

A hybrid neural network cybernetic system for quantifying cross-market dynamics and business forecasting

S. I. Ao

Published online: 11 March 2010
© Springer-Verlag 2010

Abstract The internal structure of a complex system can manifest itself with correlations among its components. In global business, the interactions between different markets cause collective lead–lag behavior having special statistical properties which reflect the underlying dynamics. In this work, a cybernetic system of combining the vector autoregression (VAR) and genetic algorithm (GA) with neural network (NN) is proposed to take advantage of the lead–lag dynamics, to make the NN forecasting process more transparent and to improve the NN’s prediction capability. Two business case studies are carried out to demonstrate the advantages of our proposed system. The first one is the tourism demand forecasting for the Hong Kong market. Another business case study is the modeling and forecasting of Asian Pacific stock markets. The multivariable time series data is investigated with the VAR analysis, and then the NN is fed with the relevant variables determined by the VAR analysis for forecasting. Lastly, GA is used to cope with the time-dependent nature of the co-relationships among the variables. Experimental results show that our system is more robust and makes more accurate prediction than the benchmark NN. The contribution of this paper lies in the novel application of the forecasting modules and the high degree of transparency of the forecasting process.

Keywords Neural network · Genetic algorithm · Vector autoregression · Hybrid system · Business forecasting

1 Introduction

In business organizations, forecasting is one of the most important activities that form the basis for strategic and operational decision (Zhang 2004). Traditionally, business forecasting has been dominated by linear methods, which are easy to implement and to understand. However, business forecasting is a very difficult task. With the progress of the globalization, the effects from other markets may affect markets in other regions. The linear models have serious limitation with problems of nonlinear relationships. It may be unsatisfactory to approximate the linear models for these nonlinear relationships. The complexity of the business forecasting problem paves way for the importance of machine learning and cybernetic paradigms. A main focus of machine learning algorithms is to automatically learn to recognize complex patterns, and the algorithms are able to change behavior based on data. A common type of machine learning algorithms is supervised learning, which can generally generate a global model that maps input objects to desired outputs. Neural network is a common tool employed to determine the structure of the learned function and corresponding learning algorithm. In our work, the vector autoregression (VAR) and genetic algorithm are proposed to assist in the learning and decision making of the neural network, and then the forecasting process can become more transparent and accurate.

The interdisciplinary cybernetic reasoning can be applied to understand, model and design systems for business forecasting. Cybernetics is interested in processes where an effect feeds back into its very cause. Such circularity has always been difficult to handle in science. In essence, cybernetics is concerned with component relations of the systems, like how the components differ from each other or connect to each other. A focus of cybernetics is to

S. I. Ao (✉)
International Association of Engineers, Hong Kong, China
e-mail: siao@comlab.ox.ac.uk; siao@harvard.edu

develop methods for modeling the relationships among measurable variables. In our work, the mechanism of learning and modeling is achieved with the tools from the econometrics and the machine learning. In this way, our cybernetic system can discover and model the various forms of interaction between systems, like the financial markets and the tourist markets in our two experiments.

The forecasting of tourism demand is important for various reasons. Hotel booking, preparing for the food and transportation, and the airline seat-reservation are a few examples that require accurate prediction. Relatively recently, neural network has been introduced into the tourist forecasting and it is found to be superior to other methods (Law and Au 1999). In their studies, the most recent historical value of the arrival number is used for the prediction, serving as the feeding data for the neural network.

Econometric tests have been employed by the financial experts for the analysis of the interdependence between stock markets. The econometric analysis includes methods like linear regression, nonlinear regression, generalized regression, VAR, panel data study, systems of regression equations and regressions with lagged variables, etc. Econometric analysis begins with a set of propositions about some parts of the economy. It is to specify a set of precise, deterministic relationships among variables. Soydemir (2002) determined the relationship between US stock market and macroeconomic factors like the interest rate and the stock markets in US and Latin America. The weekly data before the Mexico financial crisis in 1994 was used. Positive relationship was found for the stock markets while negative relationship of various degrees has been found between the US interest rate and Latin American stock market was observed. The relationship between the US and the Asian markets has also been studied. Briefly, in (He 2001), the monthly data between US stock and interest rate and the HK and SK stock markets were studied using ordinary least squares (OLS). In (Dekker et al. 2001), the US stock market and Asian markets were investigated with daily data using VAR. Our VAR models also confirm that the correlation between different stock markets is statistically significant.

Besides the applications of the econometric models in financial forecasting, there are also nonlinear systems for the forecasting tasks. The neural network is well known for its capability of pattern recognition and has been playing an increasingly active role in the forecasting field. The adoption of neural networks for time series analysis began in the mid-1980s, and then neural networks emerged as an important tool for business forecasting, especially for modeling complex nonlinear relationships. In financial applications, neural networks have been used for predicting exchange rate, interest rate, futures prices, capital market

index, property value, and many others. Lapedes and Farber (1987) were among the first to explicitly use the neural networks in this field. They demonstrated that feed-forward neural networks could be useful in modeling deterministic chaos. Since then, there have been more and more researches on the predictability of the neural networks for asset pricing movement. Among these researches, there are the comparisons of the neural networks with the classical regression models. The results show that in most cases the neural networks are better, or at least are equivalent to classical techniques.

Nevertheless, neural networks are also well known for its adduced incapacity to identify the relevance of independent variables in business forecasting. There are criticisms that, as the relationship among the variables is not known in advance, the neural network acts just like a black box. Studies like Jo et al. (1997) have stressed the lack of formal techniques for NNs to assess the relative relevance among the independent variables. This is a motivation for the proposed hybrid system of neural network, econometrics and genetic algorithm. In this paper, our automated system is exemplified to overcome this difficulty. With the VAR module, it can be assured that all of our feeding variables for the neural network have significant influences on the target variables. The GA module can enable the system to select the fittest models among the different models. Business markets are dynamics and their properties may be time-varying during different sub-periods. It is desirable to employ the most suitable model for each sub-period. This task may be achieved by the evolutionary computation. The evolutionary computation, including the genetic algorithm, can cope with the market dynamics. Combing these three procedures together, the hybrid VAR–NN–GA system is designed to automate the process of the selection of input variables, the numerical predictions and the evaluations of various prediction models. The aim is to supplement the previous studies, with additional information from other leading markets' movements.

Generally speaking, our cybernetic framework can work with multivariable time series data like indices of the regional countries and other business indicators. Input data, such as the trading volume, economic growth rate and currency exchange rate, etc., can also be tested by the VAR analysis. The input variables with the lowest significance level are used as the input variables for the neural network. In many cases, the performance of the neural network is time-dependent and unstable. In some sub-periods, some input variables may be more suitable for the prediction, while, in other sub-periods, they may be poorer. Neural networks with different inputs can be regarded as experts with different opinions of the relevant input factors. Their prediction performance may vary with time, as said. The selection and evaluation of these predictors can be made in

the evolutionary cycle. Experts with higher forecasting accuracy in each cycle are going to weigh more heavily in the coming round. The outline of our hybrid system's modules is as follows:

1. VAR analysis, which is to search for the correlated and leading indicators automatically;
2. Neural network prediction, which is to make forecasting from the relevant inputs determined by the VAR analysis;
3. Genetic algorithm, which is to cope with the time-dependent nature of the co-relationships among the variables and to adjust the weightings of each NN model (Fig. 1).

