THE VAR-NN MODEL FOR MULTIVARIATE TIME SERIES FORECASTING

Dhoriva U. Wutsqa[1]

ABSTRACT

Recently, Neural Network (NN) has been widely studied and applied in many areas, including in time series forecasting. Most existing studies have focused on the univariate cases. Here, we extend neural network application to multivariate data, particularly in time series analysis. The proposed model is Vector Autoregressive Neural Network (VAR-NN) which is generated from Vector Autoregressive (VAR) model and represents a nonlinear and nonparametric model. Evaluation of VAR-NN performance was based on train-validation sample approach. From simulation study, it can be concluded that VAR-NN yields perfect performance both in training and in testing for non-linear function approximation.

Keywords: *VAR model, VAR-NN model, multivariate time series, forecasting*

ABSTRAK

Dewasa ini, neural network (NN) secara luas telah diaplikasikan di berbagai bidang, termasuk pada peramalan time series. Pada umumnya, penelitian yang dilakukan berkaitan dengan model NN terfokus pada kasus univariat. Dalam artikel ini aplikasi NN diperluas pada data multivariat, khususnya pada masalah time series. Model yang diusulkan adalah Vector Autoregressive Neural Network (VAR-NN) yang dibentuk berdasarkan model Vector Autoregressive (VAR) yang merupakan model non linear dan nonparametrik. Evaluasi model VAR-NN dilakukan berdasarkan pendekatan sampel training dan validasi samplel testing. Hasil simulasi menunjukkan bahwa VAR-NN sangat akurat untuk pendekatan fungsi non linear baik dalam sampel training maupun testing.

Kata kunci: *model VAR, model VAR-NN, time series multivariat, peramalan*

[1] **Jurusan Matematika-FMIPA-UNY, Yogyakarta, dhoriva@yahoo.com**

INTRODUCTION

There has been a growing interest in applying neural network to many areas, particularly in time series forecasting problem. Neural network is a flexible method to approximate functional relationship. Contrary to the classical methods, neural network does not require prior knowledge of the underlying data. It is appropriate to solve problems where pre-assumption or past knowledge is difficult to be formulated. In this case, Neural network represents a non-linear and non-parametric method [see (Ripley, 1993) and (White and Swanson, 1997)].

Many studies of the forecasting performance demonstrate that neural networks have successfully estimated the non-linear model and provided high accuracy forecasts. The examples of this approach are rainfall prediction in Johor, Malaysia (Firdaus *et. al*., 2005) and airline passenger (Suhartono, *et. al*., 2005). Typical applications of neural network are in economic forecasting problems, specifically on financial data, such as stock market prediction (Chan, *et. al.*, 1999), inflation in America (Chen *et. al*., 2001), inflation in Indonesia (Brodjol Sutijo, 2005), and other financial forecasting [see (Diaz *et. al.*, 2001), (Moody, 1995), (Nikola and Jing Yang, 2000), (Ranaweera and Hubele, 1995), (Robert and Jae, 1996), and (Tkacz, 2001)].

All those studies mentioned above have focused on the univariate time series. In real world, we frequently find the time series that involve interdependency among different series of variables, which referred to as vector time series or multivariate time series. The existing method to forecast multivariate time series is Vector Autoregressive and Moving Average (VARMA) model representing a linear model and requiring tight pre-assumptions. Accordingly, we offered the more flexible approach, i.e. neural network model to estimate and to forecast the non-linear model for multivariate time series case. The proposed model is Vector Autoregressive Neural Network (VAR-NN) which generated from Vector Autoregressive (VAR) model. This model belongs to the class of a very popular neural network called Feed Forward Neural Network.

The main goal of this paper is to show that the VAR-NN model performs well in estimating and forecasting non-linear time series model. The performance is relied on mean squared error (MSE). In addition, we give the coefficient correlation to highlight the superiority of VAR-NN to predict non-linear model. To obtain the goal, we employ data simulation. Additionally, we applied the model in tourism problem including the occupation of star-hotels and non-star hotels data in Yogyakarta. To evaluate VAR-NN performance, we compare the results to those obtained from linear model VAR.

