WORKING PAPER NO.280

Current Approaches in Neural Network Modeling of Financial Time Series

By

Jang Bahadur Singh

March 2009

Please address all your correspondence to:

Jang Bahadur Singh
Doctoral Student (Quantitative Methods & Information System)
Indian Institute of Management Bangalore
Bannerghatta Road
Bangalore – 560 076
E-mail: jangb06@iimb.ernet.in
CURRENT APPROACHES IN NEURAL NETWORK MODELING OF FINANCIAL TIME SERIES

Jang Bahadur Singh

Doctoral Student (Quantitative Methods & Information Systems)

Indian Institute of Management Bangalore

e-mail: jangb06@iimb.ernet.in
ABSTRACT

Neural networks are an artificial intelligence method for modeling complex non-linear functions. Artificial Neural Networks (ANNs) have been widely applied to the domain of prediction problems. Considerable research effort has gone into ANNs for modeling financial time series. This paper attempts to provide an overview of recent research in this area, emphasizing the issues that are particularly important with respect to the neural network approach to financial time series prediction task. The purpose of this paper is to provide (1) a survey of research in this area, and, (2) synthesize insights from published research on ANN modeling issues, specifically, those that have a bearing on the design factors of neural network such as variable selection, data preprocessing, and network architecture. The paper concludes with a discussion of current and emerging research topics related to neural networks in financial time series prediction.

Keywords: Time Series, Forecasting, Finance, Neural networks, Model building
INTRODUCTION

A challenging task in financial markets such as stock market and foreign exchange market is to predict the movement direction of financial markets so as to provide valuable decision information for investors. Thus, various kinds of forecasting methods have been developed by many researchers and business practitioners. Many different techniques have been applied to financial time-series forecasting over the years, ranging from conventional, model-based, statistical approaches to data-driven, experimental ones (Harris & Sollis, 2003). Some examples of the traditional statistical models are Auto Regression (AR), ARCH, and Box-Jenkins (Box & Jenkins, 1976).

Of the various statistical forecasting models related to time series data, the exponential smoothing model has been found to be one of the effective forecasting methods. The exponential smoothing models have been widely used in business and finance (Gardner, 1985; Leung et al., 2000). Gardner (1985) introduced exponential smoothing methods into supply chain management for predicting demand, and achieved satisfactory results. Leung et al. (2000) used an adaptive exponential smoothing model to predict Nikkei 225 indices, and achieved good results. However, the popular statistical method like auto regressive model and exponential smoothing models are only a class of linear model and thus it can only capture linear feature of financial time series. But financial time series are often full of nonlinearity and irregularity. Financial time series often behave nearly like a random-walk process (Taylor and Shadbolt, 2002), and are subject to regime shifting, i.e. statistical properties of the time series are different at different points in time (the process is time-varying), (Taylor and Shadbolt, 2002). Financial time series are usually very noisy, i.e. there is a large amount of random (unpredictable) day-to-day variations (Magdon-Ismail et al., 1998). Furthermore, as the smoothing constant decreases exponentially, the disadvantage of the exponential smoothing model is that it gives simplistic models that only use several previous values to forecast the future. The linear classes of statistical models are, therefore, unable to find subtle nonlinear patterns in the financial time series data. Therefore, the approximation of linear models to complex real-world problems is not
always sufficient. Hence, it is necessary to consider other nonlinear methods to improve the predictive accuracy of the forecasting model.

Much research effort has been devoted to exploring the nonlinearity of financial time series data and to developing specific nonlinear models to improve financial time series forecasting. Parametric nonlinear models such as the autoregressive random variance (ARV) model (So et al. 1999), autoregressive conditional heteroscedasticity (ARCH) model (Hsieh, 1989), and general autoregressive conditional heteroskedasticity (GARCH) models (Bollerslev, 1990), have been proposed and applied to financial time series forecasting. While these models may be good for a particular situation, they perform poorly for other applications. The pre-specification of the model form restricts the usefulness of these parametric nonlinear models since many other possible nonlinear patterns can be considered. One particular nonlinear specification may not be general enough to capture all the nonlinearities in the data. Nonparametric method has also been proposed to financial time series (Diebold and Nason, 1990).

There has been growing interest in the adoption of the state-of-the-art artificial intelligence technologies to solve the problem. One stream of these advanced techniques focuses on the use of artificial neural networks (ANNs) to analyze the historical data and provide predictions on future movements in the financial market. An ANN is a system loosely modeled on the human brain, which detect the underlying functional relationships within a set of data and perform tasks such as pattern recognition, classification, evaluation, modeling, prediction and control. ANNs are particularly well suited to finding accurate solutions in an environment characterized by complex, noisy, irrelevant or partial information. Several distinguishing features of ANNs make them valuable and attractive in forecasting. First, as opposed to the traditional model-based methods, ANNs are data-driven self adaptive methods in that there are almost no assumptions about the models for problems under study.

