

## SHORT TERM FORECASTING OF STOCK MARKET SHARE PRICES BY ARIMA MODELS

Muhammad Iqbal and Muhammad Khalid Pervaiz<sup>1</sup>

**ABSTRACT:** *The study attempts to provide a systematic analysis of daily closing stock price of four Companies. Askari Commercial Bank, Shell, Sui Northern Gas and PIA. The weekly averages of the daily prices are used Box- Jenkins procedure of forecasting is adopted. A variety of ARIMA models are obtained and best models are proposed. The assumption of stationarity is followed and to identify the best model two sets of statistics AC's and PAS's are applied. The estimates of the parameters are derived. According to the process of validation, three components, residual diagnostic, parameter validation, and model validation are followed. To see the behaviour of residual the normal probability plot of residuals together with the plots of auto-correlation function and partial auto-correlation function are obtained. These show that residuals are independently and normally distributed.*

### 1. INTRODUCTION

Starting in the late 1950, both business firms and nonprofit organizations began to display a much greater interest in forecasting. As organizations of all types have become more complex, managerial decision makers have been forced to develop a broader and more systematic view of the future that incorporate both the dynamics of the domestic market place and the expanding importance of global economic activity.

Forecasting the future values of an observed time series is an important problem in many fields like sales forecasting and stock control. There are many types of forecasting procedures. Forecasting is based entirely on past observation in a given time series by fitting a model to the data and extrapolating. Until the 1970's there was relatively very little material available on time series analysis.

Univariate time series models and Univariate Auto-regressive moving average (ARMA) time series models have been extensively used in practice for forecasting economic time series and are generalization of the exponentially weighted moving average process which was extensively applied in inventory and sales modeling.

The objective of the study is to forecast the share prices of the following companies:

---

<sup>1</sup> Government College University, Lahore, Pakistan.

- i. Askari Commercial Bank
- ii. Shell,
- iii. Sui Northern Gas,
- iv. PIA

The models are suggested to forecast the share prices of these companies. These will provide a chance for planners of concerned companies to frame their future investment policies. Forecast should not be considered as being permanent or statistic. The dynamic nature of the market place dictates that the forecast is always be subject to review, revision and discussion. Thus forecasting should be viewed as a multi step process.

## 2. MODEL IDENTIFICATION

One of the critical assumption of Box-Jenkins methodology is the assumption of stationarity. If actual time series are non-stationary, they can be converted to a stationary time series through the process of differencing. If still the assumption of stationarity is violated the next stage is second differences or log transformation.

To identify the model that best describe the time series under consideration two sets of statistics are used AC's and PAC's. The AC's and PAC's measure how much interdependency, there is among the observation and take values that range between  $-1$  and  $+1$  depending on the pattern of relationship.

After analyzing AC's and PAC's the second stage in the Box-Jenkins model building process is to estimate the model parameters. As with regression models Box-Jenkins models are base on sample statistics that must be tested to assume their validity as estimates of the true population parameters. The process of validation has three components.

1. Residual Diagnostics.
2. Parameters Validation.
3. Model Validation.

Within the Box-Jenkins framework a key assumption is made that the residual terms are not correlated with each other. There are numerous residual diagnostic statistics, which are as follows:

1. The residual Mean and Mean percent error that measures the existence of any bias in the residuals.
2. Auto correlations of the residuals that test the validity of the random error assumption.

Model validation focuses on the reliability of the forecast generated by the model. One of the main advantages of Box-Jenkins models vis-a-vis other time series models are that systematic approaches for the measurement of uncertainty have been developed based on the use of forecast confidence interval. The forecast confidence intervals used in ARIMA programme generally rely on the residual standard error. The second procedure used as a gauge of the reliability of the forecast is the evaluation of ex-post for after the fact forecast. This procedure involves leaving out certain number of actual data points from the estimations phase and then generating forecast that can be compared with the actual values. Box-Jenkins or ARIMA models can be constructed to fit a wide variety of patterns, and this can be done with a minimum effort as long as computer programs are available.

