STAT 30010

Dr. Patrick Murphy

Time Series Project

Luke Killoran

13434418

Table of Contents

Project Outline

The first task was to simulate 50 observations each from three different AR models, and concatenate these observations to form a new time series. The parameters of the initial AR models were determined by the digits in my student number and my UCD contact number. These parameters were the same for each AR model except for the variance of the residuals. Then the main objective was to find the model (ARIMA or otherwise) that fits these data points the best.

Generation of the Time Series

The parameters for each AR model were determined by the following method:
The third, fourth and fifth digits of my student number (13434418) indicated the value of the respective phis for all three AR models. Thus, $\phi_1 = 4$, $\phi_2 = 3$, $\phi_3 = 4$. Whereas the fourth digit of my UCD contact number (0852305417) dictated the value of the mean of all three AR processes. Hence, $\mu=2$. Furthermore, the next three dissimilar non-zero digits of my UCD contact number (0852305417) specified the standard deviations of the residuals of each individual AR process. Consequently, $\sigma_1 = 3$, $\sigma_2 = 5$, and $\sigma_3 = 4$. Finally, the mean of the residuals of all three AR processes was selected to be 0.

The first task was to simulate 50 observations from each of the following models:

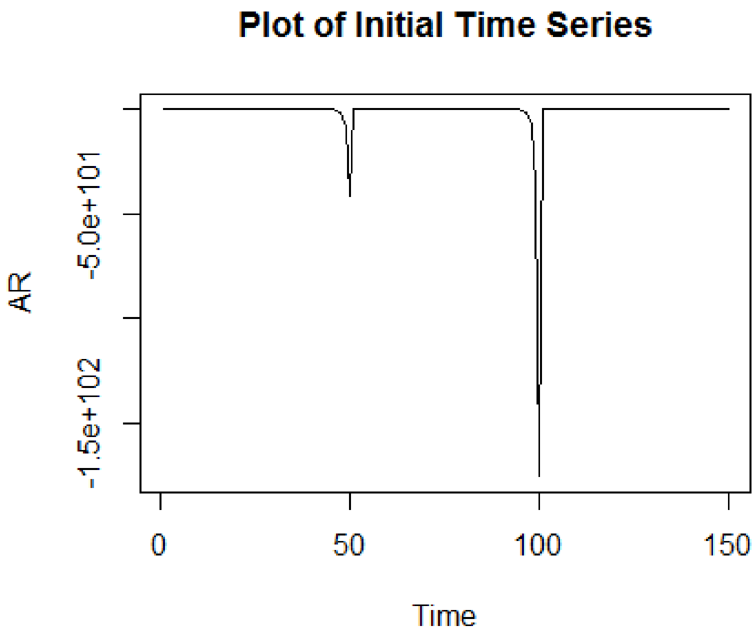
$$Y_t = 2 + 4(Y_{t-1} - 2) + 3(Y_{t-2} - 2) + 4(Y_{t-3} - 2) + \varepsilon_t \quad \text{where } \varepsilon_t \sim N(0,3^2) \text{ is white noise}$$
$$Y_t = 2 + 4(Y_{t-1} - 2) + 3(Y_{t-2} - 2) + 4(Y_{t-3} - 2) + \varepsilon_t \quad \text{where } \varepsilon_t \sim N(0,5^2) \text{ is white noise}$$
$$Y_t = 2 + 4(Y_{t-1} - 2) + 3(Y_{t-2} - 2) + 4(Y_{t-3} - 2) + \varepsilon_t \quad \text{where } \varepsilon_t \sim N(0,4^2) \text{ is white noise}$$

Firstly, 250 observations (of a normal random variable with mean 0 and standard deviation 3) were simulated for the white noise of the first AR model. I used the last 150 observations since the starting point effect would have worn off by then, and the white noise could be accepted as random. These 150 observations were used in an iterative process to determine values of the Y_t. I set $Y_1 = 2 + \varepsilon_1$, $Y_2 = 2 + 4(Y_1 – 2) + \varepsilon_2$ and $Y_3 = 2 + 4(Y_2 – 2) + 3(Y_1 – 2) + \varepsilon_3$. Then I defined $Y_t = 2 + 4(Y_{t-1} - 2) + 3(Y_{t-2} - 2) + 4(Y_{t-3} - 2) + \varepsilon_t$ for all $t \geq 3$. I adopted the strategy of generating 150 observations of the Y_t as opposed to just 50 so that I could use the last 50 observations, and the starting point of the Y_t would have much less of an effect on the time series.

The observations for the other two AR models were produced in a similar way. Then the final 50 observations of each set were concatenated to form a new time series.

