**Report on:**

**STOCKS RETURNS AND VOLATILITY MODELLING USING R**

Submitted to:

**Dr. Thota Nagaraju**

**By**

**Guduguntla Venkata Sai Sumanth**



**BITS Pilani, Hyderabad Campus**

**FIN F414 Financial Analytics and Risk Management**

Contents

[ABSTRACT 7](#_Toc38638725)

[ACKNOWLEDGEMENT 8](#_Toc38638726)

[INTRODUCTION: 9](#_Toc38638727)

[DATA: 9](#_Toc38638728)

[METHODOLOGY: 10](#_Toc38638729)

[PART-1 CESC 11](#_Toc38638730)

[CESC 11](#_Toc38638731)

[History and briefing about the company: 11](#_Toc38638732)

[Products: 11](#_Toc38638733)

[Business nature: 12](#_Toc38638734)

[Ownership: 12](#_Toc38638735)

[Industry significance: 12](#_Toc38638736)

[Overall greatness: 13](#_Toc38638737)

[References: 13](#_Toc38638738)

[ARIMA, GARCH, ARCH, VAR models for CESC 14](#_Toc38638739)

[Daily data based 14](#_Toc38638740)

[ACF&PACF of data 14](#_Toc38638741)

[Interpretation 15](#_Toc38638742)

[Fitting Arima model 15](#_Toc38638743)

[Checking residuals of model fitted: 16](#_Toc38638744)

[Forecasting 10 upcoming intervals using the model 16](#_Toc38638745)

[Volatility estimation using GARCH(1,1) and related other ARCH models 17](#_Toc38638746)

[Model parameters: 18](#_Toc38638747)

[Residual analysis: 18](#_Toc38638748)

[Predicting next observation: 20](#_Toc38638749)

[Predicting returns using both models together 20](#_Toc38638750)

[Conclusion on daily data: 22](#_Toc38638751)

[Monthly data based 22](#_Toc38638752)

[ACF&PACF of data 23](#_Toc38638753)

[Interpretation 23](#_Toc38638754)

[Fitting Arima model 23](#_Toc38638755)

[Checking residuals of model fitted: 24](#_Toc38638756)

[Forecasting 10 upcoming intervals using the model 25](#_Toc38638757)

[Volatility estimation using GARCH(1,1) and related other ARCH models 26](#_Toc38638758)

[Model parameters: 26](#_Toc38638759)

[Residual analysis: 27](#_Toc38638760)

[Predicting next observation: 28](#_Toc38638761)

[Predicting returns using both models together 29](#_Toc38638762)

[Conclusion on Monthly data: 30](#_Toc38638763)

[Weekly data based 31](#_Toc38638764)

[ACF&PACF of data 31](#_Toc38638765)

[Interpretation 32](#_Toc38638766)

[Fitting Arima model 32](#_Toc38638767)

[Checking residuals of model fitted: 33](#_Toc38638768)

[Forecasting 10 upcoming intervals using the model 33](#_Toc38638769)

[Volatility estimation using GARCH(1,1) and related other ARCH models 34](#_Toc38638770)

[Model parameters: 35](#_Toc38638771)

[Residual analysis: 35](#_Toc38638772)

[Predicting next observation: 37](#_Toc38638773)

[Predicting returns using both models together 38](#_Toc38638774)

[Conclusion on weekly data: 40](#_Toc38638775)

[VAR models on CESC: 40](#_Toc38638776)

[Daily Data Based: 40](#_Toc38638777)

[Monthly data Based 41](#_Toc38638778)

[Weekly data based: 43](#_Toc38638779)

[Identifying and interpreting ARMA models of above data as AR(p),MA(q) models; 45](#_Toc38638780)

[Part-2 CADILAHC 47](#_Toc38638781)

[Cadila Healthcare 47](#_Toc38638782)

[History and briefing about the company: 47](#_Toc38638783)

[Products: 48](#_Toc38638784)

[Business Nature: 48](#_Toc38638785)

[Ownership: 48](#_Toc38638786)

[Industry significance: 48](#_Toc38638787)

[Overall greatness: 49](#_Toc38638788)

[References: 49](#_Toc38638789)

[ARIMA, GARCH, ARCH models for CESC 50](#_Toc38638790)

[Daily data based 50](#_Toc38638791)

[ACF&PACF of data 50](#_Toc38638792)

[Interpretation 51](#_Toc38638793)

[Fitting Arima model 51](#_Toc38638794)

[Checking residuals of model fitted: 51](#_Toc38638795)

[Forecasting 10 upcoming intervals using the model 52](#_Toc38638796)

[Volatility estimation using GARCH(1,1) and related other ARCH models 53](#_Toc38638797)

[Model parameters: 54](#_Toc38638798)

[Residual analysis: 54](#_Toc38638799)

[Predicting next observation: 56](#_Toc38638800)

[Predicting returns using both models together 57](#_Toc38638801)

[Conclusion on daily data: 58](#_Toc38638802)

[Monthly data based 58](#_Toc38638803)

[ACF&PACF of data 59](#_Toc38638804)

[Interpretation 60](#_Toc38638805)

[Fitting Arima model 60](#_Toc38638806)

[Checking residuals of model fitted: 61](#_Toc38638807)

[Forecasting 10 upcoming intervals using the model 61](#_Toc38638808)

[Volatility estimation using GARCH(1,1) and related other ARCH models 62](#_Toc38638809)

[Model parameters: 63](#_Toc38638810)

[Residual analysis: 63](#_Toc38638811)

[Predicting next observation: 65](#_Toc38638812)

[Predicting returns using both models together 66](#_Toc38638813)

[Conclusion on Monthly data: 67](#_Toc38638814)

[Weekly data based 67](#_Toc38638815)

[ACF&PACF of data 68](#_Toc38638816)

[Interpretation 68](#_Toc38638817)

[Fitting Arima model 68](#_Toc38638818)

[Checking residuals of model fitted: 69](#_Toc38638819)

[Forecasting 10 upcoming intervals using the model 70](#_Toc38638820)

[Volatility estimation using GARCH(1,1) and related other ARCH models 70](#_Toc38638821)

[Model parameters: 71](#_Toc38638822)

[Residual analysis: 71](#_Toc38638823)

[Predicting next observation: 73](#_Toc38638824)

[Predicting returns using both models together 74](#_Toc38638825)

[Conclusion on weekly data: 75](#_Toc38638826)

[VAR models on CADILAHC: 75](#_Toc38638827)

[Daily Data Based: 75](#_Toc38638828)

[Monthly data Based 77](#_Toc38638829)

[Weekly data based: 78](#_Toc38638830)

[Identifying and interpreting ARMA models of above data as AR(p),MA(q) models; 80](#_Toc38638831)

[CONCLUSION 82](#_Toc38638832)

[NOTE: 82](#_Toc38638833)

[APPENDIX 83](#_Toc38638834)

[Data download from nsepy: 83](#_Toc38638835)

[Risk-free rate data: 83](#_Toc38638836)

[R Code: 83](#_Toc38638837)

# ABSTRACT

**One of the key aspects in finance is returns and its volatility. If these are predictable to a decent extent within a near but decent range, using just the available stock prices data and nothing more, by identifying trends, cycles, and seasonality in the data, then we can use it for our benefit. This is done by finding trends in the data using certain undergraduate level math. Once found these can be used to estimation with certain confidence level and of a range, and can easily be used to gain decent returns on the investments, at least prevent wary losses which could be forecasted by this method. This is the core abstract of this assignment. This is done by first finding the returns of stocks, fitting decent, and better model to the returns and use it to fit model for volatility estimations and use the knowledge of two together for our investment goals. This entire thing can be done with the help of R programming and few useful libraries in it.**

# ACKNOWLEDGEMENT

**I am extremely thankful to Dr. Thota Nagaraju for providing me with an opportunity to work on this project. This project was a wonderful way of applying what I have learnt in our course by analyzing the data, none of which would have been possible without his help and crucial inputs. I would also like to express my most heartfelt gratitude towards him for his patience and invaluable time by guiding me through each step of this project. I am also thankful to Sunny Kumar sir, for helping me understand the key concepts, without which, this assignment cannot be done. I am also thankful to the Economics Department of BITS, which has created such provisions for students to participate in fun, innovative projects that give us a hands-on experience in applying our knowledge. This project has certainly developed my skills and given us the necessary tools to excel further in the field of finance and fintech. It was challenging and laborious but the knowledge gained from it will impact me in new helpful ways. Once again, I thank you Dr. Thota Nagaraju for being a mentor and guide and hope to work with you again soon.**

# INTRODUCTION:

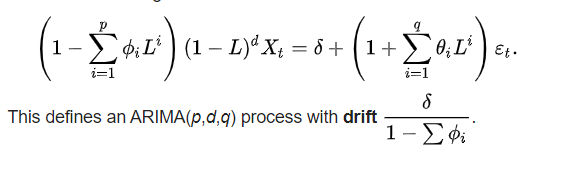
In statistics and econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. Both of these models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). ARIMA models are applied in some cases where data show evidence of non-stationarity, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied one or more times to eliminate the non-stationarity. The AR part of ARIMA indicates that the evolving variable of interest is regressed on its own lagged (i.e., prior) values. The MA part indicates that the regression error is actually a linear combination of error terms whose values occurred contemporaneously and at various times in the past. The I (for "integrated") indicates that the data values have been replaced with the difference between their values and the previous values (and this differencing process may have been performed more than once). The purpose of each of these features is to make the model fit the data as well as possible. Non-seasonal ARIMA models are generally denoted ARIMA(p,d,q) where parameters p, d, and q are non-negative integers, p is the order (number of time lags) of the autoregressive model, d is the degree of differencing (the number of times the data have had past values subtracted), and q is the order of the moving-average model. Generalized Autoregressive conditional Heteroscedasticity (GARCH) is a statistical tool to approach the estimated volatility in markets. It has its various uses in financial modelling because it provides values closer to reality when we try to predict the prices and rates of financial instruments. So, to understand in depth we need to break the GARCH process. Firstly, the heteroscedasticity tells us the asymmetrical arrangement of variation of error term where they are no more linear rather form clusters. So, with many loop holes the interpretations thus derived are not reliable and hence this is one tool which analyses the data which is today’s world is mostly used for estimating the volatility of returns for stocks, bonds, or market indices. So, we first fit the autoregressive model and then compute the sigma terms as well as residual terms and then check for their significance. While doing this we check the stationarity of the residual series too. This report consists of the steps performed to perform the GARCH Analysis followed by forecasting.

# DATA:

The data was extracted from the NSE website with the help of the nsepy library. This is one of the best techniques where we can directly download data of any duration according to the requirement. Appendix A has the code for the python for this task. Now the log returns are calculated for the closing prices for the further analysis. The companies for which this analysis is done are CESC and CADILAHC. The time period chosen for the analysis is the 5 years from April 1st 2015 to 31st march 2020 for weekly, monthly data-based models and last financial year for daily data-based models. Since they of different industries we can’t compare them directly but if at all they fall under same industry then their volatilities can be compared and conclusions are deduced.

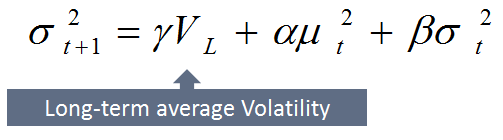
# METHODOLOGY:

Generally, need to interpret the acf and pacf graphs of our returns data, and then use the same for forecasting and modelling the returns. But as it is a tough process, we will use the auto.Arima function available in R on our returns data to find the suitable model for our data.



To model volatility in returns, we use the GARCH, and its related models. It is based on the principle explained below:

Heteroscedasticity is the change of variance with time also the conditional implies the conditional variability in the latest volatility, and the autoregressive means the positive correlation today’s and yesterday’s volatility. In GARCH the variance tends to show mean reversion and it gets pulled to a long-term volatility rate over time.



GARCH model incorporates to get rid of the least squares and also a prediction is computed for the variance of each error term. Three Models are attempted with the log returns. They are GARCH (1, 1), ARCH (1, 0), ARCH (2, 0).

# PART-1 CESC



## CESC

## History and briefing about the company:

The **Calcutta Electric Supply Corporation** or **CESC** is the Kolkata-based flagship company of the RP-Sanjiv (Goenka Group), born from the erstwhile RPG Group, under the chairmanship of businessman Sanjiv Goenka. CESC is a fully integrated power utility with its operation spanning from mining coal, generating AND distribution of power, serving 2.4 million customers for domestic, industrial and commercial users, within 567 square kilometers of Kolkata and Howrah, delivering safe, cost-effective and reliable energy since 1899. They have private participation in generation, transmission and distribution of electrical power. They own & operate three thermal power plants generating 1125 MW of power. This system comprises of 474 km circuit of transmission lines linking the company’s generating & receiving stations with 105 distribution stations, 8,211 circuit km of HT lines further linking distribution stations with LT substations, large industrial consumers and 12,269 circuit km of LT lines connecting the LT substations to LT consumers. In diurnal course, they have verged upon renewable sources. With their continuous effort to accomplish the requirement of consumers and make them easily avail all our services from their premises, they employed value added online services.

## Products:

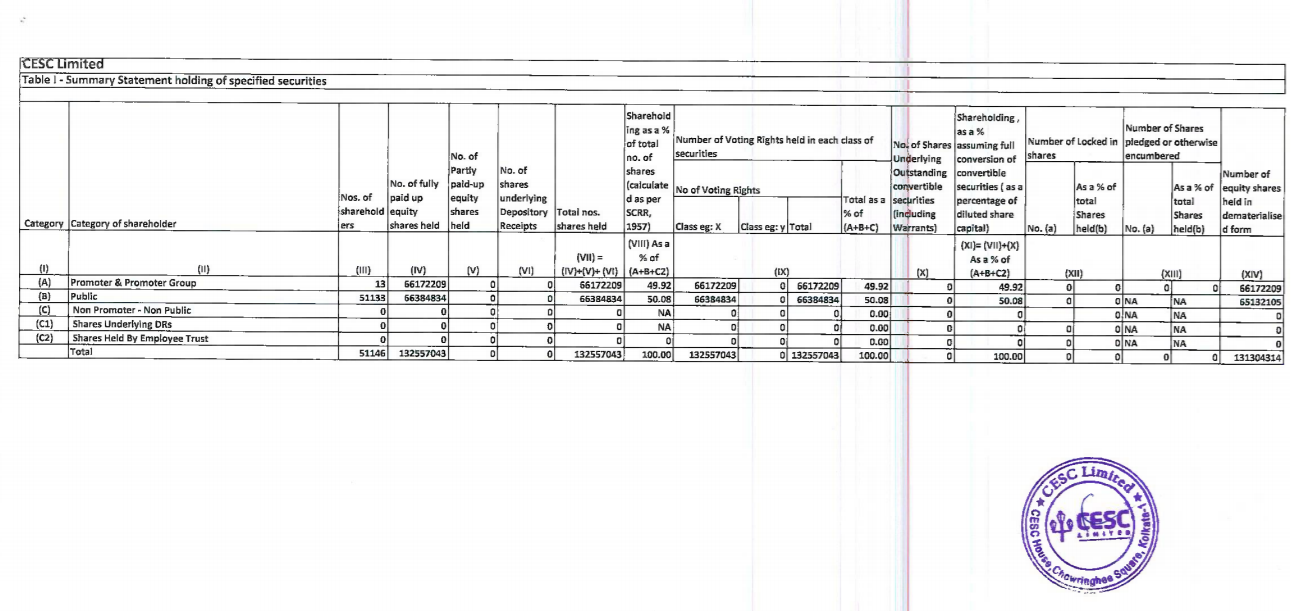
They generate and distribute power, generated by both non-renewable, and renewable sources. Throughout Kolkata, and are trying to expand their business throughout the country.

## Business nature:

**CESC** is one of India’s leading power producer and distributor. It is clearly evident from their financials and their contribution towards the total power production and distribution in the country.

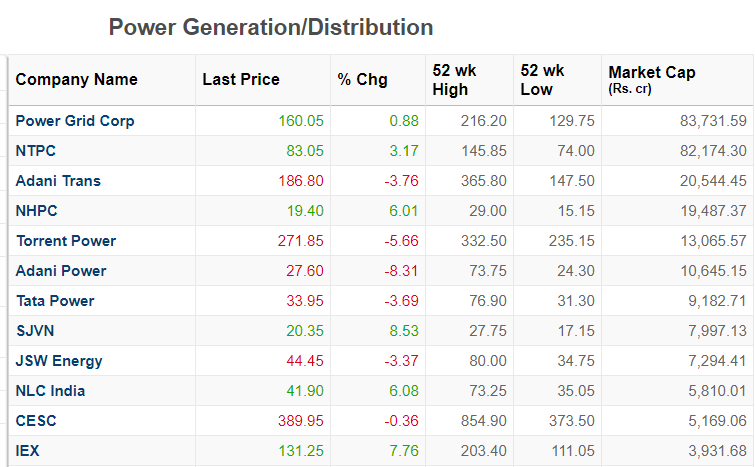
## Ownership:

CESC is a prominent energy producer and distributor, and is a Public company. This is clearly reflected in the number of shares and their market capitalization, which show the value of the company in the investor’s perspective. Below snapshot of shareholding pattern reflects the same.



## Industry significance:

The company, though has been through many ups and downs since it’s inception, tried to maintain its significance in Indian energy production and distribution industry. The market cap of the company at present, reflects the same.



## Overall greatness:

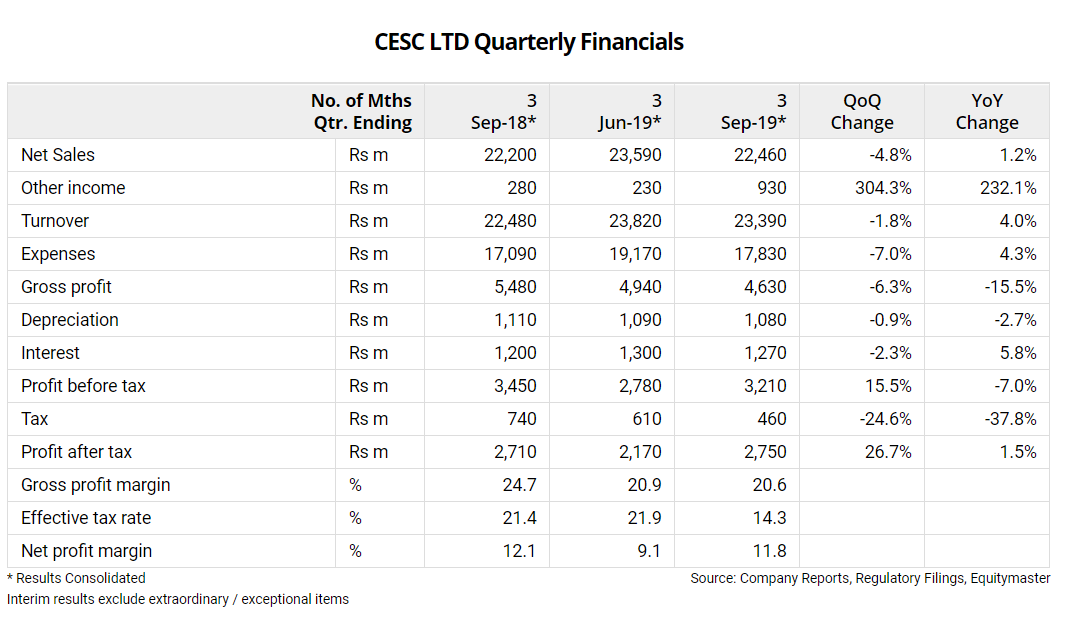
* They were the first company to produce and distribute power in Kolkata, which gave them a first mover advantage in the same in that state.
* They have received many awards since their inception for their:

🡺generation methods and efficiency

🡺distribution processes and efficiency

🡺environmental concern

* They wish to stay true and deliver exceptional services to their consumers. CESC has major plans to expand generating capacity to 7,000 MW in future. The new power generating projects will involve investments of more than ₹300 billion (US$4.2 billion or €3.9 billion).
* Their recently updated financials also try to convey the same message:



## References:

<https://www.cesc.co.in/>for company related, shareholding pattern, history, awards related data.

