Barisal University Journal Part 1, 4(1):81-94 (2017) ISSN 2411-247X

FORECASTING AND DETERMINING THE ORDER OF SELECTED PUBLISHED FINANCIAL DATA OF INTERNATIONAL MARKET

Md. Saif Uddin Rashed[1*] and Md. Mafizul Islam[2]

[1]*Department of Finance and Banking, University of Barisal, Barisal 8200, Bangladesh*
[2]*Department of Centralized Trade Service, Dutch-Bangla Bank Limited, Dhaka 1000, Bangladesh*

Abstract

This research aims at determining the appropriate order of the autoregressive (AR) and the vector autoregressive (VAR) models, estimating and then forecasting the prices of some selected commodities (rice, wheat, palm oil and soya oil) at international market for some periods ahead. Finally, we compare the two forecasted series obtained by the AR and the VAR models in terms of the forecast error. Initially all variables used in this study are tested for stationary using unit roots with the application of Dickey Fuller tests , Augmented Dickey Fuller tests and Phillips-Perron (PP) tests. Moreover, Granger causality tests are conducted to establish patterns of causality (or precedence) between the prices of the selected commodities and to determine which variables are to be included in the vector auto regression system. The Akaike information criterion (AIC), the Schwarz-Bayesian information criterion (SBIC), the final prediction error (FPE) criterion are used to determine the lag length of the AR system. Similarly the multivariate generalisations of the Akaike information criterion (AIC) and the Schwarz-Bayesian information criterion SBIC), final prediction error (FPE) and Hannan-Quinn (HQ) criterion are used to determine the lag length of the VAR system.

Keywords: Autoregressive (AR) model, Vector autoregressive (VAR) model

Introduction

In making choices between alternative courses of action, decision makers' at all structural levels often need prediction of economic variables. If time series observation are available for a variable of interest and the data from the past contain information about the future development of a variable, it is plausible to use as forecast some function

---

[*]*Corresponding author's e-mail: md.saifuddinrashed@gmail.com*

of the data collected in the past. For instance in forecasting the monthly unemployment rate, from past experience a forecaster may know that in some country or area a high unemployment rate in one month tends to be followed by a high rate in the next month.

Formally this may be expressed as follows. Let y_t denote the value of the variable of interest in period t. Then a forecast for period $T+h$, made at the end of period T, may have the form

$$\hat{y}_{T+h} = f(y_T, y_{T-1}, \ldots)$$

Where $f(.)$ denotes some suitable function of the past observations $y_T, y_{T-1}, \ldots$. For the moment it is left open how many past observations enter in to the forecast. One major goal of univariate time series analysis is to specify sensible forms of the function $f(.)$. In many applications linear functions have been used so that, for example,

$$y_{T+h} = v + \alpha_1 y_T + \alpha_2 y_{T-1} + \ldots$$

In dealing with the economic variables often the value of one variable is not only related to its predecessors in time but, in addition, it depends on past values of other variables. For instance, household consumption expenditures may depend on variables such as income, interest rates, and investment expenditures. If all these variables are related to the consumption expenditures it makes sense to use their possible additional information content in forecasting consumption expenditures. In other words, denoting the related variables by $y_{1t}, y_{2t}, \ldots, y_{kt}$, the forecast of $y_{1,T+h}$ at the end of period T may be of the form

$$\hat{y}_{1,T+h} = f(y_{1,T}, y_{2,T}, \ldots, y_{k,T}, y_{1,T-1}, y_{2,T-1}, \ldots, y_{1,T-2,\ldots})$$

Similarly, a forecast for the second variable may be based on past values of all variables in the system. More generally, a forecast of the k-th variable may be expressed as

$$\hat{y}_{k,T+h} = f_k(y_{1,T}, y_{2,T}, \ldots, y_{k,T}, y_{1,T-1}, y_{2,T-1}, \ldots, y_{1,T-2,\ldots})$$

A set of time series y_{kt}, $k = 1, \ldots, K$ and $t = 1, \ldots, T$, is called a multiple time series and the previous formula expresses the forecast $\hat{y}_{k,T+h}$ as function of multiple time series.