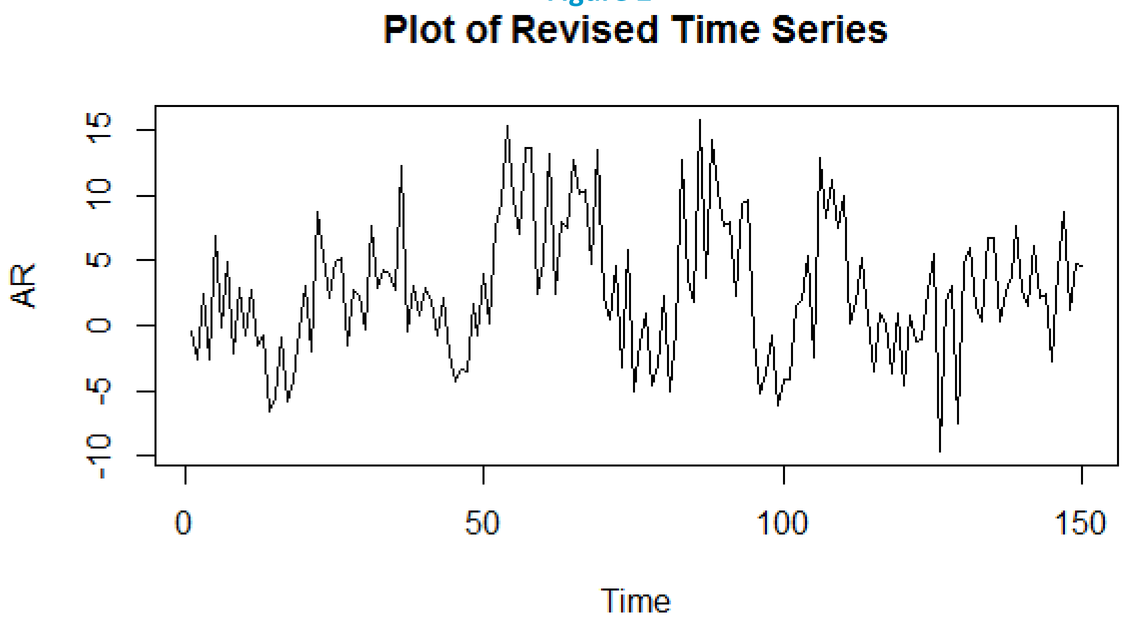
However, upon conjoining the three sets of 50 observations, the ensuing time series was unrealistic as shown by Figure 1. The time series spiked to large negative values with 101 zeros in the figure at one point. Thus, for the purposes of this report, it is understandable that I change the phis of the individual AR models.

Figure 1



Hence, I decided to reduce the three phis by inverting them. The plot of the resulting time series is given below in Figure 2.

Figure 2



The time series seems to fluctuate around a constant mean of about 2.5. However, the time series does not appear to be stationary. The variance looks to be greater between observations 50 and 100 than the variance of the series between observations 0 and 50, and even 100 and 150. Although it must be noted that, when considering random variation the time series is close to being stationary. The first model I will try to fit to the time series is a (Box-Jenkins) ARIMA model. This is because the data do not seem to exhibit signs of seasonality, and I do not know of any factors that possibly affect the time series other than past performance and random variation (which are both modelled by the ARIMA process).

ARIMA Modelling

I determined the order of the ARIMA model to use initially, by examining the ACF and PACF of the time series given below in Figure 3 and 4 respectively.

Figure 3

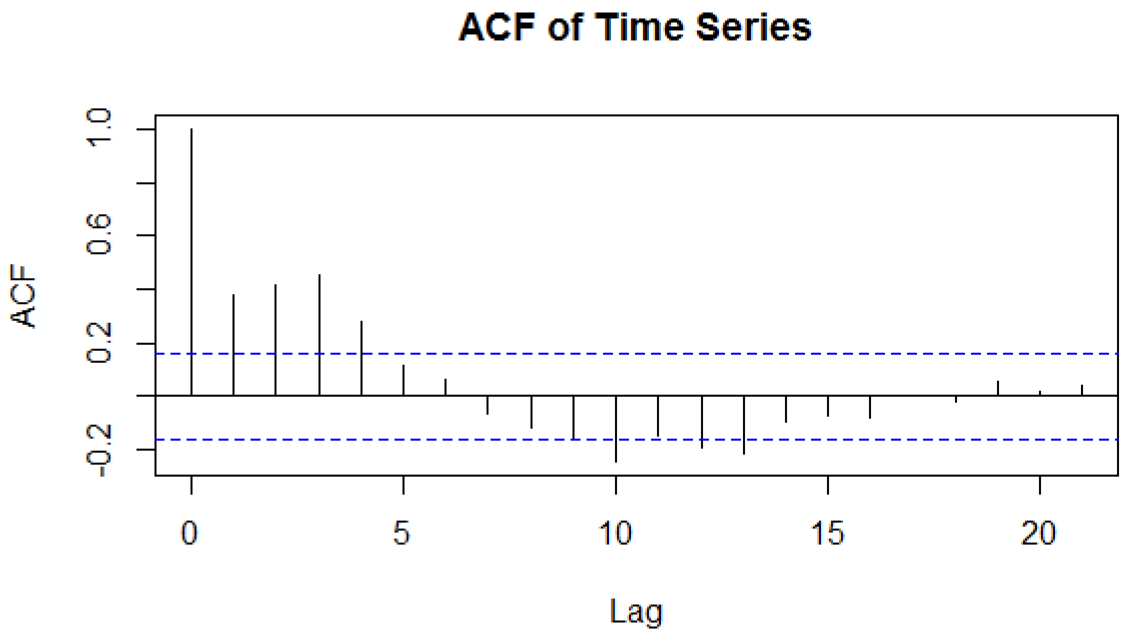
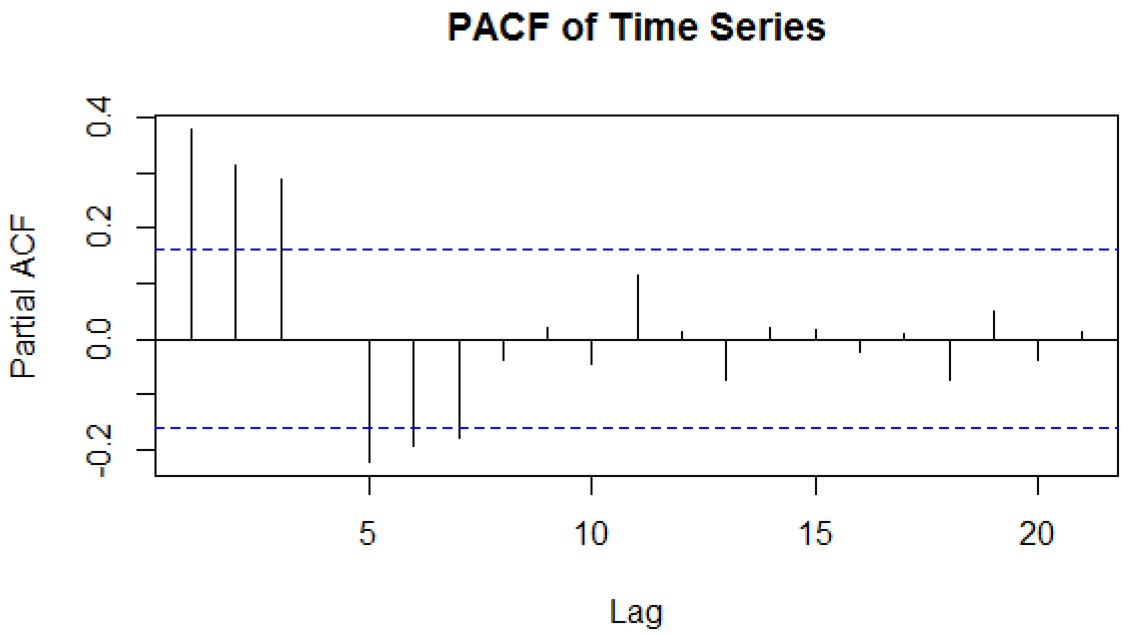


