**FINANCIAL RISK ANALYTICS & MANAGEMENT (FRAM)**

**REPORT**

**ON  
TIME SERIES ANALYSIS OF STOCK PRICES OF JSW ENERGY LIMITED**

****

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI  
HYDERABAD CAMPUS**

**(APRIL 2019)**



**JSW ENERGY LIMITED**

**An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model, GARCH Model**

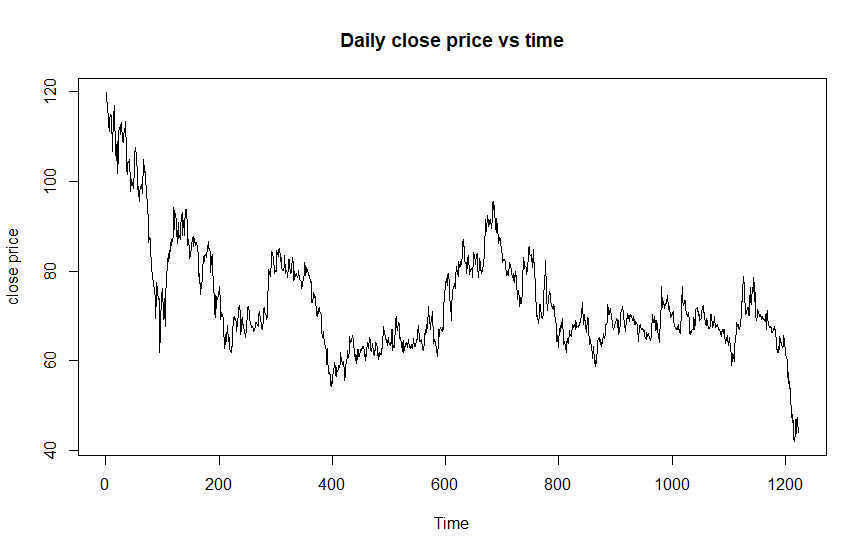
**1st April 2015-31st March 2020**

**SAMPLE DATA AND THEIR PLOTS:**

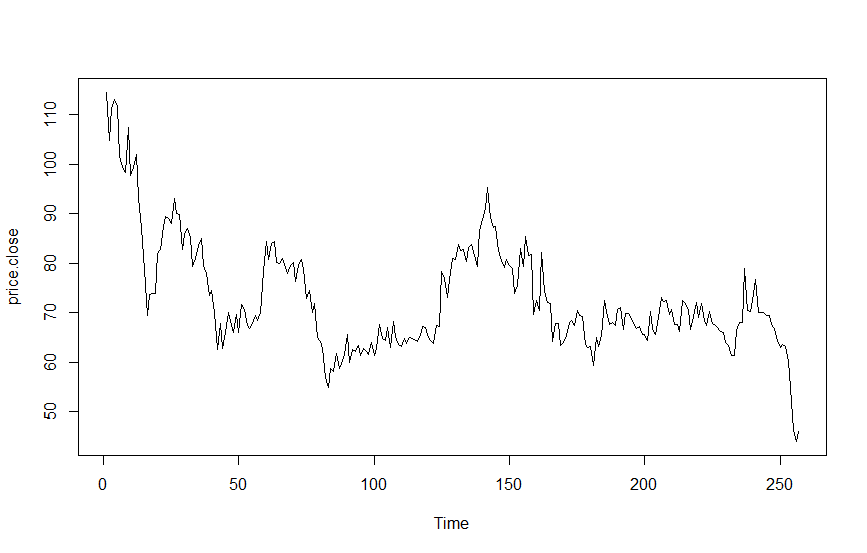
Closing stock prices (daily, weekly, monthly) for last 5 years was collected and the time series data was plotted.

Date:1st April 2015-31st march 2020

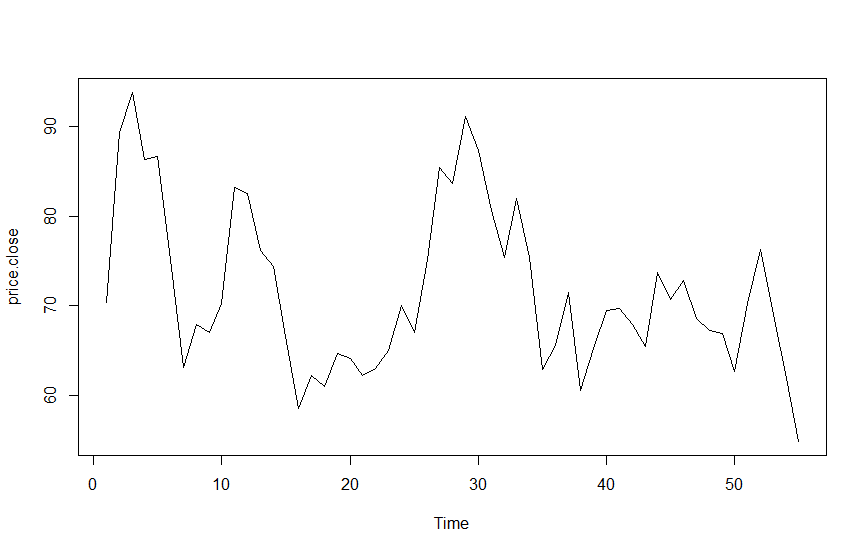
**DAILY:**



**WEEKLY:**

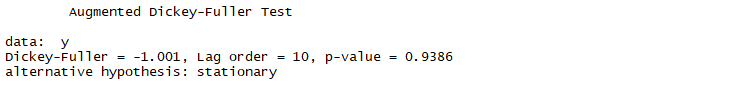


**MONTHLY:**

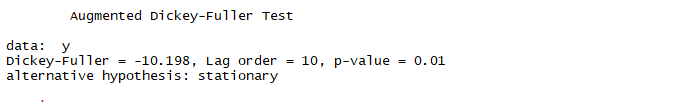


From the plot we can observe that there is a downward trend in close price of the stock, no seasonality, no obvious outliers.  For an ACF to make sense, the series must be a weakly stationary series. This means that the autocorrelation for any particular lag should be same regardless of where we are in time. Mean and variance should be same for all time, which is not in the above case. So, to make it stationary we have to do the differencing i.e take the daily, weekly, monthly returns in place of closing prices. We can also see how differencing significantly improved the p-value:

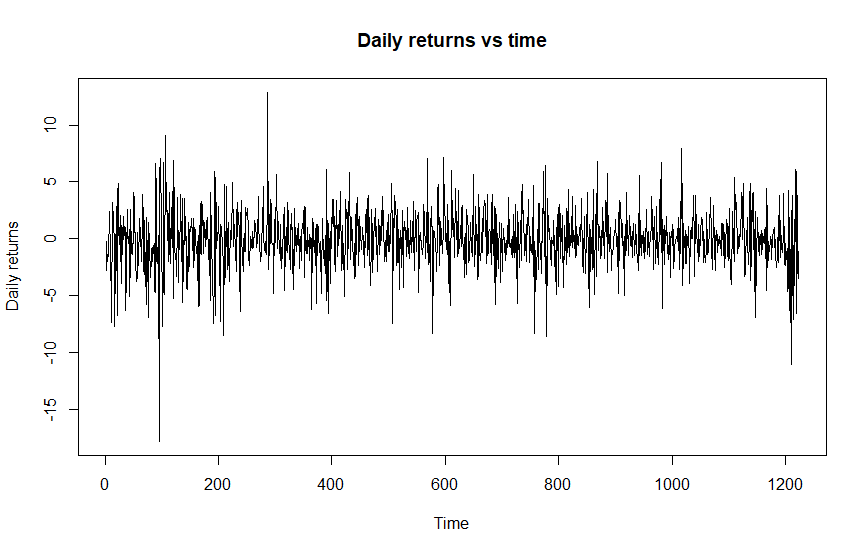
**BEFORE:**

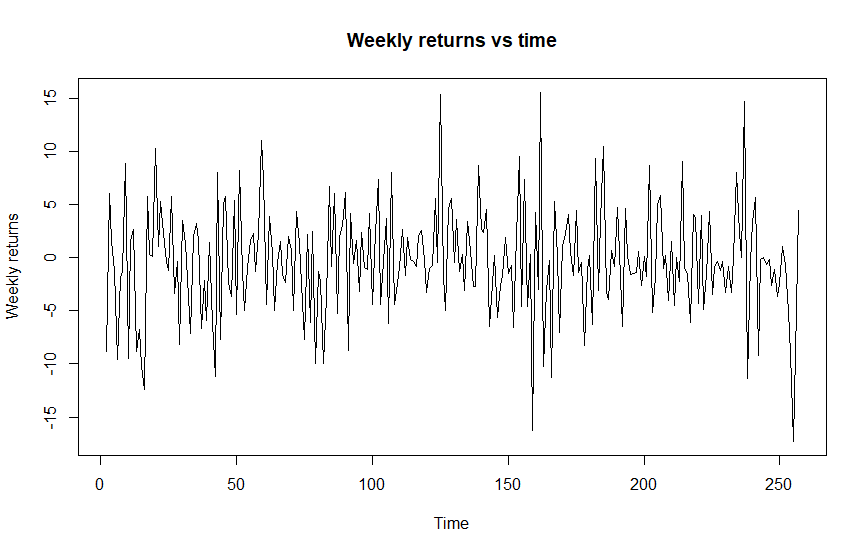


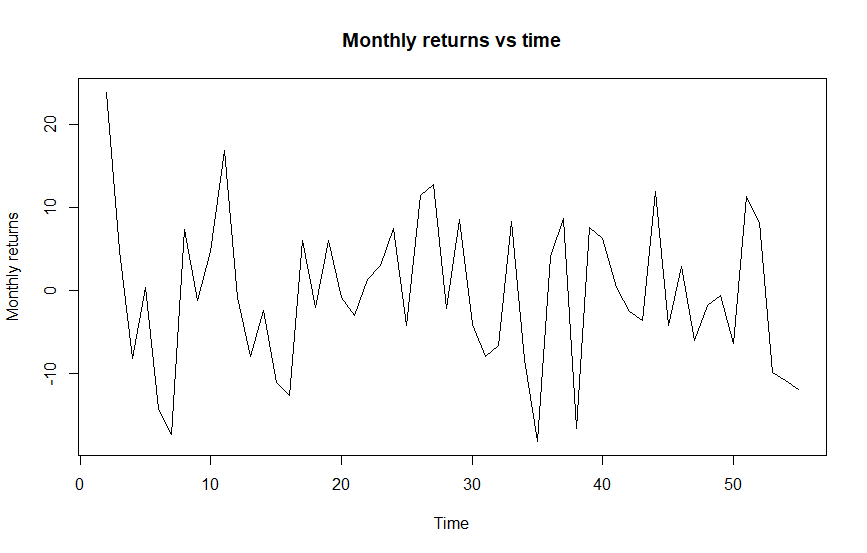
**AFTER:**



**RETURN PLOTS ARE AS FOLLOWS:**



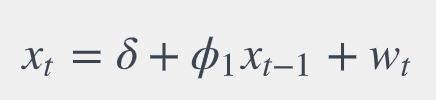




Time series models known as ARIMA models may include autoregressive terms (AR) and/or moving average (MA) terms.

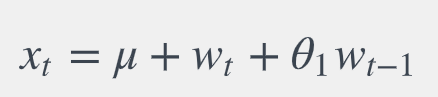
An **autoregressive (AR)** term in a time series model for the variable xtis a lagged value of xt.

AR (1) model: A lag 1 autoregressive term is xt−1(multiplied by a coefficient).



A **moving average (MA)** term in a time series model is a past error (multiplied by a coefficient).

MA (1):



**ACF** is an (complete) auto-correlation function which gives us values of auto-correlation of any series with its lagged values. For a positive value of ϕ1, the ACF exponentially decreases to 0 as the lag h increases. For negative ϕ1, the ACF also exponentially decays to 0 as the lag increases, but the algebraic signs for the autocorrelations alternate between positive and negative.