### 3. SELECTION OF COMPANIES

The stock market being a barometer of a country's economy recovered considerably during the year 2000, which is reflected by KSE-100 index. The market remained bearish for the last two years and the investors confidence was shaken. But due to some macro-economic and structural performs have restored the investors confidence and changed the perception to a reasonably good level and the investors are attracted to make investment in securities. Amongst global emerging markets Pakistani Stock Market has been on top with the highest percentage gain achieved during the year 2000. The emphasis is shifting towards developing private sectors. Information technology is being encouraged and it has planned to use the local stock exchanges to float a part of total offering of state entities through the stock exchanges. Another phenomenal taking shape with the general rise seen in the stock market is the revival of the Mutual Fund Industry. The funds are becoming more attractive and their portfolios attracting new investors to capitalize upon the current trends. With the recovery of the Pakistani Stock Market it is very good time for foreign funds to make their entry. The long term outlook is very bullish and at the same time the markets are giving enough movement to pick up blue chips at very attractive prices with good chances of reaping sizable capital gain on medium term to long term bases. In short, Pakistani Stock markets are on the roads of better performance and efficiency. Their volatility and essential part of all emerging markets creates higher risks but it also magnifies the potential of higher return for any investor who has the patience and sophistication of participate in the growth of its emerging market.

In recent years, Karachi Stock Exchange has become an important emerging stock market of the region of ever growing turn over and high potential returns. It has tremendous growth potential. In fact, KSE has experienced considerable growth during the last decade. It has become first most active stock market of Pakistan. A total of 747 companies are listed as KSE divided into fifteen major groups i.e. Cotton and Textile, Chemical and

Pharmaceuticals, Engineering, Auto and Allied, Cables and Electrical goods, Sugar and Paper, Cement, Fuel, and Energy, Transport and Communication, Banks and Financial Institutions, Mutual Funds and miscellaneous.

We have selected four companies from different sectors. These companies have maximum total turn over in their groups. These have a significant role in trading at stock exchange in terms of their demand, and volume of trading.

#### 4. NUMERICAL INFORMATION

Shares are traded at KSE on every working day. The KSE issues daily bulletin of prices of shares traded at the market. Opening price, closing price, highest price, lowest price and the volume traded are shown in the bulletin. Department of Statistics, State Bank of Pakistan collects this bulletin every day and use it as an initial data. Thereafter, DS-SBP calculates simple yearly averages. Since the main objective of this study is to look at the optimum forecast, the daily closing prices of shares of selected companies are taken and the original data are transformed into weekly averages of the daily closing prices.

A cursory look at the weekly prices from January 1995 to February 2001 gives a mixed behavior. Prices of shares went up as well as down as the time passes. What ever may be the reason of change in prices the time variable has its own role to play? The present position of multi-national company, Shell, is as follows:

Shell incorporated in Pakistan with a share value of Rs.10/- and attained a highest average share price of Rs.231.8 in 1994-95 and the lowest price of Rs.7.89 in 1973-74 perhaps this is the only company under discussion, which started with a lowest rate and rose to a very high average price of its share.

#### 5. MODEL BUILDING

##### 5.1 Askari Commercial Bank

To check the stationarity of time series the actual value were plotted and it was found that the series was not stationarity. Figure 5 shows the plot of actual data and the first differences. The plot of first differences seems to satisfy the condition of stationarity. Figure 5.1(b) and 5.1(c) shows the auto-correlation function and partial auto-correlation function of the first differences. ARIMA (0,1,1), ARIMA (1,1,0), ARIMA (2,2,0) on the first differences of the data were fitted using SPSS package. The results are summarized in Table-5.1. The study of the table reveals that ARIMA (2,2,0) is the best fitted model. The values of the AIC, SBC, log likelihood and the significant level of estimate suggest ARIMA (2,2,0). To see the behavior of the residual the normal probability plot of

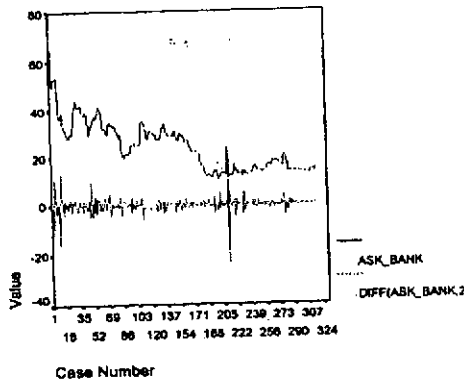
residuals together with the plots of auto-correlation function and partial auto-correlation function are shown in Figure 5.1(e), 5.1(f) and 5.1(g) respectively. These figures confirm that the residual are independently and normally distributed. The model for the data is

$$X_t = 0.0021 + 0.1549\epsilon_{t-1} + 0.8395\epsilon_{t-2} - 0.14\epsilon_{t-3}$$