Figure 4



The autocorrelations appear to decay exponentially, while the partial autocorrelations drop below the line of significance after 3 lags (despite being slightly significant for lags 5, 6 and 7). This would seem to indicate an AR(3) model. When I fitted the data to the AR(3) in R, I obtained estimates for ϕ_1, ϕ_2, and ϕ_3, which are given in Figure 5. These estimates were derived through minimizing conditional sum of squares to find starting values, followed by maximizing the likelihood.

Figure 5

```
Estimated Coefficients of AR(3):
          ar1     ar2     ar3       mean
       0.1662  0.2415  0.2868     2.6725
s.e.   0.0777  0.0765  0.0777     1.1335
```

I tested the null hypothesis $H_0 : \phi_i = 0$ given ϕ_j is in the model for i = {1,2,3}, j = {1,2,3} but $j \neq i$ against the alternative hypothesis $H_A : \phi_l \neq 0$ given the other phis are in the model. The test statistic was given by estimate of ϕ_i / standard error of the estimate of ϕ_i. If the null hypothesis is true, then the test statistic should follow a t distribution (with number of observations-number of estimated coefficients as the degrees of freedom). Since I have only 150 observations, I could assume a $t_{n-3} = t_{147}$ distribution of the test statistic. At the 5% level of significance, the critical value for this two-tailed test is approximately equal to 1.9762. Since $|t_1|$=2.14, $|t_2|$=3.16, and $|t_3|$=3.7, which are all greater than 1.9762, I reject the null hypothesis H_0 in favour of the alternative hypothesis H_A in each case. Thus, I can conclude that there was significant evidence to suggest that ϕ_1, ϕ_2, and ϕ_3 are statistically significantly different from zero at the 5% level of significance. Since all the parameters are significant, they all help explain a significant amount of the variation in Y_t and therefore should be left in the model.

I then overfitted my model to check the significance of extra terms in the model. I tried adding an extra AR term and, separately, an extra MA term. The estimated coefficients for these models are given below in Figure 6 and Figure 7 respectively.

Figure 6

```
Estimated Coefficients of AR(4):
          ar1     ar2     ar3     ar4       mean
       0.1661  0.2415  0.2867  0.0002     2.6724
s.e.   0.0812  0.0789  0.0790  0.0816     1.1336
```

Figure 7

```
Estimated Coefficients of ARMA(3,1):
          ar1     ar2     ar3      ma1       mean
       0.1664  0.2414  0.2867  -0.0003     2.6723
s.e.   0.1805  0.0871  0.0934   0.1778     1.1336
```

Since $|t_4| = 0.002 < t_{146}(0.025) = 1.9763$, I concluded that the extra AR term was not significant at the 5% level of significance. An AR(4) model with ϕ_4 set to 0 is just an AR(3) model. Therefore, the AR(3) model was preferred. Similarly, since $|t_4| = 0.002 < 1.9763$, I concluded that the extra MA term was not significant at the 5% level of significance. An ARMA(3,1) model with Θ_1 set to 0 is just an AR(3) model. Thus, the AR(3) model was preferred.

