Forecasting Wheat Futures Prices Using a Neural Network

1. Introduction

It has been observed that the major grain markets of wheat, soybeans, corn and oats move in interdependent patterns. This linking happens in part because the growing regions are greatly overlapped. Thus, the weather and general climatological factors have similar affects on all the grain crops. Also, there is some degree of substitutability in grain usage; that is, if one grain has a high price, we use another cheaper grain instead.

In spite of this observed interdependence, researchers have not had much success in forecasting short-term agricultural prices or futures. Allen [1], in a comprehensive study of all agricultural forecasting since 1900, finds, in comparing results prior to 1985 and since 1985, that there has been no significant improvement in the success rate for short-term forecasts.

Recent studies of both the short-term and long-term dynamics within the agricultural markets have indicated that wheat appears to ride the soybean oil and corn markets in the long-term [2]. That is, movements in soybean oil and corn lead the market while wheat lags slightly behind. However, statistical tests have been able to uncover this relationship only in long-term dynamics. In the short-term, say intervals of one to four days, no current tests have been able to uncover a relationship. Both econometric models and univariate methods do badly in short-term forecasting.

The focus of this proposal is the development of a neural network which can uncover the short-term relationships and thus forecast short-term wheat futures prices based on information from related agricultural markets. The backpropagation neural network has frequently been applied to classification, prediction and pattern recognition problems. Financial applications of neural networks include managing large investment funds [3], pricing gold futures [4], predicting thrift institute failure [5], estimating options prices [6], and forecasting the S&P volatility [7].

Neural networks consist of layers of connected nodes which attempt to pattern the way the brain functions. Unlike traditional methods of forecasting, neural networks do not require apriori specification of the mathematical model of the system under observation. Rather, the network itself, through repeatedly looking at the data and readjusting the importance of the data connections, attempts to discover the relationships. Since the network is data driven, its success is dependent on many data-related factors: which pieces of data to show the network, how far back in time the data should go, how many data nodes should be allowed in the model's input and hidden layers, and how much the model should be readjusted each time data prediction errors occur, to name a few. These factors are set by the researcher through an iterative process where various combinations of network architecture and parameters are constructed and tested. The appendix provides a more extensive discussion of this process for the interested readers.

2. Research Plan and Significance

The project consists of three basic components: (a) bibliographic research and data collection; (b) determination of the neural network architecture and parameters, and (c) using the resulting architecture to train networks and develop forecasts.

The bibliographic search and data collection have been completed. Daily data for a period encompassing the years 1981 through 1991 on six agricultural markets (wheat, corn, oats, soybean, soybean oil, and soybean meal), including closing prices, volume, open interest, and contract expiration times have been collected. Each series has about 2,500 observations.

The next step is to determine the network architecture and parameters. This involves a series of tests with the data in which various neural network combinations are built, tested, and progressively refined until a satisfactory model has been constructed. This should take three to four weeks. Finally, the neural network will be trained on 18 months of data and used to make daily futures forecasts for the following 6 month time period. The window of training data will then be moved forward 6 months and the training and forecasting process will be repeated. This process will continue until the entire set of data has been forecasted (6 months at a time). The results will be evaluated by calculating the mean squared error, mean absolute deviation and mean absolute percent error for each of the forecasting periods. To train, forecast and analyze the network results should take about four weeks. The results will then be prepared for publication and an article is expected by December, 1995.

The significance of this study lies in both applied and methodological areas. In the applied arena, any time prices can be predicted, market participants can make better decisions in terms of buying and selling. Thus, accurate forecasts are important to farmers, grain traders, agribusiness industries and especially to the government. The farmer uses forecasts to make decisions between immediate sale of his grain or holding the grain longer. Those in industries related to agribusiness increase profits when prices of agricultural goods can be predicted accurately. The government may choose to intervene in agricultural markets, based on forecasts, providing supports for farmers and in the interest of national well-being and security. Thus, accurate wheat futures forecasting can have implications in many areas.

In terms of methodological importance, if neural networks can be shown to be methodologically superior to current methods (time series analysis, vector autoregression, econometric models, univariate methods) in wheat futures forecasts, then this may open the door to increased short-term forecasting accuracy in other agricultural and commodity areas as well.