<https://en.wikipedia.org/wiki/CESC_Limited> <https://www.moneycontrol.com/india/stockpricequote/power-generationdistribution/cesc/CES>

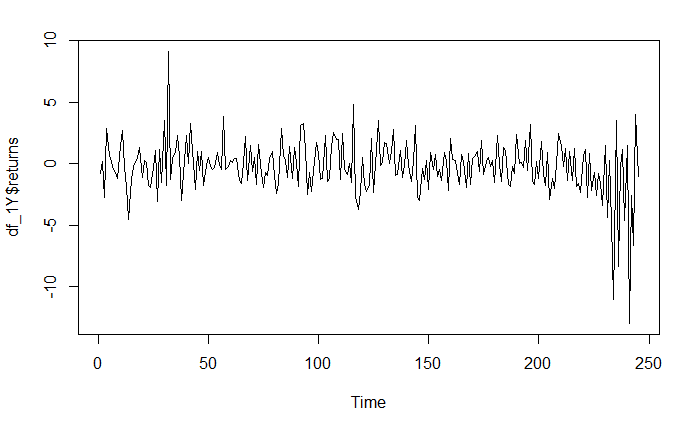
for general, products reference.

<https://www.equitymaster.com/stockquotes/complist.asp?company=cesc> for YoY comparison of quarterly financials of company.

<https://www.moneycontrol.com/stocks/marketinfo/marketcap/nse/power-generationdistribution.html> for industrial significance.

# ARIMA, GARCH, ARCH, VAR models for CESC

## Daily data based



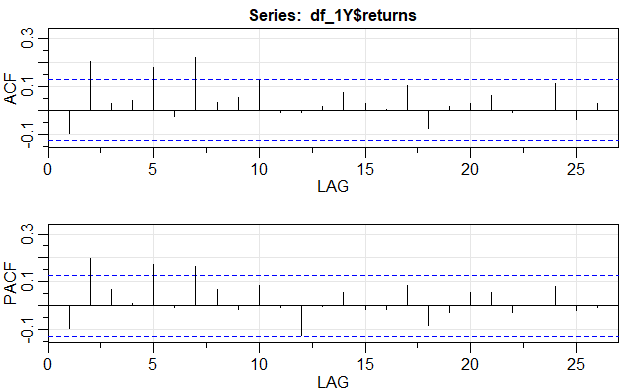
Daily returns are plotted as time series and the plot is as shown above:

Summary of above data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| -12.9606 | -1.4050 | -0.1318 | -0.2569 | 0.9981 | 9.1319 |

### ACF&PACF of data

now let us try fit an arima model to the data. Let us interpret the acf, pacf of data:



### Interpretation

as the data seems to show some trend, which is explained by the visible significant spikes at lags 2,5,7 in acf, and pacf, we shall use auto.arima to fit an approximate model to this data and use it for forcasting.

### Fitting Arima model

Using auto. arima function, arima model is fitted to the above data. The fitted model related parameters are as mentioned below.

Series: df\_1Y$returns

ARIMA(1,1,1)

|  |  |  |
| --- | --- | --- |
|  | ar1 | ma1 |
| Coefficients: | -0.2266 | -0.9117 |
| s.e. | 0.0663 | 0.0296 |

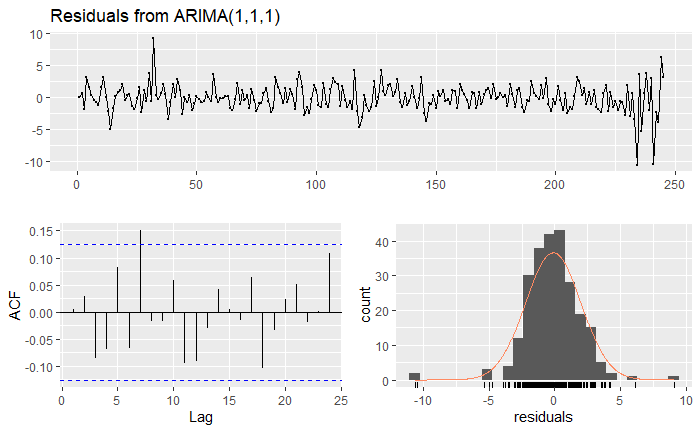
sigma^2 estimated as 4.453: log likelihood=-528.55

AIC=1063.09 AICc=1063.19 BIC=1073.58

Training set error measures:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
| -0.1189344 | 2.097331 | 1.522918 | 30.12139 | 187.2415 | 0.6499132 | 0.004117542 |

### Checking residuals of model fitted:



The fitted model’s residuals seem to do a decent interpretation of entire time series, and is clearly shown in the acf plot of model’s residuals. The residuals are following a normal distribution and seem to be not correlated at all. Hence the model does a decent job in estimating the time series. The Ljung-Box test on residuals also interpret the same.

Ljung-Box test

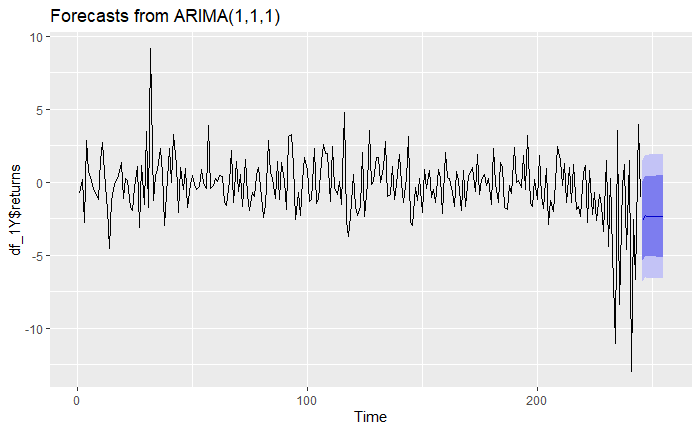
data: Residuals from ARIMA(1,1,1)

Q\* = 12.882, df = 8, p-value = 0.116

Model df: 2. Total lags used: 10

### Forecasting 10 upcoming intervals using the model

Using necessary code, forecast was done for 10 upcoming days on the same time series and is as shown below:



The dark blue region shows the possible estimated returns over a 80 percent confidence interval and light blue region shows estimated returns over a 95 percent confidence interval.

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

246 -2.655348 -5.359794 0.04909836 -6.791442 1.480746

247 -2.285285 -5.015470 0.44490053 -6.460743 1.890173

248 -2.369129 -5.118410 0.38015164 -6.573791 1.835533

249 -2.350133 -5.104387 0.40412143 -6.562401 1.862135

250 -2.354437 -5.116033 0.40716001 -6.577934 1.869061

251 -2.353462 -5.121803 0.41488045 -6.587275 1.880352

252 -2.353682 -5.128882 0.42151693 -6.597984 1.890619

253 -2.353632 -5.135643 0.42837836 -6.608351 1.901086

254 -2.353644 -5.142456 0.43516833 -6.618764 1.911476

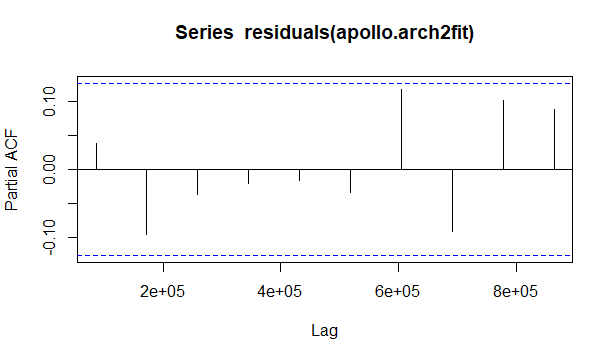
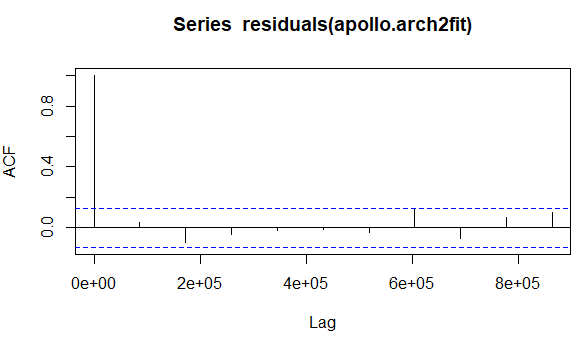
255 -2.353641 -5.149236 0.44195418 -6.629135 1.921853

The model is fitted and can be used for volatility estimations.

### Volatility estimation using GARCH(1,1) and related other ARCH models

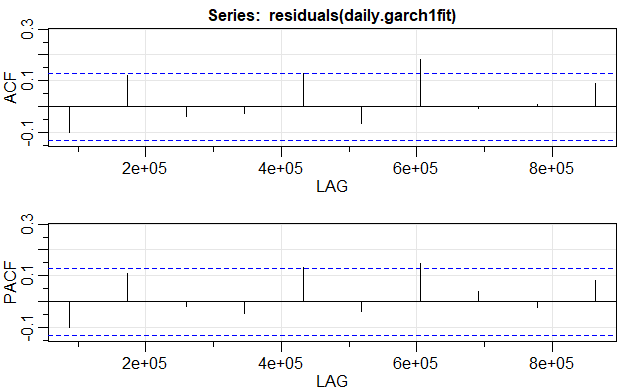
As the data is correlated, fitting garch models with known data has been difficult, since the model is needed to be differenced and used. So we will try fitting the garch model on differenced data.

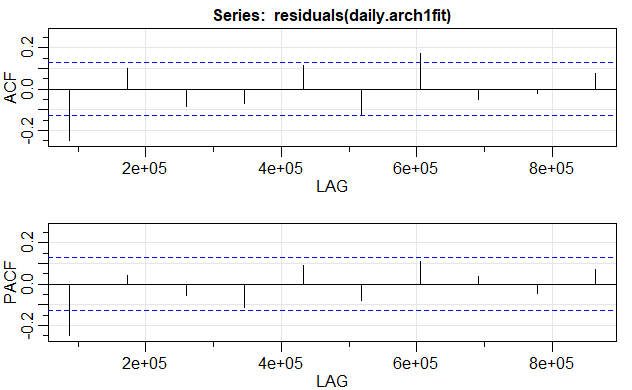
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model parameters:  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Estimate | Estimate | Estimate | | mu 0.001146 | mu 0.001976 | mu 0.001403 | | omega 0.342321 | omega 2.575582 | omega 1.658590 | | alpha1 0.231872 | alpha1 0.541786 | alpha1 0.353859 | | beta1 0.716387 |  | alpha2 0.414776 |   COMPARISION OF INFORMATION CRITERIA OF 3 MODELS:   |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Akaike 4.1838 | Akaike 4.2780 | Akaike 4.1969 | | Bayes 4.2698 | Bayes 4.3497 | Bayes 4.2829 | | Shibata 4.1826 | Shibata 4.2772 | Shibata 4.1957 | | Hannan-Quinn 4.2184 | Hannan-Quinn 4.3069 | Hannan-Quinn 4.2315 | |  | |

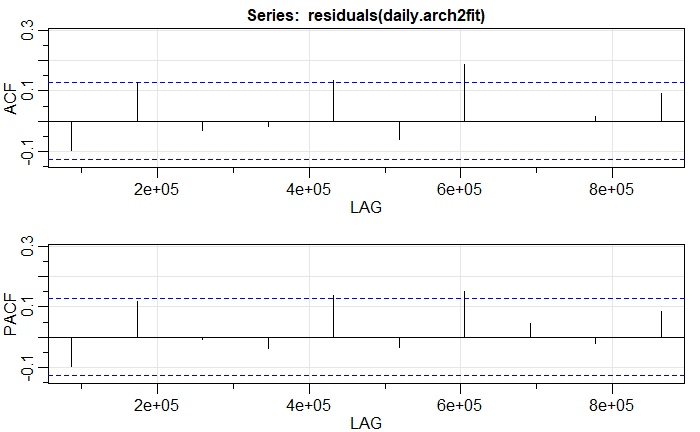
The code is given in the Appendix. The required output is tabulated above which is clearly interpreted. The Information Criteria needs to be least for the model to be best. In the above table the information criteria is the least for the GARCH (1, 1). So the garch (1,1) model is best for volatility estimation.

### Residual analysis:

We find the residuals and the sigma of the GARCH (1, 1), ARCH (1, 0) and ARCH (2, 0) then check for their stationarity by acf and pacf plots. The data needs to be stationary for various reasons because if it is not stationary the time series cannot be applied major reason being the sudden shocks can’t be absorbed by the data and get reverted back to mean. The below results clearly show that the residuals for all the models are stationary because they fall in between the blue line.







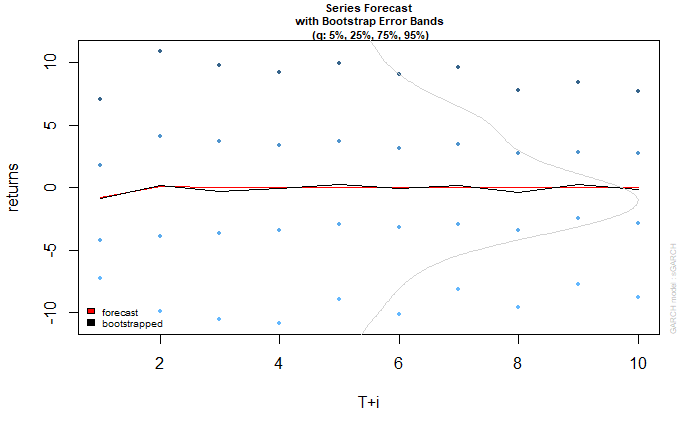
These clearly show that garch (1,1) is clearly a better fit model for volatility estimation, when compared to other models.

### Predicting next observation:

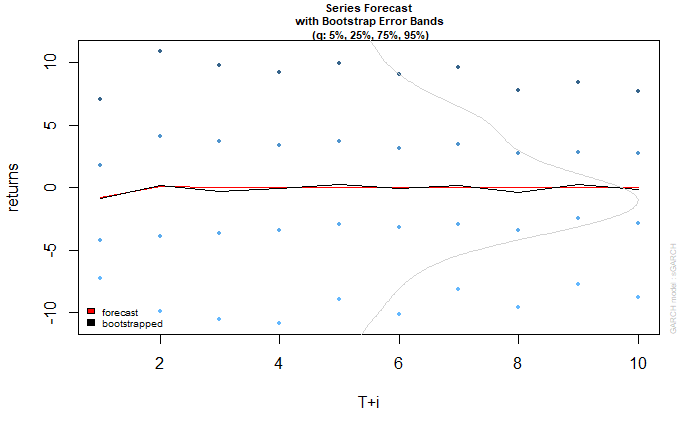
|  |  |  |
| --- | --- | --- |
| GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) |
| |  | | --- | | Sigma | | Sigma | Sigma |
| T+1 4.492083  T+2 4.413282  T+3 4.337235  T+4 4.263871  T+5 4.193117  T+6 4.124903  T+7 4.059159  T+8 3.995819  T+9 3.934813  T+10 3.876078 | T+1 1.880068  T+2 2.119105  T+3 2.237974  T+4 2.299811  T+5 2.332629  T+6 2.350218  T+7 2.359692  T+8 2.364810  T+9 2.367578  T+10 2.369076 | T+1 3.819374  T+2 2.699135  T+3 3.207361  T+4 2.884541  T+5 2.978217  T+6 2.872006  T+7 2.873385  T+8 2.828678  T+9 2.813272  T+10 2.788908 |

The forecasts for sigma show that garch (1,1) was the best at estimating the volatility pattern in the entire series. This is shown by the forecast values of sigma

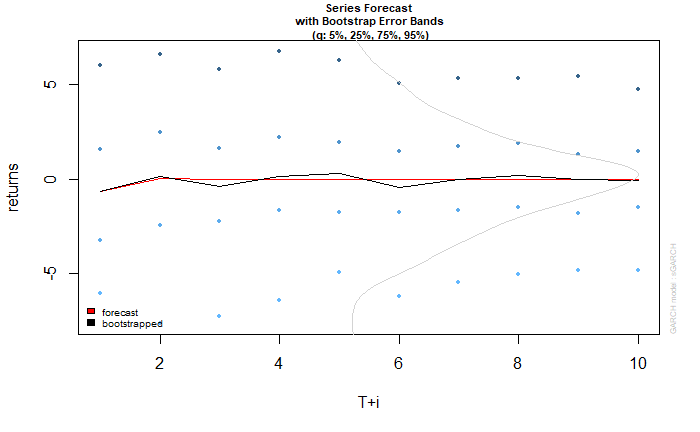
### Predicting returns using both models together



Garch(1,1) based forecast of returns, blue points showing extremes based on modelled volatility model.



Arch1 based forecast of returns, blue points showing extremes based on modelled volatility model.

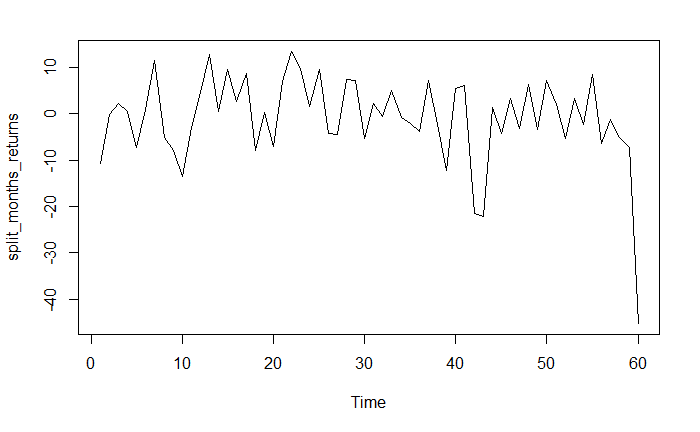


Arch2 based forecast of returns, blue points showing extremes based on modelled volatility model.

### Conclusion on daily data:

The models build above show us that arima (1,1,1) is a better fit for data, and using the same model to predict the volatility of returns, garch (1,1) gives us the best possible results in model estimation. The results are tabulated and shown above.

## Monthly data based



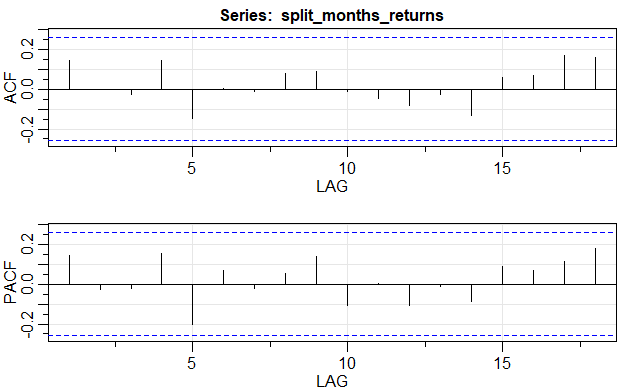
Monthly returns are plotted as time series and the plot is as shown above:

Summary of above data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| -45.1273 | -5.0405 | 0.1348 | -0.9566 | 5.6177 | 13.4819 |

### ACF&PACF of data

now let us try fit an arima model to the data. Let us interpret the acf, pacf of data:



### Interpretation

as the data seems to show no trend, which is explained by no visible significant spikes at all lags in acf, and pacf, the data is just white noise and we shall use auto.arima to show the same.

### Fitting Arima model

Using auto. arima function, arima model is fitted to the above data. The fitted model related parameters are as mentioned below.