Neural networks have been mathematically shown to be universal approximators of functions (Funahashi 1989, Hornik et al. 1989) and their derivatives. They also can be shown to approximate ordinary least squares and nonlinear least squares regression, nonparametric
regression (Kaun and White 1994), and Fourier series analysis (Gallant and White 1992). Hence, neural networks can approximate whatever functional form best characterizes a time series.

There has been a considerable application of ANNs in the business domain. According to Wong et al. (1997a), the most frequent application domains of ANN are productions/operations (53.5 %) and finance (25.4 %); Wong et al. (1997a), Zekic (1998). According to Zekic (Zekic, 1998), who analysed various ANNs in the business domain, ANNs used for

- predicting stock performance
- forecasting foreign exchange rate
- recommendation for trading
- classification of stocks
- predicting price changes of stock indexes
- stock price prediction
- modeling the stock performance
- forecasting the performance of stock prices

dominate the research. In most analyzed applications, the NN results outperform statistical methods, such as multiple linear regression analysis, discriminative analysis and others (Zekic, 1998).

Considerable research effort has gone into ANNs for financial market applications. In this paper, I attempt to provide a survey of research in this area. Modeling financial time series using ANNs is a process that can be divided into several steps. Objective of this paper is to find out current process and practices in each step. Hence, the comparisons of various methods as reported in literature are analyzed as it appears during the whole modeling process. Some guidelines are also given where there is consensus in the literature. From insights of the published research on ANN modeling issues related to design factors of neural network such as variable selection, data preprocessing, and network architecture, general considerations for neural network research in financial time series are given.

The paper is organized as follows. Section 2 covers input selection. Sec. 3 deals with data preparation. In Sec. 4, ANN architecture variation has been discussed. Section 5 describes the
different types of the neural network. The general considerations about neural network research in financial time series is given in Sec. 6. Finally, conclusions are discussed in Sec. 7.

**INPUT SELECTION**

There are two kinds of inputs — fundamental inputs and technical inputs. Technical data, this includes such figures as past stock prices, volume, volatility etc. Actually, the term financial time series usually refers to time series of technical data. Fundamental data, these are data describing current economic activity of the company or companies whose stock prices are to be predicted. Further, fundamental data include information about current market situation as well as macroeconomic parameters.

Walczak et al. (1998) claim that multivariate inputs are necessary, most neural network inputs for financial time series prediction are univariate that means only technical data. Univariate inputs utilize data directly from time series being forecast, while multivariate inputs utilize information from outside the time series (basically fundamental inputs) in addition to the time series itself. Univariate inputs rely on the predictive capabilities of the time series itself, corresponding to a technical analysis as opposed to a fundamental analysis. For a univariate time series forecasting problem, the network inputs are the past, lagged observations of the data series and the output is the future value. Each input pattern is composed of a moving window of a fixed length along the series. In this sense, the feed-forward network used for time series forecasting is a general autoregressive model. The question is how many lag periods should be included in predicting the future. Some authors designed experiments to help selecting the number of input nodes while others adopted some intuitive of empirical ideas. Mixed results are reported in the literature. Park et al. (1996) used both linear AR model to identify the input lag structure and PCA to determine the number of hidden nodes. Balkin and Ord (2000) apply the stepwise regression approach to select the inputs to the neural networks. These methods are not so popular because in these methods, they cannot capture the nonlinear structure which is a fundamental structural property of financial time series.

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) have been used as information-based in-sample model selection criteria in selecting neural networks for financial
time series forecasting (Qi and Zhang, 2001). However, the use of these methods does not give any promising result as compared to selecting number of lag structures arbitrary.

Huang et al. (2004) propose a general approach called Autocorrelation Criterion (AC) to determine lag structures in the applications of ANNs to univariate time series forecasting. They apply the approach to the determination of input variables for foreign exchange rate forecasting and conducted comparisons between AC and information-based in sample model selection criterion. Experiment results show that AC outperformed information-based in sample model selection criterion in terms of forecasting performance.

It seems its better to employ Autocorrelation Criteria in case of univariate input, because it does not require any assumptions, completely independent of particular class of model and the selection of input is data driven which is well suited for time series problems. But most of the research reported do not use these methods, and rely on previous similar reported researches.