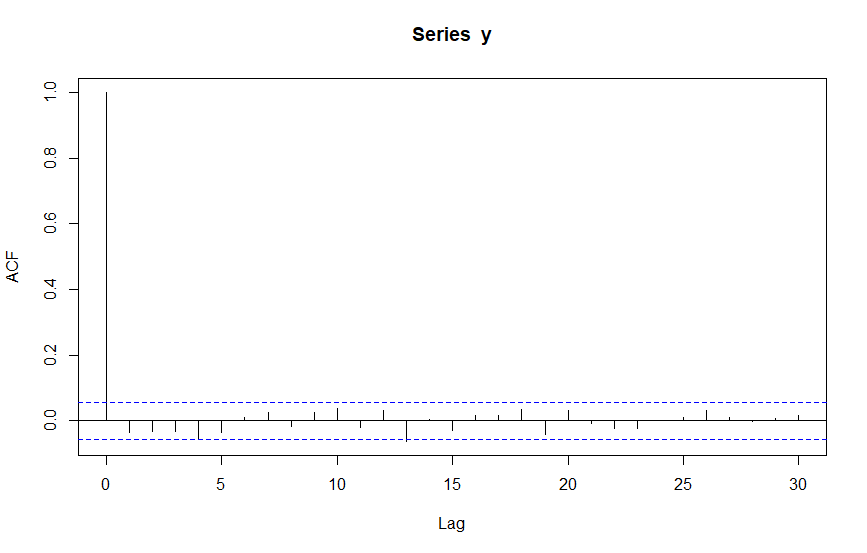
A partial correlation (**PACF**) is a conditional correlation. It is the correlation between two variables under the assumption that we know and take into account the values of some other set of variables.

**SIGNIFICANCE AND INTERPRETATION OF ACF AND PACF:**

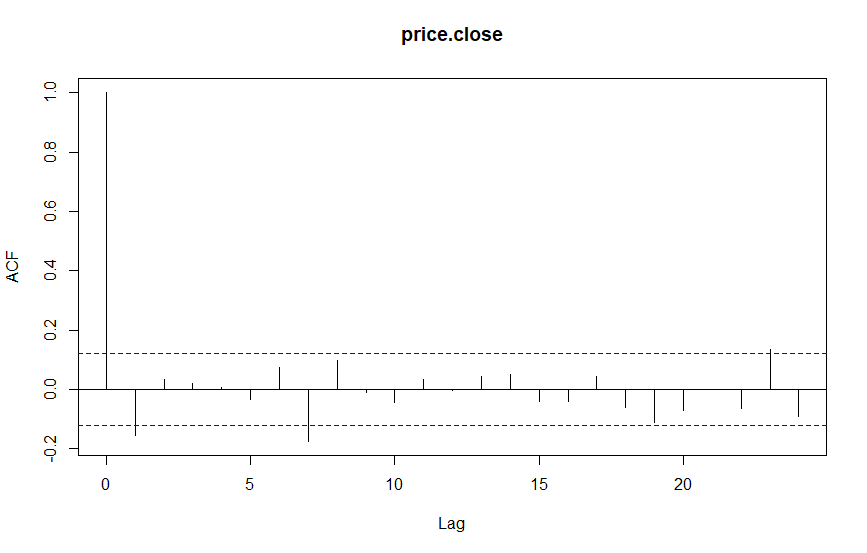
**Identification of an MA model is often best done with the ACF rather than the PACF. A sample ACF with a significant autocorrelation at lag1 is an indicator of a possible MA(1) model. Similarly, a sample ACF with significant autocorrelations at lag 1 and 2, but non-significant autocorrelation for higher lags indicates a possible MA(2) model. In this way we can identify a MA(q) model.**

Following are the ACF plots for our sample data:

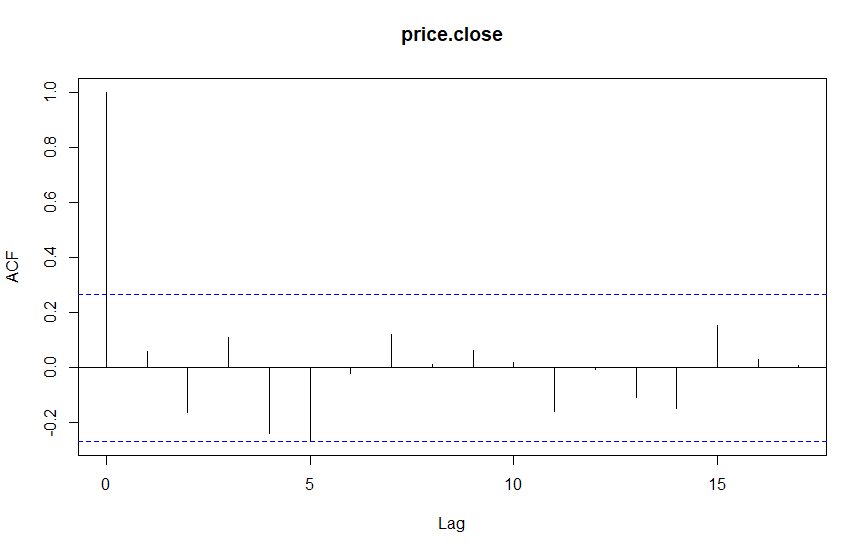
**DAILY ACF:**



**WEEKLY ACF:**



**MONTHLY ACF:**



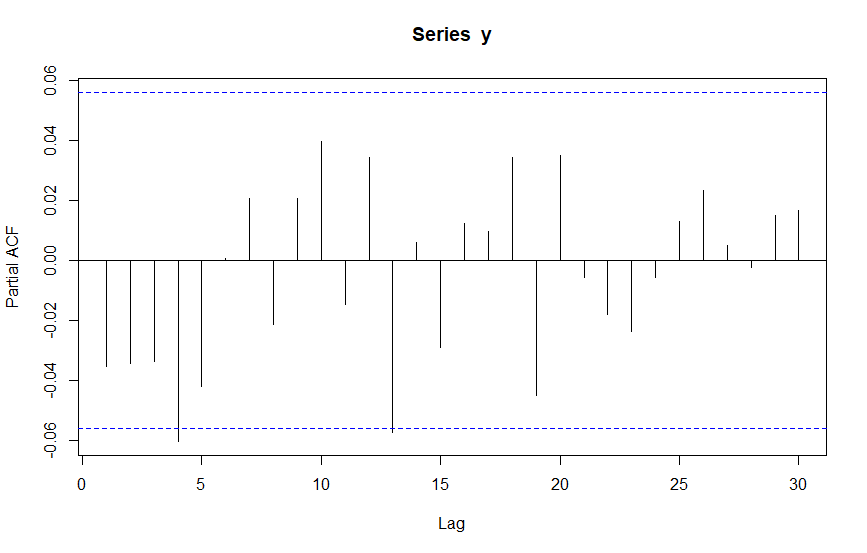
**Distinguishing AR and MA terms from simultaneously exploring an ACF and PACF-**

**From the above daily ACF plot we can see that at lag1 there is no significant spike. So, the order of MA is expected to be 0. But in weekly ACF plot we can see a significant value at lag 1 after which there is no significant value. So, order of MA() for weekly is taken as 1. Similarly for monthly ACF there is no significant spike at lag1 so it is a MA(0) model.**

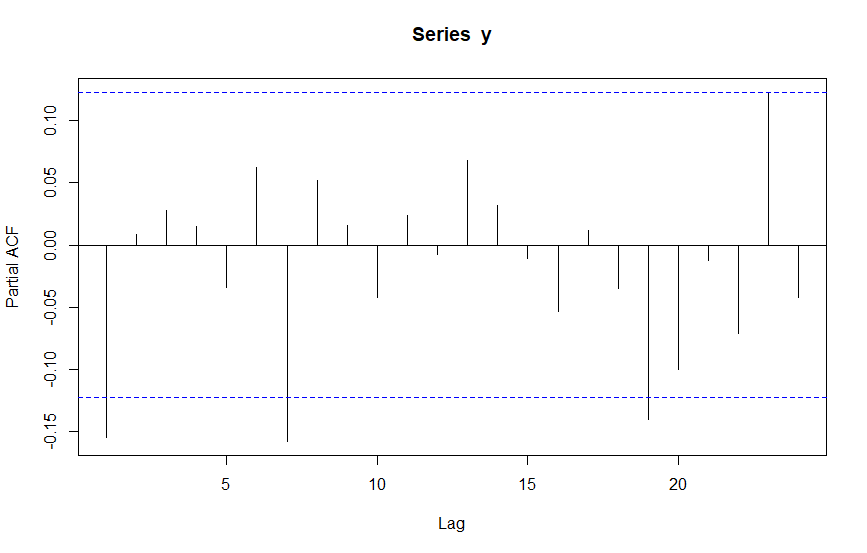
**Identification of an AR model is often best done with the PACF.** For an AR model, the theoretical PACF “shuts off” past the order of the model. If there is a sudden significant change in PACF at lag1 it may be a AR(1) model. Similarly for AR(2), AR(3) and so on..

Following are the ACF plots for our sample data:

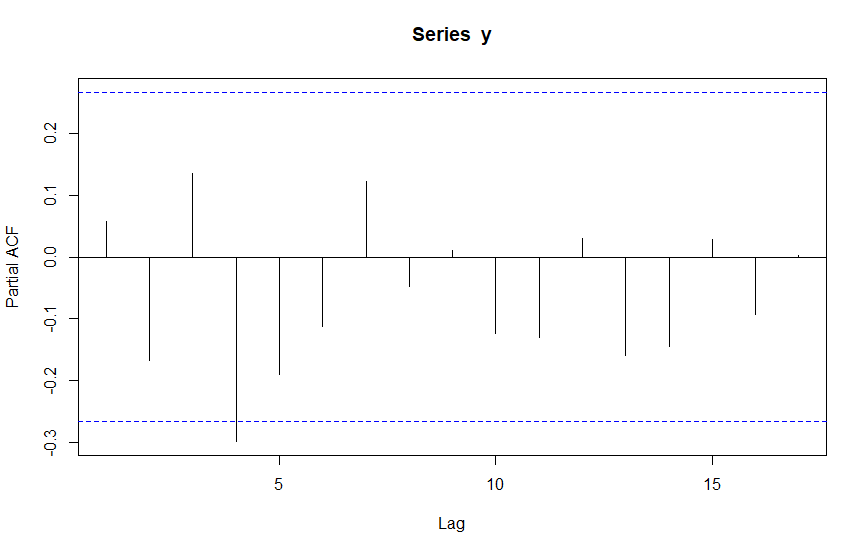
**DAILY PACF:**



**WEEKLY PACF:**



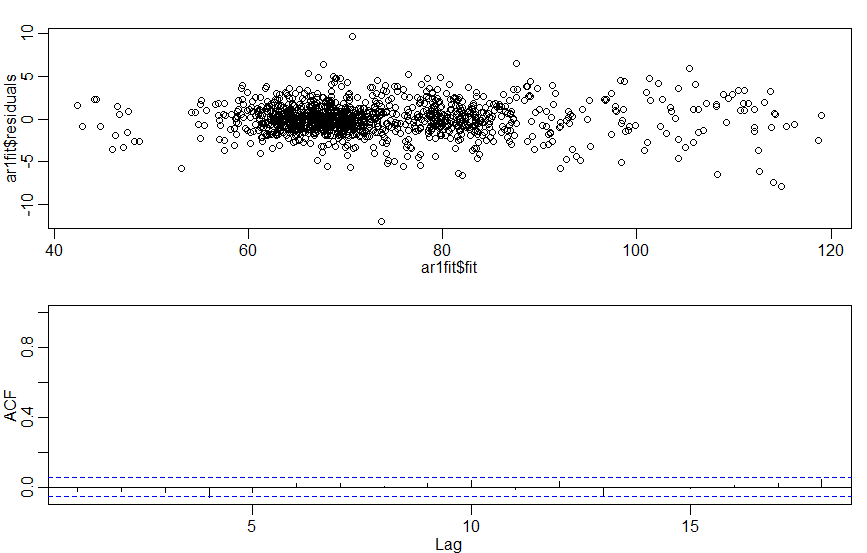
**MONTHLY PACF:**



The above PACF plots tell us about the order of AR model. We can see that in daily PACF plot there is no significant rise in lag1. So, the order of AR is 0. For weekly PACF, there is a significant value at lag 1. So, order of AR is taken as 1. And for monthly PACF due to absence of any significant value its order is also taken as o.

The above analysis can be cross checked by using a R command: “auto.arima()”, which gives the order of AR, MA with minimum error.

**Residual autocorrelations are zero:**



Blue lines in the plot indicates bounds for statistical significance. For ACF of residuals to be good, there should not be any significant terms. Theoretically all residuals autocorrelations should be zero. From the above plot we can see that the residual ACF is negligible which says that it is a good fit. The small value of ACF is due to sampling error.