THE MODEL

Vector Autoregressive Moving Average (VARMA) Model

The VARMA model is the extension of the univariate ARIMA model, which describes relationships among several time series variables. In this model, each variable depends not only on its past values, but also on the past values of other variables. The VARMA (*p*, *q*) model can be expressed as (Wei, 1990):

$$\Phi_p(B)Z_t = \Theta_q(B)\mathrm{a}_t$$

where

$$\Phi_p(B) = I - \Phi_1 B - \Phi_2 B^2 - \cdots \Phi_p B^p$$

and

$$\Theta_q(B) = I + \Theta_1 B + \Theta_2 B^2 + \cdots \Theta_q B^q$$

Are autoregressive and moving average matrix polynomials of orders *p* and *q*, respectively, and $\{a_t\}$ is sequence of white noise vector with zero mean and constant variance. Estimating and forecasting procedures relied on autocorrelation and partial autocorrelation matrix can be seen on Box *et. al*. (1994) and Brockwell and Davis (1993).

Vector Autoregressive Neural Network (VAR-NN) Model

Feed Forward Neural Network (FFNN) is the most frequently used model in applications of time series forecasting. Typical FFNN with one hidden layer for univariate case generated from Autoregressive model called Autoregressive Neural Network (ARNN). We generalize that model for multivariate cases. The proposed model is Vector Autoregressive Neural Network (VAR-NN) derived from Vector Autoregressive (AR) model. The architecture of VAR-NN is constructed similarly to that of AR-NN, with the output layer contains more than one variable (multivariate). In this case, the input layer contains the variables of the preceding lags observations of each output variable. The nonlinear estimating is processed in the hidden layer (layer between input layer and output layer) by a transfer function. The architecture of VAR-NN in general case is complex; hence, we give the architecture with single hidden layer for bivariate case (see Figure 1.). The following discussion describes the functional form of VAR-NN for architecture of the type illustrated in Figure 1.

Let $z_t = (z_{1,t}, ..., z_{m,t})'$ be the time series process consists of *m* variables influenced by the past p lags values. The input vector can be written as follow.

$$z = (z_{1,t-1}, ..., z_{1,t-p}, ..., z_{m,t-1}, ..., z_{m,t-p}) .$$

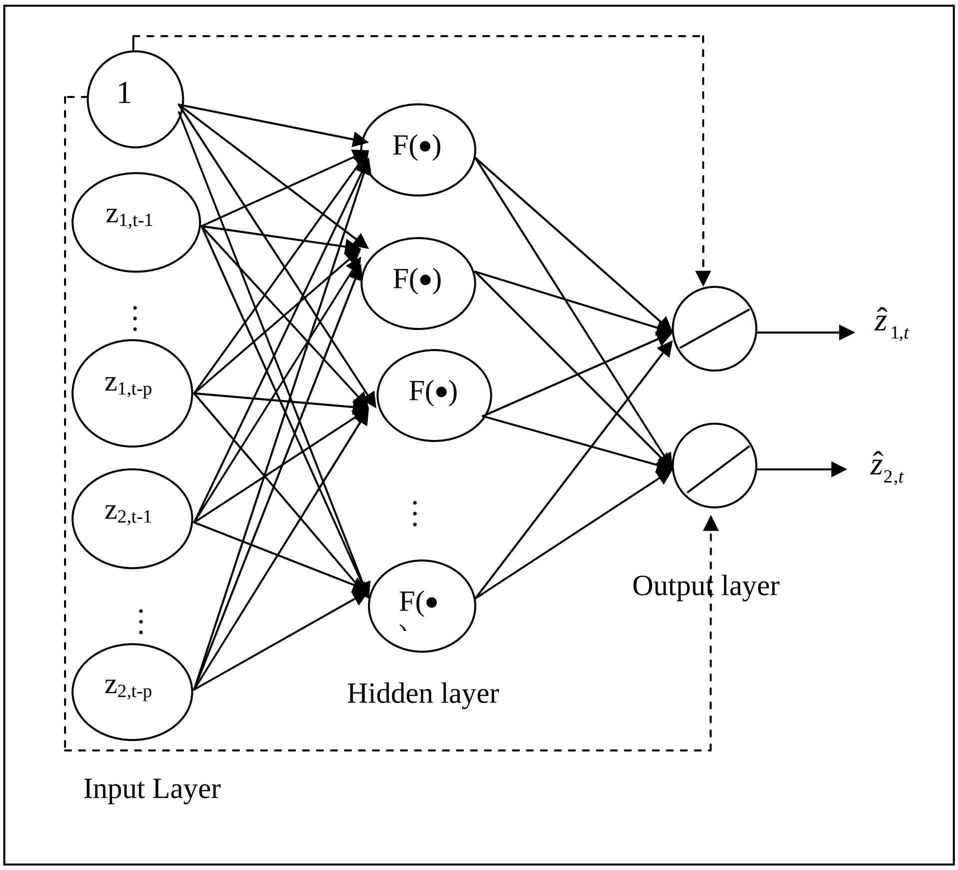

Figure 1 Architecture of VAR-NN (Bivariate Case) with Single Hidden Layer

Thus, there are $p \times m$ neurons in the input layer. If the scalar h denotes the number of hidden units then weight matrix (the network parameters) for the hidden layer has dimension $(p \times m) \times h$, where