Series: split\_months\_returns

ARIMA(0,0,0)

|  |  |  |
| --- | --- | --- |
|  | ar1 | ma1 |
| Coefficients: | 0 | 0 |
| s.e. | 0 | 0 |

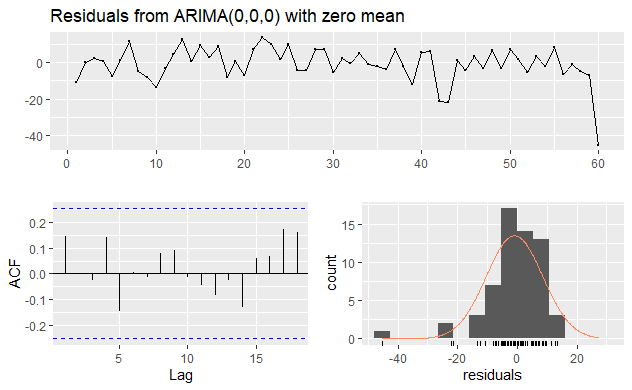
sigma^2 estimated as 89.04: log likelihood=-219.81

AIC=441.62 AICc=441.69 BIC=443.71

Training set error measures:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
| -0.9566233 | 9.436162 | 6.55695 | 100 | 100 | 0.7523192 | 0.1449289 |

### Checking residuals of model fitted:



The fitted model’s residuals seem to do a decent interpretation of entire time series, and is clearly shown in the acf plot of model’s residuals. The residuals are following a normal distribution and seem to be not correlated at all. Hence the model does a decent job in estimating the time series, ultimately saying that the entire data is just white noise, and can’t be used for predictions. The Ljung-Box test on residuals also interpret the same.

Ljung-Box test

data: Residuals from ARIMA(0,0,0) with zero mean

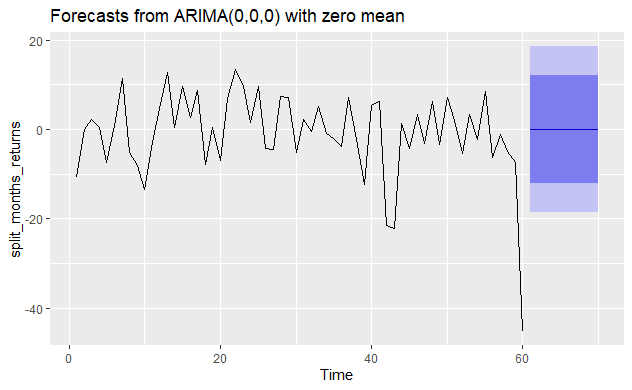
Q\* = 5.2395, df = 10, p-value = 0.8746

Model df: 0. Total lags used: 10

### Forecasting 10 upcoming intervals using the model

Using necessary code, forecast was done for 10 upcoming months on the same time series and is as shown below:

It is of no use to predict, and this is shown by the graph.



The dark blue region shows the possible estimated returns over a 80 percent confidence interval and light blue region shows estimated returns over a 95 percent confidence interval.

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

61 0 -12.09293 12.09293 -18.49454 18.49454

62 0 -12.09293 12.09293 -18.49454 18.49454

63 0 -12.09293 12.09293 -18.49454 18.49454

64 0 -12.09293 12.09293 -18.49454 18.49454

65 0 -12.09293 12.09293 -18.49454 18.49454

66 0 -12.09293 12.09293 -18.49454 18.49454

67 0 -12.09293 12.09293 -18.49454 18.49454

68 0 -12.09293 12.09293 -18.49454 18.49454

69 0 -12.09293 12.09293 -18.49454 18.49454

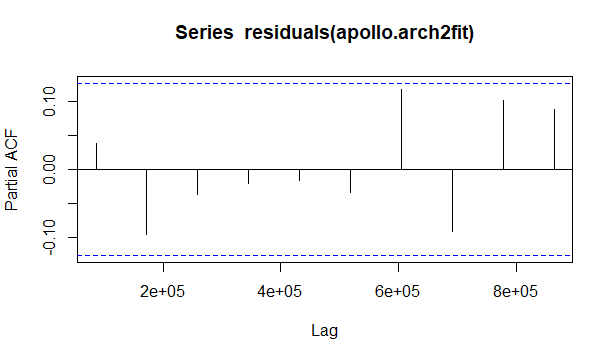
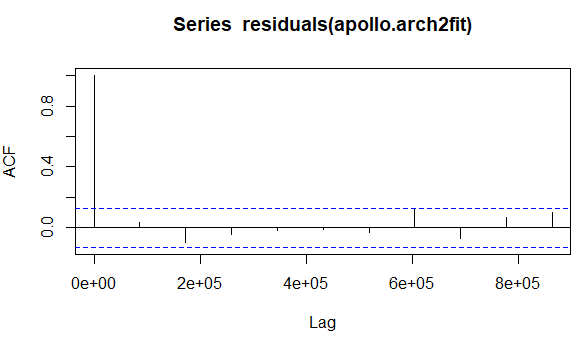
70 0 -12.09293 12.09293 -18.49454 18.49454

The model is fitted and can be used for volatility estimations.

### Volatility estimation using GARCH(1,1) and related other ARCH models

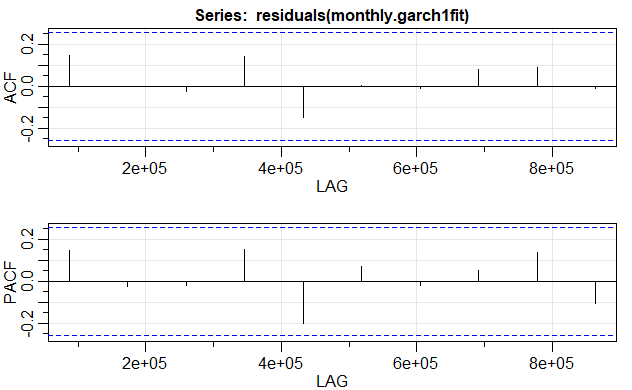
As the data is pure white noise, it cannot be used for any estimations, in addition to this problem, the no. of data points are so low, and the functions to converge the data into a model don’t seem to work. So, the models formed can be of no use for us. Nevertheless, we shall complete the asked task here.,

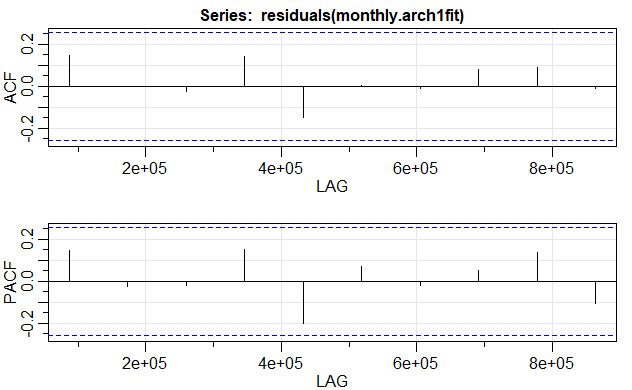
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model parameters:  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Estimate | Estimate | Estimate | | mu -0.82749 | mu 0.82993 | mu 0.85983 | | omega 0.63770 | omega 45.99828 | omega 46.41964 | | alpha1 0.00000 | alpha1 0.99900 | alpha1 0.99900 | | beta1 0.99900 |  | alpha2 0.00000 |   COMPARISION OF INFORMATION CRITERIA OF 3 MODELS:   |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Akaike 7.4243 | Akaike 7.3718 | Akaike 7.3940 | | Bayes 7.5640 | Bayes 7.4766 | Bayes 7.5336 | | Shibata 7.4162 | Shibata 7.3672 | Shibata 7.3859 | | Hannan-Quinn 7.4789 | Hannan-Quinn 7.4128 | Hannan-Quinn 7.4486 | |  | |

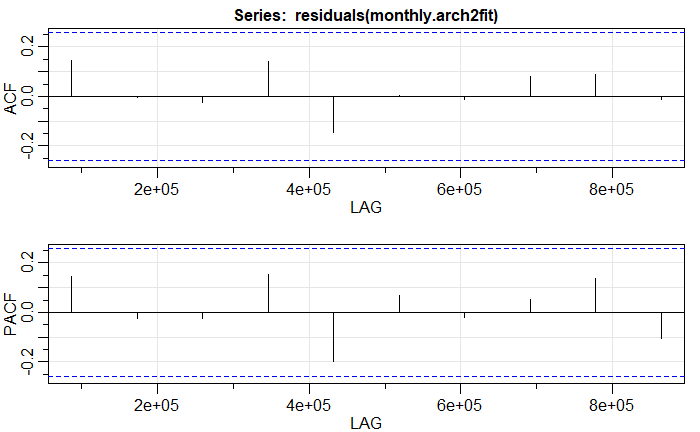
The code is given in the Appendix. The required output is tabulated above which is clearly interpreted. The Information Criteria needs to be least for the model to be best. But if we go by basic knowledge of finance, monthly volatility can never be more than at the most 20 percent. The models fitted have arch models with better aic values than of garch, but have more volatility in forecasts. So there is no use of these models. The forecasts show the same below

### Residual analysis:

We find the residuals and the sigma of the GARCH (1, 1), ARCH (1, 0) and ARCH (2, 0) then check for their stationarity by acf and pacf plots. The data needs to be stationary for various reasons because if it is not stationary the time series cannot be applied major reason being the sudden shocks can’t be absorbed by the data and get reverted back to mean. The below results clearly show that the residuals for all the models are stationary because they fall in between the blue line.







The points explained above, clearly show that no model can be used to forecast volatility of monthly intervals.

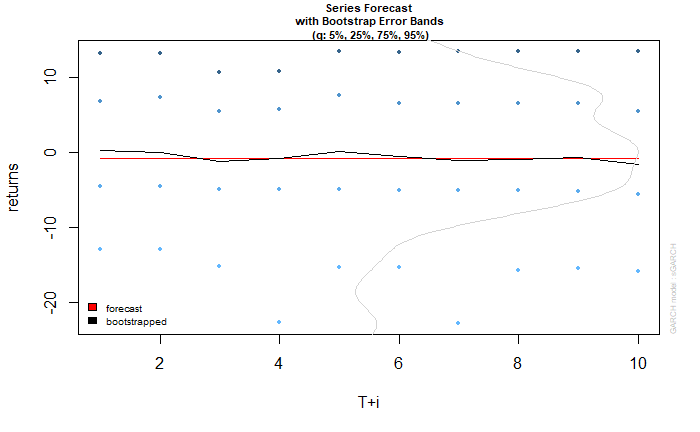
### Predicting next observation:

|  |  |  |
| --- | --- | --- |
| GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) |
| |  | | --- | | Sigma | | Sigma | Sigma |
| T+1 10.96184  T+2 10.98542  T+3 11.00893  T+4 11.03237  T+5 11.05573  T+6 11.07901  T+7 11.10223  T+8 11.12537  T+9 11.14845  T+10 11.17145 | T+1 46.43224  T+2 46.90198  T+3 47.36659  T+4 47.82623  T+5 48.28104  T+6 48.73116  T+7 49.17672  T+8 49.61784  T+9 50.05463  T+10 50.48721 | T+1 46.46634  T+2 46.94018  T+3 47.40882  T+4 47.87241  T+5 48.33110  T+6 48.78503  T+7 49.23432  T+8 49.67910  T+9 50.11950  T+10 50.55563 |

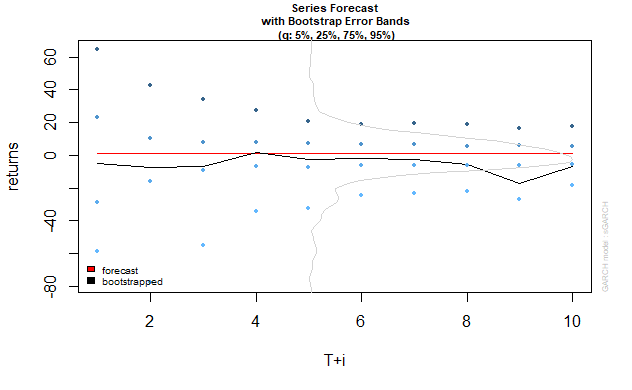
The forecasts for sigma show that no model is suitable for predicting volatility of stocks returns on monthly basis

### Predicting returns using both models together

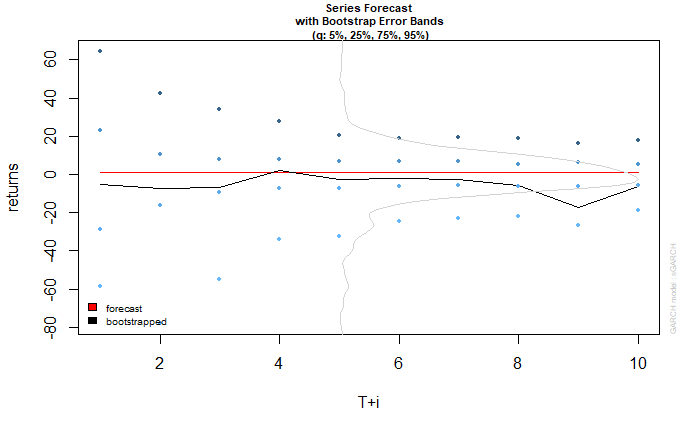
As stated above, these don’t give true picture and cannot be used. But for sake of completeness, we shall report them here.



Garch(1,1) based forecast of returns, blue points showing extremes based on modelled volatility model.



Arch1 based forecast of returns, blue points showing extremes based on modelled volatility model.

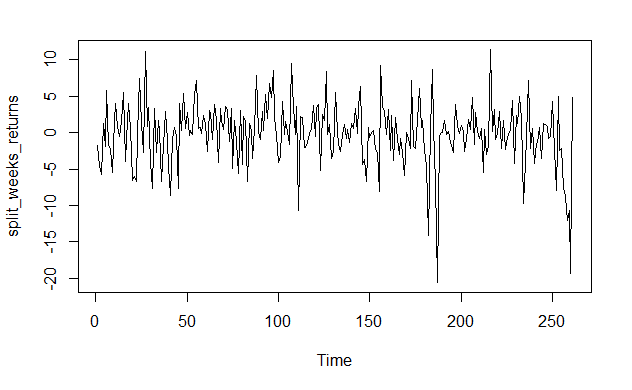


Arch2 based forecast of returns, blue points showing extremes based on modelled volatility model.

### Conclusion on Monthly data:

The models build above show us that white noise perfectly describes the data, and no model can be used for estimating the volatility trends in the same. The results calculated and shown above represent the same.

## Weekly data based



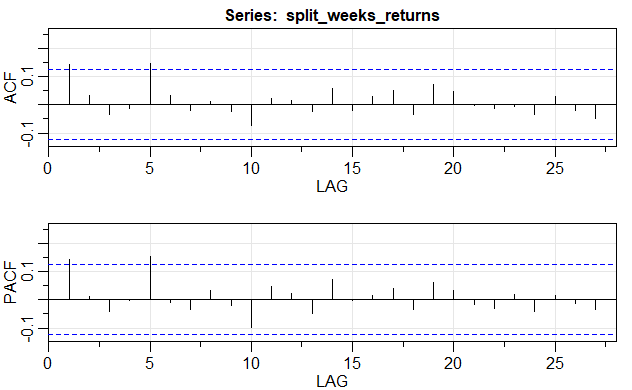
weekly returns are plotted as time series and the plot is as shown above:

Summary of above data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| -20.6149 | -2.1654 | 0.1510 | -0.1534 | 2.2397 | 11.4182 |

### ACF&PACF of data

now let us try fit an arima model to the data. Let us interpret the acf, pacf of data:



### Interpretation

as the data seems to show some trend, which is explained by the visible significant spikes at lags 1,5 in acf, and pacf, we shall use auto.arima to fit an approximate model to this data and use it for forecasting.

### Fitting Arima model

Using auto. arima function, arima model is fitted to the above data. The fitted model related parameters are as mentioned below.

Series: split\_weeks\_returns

ARIMA(0,0,1) with zero mean

|  |  |
| --- | --- |
|  | ma1 |
| Coefficients: | 0.1366 |
| s.e. | 0.0595 |

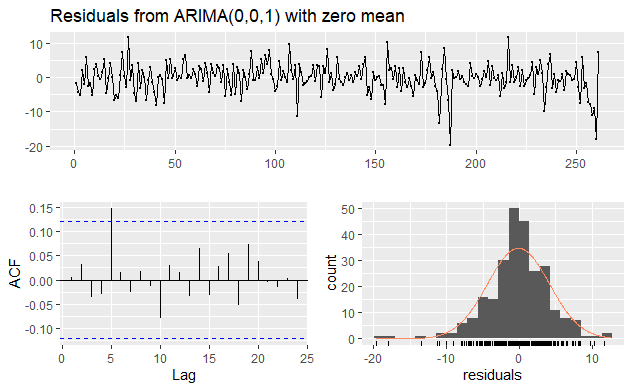
sigma^2 estimated as 18.12: log likelihood=-747.91

AIC=1499.82 AICc=1499.87 BIC=1506.95

Training set error measures:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
| 0.1315869 | -4.248569 | 3.070275 | inf | inf | 0.7139956 | 0.004332627 |

### Checking residuals of model fitted:



The fitted model’s residuals seem to do a decent interpretation of entire time series, and is clearly shown in the acf plot of model’s residuals. The residuals are following a normal distribution and seem to be not correlated at all for most lags. Hence the model does a decent job in estimating the time series. The Ljung-Box test on residuals also interpret the same.

Ljung-Box test

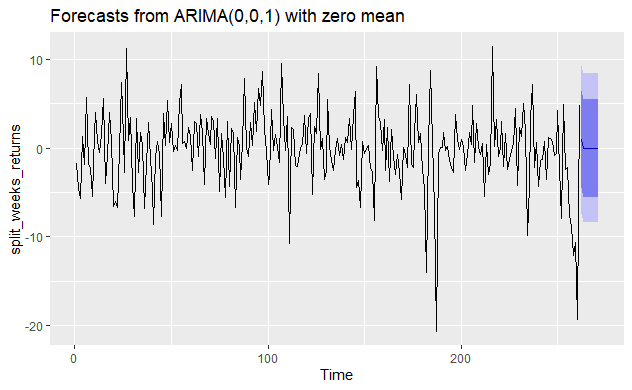
data: Residuals from ARIMA(0,0,1) with zero mean

Q\* = 8.6926, df = 9, p-value = 0.4661

Model df: 1. Total lags used: 10

### Forecasting 10 upcoming intervals using the model

Using necessary code, forecast was done for 10 upcoming weeks on the same time series and is as shown below:



The dark blue region shows the possible estimated returns over a 80 percent confidence interval and light blue region shows estimated returns over a 95 percent confidence interval.

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

262 0.9925251 -4.462695 6.447746 -7.350514 9.335565

263 0.0000000 -5.505898 5.505898 -8.420545 8.420545

264 0.0000000 -5.505898 5.505898 -8.420545 8.420545

265 0.0000000 -5.505898 5.505898 -8.420545 8.420545

266 0.0000000 -5.505898 5.505898 -8.420545 8.420545

267 0.0000000 -5.505898 5.505898 -8.420545 8.420545

268 0.0000000 -5.505898 5.505898 -8.420545 8.420545

269 0.0000000 -5.505898 5.505898 -8.420545 8.420545

270 0.0000000 -5.505898 5.505898 -8.420545 8.420545

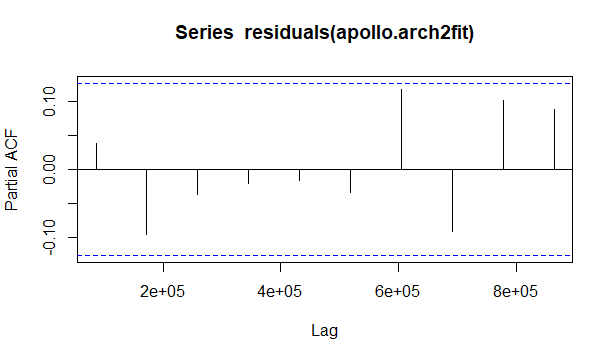
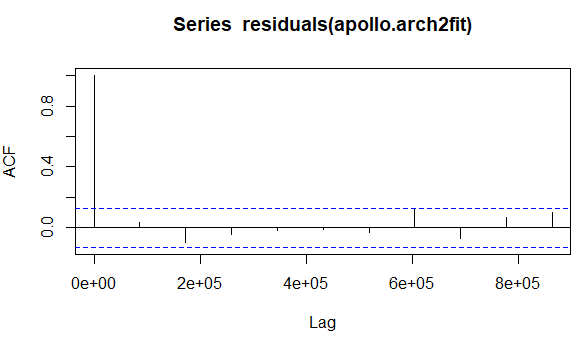
271 0.0000000 -5.505898 5.505898 -8.420545 8.420545

The model is fitted and can be used for volatility estimations.