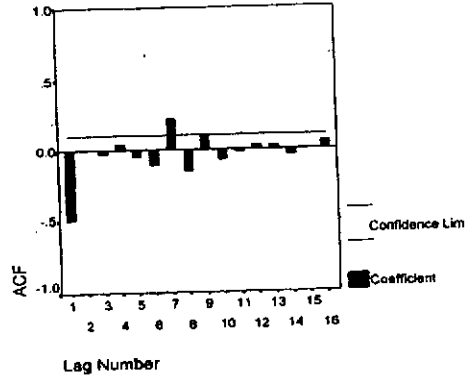
Table-5.1: Name of the Company: Askari Commercial Bank

	ARIMA (0,1,1)		ARIMA (1,1,0)		ARIMA (2,2,0)		
Number of Residuals	315		315		315		
AIC	1335.2238		1335.2237		1344.5		
SBC	1342.728		1342.728		1335.74		
LOG LIKELY-HOOD	-665.611		-665.611		-669.25		
RESIDUAL SS	1262.43		1262.43		1305.36		
Estimates	$\theta_0$	$\theta_1$	$\Phi_0$	$\Phi_1$	$\Phi_0$	$\Phi_1$	$\Phi_2$
	-0.1632	-0.0099	-0.1632	0.01	0.0021	0.1549	0.8395
STANDARD ERROR	0.1139	0.0564	0.1139	0.5049	0.00157	0.055	0.0624
T. RATIO	-1.432	-0.1769	1.432	0.17706	1.323	2.7741	13.43
P. VALUE	0.1529	0.8597	0.1529	0.8595	0.18659	0.00586	0

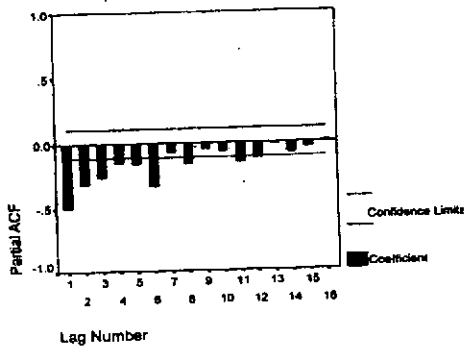
**F-5.1(a)**  
Actual and 2<sup>nd</sup> Differences



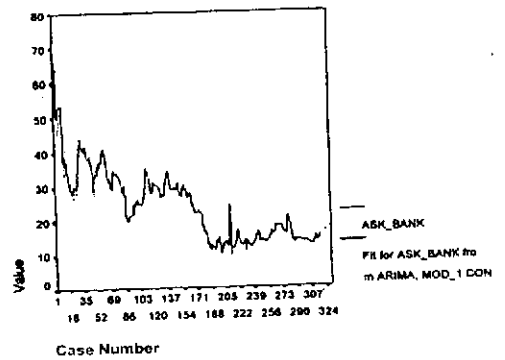
**F-5.1(b)**  
ACF of 2<sup>nd</sup> Differences  
DIFF(ASK\_BANK,2)



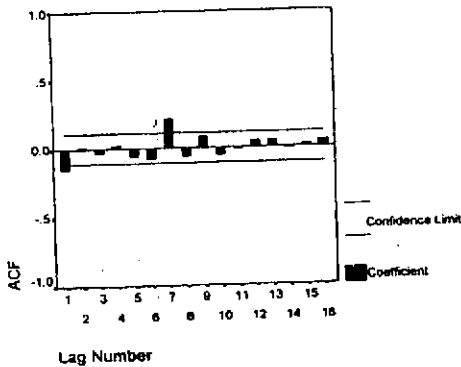
**F-5.1(c)**  
PACF of 2<sup>nd</sup> Differences  
DIFF(ASK\_BANK,2)



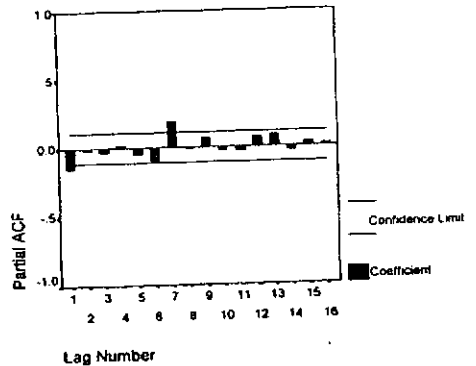
**F-5.1(d)**  
Actual and Predicted Values