**NON-SEASONAL ARIMA MODEL ANALYSIS:**

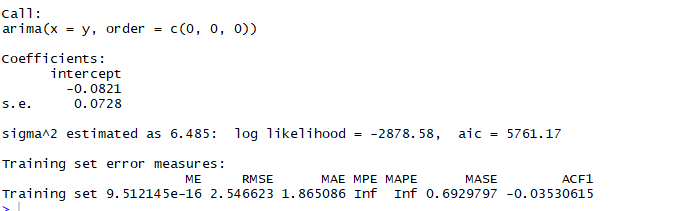
ARIMA models, also called Box-Jenkins models, are models that may possibly include autoregressive terms, moving average terms, and differencing operations. the elements in the model are specified in the order (AR order, differencing, MA order). The differencing order refers to successive first differences.

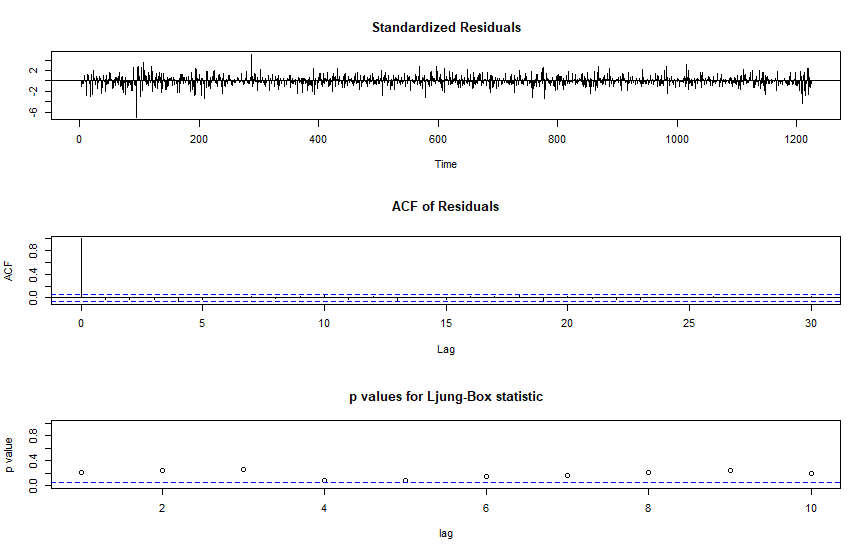
Order of AR and MA are taken from PACF and ACF plot. A model with less RMSE (Root Mean Square Error) and/or less AIC value is a good fit to the sample data. Also, we need to analyze the residual plots to check any significant correlation.

Below are some of the ARIMA models with their respective residual plots:

**DAILY:**

**ARIMA (0,0,0)**

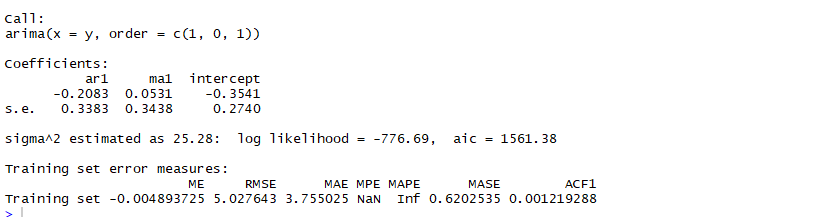


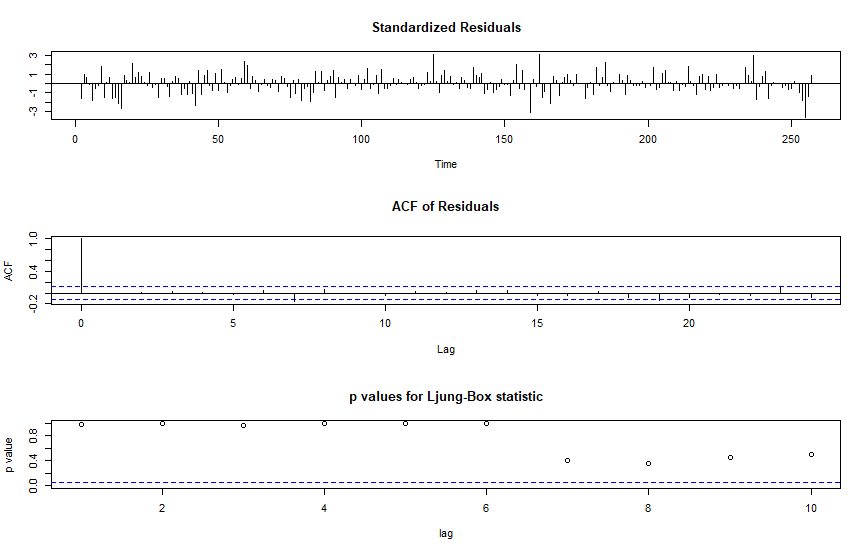


From our previous ACF and PACF analysis we have order of both AR and MA=0, also the differencing order=0. So, we simulated ARIMA (0,0,0) model and found that RMSE value comes out to be minimum =2.5466 and AIC=5761.17. Also, the residual values for ACF and PACF are not significant. So, we accept this model.

**WEEKLY:**

**ARIMA (1,0,1)**

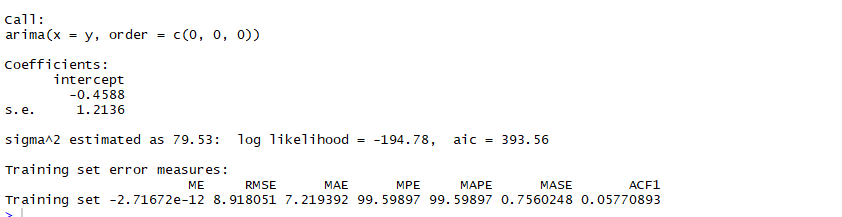


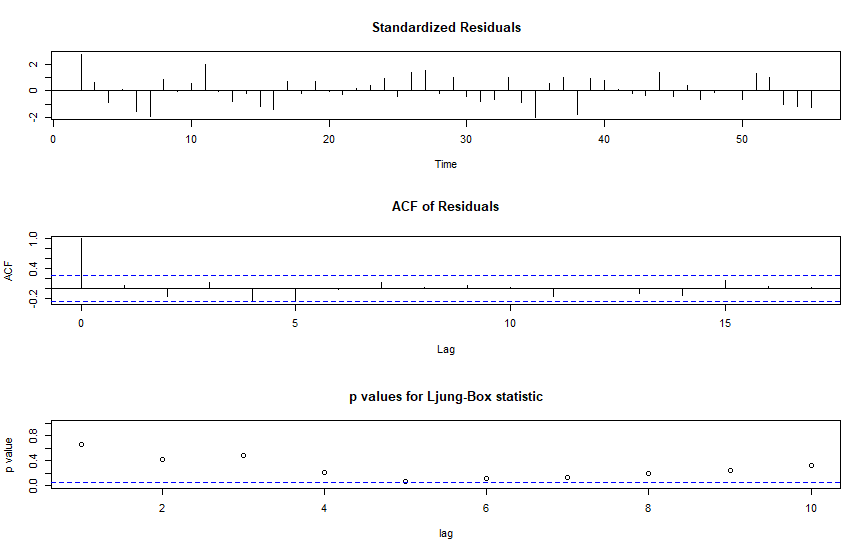


In weekly data we can see a change in the ARIMA model. ARIMA (1,0,1) was found to be more fit from the RMSE and AIC values as it had low value as compared to ARIMA (0,0,0).

**MONTHLY:**

**ARIMA (0,0,0)**





For monthly ARIMA (0,0,0) was found to be fit with minimum error terms.

**ARIMA TO INFINITE ORDER MR MODEL:**

For daily and monthly data since the order of AR and MA are 0 in ARIMA models, so their ARIMA to infinite MR is not possible. The coefficients for MA(∞) will be zero. But in weekly data we have AR and MA components with following coefficients:

ar=-0.2083, ma=0.0531

Weekly:

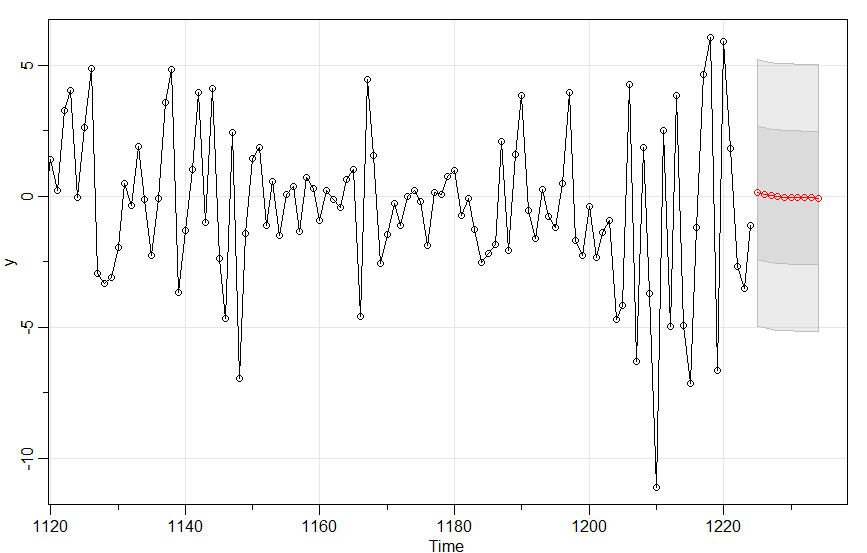


**FORECAST USING ARIMA MODEL WITH THEIR PREDICTION INTERVALS:**

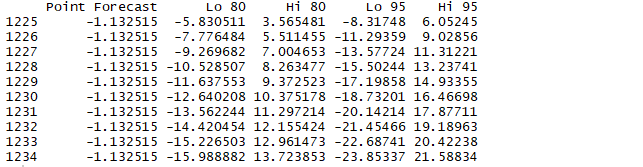
**Prediction interval is the range that likely contains the value of the dependent variable for a single new observation given specific values of the independent variables.**

Using “sarima.for()” command in R we did the following forecast for the next 10 days.10 weeks, 10 months.

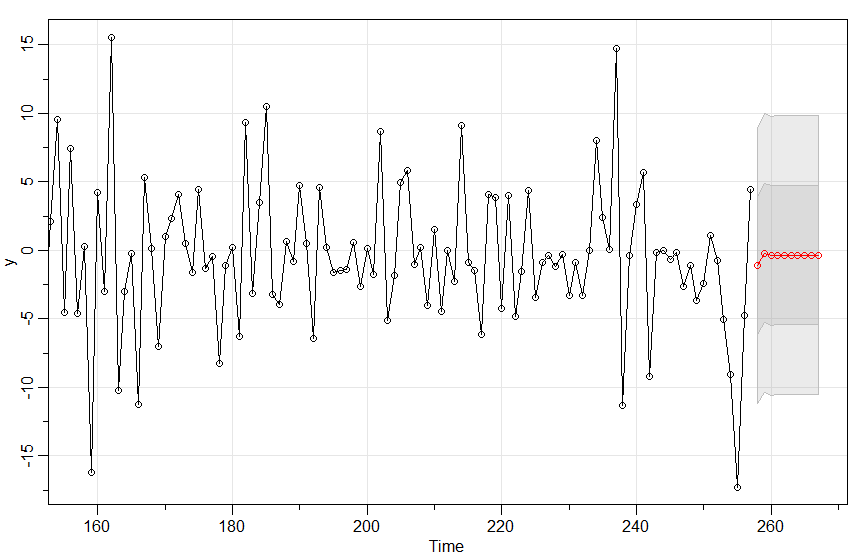
**DAILY:**



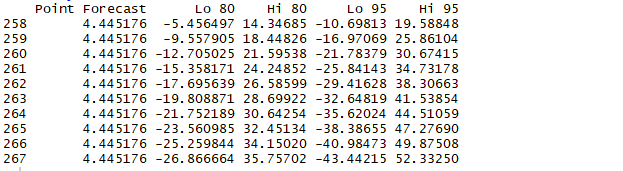
From the forecast we can see that there is a negative return trend for next 10 days. The above model was run with standard error (80% and 95%) given in below prediction interval:



