

## **Forecasting Foreign Direct Investment Inflow to Sri Lanka: Hybrid ARIMA-Neural Network model**

S.S.N. Perera, G.N. Karunathunge and G.T.B.G. Chathurdarshi  
*Department of Mathematics, University of Colombo, Sri Lanka.*

### **Introduction**

FDI is known as that “Investment made by multinational business enterprises in foreign countries to control assets and manage production activities in those countries” (Bashier and Talal, 2007). Foreign direct investment consists of three major components. Those are equity capital, reinvestment earnings and other capital used in various intra company debt transactions. In 1977, the Sri Lankan government made changes in its economic policies to establish an investment friendly economic environment. Sri Lanka reached its highest FDI inflow of US dollars 752.2 million in 2008.

The objective of the research is to build a model to forecast future FDI inflow to Sri Lanka.

Research has been carried out using different methods such as the univariate ARIMA model (Bashier and Talal, 2007), the double exponential smoothing model (Kumar *et al.*, 2009) and the neural network model (Pradhan, 2010) to forecast FDI inflow to other countries. To the best knowledge of the authors, this study could be the first to forecast FDI inflows in Sri Lanka using a combination of an ARIMA model and a neural network model. In this study, an ARIMA model is fitted through the Box-Jenkins procedure and then the back propagation neural network procedure is applied to remove the lack of accuracy due to the small sample size. The data set employed consists of annual foreign direct investment inflow to Sri Lanka from 1978 to 2010.

### **Methodology**

The ARIMA model is a combination of two processes. These are the Auto Regressive (AR) process and the Moving Average (MA) process. The Box Jenkins methodology is one of the most widely used procedures for time series forecasting using the ARIMA model. There are three primary stages in the Box Jenkins procedure. These are model identification, model estimation and model validation (Dobre and Alexandru, 2008). In addition, data preparation and forecasting are the initial and final stages.

Bashier and Talal (2007) stated that to employ the Box-Jenkins process to forecast a time series, the stationary of the series must be maintained. Transforming and differentiating can be used to convert a non stationary series into a stationary one. Model identification or determining the number of AR and MA terms can be done by looking at the plots of the Auto-Correlation Function (ACF) and the Partial Auto-Correlation Function (PACF) (Bashier and Talal, 2007). The Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) can be used to select the most suitable model. Parameter estimates are usually obtained by the least square method and the maximum likelihood method (Dobre and Alexandru, 2008). There are various diagnostic tests to check the adequacy of the model. Some of them are residual correlogram, Ljung-box chi squared statistics, normal probability plot and plot of standardized residuals vs. fitted values.

An artificial neural network is composed of many artificial neurons that are linked together according to specific network architecture to transform the inputs into a meaningful output. Feed forward neural network consists of an input layer, an output

layer and one or more hidden layers (Tseng *et al.*, 2002). Back propagation is a method which iteratively minimizes the error of the output.

The ARIMA model requires a large amount of historical data to make a good forecast. According to a previous research (Tseng *et al.*, 2002) back propagation neural network can be used to improve the forecast accuracy. The fitted value of the ARIMA model and the residuals of ARIMA model are the inputs of the neural network model. Forecasted values of the ARIMA model and the average of past residuals can be used as the inputs when the neural network model is used to forecast future values.

There are different measurements (Bashier and Talal, 2007) to compare the forecast accuracy of a set of models. MSE, MAE, MAPE and Theil's U Statistic are used in this study to compare the forecast accuracy between the ARIMA model and the hybrid model.

### Results and discussion

FDI inflow over the period 1978-2002 is used for the model fitting procedure. As the initial step, the logarithm transformation is applied to the data set to reduce its large variance. The transformed series is stated as the log (FDI) series in the rest of this paper.

Stationary test results: According to the time series plot, ACF plot, PACF plot and unit root test the log (FDI) series is not stationary. The first difference log (FDI) series is considered and tested. According to the time series plot, the ACF plot, the PACF plot, the Augmented Dickey-Fuller (ADF) and the Phillips-Perron (PP) tests log (FDI) series is stationary in it first difference form.

Model Identification: PACF is considered to identify the order of the AR part and the ACF to identify the order of the MA part. Then ARIMA (2, 1, 2) is suggested as the initial model. Then a set of tentative models is considered and based on AIC and BIC values, seven models out of eighteen are selected.

Model Estimation: Parameters are estimated for the seven selected models. There is only one model with significant parameters at 5% level. That is the ARIMA (2, 1, 1) model (Table-1).

