© Journal of Contemporary Issues in Business Research
ISSN 2305-8277 (Online), 2014, Vol. 3, No. 4, 200-210.

A TIME SERIES MODELING ON GDP OF PAKISTAN*

MANIHA ZAKAI†
University of Central Punjab

ABSTRACT

The gross domestic product (GDP), an essential gauge of an economy's economic presentation, is the market assessment of all final services and goods produced in the boundaries of a nation in a year. Within this paper, the features of annual data of Pakistan's GDP acquired from International Monetary Fund (IMF) starting from 1953 to 2012 are studied. To model the GDP, a set of Autoregressive integrated Moving Average (ARIMA) models are constructed following Box-Jenkins technique. ARIMA (1,1,0) has been obtained through expert modeler method by considering best fit model. Finally, forecast values for a few coming year have been generated applying the best fitted ARIMA model. The finding shows that the forecast values of Pakistan's GDP will be 23477Billion rupees in 2013 and 103918 billion rupees in 2025.

Keywords: ARIMA model; GDP; Pakistan; Time series modeling.

INTRODUCTION

Sustainable economic growth is of foremost concern for every economy, mainly for the developing economies which usually face tribulations. The gross domestic product (GDP), is the crucial evaluation of an economy's economic sketch, and refers to the market value of all final services and goods created in the country in a specified phase of time. Economists have thus, focused to look at the channels through which GDP can probably improve economic growth. GDP can be described with three approaches; the income, the expenditure and the product approach. The rule of income approach is that the incomes of the manufacturer must be equivalent to worth of their product, and find out the GDP by summation of every producer's incomes. The expenditure approach rule is that every product must be purchase by someone (Kennedy, 2000).

GDP is intended to measure the market worth of production that flow throughout the economy, (a) consist of only services and goods bought by their ultimate consumer, thus GDP deals with final production. (b) Keep out transfer payments and financial transactions as they do not stand for present production. (c) Count up only the services and goods produced inside the country's borders for the duration of a year, whether through foreigners or citizens. (d) Measures both productivity and income, which are equivalent.

The economy of the Islamic Republic of Pakistan is the world's 27th largest economy found on its purchasing power. Pakistan economic development faced a serious hinder in fiscal year 2009 for the reason that the dejected consumer credit market, sluggish growth of

* The views or opinions expressed in this manuscript are those of the author(s) and do not necessarily reflect the position, views or opinions of the editor(s), the editorial board or the publisher.
† Corresponding author affiliation: Faculty of Commerce, University of Central Punjab, Lahore

public sector programs, inflation, cutback in subsidies, security menace, energy crisis and instability in the state. Furthermore, no concentration was given to the agriculture division. The exports and imports turn down by six percent and 10 percent respectively. The only credible thing was the increase in remittances by 22%. Agriculture division has revealed convincing outcomes for the reason of good weather. Major crops, rice, wheat and maize proof notable growth i.e 7.7% against the goal of 4.5%. Live stock and poultry as well put in to GDP as there was no viral disease last year by (Pakistan, 2011).

This paper is resting on the time series data of Pakistan's GDP from 1953 to 2012; it is measured the class of ARIMA model (autoregressive integrated moving average). It is aimed at revealing the regularity of Pakistan's GDP Growth, via the fitted ARIMA model to carry out an out-fit-sample forecasting.

LITERATURE REVIEW

Start test from here. GDP is one the main needle exercised to find out the health of a country's economy. GDP be a sign of the total dollar value of all services and goods produced over a particular time period, one can suppose of it as the size of the economy. Therefore an extensive effort has been made to search the literature concerning with the time series modeling on GDP using ARIMA model.

Lu (2009) attempt to construct a time series model which was utilized to forecast the gross domestic product to get its future estimations up to first quarter of 2009.The study found on the figures collected through secondary sources from 1962 to 2008.ARIMA models were seek on the collected data and to conclude ARIMA (4, 1, 0) is create to be an appropriate model, which is then apply for forecasting purpose. The outcome of the future forecast explains the significant improvement in the fourth quarter of 2009.