### Volatility estimation using GARCH(1,1) and related other ARCH models

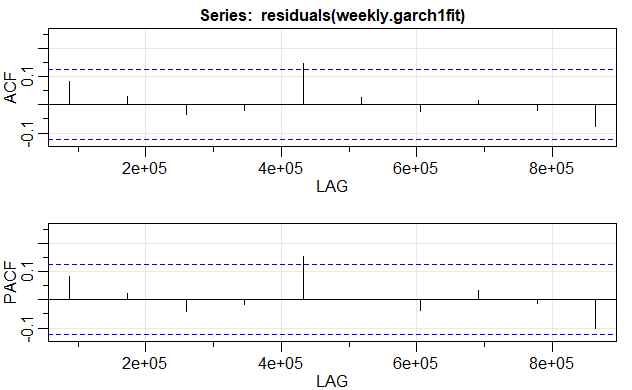
As the data is not much correlated, fitting garch models with known data can be done easily.

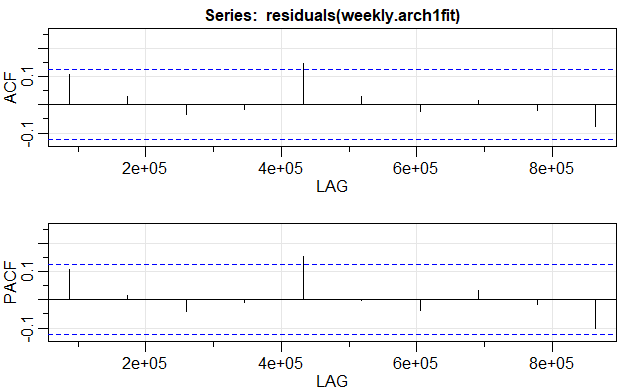
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model parameters:  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Estimate | Estimate | Estimate | | mu 0.15471 | mu 0.256699 | mu 0.253775 | | omega 4.40875 | omega 12.618932 | omega 12.177962 | | alpha1 0.19151 | alpha1 0.366360 | alpha1 0.366497 | | beta1 0.58123 |  | alpha2 0.024591 |   COMPARISION OF INFORMATION CRITERIA OF 3 MODELS:   |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Akaike 5.7066 | Akaike 5.7021 | Akaike 5.7083 | | Bayes 5.7749 | Bayes 5.7567 | Bayes 5.7766 | | Shibata 5.7059 | Shibata 5.7016 | Shibata 5.7076 | | Hannan-Quinn 5.7341 | Hannan-Quinn 5.7241 | Hannan-Quinn 5.7357 | |  | |

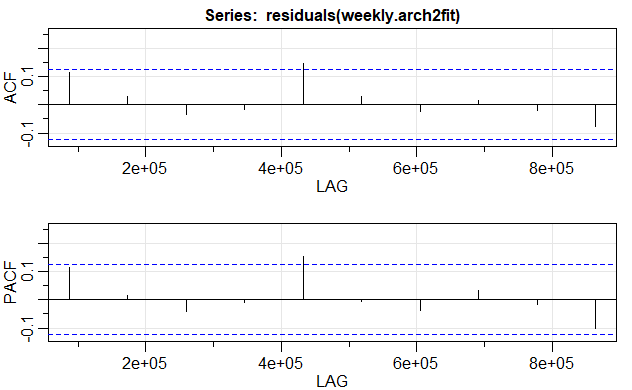
The code is given in the Appendix. The required output is tabulated above which is clearly interpreted. The Information Criteria needs to be least for the model to be best. In the above table the information criteria is the least for the ARCH (1, 0). So the arch (1,0) model is best for volatility estimation.

### Residual analysis:

We find the residuals and the sigma of the GARCH (1, 1), ARCH (1, 0) and ARCH (2, 0) then check for their stationarity by acf and pacf plots. The data needs to be stationary for various reasons because if it is not stationary the time series cannot be applied major reason being the sudden shocks can’t be absorbed by the data and get reverted back to mean. The below results clearly show that the residuals for all the models are stationary because they fall in between the blue line for almost all lags.







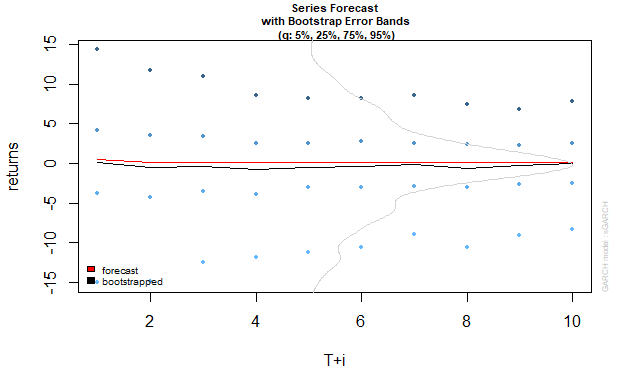
These clearly show that arch (1,0) is clearly a better fit model for volatility estimation, when compared to other models.

### Predicting next observation:

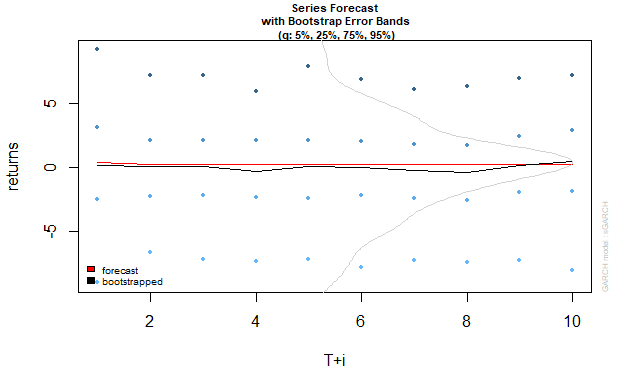
|  |  |  |
| --- | --- | --- |
| GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) |
| |  | | --- | | Sigma | | Sigma | Sigma |
| T+1 8.409571  T+2 7.684856  T+3 7.074175  T+4 6.563481  T+5 6.139827  T+6 5.791267  T+7 5.506833  T+8 5.276551  T+9 5.091477  T+10 4.943720 | T+1 4.743678  T+2 4.567597  T+3 4.501365  T+4 4.476854  T+5 4.467841  T+6 4.464534  T+7 4.463322  T+8 4.462878  T+9 4.462716  T+10 4.462656 | T+1 5.549873  T+2 4.909764  T+3 4.665843  T+4 4.555153  T+5 4.507540  T+6 4.487167  T+7 4.478493  T+8 4.474806  T+9 4.473240  T+10 4.472575 |

The forecasts for sigma show that arch (1,0) was the best at estimating the volatility pattern in the entire series. This is shown by the forecast values of sigma

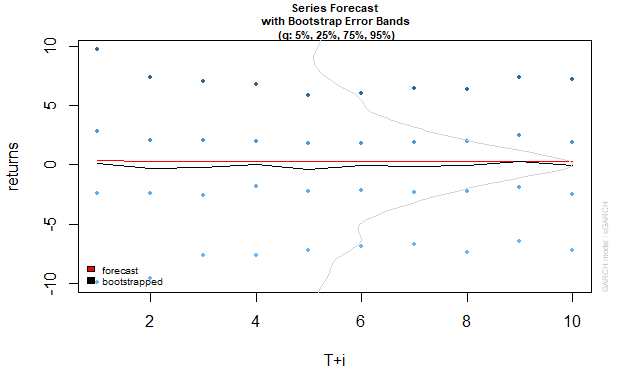
### Predicting returns using both models together



Garch(1,1) based forecast of returns, blue points showing extremes based on modelled volatility model.



Arch1 based forecast of returns, blue points showing extremes based on modelled volatility model.



Arch2 based forecast of returns, blue points showing extremes based on modelled volatility model.

### Conclusion on weekly data:

The models build above show us that arima (0,0,1) is a better fit for data, and using the same model to predict the volatility of returns, arch (1,0) gives us the best possible results in model estimation. The results are tabulated and shown above.

## VAR models on CESC:

Necessary data was taken and the results were interpreted as shown below. The factors influencing returns were: volume traded and risk-free rates. These were the data time-series used in VAR modelling. The series were taken in the following order: returns, volume, risk-free rate

### Daily Data Based:

Estimation results for equation y1:

===================================

y1 = y1.l1 + y2.l1 + const

Estimate Std. Error t value Pr(>|t|)

y1.l1 -8.508e-02 6.414e-02 -1.326 0.1860

y2.l1 -9.555e-07 5.755e-07 -1.660 0.0982 .

const 9.189e-02 2.640e-01 0.348 0.7281

---

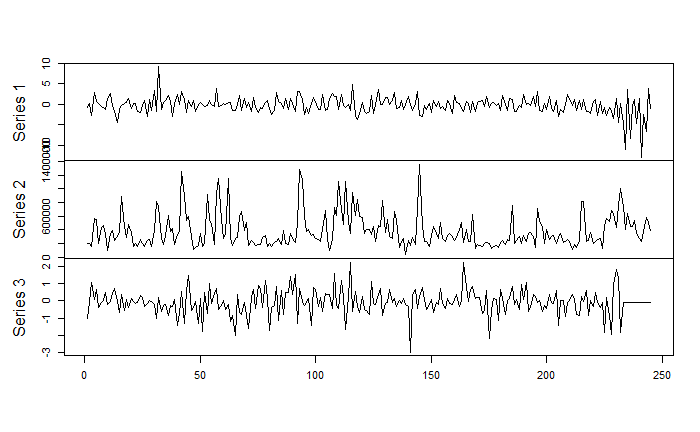
Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.178 on 241 degrees of freedom

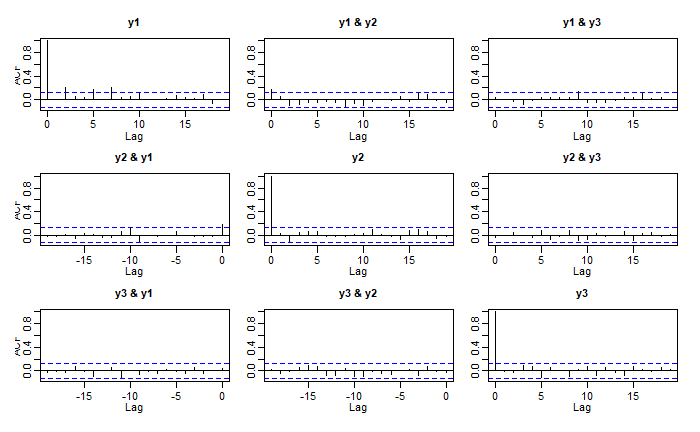
Multiple R-Squared: 0.02054, Adjusted R-squared: 0.01241

F-statistic: 2.527 on 2 and 241 DF, p-value: 0.082

series data:



Below graph shows ACF plot for residuals of model fitted.



### Monthly data Based

Estimation results for equation split\_months\_returns:

=====================================================

split\_months\_returns = split\_months\_returns.l1 + xts\_monthly\_df.xts\_monthly\_vol.1.60..l1 + L.1.60..l1 + const

Estimate Std. Error t value Pr(>|t|)

split\_months\_returns.l1 2.585e-01 1.760e-01 1.468 0.148

xts\_monthly\_df.xts\_monthly\_vol.1.60..l1 -6.381e-07 4.198e-06 -0.152 0.880

L.1.60..l1 -4.289e-01 4.208e-01 -1.019 0.313

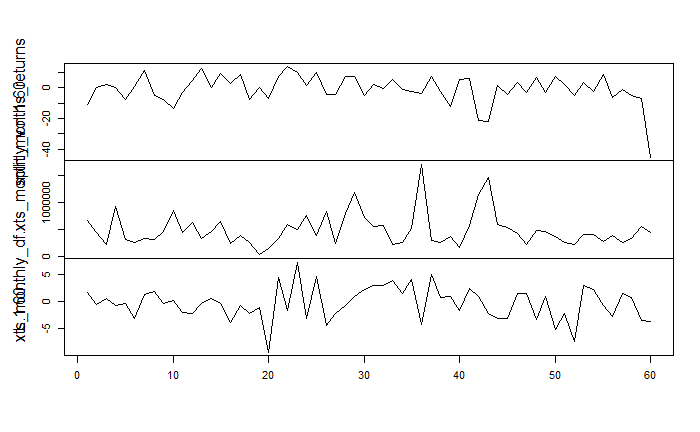
const -5.428e-01 2.415e+00 -0.225 0.823

Residual standard error: 9.46 on 55 degrees of freedom

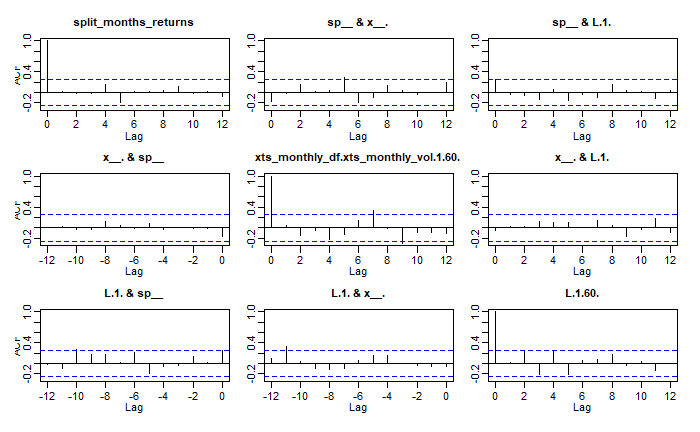
Multiple R-Squared: 0.05192, Adjusted R-squared: 0.0002087

F-statistic: 1.004 on 3 and 55 DF, p-value: 0.398

Series data:



Below graph shows acf plot of residuals for fitted model.



### Weekly data based:

Estimation results for equation split\_weeks\_returns:

====================================================

split\_weeks\_returns = split\_weeks\_returns.l1 + xts\_weekly\_df.xts\_weekly\_vol.2.262..l1 + K.1.261..l1 + const

Estimate Std. Error t value Pr(>|t|)

split\_weeks\_returns.l1 -1.747e-01 6.132e-02 -2.849 0.00474 \*\*

xts\_weekly\_df.xts\_weekly\_vol.2.262..l1 -5.832e-05 8.864e-06 -6.579 2.65e-10 \*\*\*

K.1.261..l1 1.012e+00 3.264e+00 0.310 0.75683

const 2.958e+01 6.150e+00 4.810 2.58e-06 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

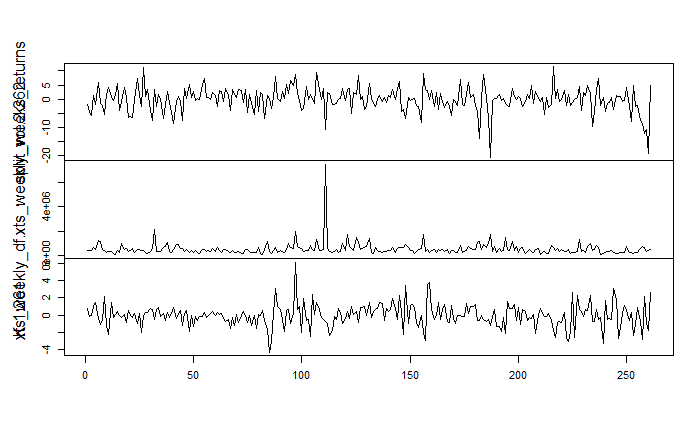
Residual standard error: 67.9 on 256 degrees of freedom

Multiple R-Squared: 0.2613, Adjusted R-squared: 0.2526

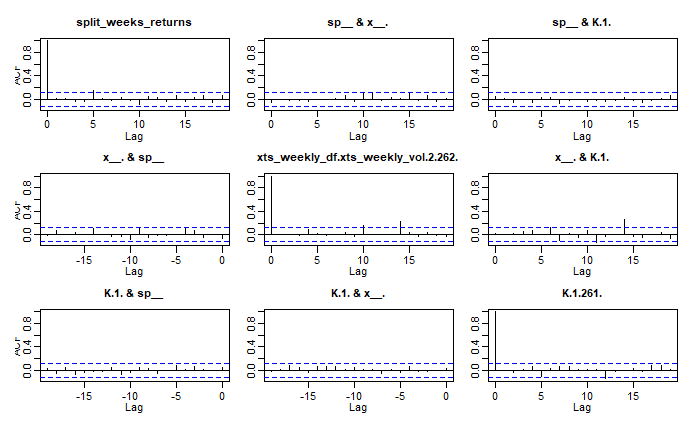
F-statistic: 30.18 on 3 and 256 DF, p-value: < 2.2e-16

F-statistic: 0.5834 on 3 and 256 DF, p-value: 0.6264

Series data.



Acf plot for residuals of fitted model:

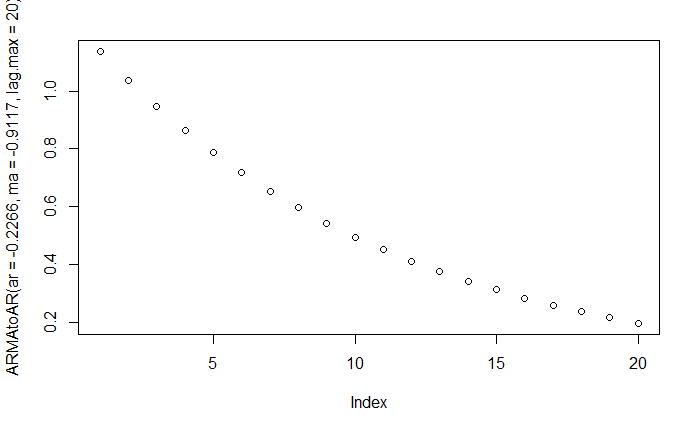


These interpretations and results conclude the VAR model interpretation of CESC returns on weekly ,monthly, daily data basis

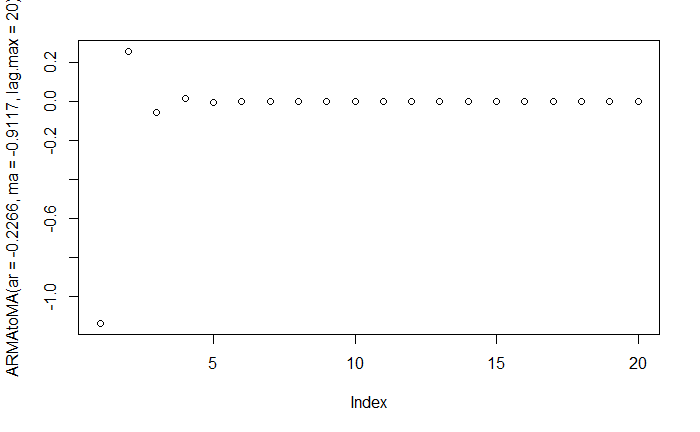
## Identifying and interpreting ARMA models of above data as AR(p),MA(q) models;

There is a function in R for the above thing to be implemented. The code is used as shown in the appendix and the required interpretations are mentioned here as points plotted:

Daily data arima(1,1,1): when used after differencing:

ARMA to AR:  


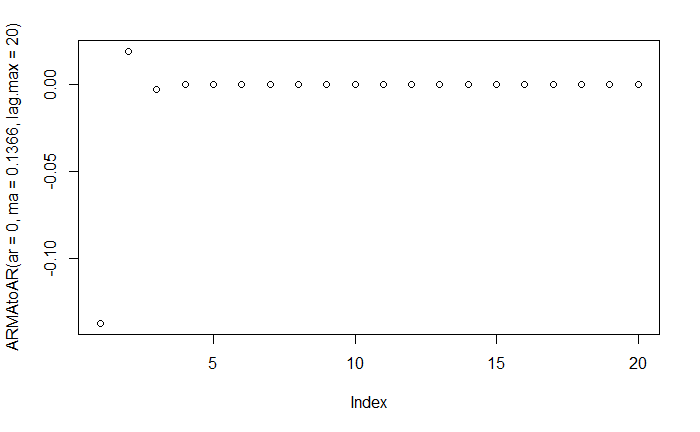
ARMA to MA



Note: as monthly data seemed to be white noise by autofit model, it is of no use to interpret the model in AR/MA terms.