DATA PREPARATION

Issued should be considered for data preparation in time series application of neural networks are:

- How much Data is sufficient?
- Does the available data cover the range of the variables?
- Splitting of training and Testing Data Set

Sample size is another factor that can affect artificial neural networks forecasting ability. Neural networks researchers have used various sizes of training sets. Walczak (2001) examines the effect of different sizes of training sample sets on forecasting exchange rates. His research results indicate that for financial time series, two years of training data is sufficient to produce optimal forecasting accuracy.
There is no consensus on how to divide the data into an in sample for learning and an out-of-sample for testing. Picard and Berk (1990) suggest that 25%–50% data are used for validation for linear regression problems and if the emphasis is on parameter estimation, fewer observations should be reserved for validation. According to Chatfield (2005), forecasting analysts typically retain about 10% of the data as holdout sample. Granger (1993) recommends that, for nonlinear modeling, at least 20% of any sample should be held back for an out-of-sample evaluation. Hoptroff (1993) suggests using 10%–25% of data as the testing sample. Many other different splitting strategies have been used in the literature. As there is no consensus on the data splitting, a typical division of 70% for training and 15% each for validation and testing could be considered on very conservative note.

Time series data are difficult or impossible to split randomly. In most of the cases splitting is done by researchers’ discretion. It is important to note that the issue of data splitting is not about what proportion of data should be put in each subsample. But rather it is about an adequate sample size in each sample to ensure sufficient learning, validation, and testing. LeBaron and Weigend (1998) evaluate the effect of data splitting on time series forecasting and find that data splitting can cause more sample variation, which in turn causes the variability of forecast performance. They caution the drawback of ignoring variability across the splits and drawing too strong conclusions from such splits.

ARCHITECTURE VARIATION

Neural networks with single hidden layer have been shown to have universal approximation ability and they are also relatively easier to train. This is the reason that two or more hidden-layered networks are not very common in the literature. However, not considering more hidden layers may cause inefficiency and poor performance in neural network training and prediction, especially when a one-layer model requires a large number of hidden nodes to give desirable performance. The number of hidden nodes determines not only the network complexity to model nonlinear and interactive behavior but also the ability of neural networks to learn and generalize. Too many or too few will cause the over fitting or under fitting problem. But there is no accepted formula that can be used to calculate this parameter before training starts. It is
Almost necessary to determine the number of hidden nodes and layers by the trial and error method. It might be the reason why neural network is called as data driven methodology because hidden nodes to a large extent determine the neural network model.

Almost all studies in financial time series forecasting use one output node for both one-step forecasting and multistep forecasting. While single output node networks are suitable for one-step forecasting, they may not be effective for multistep forecasting situations as empirical findings (Meese and Geweke, 1984) suggest that a forecasting model best for a short term is not necessarily good for a long term. It is recommended that multiple output nodes be used for multistep forecasting situations. (Zhang et al, 1998).

**TYPES OF NETWORK**

Initially as noted by (Chakraborty et al., 1992), a lot of attention has been focused in MLPs in the time series forecasting domain. There have been a variety of neural networks that have been used in the domain of time series and they are listed here:

1. Recurrent (Pham & Liu, 1992; Tenti, 1995),
2. Probabilistic (Zaknich & Attikiouzel, 1991; Tan et al., 1995),
3. Radial Basis Functions (Wedding & Cios, 1996),
4. Cascade Correlation (Ensley & Nelson, 1992),
5. General Regression Neural Net (Chen, 1992),
6. Group Method of Data Handling (Pham & Liu, 1995),
7. Modular ANNs (Kimoto et al., 1991),

A very good description of the most of the listed networks can be found in Sitte and Sitte (2004) and Huang et al (2004). Following discussion will focus mainly about the results reported and the current approaches of ANN model in financial time series domain.

There are mixed result of the success of different types of the neural networks in application to the financial time series analysis. Time Delay Neural Network is the most commonly used
architecture in the domain of the time series, but there is also mixed result about its predictive ability as R Sitte and J Sitte (2000), asserts that reported work on financial time series prediction using neural networks often shows a characteristic one step shift relative to the original data. This seems to imply a failure of the neural network (NN), because a shift corresponds to a random walk prediction. Sitte and Sitte (2000) systematic analysis of different time delay neural networks predictors applied to the detrended S&P 500 time series, indicates that this prediction behavior is not a limitation of the network, but may be a characteristic of the time series.

As discussed in the characteristics of the financial time series, financial time-series data is characterized by nonlinearities, discontinuities, and high-frequency multi polynomial components, conventional ANN models are incapable of handling discontinuities in the input training data. They suffer from two further limitations: (Zhang et al. 2000)

(i) they do not always perform well because of the complexity (higher frequency and higher order nonlinearity) of the economic data being simulated, and

(ii) the neural networks function as “black boxes”, and thus are unable to provide explanations for their behaviour.

In an effort to overcome these limitations, interest has recently been expressed in using Higher Order Neural Network – HONN - models for economic data simulation (Hornik 1991; Redding 1993). Such models are able to provide information concerning the basis of the economic data they are simulating, and hence can be considered as ‘open box’ rather than ‘black box’ solutions. Furthermore, HONN models are also capable of simulating higher frequency and higher order nonlinear data, thus producing superior economic data simulations, compared with those derived from traditional ANN-based models.