It is also often of interest to learn about the dynamic interrelationships between a number of variables. For instance, in a system consisting of investment, income and consumption one may want to know about the likely impact of an impulse in income. What will be the present and future implications of such an event for consumption and for investment? Under what conditions can the effect of such an impulse be isolated and trace through the system? Alternatively, given a particular subject matter theory, is it consistent with the relations implied by a multiple time series model which is developed with the help of statistical tools? These and other questions regarding the structure of the relationships between the variables involved are occasionally investigated in the context of multiple time series analysis. Thus, obtaining insight in to the dynamic structure of a system is a further objective of multiple time series analysis. One class of multiple time series models which has received much attention recently is the class of Vector Autoregressive (VAR) models. VAR models constitute a special case of the more general class of Vector Autoregressive Moving Average (VARMA) models. Although VAR models have been used primarily for macroeconomic models, they offer an interesting alternative to either structural econometric (market share) or univariate (e.g., Box-Jenkins ARIMA or exponential smoothing) models for problems in which simultaneous forecasts are required for a collection of related microeconomic variables, such as industry and firm sales forecasting. The use of VAR models for economic forecasting was proposed by Sims (1980), motivated in part by questions related to the validity of the way in which economic theory is used to provide *a priori* justification for the inclusion of a restricted subset of variables in the "structural" specification of each dependent variable 1.

This study adopts two methodologies to forecast the prices of some selected commodities (Rice, Wheat, at international market for some period ahead. The two methodologies are: the autoregressive process (AR) and the vector autoregression (VAR) system. Initially all variables used in this study are tested for stationarity (so that the mean, variance and autocovariances are independent of time), that is, the variables are tested for unit roots. The main statistical reason for this is that stationary series are required by the autoregressive model and for valid application of and inference under the least squares method in vector autoregressions.

Subsequently, Granger causality tests (Granger, 1969) are applied to examine whether the relationships between the prices of the selected commodities and the variables conform to theoretical intuition and to determine which variables are to be included in the vector auto regression system. If a variable does not have a significant impact on the prices of the selected commodities under this test, it is excluded from the system.

In the estimation of a AR model, statistical tests are used to decide upon the appropriate number of lags for each equation. The Akaike information criterion (AIC), the Schwarz-Bayesian information criterion (SBIC) and the final prediction error criterion (FPE) are used to determine the lag length of the AR system. Similarly, in the estimation of a VAR model, statistical tests are used to decide upon the appropriate number of lags for each equation.

In this study, the multivariate generalisations of the Akaike information criterion (AIC) and the Schwarz-Bayesian information criterion (SBIC), final prediction error (FPE) and Hannan-Quinn (HQ) criterion are used to determine the lag length of the VAR system. After selecting the appropriate AR and the VAR models, we forecast the prices of the selected commodities at international market by the selected AR models and the VAR models. Finally, we compare the two forecasted series of values obtained by the AR and the VAR models in terms of the forecast error.

Methodology

Four series of the prices of some selected commodities (rice, wheat, palm oil and soya oil) at international market are employed in this study. This data set consists of monthly average prices of 126 months on rice, wheat, palm oil and soya oil from January, 1997 to September, 2007. We have used only 123 observations for each of the selected commodity prices and 3 observations are retained for comparison purpose with the forecasted values in terms of the forecast error. The variables are in this study are: the price of rice(R), (Thailand, Bangkok), wheat(W), (United States, Gulf Ports), palm oil(PO), (Malaysia) and soya oil(SO), (All Origins) in US Dollar per metric ton. The Statistics Department of Bangladesh Bank is the source of the data. For the computation and analysis for the data we use SPSS version-16, Eviews, software. Subsequently the selected models will be estimated by using some computer packages (JMulTi and Gauss) for each of the variables (rice, wheat, palm oil and soya oil)

Result

Test of Stationary

Graphical Analysis

The ACF for the price of rice at international market, shown in the **Fig. 1** suggests that the price of rice series is non stationary because the ACF decline very slowly as increasing in lag and most of the spikes are outside the 95% confidence band. And it is also observed that after the first lag the partial ACF, shown in the **Fig. 2.** drops very quickly.

