International Journal of Economics and Management Sciences
Vol. 1, No. 9, 2012, pp. 52-58

MANAGEMENT JOURNALS
managementjournals.org

FORECASTING GDP GROWTH RATES OF INDIA: An Empirical Study

Bipasha Maity, *PhD*[*1]; Bani Chatterjee, *PhD*[2]

*[*1] Corresponding author: Assistant Professor of Economics, Department of MBA*
Future Institute of Engineering and Management, Kolkata-700150, India
E-mail: bipasha812@yahoo.com
[2] Professor of Economics, Department of HSS
Indian Institute of Technology, Kharagpur -721302, India

ABSTRACT

This study has been attempted to shed light on the issues such as forecasting growth rates of GDP of India. Data on GDP have been collected over a period of 60 years from various publications of Reserve Bank of India. A very simple tentative ARIMA (1, 2, 2) model has been fitted on data to estimate the parameters of autoregressive and moving average components of this model. Results suggest that only one period of autoregressive and moving average terms are statistically significant. Further, absolute values of forecasted GDP indicate an increasing trend and its respective growth rates reveal an opposite trend in future. Findings will assist the policy makers and managers to formulate economic and business strategies in turn more precisely.

Keywords: *GDP, ARIMA, Forecasts, Growth Rates, Trend*

1. INTRODUCTION

GDP indicates the financial health of a country as a whole-which is actually a hunting ground of researchers in the field of business in general and of economics in particular. The issues of GDP has become the most concerned amongst macro economy variables and data on GDP is regarded as the important index for assessing the national economic development and for judging the operating status of macro economy as a whole (Ning et al. 2010)

GDP is the aggregate statistic of all economic activity and captures the broadest coverage of the economy than other macro economic variables. It is the market value of all final goods and services produced within the borders of a nation in a year. It is often considered the best measure of how well the economy is performing. GDP can be measured in three ways. First, the Expenditure approach, it consists of household, business and government purchases of goods and services and net exports. Second the Production approach, it is equal to the sum of the value added at every stage of production (the intermediate stages) by all industries within the country, plus taxes and fewer subsidies on products in the period. Third is Income approach, it is equal to the sum of all factor income generated by production in the country (the sum of remuneration of employees, capital income, and gross operating surplus of enterprises i.e. profit, taxes on production and imports less subsidies) in a period (Yang 2009; Ard 2010)

Besides these, it is also the vital basis for government to set up economic developmental strategies and policies. Therefore, an accurate prediction of GDP is necessary to get an insightful idea of future health of an economy since the data on GDP is actually represented the past activities in summary form, which is not very helpful to frame suitable economic development strategies, economic policies and allocation of funds on different priorities for government as well as individual firms in a particular industry. It needs a reliable estimate of GDP in some period ahead, which is only possible by forecasting GDP as accurately as possible by suitable sophisticated time series modeling, since it is not easy to identify the variables those effects on GDP precisely.

In a study, Tsay and Tiao (1984, 1985) used ARIMA model, which is in fact fitted on non-seasonal data by identifying autoregressive and moving average terms with the help of partial autocorrelation and autocorrelation functions (Box and Jenkins 1970:1976,Pankratz 1991).However, in the case of seasonal data, a number of studies used filtering approach, which in fact very helpful in case of weekly, monthly, quarterly and semiannual data to estimate a model to forecast any macro variable (Liu 1989; Liu and Hudak 1992; Liu 1999).
In another research, Reynolds et al. (1995) developed automatic methods to identify as well as estimate the parameters of ARIMA model by utilizing time-series data for a single variable. In another study, Reilly (1980) used similar methodology to model macroeconomic variable like GDP. However, both the studies confined themselves only on non-seasonal time series data and restrained to predict the variable in future. However, the above mentioned methods need a long time-series data on the macroeconomic variable in question. To estimate the model for prediction of a macro variable, a number of studies imply analytical neural network techniques, which is very effective in the case of seasonal data (Chiu et al. 1995; Cook and Chiu 1997; Geo et al. 1997; Saad et al. 1998). These types of models have got pace since the seminal paper of Granger and Joyeux (1980) and Hosking (1981). However, this neural networking approach is very difficult to applying in real life situation by the policy makers /managers due to difficult network design, training and testing are required to build the model as well as to estimate the parameters.