The outline of this paper is as follows. In Sect. 2, it is about the methodology of our hybrid modules. In Sect. 3, we will talk about the details of the application of our proposed system for tourism demand forecasting. In Sect. 4, we will describe about the application of our system for cross-market financial forecasting. Section 5 is the summary.

2 Methodology

There is a well-established tradition in forecasting research of comparing techniques on the basis of empirical results (Adya and Collopy 1998). Comparison of the business forecasts should be based on out-of-sample performance, with the testing sample different from the training sample. This practice matches with the conditions in real-world business forecasting. In a survey of 22 studies that effectively implemented the NN models for business forecasting, the NN models outperformed alternative approaches in 19 (86%) of these studies. Kamel et al. (2009) conducted a large-scale comparison study for the major machine

learning models for time series forecasting. The comparison included multilayer perception, Bayesian neural networks, radial basis functions, generalized regression neural networks, K-nearest neighbor regression, CART regression trees, support vector regression, and Gaussian processes. The best method was found to be the multilayer perception. In this paper, out-of-sample performance of a standard NN model is used as a benchmark to gauge the performance of the proposed cybernetic system.

2.1 Vector autoregression

The VAR model is a useful modeling method for business forecasting (Masih and Masih 2001), and is widely used for forecasting macroeconomic time series. VAR techniques (Enders 1995; Greene 2000) are used to understand their interactions among different variables. Masih and Masih (2001) considered the dynamic causal linkages amongst nine major international stock price indexes over the period 1982 to 1994. The results showed significant interdependencies between Asian markets and the leadership of the US markets. At the global level, the results showed the leadership of the US over both the short- and long-term. It also showed the existence of a significant short- and long-term relationship between the established OECD and the emerging Asian markets. At the regional level in Southeast Asia, the results showed the leading role of Hong Kong.

The basic idea of VAR is to treat all variables symmetrically. VAR is a multivariate system of equations that we do not need to take the dependence versus independence into account. Mathematically, a VAR model is expressed as follows (Sims 1980):

$$\vec{y}_t = \vec{C} + A(L)\vec{y}_t + \vec{e}_t \quad (1)$$

where \vec{y} is a $(n \times 1)$ vector of variables, $A(L)$ is a $(n \times n)$ polynomial matrix in the backshift operator L with lag

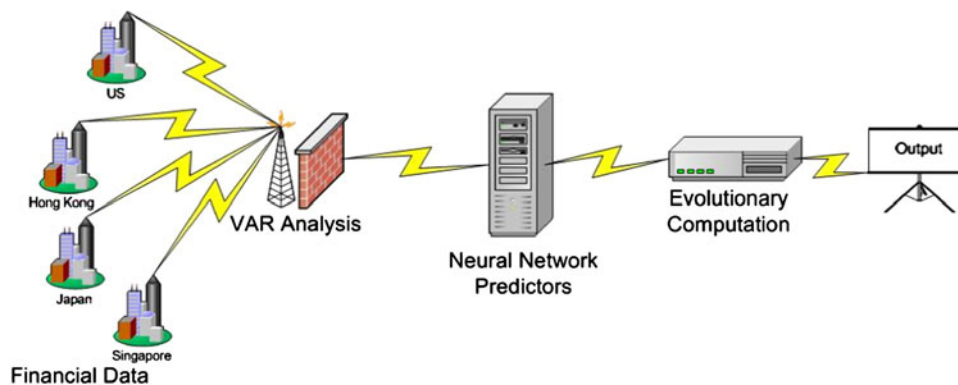


Fig. 1 System diagram of the proposed cybernetic mechanism. The multivariable time series data is firstly fed into the VAR analysis, and then the neural network prediction is to make forecasting from the relevant inputs determined by the VAR analysis. Lastly, genetic

algorithm is used to cope with the time-dependent nature of the co-relationships among the variables and to adjust the weightings of each neural network model

length p , such that $A_1L + A_2L^2 + \dots + A_pL^p$. \vec{C} is a $(n \times 1)$ vector of constant terms \vec{e} is a $(n \times 1)$ vector of white-noise error terms. For a model of n variables and n separate equations, the coefficients can be estimated by OLS. Using the same lag length for all variables produce an equation of the model which have $(n \times p) + 1$ coefficients. Statistically insignificant variables and lags can then be excluded in the model. In our experiments, the VAR models are implemented with the RATS, a fast and efficient econometrics and time series analysis software package.

Refenes et al. (1994) indicated that the conventional statistical techniques for forecasting have reached their limitation in applications with nonlinearities in the data set such as stock indices. Studies like Zhang (2003) proposed a hybrid methodology that combines both ARIMA and ANN models to take advantage of the unique strength of ARIMA and ANN models in linear and nonlinear modeling. The experimental results show that the combined model can be an effective way to improve forecasting accuracy achieved by either of the models separately. Yu et al. (2005) proposed a novel nonlinear ensemble forecasting model integrating generalized linear autoregression with artificial neural networks. Empirical results with the exchange rate data reveal that the prediction with the proposed hybrid nonlinear ensemble model can be improved. Guajardo et al. (2010) described a strategy to update support vector regression based forecasting models for time series with seasonal patterns. The most recent data was added to the training set every time a predefined number of observations take place.

It is suggested that the relationship between neural networks and the traditional statistical approaches for time series forecasting is complementary by White (1989). An advantage of VAR is that multiple variables can be investigated at the same time. Results from the VAR analysis provide the direction of interaction and the quantitative amount of interdependence among specific markets. This characteristic is suitable for our hybrid system to study the interactions among the Asian Pacific markets and the tourist markets without predefined assumption.

2.2 Neural network

The advantages of neural network include the capability to implicitly detect complex nonlinear relationships and interactions between dependent and independent variables, the requirement of less formal statistical training, and the availability of multiple training algorithms (Smith and Gupta 2000). The users may not necessarily need to predetermine the relationship between inputs and outputs with the exact functional form. Instead, it is decided by the data. Neural networks are being applied to a wide variety of tasks in many different business fields. Theoretically, it can

approximate any functional forms of the input–output pair and can be used in the regression analysis.

There is a rapid acceptance of the neural networks in the traditional domain of the operations researchers, like forecasting, modeling, clustering, and classification. The neural network is a backbone of many data mining products available. Neural networks have been successfully applied to business forecasting, a main area of the banking and finance industry, while neural networks have also been applied in various operation planning and control activities (Garetti and Taisch 1999). This nonlinear method can provide more accurate numerical forecasting. In business forecasting, the neural networks provide an attractive alternative tool for both forecasting researchers and practitioners. The traditional approaches to time series prediction, such as the Box–Jenkins or ARIMA method (Box and Jenkins 1976), need to assume that the time series under study are generated from linear process. Linear models have advantages like the ease of explaining the results. However, they may be inappropriate if the underlying mechanism is nonlinear. The formulation of a nonlinear statistical model to a particular data set is difficult as there can be too many possible nonlinear patterns.