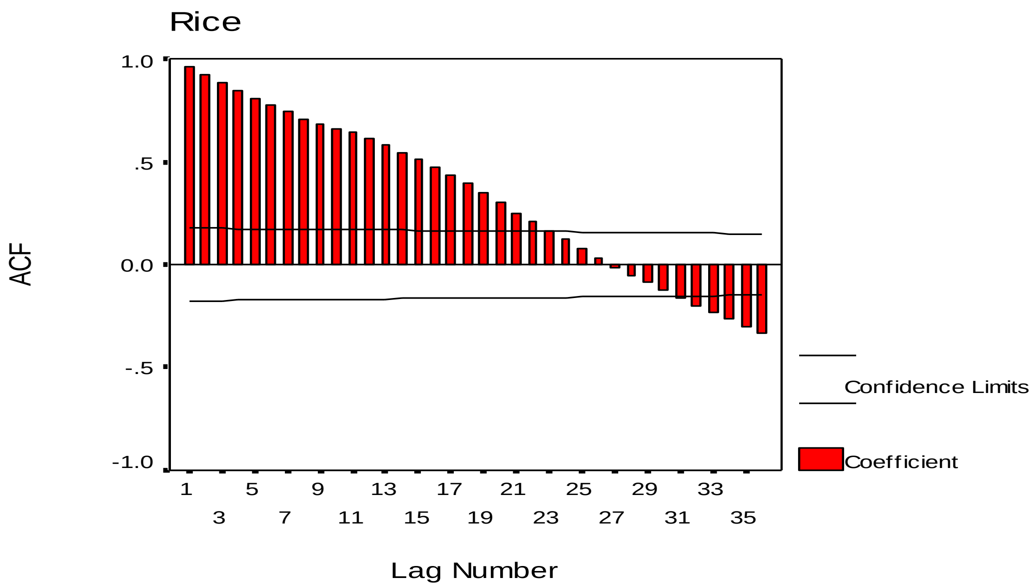

Fig. 1. The autocorrelation function for the price of rice at international market of Thailand (Bangkok), January, 1997 to July,

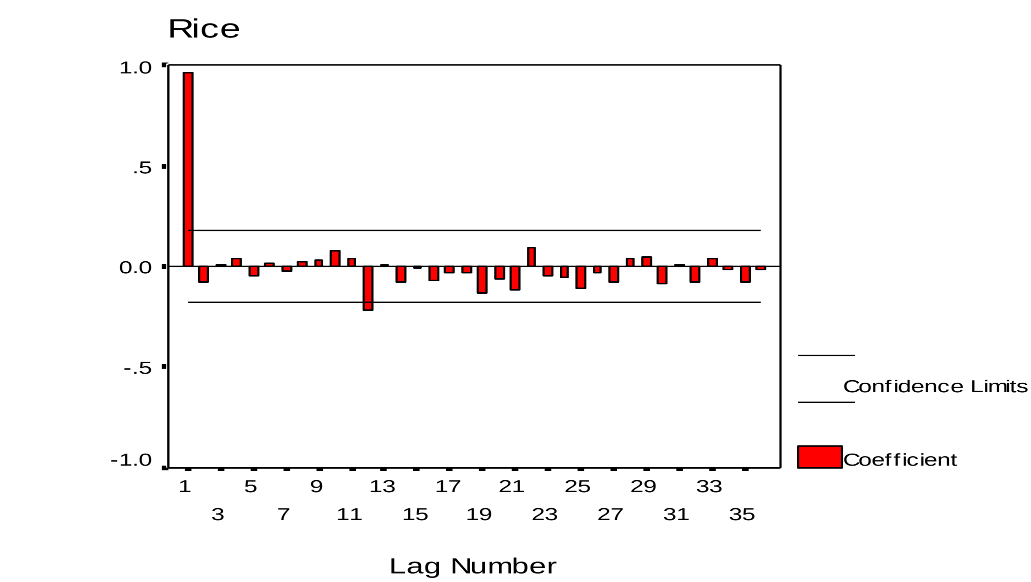

Fig. 2. The partial autocorrelation function for the price of rice at international market of Thailand (Bangkok), January, 1997 to July, 2007.

Unit root test

It is observed from the **Table 1.** that the hypothesis of a unit root in all variables are not rejected when Dickey Fuller (DF) tests, augmented Dickey Fuller (ADF) tests and the Phillips-Perron (PP) test are applied to the level of all variables. Subsequently, DF, ADF tests and PP test are carried out on the first differences of all series. The results are reported in **Table 2**. All differenced series appear to be stationary not only at the 5% level of significance but also 1% level of significance. Therefore, all variables are included in the Granger causality tests.

Table 1. Tests for stationary of the original series.

Variable	Computed DF statistic	Computed ADF statistic	Computed PP statistic
Rice	-1.02	-1.12	-1.08
Wheat	-1.34	-1.03	-1.22
Palm Oil	-0.42	-0.82	-0.99
Soya-Oil	-0.44	-0.61	-0.74

Table 2. Tests for stationary of the first difference of the series.

