

A Neural Network Supports the Chaotic Paradigm for the S&P 500 Index

Mary E. Malliaris

Abstract

The October 1987 stock market crash has challenged the efficient market hypothesis and led to the emergence of a new and competing model of stock price time series. This new approach supports a non-random underlying structure and is labeled chaotic dynamics. If a neural network can be constructed which determines market prices better than the random walk model, it would challenge the efficient market hypothesis and support those who claim that they have found statistical evidence that a chaotic dynamics structure underlies the market. This paper constructs a neural network which lends support to the deterministic paradigm.

1. Introduction

On Black Monday, October 19, 1987, in less than four hours, the US stock markets lost 1.5 trillion dollars -- five times the amount of the US annual budget deficit at that time. Since this dramatic and surprising 23% decline, as measured by the S&P 500 Index, researchers have intensified their efforts to characterize the behavior of stock market prices. Is the behavior of the market structured? For a long time, it was believed that price changes were, to a large extent, random [Kendall (1953)]. However, the October 1987 stock market crash challenged the then prevailing financial models of the random walk and led to the emergence of a new and competing model of stock price time series. This new approach supports a non-random underlying structure and is labeled chaotic dynamics.

The chaotic dynamics approach is a deterministic model which yields a time series behavior that appears random when in fact such a series is generated by a nonlinear deterministic equation. Sometimes small input changes can produce very divergent outputs, causing a seemingly orderly system to become chaotic. The chaos appears to be random and unpredictable while it is actually following strict mathematical principles. Although deterministic, chaotic dynamics, when graphed, resembles a random walk. Preliminary statistical evidence has not succeeded in rejecting the presence of chaos in the S&P 500 Index series [Scheinkman and LeBaron (1989), Peters (1989)]. The possibility that the underlying dynamics of the S&P 500 might follow, over some time-frame, a model of low order, nonlinear, deterministic chaos, motivates the search for a neural network which can detect the existence of such a structure [Malliaris (1994)].

Neural networks, built to pattern the way a brain

learns, can develop a representation of variables and their relationships without requiring that this relationship be specified in advance. If a neural network can be constructed which predicts market prices, this would imply that the network has discovered an underlying deterministic structure in the data. Such a result would challenge the random walk hypothesis since any deterministic forecasting technique should do no better than chance if the random walk model is true.

The implications of this investigation are quite significant for several reasons. First, the results obtained utilizing neural networks should provide evidence as to the appropriateness of one of the two alternative paradigms: random walk or chaotic. As White (1989) has suggested, neural networks can be applied to advance our empirical understanding of applied disciplines. Second, support for the chaotic paradigm would imply that active management of an S&P 500 portfolio is possible since it would have been demonstrated that the S&P follows a non-linear deterministic model, at least during certain time-frames. And third, if a neural network can be developed to outperform the random walk, then researchers would be encouraged to search for expressions linking the unknown but deterministic pattern of the S&P 500 to the explanatory variables.

2. The Random Walk Model

Random walk is a statistical term used to describe dynamic behavior. In its simplest formulation we define the sequence of prices, denoted by $\{p(t): t = 0, 1, 2, \dots\}$, to follow a random walk if

$$p(t+1) = p(t) + \epsilon(t+1) \quad (1)$$

where $\epsilon(t+1)$ is the value obtained from sampling with replacement from a certain distribution with a given population mean μ and a variance of σ^2 . Equation (1) expresses tomorrow's price as a random departure from today's price, or equivalently, the price change between today and tomorrow, i.e., $p(t+1) - p(t)$, as random. It is usually assumed that $\mu = 0$.

The notion of a random walk has its methodological foundation in probability theory. Probability theory analyzes events whose outcome is uncertain in contrast to deterministic calculus where a relationship between the dependent and independent variables is exact. Long before the efficient market hypothesis was conceived, formulated

and tested, the random walk model was utilized to convey the notion that stock prices cannot be systematically forecasted [Roberts (1959)].