Werbos (1974) compared the performance of the neural network trained with back-propagation with the traditional statistical methods such as regression and Box–Jenkins approaches. The neural networks outperformed the other traditional methods. Since then, there were different studies of how to implement neural networks for forecasting. Neural and neurofuzzy techniques are widely accepted to studying and evaluating stock market behavior (Atsalakis and Valavanis 2009). Most studies use the straightforward MLP networks (Kang 1991). Yao and Tan (2000) applied the neural networks in foreign exchange rates for forecasting between American Dollar and five other major currencies, Japanese Yen, Deutsch Mark, British Pound, Swiss Franc and Australian Dollar. Technical indicators and time series data were fed to neural networks to capture the underlying rules of the movement in currency exchange rates. The experiment showed that useful prediction can be made. Swanson and White (1997) applied NN for predicting future values of nine macroeconomic variables. Abraham et al. (2003) applied an artificial neural network trained using Levenberg–Marquardt algorithm, support vector machine (SVM), Takagi–Sugeno neurofuzzy model and a difference boosting neural network (DBNN) for the prediction of the Nasdaq-100 index and S&P CNX NIFTY stock index. Experimental results reveal that all the connectionist paradigms considered could represent the stock indices behavior very accurately. The results indicate that the neural network often outperforms the non-adaptive models, linear models like the VAR models, and even professional forecasters as well. Faria et al. (2009) showed that the neural networks

outperformed the adaptive exponential smoothing method in forecasting of the market movement.

The inspiring idea for the neural network is to mimic the working of our brain. It consists of axons for inputs, synapses, soma, and axons for outputs. In the typical neural network, there are three layers: the input layer, the hidden layer and the output layer. All these layers are connected and the architecture of the neural network design is itself a worthy field. The typical three-layer neural network architecture is employed in our study. The above layers correspond to the axons for inputs, synapses, soma, and axons for outputs. The mathematical structure for the neural network structure can be expressed as follows (Principe et al. 2000):

$$y = g \left(\sum_{j=1}^J w_j^{(2)} f \left(\sum_{i=1}^I w_{ji}^{(1)} x_i \right) \right) \tag{2}$$

where I denotes the number of inputs, J the number of hidden neurons, x_i the i th input, $w^{(1)}$ the weights between the input and hidden layers, $w^{(2)}$ the weights between the hidden and output layers.

The training of neural network by back-propagation algorithm has become a standard of operation in business applications, sometimes even a byword for supervised neural networks. To make comparison easier with other studies, the training of neural network in the study also uses the back-propagation algorithm:

Algorithm 1 Back-propagation algorithm of NN

1. Present the input vector patterns to the network;
2. Propagate the signals forwards, and calculate

$$u_j = a_{0j} + \sum_{i=1}^I a_{ij}x_i, v_k = b_{0k} + \sum_{j=1}^J b_{jk}y_j,$$

$$y_j = g(u_j), j = 1, \dots, J, z_k = g(v_k), k = 1, \dots, K \tag{3}$$

3. Calculate the mean square error

$$E = \frac{\frac{1}{2} \sum_{n=1}^N \sum_{k=1}^K (z_{kn} - t_{kn})^2}{NK} \tag{4}$$

4. Update the weights according to the delta rule:

$$w^{m+1} = w^m - \lambda d^m, d^m = \sum_{n=1}^N \left(\frac{\partial y}{\partial x} \Big|_m \right)_n \tag{5}$$

5. Repeat the above steps 2, 3, 4 until the error is less than the predefined value or for a predefined number of iterations.

The business applications show that neural networks are a powerful tool on their own. Despite the success of these business applications, the neural network has been criticized for its “black box” nature, proneness to overfitting, and the empirical nature of model development (Tu 1996). In the neural network which variables are the most important contributors to a particular output cannot easily be determined. A NN model may also contain a number of insignificant predictor variables that the developer fails to figure out. There are no well-established criteria for interpreting the weights in a connection weight matrix. It is an actively developing field to investigate techniques to increase the understanding of the internal logic of neural networks. One proposed technique (Baxt 1992) is to train the neural network with each input variable node removed one at a time and then to observe the effect on network performance. Others (Gnfnth et al. 1992) started to develop regression-like technique to examine the connection weights of various input variables and then determine which variables can be removed from a model without affecting its performance. Nevertheless, none of these techniques has achieved widespread use, as they do not offer the ease of interpretation of the odds ratios associated with the coefficients of a regression model. It will likely hasten their acceptance with the reducing of the black box nature of neural networks.

The performance of the neural network is expected to benefit from the assistance of other techniques in complex models (Vellido et al. 1999). In other words, they can be integrated into a more general schema of a hybrid system composed of a mixture of models. Synergy can be gained from the integration of neural networks within more general systems. The mixture of the predictions of different neural networks can improve their individual performance (Markham and Ragsdale 1995). It is these potential benefits of the synergy that motivated our works of the hybrid neural network system.

2.3 Genetic algorithm

The idea of genetic algorithm (GA) is inspired by the concept of natural evolution, which was formulated by Charles Darwin in the nineteenth century. Genetic algorithms were first developed by John Holland in the 1970s (Holland 1975). GAs are parallel, adaptive search algorithms inspired by the mechanisms of biological evolution. GA can be regarded as a broad collection of stochastic optimization algorithms that let the fittest to survive and the weak to die. The algorithms have been applied to solve a variety of optimization problems (Goldberg 1989). Genetic algorithms were used to find an analytical function

that best approximated the time variability of the studied exchange rates (Alvarez-Diaz and Alvarez 2003). In all cases, the mathematical models found by the GA predicted slightly better than the random walk model. Jeong et al. (2002) built a generic forecasting model applicable to supply chain management. A linear causal forecasting model was proposed and its coefficients were efficiently determined using the genetic algorithms. The results showed that it greatly outperformed the regression analysis.

GAs start with a population of solutions to a problem and then attempt to produce new generations of solutions that are better than the previous ones. Dual representation of individuals is needed, one for the 'genotype' (representation space), and another for the 'phenotype' (problem space). A mapping function is also needed between these two representations. In biological systems, a genotype is made up of chromosomes, and the phenotype is the actual organism formed by the interaction of the genotype with its environment. Based on evolutionary theory of natural selection, the chance of producing optimal structures is greatly enhanced by GAs. In GA, the whole solution sets are called the population while an individual solution is referred to as a chromosome. In a chromosome, there are different characteristics represented as the gene. They correspond to the different properties of an individual. There are many generations in GA. The individuals try to reproduce in each generation. The procedure codes of GA are as follows (Tettamanzi and Tomassini 2001):

Algorithm 2 The procedure codes of the genetic algorithm

1. Initialize $P(t)$;
2. Evaluate $P(t)$;
3. Recombine $P(t)$ to yield $C(t)$;
4. Evaluate $C(t)$;
5. Select $P(t + 1)$ from $P(t)$ and $C(t)$;
6. Repeat the above three steps (3, 4, 5) in the next generation $t + 1$ until the termination condition is met, where t is the order of generation, $P(t)$ is the population set at the generation t , $C(t)$ is the population set after reproduction in the generation t .