Variable	Computed DF statistic	Computed ADF statistic	Computed PP statistic
Δ Rice	-9.89	-6.87	-9.88
Δ Wheat	-11.36	-6.73	-11.37
Δ Palm Oil	-9.94	-5.61	-9.90
Δ Soya Oil	-8.43	-5.68	-8.34

Granger causality

The Granger (1969) approach to the question of whether x causes y is to see how much of the current y can be explained by past values of y and then to see whether adding lagged values of x can improve the explanation. The y is said to be Granger-caused by x if x helps in the prediction of y, or equivalently if the coefficients on the lagged x's are statistically significant. It is important to note that the statement “x Granger causes y” does not imply that y is the effect or the result of x. Granger causality measures precedence and information content but does not by itself indicate causality in the more common use of the term.

From the Granger Causality test, presented in **Table 3** it is evident that rice Granger causes soya oil. palm oil Granger Causes rice but palm oil and soya oil does not Granger causes wheat in two way direction. Granger causality test indicates that wheat does not include in the VAR system. So our included variables are rice, palm oil and soya oil for the VAR system. We also consider wheat, palm oil and soya oil for the VAR system in order to compare.

Table 3. Granger causality tests.

Null Hypothesis	F Value	p value	Judgment
Rice does not Granger Cause Wheat	2.54713	0.04326	Rejected
Wheat does not Granger Cause Rice	0.98626	0.41812	Accepted
Rice does not Granger Cause Palm Oil	0.72291	0.57806	Accepted
Palm Oil does not Granger Cause Rice	3.64721	0.00784	Rejected
Rice do not Granger Cause Soya Oil	2.13891	0.08062	Rejected
Soya Oil does not Granger Cause Rice	0.50920	0.72906	Accepted
Wheat does not Granger Cause Palm Oil	0.75080	0.55952	Accepted
Palm Oil does not Granger Cause Wheat	1.45984	0.21921	Accepted
Soya Oil does not Granger Cause Palm Oil	4.46415	0.00220	Rejected
Palm Oil does not Granger Cause Soya Oil	1.72251	0.14996	Accepted
Wheat does not Granger Cause Soya Oil	1.38248	0.24456	Accepted
Soya Oil does not Granger Cause Wheat	0.85705	0.49220	Accepted

Forecasting the AR and VAR models

Forecasting the price variables with AR models

The autoregressive model of order one, that is, AR(1), for the Rice series and it is also evident from the values of Akaike information criterion (AIC), the Schwarz-

Bayesian information criterion (SBIC), and the final prediction error (FPE)(clear minimum at $p=1$), shown in **Table 4.** Thus, our selected model is AR(1) of the following form:

$$y_t = v + \alpha_1 y_{t-1} + u_t \ ,$$

where $\{u_t\} \sim WN\left(0,\sigma^2\right)$, $|\alpha_1| < 1$ and u_t is uncorrelated with y_S for each $s < t$.

Table 4. Order selection of the AR model for the price of rice.

	AR Order P	AIC	SBC	FPE
Rice	1	970.199*	976.055*	137.74*
	2	971.981	980.466	140.63
	3	972.991	984.274	142.88
	4	974.876	989.018	146.29

*Minimum

For the price of rice the model becomes

$$R_t = v + \alpha_1 R_{t-1} + u_t \ .$$

Using the computer package JMulTi and Gauss, the estimated AR(1) model for the price of rice is as follows:

$$\hat{R}_t = 0.1997 + 0.1081 R_{t-1}$$

Table 5. The order selection and estimated autoregressive model for wheat, palm oil, palm oil, soya oil.

Product name	Order selection for AR model			Estimated model
	AIC	SBC	FPE	
Wheat	AR(1)	AR(1)	AR(1)	$\hat{W}_t = 0.374 - 0.055 W_{t-1}$
Palm Oil	AR(4)	AR(1)	AR(1)	$PO_t = 1.468 - 0.105 PO_{t-1}$
Soya Oil	AR(4)	AR(2)	AR(2)	$SO_t = 1.903 + 0.315 SO_{t-1} - 0.221 SO_{t-2}$

In the **Table 5.** shown that**,** If we consider for the wheat series the values of AIC , SBC and FPE which have minimum value at $p = 1$, so the wheat series select the autoregressive model of order one, that is, AR(1). Moreover, for the palm Oil series the minimum values of SBC and FPE shows at p =1, so the autoregressive model is order one, that is, AR(1). But the values of AIC have a clear minimum at $p = 4$ suggesting AR(4). Finally, for the soya oil series the minimum values of SBC and FPE shows at p =2, so the autoregressive model is order two, that is, AR(2). But the values of AIC have a clear minimum at $p = 4$ suggesting AR(4).