Weekly data: arima (0,0,1)

As it is already MA model, we will interpret it as AR

ARMA to AR 

These finally conclude all the necessary interpretations for CESC data. We will continue about the other company in the next part of the report.

# Part-2 CADILAHC



## Cadila Healthcare

## History and briefing about the company:

**Cadila Healthcare Ltd (Zydus Cadila)** is an Indian pharmaceutical company headquartered at Ahmedabad in Gujarat state of western India. With in-depth domain expertise in the field of healthcare, it has strong capabilities across the spectrum of the pharmaceutical value chain. From formulations to active pharmaceutical ingredients and animal healthcare products to wellness products, Zydus has earned a reputation amongst Indian pharmaceutical companies for providing comprehensive and complete healthcare solutions. The company is one of the leading pharmaceutical companies in India, with INR 119.05 Billion revenue (2018). Cadila was founded in 1952 by Ramanbhai Patel (1925–2001), formerly a lecturer in the L.M. College of Pharmacy, and his business partner Indravadan Modi. It evolved over the next four decades into an established pharmaceutical company. In 1995 the Patel and Modi families split; the Modi family's share was moved into a new company called Cadila Pharmaceuticals Ltd., and Cadila Healthcare Ltd became the Patel family's holding company. Cadila has launched many of first a kind generic drugs many times

## Products:

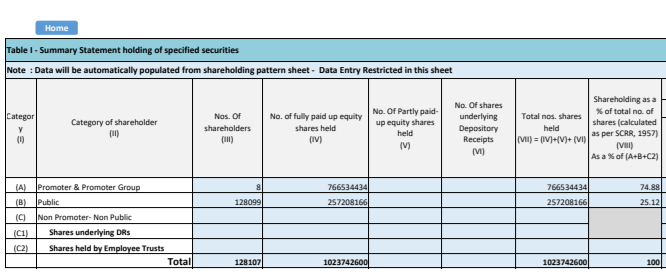
From nine pharmaceutical production operations in India as well as a Zydus Cadila develops and manufactures a large range of pharmaceuticals as well as diagnostics, herbal products, skin care products and other OTC products. Starting from late 2015, having concluded a voluntary license agreement with Gilead, the company also produces the generics for hepatitis C treatment (i.e. sofosbuvir, distributed under the brand name SoviHep).

## Business Nature:

Cadila healthcareis one of India’s leading generic drug manufacturer and formula developer. It is clearly evident from their financials and their contribution towards the pharma industry in the country and even in the world.

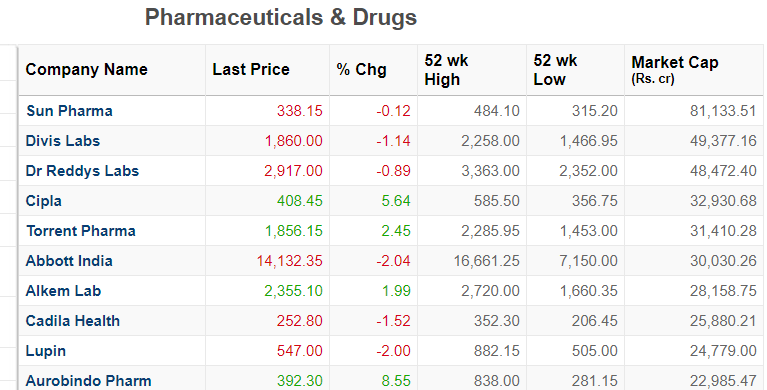
## Ownership:

Cadila healthcare is a prominent pharmaceutical and a Public company. This is clearly reflected in the number of shares and their market capitalization, which show the value of the company in the investor’s perspective. Below snapshot of shareholding pattern indicate the same.



## Industry significance:

The company, though has been through many ups and downs since its inception, tried to maintain its significance in Indian energy production and distribution industry. The market cap of the company at present, reflects the same.



## Overall greatness:

* They were given many awards for their excellence in

🡺innovation

🡺emerging player in the sector

🡺global manufacturer

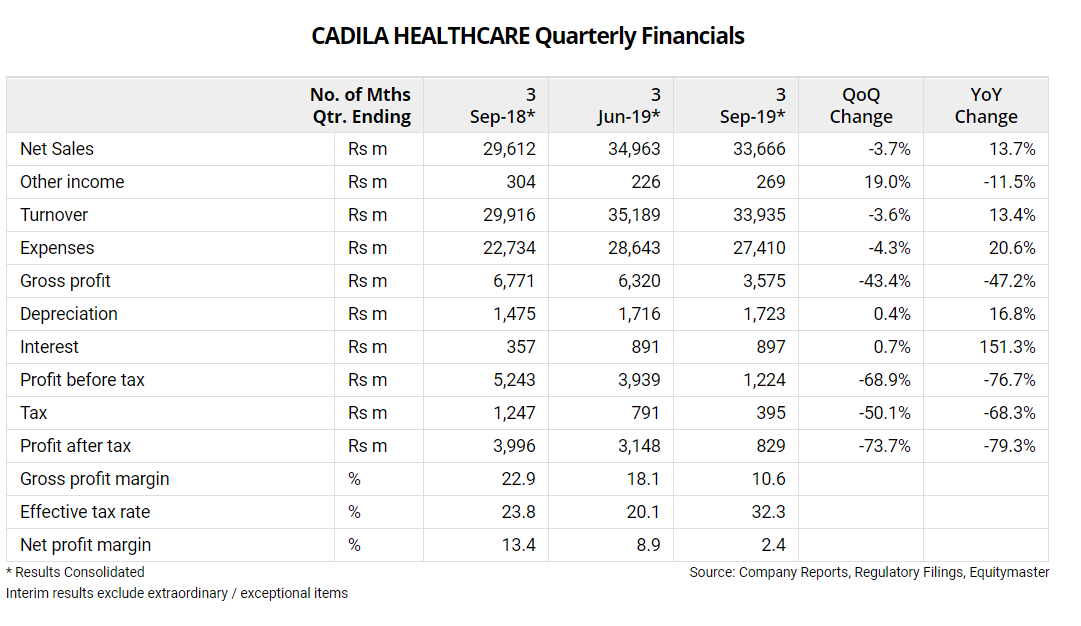
* They wish to:

🡺create healthier communities globally

🡺open up new pathways through innovation and be a research-based

company by 2020

* Though they are passing through a rough phase, their recently updated quarterly financials also convey the same.



## References:

<https://zyduscadila.com/> for company related, shareholding pattern, history, awards related data.

<https://en.wikipedia.org/wiki/Cadila_Healthcare> , <https://www.moneycontrol.com/india/stockpricequote/pharmaceuticals/cadilahealthcare/CHC>

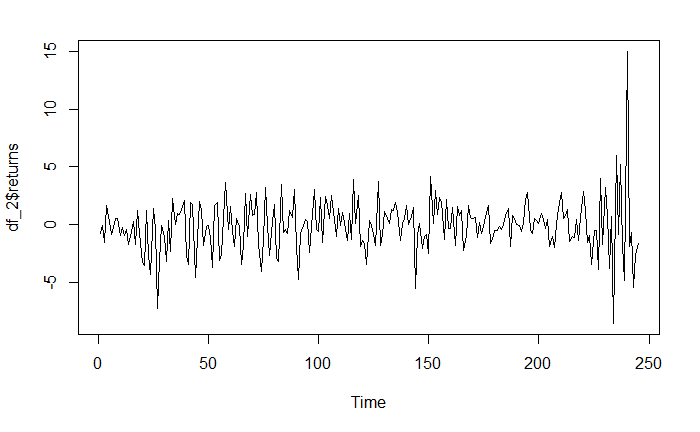
for general, products reference.

<https://www.equitymaster.com/stockquotes/complist.asp?company=cadilahc> for YoY comparison of quarterly financial of company.

<https://www.moneycontrol.com/stocks/marketinfo/marketcap/bse/pharmaceuticals-drugs.html> for industrial significance.

# ARIMA, GARCH, ARCH models for CESC

## Daily data based



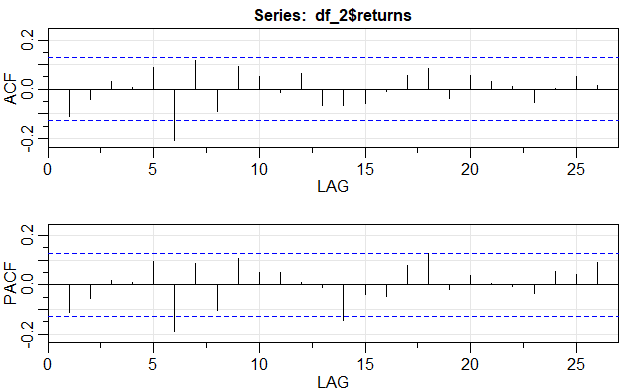
Daily returns are plotted as time series and the plot is as shown above:

Summary of above data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| -8.5709 | -1.3353 | -0.0814 | -0.1295 | 1.1236 | 15.0571 |

### ACF&PACF of data

now let us try fit an arima model to the data. Let us interpret the acf, pacf of data:



### Interpretation

as the data seems to show some trend, which is explained by the visible significant spikes at lags 6 in acf, and pacf, we shall use auto.arima to fit an approximate model to this data and use it for forecasting.

### Fitting Arima model

Using auto. arima function, arima model is fitted to the above data. The fitted model related parameters are as mentioned below.

Series: df\_2$returns

ARIMA(1,0,2) with zero mean

|  |  |  |  |
| --- | --- | --- | --- |
|  | ar1 | ma1 | ma2 |
| Coefficients: | -0.7943 | 0.7224 | -0.1751 |
| s.e. | 0.0786 | 0.0918 | 0.0639 |

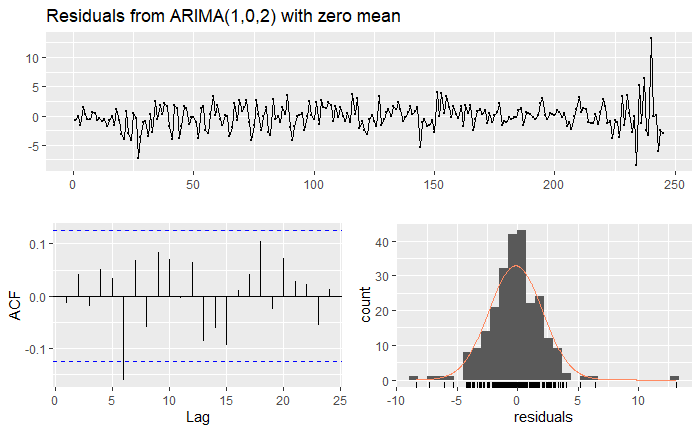
sigma^2 estimated as 4.942: log likelihood=-541.97

AIC=1091.93 AICc=1092.1 BIC=1105.94

Training set error measures:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
| -0.1189344 | 2.097331 | 1.522918 | 30.12139 | 187.2415 | 0.6499132 | 0.004117542 |

### Checking residuals of model fitted:



The fitted model’s residuals seem to do a decent interpretation of entire time series, and is clearly shown in the acf plot of model’s residuals. The residuals are following a normal distribution and seem to be not correlated at all for most lags. Hence the model does a decent job in estimating the time series. The Ljung-Box test on residuals also interpret the same.

Ljung-Box test

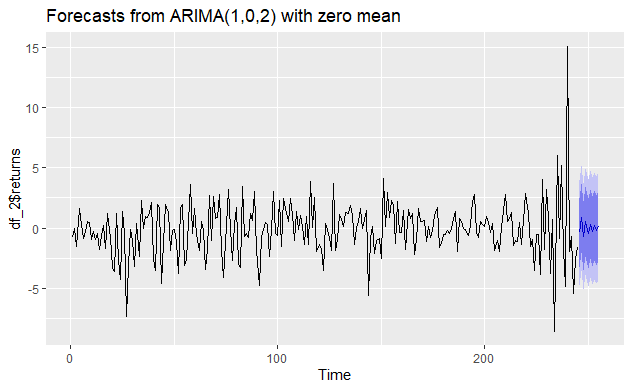
data: Residuals from ARIMA(1,0,2) with zero mean

Q\* = 13.196, df = 7, p-value = 0.06748

Model df: 3. Total lags used: 10

### Forecasting 10 upcoming intervals using the model

Using necessary code, forecast was done for 10 upcoming days on the same time series and is as shown below:



The dark blue region shows the possible estimated returns over a 80 percent confidence interval and light blue region shows estimated returns over a 95 percent confidence interval.

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

246 -0.4038074 -3.252809 2.445194 -4.760979 3.953364

247 0.8395651 -2.016785 3.695915 -3.528846 5.207976

248 -0.6668575 -3.542941 2.209226 -5.065448 3.731733

249 0.5296777 -2.358787 3.418142 -3.887847 4.947203

250 -0.4207172 -3.316965 2.475531 -4.850146 4.008712

251 0.3341711 -2.566977 3.235319 -4.102752 4.771094

252 -0.2654285 -3.169663 2.638806 -4.707072 4.176215

253 0.2108270 -2.695354 3.117008 -4.233793 4.655447

254 -0.1674576 -3.074865 2.739950 -4.613954 4.279039

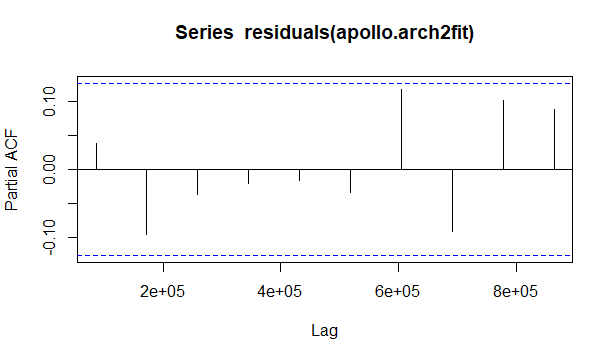
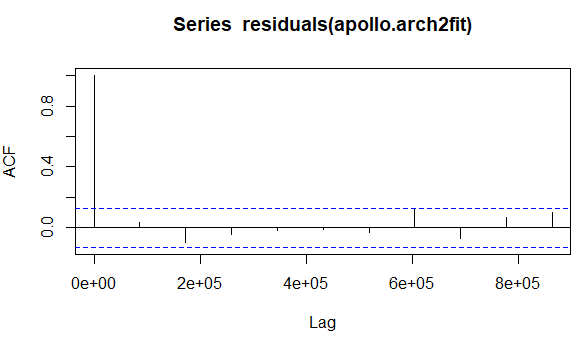
255 0.1330097 -2.775172 3.041191 -4.314670 4.580690

The model is fitted and can be used for volatility estimations.

### Volatility estimation using GARCH(1,1) and related other ARCH models

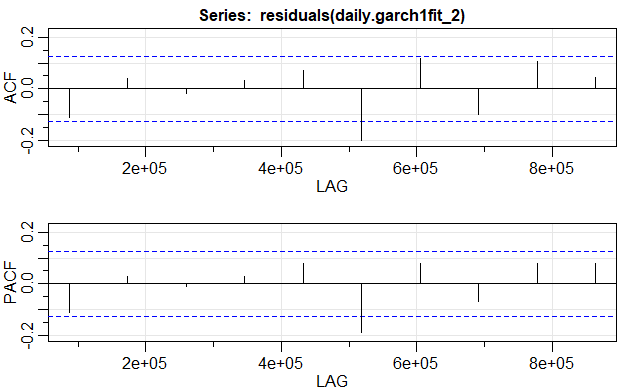
As the data is not much correlated, fitting garch models with known data can be done easily.

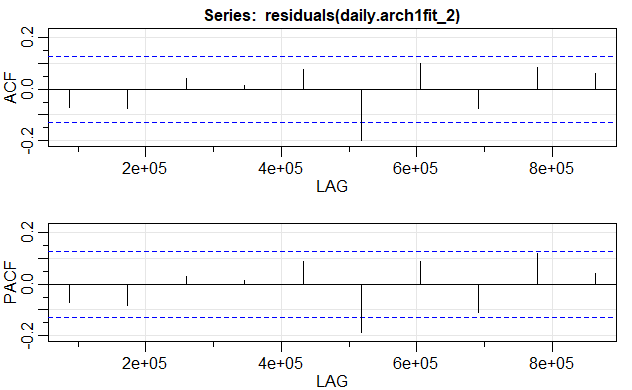
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model parameters:  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Estimate | Estimate | Estimate | | mu -0.048857 | mu -0.064726 | mu -0.675510 | | omega 0.165845 | omega 3.788342 | omega 1.697005 | | alpha1 0.168341 | alpha1 0.275381 | alpha1 0.374301 | | beta1 0.821250 |  | alpha2 0.624698 |   COMPARISION OF INFORMATION CRITERIA OF 3 MODELS:   |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Akaike 4.2758 | Akaike 4.4368 | Akaike 4.3720 | | Bayes 4.3759 | Bayes 4.5226 | Bayes 4.4720 | | Shibata 4.2743 | Shibata 4.4357 | Shibata 4.3704 | | Hannan-Quinn 4.3161 | Hannan-Quinn 4.4714 | Hannan-Quinn 4.4123 | |  | |

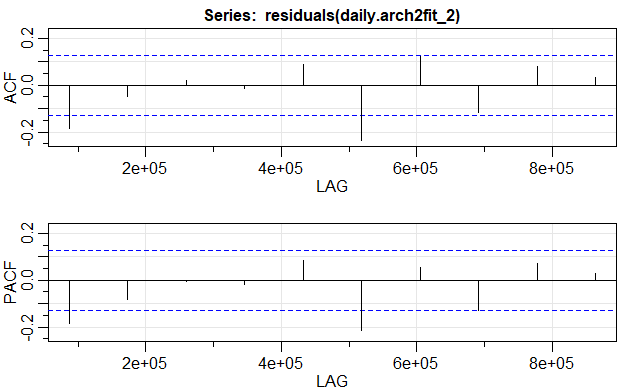
The code is given in the Appendix. The required output is tabulated above which is clearly interpreted. The Information Criteria needs to be least for the model to be best. In the above table the information criteria is the least for the GARCH (1, 1). So the arch (1,0) model is best for volatility estimation.

### Residual analysis:

We find the residuals and the sigma of the GARCH (1, 1), ARCH (1, 0) and ARCH (2, 0) then check for their stationarity by acf and pacf plots. The data needs to be stationary for various reasons because if it is not stationary the time series cannot be applied major reason being the sudden shocks can’t be absorbed by the data and get reverted back to mean. The below results clearly show that the residuals for all the models are stationary because they fall in between the blue line for almost all lags.







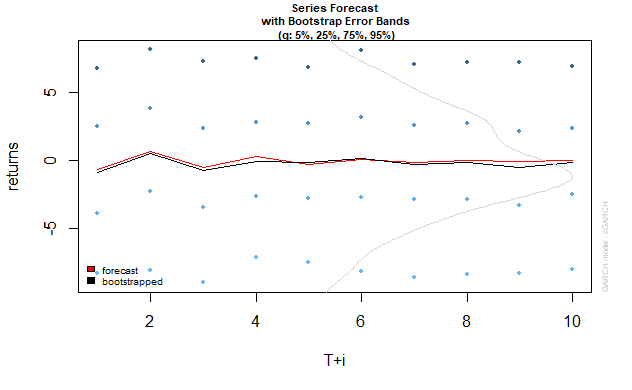
These clearly show that garch (1,1) is clearly a better fit model for volatility estimation, when compared to other models.