**F-5.1(e)**  
Error for ASK\_BANK from ARIMA, MC



**F-5.1(f)**  
Error for ASK\_BANK from ARIMA, MC



## 5.2 Shell Company

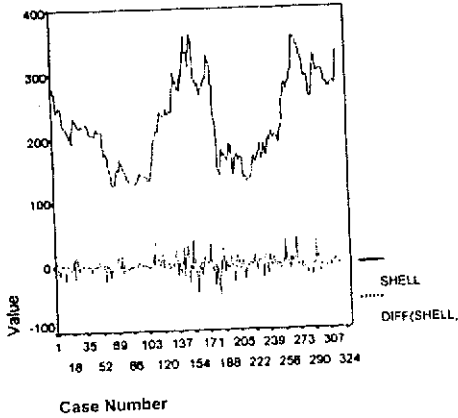
To find which model is appropriate ACF and PACF were calculated of the first differences of the actual data and are shown in Figure 5.2b & 5.2c. This plot suggests ARIMA (0,1,1), however, we have tried ARIMA (1,1,0) and ARIMA (1,1,1) as well. Table 5.2 shows the summary of the results.

The results clearly support ARIMA (0,1,1). The value of AIC and SBC for ARIMA (0,1,1), are smaller than the other two models ARIMA (1,1,0) and ARIMA (1,1,1). The values of log likelyhood and residual SS for ARIMA (1,1,1) are smaller but the estimates of  $\theta_1$  and  $\Phi_1$  are highly insignificant where as for the model ARIMA (0,1,1) the estimates of  $\theta_0$  and  $\theta_1$  are highly significant with smaller standard errors. Therefore, our choice of the model is ARIMA (0,1,1). To see the behavior of the residual, the normal probability plot of the residual together with the plot of the ACF and PACF are shown in Figure 5.2e and 5.2f respectively. The model fitted to the data is  $X_t = 0.2069 - 0.217 \epsilon_{t-1}$

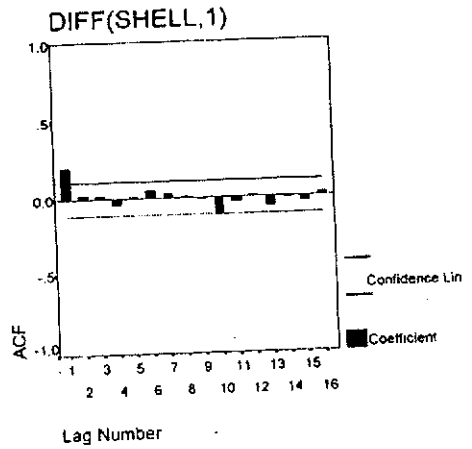
Table-5.2: Name of the Company: Shell

	ARIMA (0,1,1)		ARIMA (1,1,0)		ARIMA (1,1,1)		
Number of Residuals	318		318		318		
AIC	2440.338		2440.371		2442.241		
SBC	2447.862		2447.895		2453.527		
LOG LIKELY-HOOD	-1218.169		-1218.185		-1218.12		
RESIDUAL SS	39561.487		39565.546		39548.816		
Estimates	$\theta_0$	$\theta_1$	$\Phi_0$	$\Phi_1$	$\Phi_0$	$\Phi_1$	$\Phi_2$
	0.2069	-0.127	0.2166	0.2157	0.21169	-0.1135	0.1069
STANDARD ERROR	0.7613	0.0558	0.7969	0.0558	0.7805	0.2635	0.2636
T. RATIO	0.2718	-3.7026	0.2718	3.862	0.2712	-0.43	0.4058
P. VALUE	0.6858	0.0001	0.7859	0.0001	0.7863	0.6668	0.6851

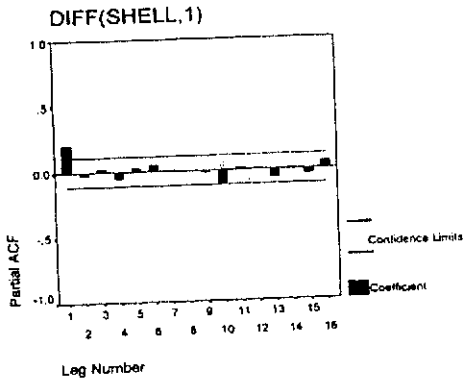
F-5.2(a)