Earlier observations of the random behavior of stock market prices and their modeling using the random walk paradigm eventually directed researchers to seek explanations for such a statistical phenomenon. Thus, the efficient market hypothesis was developed to rationalize the random walk behavior. This hypothesis claimed that the current price $p(t)$ fully and correctly reflects all relevant information and because the flow of information between now and next period cannot be anticipated, price changes are serially uncorrelated.

Though numerous studies have confirmed market efficiency, several studies have rejected it [Pardo (1992), Bauer (1994)]. The rejections of the random walk paradigm were considered to be anomalies by efficient hypothesis researchers. However, the October 1987 stock market crash caused a serious reevaluation of the efficient market hypothesis. Actually, numerous authors are currently expressing their skepticism about the ability of the theory to explain such a major decline in the absence of any fundamental change. Shleifer and Summers (1990) are quite critical and they claim that "the stock in the efficient market hypothesis — at least as it has been traditionally formulated — crashed along with the rest of the market on October 19, 1987."

3. Determinism: The Chaotic Dynamics Methodology

If the random walk model is not a satisfactory description of stock price behavior and if prices move without any obvious change in the fundamentals of the economy, what methodological alternatives exist to explain the observed price patterns? To answer this question, a handful of quantitative economists investigated the deterministic methods of Ruelle and Takens (1971) who studied the physical problem of turbulence. These authors and the numerous physicists who followed them developed a very active field of current research called chaotic dynamics.

We do not wish to diverge into the details of chaos theory. They are skillfully presented in Gleick (1987), Vaga (1990) and elsewhere. Instead, we will offer rapidly a definition and move on to the contribution of neural networks. Consider a real-valued function $f:R \rightarrow R$. Denote by $f^2(x) = f(f(x))$ (i.e., the function applied to itself), $f^3(x) = f(f(f(x)))$, and in general let $f^n(x)$ mean the function $f(f^{n-1}(x))$. We are interested in the time series generated by this function starting from some arbitrary $x_0 \in R$. The time series which begins at x_0 takes the values

$$x_0, f(x_0), f^2(x_0), \dots, f^n(x_0), \dots \quad (2)$$

for $t = 0, 1, 2, \dots, n$. For (2) to describe a chaotic function it must satisfy three requirements [Devaney (1986)].

The most important requirement is sensitive dependence on initial conditions. This condition says that there are time series that start very close to each other but diverge exponentially fast from each other. Put differently, given a state, x , and some very close neighbor, x' , $f^n(x)$ and $f^n(x')$ will diverge exponentially from each other over time as you increase n .

The methodology for detecting chaotic dynamics in stock price changes frequently uses the Grassberger and Procaccia (1983) correlation integral to compute the correlation dimension. Several studies have computed the correlation dimension for the S&P 500 Index. For example, Scheinkman and LeBaron (1989) concluded that the correlation dimension for the S&P 500 Index appeared to be about 6 or 7, implying that such an index has nonlinear structure. The market may follow several different modes (stochastic, chaotic, etc.) and there is evidence that all of these modes exist. But the evidence of some unknown underlying structure motivates the search for nonlinear behavior through neural networks [White (1988)].

4. The Backpropagation Neural Network Approach

A neural network uses an abundance of input data that require categorizing and interpreting. It is not necessary to specify an underlying structure, as the network infers the patterns by generalizing from the interaction of the inputs. One type of architecture, or network structure, contains layers of nodes (also called neurons or processing elements) with weights and connections. The layers include an input layer, an output layer, and one or more interior layers called the middle or hidden layers. Signals travel via connections through the network from the input layer to the hidden layers to the output layer. Each node of each layer is connected to each node of the next layer.