Li, Liu and Zhao (2007) focused on Monetary policy, have an effect on the economy among long and variable lags, and meant for this explanation policy-makers have need of reliable forecasts of economic activity. Thus, forecasts of real GDP growth have turn into more and more needed. Projected a new modified ARIMA model and make use of it to forecast the GDP growth of China since 1978 to 2004. Within this paper, they propose a new genetic programming technique to forecast the GDP time series of United States, Japan and China from 1980 to 2006. In addition, they apply the proposed technique to forecast the prospect GDP growth of United States, Japan and China from 2007 to 2020, and they come across that the GDP of Japan vary periodically, though the GDP of United States and China raise steadily in the near future. With the predicted data it is found that by 2020 the GDP of China will go above the GDP of Japan for the first time.

The evaluation on the state of the Indian economy throughout recession by analyzing different macro-economic factors such as GDP, inflation, exchange rate, fiscal deficit and capital markets. This study forecast some of the major economic variables by ARIMA modeling and presents a depiction of the Indian economy in the coming years. The results indicate that Indian economy is stimulating after a slowdown in the phase of global recession. It is forecasted that GDP, fiscal deficit, capital markets and foreign investments will increase in 2010-11 (Sinha, 2010).

The study of Mamun (2009) focused on ARIMA model to evaluate and forecast GDP, figures are extracted from Bangladesh Bureau of Statistics .This paper made a evaluation of two forecasting methods to observe which suits in forecasting the GDP and its divisions. ARIMA model plays quite an admirable part in uni-variate econometric time series data study of GDP of Bangladesh and its divisions.

Ning (2010) made an effort to build a time series model which was then utilize to forecast the GDP of Shaanxi to achieve its future assessment up to the year 2013.This paper based on the figures collected from secondary sources as of the year 1952-2007. ARIMA models were seek on the collected figures and at last ARIMA (1, 2, 1) is found to be suitable model, which is then employ for forecasting purpose. The outcomes of the prospect forecasts showed the considerable progress in GDP of Shaanxi Province up to the year 2013.

RESEARCH METHODOLOGY

In this paper the aim is to build a suitable and the most appropriate model which can be used to forecast the GDP of Pakistan. In this chapter: Box-Jenkins Methodology, Unit Root test, Dickey Fuller Test, The Augmented Dickey Fuller Test, ARIMA model and diagnostic checks are described.

The Box-Jenkins Methodology

Box and Jenkins (1976) admired method intended at opt for a suitable model for the reason of forecasting and estimation a uni-variate time series. To exercise the Box-Jenkins Methodology, one should have either a stationary time series otherwise a time series that could be stationary after one or two differencing. The explanation for necessitate stationary data is that every model which is inferred by these data be capable interpreted as stable, thus provided that convincing basis for forecasting. The Box Jenkins methodology consist of the four steps: 1) Identification,2) Estimation, 3)Diagnostic Check, and 4) Forecasting.

Unit Root test

A test of stationary that has become extensively admired over the past several years is the unit root test. Considering the unit root (stochastic) method:

$$y_t = p y_{t-1} + e_t \qquad -1<p<+1$$

where "e" is a white-noise error term .As if p = 1, that is in the unit root, the above turn into:

$$y_t = y_{t-1} + e_t$$

i.e a random walk model exclusive of drift, which is well-known as a non-stationary stochastic process. Thus if the estimates p is statistically equal to one, then y_t is stationary. This is the basic idea at the back the unit root test of stationary.

Mostly time series data have problem of unit root, so application of unit root test is valuable in time series data related studies (Bhunia, 2013; Husin, 2013; Mehmood & Ahmad, 2013; Mehmood (2012); Naz (2012).

Dickey Fuller Test

The most famous test, recognized as Dickey Fuller Test (DF) test was projected in 1979 by D.A Dickey and W.A Fuller. To verify the stationary of any procedure, this test is extensively used in lots of researches.

Test for a unit root:

$$\nabla y_t = \delta y_{t-1} + u_t$$

Test for a unit root with drift:

$$\nabla y_t = a_0 + \delta y_{t-1} + u_t$$

Test for a unit root with drift and deterministic time trend:

$$\nabla y_t = a_0 + a_1 t + \delta y_{t-1} + u_t$$

Each version of the test possesses critical value which depends on the size of the sample. In each of the above, the null hypothesis is that there is a unit root, $\delta = 0$. ($\delta = p-1$). The tests have little statistical power in that they cannot differentiate among true unit-root processes ($\delta = 0$) and near unit-root processes (δ is close to zero) means the unit root and time series is non stationary. The alternative hypothesis is that; $\delta < 0$; that is the time series is stationary. The possibility of the hypothesis of $\delta > 0$ is deficient, as in that case p is greater than one, which is not achievable.