$$w = \begin{bmatrix} w_{1,t-1,1} & w_{1,t-1,2} & \vdots & w_{1,t-1,h} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1,t-p,1} & w_{1,t-p,2} & \vdots & w_{1,t-p,h} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m,t-1,1} & w_{m,t-1,2} & \vdots & w_{m,t-1,h} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m,t-p,1} & w_{1,t-p,2} & \vdots & w_{1,t-p,h} \end{bmatrix}$$

The constant input unit is involved in the architecture, and connected to every neuron in the hidden layer, and in the output layer. This yields bias vectors $\alpha = (\alpha_1, \alpha_2, ..., \alpha_h)'$ in the hidden layer and $\beta = (\beta_1, \beta_2, ..., \beta_m)'$ in the output layer. As there are m variables in the output layer, the weight matrix for output layer takes the form

$$\lambda = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,h} \\ \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,h} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m,1} & \lambda_{m,2} & \cdots & \lambda_{m,h} \end{pmatrix}$$

Then the output of the model of VAR-NN can be defined as

$$z_t = \lambda F((zw)' + \alpha) + \beta + \varepsilon_t \tag{1.a}$$

or more explicitly

$$z_{l,t} = \sum_{k=1}^{h} \lambda_{lk} F_k (\sum_{i=1}^{m} \sum_{j=1}^{p} w_{i,t-j,k} z_{i,t-j} + \alpha_k) + \beta_l + \varepsilon_{l,t} \ , \quad l = 1, ..., m, \tag{1.b}$$

where $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t} ..., \varepsilon_{m,t})'$ is error vector and F is a transfer function operated to vector element of $zw + \alpha$. A commonly used function is logistic sigmoid function. Function F in the (1.b) model becomes

$$F_k \left(\sum_{i=1}^{m} \sum_{j=1}^{p} w_{i,t-j,k} z_{i,t-j} + \alpha_k \right) = \frac{1}{1 + \exp\left(-\left(\sum_{i=1}^{m} \sum_{j=1}^{p} w_{i,t-j,k} z_{i,t-j} + \alpha_k \right) \right)} \tag{2}$$

We apply back propagation algorithm to estimate the parameters of the neural network model, or as learning method, in term of neural network. The explanation of the learning procedure is described completely in (Haykin, 1999).

Simulation Design

The simulation study designed to highlight the superiority of VAR-NN in modeling and forecasting the nonlinear multivariate time series data. We want to show that the proposed model performs very well which indicated from the small MSE value and the coefficient correlation value that close to unity.

The Monte Carlo experiment applied to generate non-linear multivariate time series data. The data processing is done by using S-Plus soft ware. In this study, we implement the experiment only for bivariate case. The model is generated based on the Exponential smooth transition autoregressive (ESTAR) model. Since it is for univariate case, we manipulate the model for bivariate data. The functional relationship is of the form:

$$\begin{aligned} x_t &= 6,5x_{t-1}\exp[-0,25x_{t-1}^2] + 3,5y_{t-1}\exp[-0,45y_{t-1}^2] + \text{rnorm}[0,0,5] \\ y_t &= 4,5x_{t-1}\exp[-0,15x_{t-1}^2] + 5,5y_{t-1}\exp[-0,25y_{t-1}^2] + \text{rnorm}[0,0,5] \end{aligned} \quad (3)$$

The simulated data consist of 300 observations, divided into 250 observations for in-sample set (training data) and the rest for out-of-sample set (testing data). Then, VAR-NN model is implemented to in-sample set to estimate the model. Two criteria are employed to evaluate the model performance: MSE and the coefficient correlation values of both in-sample and out-of-sample data. We show the out performed of the VAR-NN models based on those values compared with the known out-of-sample MSE of the simulation model which will be referred to as the true model in the next discussion.