Genetic algorithms have been employed to improve the performance of neural networks. Williamson (1995) applied the GA to select the optimum topology of a neural network. GA, as a type of unstructured search, can be used to assist neural networks in the task of variable selection (Back et al. 1996). In Hansen et al. (1999), genetic algorithms were used to evolve connection topologies for NNs having a fixed number of hidden layers and a fixed number of computational units in each layer. The performance of the proposed model and autoregressive integrated moving average forecasting models is evaluated within six different time series examples. Refinements to the autoregressive

integrated moving average model improve forecasting performance over standard OLS estimation by 8–13%. In contrast, neural networks achieve dramatic improvements of 10–40%. Goh (2000) combined NNs and GAs to allow for the search of optimum NN structures for construction demand forecasting in Singapore. The combined model outperformed the basis NN model remarkably. Leigh et al. (2002) used a genetic algorithm to determine the subset of input variables to improve the R^2 correlation between the neural network estimated price increase and the actual, experienced price increase.

2.4 A cybernetic framework of hybrid VAR, neural network and genetic algorithm

Cybernetics is a science that studies the abstract principles of organization in complex systems (Heylighen and Joslyn 2001). The focus is on how systems use information, models, and control actions to steer towards and maintain their goals amongst various disturbances. This goal-directed behavior can be done with negative feedback control loops which try to achieve and maintain goal states. Cybernetics had a crucial influence on the birth of various modern sciences like artificial neural networks, computer modeling and simulation science, and dynamical systems. A simple example can be a domestic central heating system, composed of an appliance and a regulator. The regulator can adjust the heating level in accordance with the detections of changes in temperature inside the room. This negative feedback can achieve its predetermined goal.

A cybernetic system can dynamically match acquired information to selected actions relative to a computational issue that defines the essential purpose of the system or machine (Fry 2002). It is necessary that the information and control need to be quantified. In terms of automation, a simple cybernetic model can be regarded as consisting of automata, a system that transforms one set of signals into another one. Each automation receives one or more input signals which it transforms into one or more output signals. An automation represents a simple function operating on input signal or signals, while the number of input signals received and output signals sent may vary among automata. In business domain, Morgan and Hunt (2002) discussed with a conceptual and theoretical view about cybernetic systems of scenario planning in marketing strategy.

Modular hybrid systems refer to those systems that are modular in nature, i.e. they comprise several modules which can have different degrees of coupling and integration. An important feature is that they do not involve any changes regarding the conceptual operation of the individual modules (McGarry et al. 1999). A vast majority of hybrid systems fall into this category. The main reason is

that these systems are powerful processors of information and are relatively easy to implement. The hybrid VAR, neural network and genetic algorithm (VAR–NN–GA) framework can supplement its separate stand-alone components. The framework of the VAR–NN–GA is as follows:

Algorithm 3 Framework of the proposed VAR–NN–GA system

Input: multivariable time series data (MTS) like index, visitor number, GNP, etc.

Implementation procedure:

1. Pass the MTS data through VAR module;
2. Test the variables against each other to see their respective significance levels;
3. Select the variables and lag terms that are within the confidence interval;
4. Formulate the input vectors for the neural network from the above selected MTS;
5. Make one-step ahead forecasting with the NN models for the first out-of-sample time point;
6. Apply the GA module to assign dynamic weights to the NN models based on the prediction performance of the time point in step 5.

Steps 5 and 6 will then repeat with the forecasting for next out-of-sample time point, and so on, until finishing all out-of-sample time points.

The objective here is to develop a system that can take over the task of the model selection process. The system has the capability to learn which model should be employed for different sub-periods. Previous studies have recognized the close relationships between the economic learning models and models of evolutionary computation. The goal of the GA is to let the strategy with higher payoff to remain while the ones with lower payoff more likely to disappear. This methodology is suitable for our evolutionary purpose. We can image that there exist many prediction models. The evolutionary process comes into play to let the model with better prediction to survive more likely than the ones with poor records. The objective function is defined to minimize the sum of absolute errors of the predictions.

The methodology can work with input data like indices of the regional countries. Input data, like the trade volume, economic growth rate and currency exchange rate, etc., can also be tested by the VAR analysis. For the input variables with the lowest significance level, they will be used as the input variables for the neural network. The performance of the predictions made by the neural network with these input variables is time-dependent and unstable. These neural network prediction models with different inputs can be regarded as experts of different opinions of the relevant input factors. Their predictions may vary with time, as said.

The selection of these predictors can be made in the evolutionary cycle. Experts with higher forecasting accuracy in each cycle are going to weigh more heavily in the coming round.

In the following two business applications, we will show the details of how to implement this cybernetic approach to quantify the cross-market dynamics and to make accurate forecasting.

3 Business intelligence case study I: tourism demand forecasting

It is important to make accurate forecasting of tourist numbers for various reasons. Hotels, restaurants, ground transportation companies, and the airline corporations are a few examples. Because of the perishable nature of the tourism industry, the need for accurate forecasts is crucial. It is not possible to stock the unfilled airline seats, unoccupied hotel rooms, or unused facilities. Forecasting of tourism demand provides a guide to the efficient allocation of resources for the tourism industry. Tourism demand is usually measured by the number of tourist arrivals. The analysis is an integral part of sound decision making regarding investments in both the public and private sectors. It is also important for the short-term marketing decisions to promote tourism products and services. Nevertheless, there is no standard supplier of tourism forecasts (Witt and Witt 1995).

Uysal and Roubi (1999) applied the NN in tourism demand studies. The study used Canadian tourism expenditures in the US as a measure of demand to demonstrate its application. The results revealed that the use of NNs in tourism demand studies may result in better estimates in terms of prediction bias and accuracy than the multiple regression. Multiple regression using OLS has been the most widely used approach in international tourism demand analysis. Palmer et al. (2006) designed a neural network for tourism time series forecasting. The time series corresponding to tourism expenditure in the Balearic Islands (Spain), one of the world's major tourist destination, has been used. The experimental results support that the NN model can be applied successfully to tourism data forecasting. Law and Au (1999) proposed the application of a supervised feed-forward NN model to forecast Japanese tourist arrivals in Hong Kong. The experimental results showed that the NN model outperforms multiple regression, naive, moving average, and exponent smoothing.