The Augmented Dickey Fuller Test

The main supposition closed with the DF test that the error terms are uncorrelated. But while these terms are interrelated, Dickey and Fuller developed one more test acknowledged as Augmented Dickey Fuller test in this they deal with this difficulty by including the lag terms of the dependent variable. The testing procedure for the ADF test is the same as for the Dickey–Fuller test but it is applied to the model

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t,$$

Where "a" is a constant, *p* the lag order of the autoregressive process and β the coefficient on a time trend. Imposing the limitation "a" = 0 and "b" = 0 match up to modeling a random walk and using the limitation "b" = 0 match up to modeling a random walk with a drift. By including lags of the order *p* the ADF formulation permit for higher-order autoregressive processes.

ARIMA model (Auto-Regressive Integrated Moving Average)

Box and Jenkins (1976) united the two autoregressive and moving average terms in order to shape an ARMA model.

An autoregressive model;

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots + a_p Y_{t-p} + e_t$$

A moving average model is:

$$y_t = et + b_1 e_{t-1} + b_2 e_{t-2} + \ldots + b_q e_{t-q}$$

Then ARMA (p,q) model is specified as

$$Yt = a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots + a_p Y_{t-p} + e_t + et + b_1 e_{t-1} + b_2 e_{t-2} + \ldots + b_q e_{t-q}$$

As mainly the economic time series are non-stationary in nature afterward, the order of "integration" at which the series develop into stationary is built-in, in the model. If y_t is not stationary moreover it is stationary at "d" difference, the ARIMA (p,d,q) model;

$$\Delta_d\, y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots + a_p Y_{t-p} + e_t + et + b_1 e_{t-1} + b_2 e_{t-2} + \ldots + b_q e_{t-q}$$

Diagnostic Checks

Start test from here. Subsequent are few diagnostic checks which each estimated model has to fulfill.

- The residuals are normally distributed.
- The model is stable.
- The residuals of the projected model are not serially interrelated.

The q-statistic. In any huge group of autocorrelations, some surpass two-standard deviation because of pure possibility even while the true values in the data generating process are "0". The Q-statistic is exercise to test whether a set of autocorrelations is significantly diverse from zero. Box and Pierce (1970) make use of sample autocorrelation to form the statistics.

$$Q = T \sum_{k=1}^{s} r_k^2.$$

In the null hypothesis that every values of = 0, is asymptotically distributed by degree of freedom. A difficulty by way of the Box-Pierce Q-statistic is that it workings inadequately even in moderate huge samples.

Histogram-normality test and jarque-bera statistic. Histogram-Normality test can be applied to check the normality of the residuals. It shows the histogram of the standardized residuals and descriptive statistics. Jarque-Bera statistic is utilized to check the normality of the residuals

Model selection Criteria

Schwaarz information criteria and akaike information criteria. SIC is applied to enforce harsher penalty than AIC for adding huge number of regressors in the model. AIC is exercised to enforce penalty for adding increasingly huge number of regressors in the model. Both the criteria can be utilizing in sample as well out-of-sample forecasting lower values of SIC and AIC is chosen.

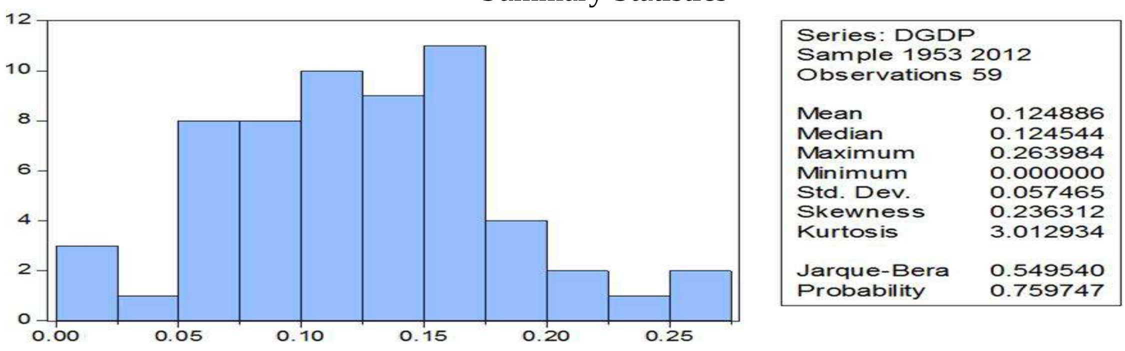
Goodness of Fit. The major concern is to approximate a model that can elucidate the data well and such a model is measured as a superior projected model. Within regression theory, two measures; R ^2, and sum of squares of residuals, are the vital measures to verify the goodness of fit of any model. Although there is one weakness linked by these measures is that they depend on the number of the parameters incorporated in the model. The raise in the number of parameters reason the loss of degree of freedom. Because of this weakness, two other criterion, known as SBC (Schwartz Bayesian Criteria) and AIC (Akaike Information Criteria) are preferred for the purpose of model selection.