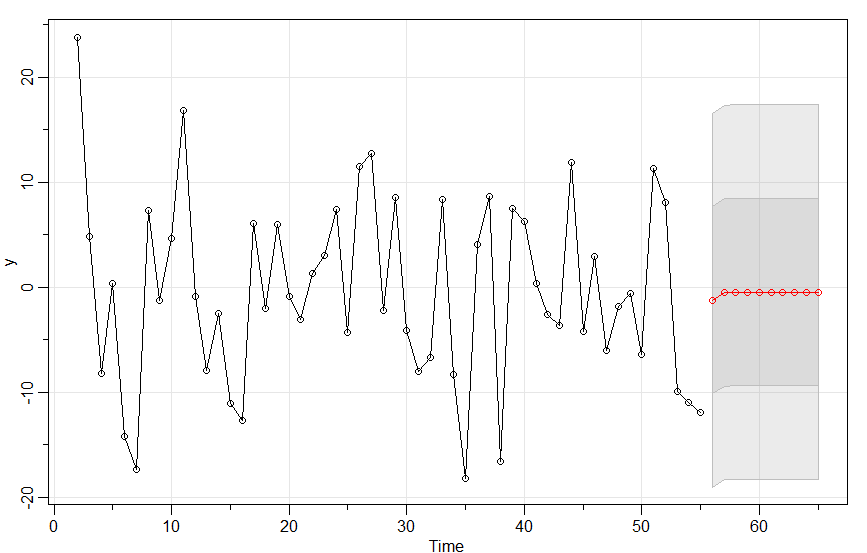
**WEEKLY:**



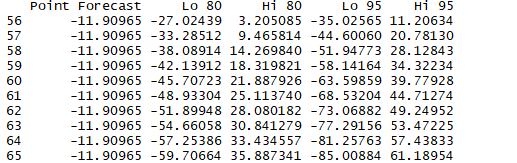
From the summary of forecast we found that the predicted returns for the next 10 weeks is in negative trend. The above model was run with standard error (80% and 95%) given in below prediction interval:



**MONTHLY:**



Also, for the monthly data the forecast results in a negative return trend for next 10 months. The above model was run with standard error (80% and 95%) given in below prediction interval:



**ARCH AND GARCH MODEL ANALYSIS:**

The autoregressive conditional heteroscedasticity (ARCH) model is a [statistical model](https://en.wikipedia.org/wiki/Statistical_model) for [time series](https://en.wikipedia.org/wiki/Time_series) data that describes the [variance](https://en.wikipedia.org/wiki/Variance) of the current [error term](https://en.wikipedia.org/wiki/Errors_and_residuals_in_statistics) or [innovation](https://en.wikipedia.org/wiki/Innovation_(signal_processing)) as a function of the actual sizes of the previous time periods' error terms; often the variance is related to the squares of the previous [innovations](https://en.wikipedia.org/wiki/Innovation_(signal_processing)). The ARCH model is appropriate when the error variance in a time series follows an [autoregressive](https://en.wikipedia.org/wiki/Autoregressive) (AR) model; if an [autoregressive moving average](https://en.wikipedia.org/wiki/Autoregressive_moving_average_model) (ARMA) model is assumed for the error variance, the model is a generalized autoregressive conditional heteroskedasticity (GARCH) model. ARCH models are commonly employed in modelling [financial](https://en.wikipedia.org/wiki/Mathematical_finance) [time series](https://en.wikipedia.org/wiki/Time_series) that exhibit time-varying [volatility](https://en.wikipedia.org/wiki/Volatility_(finance)) and [volatility clustering](https://en.wikipedia.org/wiki/Volatility_clustering), i.e. periods of swings interspersed with periods of relative calm. For larger number of terms, ARCH models are often not significant or the constraints on parameters are not satisfied. So, GARCH comes into picture.

**Estimation of ARCH model:**

We need to check-

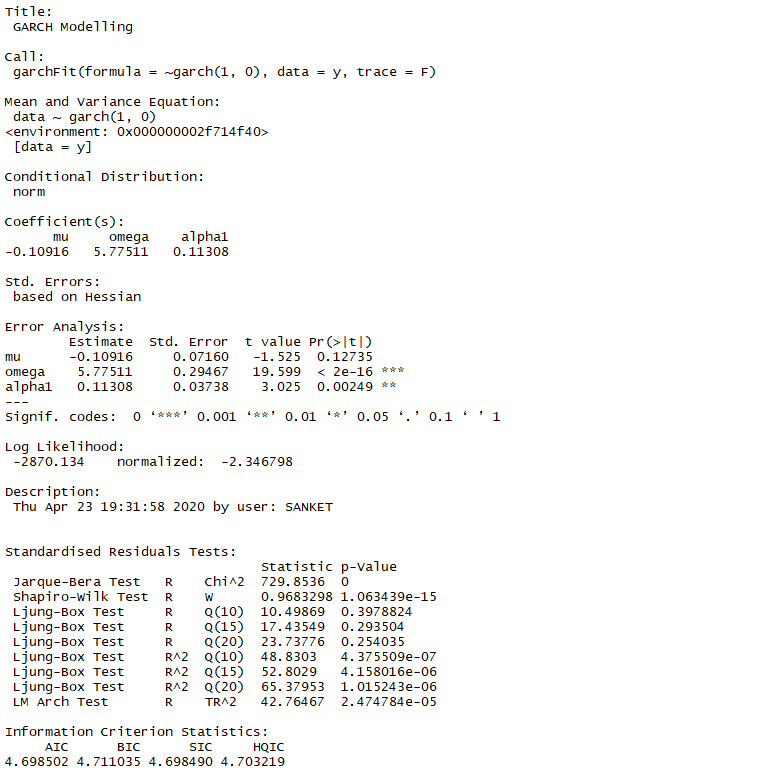
* ACF of standardized residuals
* ACF of squared standardized residuals
* Summary with tests about standardized residuals and their squares

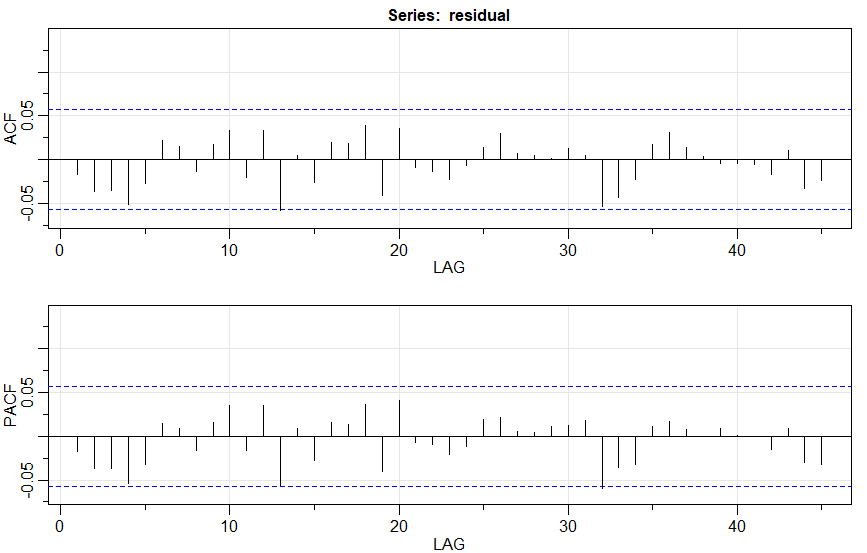
We ran simulation for the above two models and got the following results:

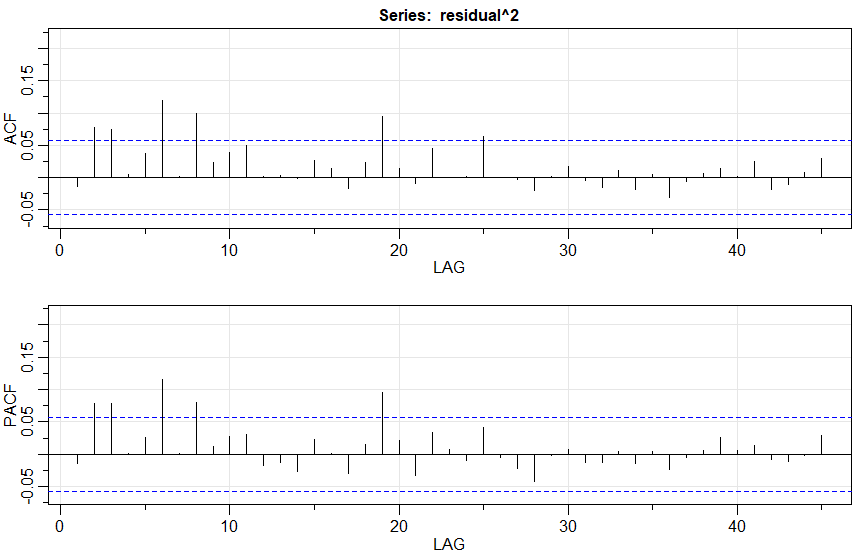
Below are the attached summary of simulation and plots of ACF and PACF for residuals to analyze which model is better

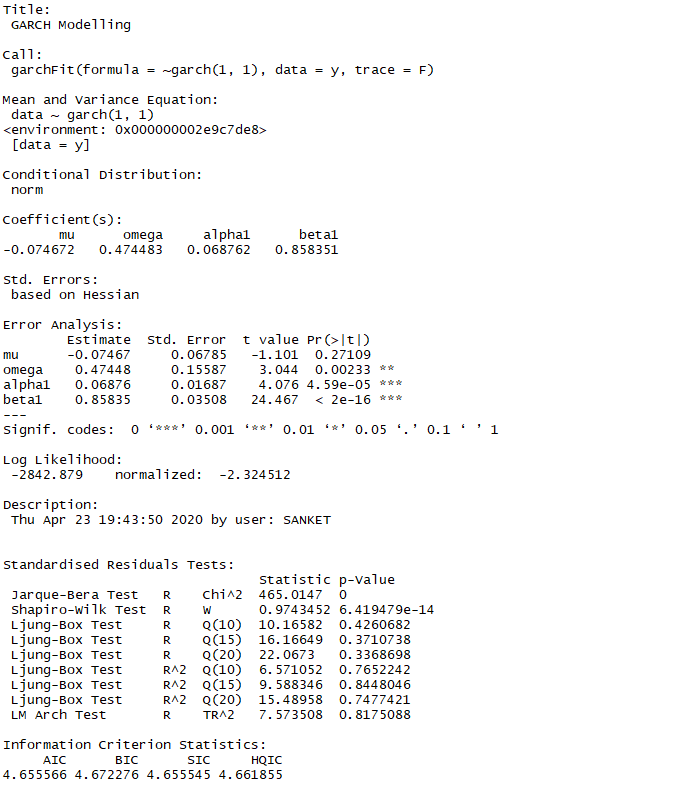
**DAILY:**

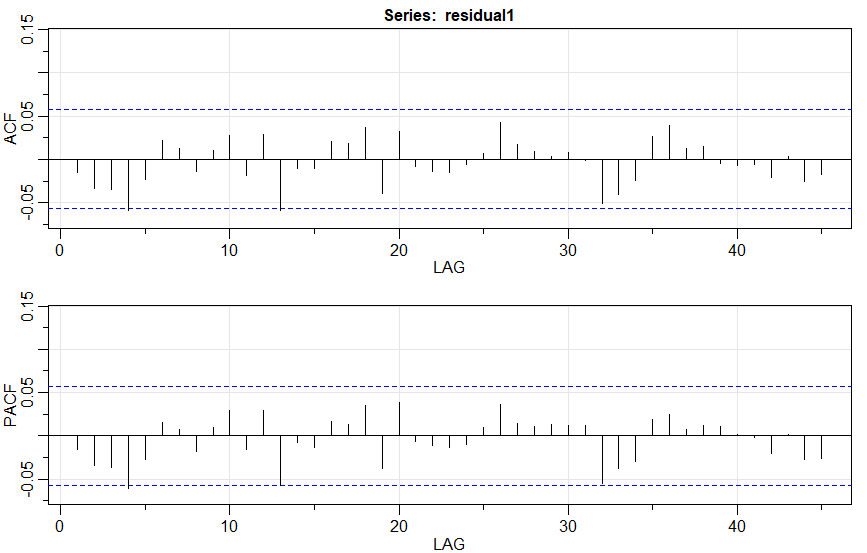
**Simulation of GARCH(1,0) and GARCH(1,1)**