THE RESULT

Simulation Results

In this study, we examine the neural network model with one hidden layer, whose number of hidden units ranging from 1 to 10. Based on the model (3), the input of neural network consists of two variables, i.e. one past lag of each output variable. The neural network model estimation is employed by back propagation algorithm with transfer function (2). The results of the simulation study is illustrated in Figure 2 displaying the Mean Square Error (MSE) related to the in-sample (training) set and out-of-sample (testing) set of two variables. It appears that the MSE values approach the mean squared error (MSETM) of true model, and become stably after the size of the model is three. In addition, as shown in Figure 3, the correlation coefficient of two data sets is close to unity. This suggests that the predicted values almost exactly equal to the values of the true model. Based on this observation, we show that the neural network performs very well in forecasting the non-linear model.

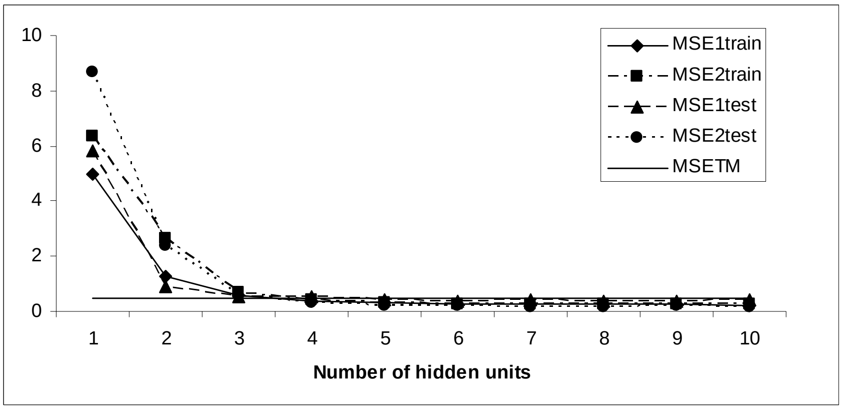

Figure 2 The Mean Square Error Values Related
to the Training and Testing Sets of Two Variables

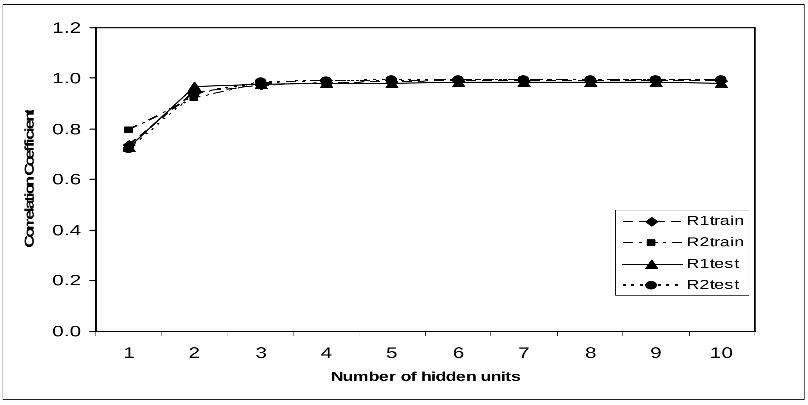

Figure 3 The Correlation Coefficients Related
to the Training and Testing Sets of Two Variables

Empirical Study in Tourism

We use the hotel occupation data of star hotels and non-star hotels in Yogyakarta started in January 1991 and ended in December 2003, so there are 156 observations. Time series plot of the standardized data indexed by month as illustrated in Figure 4. We employ standardization process on the observed data. As in simulation study, the data was divided into in-sample set (training data) and the rest for out-of-sample set (testing data). The training data was intend for estimating model, which is based on the first 120 data, and the testing data is intended for forecasting evaluation, which uses the rest data. The performance of the model is evaluated based on MSE value. Here, we use MSE value of in sample data to evaluate the accuracy of the model estimation and MSE value of out-of- sample data to evaluate the accuracy of forecasting. Thus, we compare the performance of the VARMA model and VAR-NN model based on those values. The best model provides the smallest MSE values, emphasized on MSE of out sample.

The model building of VARMA is process by implementing PROC STATESPACE program in part of SAS package. The steps of the model estimation is done through the following procedures: the identification by using Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), Matrix of Autocorrelation Function (MACF) , Matrix of Partial Autocorrelation Function (MPACF), and Akaike Information Criterion (AIC) values, and parameter estimation.