EMPIRICAL RESULTS AND CONCLUSION

The foremost purpose of this study is to construct a forecast model meant for the GDP of Pakistan. The annual data of variable for the period 1953-2012 comprising of 60 observations are obtained from the IMF database .The basis for the selection of this time period is that it reveal the pattern of GDP of Pakistan in six decades.

FIGURE 1

Summary Statistics



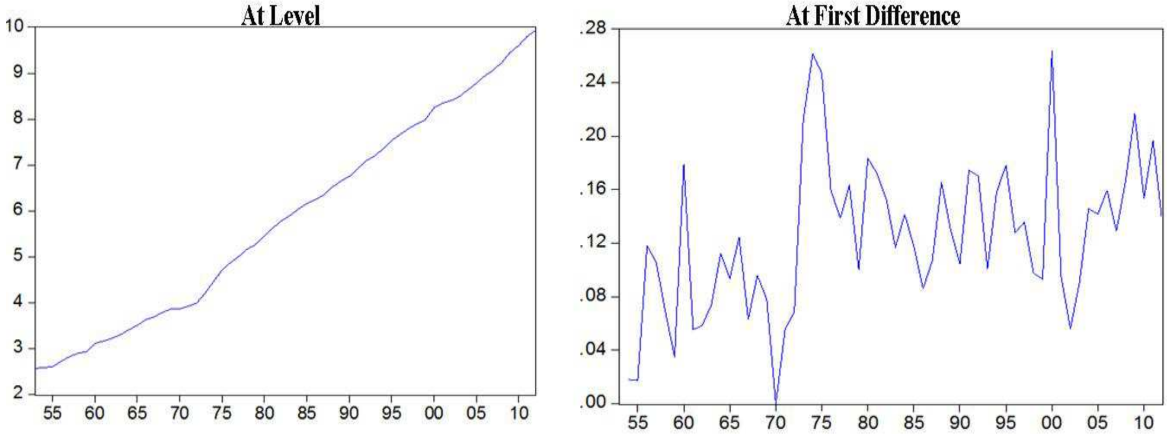
Summary statistics in Figure 1 contains values in natural logarithm which indicated the Pakistan average GDP was .124 billion rupees during the last 60 years from 1953 to 2012. And the most balanced point of the GDP during this period was 0.124 billion rupees. In this duration the maximum GDP was 0,263 billion rupees. This summary tells the story of an increasing trend in the GDP of Pakistan. The statistics also point out the low variation of 0.057 billion rupees from the average of GDP. The skewness is greater than zero reveal the right skewed distribution, which is 0.236 while kurtosis is 3.012, shows the leptokurtic distribution.

The basic prerequisite of time series analysis is that the utilized data have to be stationary. Here two measures presented to verify the stationary of the data (the subjective and the objective). Each of these process has its due importance and merits and demerits. The subjective method is based on the idea of illustration assessment of the correlogram of auto-correlations and partial-autocorrelations at different lag ideals. Though in objective method is the mathematical measures these measures incorporated the Philips Perrin (PP) and of Unit-Root Tests of Augmented Dickey Fuller (ADF) a type. Within this study both the objectives and subjective measures are used for checking the stationarity of the observed figures.

Figure 2 shows the line graphs of GDP at level and first difference respectively. It is observed that the series of GDP is not stationary at level however it becomes stationary at the first difference.