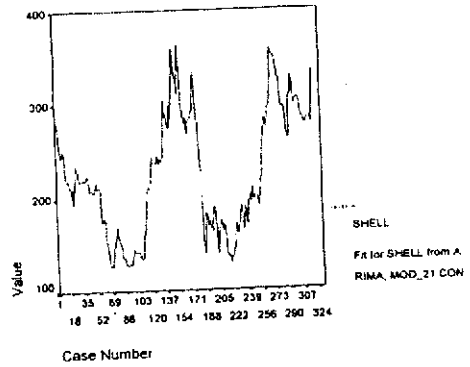
F-5.2(b)



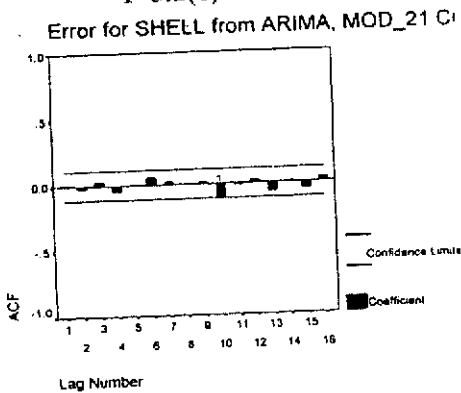
F-5.2(c)



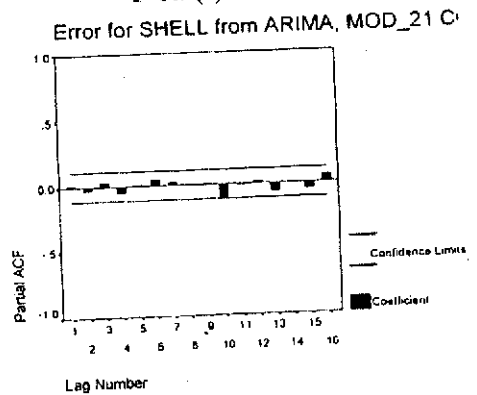
F-5.2(d)



F-5.2(e)



F-5.2(f)





### 5.3 Sui Northern Gas Company

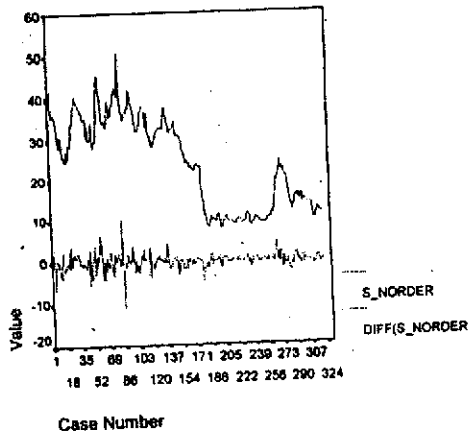
The plots of the ACF and PACF of the first differences suggests ARIMA (2,1,2) model. Table-5.3 shows the summary of the results. The values of AIC, SBC and residual SS are slightly greater for ARIMA (2,1,2) but the value of the log likelihood is smaller. At the same time the estimates of  $\theta_1$ ,  $\theta_2$ ,  $\Phi_1$ , and  $\Phi_2$  are highly significant with smaller value of standard error. To see the behavior of residuals the normal probability plot of residual together with the plot of ACF and PACF are shown in Figure 5.2e and 5.2f respectively. Their figure confirms that the residual are independently and normally distributed. The model fitted to the data is

$$X_t = -0.095 + 0.725 \epsilon_{t-1} - 0.965 \epsilon_{t-2} + 0.703 Y_{t-1} - 0.931 Y_{t-2}$$

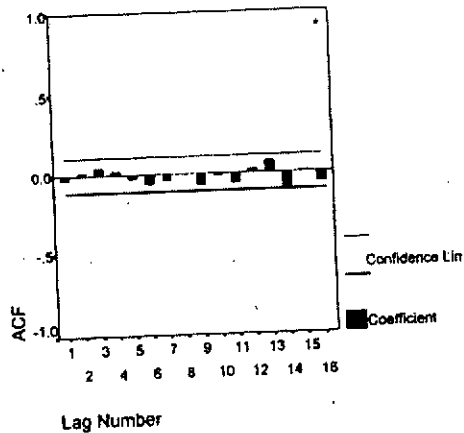
Table-5.3: Name of the Company: Sui Northern Gas