In summary, all the studies reported in the western world, which in turn motivated the researcher to carry out this research which deals with the GDP issues of India. Further, it is not clear why there is not many studies attempted to forecast the GDP as well as predicts the growth rates in various forms in India. Al though some studies attempted to forecast this macro variable only as point estimates which has very little help for the policy makers/ managers since variability is the key in decision making when a certain level of risk is involved.

In the above backdrop, this study is a modest attempt to fill the gap by identifying the following two research questions with respect to GDP forecasting issues in India

1. What are the year-wise forecasted GDP values in various forms over a period 2012-2021?
2. What are the year-wise GDP growth rates in different forms over a period 2012-2021?

2. MODEL

To uncover the hidden structure as well as to forecast GDP, the following equation has been taken into consideration in this work:

$$Y_t = \alpha + \sum_{p=1}^{n} \emptyset_p Y_{t-p} + \mathcal{E}_t + \sum_{q=1}^{n} \theta_q \mathcal{E}_{t-q} \qquad (1)$$

where Y_t = GDP in billion of Rupees at period t,

Y_{t-p} = GDP in billions of Rupees at period (t-p),

θ_{t-q} = Random shock at period t- q,

$\mathcal{E}_t$ = Random error term at period t and

(α, $\emptyset_p$ *and* θ_q) are parameters to be estimated.

The above equation is nothing but an ARIMA model, which is developed by Box and Jenkins (1976). Since then, this model has gained enormous popularity due to its versatility in many areas in business and in economics in particular. Further, several research studies confirm its power and flexibility to capture variability in any single variable collected over a period of time (Box et al. 1994; Enders 2004; Hoff 1983; Hossain et al. 2006; McDowall et al. 1980; Melard 1984; O' Donovan 1983; Pankratz 1983; Vandaele 1983). The following hypotheses have been formulated with respect to equation (1) which is in fact, the precondition to estimate the most parsimonious model to forecast GDP and its growth rates in some periods ahead.

$H_{1.1}$: GDP in (t- p) period would have an impact on GDP in current period. The mathematical hypothesis would be as follows:

$$H_0: \emptyset_p = 0 \ \ against \ \ H_a : \emptyset_p \neq 0$$

$H_{1.2}$: Random shock in (t- q) period would have an effect on GDP in period t. The mathematical hypothesis would be as

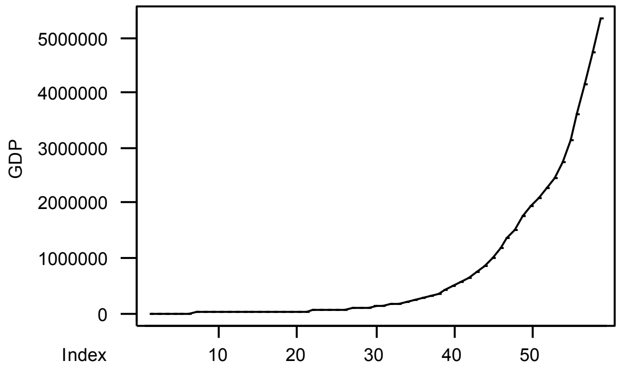
$$H_0: \theta_q = 0 \ against \ H_a: \theta_q \neq 0$$

3. DATA and METHODS of ANALYSIS

To forecast GDP, data on this macro-economic variable have been collected over a period 1959-2011 from the publication of Reserve Bank of India (2010).The data file consists of 59 observations of GDP (in billion rupees) in constant cost terms.

A graphical representation of data reveals that GDP series follows an exponential pattern over the period 1959-2010 (Fig 1).

Fig: 1 Exponential curve of GDP of India over a period 1959-2011



That is, GDP series does not follow stationary pattern. Hence, data have been differenced twice keeping in view of theory to convert them from non-stationary to stationary. Further, an ADF test has been conducted on this differenced GDP series of 57 observations.

Table 1 Results of ADF test on GDP

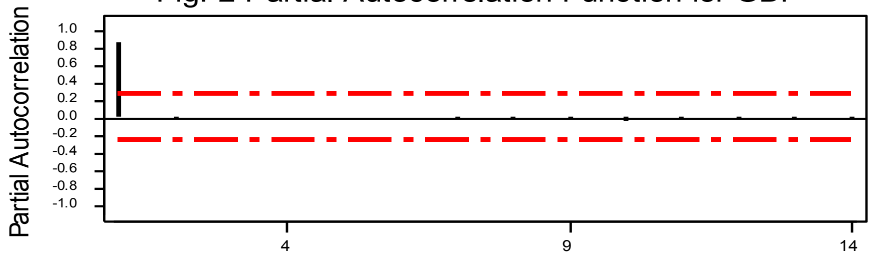
Variable	Hypothesis	t-statistic	P*
GDP	GDP has a unit root	6.011	1.00

*MacKinnon (1996) one-sided p-values.