FIGURE 2

Line Graphs of Log of GDP at level and at First Difference

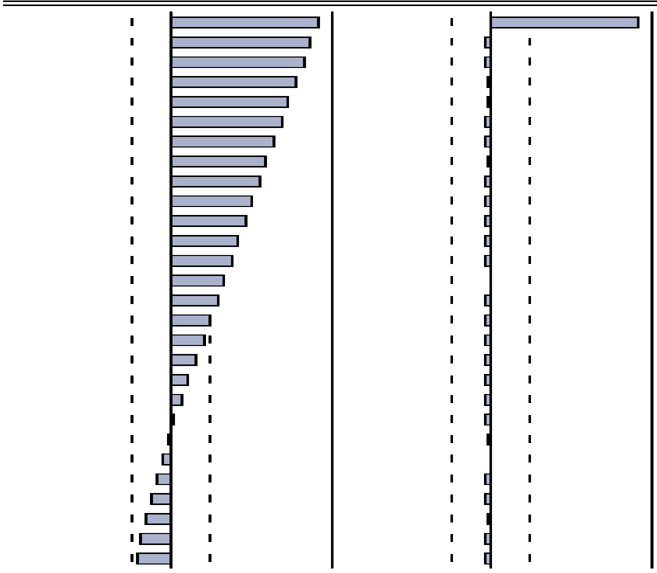


The same results are shown by correlograms at level and at first difference are displayed in Figure 3 and Figure 4.

FIGURE 3

Correlogram at Level

Date: 02/13/13 Time: 15:12
Sample: 1953 2012
Included observations: 60



	AC	PAC	Q-Stat	Prob
1	0.953	0.953	57.293	0.000
2	0.905	-0.038	109.85	0.000
3	0.857	-0.028	157.77	0.000
4	0.809	-0.022	201.25	0.000
5	0.762	-0.014	240.55	0.000
6	0.716	-0.025	275.83	0.000
7	0.668	-0.035	307.17	0.000
8	0.622	-0.013	334.88	0.000
9	0.576	-0.031	359.09	0.000
10	0.530	-0.031	379.97	0.000
11	0.483	-0.036	397.68	0.000
12	0.435	-0.038	412.38	0.000
13	0.388	-0.035	424.28	0.000
14	0.342	-0.008	433.77	0.000
15	0.297	-0.040	441.04	0.000
16	0.251	-0.037	446.36	0.000
17	0.205	-0.034	449.99	0.000
18	0.159	-0.045	452.22	0.000
19	0.113	-0.033	453.37	0.000
20	0.068	-0.034	453.80	0.000
21	0.024	-0.025	453.86	0.000
22	-0.016	-0.009	453.88	0.000
23	-0.054	-0.008	454.17	0.000
24	-0.090	-0.027	455.01	0.000
25	-0.125	-0.026	456.67	0.000
26	-0.157	-0.016	459.38	0.000
27	-0.189	-0.031	463.41	0.000
28	-0.219	-0.026	469.00	0.000

FIGURE 4

Correlogram at First Difference

Date: 02/13/13 Time: 15:13
Sample: 1953 2012
Included observations: 59

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.408	0.408	10.352	0.001
		2	0.152	-0.018	11.813	0.003
		3	0.132	0.091	12.935	0.005
		4	0.112	0.033	13.749	0.008
		5	0.035	-0.036	13.830	0.017
		6	0.049	0.046	13.996	0.030
		7	0.063	0.023	14.273	0.047
		8	0.111	0.087	15.149	0.056
		9	0.134	0.066	16.439	0.058
		10	-0.080	-0.212	16.909	0.076
		11	-0.020	0.086	16.938	0.110
		12	-0.049	-0.104	17.121	0.145
		13	0.003	0.084	17.121	0.194
		14	0.058	0.070	17.388	0.236
		15	0.084	0.019	17.960	0.265
		16	0.095	0.070	18.716	0.284
		17	0.081	-0.021	19.276	0.313
		18	-0.005	-0.057	19.277	0.375
		19	-0.146	-0.129	21.203	0.326
		20	-0.079	-0.008	21.772	0.353
		21	-0.144	-0.109	23.750	0.305
		22	-0.070	0.024	24.232	0.335
		23	-0.099	-0.089	25.206	0.340
		24	-0.072	0.006	25.737	0.367

Then the ADF test and PP test (both at level and at first difference) are applied. .The null hypothesis of having unit root is not rejected at level if the p-value of the test statistic exceeds the observed level of significance α=0.05. In Table 1 to Table 4, results of ADF test and PP test on GDP indicates that the hypothesis: GDP has a unit root, is acknowledged that means the series is non stationary at level. Whereas results from Figure 2 and Figure 4 evidently indicate that the series is stationary at first difference.