The hybrid model of support vector machines and genetic algorithms was employed for accurately forecasting arrivals in Barbados (Pai and Hong 2005). The experimental results showed that the proposed models

outperformed the ARIMA approaches. Chen and Wang (2007) proposed a novel neural network technique, support vector regression (SVR), to tourism demand forecasting. To build an effective SVR model, SVR parameters must be set carefully. The genetic algorithm is used to search for SVR's optimal parameters. The tourist arrivals to China during 1985–2001 were employed as the data set, and the experimental results are promising.

In this work, the relationships between the current visitor number and its previous numbers were studied. The multivariate regression techniques are suitable to assist the understanding of their interactions with each other. The time series of the visitor number data is auto-regressed against its own lagged terms as well as against the time series of population and GDP. The VAR analysis assists us to identify lead-lag dynamics among the variables. Our works are the first attempt to combine the linear autoregression models with the nonlinear hybrid NN and GA modules for the tourism demand forecasting.

3.1 Quantifying cross-market dynamics

In our works, the VAR models were employed to test the interdependence of tourist numbers from different countries to Hong Kong. The monthly and quarterly tourist data of the number of visitors to Hong Kong were used. The data is from January 1978 to December 2002, available from the DataStream. Separate origins of the visitors are available for those from Japan, Southeast Asia, West Europe, USA, Australia and New Zealand, Canada, and others, denoted by HKVISJAF, HKVISASIF, HKVISWEF, HKVISUSA, HKVISANZF and HKVISCANF, respectively. The total arrival number is denoted by HKARRIVL. The GDP and population data are from the International Financial Statistics.

The tourism demand time series data of various source countries are tested with its own lagged terms as well as against the time series of population and GDP. Our results show that the patterns of these time series data are not clear. Even though some lead-lag relationship is statistically significant, its importance on the forecasting of future visitor number is not as useful as the time series data of that country itself. This is mainly in agreement with the study (Law and Au 1999). The hypothesis is that the variables between markets are independent. The significance level tells us the probability that the hypothesis is found invalid. For example, with significance level at 0.022 for the second lag term of HKVISJAPF, denoted by HKVISJAPF{2}, on HKVISJAPF, it means that there is only statistically 2.8% of the independence of these two variables. In the order words, they are likely to be correlated.

Figures 2 and 3 show us two examples of the variables: the US visitors to Hong Kong and the US population.

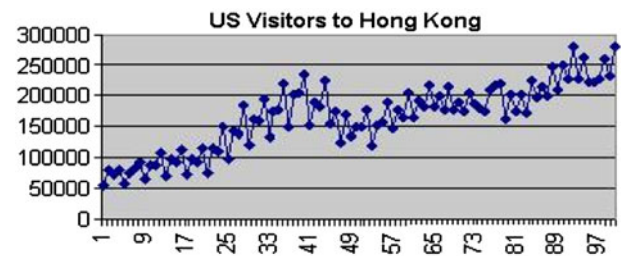


Fig. 2 Number of US Visitors to HK (per quarter, *vertical axis*), with the quarter data from January 1978 to December 2002 (*horizontal axis*)

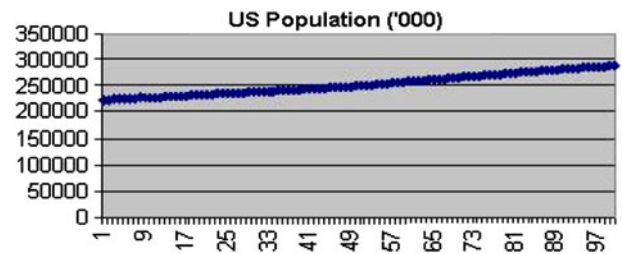


Fig. 3 US population (in thousand, per quarter, *vertical axis*), with the quarter data from January 1978 to December 2002 (*horizontal axis*)

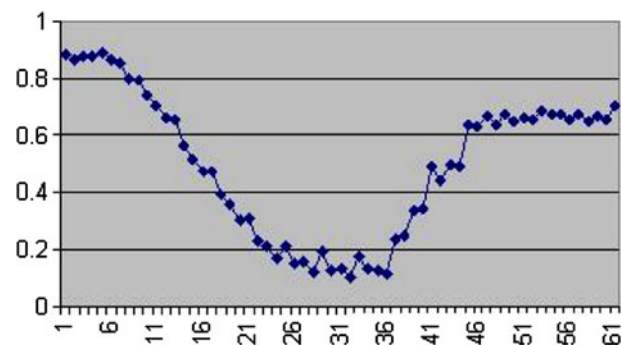


Fig. 4 Correlation of US visitor number to Hong Kong with US population (quarter data from 1978 to 2002), with the correlation formulas based on 40 nearby sample data. To compute the correlation value at a certain time, the correlation analysis is applied to the data point with its 40 nearby quarter data. The magnitude of the Pearson's correlation is represented in the *vertical axis*, with the values 0–1 for positive correlation while –1 to 0 for negative correlation. A correlation of 1 or –1 shows that the two variables studied are equivalent modulo scaling

During our modeling period, we can see that relationships among the significant variables are still far from static and it can be observed that the correlation is time-dependent. Figure 4 shows the correlation of the US population against the US visitors to Hong Kong. We can observe that, at the time of low correlation, their correlation value can be less than 0.2 and the visitor number is mainly independent of the US population, while, during other periods, with their correlation value can be larger than 0.8 and the two factors are highly correlated. Correlation value close to 1 means

two variables are highly positively correlated, while 0 means not correlated.

3.2 Experimental results

From the VAR tests performed, it is found out that the first lag term is the most appropriate feeding variable for the NN module, while the first lag term of the population size of the source market is also statistically significant. This is different from the trial-and-error approach of training and predicting with the NN and can improve the transparency of the NN module. Then, these selected lags of the variables are fed to the neural networks for the predictions. It consists of two periods 1 and 2 of different cross-market dynamics. In the period 1, the first 19th to 49th data are used for training the neural network and the next 10 are for testing. In the period 2, the first 59th to 89th data are used for training the neural network and the next 10 are for testing. It can be observed that no unique model can outperform the others over all the sub-periods. At the time of low correlation among visitor number and population (period 1), predictions with inputs of both US visitor and population (US_POP) under-perform. While at the period 2, the situation is reverse.

As seen above, the suitability of feeding variables for network may change over time, because the cross-market relationships are dynamic. A problem for the business forecasting is that we cannot know the coming correlations among the markets in advance, to automate the process to decide which model will be employed with the VAR and NN modules alone. Human expert's opinion is needed to make the decision. Here instead, the genetic algorithm is employed to simulate this process of human decision making. We are going to represent the corresponding weightings of the input variables by GA. Their respective weightings at each sub-period are decided by the number of its genes in the chromosome. For example, we have an expert of inputs of its own lag 1 term, and, another one of inputs lag 1 term plus the population lag 1 term. They are represented as gene type 1 and gene type 2, respectively. In parent chromosome 1, the ratio is as 1:1. In parent 2, the ratio of the opinion of importance is 3:5. The child 1 is reproduced from the crossover of parent 1 and parent 2, by inheriting the first half of genes from parent 1 and the second half from parent 2. The child 2 is formed from the mutation process, by stochastically selecting the third and fourth genes of parent 1 to change to other values. This mutation is to ensure that the population will be able to cover all possible opinions and has a globally suitable solution. After the cycles of reproduction and selection, we can obtain modules that fit better with the cross-market situations.