HONNs have traditionally been characterized as those in which the input to a computational neuron is a weighted sum of the products of its inputs (Lee et al., 1986). Such neurons are sometimes called higher-order processing units (HPUs) (Lippmann, 1989). Zhang et al. (2000), Fulcher et al. (2006) have developed several different HONN models and these models are
termed as polynomial, trigonometric, and similar HONN models. All HONN model described in their paper utilize various combinations of linear, power, and multiplicative (and sometimes other) neuron types and are trained using standard back-propagation. The generic HONN architecture, where there are two network inputs (independent variables) $x$ and $y$, and a single network output (dependent variable) $z$. In the first hidden layer, neurons are either $\cos(x)$ or $\sin(y)$, and either $\cos^2(x)$ or $\sin^2(y)$. All neurons in the second hidden layer are multiplicative neuron. By using the functions of the neuron in the sine-cosine format and also multiplicative neurons, authors have developed the model in visible equation form, and claim to open up the "black box" or closed architecture normally associated with ANNs. In other words, they are able to associate individual network weights with polynomial coefficients of model equation. This is a significant work in the area of neural network research, since users particularly in financial domain would always like to know the explanations or justifications for the obtained results because on the basis of those results, decisions are generally made.

**GENERAL CONSIDERATIONS FOR NEURAL NETWORK RESEARCH IN FINANCIAL TIME SERIES**

Neural networks are often treated and used as black boxes. In a survey by Vellido et al. (1999), reported that the lack of explanatory capability is considered by researchers as the main shortcoming in the application of neural networks. The nature of the neural network model is such that it does not have as interpretability as some other statistical models like regression. Often, "black box" is used either as an excuse to relieve researchers from exploring further inquiries and examining the established model more rigorously or as a justification for automatic modeling so that people with little knowledge of neural networks and subject matter can do the modeling easily (Hoptroff, 1993). Current research in the neural network modeling is focusing on the higher order neural network and there has been successful attempt to put the ANN model in a visible equation form. Future researcher should focus on this area to do away with this major limitation of the neural network modeling.

In conventional financial time series forecasting, it is well known that nonstationarity can have significant impact on the analysis and forecasting and preprocessing are often necessary to make
data stationary. In the neural network literature, most studies do not consider the possible effect of nonstationarity. One should not always think that ANN model will take care of everything, certain data sets have its own characteristics, so issue like error in data set, data representativeness and shift in its characteristics during the course of time should thoroughly be analyzed.

Another general drawback in the published research is the lack of details about the model building process. As replicability is a critical principle of scientific research, lack of detail is not healthy for the neural network field itself. By reporting Heuristics involved in model building process will give empirical rules for future researcher in this domain. Zhang (2007) calls for the following minimum details in the published research in neural network modeling domain:

- The architectures experimented
- Data splitting and preprocessing,
- Training settings such as weight initialization method, learning rate, momentum, training length, stopping condition,
- Algorithm used, and model selection criterion.

If special procedures such as regularization, weight decay, or node pruning are employed, it is necessary to give the detail.

Other issue is the model evaluation and comparison, in particular financial time series literature, the random walk model often emerges as the dominant one among many linear and nonlinear methods. Random Walk model should be used as a benchmark model for comparing the results obtained from the neural network models. Statistical testing should be considered in most of the comparisons. As many comparisons are based on the same holdout sample, special matched sample statistical procedures can be used. Zhang (2007) suggests that for time series forecasting, researchers should consider more rigorous statistical test such as the Diebold–Mariano test (Diebold and Mariano, 1995).
CONCLUSION

This paper reviews applications of artificial neural networks in financial time series analysis seeks to discern trends in this research area and suggests promising avenues for future research. ANNs have been shown to be a promising tool for forecasting financial time series. Several modeling factors significantly impact the accuracy of neural network prediction. These factors include selection of input variables, data preparation, ANNs architecture and the type of network. The issues related to these factors have been discussed in this paper. General guideline is also summarized for the effective ANN modeling process in the financial time series domain.

There is no doubt that ANNs can be useful as promising modeling for financial time series modeling. However, they cannot replace all other data analysis methods from statistics. ANNs should not be seen as competing approach to statistical modeling, they can be treated as statistical method itself or can be paired with statistical analysis. For example, Zhang (2003) combined ARIMA and neural network model and reported improvement in the performance. General guidelines and good practices from statistics can be and should be followed in ANN research. Future research should attempt more on hybrid approach with more complementary techniques.
References


Fulcher, J., Zhang, M., and Xu, S. "Application of Higher-Order Neural Networks to Financial


Peter Zhang, G. "Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model"


Artificial Intelligence Applications on Wall Street, New York) 1995.