Results (Table 1) confirm that GDP series (of 57 observations) follow stationary pattern, which in turn help to specification of Integrated (I) term in the ARIMA model

Researchers have to go further for the specification of two more terms Autoregressive (AR) and moving average (MA) in the tentative ARIMA model. A correlogram of partial autocorrelations of GDP and its different lags along with t- statistic has been presented in Fig 2.
That is, GDP data are non-stationary and need appropriate

Fig: 2 Partial Autocorrelation Function for GDP

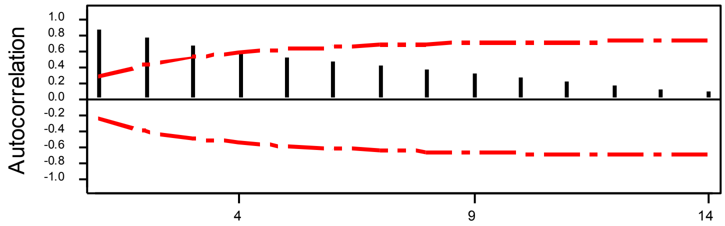


Lag	PAC	T	Lag	PAC	T
1	0.88	6.74	8	-0.02	-0.17
2	-0.01	-0.09	9	-0.02	-0.17
3	0.00	0.02	10	-0.03	-0.23
4	0.01	0.07	11	-0.02	-0.15
5	0.00	0.03	12	-0.01	-0.09
6	0.01	0.07	13	-0.02	-0.18
7	-0.00	-0.00	14	-0.01	-0.10

Results clearly indicate that the partial autocorrelation coefficient between GDP and its one period lag is only statistically significant.

In addition, a correlogram of GDP series (of 57 observations), which deals with autocorrelation of GDP and its different lags with respective t-statistics is shown in Fig 3.

Fig: 3 Autocorrelation Function for GDP



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.88	6.74	47.83	8	0.35	1.06	189.75
2	0.77	3.70	85.07	9	0.30	0.89	196.41
3	0.67	2.68	114.13	10	0.25	0.73	201.14
4	0.59	2.11	136.98	11	0.21	0.59	204.34
5	0.52	1.73	155.03	12	0.16	0.47	206.41
6	0.46	1.46	169.41	13	0.12	0.35	207.63
7	0.41	1.25	180.87	14	0.09	0.25	208.27

The evidence suggests three autocorrelation coefficients between GDP and its first, second and third lags are statistically significant. However, the coefficient of correlation between GDP and its lag-3 series is marginally significant that is why, researchers have considered only two significant coefficients. In summary, the researchers have specified the provisional model as ARIMA (1, 2, 2) with the help of PACF (in the case of AR issue) and ACF (in the case of MA issue). Finally, this model has been fitted on data of GDP (of 57 observations) to estimate all the parameters as defined earlier.

4. RESULTS

Estimation of parameters in the ARIMA model is done by maximum likelihood estimation (MLE) algorithm by assuming that the error terms are normally distributed (Ansley 1979). Results are presented in Table 2.

Table 2 Estimates of Parameters of AR and MA terms

Predictor	Coefficient	Std.E	t	p
AR1	-0.761	0.230	-3.30**	0.002
MA1	-1.145	0.266	-4.30*	0.000
MA2	-0.206	0.192	-1.07	0.288

Notes: AR= Autoregressive, MA= Moving Average,*p<.001,**p<.01

4.1 Effects of AR and MA terms:

It has been hypothesized that GDP in (t-1) period is related to GDP in current period. Support of this hypothesis($H_{1.1}$) is strongest. That is, GDP in present period has a significant negative relationship with GDP in immediate past period in the case of this macro variable.

It is hypothesized that random shock in the (t-1) and (t-2) periods are related to GDP in present period. Results support only the hypothesis ($H_{1.2}$) of (t-1) period since the t- statistic is highly significant rather than the hypothesis ($H_{1.2}$) of (t-2) period because the t-statistic is not significant at all. Hence, to forecast GDP random shock of only one period lag should be taken into consideration in the case of GDP of India.