Each connection between neurons has a numerical weight, either positive or negative, associated with it which is multiplied times the data coming from the previous layer node to which it is connected. This weight reinforces or inhibits the effect of the previous node on the next layer node. The node applies a function to the sum of the weighted inputs and computes one output signal. This function, called a transfer function, is specified before running the network and cannot be modified during the running of the neural network. The most commonly used transfer function in a backpropagation-trained network is the sigmoid function. This function is a monotonic, semi-linear function which is continuously differentiable.

The backpropagation training process involves two repeated phases: the forward phase and the backward phase. During the forward phase, the input is sent forward through the network, generating an output value for the final layer node. The difference between what the actual output should be and the network's output is computed. In the backward phase, the computed error is used

and weights are changed in proportion to the error times the input node signal. After weight adjustment, the data is again sent through the network, and the process continues until the difference of each output of the network and the actual value is below some specified level. For a detailed explanation of the training process, see Rumelhart, et. al. (1986), and for a collection of papers applying such methods to finance, see Trippi and Turban (1993).

A generally recommended method for assessing the true unbiased amount of error in a neural network model is the use of cross-validation [Weiss and Kulikowski (1991); Kuhn and Herzberg (1991)]. To use cross-validation, the data set must be divided into k distinct sets of about the same size. Each set is used independently for testing while the remaining data is utilized for training the network. Each training and testing of the network will result in a final error amount in each set. The average of these errors over all the k sets is an almost unbiased estimator of the true error rate [Weiss and Kulikowski (1991)].

5. Data

Weekly data have been collected from each Friday for two years, 1989 and 1990, on ten variables, including: the S&P 500 closing Index, the three month Treasury Bill interest rate, the thirty year Treasury Bond interest rate, weekly New York Stock Exchange volume, Money Supply as measured by both M1 and M2, Price/Earnings ratio, Gold price, Crude Oil price, and the CBOE put/call ratio. The stock market is influenced by expectations of the traders, fundamental measures of economic activity and technical factors such as trading volume. Some representative references that discuss the selection of these variables are Estrella and Hardouvelis (1991), Fama (1990) and Malliaris and Malliaris (1992).

6. Comparison Process

The efficient market hypothesis claims that the best estimate of a value for a following period is expected to be the same as the value in the current period. Since the best estimate of tomorrow's expected price is today's actual price, to have a baseline amount against which to compare the performance of the neural network, we calculated the difference between these two as the numbers to "beat" if the random walk hypothesis is to be challenged. The results are calculated for the same ten testing periods used for the network cross-validation and shown as the Mean Absolute Deviation (MAD), Mean Square Error (MSE) and the Pearson product-moment correlation between expected and actual output. After using the constructed neural network to generate output for the same weeks, the same statistics are also calculated for the network and actual differences. For the neural network to do a better job predicting than the random walk hypothesis, these comparison statistics must outperform those gener-

ated previously more than 50% of the time.

7. Neural Network Methodology

The neural network was built and refined using a commercial neural network development system which implements a multilayer perceptron (MLP) using the backpropagation training algorithm. It includes a feature that utilizes genetic algorithms to assist in the selection of optimal network architectures during the training process. When setting up a neural network, one must decide on how many nodes to have in the input and hidden layers. Because a standard MLP cannot "remember" data from a previous input data vector in the sense of a time series, it is important to include in each vector additional samples of the time series to give the network this knowledge. Each of the data input variables included two additional samples of its series, tripling the number of input nodes. Input nodes were also included for the week of the month and the month of the year. The output node was the following Friday's value of the S&P Closing Index. That is, the network was structured to give an estimate for the following Friday's S&P 500 Index.

In order to determine the optimal number of nodes in the hidden layer or layers, the program's genetic algorithm features were utilized. This program tests all possible networks within user-specified parameters. The number of hidden layers was varied between 1 and 2, with the number of nodes in each layer varying from 2 to 45. The training tolerance, originally set at the default value of .1, is a number used to compare network output to actual values (data has been scaled to be between 0 and 1). If the difference between the two is less than the tolerance, the fact is classified as "good". When the network can judge all the training facts as good, it stops training, i.e., it has converged. Lowering the tolerance forces the network to work a little harder: Convergence will take longer but the final error should be smaller and the correlation between output and actual values in the testing set may be higher. The training tolerance was lowered to 0.07. That is, no network output would be classified as good if its scaled absolute difference from the scaled actual output was greater than 0.07.