Consequently, my model for the data was given by the following equation:

$$Y_t = 2.6725 + 0.1662*(Y_{t-1} - 2.6725) + 0.2415*(Y_{t-2} - 2.6725) + 0.2868*(Y_{t-3} - 2.6725) + \varepsilon_t$$

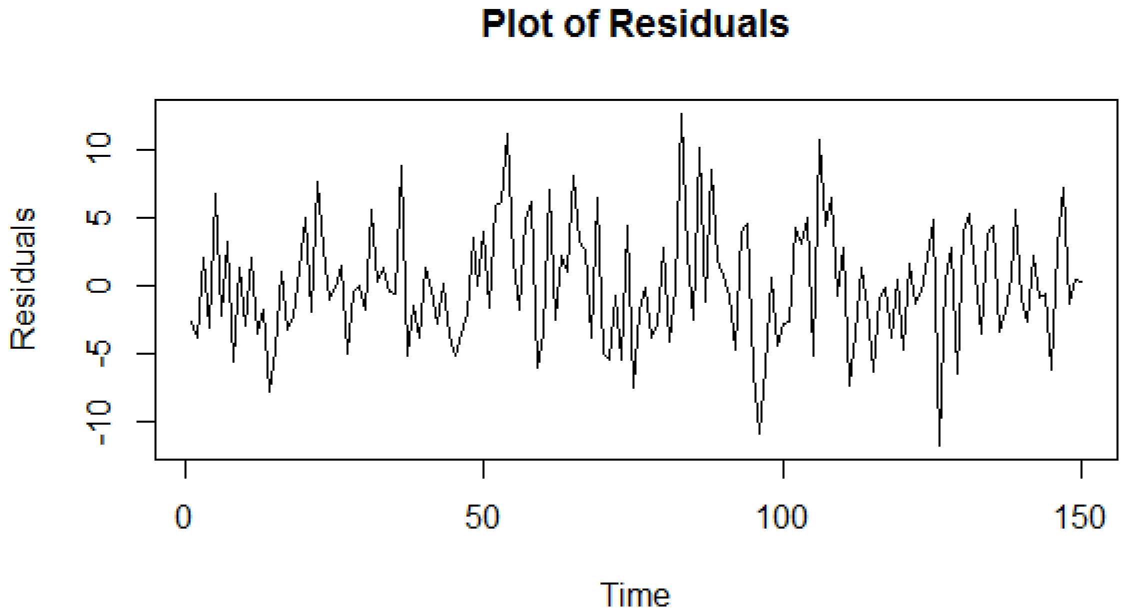
ARIMA Diagnostics

After finding what I believed to be the best ARIMA model for the original series, I had to test how well this model fitted the data. If the model was a perfect fit then it would be expected that the residuals would be distributed as white noise. In other words, they would be independently and identically distributed with mean 0 and constant variance. The residuals of my model were given by the following equation:

$$\hat{\varepsilon}_t = Y_t - 0.1662*(Y_{t-1} - 2.6725) - 0.2415*(Y_{t-2} - 2.6725) - 0.2868*(Y_{t-3} - 2.6725) - 2.6725$$

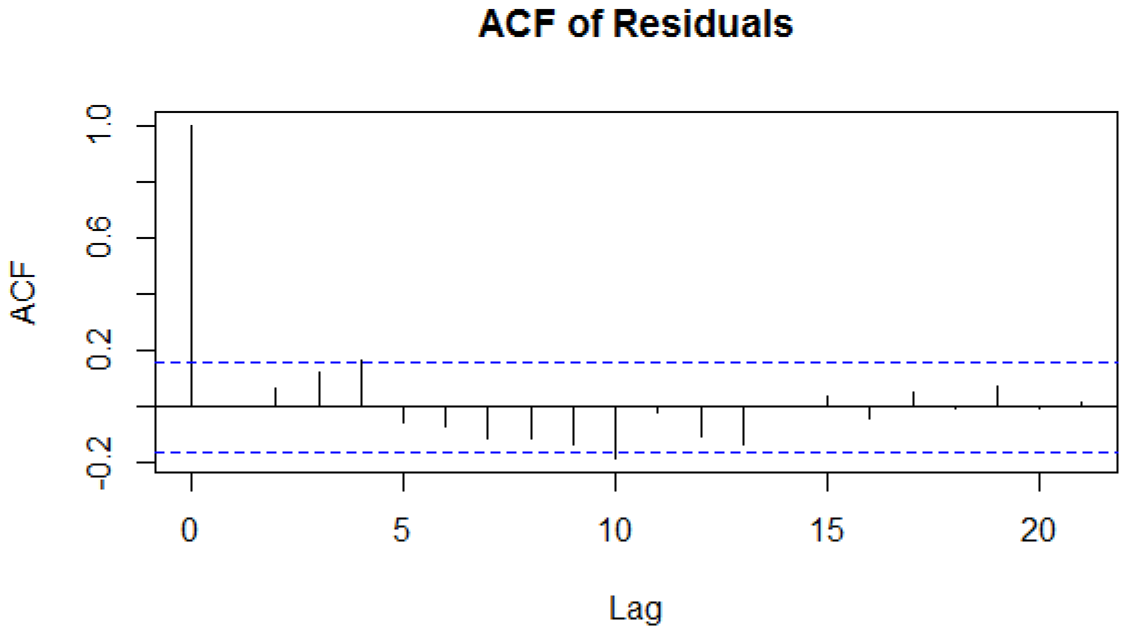
The time series plot of the residuals, which is given below in Figure 8, indicates that the residuals seem to fluctuate around a constant mean of 0. However, the residuals exhibit signs of non-stationarity, as they appear to be somewhat heteroscedastic. This can be seen in how the residuals vary more between observations 100 and 150 than between 0 and 50.