	ARIMA (0,1,1)		ARIMA (1,1,0)		ARIMA (2,1,2)			
Number of Residuals	319		319		319			
AIC	1272.84		1272.836		1276.219			
SBC	1280.87		1280.366		-1295.045			
LOG LIKELYHOOD	-634.92		-634.418		-633.109			
RESIDUAL SS	997.11		997.097		998.889			
Estimates	$\theta_0$	$\theta_1$	$\Phi_0$	$\Phi_1$	$\Phi_0$	$\theta_1$	$\Phi_1$	$\Phi_2$
	-0.953	0.0213	-0.0953	-0.02201	-0.0951	0.7258	0.7037	-0.9316
STANDARD ERROR	0.0969	0.0512	0.0969	0.0561	**0.0999	0.0483	0.0649	0.0645
T. RATIO	-0.983	0.3798	0.9837	-0.3921	-0.9512	-15.0073	10.8315	-14.4263
P. VALUE	0.3261	0.7043	0.3259	0.6951	0.3422	0	0	0

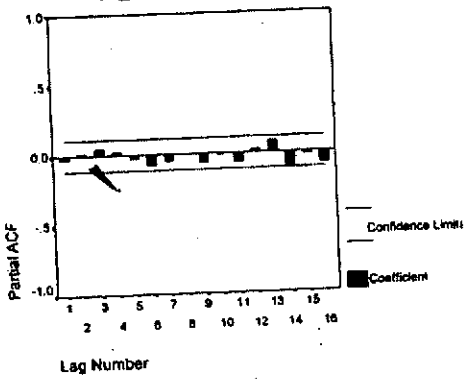
F-5.3(a)



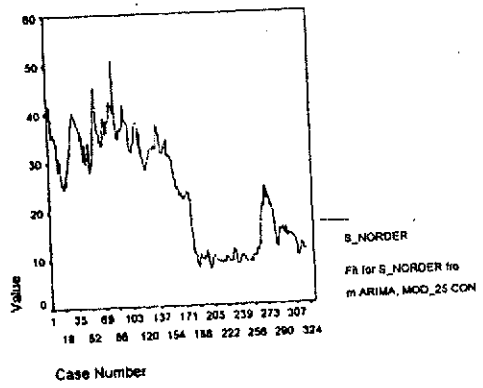
F-5.3(b)  
DIFF(S\_NORDER,1)



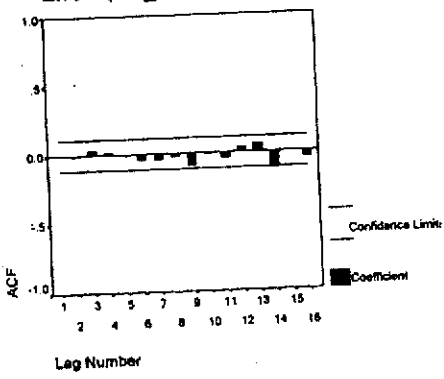
F-5.3(c)  
DIFF(S\_NORDER,1)



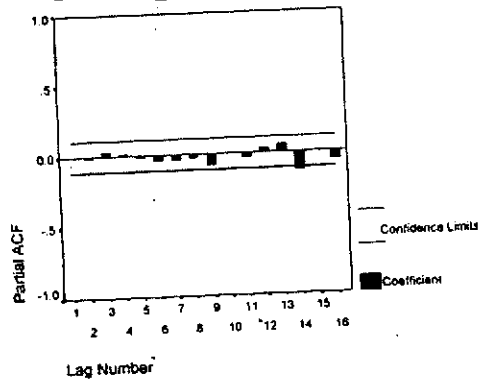
F-5.3(d)



F-5.3(e)  
Error for S\_NORDER from ARIMA, MOD\_



F-5.3(f)  
Error for S\_NORDER from ARIMA, MOD\_



#### 5.4 P.I.A.

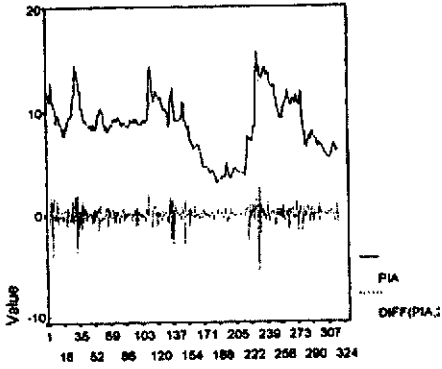
The plot of actual data shows that the series was not stationary. First differences were calculated to satisfy the condition of stationarity. The plot of first differences shows that the data has not constant variance and there is a parabolic trend. Second differences were calculated and the condition of stationarity was obtained. Figure 10-B and 10C shows the ACF and PACF of the second differences. This suggests ARIMA (2,2,0) as the best model. The results are summarized in Table-10. The study of the table reveals that the value of AIC, SBC and log likelihood are smaller in ARIMA (0, 1, 1) and ARIMA (1, 1, 0) but estimates of  $\phi_0$ ,  $\phi_1$  and  $\phi_2$ , are highly significant with smaller standard error. The behavior of the residual is observed by the normal probability plot of residual together with the plot of ACF and PACF of the residual, which are shown in figure 10-Em10-F and 10-G respectively. Their figure confirms that the residual are independently and normally distributed. The model fitted to the data is