From the following table of the tourism forecasting, our hybrid NN-GA is better than the benchmark stand-alone

NN models. The overall percentage error of the proposed model is 13.47%, while the prediction errors of the two benchmark NN models are 14.49 and 13.74%, respectively. The data covers both periods 1 and 2. Results of labels 1 and 2 are the sub-period results, respectively. When the domains concerned become more complicated, for example, the stock markets, the advantage of the system will become clearer as shown in the following section.

While the cross-market relationships are highly nonlinear and dynamic, with our cybernetic model, significant cross-market input variables are identified in our VAR component, and fed into our NN model. The above results in Table 1 support that our cybernetic model is more robust and makes more accurate prediction than the stand-alone neural network. In this example, it is exemplified how an additional leading variable can be identified and employed in the neural network module, and then how the GA module can make the hybrid system robust to the changes of the cross-market dynamics.

4 Business intelligence case study II: cross-market financial forecasting

With the flow of new information into asset markets, the market prices for the assets concerned readjust to such news flows. If two markets do not react at the same time, one market will then lead the other (Bose 2007). The leading market can be viewed as contributing a price discovery function for the lagging market, and this motivates the studies of the cross-market financial forecasting studies. The stock prediction analysis derives the future stock movement from its historical movement, based on the assumption that there exists strong enough correlation for prediction. The historical data can be used directly to form the support level and the resistance or they can be plugged into many technical indicators for further investigation. There have been reports of the network's superiority on predictions. Most of these studies use the historical data of one market only, and the neural network uses the time series data for training the network to make prediction. Liao and Wang (2010) studied the statistical properties of

Table 1 Prediction performance of the US visitor number by the benchmark NN models

Prediction	US (%)	US_POP (%)	VAR-NN-GA (%)
Error (period I)	15.45	18.64	15.59
Error (period II)	13.85	10.45	12.04
Overall percentage error	14.49	13.74	13.47

One fed with US visitor time series alone, another fed with US visitor number and population), and by the proposed hybrid system, for the combined time periods

the fluctuations of the stock index and proposed the stochastic time effective neural model. In the model, the historical data are given weights depending on their time, i.e. the nearer the time of the historical data is to the present, the stronger impact the data have on the predictive model.

In the stock markets, there exist many different prediction models basing on various opinions and assumptions. For example, one may say that the market in Hong Kong is influenced heavily by its own previous movements, while others may argue that US may also have a strong influence. The studies like (He 2001; Ao 2003a) supports the latter one. There is strong correlation between the US market and the Asian markets in the long run. The VAR analysis shows that the US indices lead the Asian indices. But, in our works (Ao 2003b), such correlation is found to be time-dependent/unstable, which affects the performance of using the historical US data to predict the Asian markets by neural network.

In this case study, other Asian Pacific markets' historical data are incorporated into the neural network prediction models. In our VAR analysis, it is shown that the lead-lag relationship does exist between the markets. The neural network is employed to incorporate other leading markets' information into the prediction of the lagging markets. The result has been found to be positive in general. But, it is also found that such correlation is unstable and time-dependent. A problem that has not yet received proper attention is how to update the forecasting models when new data arrives, that is, when new events occur (Guajardo et al. 2010). Additional information of the current correlation magnitude is needed to determine which model is to be adopted. Here, a system with the neural network working together with the genetic algorithm is built to see if such a process can be automated. It is found that our inclusion of highly correlated market movements can assist the forecasting process. This is a first attempt to apply nonlinear hybrid system for the forecasting of Asian Pacific stock markets as a whole.

4.1 Quantifying the cybernetic lead-lag dynamics across different markets

The stock data cover the stock markets in US and East Asia region, namely Hong Kong (HK), Japan (JP), Australia (AU), Singapore (SG), Nasdaq (NASDAQ), S&P (PSCOMP) and Dow Jones DJINDUS (DJ) (He 2001; Baek et al. 2002; Oscar et al. 2002). The data is in daily format from 3rd May 1990 to 3rd May 2002, available from the DataStream and Reuters. The multivariable time series data are quantitatively analyzed with the VAR analysis. The econometric software RATS is used to compute the VAR results. The right-hand sides of the VAR equations contain

the lag terms of the dependent variables while the right-hand sides of the VAR equations are the independent variables of the different market indices. Each equation in the system can be estimated using OLS with the RATS. OLS estimates are consistent and asymptotically efficient. The results of the VAR include the significance test of the hypothesis of the dependence relationship between the variables. A popular threshold for the significance level is 5%, and this threshold value is used in our experiments. Our VAR results of the correlations among the Asian Pacific markets are summarized as follows:

- (1) HK depends on its past price, JP, Nasdaq, S&P and DJ;
- (2) AU depends on its past price, S&P and DJ;
- (3) SG depends on its past price, HK, Nasdaq, S&P and DJ;
- (4) JP depends on its past price, Nasdaq, S&P and DJ;
- (5) Nasdaq depends on its past price only;
- (6) DJ depends on its past price and Nasdaq;
- (7) S&P depends on its past price and Nasdaq.

With the above VAR results, we can know which variables are most suitable inputs for the neural network. For the Asian markets, the relevant information is its own historical value as well as the stock movements from the US markets. But, such relationship is far from static and it can be observed that the correlation is time-dependent. Various correlation tests have been employed by the economics experts.

In global terms, the correlations of the financial time series can be classified as correlation in space and correlations in time. It is important to study the explicit correlations in time, as they directly reflect a nature of the financial dynamics. Drozd et al. (2001b) applied the correlation matrix formalism to study dynamics of the financial evolution. In the study, the memory effects as well as some potential repeatable intra-daily structures in the financial time series were quantified. Drozd et al. (2001a) applied the correlation matrix to study the 60 companies of the Deutsche Aktien Index (DAX) and the Dow Jones (DJ) industrial average. It is found that both these markets largely merge into a single one, with the DJ taking a leading role in this emerging global market. Copikrishnan et al. (2001) analyzed the US stocks with two different cross-correlation matrices. It is found that they can partition the set of all stocks into distinct subsets. These subsets are similar to business sectors, and are stable for extended periods of time.