### Predicting next observation:

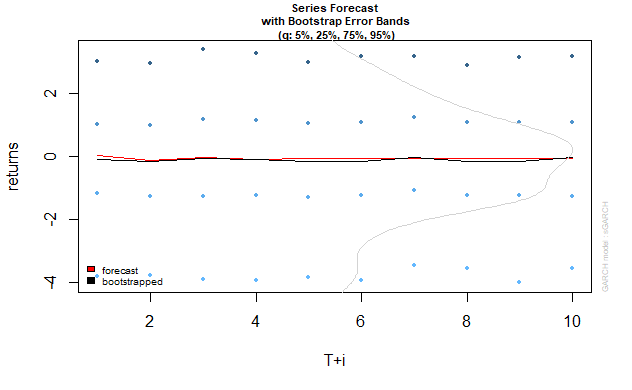
|  |  |  |
| --- | --- | --- |
| GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) |
| |  | | --- | | Sigma | | Sigma | Sigma |
| T+1 4.979644  T+2 4.970370  T+3 4.961177  T+4 4.952062  T+5 4.943025  T+6 4.934067  T+7 4.925185  T+8 4.916380  T+9 4.907651  T+10 4.898998 | T+1 2.089581  T+2 2.233999  T+3 2.272158  T+4 2.282554  T+5 2.285408  T+6 2.286194  T+7 2.286410  T+8 2.286470  T+9 2.286486  T+10 2.286491 | T+1 2.417632  T+2 2.282243  T+3 2.701467  T+4 2.771723  T+5 3.021845  T+6 3.148675  T+7 3.333517  T+8 3.471270  T+9 3.626165  T+10 3.761136 |

The forecasts for sigma show that garch (1,1) was the best at estimating the volatility pattern in the entire series. This is shown by the forecast values of sigma

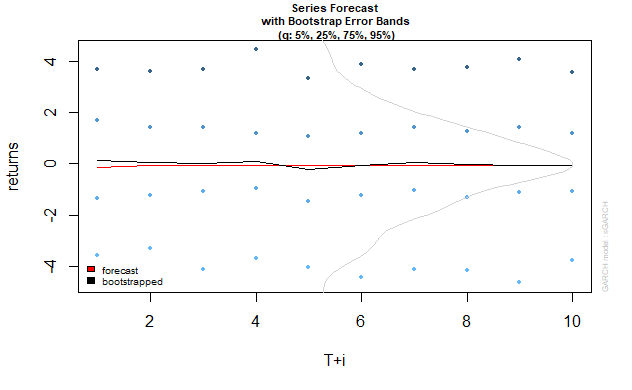
### Predicting returns using both models together



Garch(1,1) based forecast of returns, blue points showing extremes based on modelled volatility model.



Arch1 based forecast of returns, blue points showing extremes based on modelled volatility model.

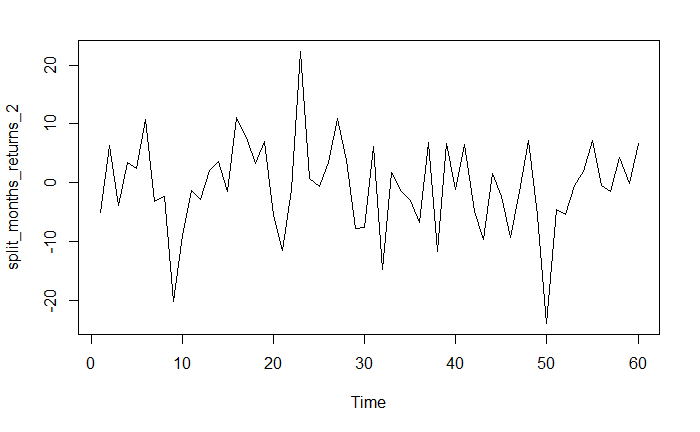


Arch2 based forecast of returns, blue points showing extremes based on modelled volatility model.

### Conclusion on daily data:

The models build above show us that arima (1,0,2) is a better fit for data, and using the same model to predict the volatility of returns, garch (1,1) gives us the best possible results in model estimation. The results are tabulated and shown above.

## Monthly data based



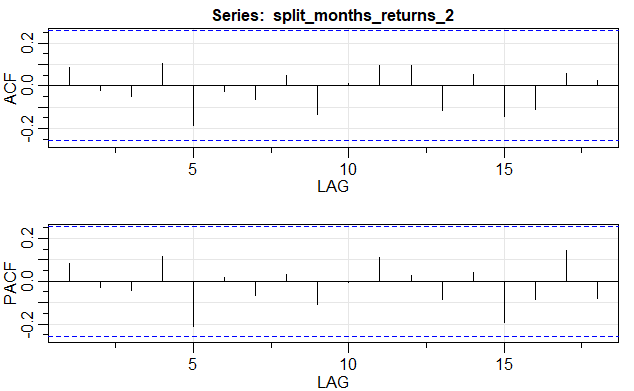
Monthly returns are plotted as time series and the plot is as shown above:

Summary of above data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| -23.8826 | -4.9621 | -0.8603 | -0.6134 | 3.8840 | 22.3570 |

### ACF&PACF of data

now let us try fit an arima model to the data. Let us interpret the acf, pacf of data:



### Interpretation

as the data seems to show no trend, which is explained by no visible significant spikes at all lags in acf, and pacf, the data is just white noise and we shall use auto.arima to show the same.

### Fitting Arima model

Using auto. arima function, arima model is fitted to the above data. The fitted model related parameters are as mentioned below.

Series: split\_months\_returns\_2

ARIMA(0,0,0) with zero mean

sigma^2 estimated as 59.04: log likelihood=-207.48

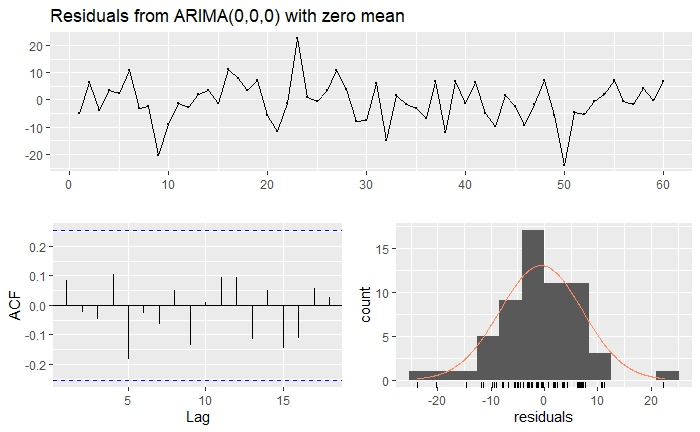
AIC=416.97 AICc=417.04 BIC=419.06

|  |  |  |
| --- | --- | --- |
|  | ar1 | ma1 |
| Coefficients: | 0 | 0 |
| s.e. | 0 | 0 |

Training set error measures:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
| -0.6134326 | 7.683855 | 5.795936 | 100 | 100 | 0.680223 | 0.08477927 |

### Checking residuals of model fitted:



The fitted model’s residuals seem to do a decent interpretation of entire time series, and is clearly shown in the acf plot of model’s residuals. The residuals are following a normal distribution and seem to be not correlated at all. Hence the model does a decent job in estimating the time series, ultimately saying that the entire data is just white noise, and can’t be used for predictions. The Ljung-Box test on residuals also interpret the same.

Ljung-Box test

data: Residuals from ARIMA(0,0,0) with zero mean

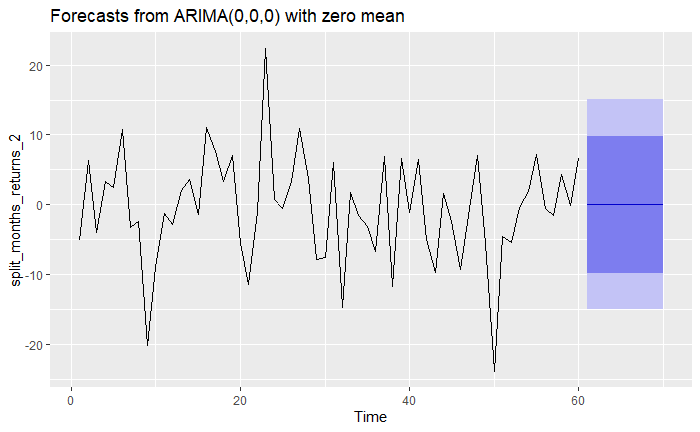
Q\* = 5.5214, df = 10, p-value = 0.8537

Model df: 0. Total lags used: 10

### Forecasting 10 upcoming intervals using the model

Using necessary code, forecast was done for 10 upcoming months on the same time series and is as shown below:

It is of no use to predict, and this is shown by the graph.



The dark blue region shows the possible estimated returns over a 80 percent confidence interval and light blue region shows estimated returns over a 95 percent confidence interval.

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

61 0 -9.847256 9.847256 -15.06008 15.06008

62 0 -9.847256 9.847256 -15.06008 15.06008

63 0 -9.847256 9.847256 -15.06008 15.06008

64 0 -9.847256 9.847256 -15.06008 15.06008

65 0 -9.847256 9.847256 -15.06008 15.06008

66 0 -9.847256 9.847256 -15.06008 15.06008

67 0 -9.847256 9.847256 -15.06008 15.06008

68 0 -9.847256 9.847256 -15.06008 15.06008

69 0 -9.847256 9.847256 -15.06008 15.06008

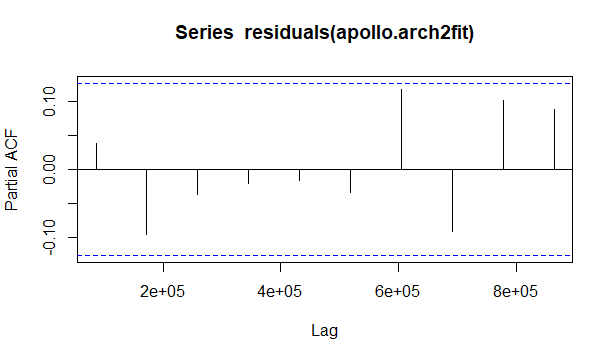
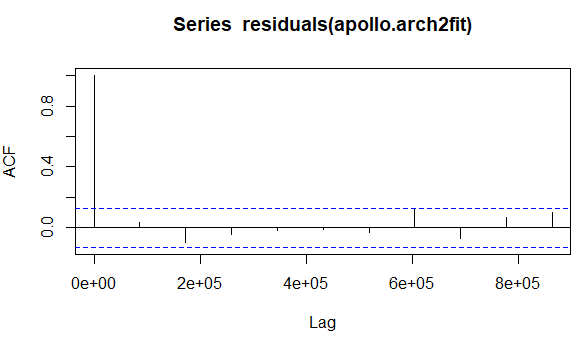
70 0 -9.847256 9.847256 -15.06008 15.06008

The model is fitted and can be used for volatility estimations.

### Volatility estimation using GARCH(1,1) and related other ARCH models

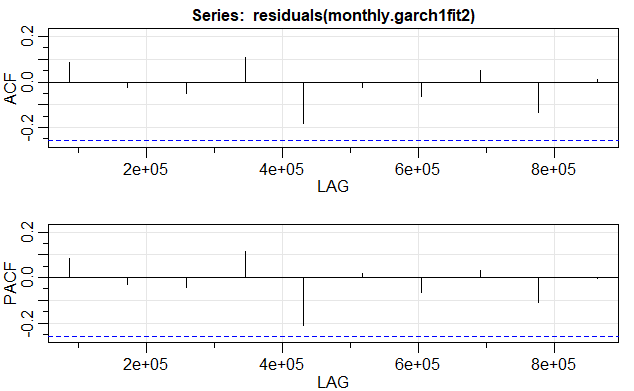
As the data is pure white noise, it cannot be used for any estimations, in addition to this problem, the no. of data points are so low, and the functions to converge the data into a model don’t seem to work. So, the models formed can be of no use for us. Nevertheless, we shall complete the asked task here.,

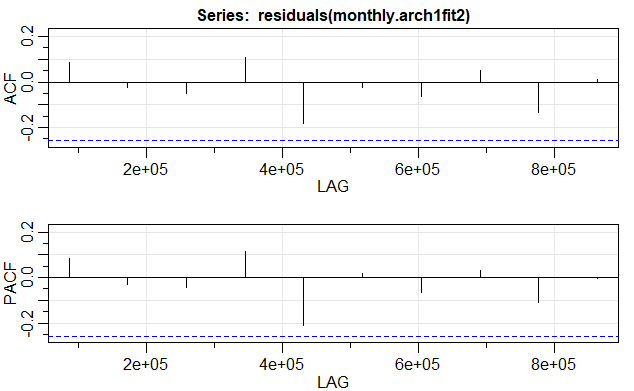
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model parameters:  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Estimate | Estimate | Estimate | | mu -0.62402 | mu -0.61428 | mu -0.61283 | | omega 0.00004 | omega 59.32354 | omega 59.52024 | | alpha1 0.00000 | alpha1 0.99900 | alpha1 0.00000 | | beta1 0.99889 |  | alpha2 0.00000 |   COMPARISION OF INFORMATION CRITERIA OF 3 MODELS:   |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Akaike 7.0424 | Akaike 7.0097 | Akaike 7.0430 | | Bayes 7.1820 | Bayes 7.1144 | Bayes 7.1826 | | Shibata 7.0342 | Shibata 7.0050 | Shibata 7.0348 | | Hannan-Quinn 7.0970 | Hannan-Quinn 7.0506 | Hannan-Quinn 7.0976 | |  | |

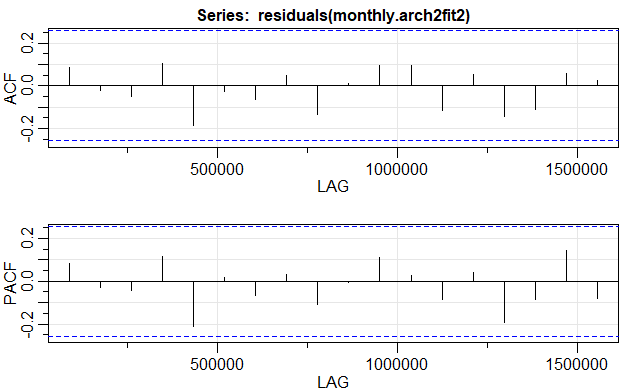
The code is given in the Appendix. The required output is tabulated above which is clearly interpreted. The Information Criteria needs to be least for the model to be best. But if we go by basic knowledge of finance, monthly volatility can never be more than at the most 20 percent. The models fitted have arch models with better aic values than of garch, but have more volatility in forecasts. So there is no use of these models. The forecasts show the same below

### Residual analysis:

We find the residuals and the sigma of the GARCH (1, 1), ARCH (1, 0) and ARCH (2, 0) then check for their stationarity by acf and pacf plots. The data needs to be stationary for various reasons because if it is not stationary the time series cannot be applied major reason being the sudden shocks can’t be absorbed by the data and get reverted back to mean. The below results clearly show that the residuals for all the models are stationary because they fall in between the blue line.







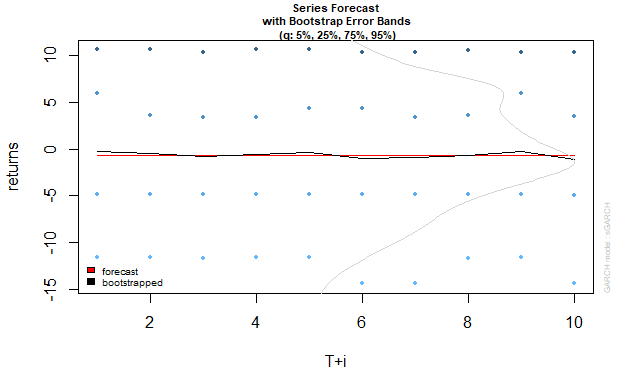
The points explained above, clearly show that no model can be used to forecast volatility of monthly intervals.

### Predicting next observation:

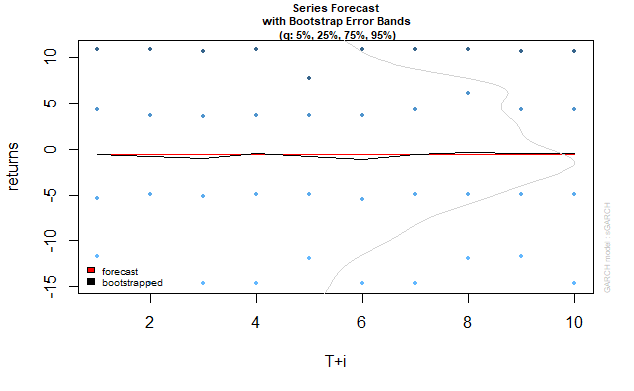
|  |  |  |
| --- | --- | --- |
| GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) |
| |  | | --- | | Sigma | | Sigma | Sigma |
| T+1 7.408007  T+2 7.403889  T+3 7.399773  T+4 7.395660  T+5 7.391549  T+6 7.387440  T+7 7.383333  T+8 7.379229  T+9 7.375127  T+10 7.371027 | T+1 7.702178  T+2 7.702178  T+3 7.702178  T+4 7.702178  T+5 7.702178  T+6 7.702178  T+7 7.702178  T+8 7.702178  T+9 7.702178  T+10 7.702178 | T+1 7.714936  T+2 7.714936  T+3 7.714936  T+4 7.714936  T+5 7.714936  T+6 7.714936  T+7 7.714936  T+8 7.714936  T+9 7.714936  T+10 7.714936 |

The forecasts for sigma show that no model is suitable for predicting volatility of stocks returns on monthly basis

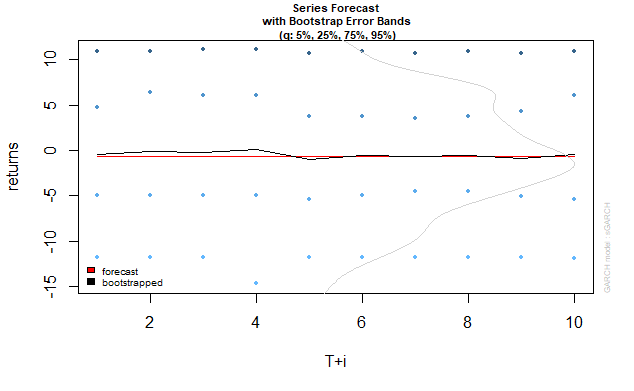
### Predicting returns using both models together

As stated above, these don’t give true picture and cannot be used. But for sake of completeness, we shall report them here

Garch(1,1) based forecast of returns, blue points showing extremes based on modelled volatility model.



Arch1 based forecast of returns, blue points showing extremes based on modelled volatility model.

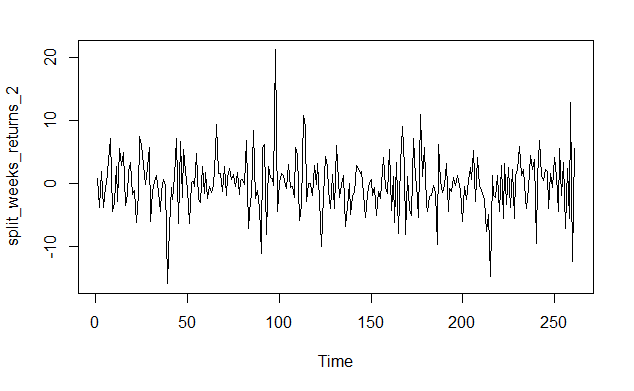


Arch2 based forecast of returns, blue points showing extremes based on modelled volatility model.

### Conclusion on Monthly data:

The models build above show us that white noise perfectly describes the data, and no model can be used for estimating the volatility trends in the same. The results calculated and shown above represent the same.

## Weekly data based



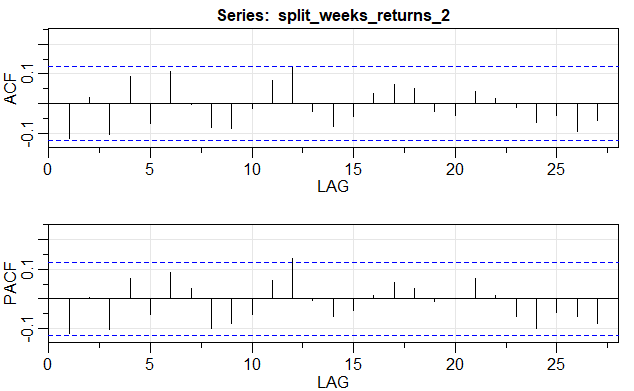
weekly returns are plotted as time series and the plot is as shown above:

Summary of above data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| -15.9034 | -2.3658 | -0.1021 | -0.1012 | 2.2998 | 21.1930 |

### ACF&PACF of data

now let us try fit an arima model to the data. Let us interpret the acf, pacf of data:



### Interpretation

as the data seems to show some trend, which is explained by the visible significant spikes at lags 12 in acf, and pacf, we shall use auto.arima to fit an approximate model to this data and use it for forecasting.