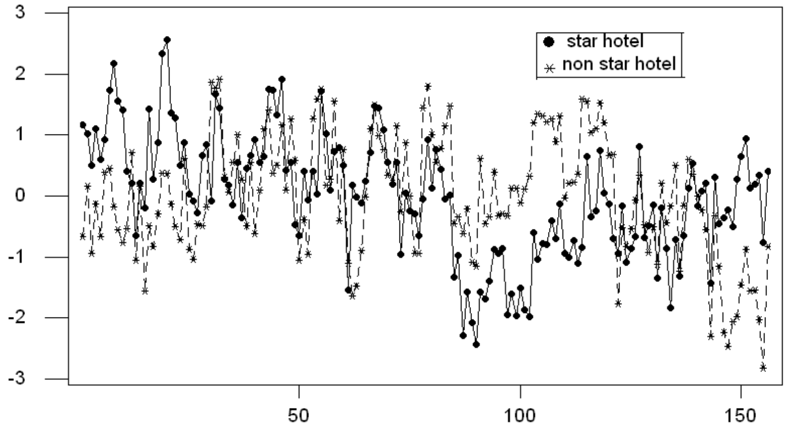

Figure 4 Time Series Plot of the Hotel Occupation Data in Yogyakarta

The result of ACF and PACF plots show that data are not stationer. Since the PACF plot yields significant values at lag 1, 2, 12, and 13, the data should be differenced non-seasonal (d = 1) and seasonal (D = 1, S = 12). After differencing, the data tends to be stationer, and then the procedure is continued by using PROC STATESPACE program. The identification of autoregressive from both MACF and the PACF pattern that tend to cut off after lag 2 leads to the tentative model is VARMA (2,2).

Then, the estimation of the parameters is performed by using *maximum likelihood estimation* (MLE). In this step, the parameters are selected by significance test. The insignificant parameters are removed from the model, and the process is continued by estimating the parameters without involving the insignificant parameters. The process is stopped until all parameters are significant. The final result is presented in Table 1.

Table 1 The Final Result of Parameter Estimation
on the Best VARMA Model

Variance Matrix for Innovation

0.398501	0.142722
0.142722	0.661674

Parameter Estimates

Parameter	Estimate	Standard Error	t Value
F(1,1)	-0.37890	0.085290	-4.44
G(3,2)	-0.57205	0.077109	-7.42

Based on that result, we obtain the best VARMA model for this case is of the form

$$\left(\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} - \begin{pmatrix} -0.38 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} z_1(t-1) \\ z_2(t-1) \end{pmatrix}\right) = \begin{pmatrix} 0 & 0 \\ 0 & -0.57 \end{pmatrix}\begin{pmatrix} a_1(t-1) \\ a_2(t-1) \end{pmatrix} + \begin{pmatrix} a_1(t) \\ a_2(t) \end{pmatrix}$$

where, z_1(t) and z_2(t) are non-seasonal(d = 1) and seasonal (D = 1, S = 12) differenced series, respectively. This model delivers the values of MSE in training sample as 0,556 and in testing sample as 0,546.

Model estimation of the type VAR-NN illustrated in Figure 1 is applied using the same data set as in model building of VARMA. To develop the VAR-NN, we consider three models with different input variables. The first model uses the same inputs as those of VARMA model, i.e. inputs lag 1, 2, 12, 13, and 14. The input pattern of the second and the third models are (1, 12, 13, 14), and (1, 12, 13), respectively. Then, we investigate all those different models with different number of hidden units. We concentrate on the MSE value of out-of sample set to evaluate the forecast performance of the models and to choose the best model in each model investigated.

The results of the VAR-NN together with the corresponding values for VARMA model are summarized in Table 2. Note that the MSE V1 represents the MSE value of the star hotels variable, similarly, MSE V2 represents the MSE value of the non-star hotels variable. The average of MSE V1 and MSE V2 corresponds to the MSE value of model. Not surprisingly that the performance of VAR-

NN models in training data are satisfy, indicated by their smaller MSE values than that achieved from VARMA model. The out-of-sample performance of VAR-NN models tend to over fitting. However, we find that one of the models observed (VAR-NN with input lag 1, 12, 13) yields better performance than VAR model in out-of-sample set. For this reason, we recommended VAR-NN with input lag 1, 12, 13 as the best model to forecast the hotel occupation in Yogyakarta.