Statistics were recorded on each network configuration which allowed for comparison between the configurations. The network with the lowest Root Mean Square Error and the highest correlation had the configuration of 2 hidden layers with 24 nodes in the first hidden layer and 8 nodes in the second hidden layer.

Once the number of layers and nodes per layer was decided, the weights assigned to each node were adjusted to give the best performance. Initially, weights are random; they are refined through an evolutionary process. This was done using the program's genetic algorithm features to evolve the trained network. In this process, the weights are mutated and crossed over, based on theories of genetic evolution, to see whether a child network can

Data Set	Model	MAD (normalized)	MSE (normalized)	Correlation
A	network	2.633 (.126)	14.083 (0.675)	0.9895
	random walk	6.155 (.295)	53.203 (2.547)	0.9629
B	network	2.123 (.102)	5.897 (0.283)	0.9965
	random walk	4.887 (.234)	30.567 (1.465)	0.9580
C	network	2.506 (.120)	10.236 (0.490)	0.9848
	random walk	6.552 (.314)	61.250 (2.935)	0.9020
D	network	2.905 (.193)	10.449 (0.501)	0.9894
	random walk	3.716 (.178)	21.952 (1.052)	0.9876
E	network	4.072 (.195)	28.987 (1.389)	0.9710
	random walk	5.653 (.271)	46.875 (2.246)	0.9644
F	network	3.087 (.148)	13.132 (0.629)	0.9945
	random walk	3.496 (.167)	22.869 (1.096)	0.9836
G	network	4.000 (.192)	21.396 (1.025)	0.9595
	random walk	6.539 (.313)	50.489 (2.419)	0.9385
H	network	3.679 (.176)	21.184 (1.015)	0.9723
	random walk	4.806 (.230)	29.089 (1.398)	0.9708
I	network	3.548 (.170)	17.959 (0.861)	0.9750
	random walk	6.111 (.293)	61.577 (2.950)	0.9202
J	network	3.060 (.147)	12.718 (0.609)	0.9810
	random walk	5.036 (.241)	35.194 (1.686)	0.9465

Table 1. Comparison statistics for output Data Model MAD (normalized) MSE (normalized) Correlation Set

be evolved with the same node structure as the parent, but with superior weights. The mutation rate was set at 10%, the crossover rate at 50% and each child network was run 100 times to polish the weights. The resulting best weights were saved and used as the weights of the network.

After the structure of the network was determined, the data set was randomized and divided into 10 sets of data, each with 10 rows. This was done in order to use cross-validation to better estimate the true error. Each set was used, in turn, as a testing set, while the network was trained on the remaining data. Using cross-validation techniques, the RMS errors and the correlations between network output and actual S&P Index values were calculated for each of the ten testing sets. The average RMS error, the almost unbiased estimate of the true error, across all the sets was 0.046 (on scaled data), with an average correlation of 0.964.

8. Results

The MAD, MSE and correlation for each of the ten non-overlapping sets were calculated, comparing the actual value of the S&P to the efficient market hypothesis, which expects the next period's value to be equal to today, and to the output from the neural network. The neural network output for each of the sets was generated on the testing set using the network which had been trained on the corresponding training set. The results are shown in

Table 1 as raw score values and, in parentheses, as normalized by the standard deviation of the targets. Figures 1 and 2 plot the values generated by each model and the S&P 500 Index. As can be seen for each testing set, the neural network outperformed the random walk model in each category. The average MAD for the network was 3.167, as opposed to 6.188 for the random walk. The average MSE for the network was 15.609, as opposed to 41.32 for the random walk. We interpret such results to offer further evidence that the S&P Index may be driven by low dimensional noisy chaotic dynamics.