### Fitting Arima model

Using auto. arima function, arima model is fitted to the above data. The fitted model related parameters are as mentioned below.

Series: split\_weeks\_returns\_2

ARIMA(1,0,1) with zero mean.

sigma^2 estimated as 19.09: log likelihood=-754.22

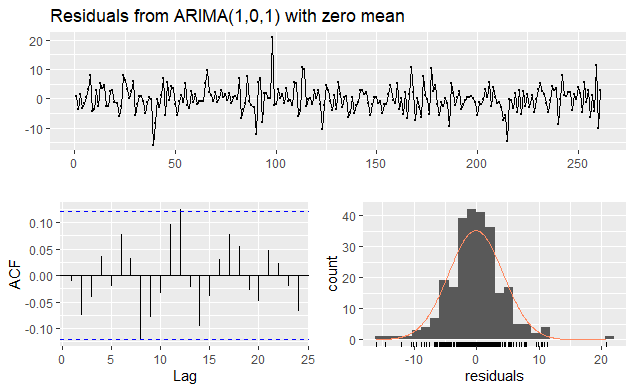
AIC=1514.44 AICc=1514.54 BIC=1525.14

|  |  |  |
| --- | --- | --- |
|  | ar1 | ma1 |
| Coefficients: | -0.8739 | 0.7854 |
| s.e. | 0.1076 | 0.1297 |

Training set error measures:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ME | RMSE | MAE | MPE | MAPE | MASE | ACF1 |
| 0.1315869 | -0.1120408 | 4.352024 | 3.204126 | 72.94993 | 0.6462954 | -0.01200717 |

### Checking residuals of model fitted:



The fitted model’s residuals seem to do a decent interpretation of entire time series, and is clearly shown in the acf plot of model’s residuals. The residuals are following a normal distribution and seem to be not correlated at all for most lags. Hence the model does a decent job in estimating the time series. The Ljung-Box test on residuals also interpret the same.

Ljung-Box test

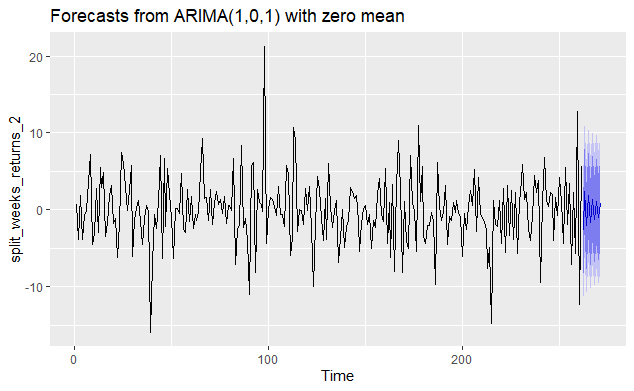
data: Residuals from ARIMA(1,0,1) with zero mean

Q\* = 10.215, df = 8, p-value = 0.2502

Model df: 2. Total lags used: 10

### Forecasting 10 upcoming intervals using the model

Using necessary code, forecast was done for 10 upcoming weeks on the same time series and is as shown below:



The dark blue region shows the possible estimated returns over a 80 percent confidence interval and light blue region shows estimated returns over a 95 percent confidence interval.

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

262 -2.7348429 -8.333679 2.863994 -11.297524 5.827839

263 2.3898927 -3.230816 8.010602 -6.206240 10.986026

264 -2.0884516 -7.725807 3.548903 -10.710042 6.533139

265 1.8250317 -3.825002 7.475065 -6.815949 10.466012

266 -1.5948375 -7.254534 4.064859 -10.250596 7.060921

267 1.3936780 -4.273386 7.060742 -7.273348 10.060704

268 -1.2178911 -6.890575 4.454793 -9.893513 7.457730

269 1.0642765 -4.612695 6.741248 -7.617903 9.746456

270 -0.9300376 -6.610282 4.750207 -9.617221 7.757146

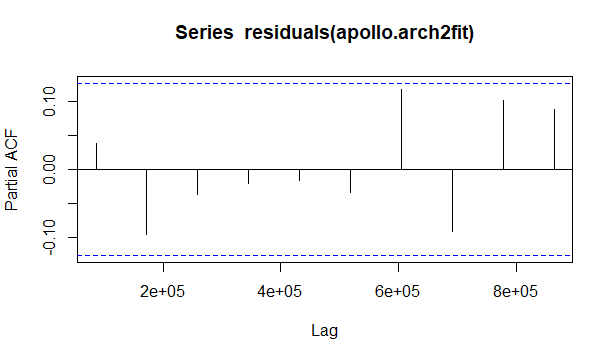
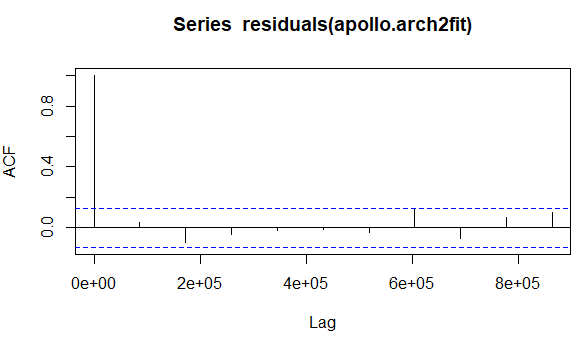
271 0.8127304 -4.870011 6.495472 -7.878273 9.503734

The model is fitted and can be used for volatility estimations.

### Volatility estimation using GARCH(1,1) and related other ARCH models

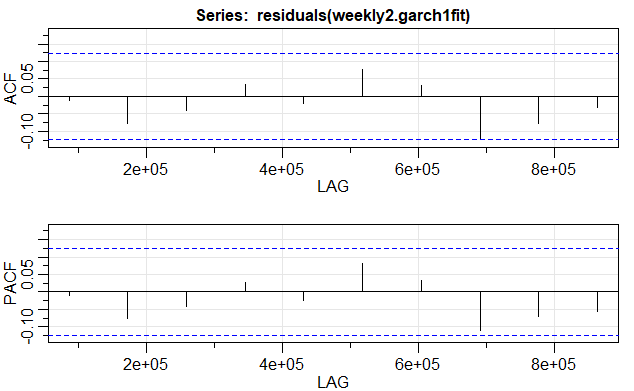
As the data is not much correlated, fitting garch models with known data can be done easily.

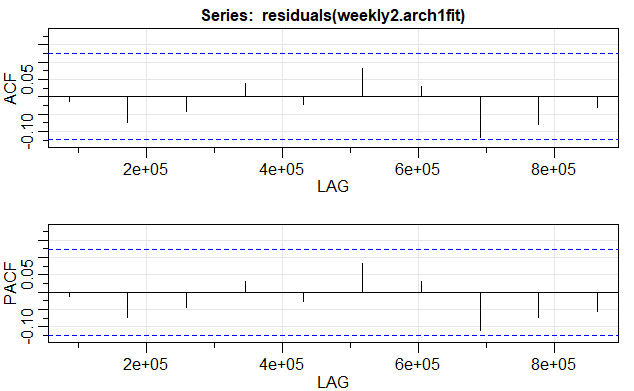
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Model parameters:  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Estimate | Estimate | Estimate | | mu -0.108060 | mu -0.094688 | mu -0.083992 | | omega 0.019708 | omega 17.918042 | omega 17.828910 | | alpha1 0.000000 | alpha1 0.058471 | alpha1 0.088547 | | beta1 0.998939 |  | alpha2 0.000000 |   COMPARISION OF INFORMATION CRITERIA OF 3 MODELS:   |  |  |  | | --- | --- | --- | | GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) | | Akaike 5.8046 | Akaike 5.8128 | Akaike 5.8407 | | Bayes 5.8066 | Bayes 5.8811 | Bayes 5.9227 | | Shibata 5.8036 | Shibata 5.8121 | Shibata 5.8397 | | Hannan-Quinn 5.8376 | Hannan-Quinn 5.8403 | Hannan-Quinn 5.8737 | |  | |

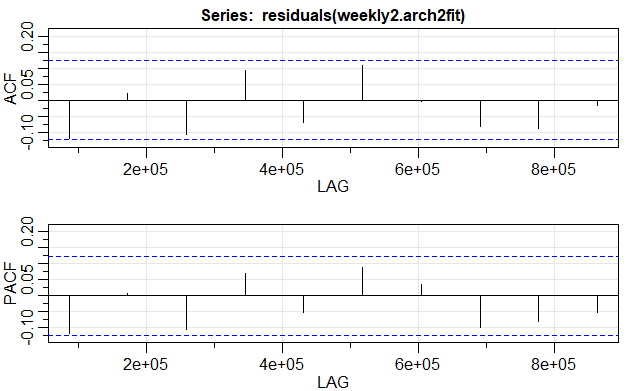
The code is given in the Appendix. The required output is tabulated above which is clearly interpreted. The Information Criteria needs to be least for the model to be best. In the above table the information criteria is the least for the ARCH (1, 0). So the arch (1,0) model is best for volatility estimation.

### Residual analysis:

We find the residuals and the sigma of the GARCH (1, 1), ARCH (1, 0) and ARCH (2, 0) then check for their stationarity by acf and pacf plots. The data needs to be stationary for various reasons because if it is not stationary the time series cannot be applied major reason being the sudden shocks can’t be absorbed by the data and get reverted back to mean. The below results clearly show that the residuals for all the models are stationary because they fall in between the blue line for almost all lags.







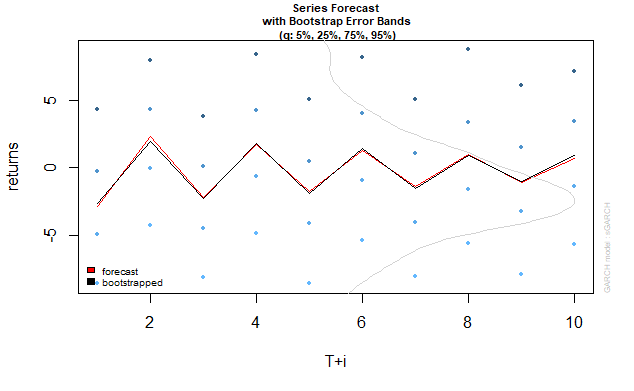
These clearly show that garch (1,1) is clearly a better fit model for volatility estimation, when compared to other models.

### Predicting next observation:

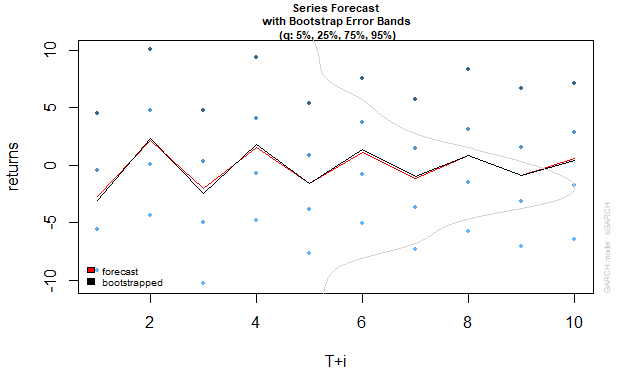
|  |  |  |
| --- | --- | --- |
| GARCH (1, 1) | ARCH (1, 0) | ARCH (2, 0) |
| |  | | --- | | Sigma | | Sigma | Sigma |
| T+1 4.341148  T+2 4.341116  T+3 4.341083  T+4 4.341051  T+5 4.341018  T+6 4.340986  T+7 4.340953  T+8 4.340921  T+9 4.340888  T+10 4.340856 | T+1 4.291749  T+2 4.358329  T+3 4.362190  T+4 4.362416  T+5 4.362429  T+6 4.362430  T+7 4.362430  T+8 4.362430  T+9 4.362430  T+10 4.362430 | T+1 4.563192  T+2 4.435392  T+3 4.423898  T+4 4.422878  T+5 4.422788  T+6 4.422780  T+7 4.422779  T+8 4.422779  T+9 4.422779  T+10 4.422779 |

The forecasts for sigma show that garch (1,1) was the best at estimating the volatility pattern in the entire series. This is shown by the forecast values of sigma

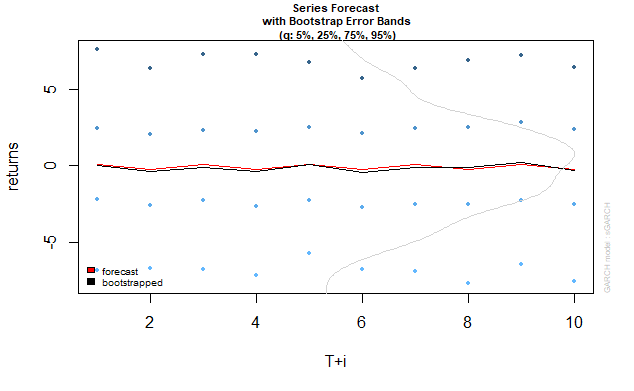
### Predicting returns using both models together



Garch(1,1) based forecast of returns, blue points showing extremes based on modelled volatility model.



Arch1 based forecast of returns, blue points showing extremes based on modelled volatility model.



Arch2 based forecast of returns, blue points showing extremes based on modelled volatility model.

### Conclusion on weekly data:

The models build above show us that arima (1,0,1) is a better fit for data, and using the same model to predict the volatility of returns, garch (1,1) gives us the best possible results in model estimation. The results are tabulated and shown above.

## VAR models on CADILAHC:

Necessary data was taken and the results were interpreted as shown below. The factors influencing returns were: volume traded and risk-free rates. These were the data time-series used in VAR modelling. The series were taken in the following order: returns, volume, risk-free rate

### Daily Data Based:

Estimation results for equation y1:

===================================

y1 = y1.l1 + y2.l1 + y3.l1 + const

Estimate Std. Error t value Pr(>|t|)

y1.l1 -1.160e-01 6.438e-02 -1.802 0.0728 .

y2.l1 3.315e-09 7.210e-08 0.046 0.9634

y3.l1 2.312e-01 1.997e-01 1.157 0.2483

const -1.313e-01 2.022e-01 -0.649 0.5169

---

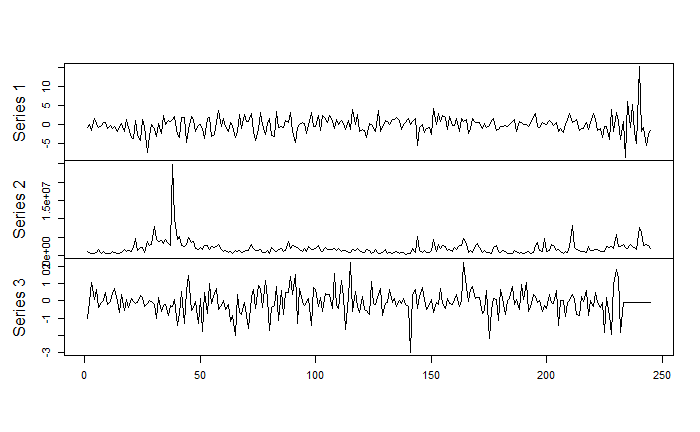
Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.258 on 240 degrees of freedom

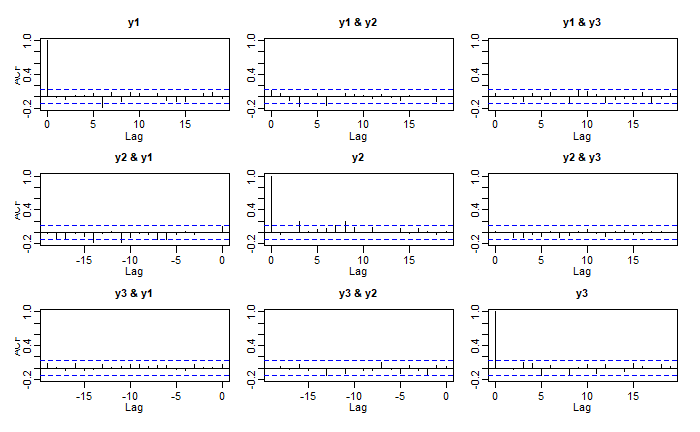
Multiple R-Squared: 0.01808, Adjusted R-squared: 0.005809

F-statistic: 1.473 on 3 and 240 DF, p-value: 0.2224

series data:



Below graph shows ACF plot for residuals of model fitted.



### Monthly data Based

Estimation results for equation split\_months\_returns\_2:

=======================================================

split\_months\_returns\_2 = split\_months\_returns\_2.l1 + xts\_monthly\_df.xts\_monthly\_vol.1.60..l1 + L.1.60..l1 + const

Estimate Std. Error t value Pr(>|t|)

split\_months\_returns\_2.l1 7.150e-03 1.404e-01 0.051 0.9596

xts\_monthly\_df.xts\_monthly\_vol.1.60..l1 -4.832e-07 2.602e-07 -1.857 0.0687 .

L.1.60..l1 -2.482e-01 3.363e-01 -0.738 0.4637

const 3.869e-01 1.124e+00 0.344 0.7320

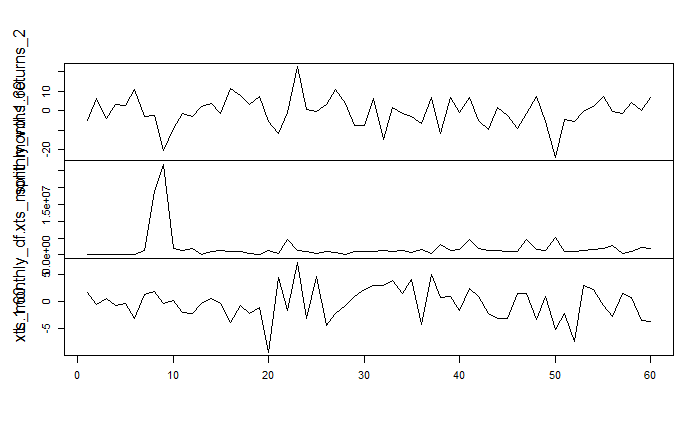
---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

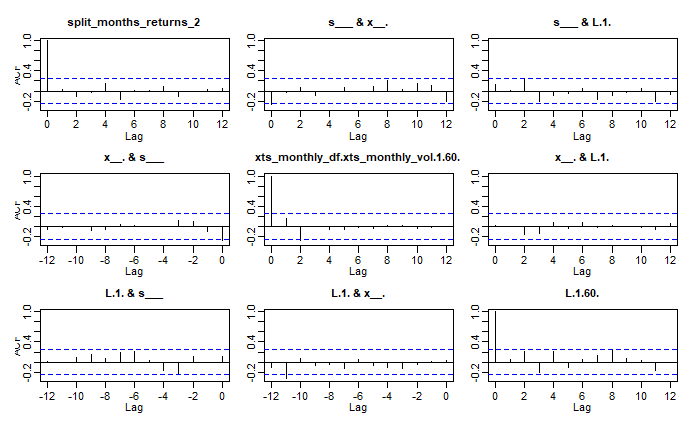
Residual standard error: 7.662 on 55 degrees of freedom

Multiple R-Squared: 0.07747, Adjusted R-squared: 0.02715

F-statistic: 1.539 on 3 and 55 DF, p-value: 0.2146



Below graph shows acf plot of residuals for fitted model.