4.2 Correlograms and Model Validation Statistic:

Correlogram of PACF of residuals (Fig 4) deals with the partial correlation coefficients between ε_t and ε_{t-p}(p represents the period of different lags. Another correlogram of ACF of residuals (Fig 5) suggests similar interpretation that there is no substantial spike has been observed as the case of correlogram of PACF of residuals. In turn, it has been concluded that the error terms become white noise so no more information is available in them leading to confirm the face validity of the model.

Fig 4 PACF of Residuals for GDP

(with 95% confidence limits for the partial autocorrelations)

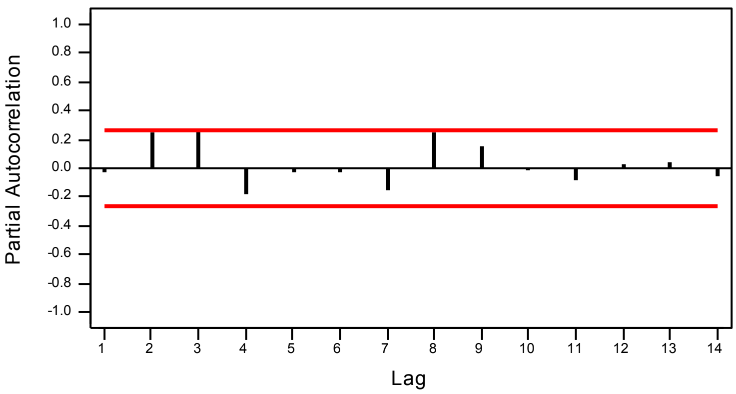
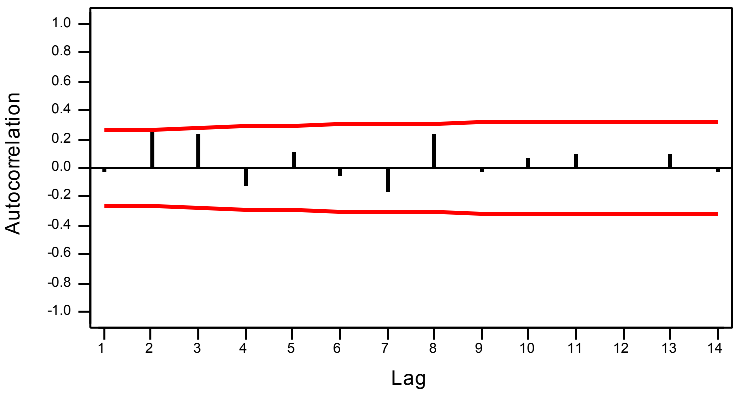


Fig: 5 ACF of Residuals for GDP

(with 95% confidence limits for the autocorrelations)



In addition of the subjective validation of model, the researcher attempted to test the validity of the model statistically. A modified portmanteau test as suggested by Ljung and Box (1978) has been performed on residuals of the model and results are reported in Table 3.

Table 3 Modified Box-Pierce (Ljung-Box) Chi-Square Statistics

Lag	Hypotheses	Chi-Square Statistic	DF	p
12	$H_0\colon \rho_{u,1} = \rho_{u,2} \dots\dots\dots\dots\dots\dots\dots. = \rho_{u,h} = 0\,, H_a = \rho_{u,i} \neq 0$	16.0	9	0.068
24	$H_0\colon \rho_{u,1} = \rho_{u,2} \dots\dots\dots\dots\dots\dots\dots.. = \rho_{u,h} = 0\,, H_a\colon \rho_{u,i} \neq 0$	17.9	21	0.653
36	$H_0\colon \rho_{u,1} = \rho_{u,2} \dots\dots\dots\dots\dots\dots\dots.. = \rho_{u,h} = 0 \quad, H_a\colon \rho_{u,i} \neq 0$	20.4	33	0.957
48	$H_0\colon \rho_{u,1} = \rho_{u,2} \dots\dots\dots\dots\dots\dots\dots.. = \rho_{u,h} = 0 \quad H_a\colon \rho_{u,i} \neq 0$	25.7	45	0.991

Evidence suggests that the Chi-square statistic is not significant even for a small sample of 13 observations from the population of 57 of GDP. That is the ARIMA model seems to be quite adequate to capture the issues of GDP forecasting and its growth rates in India.