Conclusions

This paper has been motivated by the recent difficulties of the efficient market hypothesis to explain significant changes in stock returns that have occurred without corresponding major changes in the fundamentals of the economy. More specifically, we have argued, as have other financial researchers, that the October 1987 stock market crash, to this date, has not been reconciled with the efficient market hypotheses. While numerous researchers have criticized the random walk behavior and have sought evidence of a deterministic structure in stock market returns, no theoretical framework currently exists that describes the stock returns in a deterministic way. If such a structure were to be identified, phenomenal arbitrage opportunities would become available.

Using two years worth of weekly data of stock returns,

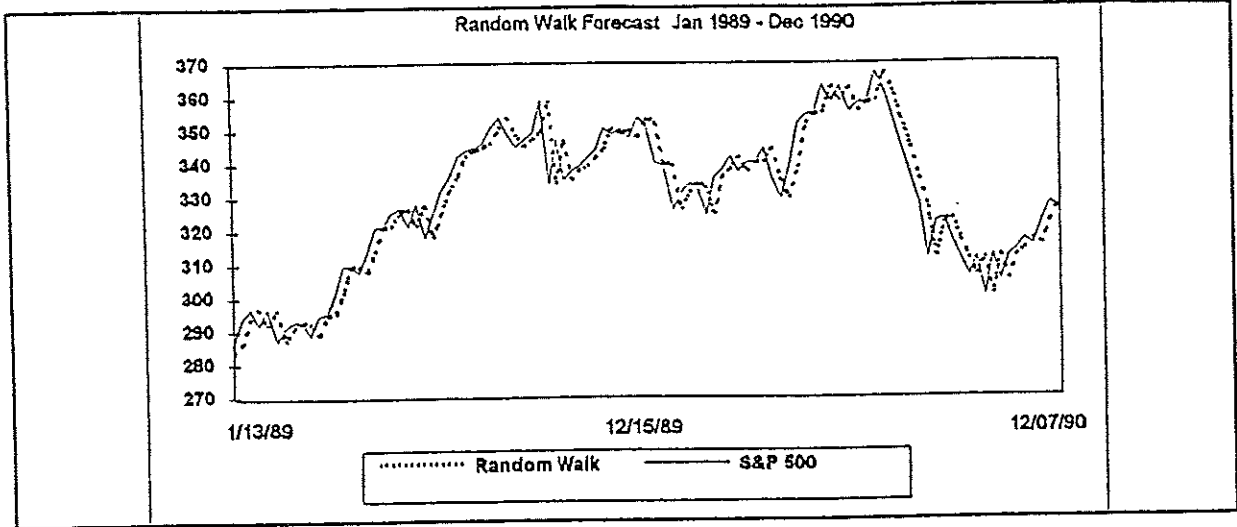


Figure 1. Random Walk Forecast

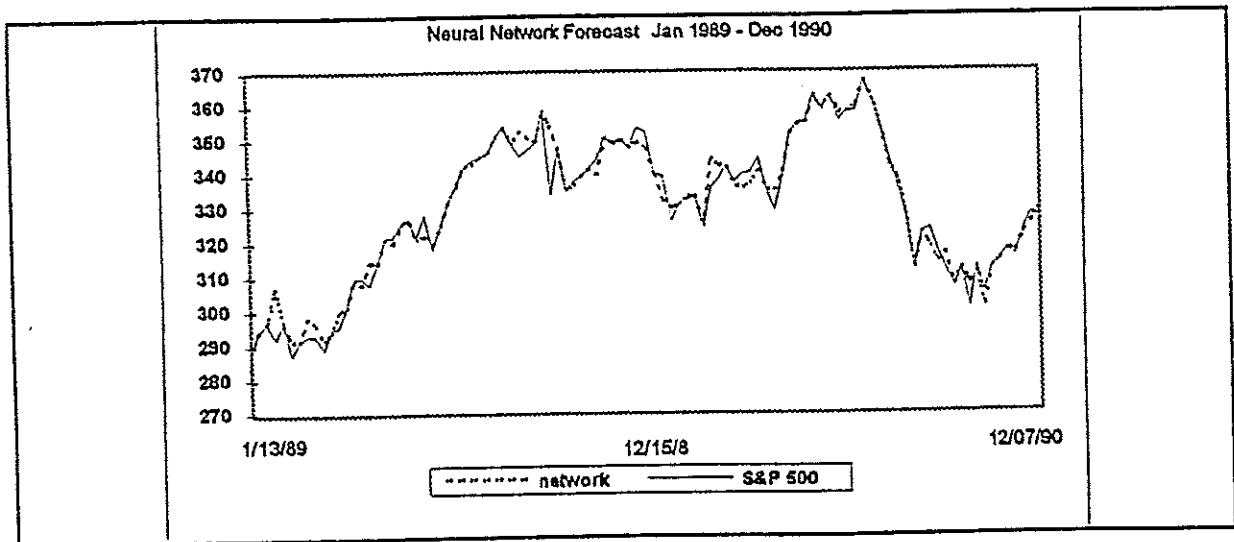


Figure 2. Neural Network Forecast

along with several carefully selected variables, we have used backpropagation neural networks and cross-validation techniques to help us gain some evidence in support of either the deterministic or random walk paradigm. The strong evidence that in ten out of ten sets of data, neural networks have outperformed the random walk model can be interpreted to be supportive of the deterministic structure of the stock market returns during our sample period. Results of this nature are encouraging to researchers who wish to develop deterministic theories which eventually may replace the existing probabilistic paradigm. For traders, who wish to build predictive networks, the set of input data may need to be longer [Thomason (1994) suggests a minimum of four years.]

Although the development of such deterministic models may take a long time to be developed, results such as

these shown here provide partial support from the neural network methodology for deterministic chaos.

Acknowledgment

I wish to acknowledge my gratitude to the Kemper Foundation for a summer research grant to the Business School of Loyola University Chicago that made this research possible. A earlier version of this paper was presented at the IEEE Tenth Conference on Artificial Intelligence for Applications, San Antonio, Texas, March, 1994.

References

- Bauer, R. [1994] *Genetic Algorithms and Investment Strategies*, John Wiley & Sons, Inc., New York.
- Devaney, R. [1986] *An Introduction to Chaotic Dynamical Sys-*

tems. Benjamin/Cummings Publishing, Menlo Park, California.