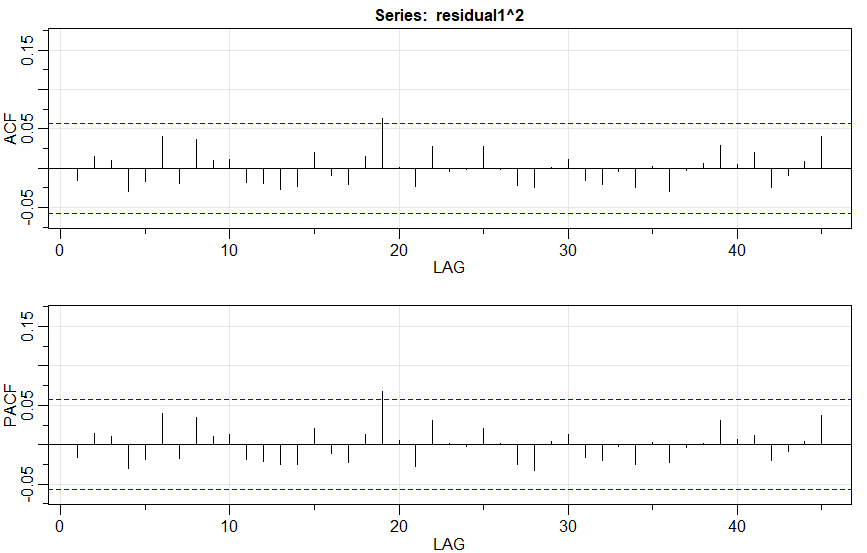


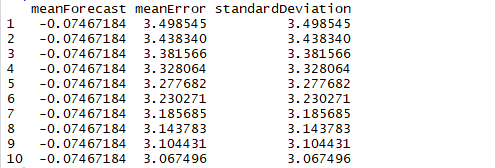








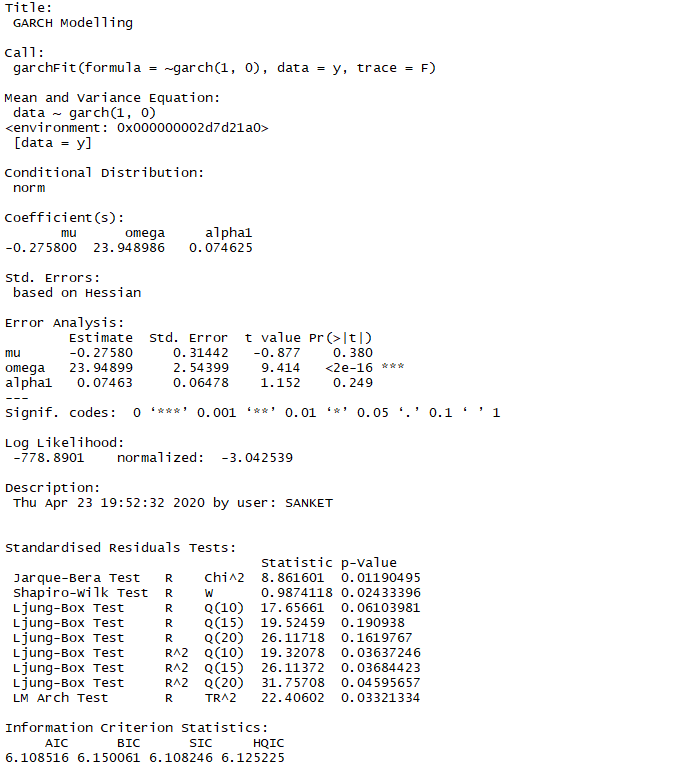


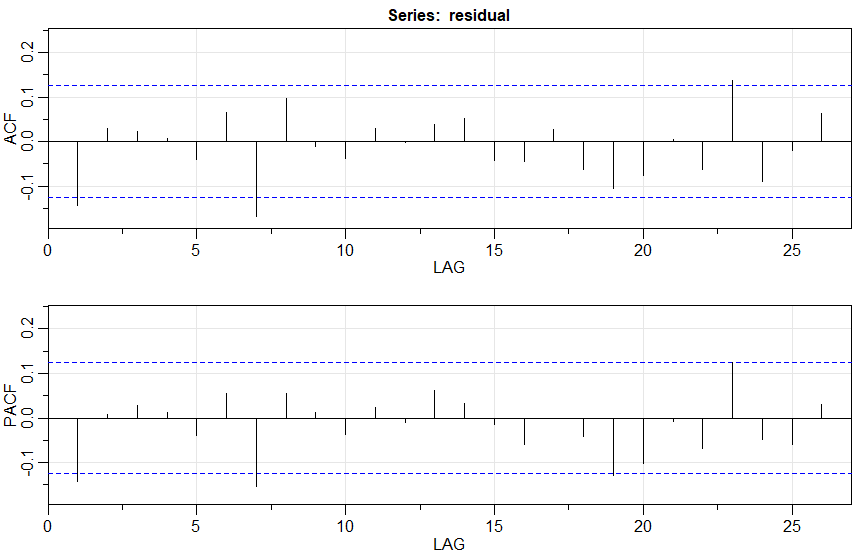


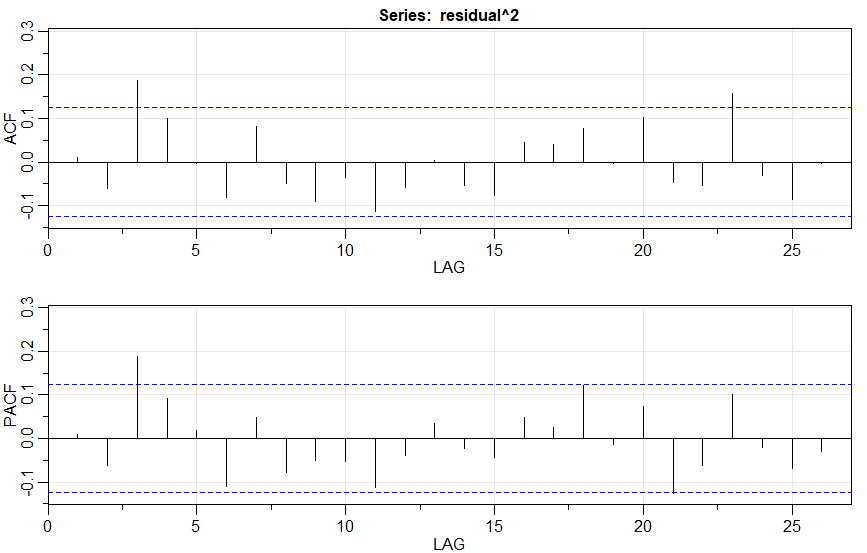
As we can see from the summary that GARCH(1,1) has less standard deviation than GARCH(1,0)model. Also the AIC value for GARCH(1,1) is less which makes it less volatile. Analysis can also be done from the residual plots. We know that ACF and PACF for residuals should be non-significant for the model to be good ( i.e errors should be random without any correlation). GARCH(1,0) model has significant ACF and PACF value for the residuals. Hence we use GARCH(1,1), for forecast purpose as it will give less error results.

**WEEKLY:**

Same model was run for weekly data and the summary was interpreted:

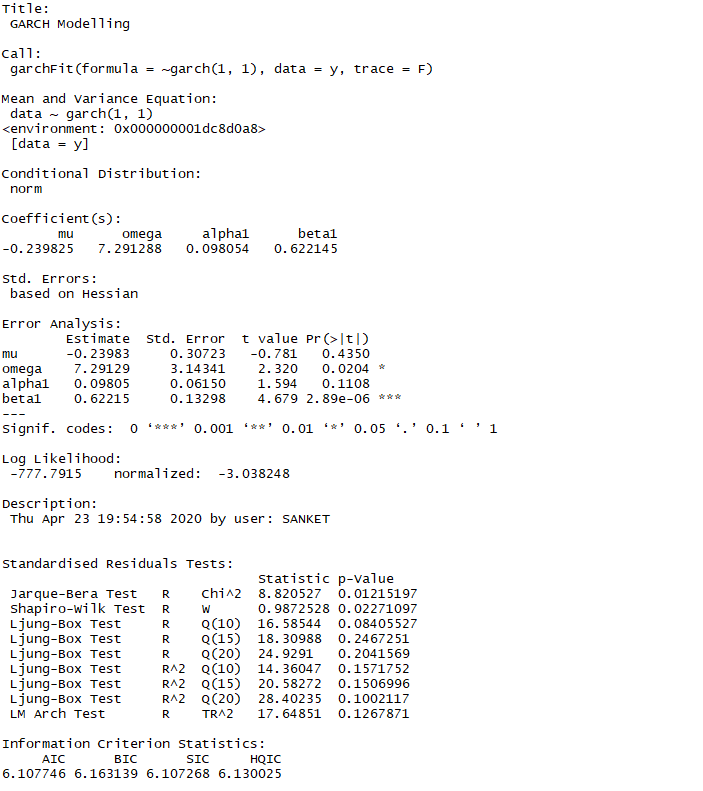






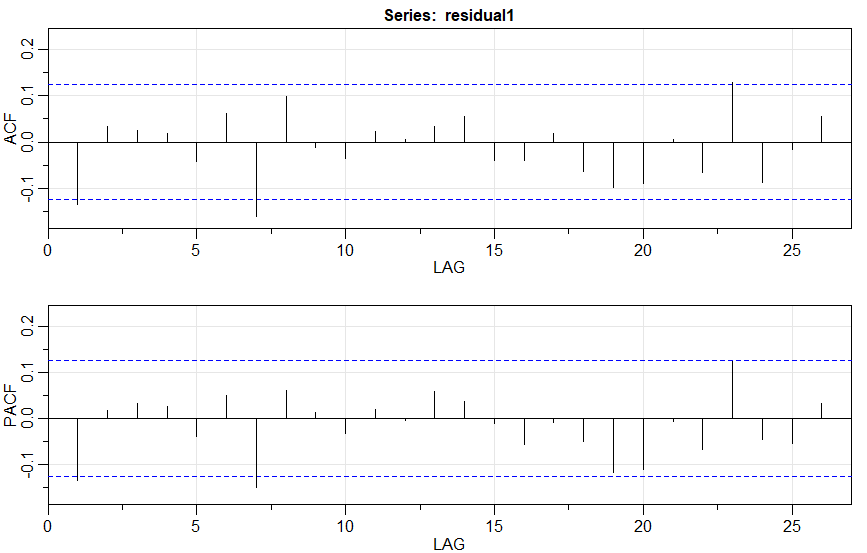
We get the same trend in results for the weekly data as that of daily data. We can see from the ACF and PACF plots that there is a significant correlation in residuals for GARCH(1,0) model which is relatively less in GARCH(1,1) model. So, for weekly data we should go for some other time series model for predictions.

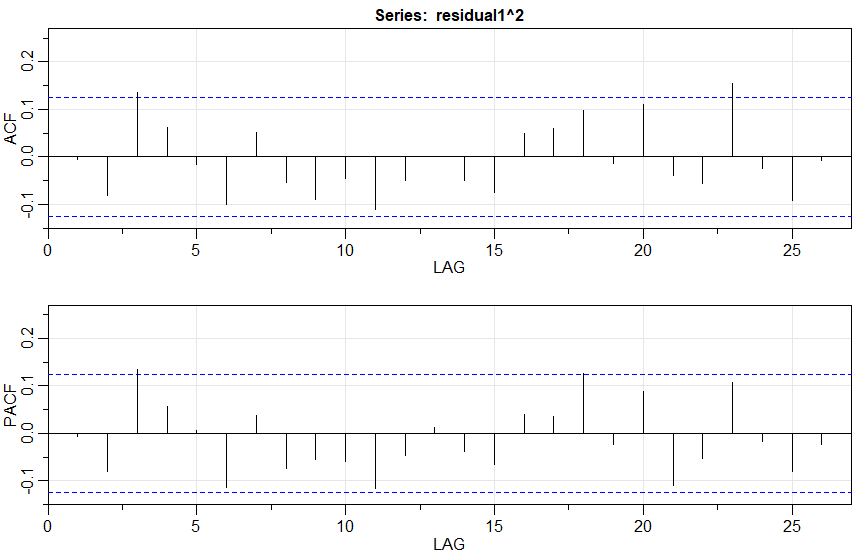
Summary of GARCH(1,1) model for weekly data:

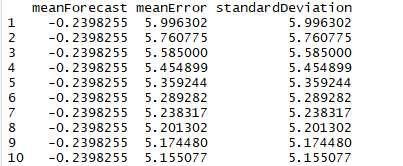


P-value should be less than 0.05 to reject the null hypothesis. Model with less p-value, with less AIC is always favoured. For the above simulation we can see that AIC value for GARCH(1,1) is less. So, it is more fit model.