Figure 8



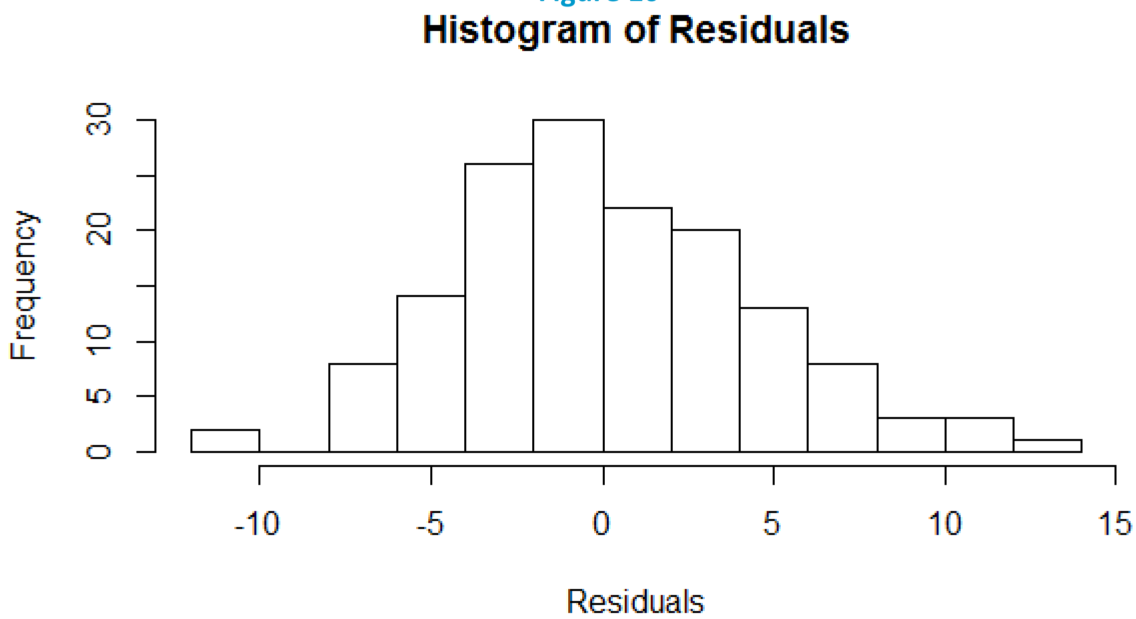
The ACF of the residuals (given in Figure 9) shows that autocorrelations for any lag up to 20 are below the level of significance (except for the trivial case of lag 0, and lags 4 and 10, which are slightly significant – this can be attributed to random variation). This indicates that the residuals do not appear to deviate from being white noise variables.

Figure 9



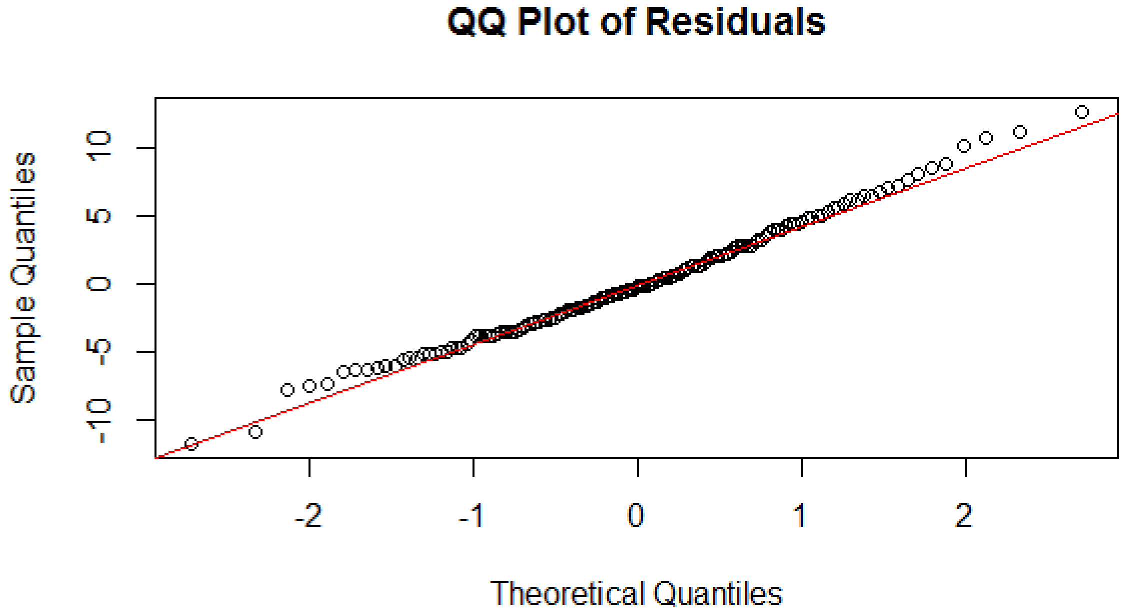
The histogram of the residuals, given below in Figure 10, indicates that the residuals seem to have a mean of 0 and are normally distributed. It slightly resembles a bell-shaped curve, despite being somewhat positively skewed.