4. DISCUSSION

The goal of this research is twofold: (1) to uncover the hidden pattern in the data on GDP of India and (2) to forecast GDP and its growth rates for a period of 10 years (2012-2021). First goal has been realized by estimating a parsimonious model of GDP, which reveals that GDP of present period is related to one period lag of its own value and one period lag of error term. To fulfill the second goal, forecasted GDP of 10 years has been estimated by this model. Results along with growth rates GDP are shown in Table 4.

Table 4 Results of forecasts and Growth Rate of GDP over a period (2012-2021).

Year	PBF	PBGR	LLBF	LLBGR	ULBF	ULBGR
2012	5925078	10.21	5867767	8.25	5982389	11.75
2013	6525433	10.10	6377268	8.70	6673598	11.55
2013	7128047	9.20	6870197	7.73	7385897	10.67
2015	7728943	8.45	7341506	6.90	8116380	9.90
2016	8331146	7.80	7799221	6.23	8863071	9.20
2017	8932354	7.21	8240801	5.66	9623907	8.60
2018	9534319	6.74	8670431	5.21	10398207	8.00
2019	10135709	6.30	9086852	4.80	11184565	7.56
2020	10737536	5.94	9492451	4.50	11982621	7.10
2021	11339030	5.60	9886659	4.20	12791401	6.75

Notes: PBF- point based forecast, PBGR- Point based growth rates, LLBF- Lower Limit based forecast, LLBGR-Lower limit based growth rates, ULBF-upper limit based forecast, ULBGR-Upper limit based growth rate.

Researchers have estimated three categories of forecast such as lower- limit based forecast (LLBF), point based forecast (PEBF) and upper limit based forecast (ULBF). These forecasts help to calculate three types of growth rates: LLBGR, PBGR and ULBGR of GDP over a period 2012- 2021. Results reveal a number of interesting findings of issues of GDP.

Forecasts and their respective growth rates are absolutely necessary to capture the variation in GDP due to the influence of uncontrollable variables in the economy. If the uncontrollable variables are assumed to be similar as were in earlier years, then it is highly suggested to use point based forecasts and its corresponding growth rates to handle the issues of GDP in near future. Similar qualitative interpretation can be drawn in the cases of other two types of forecasts and their respective growth rates

Results suggest that there is an increasing trend in the forecasted values of GDP (lower limit based, point based and upper limit based) over the period 2012-2021. However, GDP growth rates suggest that there is a decreasing trend in all the categories (lower limit based, point based and upper limit based. Hence, forecasted GDP alone cannot represent the economic health of any country. It is highly suggested that their growth rates have to be reported along with forecasted values to capture a clear picture of this important macroeconomic variable.

5. CONCLUSION

In this study, an ARIMA model has been estimated to forecast GDP and its growth rates for a couple of years ahead by utilizing Time – series data over a period 1959-2011.

Instead of using ordinary least square algorithm, the maximum likelihood estimation (MLE) algorithm has been used in this work. Results comply with the norms that suggested by the MLE algorithm such as the coefficients of AR terms is negative and less than 1 and coefficients of MA is more than 1. This aspect surely confirms the face validity of the results reported in this study. Further, statistical validity of the model is also checked by modified Ljung-Box statistic. In the end, results are very impressive since only one AR coefficient and one MA coefficient is statistically significant, which in fact estimated a very parsimonious model, which is very effective to forecast GDP and its growth rates in India.

The findings of this study have some important implications for policy makers and managers. Policy makers- who deal with macro variables such as foreign direct investment (FDI), foreign institutional investment (FII), etc., would find the results of this work are very helpful to formulate better policy. On the other hand, the manager who is planning to invest in the expansion of existing business or in new project will be benefitted tremendously since the findings will help them to portrait the picture of economic condition of India more precisely in advance. Further, findings may not be best one since the researchers do not have taken into consideration of the models such as Regression analysis, VAR, ECM etc. to forecast GDP and its growth rates in India.

However, it would be interesting to expand this research by including the factors, which may influence on GDP such as the growth rates of the population, the industrial growth rate, the immigration rate, etc. in the model in future.

REFERENCES

Ansley, C.F. (1979), An algorithm for the exact Likelihood of a mixed autoregressive-moving average process, Biometrika, Vol. 66, pp. 59-65.

Ard, H.J, den Reijer. (2010), Macroeconomic Forecasting using Business Cycle leading indicators, Stockholm: US-AB.

Box, G.E.P. and Jenkins, G.M. (1970), Time Series Analysis: Forecasting and Control, San Francisco: Holden-Day.