Residual Plots for weekly data:

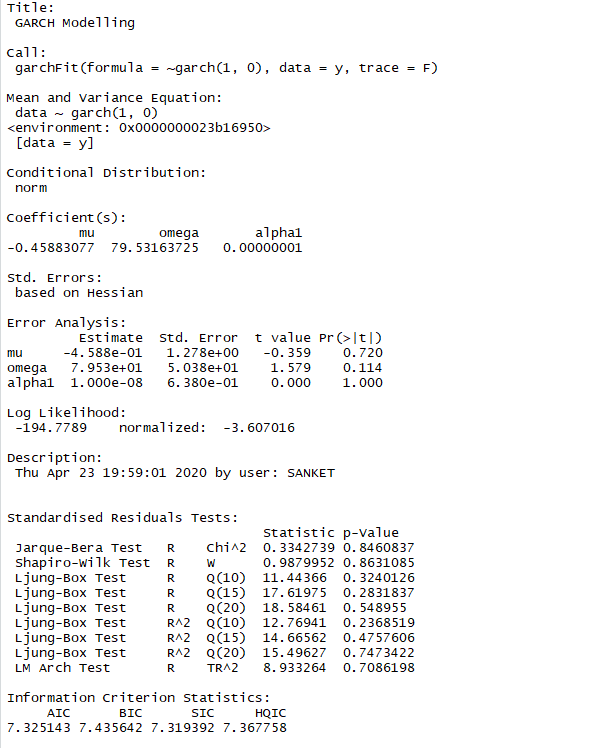


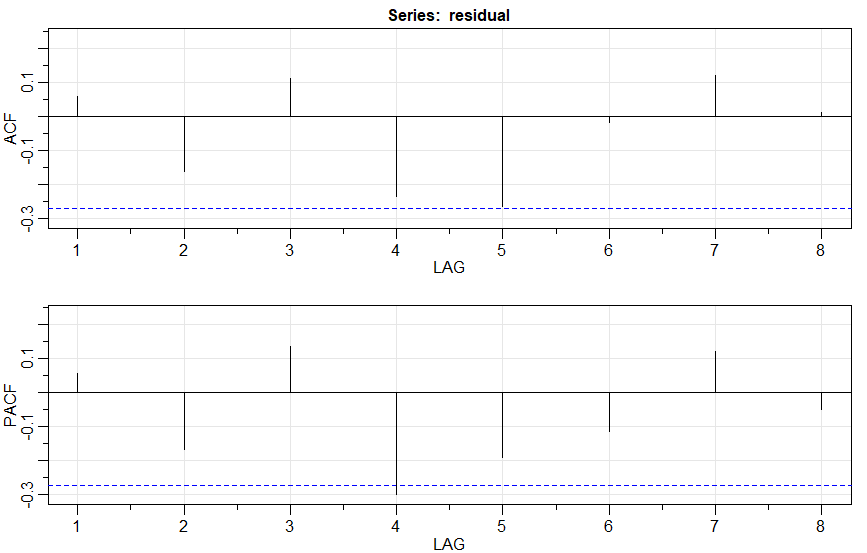


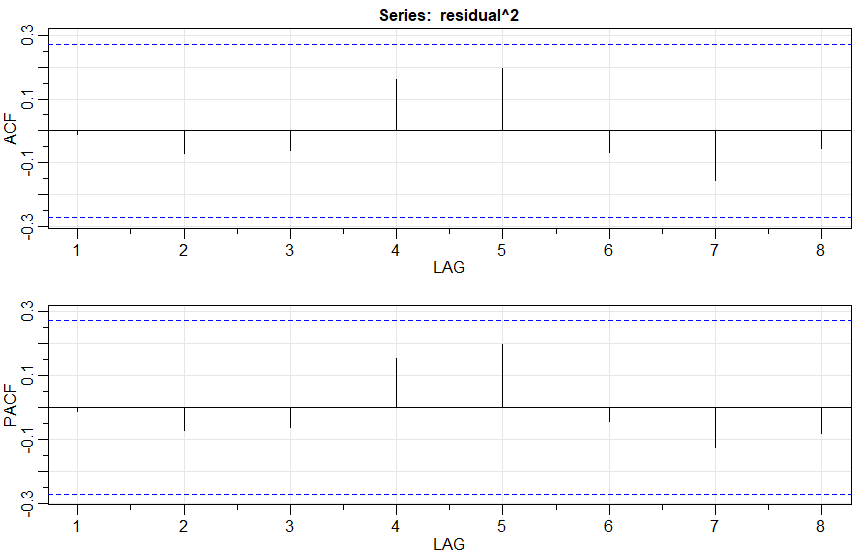


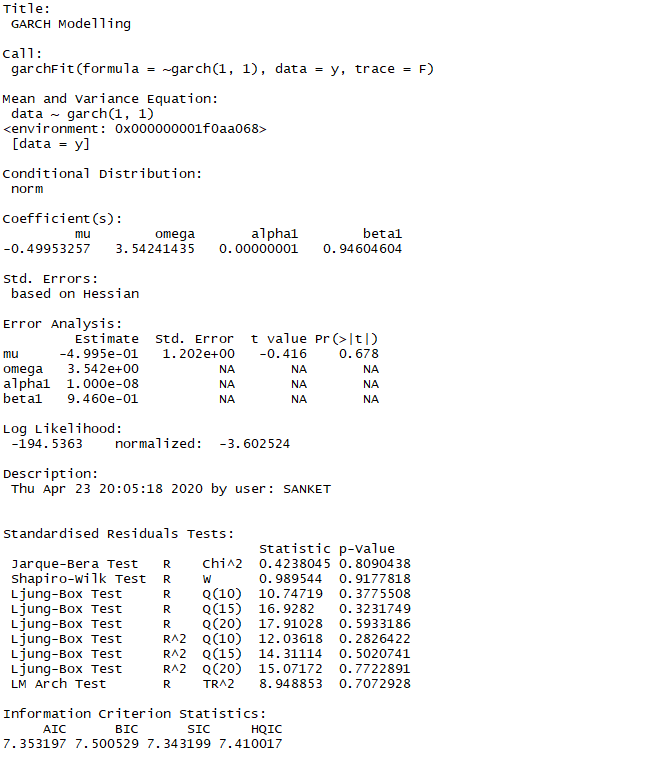
**MONTHLY:**

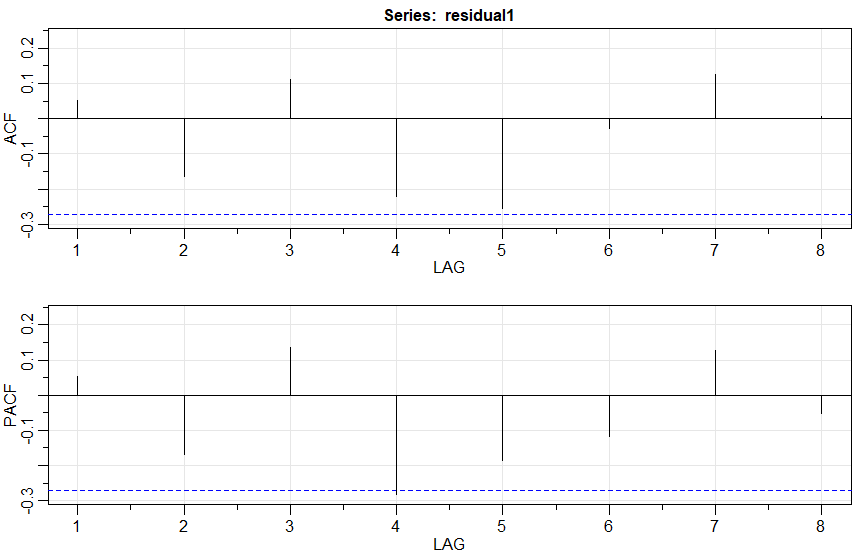
Same model was run for monthly data and the summary was interpreted:

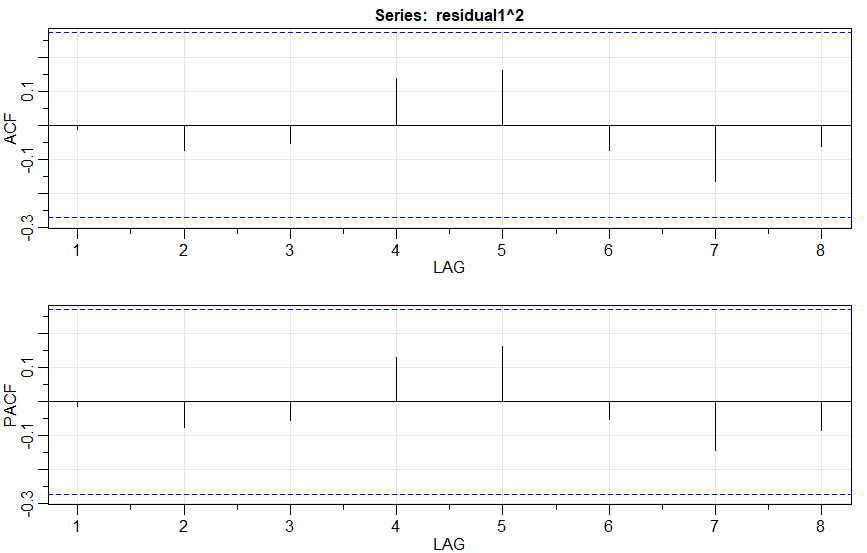












In case of monthly data we can see that both model can be accepted as ACF and PACF of residuals lie below the significant line

(Note: Daily, weekly, monthly all followed the same trend)

**VECTOR AUTOREGRESSIVE MODELS VAR(P) MODELS:**

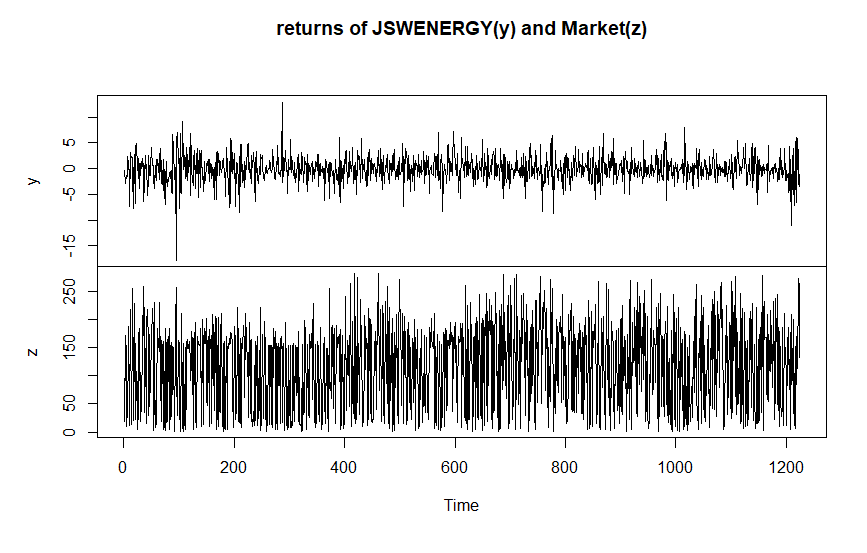
VAR models (vector autoregressive models) are used for multivariate time series. The structure is that each variable is a linear function of past lags of itself and past lags of the other variables. In general, for a VAR(p) model, the first p lags of each variable in the system would be used as regression predictors for each variable. VAR models are a specific case of more general VARMA models. VARMA models for multivariate time series include the VAR structure above along with moving average terms for each variable.

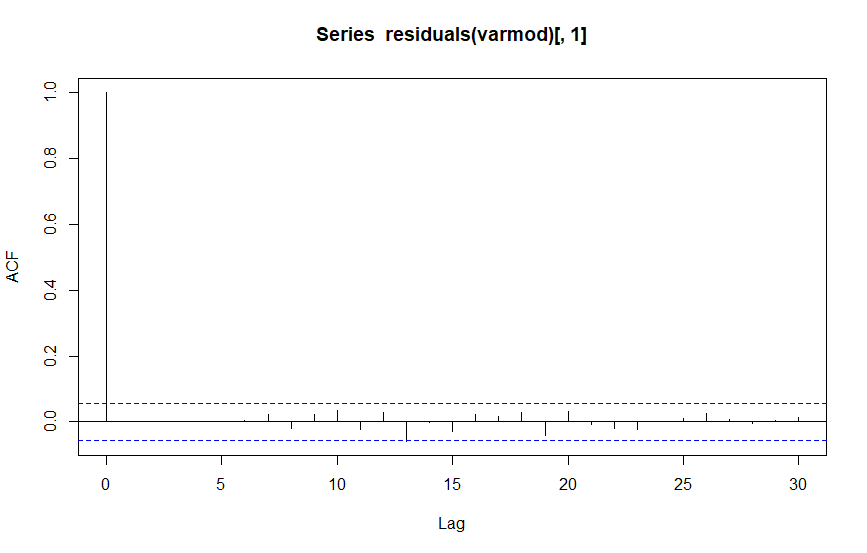
****

The coefficients to the corresponding lag components helps to build the equation which will describe how and to what extent the variables are related.