$$X_t = 0.9741 + 0.00 Y_{t-1} + 0.0001 Y_{t-2}$$

Table-5.4: Name of the Company: P.I.A.

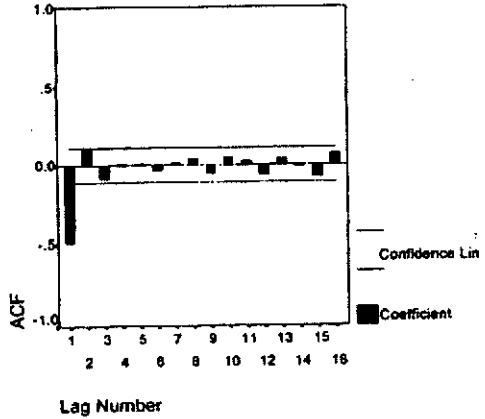
	ARIMA (0,1,1)		ARIMA (1,1,0)		ARIMA (2,2,0)		
Number of Residuals	317		317		317		
AIC	642.08178		641.55357		721.62173		
SBC	649.59958		649.07137		732.888		
LOG LIKELY-HOOD	-319.04089		-318.77678		-357.81086		
RESIDUAL SS	138.91935		138.6881		178.12391		
Estimates	$\theta_0$	$\theta_1$	$\Phi_0$	$\Phi_1$	$\Phi_0$	$\Phi_1$	$\Phi_2$
	0.01981	0.07577	0.01987	0.09408	0.00103	-0.63502	-0.21564
STANDARD ERROR	0.04	0.05823	0.041	0.05747	0.02288	0.05756	0.0551
T. RATIO	0.49512	1.3012	0.48453	1.63746	0.04493	-11.0312	-3.91345
P. VALUE	0.62086	0.19414	0.62835	0.10263	0.96419	0	0.00011

F-5.4(a)



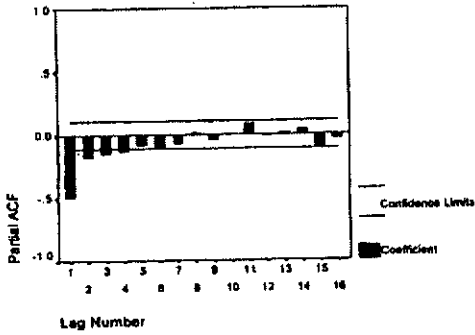
F-5.4(b)

DIFF(PIA,2)

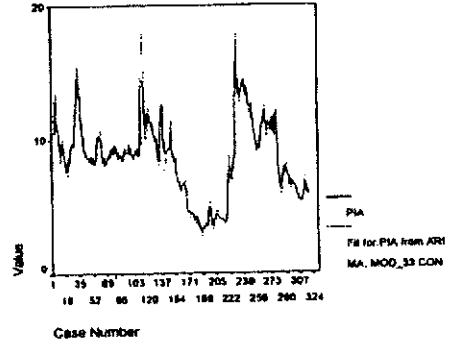


F-5.4(c)

DIFF(PIA,2)

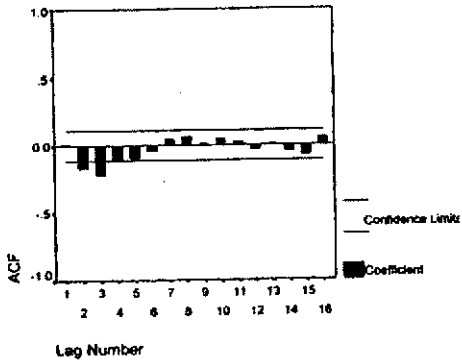


F-5.4(d)