Figure 10



The QQ plot of the residuals, given in Figure 11, provides further evidence that the residuals follow a Normal distribution. The sample quartiles of the residuals mirror the theoretical quartiles of the Normal distribution, as all the points lie approximately along the line and do not deviate much from it, even in the tails.

Figure 11



The same inference was concluded when I conducted a Shapiro-Wilk Normality test on the residuals. This tests the null hypothesis that the data is Normally distributed against the alternative hypothesis that the data is not Normally distributed. This test resulted in a test statistic of W = 0.99134 with p-value = 0.4927. Since the p-value > α = 0.05, the test failed to reject H_0 in favour of H_A. Therefore, there is insufficient evidence to suggest that the residuals do not follow a Normal distribution at the 5% level of significance.

I also ran a Ljung-Box test on the residuals. This tests the null hypothesis that the residuals are independent versus the alternative hypothesis that the residuals are dependent. Moreover, in order to check whether the residuals depend on just the recent values or all the past values, this test was carried out with lags equal to 30, 60, 90, 120 and 149. The resulting p-values were 0.1748, 0.1184, 0.2002, 0.4136, and 0.8962 respectively. If the residuals do depend on each other, they certainly depend on more recent values. However, since all the p-values > α = 0.05, I failed to reject H_0 in favour of H_A. Therefore, there was insufficient evidence that the value of residuals depended on each other at the 5% level of significance.

Thus, in summary, the model seems to fit the time series well. The residuals seem to be independent and normally distributed with zero mean. Although, the residuals may be heteroscedastic.

Non-Stationarity

As I have previously alluded to, the time series appeared to display signs of non-stationarity. There are three main reasons why the time series could be non-stationary: It could have d unit roots ($d \geq 1$). It could have a polynomial trend. It could be heteroscedastic.

Initially, I checked for the presence of a polynomial trend. To test this, I regressed the time series data to include a drift term α, linear term βt, quadratic term γt^2 , and cubic term ωt^3 and yielded estimates (given in Figure 12) through minimising the sum of squares. Using a t test with degrees of freedom equal to 148, I can test the null hypothesis that a particular term is equal to 0 given all the other terms are in the model against the alternative hypothesis that the term is not equal to 0. The p-values for the linear, quadratic and cubic terms were 0.00241, 0.01010 and 0.02820 respectively. Since each p-value < α = 0.05, H_0 is rejected in favour of H_A. Thus, there is sufficient evidence to suggest the time series contains a linear, quadratic and cubic trend term, given that the other terms are in the model (at the 5% level of significance). I then conducted an F-test (with degrees of freedom equal to 146) which tests the null hypothesis that all three terms are simultaneously zero against the alternative hypothesis that at least one of the three terms is not zero. This yielded a test statistic of 3.996 and subsequently a p-value of 0.009053. Since the p-value < α = 0.05, H_0 was rejected in favour of H_A. Hence, there is sufficient evidence that at least one of the terms is significant at the 5% level of significance. However, I chose not to de-trend the time series. This decision was made for three reasons. Firstly, when testing using the significance of the terms with the drift as the only other term in the model, they all reported p-values > α = 0.05. This was the same t-test procedure as before except for the different estimates and standard errors. For the model with only the drift term and the linear term, the linear term produced a p-value of 0.443. For the model with only the drift term and the quadratic term, the quadratic term produced a p-value of 0.894. For the model with only the drift term and the cubic term, the cubic term produced a p-value of 0.904. Therefore, given that the only other trend term in the model was the drift, each of these polynomial terms were insignificant at the 5% level of significance. Secondly, the estimates of the coefficients of these trend terms are so small (very close to 0) that they have minimal effect on the value of the Y_t. Thirdly, the trend terms explain very little in the variation of Y_t. This can be seen in the way they contribute a coefficient of determination (R^2) of only 0.07587702 and an adjusted coefficient of determination of only 0.05688819. Therefore, since they have such a small effect on Y_t, it would not be worth it to add them to the model.

Figure 12

```
Estimated Trend Coefficients:
              Estimate Std. Error t value Pr(>|t|)
α            -2.777e+00  1.703e+00  -1.630  0.10527
β             3.007e-01  9.737e-02   3.088  0.00241
γ            -3.898e-03  1.496e-03  -2.606  0.01010
ω             1.443e-05  6.513e-06   2.216  0.02820
```

Secondly, I checked for the presence of any unit roots. The time series plots of the differenced time series and twice differenced time series are given in Figure 13 and 14 respectively. Differencing once did not make the time series become stationary. Despite fluctuating about a constant mean of 0, the differenced time series still exhibits signs of heteroscedasticity. Differencing twice did not make the time series become stationary for the same reason as before.