**DAILY:**

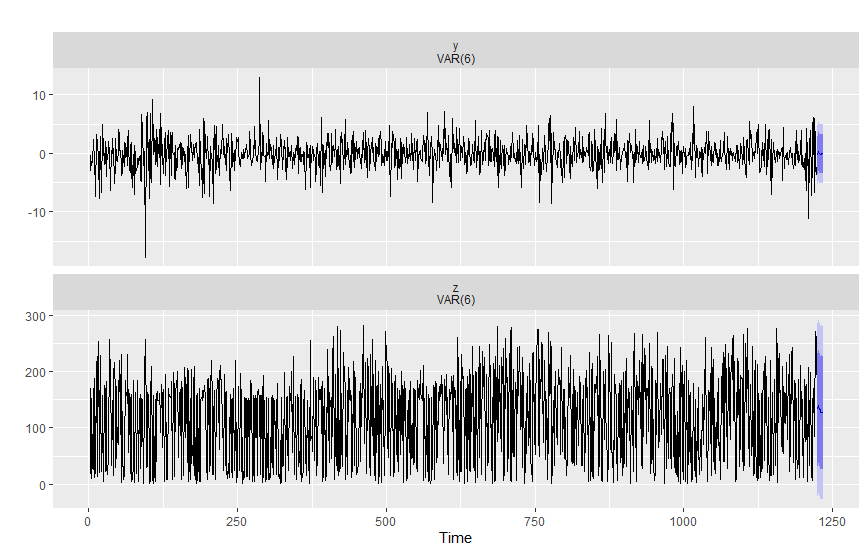
Returns of JSWENERGY(y) and Market(z1)

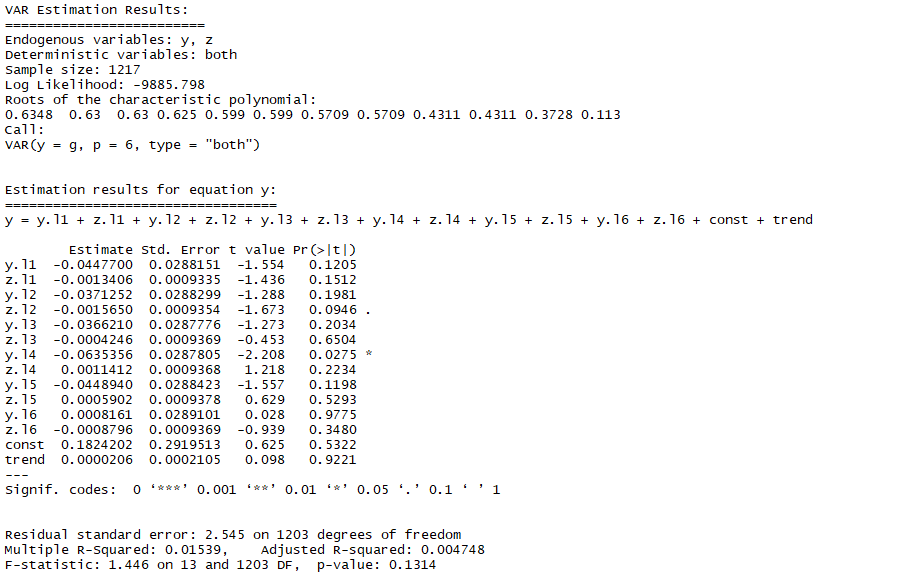


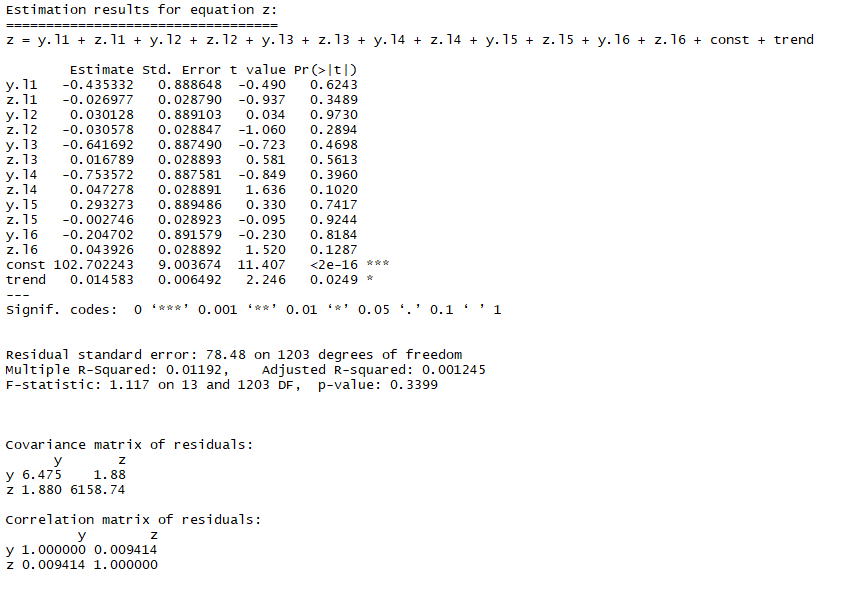


ACF for residuals are non-significant which tells that it is a good fit model.

Forecasting using VAR model:

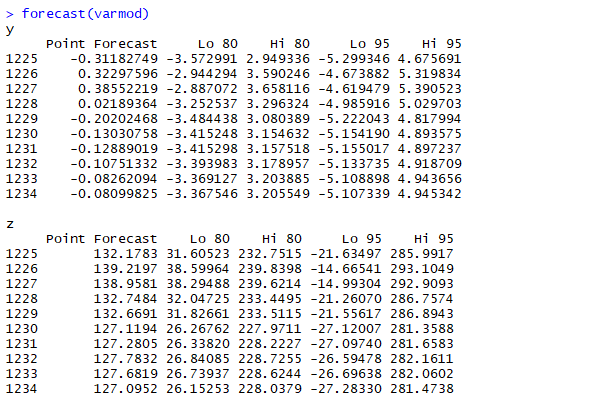






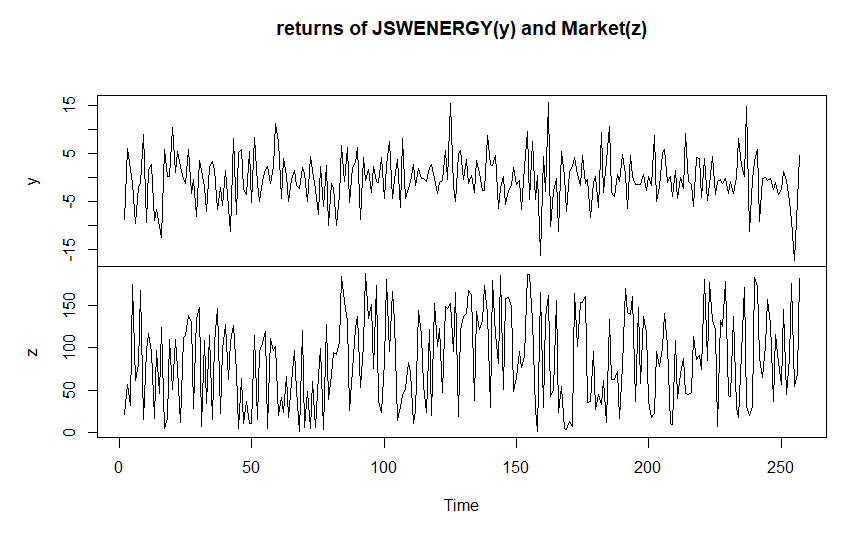
From the above summary we can find the coefficients for variables with lag1, lag2...upto lag16.

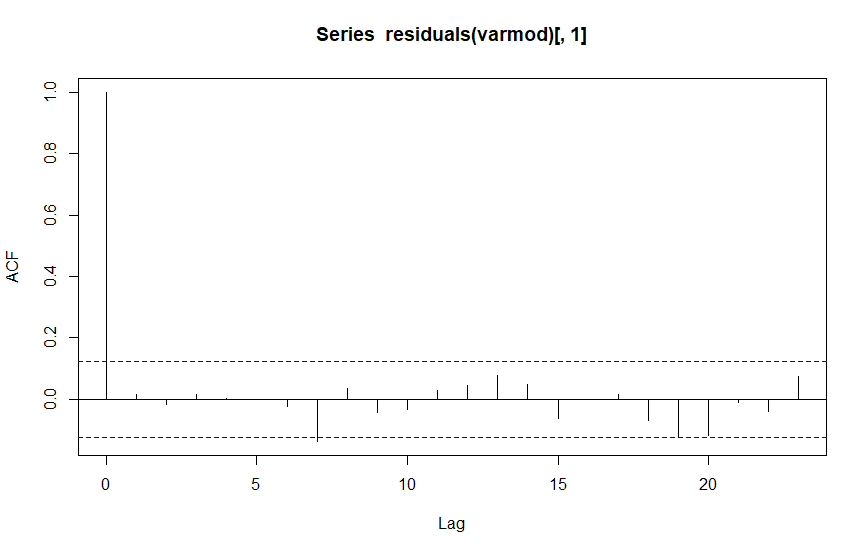
P-value should be less than 0.05 for the model to be a good fit.

****

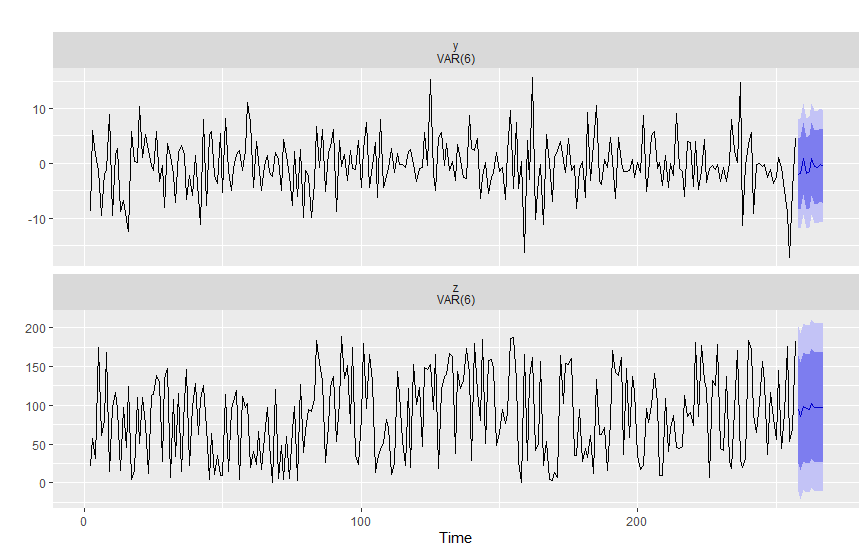
**WEEKLY:**

Returns of JSWENERGY(y) and Market(z1)