3. Appendix --Discussion of the Neural Network Paradigm

Neural networks are an information processing technology which model mathematical relationships between inputs and outputs. Nonlinear, multilayer, feedforward networks differ from traditional modelling techniques in several ways. Relationships between inputs and outputs are learned during a training process in which the network is repeatedly presented with historical examples. Neural networks possess the ability to approximate arbitrary mappings with no apriori assumptions about the nature of the underlying model required. Also, no assumptions about the distributions of the variables are required and the variables may be highly correlated. Based on the architecture of the human brain, a set of processing elements or neurons (nodes) are interconnected and organized in layers. These layers of nodes can be structured hierarchically, consisting of an input layer, an output layer, and middle (hidden) layers. Each connection between neurons has a numerical weight associated with it which models the influence of an input cell on an output cell. Positive weights indicate reinforcement; negative weights are associated with inhibition. Connection weights are "learned" by the network through a training process, as examples from a training set are presented repeatedly to the network. Feedforward networks map inputs into outputs with signals flowing in one direction only, from the input layer to the output layer. Each node has an activation level, specified by continuous or discrete values. If the neuron is in the input layer, its activation level is determined in response to input signals it receives from the data set. For cells in the middle or output layers, the activation level is computed as a function of the activation levels on the cells connected to it and the associated connection weights. This function is called the transfer function or activation function. The most commonly used for backpropagation is the sigmoidal or logistic function

$$f(x)=\frac{1}{1+e^{-\gamma x}} \tag{1}$$

where γ is a constant which controls the slope.

The term backpropagation technically refers to the method used to train the network, although it is commonly used to characterize the network architecture. In this learning algorithm, mean squared error and gradient descent are employed to determine a set of weights for the trained network. At each iteration, current weights are updated by minimizing the mean squared differences between the actual response of the system to a given example and the desired response. The nonlinear response functions generate gradients of the error function with respect to the weights and the chain rule is used to determine the appropriate weight changes which propagate back through the layers of the network. For more details of this method, see [8].

To develop a neural network which is capable of generalizing a relationship between inputs and outputs, the training set selected must contain a sufficient number of examples which are representative of the process which is being modelled. There is no well-defined theory to assist with the selection of input variables. Generally, heuristic method are employed. One approach is to include all the variables in the network and perform an analysis of the connection weights or a sensitivity analysis to determine which may be eliminated without reducing predictive accuracy. An alternative is to begin with a small number of variables and add new variables which improve network performance. Many decisions regarding model parameters and network topology can affect the performance of the network. The best networks are the most parsimonious, so it is important to determine the smallest number of input and hidden nodes, and therefore, the smallest number of connection weights which give satisfactory results for a given data set.

Bibliography

[1] Allen, P., "Economic Forecasting in Agriculture", *International Journal of Forecasting*, vol. 10, 1994, pp. 81-135.

[2] Malliaris, A. and J. Urrutia, "Short-term and Longterm Linkages Between Agricultural Futures Contracts", *Working Paper, School of Business, Loyola University Chicago*, June, 1993.

[3] Fishman, M., D. Barr, and W. Loic, "Using Neural Nets in Market Analysis", *Technical Analysis of Stocks and Commodities*, Vol. 9, No. 4, April, 1991, pp. 18-22.

[4] Grudnitski, G. and L. Osburn, "Forecasting S&P and Gold Futures Prices: An Application of Neural Networks", *Journal of Futures markets*, Vol 13, No. 6, September, 1993, pp. 631-643.

[5] Salchenberger, L., M. Cinar, and N. Lash, "Predicting Savings and Loans Failures Using Neural Networks", *Decision Sciences*, Vol. 23, No. 4, 1992.

[6] Malliaris, M. and L. Salchenberger, "A Neural Network Model for Estimating Option Prices", *Journal of Applied Intelligence,* No. 3, 1993, pp. 193-206.

[7] Malliaris, M. and L. Salchenberger, "Do-Ahead Replaces Run-Time: A Neural Network Forecasts Options Volatility", *Proceedings of the Tenth Conference on Artificial Intelligence for Applications*, San Antonio, Texas, March 1-4, pp. 480-481.

[8] Rumelhart, D. and J. McClelland, *Parallel Distributed Processing*, MIT Press: Cambridge, MA, 1986.