Forecasting the price variables with VAR models

Order selection of the VAR model of the palm oil/soya oil/rice

If forecasting is the objective it makes sense to choose the order such that a measure of forecast precision is minimized. The AIC, FPE, SC and HQ values are presented in **Table 6.** for the palm oil/soya oil/rice system. All these criteria reach their minimum for $\hat{p} = 1$, that is, $\hat{p}(FPE) = \hat{p}(AIC) = \hat{p}(HQ) = \hat{p}(SC) = 1$, suggesting the vector autoregressive model of order one, that is, VAR(1).

Table 6. Estimation of the VAR order of the palm oil/soya oil/rice

VAR order m	AIC(m)	FPE(m) $\times 10^{-6}$	SC(m)	HQ(m)
1	18.467*	104.838*	18.741*	18.579*
2	18.501	108.423	18.981	18.696
3	18.503	108.734	19.193	18.783
4	18.501	108.682	19.403	18.867
5	18.546	113.894	19.961	18.999

*Minimum.

Using the likelihood ratio statistics, presented in **Table 7.** at individual significance levels of .05 in each test, $H_0^1 : A_4 = 0$ is the first null hypothesis that is rejected. Assuming that $M = 4$ is the upper bound for the VAR order. Thus, the estimated order from both tests (F-test and χ^2-test) is $\hat{p} = (M - i + 1) = (4 - 1 + 1) = 4$, that is VAR(4). This does not support the order chosen by the selection criteria.

Table 7. Likelihood Ratio (LR) statistics for the price of some selected commodities (Palm Oil, Soya Oil and Rice) at international market.

i	H_0^i	VAR order under H_0^i	λ_{LR^a}	$\lambda_{LR}/9^b$
1	$A_4 = 0$	3	19.34	2.15
2	$A_3 = 0$	2	18.38	2.04
3	$A_2 = 0$	1	14.09	1.57
4	$A_1 = 0$	0	17.83	1.98

[a]Critical value for individual 5% level test: $\chi^2(9)_{.95} = 16.92$.

[b]Critical value for individual 5% level test: $F(9,125-3(6-i)-1)_{.95} \approx 2.02$

Thus, however, for forecasting purpose we select the VAR (1) models and our model is of the following form:

$$y_t = v + A_1 y_{t-1} + u_t.$$

Where, $y_t = (PO, SO, R)', v = (v_1, v_2, v_3)'$ and

$$A_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$ PO = Palm oil, SO = Soya oil, R = Rice

For the palm oil/soya oil/rice system our selected models becomes:

$$\begin{pmatrix} PO \\ SO \\ R \end{pmatrix}_t = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{pmatrix} PO \\ SO \\ R \end{pmatrix}_{t-1} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}_t .$$

Thus, the estimated VAR(1) model for the palm oil/soya oil/rice system is given by:

$$y_t = \hat{v} + \hat{A}_1 y_{t-1}$$

or

$$\begin{pmatrix} PO \\ SO \\ R \end{pmatrix}_t = \begin{pmatrix} 0.65 \\ 1.59 \\ 0.12 \end{pmatrix} + \begin{bmatrix} -0.12 & 0.47 & 0.17 \\ -0.01 & 0.28 & 0.08 \\ 0.05 & -0.01 & 0.09 \end{bmatrix} \begin{pmatrix} PO \\ SO \\ R \end{pmatrix}_{t-1}$$

For this process the modulus of the eigen-values of the reverse characteristic polynomial is $|z| = (8.01,\ 11.57,\ 3.65)$. Thus, the process satisfies the stability condition since all roots are greater than one.