Box, G.P. and Jenkins, G.M. (1976), Time Series Forecasting and Control, San Francisco: Holden-Day.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994), Time Series Analysis: Forecasting and Control, Englewood Cliffs, NJ: Prentice Hall.

Chiu, C.C., Cook, D.F. and Pignatiello, J.J. (1995), Radial basis function neural network for kraft pulping forecasting, International Journal of Industrial Engineering, Vol. 2 No. 2, pp. 209-215.

Cook, D.F. and Chiu, C.C. (1997), Predicting the internal bond strength of particleboard utilizing a radial basis function neural network, Engineering Applications AI, Vol.10 No.2, pp. 171-177.

Gao, X.M., Gao, X.Z., Tanskanen, J. and Ovaska, S.J. (1997), Power prediction in mobile communications systems using an optimal neural structure, IEEE Transportation Neural Networks, Vol. 8 No. 6, pp. 1446-1455.

Granger, C.W.J. and Joyeux R. (1980), An introduction to long-memory time series models and fractional differencing, Journal of Time Series Analysis, Vol. 1, pp. 15-39.

Enders, W. (2004), Applied Econometric Time Series, New York: John Wiley and Sons.

Hoff, J.C. (1983), A Practical guide to Box-Jenkins Forecasting, London: Lifetime Learning Publications.

Hosking J.R.M. (1981), Fractional differencing, Biometrica, Vol. 68, pp. 165-176.

Hossain, M.Z., Samad, Q.A. and Ali, M.Z. (2006), ARIMA model and forecasting with three types of pulse prices in Bangladesh: A case study, International Journal of Social Economics, Vol. 33 No. 4, pp. 344-353.

Liu, L.M. (1989), Identification of seasonal ARIMA models with using a filtering method, Communications in Statistics, Vol. A18, pp. 2279-2288

Liu, L.M. (1999), Forecasting and time series analysis using the SCA Statistical System: Vol. 2, Chicago: Scientific Computing Associates Corp.

Liu, L.M. and Hudak, G.B. (1992), Forecasting and time series analysis uses the SCA Statistical System: Vol. 1 Chicago: Scientific Computing Associates Corp.

Ljung, G.M. and Box, G.E.P. (1978), On a measure of lack of fit in time series models, Biometrika, Vol.65 No. 1, pp. 297-303

McDowall, D., McCleary, R., Meidinger, E.E. and Hay, R. A. (1980), Interrupted time series Analysis, Baverly Hills, CA: Sage Publications.

Melard, G. (1984), A fast algorithm for the exact likelihood of auto regressive- moving average models, Applied Statistics, Vol. 33, pp. 104-119

Ning, W., Kuan-jiang, B. and Zhi-fa, Y. (2010), Analysis and forecast of Shaanxi GDP based on the ARIMA Model, Asian Agricultural Research, Vol. 2 No. 1, pp. 34-41.

O'Donovan, T.M. (1983), Short Term Forecasting: An Introduction to the Box-Jenkins Approach, New York: John Wiley and Sons.

Pankratz, A. (1983), Forecasting with Univariate Box-Jenkins models. Concepts and Cases, New York: John Wiley and Sons.

Pankratz, A. (1991), Forecasting with dynamic regression models, New York: John Wiley and Sons.

Reilly, D.P. (1980), Experiences with an automatic Box-Jenkins modeling algorithm. Time Series Analysis- Proceedings of Houstan Meeing on Time Series Analyis, Amsterdam: North-Holland publishing.

Reynolds, S.B., Mellichamp, J.M. and Smith, R.E. (1995), Box-Jenkins forecast model identification. AI Expert, June, pp. 15-28.

Saad, E.W., Prokhorvo, D.V. and Wunsch, D.C. (1998), Comparative study of stock trend prediction using time delay. Recurrent and probabilistic neural networks, IEEE Transportation Neural Networks, Vol. 9 No. 6, pp. 1456-1469

Tsay, R.S. and Tiao, G.C. (1984), Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and non-stationary ARMA models, Journal of American Statistical Association, Vol. 79, pp. 84-96.

Tsay, R.S. and Tiao, G.C. (1985), Use of canonical analysis in time series model identification, Biometrika, Vol. 72, pp. 299-315.

Yang, Lu. (2009), Modeling and a forecasting China's GDP data with time series model. Thesis unpublished.

Vandaele, W. (1983), Applied time series and Box-Jenkins models, New York: Academic Press.