Table 1: Parameter estimation for ARIMA (2, 1, 1)

Variable	Coefficient	t value	Probability
AR 1	0.4547	2.67	0.015
AR 2	-0.3672	-2.31	0.031
MA 1	1.0923	11.06	0.000
Constant	0.08549	12.6	0.000

Model Validation: Residuals are white noise or random according to the residual correlogram and Ljung Box chi squared statistic.

Also, residuals are normally distributed. There is an approximately linear pattern in the normal probability plot of residuals. Except one all the other standardized residuals lie between +2 and -2. It implies there are no outliers. Also, there is no systematic pattern in the plot of standardized residual vs. fitted values. According to the diagnostic check, ARIMA (2,1,1) is an adequate model.

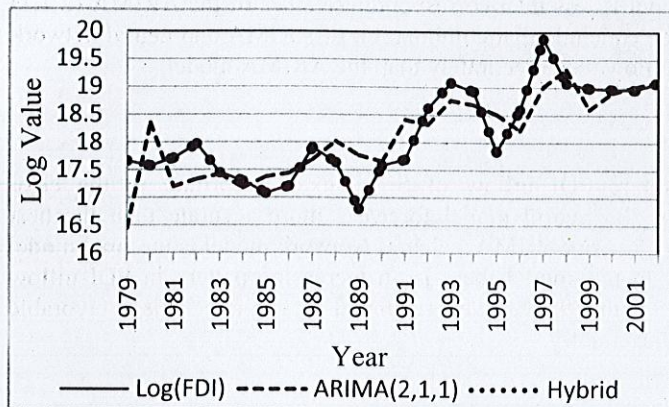


Figure-1: Actual VS. Fitted and Forecasted

Forecasting: The ARIMA (2,1,1) model is used to forecast FDI inflow over the period 2011-2015 .

Only 25 annual data is used to build the ARIMA model. However small sample size is not sufficient to build an ARIMA model with a high forecasting accuracy. Therefore, back propagation neural network approach is applied.

Table 2: Forecasts of FDI-Hybrid model

Year	FDI (\$)
2011	504460450
2012	550864819
2013	599136494
2014	647740039
2015	696097295

Results of the hybrid ARIMA (2,1,1)-NN approach: Fitted values and residuals of the ARIMA (2,1,1) model are given as the inputs. The output is actual log (FDI) values.

The Net is created by using logsig, tansig and purelin transfer functions. The corresponding numbers of nodes are 2, 21 and 1. Fitted values of the Hybrid model are more accurate than that of the ARIMA model. Because these are closer to the actual Log (FDI). Forecasting is done by using the forecasted values of the ARIMA model over the period 2003-2015 and the average of the historical residuals (Table-2)

Table 3: Measurements for Accuracy Comparison

Error Measurement	ARIMA(2,1,1)	Neural Network
MSE	0.14008	0.12512
MAE	0.25070	0.24697
MAPE	0.01243	0.01091
Theil's	0.00950	0.00897

Test data (2003-2010) is used to find the forecast error. MSE, MAE, MAPE and Theil's U Statistic of the hybrid model is less than corresponding values of the ARIMA (2,1,1) (Table-3). Therefore it can be concluded that combination of ARIMA and neural network models forecast future FDI inflow more accurately than the ARIMA model.

### **Conclusions**

This study forecasts the future FDI inflow of Sri Lanka. According to the error measurement of the models, the hybrid model forecasts more accurately than the best ARIMA model. Therefore the hybrid ARIMA - Neural Network model is the final model used to forecast. According to this model there is an increasing pattern in FDI inflow over the period 2011 to 2015 with a 9.18% average annual growth rate. This is favorable for the economic growth of the country.

### **References**

- Al-Abdulrazag Bashier and Bataineh Talal 2007. Forecasting Foreign Direct Investment Inflow in Jordan: Univariate ARIMA Model, *Journal of Social Sciences*. 3: 1-6.
- Fang-Mei Tseng,, Hsiao-Cheng Yu and Gwo-Hsiung Tzeng 2002. Combining neural network model with seasonal time series ARIMA model, *Technological Forecasting & Social Change*. 69:71-87.
- Gulshan Kumar, Neerja Dhingra and Avtar Singh 2009. Growth and Forecasts of FDI Inflows to Latin America and the Caribbean, *European Journal of Social Sciences*. 12: 216-234.
- Ion Dobre and Adriana Ana Maria Alexandru 2008. Modeling Unemployment Rate using Box-Jenkins procedure, *Journal of applied quantitative methods*.3: 156-166.
- Rudra Prakash Pradhan 2010. Forecasting FDI in the Asian economy, *The IUP Journal of computational mathematics*. 3: 39-53.