Table 2 The Results of VAR-NN and VARMA Models

Model	MSE of In sample data			MSE of Out-of- sample data		
	V1	V2	Average	V1	V2	Average
VAR-NN with input lag 1, 2, 12, 13, 14	0,200	0,15	0,18	0,436	1,036	0,786
VAR-NN with input lag 1, 12, 13, 14	0,269	0,376	0,323	0,308	1,327	0,817
VAR-NN with input lag 1, 12, 13	0,242	0,290	0,266	0,359	0,637	0,498
VARMA	0, 399	0,662	0,556	0,308	0,785	0,546

CONCLUSIONS

It has been demonstrated that the proposed VAR-NN can estimate non-linear function perfectly and gives forecasts accurately in Monte Carlo simulation study. The empirical results on tourism problem show the superiority of VAR-NN in training compared with VARMA model, but the models tend to over fitting in testing. However, the best model of VAR-NN with input lag 1, 12, 13 is better than VARMA model in both training and testing. We suggest VAR-NN with input lag 1, 12, 13 as the best model to forecast the hotel occupation in Yogyakarta. From the results, we may also see that the input pattern influences the forecast accuracy. Further research is necessary to develop inputs selection procedure that significantly including in the model to gain appropriate forecast performance for multivariate time series case.

REFERENCES

Box, G.E.P., Jenkins G.M., and Reinsel G.C. 1994. *Time Series Analysis, Forecasting and Control.* 3rd Edition. Englewood Cliffs: Prentice Hall.

Brockwell, P.J. and Davis R.A. 1993. *Time Series: Theory and Methods*. 2nd Edition. New York: Springer Verlag.

Chen, X., Racine J., and Swanson N. R. 2001. “Semiparametric ARX Neural-Network Models with an Application to Forecasting Inflation,” *IEEE Transaction on Neural Networks*, Vol. 12, p. 674-683.

Chan, M., Wong C., and Lam C. 1999. “Financial Time Series Forecasting by using Conjugate Gradient Learning Algorithm and Multiple Linier Regression Weight Initialization,” *Department of Computing, The Hongkong Polytechnic University*, Hongkong.

Diaz, Borrajo F.L., Riverola F.F., Usero A., and Corchado J.M. Negative. 2001. “Feedback Network for Financial Prediction,” *Artificial Intelligence Research Group. Universidad de Vigo*, Spain.

Firdaus, N., Shukor M., Azlan Roselina, and Nuradibah S. 2005. Back Propagation Neural Network (BPNN) Model as a Solution of Short-Term Rainfall Prediction for Johor Catchment Area, *IRCMSA Proceedings,* Medan, Indonesia.

Haykin, H. 1999, *Neural Networks: A Comprehensive Foundation*. Second Edition. Oxford: Prentice-Hall.

Moody, J. 1995. “Economic Forecasting Challenger and Neural Network Solutions,” *In Proceedings of the International Symposium on Artificial Neural Networks,* Taiwan.

Nikola, G. and Jing Yang. 2000. The Application of Artificial Neural Networks to Exchange Rate Forecasting: *The Role of Market Microstructure Variables*, Financial Markets Department Bank of Canada.

Ripley, B.D. 1993. *Statistical Aspects of Neural Networks*, in O.E. , J.L Barndorff- Nielsen. Jensen and W.S. Kendall, eds., Networks and Chaos: Statistical and Probabilistic Aspects, Chapman & Hall.

Ranaweera, D.K. and Hubele N.E. 1995. Application of radial Basis Function Neural Network Model for Short-Term Load Forecasting, *IEE Proc.-Gener, Transm. Distrib.*, Vol.142, No. 1.

Robert, R. and Jae LL.1996, Artificial Intelligence in Finance & Investing. *Ch. 0, IRWIN*.

Suhartono, Subanar and Sri Rejeki, 2005, “Feedforward Neural Networks Model for Forecasting Trend and Seasonal Time Series,” *IRCMSA Proceedings*, Sumatra Utara Indonesia.

Sutijo, Brodjol. 2005. “Radial Basis Function as Statistical Modeling for Financial Data,” *Proceedings International Conference on Applied Mathematics*, Bandung.

Tkacz, G. 2001. Neural Network Forecasting of Canadian GDP Growth,” *International Journal of Forecasting*, Vol 17, p. 57-69.

Wei, W.W.S. 1990. *Time Series Analysis: Univariate and Multivariate Methods.* USA: Addison-Wesley Publishing Co.

White, H. and Swanson N. R. 1997. “A Model-Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks,” *Review of Economic an Statistics*, 79, 540–550.