Therefore, the forecast values for the palm oil/soya oil/rice system for three periods ahead with the forecast error are given by the **Table 8.**

Table 8. The forecast values for the palm oil/soya oil/rice system

Model	Prices	Forecast			Forecast Error		
		1-step	2-step	3-step	1-step	2-step	3-step
VAR(1)	Palm Oil	761.72	765.18	767.40	2.78	-35.58	-22.20
	Soya Oil	792.00	795.94	798.59	27.20	1.16	54.11
	Rice	326.98	327.67	327.96	5.62	3.83	2.04

Table 9. The forecast values for the palm oil/soya oil/wheat system

Model	Prices	Forecast			Forecast Error		
		1-step	2-step	3-step	1-step	2-step	3-step
VAR(1)	Palm Oil	753.60	757.64	759.45	10.90	-28.04	-14.25
	Soya Oil	791.07	794.81	797.36	28.13	2.29	55.34
	Wheat	223.56	224.29	224.80	14.84	35.41	11.70

Using the same procedure we get the forecast value for the palm oil/soya oil/ wheat for VAR (1) model which accepted and shown in **Table 9.**

Comparison of the AR and VAR system

Table 10. The forecast values of auto regressive (AR) and vector autoregressive (VAR) system

Model	Prices	Forecast			Forecast Error		
		1-step	2-step	3-step	1-step	2-step	3-step
Reject Granger Test VAR(1)	Palm Oil	761.72	765.18	767.40	2.78	-25.58	-22.20
	Soya Oil	792.00	795.94	798.59	27.20	1.16	54.11
	Rice	326.98	327.67	327.96	5.62	3.83	2.04
Accept Granger Test VAR(1)	Palm Oil	753.60	757.64	759.45	10.90	-28.04	-14.25
	Soya Oil	791.07	794.81	797.36	28.13	2.29	55.34
	Wheat	223.56	224.29	224.80	14.84	35.41	11.70
AR(1)	Rice	327.12	327.40	327.63	5.48	4.10	2.37
AR(1)	Wheat	222.44	222.88	223.23	15.96	36.82	13.27
AR(1)	Palm Oil	750.69	752.39	754.04	13.81	-22.79	-8.84
AR(4)	Palm Oil	750.13	775.15	798.01	14.38	-45.55	-52.81
AR(2)	Soya Oil	783.98	780.30	780.80	35.92	16.80	71.90

It is clear from the forecast values of the VAR models, shown in **Table 10.** That all the forecast errors of the soya oil in palm oil/soya oil/rice system are smaller than that of the soya oil in the palm oil/soya oil/wheat system.

It is also clear from the **Table 10.** That almost all of the forecast errors from both the VAR system (pass/do not pass Granger test) are smaller than that of the AR system except the palm oil. If we consider the AR(4) model for the palm oil, it produces the more forecast error. This Indicate that AR models always are not the forecasting tools if forecasting is our main concern.

Conclusion

In our research we test the time series data of different commodities for the purpose of stationary assumption by unit root test (Dickey Fuller (DF) tests, augmented Dickey Fuller (ADF) tests and the Phillips-Perron (PP) test). The outcomes are reported that all differenced series appear to be stationary not only at the 5% level of significance but also 1% level of significance. So, all variables are included in the Granger causality test which indicates that one product price (wheat) does not include in the VAR system. From the comparison of forecast error for the VAR models and AR model, it is clear that the VAR model have smaller values than AR model. We carry out that AR models always are not the forecasting tools if forecasting is our main goal because sometimes it produce more forecast error. We expect using VAR procure we get the better forecast for time series analysis. Ignoring this leads to different solution. So we recommend to forecast we need to apply not only the autoregressive model but also use the vector autoregressive model.

References

Box, G. E. P. and G. M. Jenkins. 1978. *Time Series Analysis, Forecasting and Control.* Revised Edition, Holden-Day San Francisco, CA.

Brockwell, P. J and R. A. Davis. 2002. *An Introduction to Time Series and Forecasting* 2nd ed. Springer-Verlag.

Chen, A. and T. L. Mark. 2003. A Bayesian vector error correction model for forecasting exchange rates. *Computers & Operations Research*. **30**(**6**): 886-888.

Dickey, D. A. and W. A. Fuller. 1979. Distribution of the estimator for autoregressive time series with a unit root. *Journal of American Statistical Association.* **74**:34-37

Gujarati, D. N. 2002. *Basic Econometrics*. 4th ed. McGraw-Hill; New York.

Johnston, J. 1997. *Econometric Methods.* 4th ed. McGraw-Hill; New York.

Ljung, G. M. and G. E. P. Box. 1978. On the measure of lack of fit in the time Series models. *Biometrica.* **65**:297-299.

Neely, J. C. and L. Sarno. 2003. How well do monetary fundamentals forecast exchange rates. *Federal Reserve Bank of Saint Louis Review.* **84(5):**20-22.