The following figure shows the changes of Hong Kong's correlation with S&P over the recent 10 years. Further investigation tells us that, at the time of low correlation like the late 1990s of the Asian Financial crisis, the Hong Kong market (and similarly other Asian markets) is dominated by

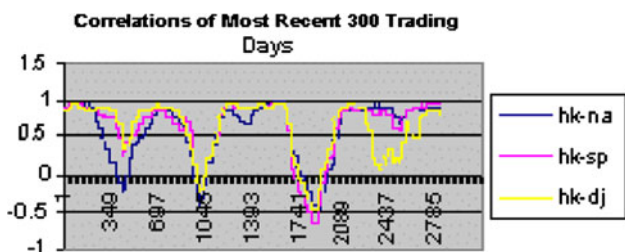


Fig. 5 Correlation of Hong Kong index with US S&P (daily data from 1990 to 2002, horizontal axis). To compute the correlation value at a certain time, the correlation analysis is applied to the data point with its 300 nearby trading days. The magnitude of the Pearson’s correlation is represented in the vertical axis, with the values 0–1 for positive correlation while –1 to 0 for negative correlation. A correlation of 1 or –1 shows that the two variables studied are equivalent modulo scaling

the local events like the currency problems. At other periods, the local markets are greatly correlated with the US markets. Numerically, at the time when highly positive correlated, the correlation value can rise close to 1, while at period with low correlation, the value can decrease to less than –0.5 (Fig. 5).

4.2 Experimental results with the benchmark stand-alone neural network

The interrelationship of the opening price of Hong Kong (HK) and Japan (JP) was analyzed with the VAR model. It can be seen that their opening price depends on its previous opening price and closing price with statistical significance. It is found that the opening price of Hong Kong’s market depends on its own previous opening and closing price over a long period (about 2,000 trading days). Similar result is identified for the Japan market too. From the VAR tests, we have found that there is a significant relation between the $t - 1$ value and the current t value of the opening price. The $t - 1$ value of closing price is also significant. NN model is set up and the results are shown below. HK_1 is the opening price (OP) prediction by its $t - 1$ price only, while HK_2 by its $t - 1$ opening and closing price. Similarly, JP_1 is by its $t - 1$ opening price while JP_2 by both its $t - 1$ opening and closing price. From the following table, it is observed that the percentage errors are reduced by about one-third by introducing one more variable identified with the VAR method (Table 2).

Table 2 Prediction performance of the opening price by NN models 1 and 2 for Hong Kong and Japan markets

Prediction	HK_1	HK_2	JP_1	JP_2
Percentage error	1.34	0.96	1.27	0.83

The forecasting results by the neural network for the Hong Kong and US S&P markets are shown in the following table. It was tested in two different sub-periods 1 and 2, respectively. Period 1 is the time with high correlation between these two markets, while period 2 is the time of low correlation. It is observed that at different periods, the prediction performances of the neural networks can vary. The NN model feeding with both the HK and S&P data can outperform the other model that was fed with HK data at the period 1, while the situation is reverse for the period 2.

These results reflect the fact that we have chosen two periods of different cross-market dynamics. But, as said in previous case study, the problem is that we cannot know the coming correlation patterns in advance. It is not possible to automate the process to decide which model will be employed. Human expert’s opinion is needed to make the decision. In our proposed system, the genetic algorithm is employed to play the role of the human expert (Table 3).

4.3 Experimental results with the VAR–NN–GA system and results comparison

From the following results of the real-world financial forecasting, it is observed that the model with our hybrid system is better than the benchmark stand-alone NN model. The data here covers both periods 1 and 2. Results with labels 1 and 2 refer to the sub-period results, respectively. The overall average error of the proposed model in this case study is 1.54%, while the prediction errors of the two benchmark NN models are 3.11 and 2.37%, respectively. The results show that the performance of the proposed system is 35% better than that of the benchmark NN model, the one feeding with HK data (Table 4).

Table 3 Prediction performance of the Hong Kong index price by NN models, one feeding with HK data, another one feeding with HK and SP data, for the two periods 1 and 2

Prediction	HK_SP_1	HK_1	HK_SP_2	HK_2
Percentage error	1.26	2.63	4.95	2.11

Table 4 Prediction performance of the Hong Kong index price by NN models, one feeding with HK and SP data, another one with HK data, versus the hybrid system for the periods 1 and 2

Prediction	HK_SP	HK	VAR–NN–GA
Period I error (%)	1.26	2.63	1.15
Period II error (%)	4.95	2.11	1.92
Overall average error (%)	3.11	2.37	1.54

5 Summary

Our proposed hybrid system is designed to automate the process of selecting input variables, to make the neural network module more transparent and to improve the prediction performance with dynamical cross-market lead-lag interactions. For the tourist demand forecasting, different models have been developed separately for making prediction. There are the nonlinear systems like neural network, which has been known for its capability of pattern recognition and has played a more and more active role in the forecasting field. Our works are the first attempt to combine the linear autoregression models with the nonlinear hybrid NN and GA modules for the tourism demand forecasting. Results show that the hybrid system is robust to the changes of the cross-market dynamics.

In the financial forecasting, previous studies have focused more or less on the historical prices and the trading volume of one market only, with either the linear statistical or the nonlinear machine learning approach. Our cybernetic system combines the advantages of both the econometrics, which can offer clear explanation and significant testing for the correlations among the variables, and the machine learning modules NN and GA, which can make our system adaptive to the changing dynamics of cross-market dependences. This is a first attempt of applying this hybrid cybernetic system for the forecasting of Asian Pacific stock markets as a whole. The results show that the performance of the proposed system is 35% better than that of the benchmark NN model.

Acknowledgments The author would like to thank Professor K.P. Lam for his advise on this topic.