Figure 13

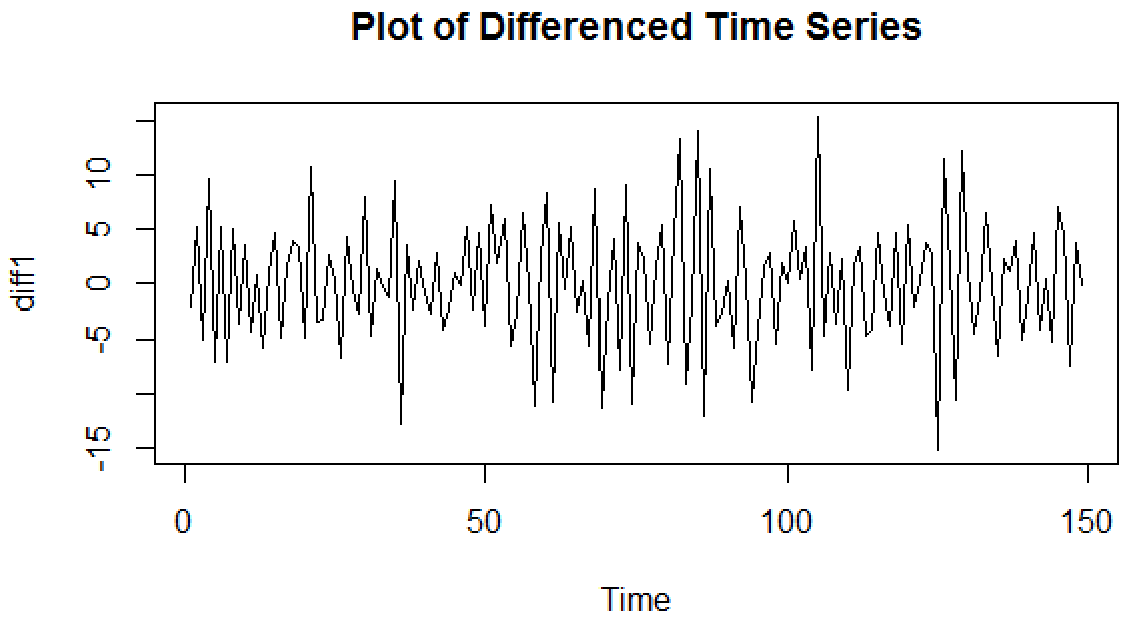
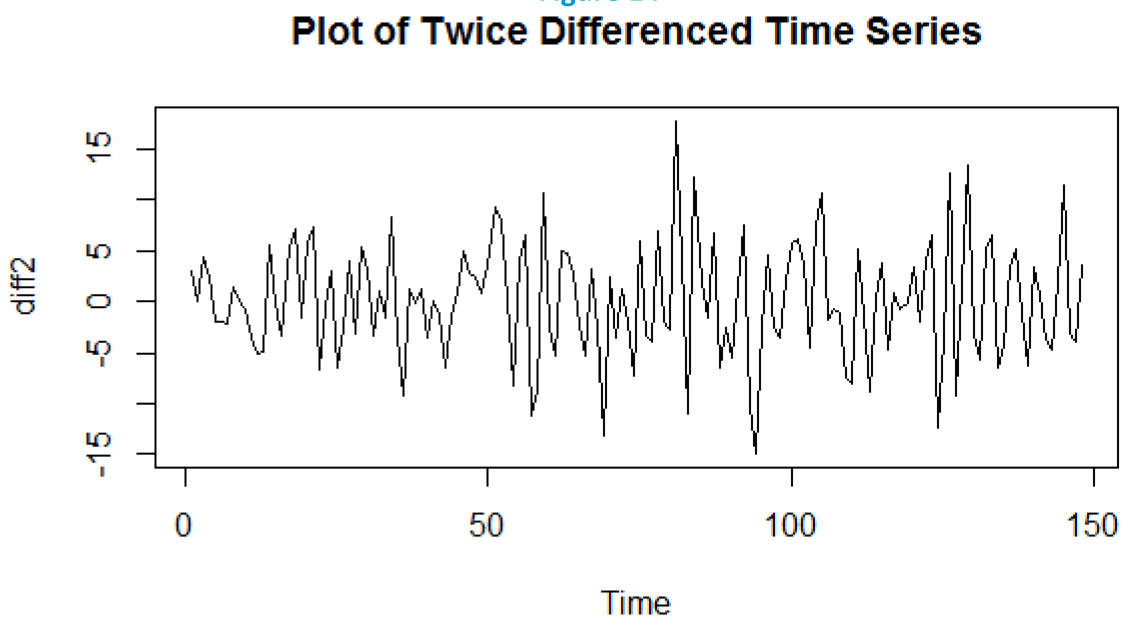
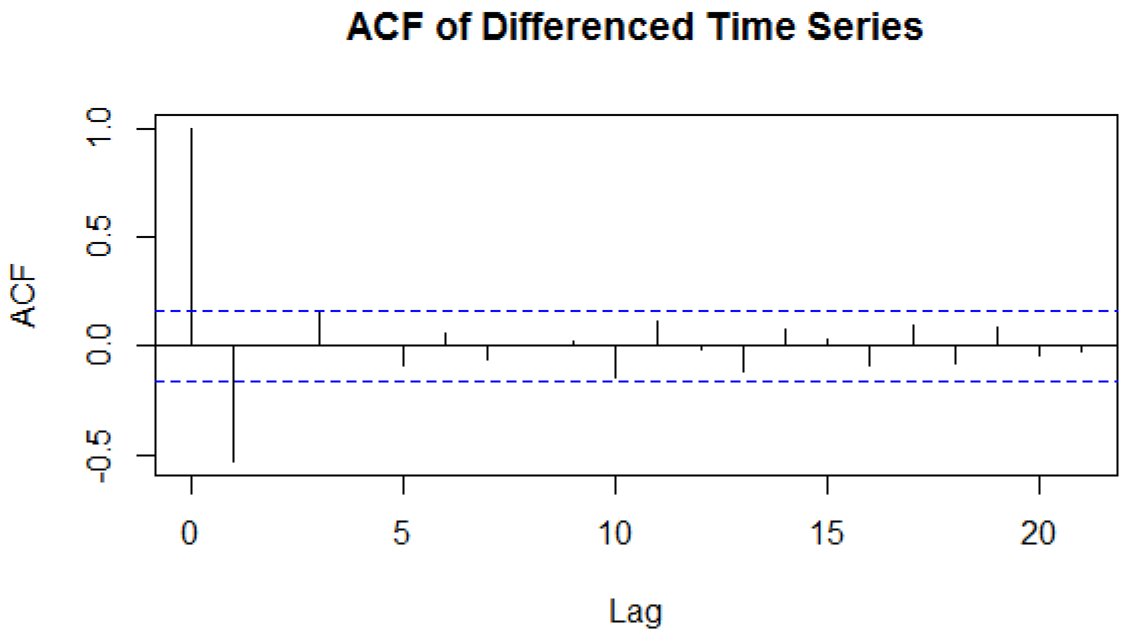