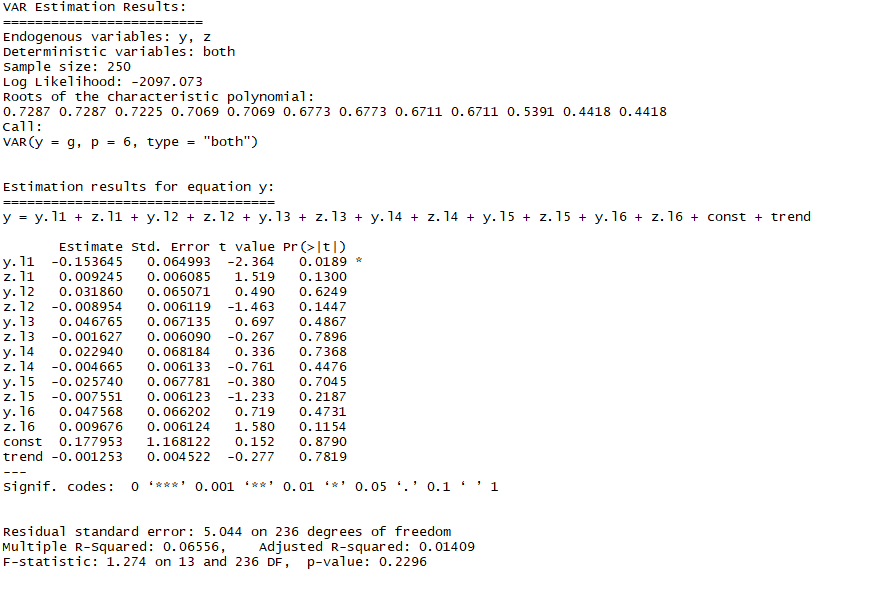


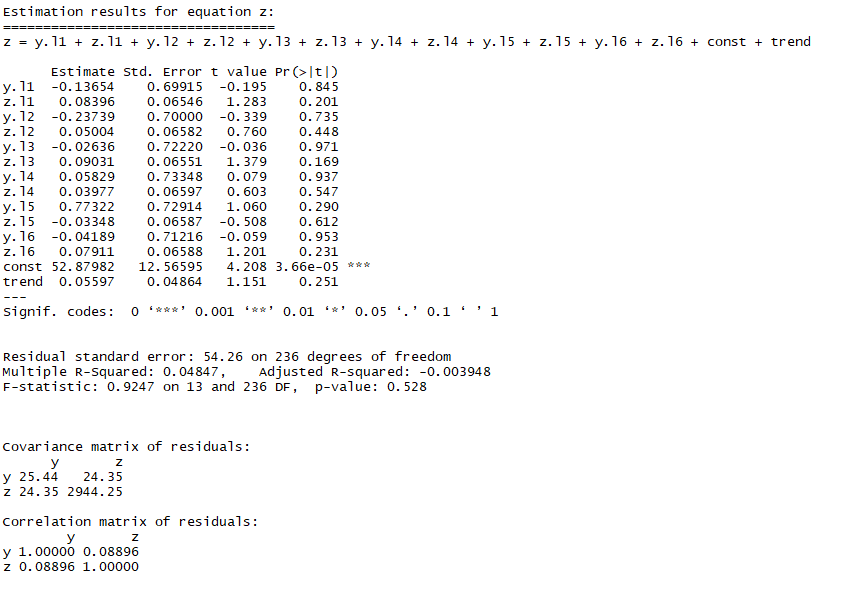


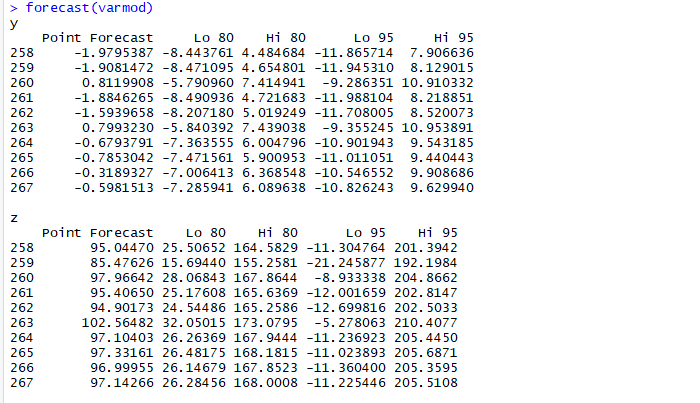
Here also there is no significant correlation in residuals which tells it is a good model.



The forecast shows a negative return trend, So, the prices may fall down on coming days according the given model.

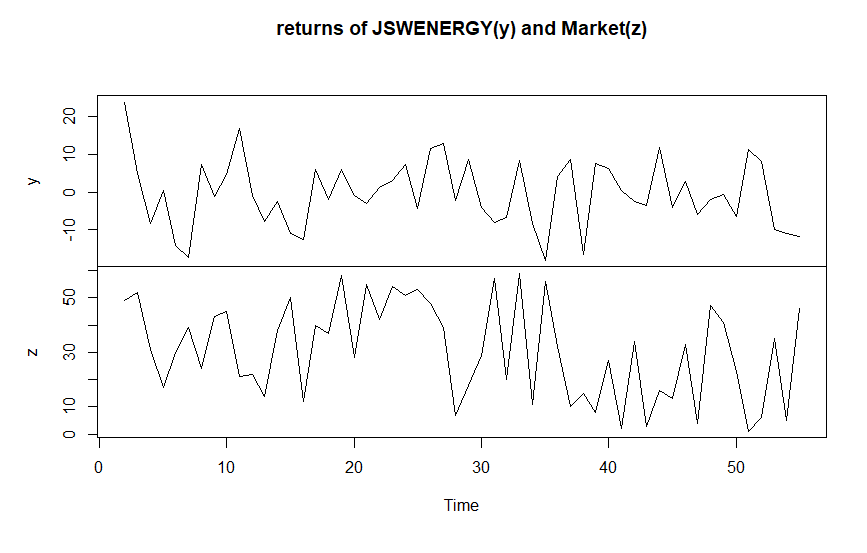
****

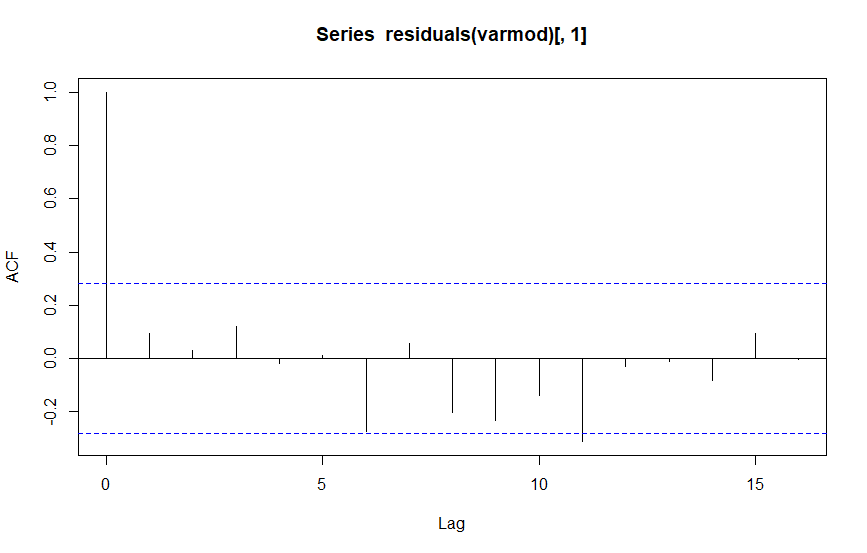
****

****

**MONTHLY:**

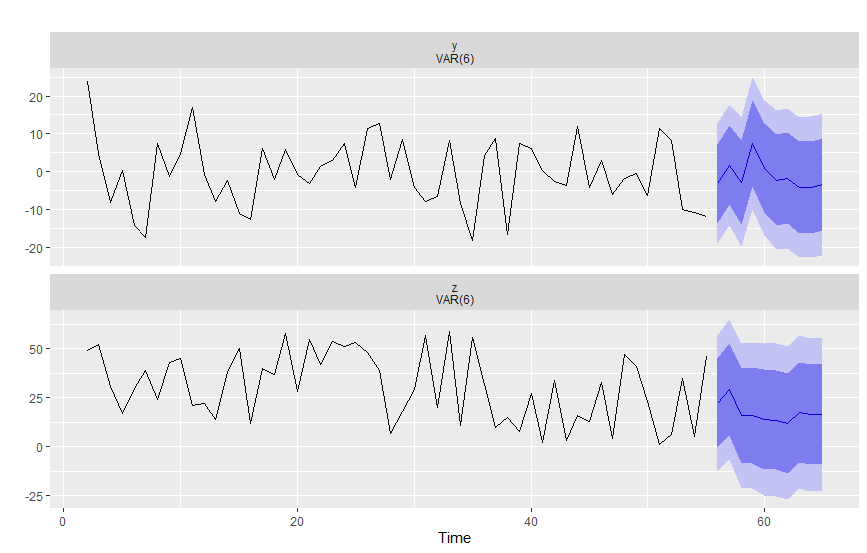
Returns of JSWENERGY(y) and Market(z1)

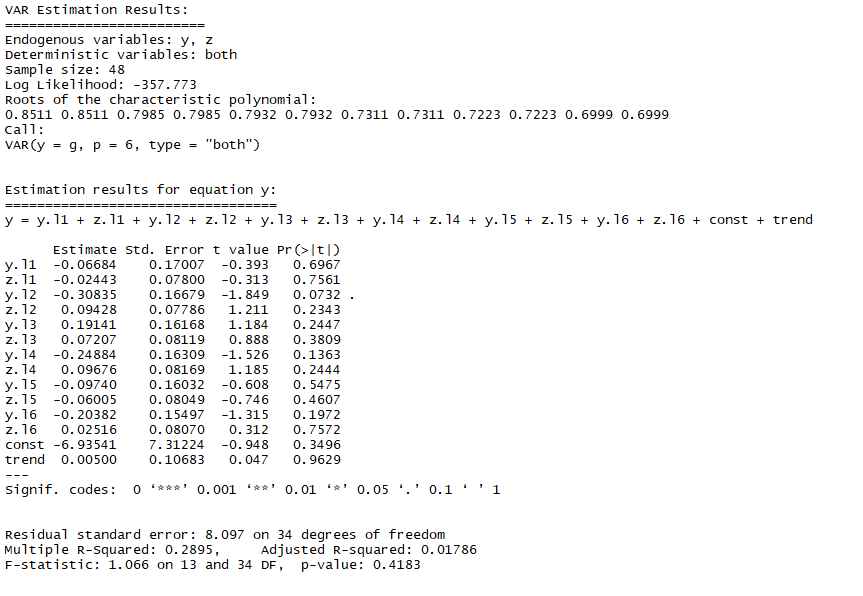


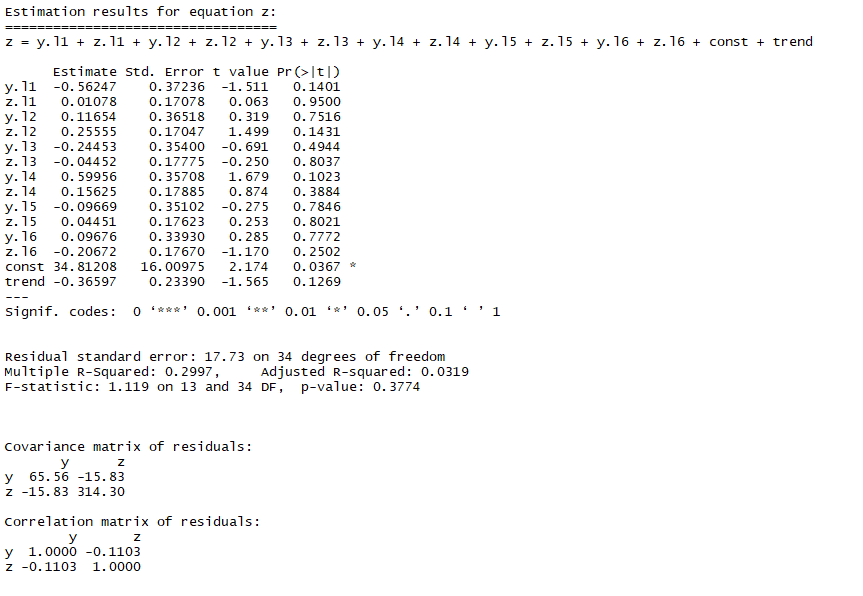


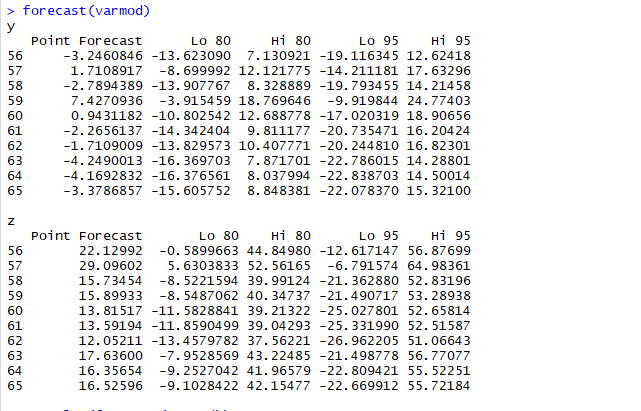
Here also there is no significant correlation in residuals which tells it is a good model.

Forecast using VAR model:



****

****

****

The forecast shows a negative return trend, So, the prices may fall down on coming days according the given model. We build different models and ran simulation and the most fit model always gave a negative stock returns trend.

**CONCLUSION-**

Financial stock price data are generally non-seasonal in reason which made the forecasting job easy. ACF and PACF played an important role in determining correct time series model for forecasting and for residual analysis.  The longer the time span, the less sensitive, or more smoothed out, the moving average becomes. The forecasted data had a negative return trend which we can validate from the downward moving close price plot of the stock. Daily, weekly and monthly data act like different time series. This results in different ARIMA model for same time duration due to change in frequency. GARCH is a better fit for modelling time series data when the data exhibits heteroskedacisticity and volatility clustering. ARCH(1) = GARCH(1,0) versus GARCH(1,1), the latter always fits financial data better than does the former. Neither ARCH nor GARCH can capture asymmetry or leverage.