References

- Abraham A, Philip N, Saratchandran P (2003) Modeling chaotic behavior of stock indices using intelligent paradigms. *Int J Neural Parallel Sci Comput* 11(1–2):143–160
- Adya M, Collopy F (1998) How effective are neural networks at forecasting and prediction? A review and evaluation. *J Forecast* 17:481–495
- Baek J et al (2002) An up-trend detection using an auto-associative neural network: KOSPI 200 futures. In: *Proc intelligent data engineering and automated learning*, Hong Kong
- Alvarez-Diaz M, Alvarez A (2003) Forecasting exchange rates using genetic algorithms. *Appl Econ Lett* 10:319–322
- Ao S (2003a) Analysis of the interaction of Asian Pacific indices and forecasting opening prices by hybrid VAR and neural network procedures. In: *Proc international conf on computational intelligence for modelling, control and automation*, Feb 2003, Vienna, Austria
- Ao S (2003b) Automating stock prediction with neural network and evolutionary computation. In: *Proc fourth international conference on intelligent data engineering and automated learning*, Hong Kong
- Atsalakis G, Valavanis K (2009) Surveying stock market forecasting techniques—Part II: soft computing methods. *Expert Syst Appl* 36(2):5932–5941
- Back B, Laitinen T, Sere K (1996) Neural networks and genetic algorithms for bankruptcy predictions. *Expert Syst Appl* 11(4):407–413
- Baxt W (1992) Analysts of the clinical variables driving decision in an artificial neural network trained to identify the presence of myocardial infarction. *Ann Emerg Med* 21:1439–1444
- Bose S (2007) Contribution of Indian index futures to price formation in the stock market. *Money & Finance* 39–56
- Box G, Jenkins G (1976) *Time series analysis: forecasting and control*. Holden-Day, San Francisco
- Chen K, Wang C (2007) Support vector regression with genetic algorithms in forecasting tourism demand. *Tour Manag* 28(1):215–226
- Copikrishnan P, Rosenow B, Plerou V, Stanley H (2001) Quantifying and interpreting collective behavior in financial markets. *Phys Rev E* 64:035106(R)
- Dekker A, Sen K, Young M (2001) Equity market linkages in the Asia Pacific region—a comparison of the orthogonalised and generalised VAR approaches. *Glob Finance J* 12:1–33
- Drozd S, Grummer F, Ruf F, Speth J (2001a) Towards identifying the world stock market cross-correlations: DAX versus Dow Jones. *Phys A* 294(1):226–234
- Drozd S, Kwapien J, Grummer F, Ruf F, Speth J (2001b) Quantifying dynamics of the financial correlations. *Phys A* 299(1–2):144–153
- Enders W (1995) *Applied econometric time series*. Wiley, USA
- Faria E, Albuquerque M, Gonzalez J, Cavalcante J, Albuquerque M (2009) Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. *Expert Syst Appl* 36(10):12506–12509
- Fry R (2002) The engineering of cybernetic systems. In: *Bayesian inference and maximum entropy methods in science and engineering*, AIP Conf Proc 617:497–528
- Garetti M, Taisch M (1999) Neural networks in production planning and control. *Prod Plan Control* 10(4):324–339
- Gnfnth J, D'Agostino R, Selker H (1992) Statistical regression techniques for the construction, interpretation and resting of computer neural networks. *Med Decis Mak* 12:343
- Goh B (2000) Evaluating the performance of combining neural networks and genetic algorithms to forecast construction demand: the case of the Singapore residential sector. *Constr Manag Econ* 18:209–217
- Goldberg D (1989) *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Reading, MA
- Greene W (2000) *Econometric analysis*. Prentice Hall, USA
- Guajardo J, Weber R, Miranda J (2010) A model updating strategy for predicting time series with seasonal patterns. *Appl Soft Comput* 10(1):276–283
- Hansen J, McDonald J, Nelson R (1999) Time series prediction with genetic-algorithm designed neural networks: an empirical comparison with modern statistical models. *Comput Intell* 15:3
- He L (2001) Time variation paths of international transmission of stock volatility—US vs Hong Kong and South Korea. *Glob Finance J* 12:79–93
- Heylighen F, Joslyn C (2001) Cybernetics and second-order cybernetics. In: *Encyclopedia of physical science and technology*, 3rd edn. Academic Press, New York
- Holland J (1975) *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, MI
- Jeong B, Jung H, Park N (2002) A computerized causal forecasting system using genetic algorithms in supply chain management. *J Syst Softw* 60:223–237

- Jo H, Han I, Lee H (1997) Bankruptcy prediction using case-based reasoning, neural network and discriminant analysis. *Expert Syst Appl* 13(2):97–108
- Kamel N, Atiya A, Gayar N, El-Shishiny H (2009) An empirical comparison of machine learning models for time series forecasting. *Econ Rev* (in press)
- Kang S (1991) An investigation of the use of feedforward neural networks for forecasting. PhD thesis, Kent State University
- Lapedes A, Farber R (1987) Non-linear signal processing using neural networks: prediction and system modeling. Technical Report LA-UR-87, Los Alamos National Laboratory
- Law R, Au N (1999) A neural network model to forecast Japanese demand for travel to Hong Kong. *Tour Manag* 20(1):89–97
- Leigh W, Purvis R, Ragusa J (2002) Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support. *Decis Support Syst* 32:361–377
- Liao Z, Wang J (2010) Forecasting model of global stock index by stochastic time effective neural network. *Expert Syst Appl* 37(1):834–841
- Markham I, Ragsdale C (1995) Combining neural networks and statistical predictions to solve the classification problem in discriminant analysis. *Decis Sci* 26(2):229–241
- Masih A, Masih R (2001) Long and short-term dynamic causal transmission amongst international stock markets. *J Int Money Finance* 20(4):563–587
- McGarry K, Wermter S, MacIntyre J (1999) Hybrid neural systems: from simple coupling to fully integrated neural networks. *Neural Comput Surv* 2:62–93
- Morgan R, Hunt S (2002) Determining marketing strategy—a cybernetic systems approach to scenario planning. *Eur J Mark* 36(4):450–478
- Oscar B, Simon S, Fernando F (2002) Non-linear forecasting methods: some applications to the analysis of financial series. In: *Progress in economics research II*, Nova Science, USA, pp 77–96
- Pai P, Hong W (2005) An improved neural network model in forecasting arrivals. *Ann Tour Res* 32(4):1138–1141
- Palmer A, Montano J, Sese A (2006) Designing an artificial neural network for forecasting tourism time series. *Tour Manag* 27(5):781–790
- Principe J, Euliano N, Lefebvre W (2000) *Neural and adaptive systems: fundamentals through simulations*. Wiley and Sons, USA
- Refenes A, Zapranis A, Francis G (1994) Stock performance modeling using neural networks: a comparative study with regression models. *Neural Netw* 5:961–970
- Sims C (1980) Macroeconomics and reality. *Econometrica* 48:1–48
- Smith K, Gupta J (2000) Neural networks in business: techniques and applications for the operations researcher. *Comput Oper Res* 27:1023–1044
- Soydemir G (2002) The impact of the movements in US three month treasury bill yields on the equity markets in Latin America. *Appl Financial Econ* 12(2):77–84
- Swanson N, White H (1997) A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks. *Rev Econ Stat* 79(4):540–550
- Tettamanzi A, Tomassini M (2001) *Soft computing, integrating evolutionary, neural, and fuzzy systems*. Springer, Germany
- Tu J (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol* 49(11):1225–1231
- Uysal M, Roubi M (1999) Artificial neural networks versus multiple regression in tourism demand analysis. *J Travel Res* 38:111–118
- Vellido A, Lisboa P, Vaughan J (1999) Neural networks in business: a survey of applications (1992–1998). *Expert Syst Appl* 17:51–70
- Werbos P (1974) *Beyond regression: new tools for prediction and analysis in the behavioral sciences*. Ph.D. thesis, Harvard University
- White H (1989) Learning in artificial neural networks: a statistical perspective. *Neural Comput* 1:425–464
- Williamson A (1995) Refining a neural network credit application vetting system with a genetic algorithm. *J Microcomput Appl* 18:261–277
- Witt S, Witt C (1995) Forecasting tourism demand: a review of empirical research. *Int J Forecast* 11(3):447–475
- Yao J, Tan C (2000) A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing* 34:79–98
- Yu L, Wang S, Lai K (2005) A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rate. *Comput Oper Res* 32(10):2523–2541
- Zhang G (2003) Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50:159–175
- Zhang G (2004) *Neural networks in business forecasting*. IGI Publishing, USA