### Weekly data based:

Estimation results for equation split\_weeks\_returns\_2:

======================================================

split\_weeks\_returns\_2 = split\_weeks\_returns\_2.l1 + xts\_weekly\_df.xts\_weekly\_vol.2.262..l1 + K.1.261..l1 + const

Estimate Std. Error t value Pr(>|t|)

split\_weeks\_returns\_2.l1 -1.013e-01 6.267e-02 -1.616 0.107

xts\_weekly\_df.xts\_weekly\_vol.2.262..l1 -1.867e-07 1.952e-07 -0.956 0.340

K.1.261..l1 3.088e-01 2.124e-01 1.454 0.147

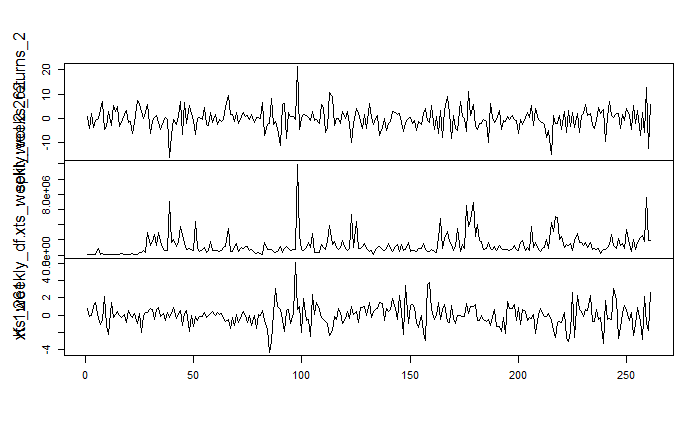
const 1.556e-01 3.768e-01 0.413 0.680

Residual standard error: 4.395 on 256 degrees of freedom

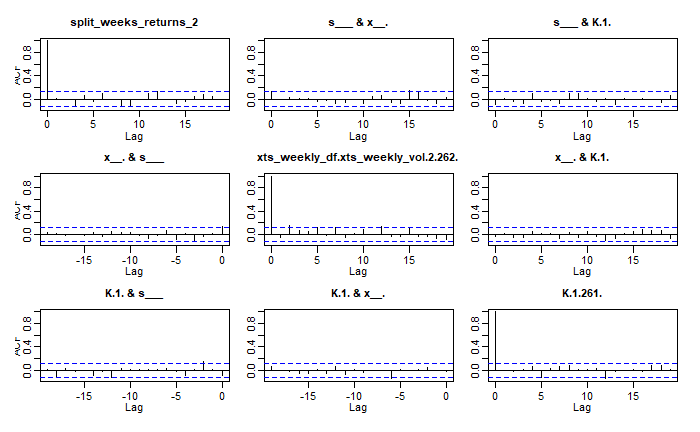
Multiple R-Squared: 0.02556, Adjusted R-squared: 0.01414

F-statistic: 2.238 on 3 and 256 DF, p-value: 0.08425

Series data.



Acf plot for residuals of fitted model:

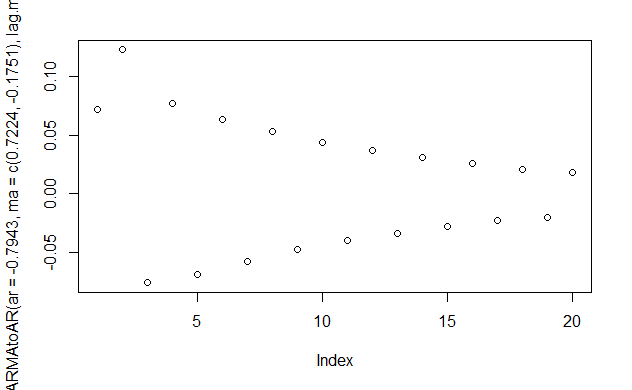


These interpretations and results conclude the VAR model interpretation of CESC returns on weekly ,monthly, daily data basis.

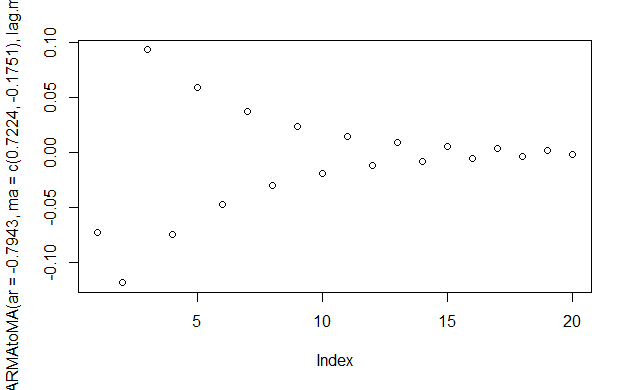
## Identifying and interpreting ARMA models of above data as AR(p),MA(q) models;

There is a function in R for the above thing to be implemented. The code is used as shown in the appendix and the required interpretations are mentioned here with coefficients as plots:

Daily data arima(1,0,2): when

ARMA to AR:  


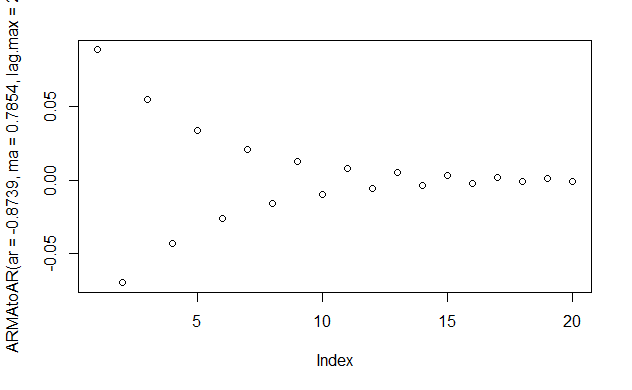
ARMA to MA



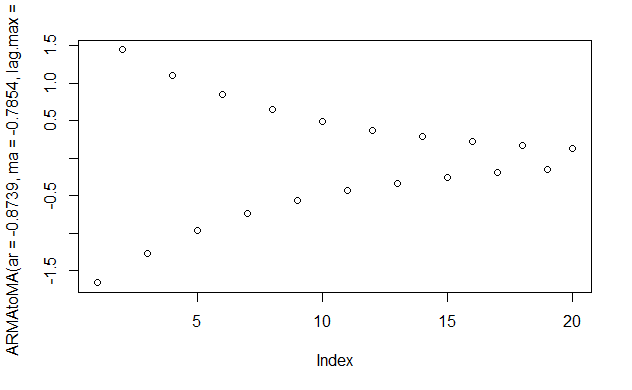
Note: as monthly data seemed to be white noise by autofit model, it is of no use to interpret the model in AR/MA terms.

Weekly data: arima (1,0,1)

ARMA to AR:



ARMA to MA



These finally conclude all the necessary interpretations for CADILAHC data.

# CONCLUSION

All the necessary calculations were done wherever required in R studio only and the necessary results were tabulated, plotted, explained, and interpreted whenever and wherever necessary. Trends were identified wherever present in returns data, and were modelled accordingly. The same models were used to estimate volatility in returns, which was difficult in the case of monthly returns as the number of datapoints used were not enough for the series to converge and moreover the data was just white noise which means that it cannot be modelled. All of this has been put in an order above and explained in detail, this concludes the report for this assignment. Important note points and appendix is attached below for your reference.

# NOTE:

For these models, entire code necessary and used to work is provided in the appendix. All the necessary models were built using libraries, some decent outlier removals, approximations. Explanation and summary statistics are provided wherever and whenever necessary. R notebook is used in R studio to solve the assignment code part block by block. The code is the same for both the data. Below one in the appendix is provided for sake of understanding.

## APPENDIX

### Data download from nsepy:

from datetime import date

from nsepy import get\_history

cesc = get\_history(symbol='CESC',

start=date(2015,4,1),

end=date(2020,3,31))

cesc.to\_csv('C:/Users/SUMANTH/Downloads/fram\_assignment/cesc%205%20years.csv',header=True)

cadilahc = get\_history(symbol='CADILAHC',

start=date(2015,4,1),

end=date(2020,3,31))

cadilahc.to\_csv('C:/Users/SUMANTH/Downloads/fram\_assignment/cadilahc%205%20years.csv',header=True)

### Risk-free rate data:

Downloaded from the internet.

Daily data: alpha.csv

Monthly data: gamma.csv

Weekly data: beta.csv

### R Code:

```{r}

library(astsa)

library(tseries)

library(xts)

library(zoo)

library(rugarch)

library(quantmod)

library(timeSeries)

library(forecast)

library(xts)

library(fGarch)

library(vars)

```

```{r}

#daily data working on cesc 1 year

df\_1Y<-read.csv("C:/Users/SUMANTH/Downloads/fram\_assignment/1\_year\_data/cesc.csv")

df\_1Y<-transform(df\_1Y,returns=round((100\*log(Close/Prev.Close)),4))

outliers\_free\_data<-c()

for (i in 1:length(df\_1Y$Date)){

if(df\_1Y$returns[i]<3\*sd(df\_1Y$returns)|df\_1Y$returns[i]>-3\*sd(df\_1Y$returns))

outliers\_free\_data<-c(outliers\_free\_data,df\_1Y$returns[i])

}

acf2(df\_1Y$returns)

summary(df\_1Y$returns)

xts\_1<-xts(outliers\_free\_data,order.by=as.Date(df\_1Y$Date),na.locf=TRUE)

modelfit<-auto.arima(diff(df\_1Y$returns))

modelfit

summary(modelfit)

ts.plot(df\_1Y$returns)

checkresiduals(modelfit)

daily\_forecast<-forecast(modelfit,10)

daily\_forecast\_df<-data.frame(daily\_forecast)

autoplot(daily\_forecast)

daily\_forecast

df<-read.csv("C:/Users/SUMANTH/Downloads/alpha.csv")

nrow(df)

xts\_rfd<-xts(df$change\_percent,order.by=as.Date(df\_1Y$Date),na.locf=TRUE)

multi\_v<-cbind(df\_1Y$returns,df\_1Y$Volume,df$change\_percent)

multi\_v

plot.ts(multi\_v , main = "", xlab = "")

fitvar1= VAR(multi\_v, p=1, type="const", ic="AIC")

summary(fitvar1)

acf(residuals(fitvar1))

```

```{r}

#monthly data working on cesc 5 years

df<-read.csv("C:/Users/SUMANTH/Downloads/fram\_assignment/cesc\_5\_years.csv")

xts\_1\_5Y<-xts(df$Close,order.by=as.Date(df$Date))

split\_months\_end\_index<-endpoints(xts\_1\_5Y,"months")

split\_months<-xts\_1\_5Y[xts:::startof(xts\_1\_5Y,"months")]

split\_months\_end<-xts\_1\_5Y[split\_months\_end\_index,1]

names(split\_months)<-"Close"

names(split\_months\_end)<-"Close\_end"

split\_months\_df<-data.frame(split\_months)

split\_months\_end\_df<-data.frame(split\_months\_end)

split\_months\_returns<-c()

for(i in 1:length(coredata(split\_months))){

split\_months\_returns<-c(split\_months\_returns,round((100\*log(split\_months\_end\_df$Close\_end[i]/split\_months\_df$Close[i])),4))

}

model\_monthly<-auto.arima(split\_months\_returns)

model\_adjusted\_monthly<-split\_months\_returns-residuals(model\_monthly)

summary(model\_monthly)

ts.plot(split\_months\_returns)

summary(split\_months\_returns)

acf2(split\_months\_returns)

points(model\_adjusted\_monthly,type="lty",col=2)

checkresiduals(model\_monthly)

forecast(model\_monthly,10)

autoplot(forecast(model\_monthly,10))

garch1.spec\_m<-ugarchspec(variance.model = list(garchOrder=c(1,1)), mean.model = list(armaOrder=c(0,0)))

monthly.garch1fit<-ugarchfit(spec=garch1.spec\_m, data=split\_months\_returns)

garchforecast\_m<-ugarchboot(monthly.garch1fit, n.ahead=10,method = c("Partial","Full")[1])

plot(garchforecast\_m,which=2)

gf\_m<-ugarchforecast(monthly.garch1fit,n.ahead=10)

gf\_m

sigma(gf\_m)

sigma(monthly.garch1fit)

residuals(monthly.garch1fit)

garch1.spec\_m

monthly.garch1fit

acf2(residuals(monthly.garch1fit),max.lag = 10)

arch1.spec\_m<-ugarchspec(variance.model = list(garchOrder=c(1,0)), mean.model = list(armaOrder=c(0,0)))

monthly.arch1fit<-ugarchfit(spec=arch1.spec\_m, data=split\_months\_returns)

arch1forecast\_m<-ugarchboot(monthly.arch1fit, n.ahead=10,method = c("Partial","Full")[1])

plot(arch1forecast\_m,which=2)

a1f\_m<-ugarchforecast(monthly.arch1fit,n.ahead=10)

a1f\_m

sigma(a1f\_m)

sigma(monthly.arch1fit)

residuals(monthly.arch1fit)

arch1.spec\_m

monthly.arch1fit

acf2(residuals(monthly.arch1fit),max.lag = 10)

summary()

arch2.spec\_m<-ugarchspec(variance.model = list(garchOrder=c(2,0)), mean.model = list(armaOrder=c(0,0)))

monthly.arch2fit<-ugarchfit(spec=arch2.spec\_m, data=split\_months\_returns)

arch2forecast\_m<-ugarchboot(monthly.arch2fit, n.ahead=10,method = c("Partial","Full")[1])

plot(arch2forecast\_m,which=2)

a2f\_m<-ugarchforecast(monthly.arch2fit,n.ahead=10)

sigma(a2f\_m)

arch2.spec\_m

sigma(monthly.arch2fit)

residuals(monthly.arch2fit)

monthly.arch2fit

acf2(residuals(monthly.arch2fit),max.lag = 10)

length(split\_months\_returns)

xts\_monthly<-to.monthly(xts\_1\_5Y,OHLC=FALSE)

xts\_monthly\_vol<-to.monthly(xts(df$Volume,order.by=as.Date(df$Date)),OHLC=FALSE)

xts\_monthly\_df<-data.frame(xts\_monthly\_vol)

lapply(xts\_monthly\_df,cumsum)

xts\_monthly\_df

df\_m<-read.csv("C:/Users/SUMANTH/Downloads/gamma.csv")

L<-df\_m$change\_percent

length(L)

df\_wr<-data.frame(split\_months\_returns,xts\_monthly\_df$xts\_monthly\_vol[1:60],L[1:60])

plot.ts(df\_wr , main = "", xlab = "")

fitvar2= VAR(df\_wr, p=1, type="const", ic="AIC")

summary(fitvar2)

acf(residuals(fitvar2))

```

```{r}

#weekly data working on cesc 5 years

xts\_weekly<-to.weekly(xts\_1\_5Y,OHLC=FALSE)

xts\_weekly\_vol<-to.weekly(xts(df$Volume,order.by=as.Date(df$Date)),OHLC=FALSE)

xts\_weekly\_df<-data.frame(xts\_weekly\_vol)

lapply(xts\_weekly\_df,cumsum)

xts\_weekly\_df

colnames(xts\_weekly\_df[1])<-"Date"

split\_weeks\_vol<-split(xts\_1\_5Y,"weeks")

split\_weeks\_vol

df\_w<-read.csv("C:/Users/SUMANTH/Downloads/betaa.csv")

K<-df\_w$change\_percent

length(K)

split\_weeks\_returns<-c()

for(i in 1:length(coredata(xts\_weekly))-1){

split\_weeks\_returns<-c(split\_weeks\_returns,round((100\*log(xts\_weekly\_df[i+1,1]/xts\_weekly\_df[i,1])),4))

}

split\_weeks\_returns\_df<-data.frame(split\_weeks\_returns)

new\_weekly\_model<-auto.arima(split\_weeks\_returns)

new\_weekly\_model\_adj<-split\_weeks\_returns-residuals(new\_weekly\_model)

ts.plot(split\_weeks\_returns)

points(new\_weekly\_model\_adj,type="lty",col=2)

acf2(split\_weeks\_returns)

new\_weekly\_model

checkresiduals(new\_weekly\_model)

x<-forecast(new\_weekly\_model,10)

summary(split\_weeks\_returns)

x\_df<-data.frame(x)

autoplot(x)

summary(new\_weekly\_model)

df\_wr<-data.frame(split\_weeks\_returns,xts\_weekly\_df$xts\_weekly\_vol[2:262],K[1:261])

fitvar2= VAR(df\_wr, p=1, type="const", ic="AIC")

plot.ts(df\_wr , main = "", xlab = "")

summary(fitvar2)

acf(residuals(fitvar2))

```

```{r}

garch1.spec<-ugarchspec(variance.model = list(garchOrder=c(1,1)), mean.model = list(armaOrder=c(1,1)))

daily.garch1fit<-ugarchfit(spec=garch1.spec, data=diff(df\_1Y$returns))

garchforecast<-ugarchboot(daily.garch1fit, n.ahead=10,method = c("Partial","Full")[1])

plot(garchforecast,which=2)

gfd<-ugarchforecast(daily.garch1fit,n.ahead=10)

plot(gfd,which=3)

garch1.spec

sigma(gfd)

daily.garch1fit

acf2(residuals(daily.garch1fit),max.lag = 10)

```

```{r}

arch1.spec<-ugarchspec(variance.model = list(garchOrder=c(1,0)), mean.model = list(armaOrder=c(1,1)))

daily.arch1fit<-ugarchfit(spec=arch1.spec, data=diff(df\_1Y$returns))

arch1forecast<-ugarchboot(daily.arch1fit, n.ahead=10,method = c("Partial","Full")[1])

plot(arch1forecast,which=2)

a1fd<-ugarchforecast(daily.arch1fit,n.ahead=10)

a1fd

sigma(a1fd)

plot(a1fd,which=3)

arch1.spec

daily.arch1fit

acf2(residuals(daily.arch1fit),max.lag = 10)

```

```{r}

arch2.spec<-ugarchspec(variance.model = list(garchOrder=c(2,0)), mean.model = list(armaOrder=c(1,1)))

daily.arch2fit<-ugarchfit(spec=arch2.spec, data=diff(df\_1Y$returns))

arch2forecast<-ugarchboot(daily.arch2fit, n.ahead=10,method = c("Partial","Full")[1])

plot(arch2forecast,which=2)

a2fd<-ugarchforecast(daily.arch2fit,n.ahead=10)

a2fd

sigma(a2fd)

arch2.spec

daily.arch2fit

acf2(residuals(daily.arch2fit),max.lag = 10)

```

```{r}

garch1.spec\_w<-ugarchspec(variance.model = list(garchOrder=c(1,1)), mean.model = list(armaOrder=c(0,1)))

weekly.garch1fit<-ugarchfit(spec=garch1.spec\_w, data=split\_weeks\_returns)

sigma(weekly.garch1fit)

residuals(weekly.garch1fit)

garchforecast<-ugarchboot(weekly.garch1fit, n.ahead=10,method = c("Partial","Full")[1])

plot(garchforecast,which=2)

gfd\_w<-ugarchforecast(weekly.garch1fit,n.ahead=10)

gfd\_w

sigma(gfd\_w)

garch1.spec\_w

weekly.garch1fit

acf2(residuals(weekly.garch1fit),max.lag = 10)

```

```{r}

arch1.spec\_w<-ugarchspec(variance.model = list(garchOrder=c(1,0)), mean.model = list(armaOrder=c(0,1)))

weekly.arch1fit<-ugarchfit(spec=arch1.spec\_w, data=split\_weeks\_returns)

arch1forecast\_w<-ugarchboot(weekly.arch1fit, n.ahead=10,method = c("Partial","Full")[1])

plot(arch1forecast\_w,which=2)

a1fd\_w<-ugarchforecast(weekly.arch1fit,n.ahead=10)

a1fd\_w

sigma(a1fd\_w)

arch1.spec\_w

weekly.arch1fit

acf2(residuals(weekly.arch1fit),max.lag = 10)

```

```{r}

arch2.spec\_w<-ugarchspec(variance.model = list(garchOrder=c(2,0)), mean.model = list(armaOrder=c(0,1)))

weekly.arch2fit<-ugarchfit(spec=arch2.spec\_w, data=split\_weeks\_returns)

arch2forecast\_w<-ugarchboot(weekly.arch2fit, n.ahead=10,method = c("Partial","Full")[1])

plot(arch2forecast\_w,which=2)

a2fd\_w<-ugarchforecast(weekly.arch2fit,n.ahead=10)

a2fd\_w

sigma(a2fd\_w)

arch2.spec\_w

weekly.arch2fit

acf2(residuals(weekly.arch2fit),max.lag = 10)

```

```{r}

plot(ARMAtoAR(ar= -0.2266 ,ma= -0.9117,lag.max=20),type="bar")

plot(ARMAtoMA(ar=-0.2266,ma=-0.9117,lag.max=20))

plot(ARMAtoAR(ar=0,ma=0.1366,lag.max=20))

```