- Estrella, A. and G. Hardouvelis [1991] "The Term Structure as a Predictor of Real Economic Activity," *The Journal of Finance*, XLVI, 2, pp. 555-576.
- Fama, E. [1990] "Stock Returns, Expected Returns, and Real Activity," *The Journal of Finance*, XLV, 4, pp. 1089-1108.
- Gleick, J. [1987] *Chaos*, Penguin Books, New York.
- Grassberger, P. and I. Procaccia [1983] "Measuring the Strangeness of Strange Attractors," *Physics*, Vol. 9-D, pp. 189-208.
- Kendall, M. [1953] "The Analysis of Economic Time Series, I," *Journal of the Royal Statistical Society (Series A)*, CXVI, pp. 11-25.
- Kuhn, G. and N. Herzberg [1991] "Some Variations on Training of Recurrent Networks," in *Neural Networks, Theory and Applications*, Academic Press, Inc., Boston.
- Malliaris, A. and M. Malliaris [1992] "Interest Rates and Inflation, A Continuous Time Stochastic Approach," *Economics Letters*, Vol. 37, pp. 351-356.
- Malliaris M. [1994] "Modelling the Behavior of the S&P 500 Index: A Neural Network Approach," *Proceedings of the Tenth IEEE Conference on Artificial Intelligence for Applications*, San Antonio, Texas, March 1-4, pp. 86-90.
- Pardo, R. [1992] *Design, Testing and Optimization of Trading Systems*, John Wiley & Sons, Inc., New York.
- Peters, E. [1989] "Fractal Structure in the Capital Markets," *Financial Analysts Journal*, July/August.
- Roberts, V. [1959] "Stock-Market 'Patterns' and Financial Analysis: Methodological Suggestions," *The Journal of Finance*, Vol. 14, No. 1, pp. 1-10.
- Rumelhart, D., J. McClelland and the PDP Research Group [1986] *Parallel Distributed Processing*, MIT Press, Cambridge, Massachusetts.
- Ruelle, D., and F. Takens [1971] "On the Nature of Turbulence," *Communications of Mathematical Physics*, Vol. 20, pp. 167-192.
- Scheinkman, J., and B. LeBaron [1989] "Nonlinear Dynamics and Stock Returns," *Journal of Business*, Vol. 62, pp. 311-337.
- Shleifer, A., and L. Summers [1990] "The Noise Trader Approach to Finance," *The Journal of Economic Perspectives*, 4, pp. 19-33.
- Thomason, M. [1994] "A Basic Neural Network-based Trading System Project #1," *NEUROVEST JOURNAL*, November/December, pp. 23-27.
- Trippi, R. and E. Turban [1993] *Neural Networks in Finance and Investing*, Probus Publishing Company, Chicago.
- Vaga, T. [1990] "The Coherent Market Hypotheses," *Financial Analysts Journal*, November/December, pp. 36-49.
- Weiss, S., and C. Kulikowski [1991] *Computer Systems That Learn*, Morgan Kaufmann Publishers, Inc., San Mateo, California.
- White, H. [1988] "Economic Prediction Using Neural Networks: The case of IBM Stock Prices," *Proceedings of the IEEE Second International Conference on Neural Networks*, II, pp. 451-458.
- White, H. [1989] "Neural-network Learning and Statistics," *AI Expert*, 4, 12, p. 52.

Dr. Mary E. Malliaris is with the Department of Management Science at Loyola University Chicago.

Call for Input

Special Focus Issue

Visualization Tools for Complexity and Finance

NEUROVEST JOURNAL is planning to publish a special issue in 1996 to address important subjects related to visualization tools in regards to their application to complex financial environments, particularly neural network-based trading systems.

Potential subjects include domain (time, frequency, phase, space) transformation and visualization, multidimensional graphs, neural network development, prediction and trading system performance analysis, and others. We are seeking:

- Original papers that focus on specific applicability issues of interest to readers of the Journal,
- Letters that address items from the literature or personal experiences regarding the subject,
- Reviews of books which address the subject,
- Reviews of publicly available products that incorporate or provide analysis of the subject.

Motivation:

- To address important subjects, for the purpose of examination and clarification, in an unbiased and open forum.
- Systems developed for financial applications are becoming increasingly complex. Neural networks and other nonlinear or nonparametric tools represent an additional level of complexity. In finance, processing is almost exclusively restricted to the time domain. However, relationships between complex, nonlinear and dynamic variables can often be better represented when visualized in other domains. Therefore, the subjects are deemed worthy of a special focus issue.

Important Dates

Proposal deadline (papers, book and product reviews):

August 15, 1995

Final submission deadline for manuscripts, reviews, letters:

September 30, 1995

Submission Requirements

Send Proposals or Letters to either the email or postal address, or by Fax.

Postal Mail: NEUROVEST JOURNAL

P.O. Box 764, Haymarket, VA 22069 USA

Email: rbcaldwell@delphi.com. Fax: 703-754-0696

Send final manuscripts in both hardcopy and softcopy (papers, book and product reviews) to the postal address using the Journal's submission guidelines. Papers: 3000 to 4000 words, book and product reviews: 1500 words. See guidelines elsewhere in this or previous issues of the Journal, or request a copy.