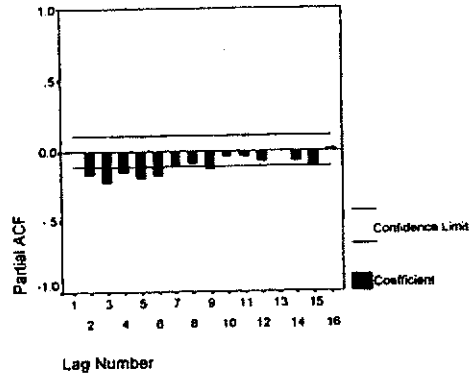


F-5.4(e)

Error for PIA from ARIMA, MOD\_33 CO



F-5.4(f)  
Error for PIA from ARIMA, MOD\_33 CO



## 5. SUMMARY AND CONCLUSION

The basic objective of the study is to forecast the data under consideration by fitting the appropriate model. In any natural phenomenon having statements by business conditions forecasters, one is quickly struck by the probabilistic nature of these projects. Rather than making absolute assertions concerning the likely course of GDP, corporate profits, the probability of selected industries, the prime state, domestic auto-mobile sales, or the labor costs associated with a strike, the projections are usually hedged in its references to alternative scenarios that are possible but less likely than the business forecasters, best guess. Thus even though economic forecasters may argue that the most likely outcome for the coming years is 3% real GDP growth and 4% inflation, they will admit that there is an outside chance of 1% real growth and 6% inflation. Whether they are in the political or business sector, all forecasters produce contingency projections based on a subjective probability assessment of likely exogenous assumptions. Real process cannot be observed because many different factors are involved, in the case of stock exchange price. Many economic indicators are affecting the price of share but with the help of model, the real generating Mechanism can be approximated.

For the present study, data is collected for different companies, for different groups listed at KSE. Karachi Stock Exchange being the premiere stock company and the largest in terms of stock listed and volume of trading witnessed every day. KSE is regarded as an economic HUB of Pakistan where industrial and commercial share of different registered companies are sold freely. The long term and short-term movements of investment at KSE is the key indicator of economic health of Pakistan. In this sense, stock exchange provides the long-term investment and in commerce and industry but also provide the finance and capital to the Government for the completion of its mega projects.

Four companies are selected from different sectors on the basis of maximum turn over. Those companies have been selected which have possessed dynamic change and maximum volume of trading and alive/active at the platform of stock exchange. The weekly average share prices of these companies are considered during the time period January 1995-January 2001. Data have been collected from the publication of State Bank of Pakistan daily bulletin of the share prices and calculated average of these share prices for every financial year by the SBP.

Since short term forecasting is required, therefore, daily data is selected and converted into weekly average data. Short-term forecast are required for the

period of 10 weeks ahead, so the knowledge about the trend of the data is essential for forecasters. The use of trend curves depends upon the assumption that the most probable course in the future is a projection of that which prevailed in the past. The first requirement, therefore, is to determine whether they lie on a smooth curve, which can be defined mathematically. If so, the curve then be projected into the future consumption. If something goes different or wrong in terms of changes then forecasts have to be revised. Therefore, no forecast can be taken as final instead they must be revised by time again and new inputs must be used for fresh forecast.

In order to study the data a comprehensive variety of models are fitted to the data. Appropriate model for the data is selected on the basis of many considerations which are as follows:-

1. The ACF and PACF of the first differences, second differences and log transformation are taken which suggests an appropriate model.
2. Akaike Information criterion (AIC) and Schwarz Bayesian criterion (SBC) are applied on the model. Those models are suggested which have minimum value of AIC and SBC.
3. Log likelihood and residual sum of the square are also computed for the model and got the direction for the selection of an appropriate model.
4. Parameters are tested either they are significant or insignificant by getting the P-value. It is also compared with the value of standard error.
5. Also model adequacy of the selected model, which determine through the checking of normality assumption plotting residual on normal probability plots and ACF and PACF of residual plots. It becomes evident, that is independently and normally distributed.

In Pakistan, information about the performance of any company for the general public is available only on two occasions; half yearly (unaudited accounts and year end (Audited accounts). On another, side, in advance countries this information is available on quarterly basis. Such availability of data and ingenuity of analyst is obviously in dearth in this part of the world. We have more of spectaculars and punters than competent analysts. So cut of period for dissemination of information is necessary for drawing conclusions.

#### ACKNOWLEDGEMENT

We are thankful to Prof. Dr. Muhammad Akram, Department of Business Studies, University of the Bahrain and Prof. Dr. A.Z. Mamon, Dean, Pak-AIMS, Lahore for their guidance in connection with this research.