TABLE 1

ADF Test at Level

Null Hypothesis: GDP has a unit root			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		1.697985	0.9996
Test critical values:	1% level	-3.548208	
	5% level	-2.912631	
	10% level	-2.594027	

TABLE 2

ADF Test at First Difference

Null Hypothesis: D(GDP) has a unit root			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.033062	0.0001
Test critical values:	1% level	-3.548208	
	5% level	-2.912631	
	10% level	-2.594027	

TABLE 3

PP Test at Level

Null Hypothesis: GDP has a unit root

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		2.626420	1.0000
Test critical values:	1% level	-3.546099	
	5% level	-2.911730	
	10% level	-2.593551	

TABLE 4

PP Test at First Difference

Null Hypothesis: D(GDP) has a unit root			
		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-5.033062	0.0001
Test critical values:	1% level	-3.548208	
	5% level	-2.912631	
	10% level	-2.594027	

The initial stair in Box-Jenkins methodology is recognition: to be come across the appropriate principles of p (Auto-regressive term), d (no of difference taken) and q (moving average term), which depend on the presentation of correlograms of auto-correlations and partial auto correlations. The significance of p is obtained from correlogram of partial auto correlations and the significance of q is attaining from the correlogram of auto-correlation.

TABLE 5

Best Fit Model Description Based on SPSS Expert Modeler Method

			Model Type
Model ID	GDP	Model_1	ARIMA(1,1,0)

Table 5 indicate best fit model which has been obtained by using SPSS expert modeler method. Furthermore, Table 6 contains the results of best fit model statistics, for this it is considered that the value of R-squared must be maximum and remaining values of Stationary R-squared, RMSE, MAPE, MaxAPE, MAE, MaxAE and Normalized BIC must be minimum. While analyzing the above Table 7, it has been noticed that the parameters of ARIMA (1, 1, 0) are acceptable and recommended. The parameters of ARIMA (1, 1, 0) have been suggested by experts modeler using by SPSS software (version 16).

TABLE 6

Best Fit Model Statistics

Model Fit statistics						Ljung-Box	Number
Stationary R-squared	R-squared	RMSE	MAPE	MAE	Normalized BIC	Q Statistics	of Outliers
.167	.999	.053	.864	.041	-5.740	6.792	0

TABLE 7

Best fit model based on SPSS expert modeler method

Model	Parameters	Estimate	S.E.	t-ratio	P-value
ARIMA	C	.124	.012	10.487	.000

(1,1,0)	AR1	.427	.119	3.577	.001

Econometric equation of best fit ARIMA model is as follow:

