

G. H. Dash Jr., C. Hanumara and N. Kajji (2003). Neural Network Architectures for Modeling FX Futures Options Volatility. Annual Meetings of the Northeast Decision Sciences Institute, Providence, Rhode Island

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