$$\Delta GDP_t = 0.124 + 0.427\ AR_{t-1} + e_t$$

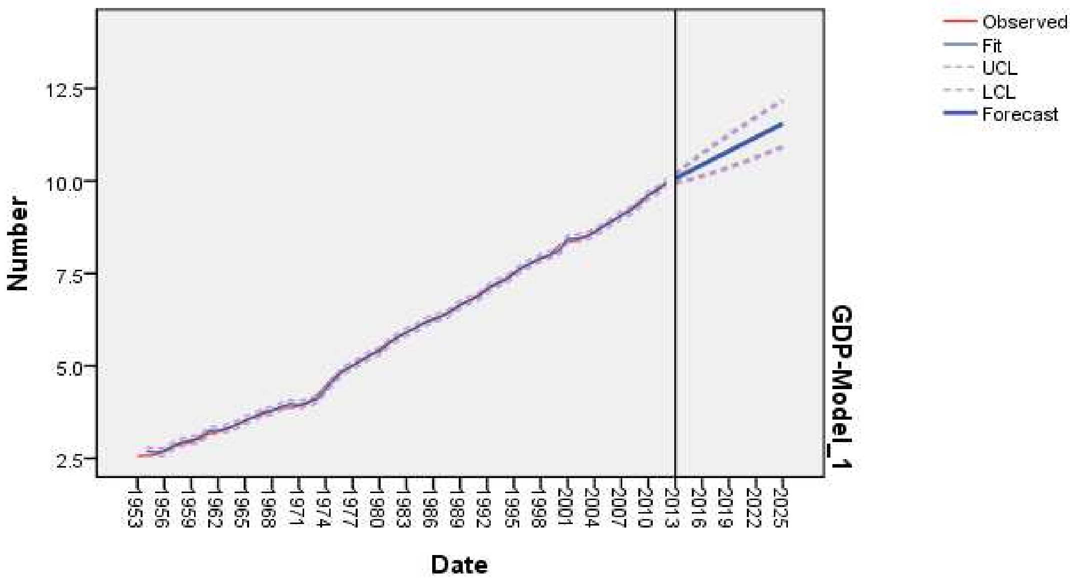
TABLE 8

Forecast Values of Pakistan's GDP in Billions Rupees

Year	Forecasts Values	Forecasts Values with Lower Class Limits (LCL)	Forecasts Values with Upper Class Limits (UCL)
2012-2013	23477	21131	26083
2013-2014	26618	22157	31977
2014-2015	30147	23486	38697
2015-2016	34128	25121	46364
2-16-2017	38627	27060	55140
2017-2018	43716	29303	65218
2018-2019	49473	31863	76816
2019-2020	55988	34759	90184
2020-2021	63360	38015	105603
2021-2022	71703	41665	123395
2022-2023	81144	45747	143928
2023-2024	91828	50305	167624
2024-2025	103918	55389	194967

FIGURE 5

Visual presentation of Forecast values of Pakistan's GDP



CONCLUSIONS AND RECOMMENDATIONS

Sample forecasting is done for the period of 1953 to 2009, visual presentation of forecast values reveal a good behavior. ARIMA (1,1,0) has been obtained through expert modeler method by considering best fit model. Finally, forecast values for a few coming year have been generated applying the best fitted ARIMA model. The finding shows that the forecast values of Pakistan's GDP will be 23477Billion rupees in 2013 and 103918 billion rupees in 2025. Recommendations for future Study are as follow:

- Due to certain limitations the study only focused ARIMA model is used to forecast the GDP of Pakistan but other models like ARCH and GARCH models could be consider better as well
- Furthermore the forecasting of GDP indicating factors can be conducted
- A very important study can be conducted by forecasting and comparing the economy indicators like GDP,CPI(consumer price index), GNP and unemployment can be study to check the future trend an pattern of economy indicator of Pakistan

REFERENCES

Box, G. E., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, *65*(332), 1509-1526.

Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control, revised ed*. Holden-Day.

Bhunia, A. (2013). Cointegration and Causal Relationship among Crude Price, Domestic Gold Price and Financial Variables: An Evidence of BSE and NSE. *Journal of Contemporary Issues in Business Research , 2* (1), 01-10.

Husin, M. M. (2013). Efficiency of Monetary Policy Transmission Mechanism via Profit Rate Channel for Islamic Banks in Malaysia. *Journal of Contemporary Issues in Business Research , 2* (2), 44-55.

Kennedy, P. (2000). *Macroeconomic Essentials:Understanding economics in the news.* Massachusetts Institute of Technology.

Li, M., Liu, G., & Zhao, Y. (2007, August). Forecasting GDP growth using genetic programming. In *Natural Computation, 2007. ICNC 2007. Third International Conference on* (Vol. 4, pp. 393-397). IEEE.

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297-303.

Lu, Y. (2009). Modelling and forecasting China's GDP data with time series models. *D-level Essay in Statistics Department of Economics and Society, Högskolan Dalarna, Sweden*.

Mamun, M. A. (2009). A study of GDP sectors.

Mehmood, S. (2012c). Forecasting Pakistan's Exports to SAARC: An Application of Univiriate ARIMA Model. *Journal of Contemporary Issues in Business Research , 1* (3), 41-54.

Mehmood, S., & Ahmad, Z. (2013). Time Series Model to Forecast Area of Mangoes from Pakistan: An Application of Univariate Arima Model. *Academy of Contemporary Research Journal , 2* (1), 10-15.

Naz, F. (2012). A Univariate Time Series Modelling of Dates Exports in Pakistan. *Journal of Contemporary Issues in Business Research* , *1* (2), 37-48.

Ning, W., Kuan-jiang, B., & Zhi-fa, Y. (2010). Analysis and forecast of Shaanxi GDP based on the ARIMA Model. *Asian Agricultural Research*, *2*(01).

Pakistan, S. B. (2011). *The State of Pakistan's Economy - Second Quarterly Report 2010 - 2011*.