Figure 14



As shown in Figure 15, the autocorrelations of the differenced time series die out quickly, with only one autocorrelation being significant up to lag 20 (apart from the trivial case of lag 0). This is an indication that the time series is not non-stationary due to a unit root.

Figure 15



For further evidence of this, an Augmented Dickey Fuller test was conducted, which can deal with the case of AR(p) models and is still valid if the residuals are not white noise variables. The tests null hypothesis is that the model contains a unit root against the alternative hypothesis that the model is stationary. So a unit root is accepted unless there is sufficient evidence that the model is stationary. This is the case because of the relative importance of the two errors in the testing procedure. Differencing a stationary series would lead to more conservative forecast intervals. However, not differencing an otherwise non-stationary series can lead to over-confident and inaccurate forecast intervals, not to mention spurious regression if used in conjunction with another time series. The result of this test yielded a p-value of less than 0.01. Since the p-value < $\alpha = 0.05$, the null hypothesis is rejected in favour of the alternative hypothesis. There I could conclude that there sufficient evidence to suggest that the time series is not difference-stationary at the 5% level of significance.

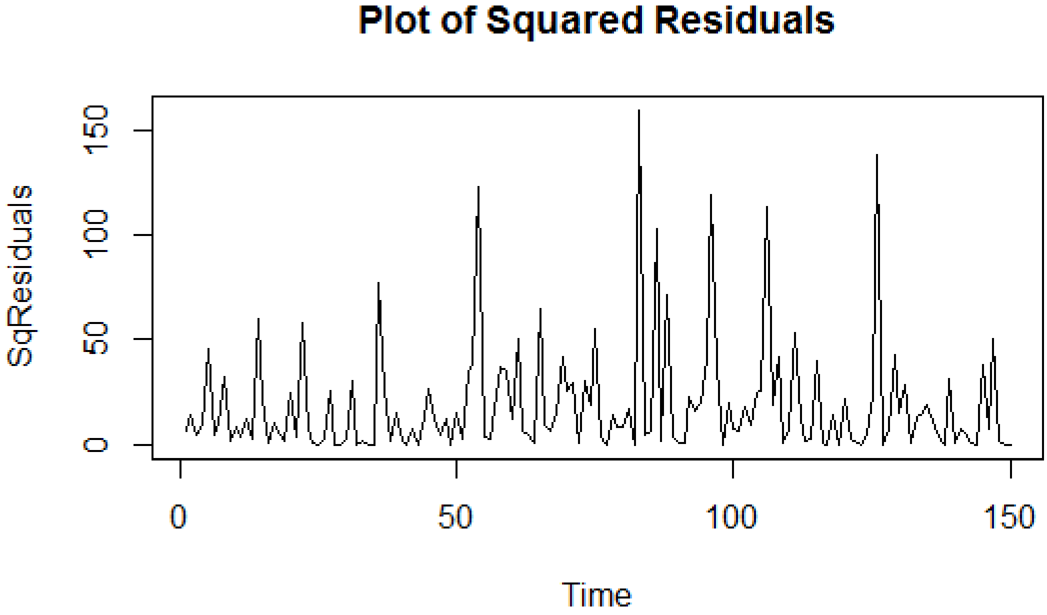
The result of this analysis for potential unit roots supports my decision not to include a linear, quadratic or cubic term. Differencing once would remove the linear term; differencing twice would remove the quadratic term and differencing three times would remove the cubic term. Therefore if the trend terms were influential, differencing the time series would have been recommended by the ADF test.

Since I have concluded that the time series is not difference-stationary, and the trend terms are not influential. The heteroscedasticity of the residuals must be main reason why the time series exhibits signs of non-stationarity.

GARCH/ARCH Modelling

I have deduced that the reason the time series exhibits signs of heteroscedasticity is largely down to the non-constant variance of the residuals. This can be seen in the time series plot of the residuals in Figure 8. However, these clusters of volatility can be seen more clearly in the time series plot of the squared residuals, given below in Figure 16.

Figure 16



This volatility is expected to have been a combination of random variation and a dependence of the residuals on time and/or themselves. In order to model this heteroscedasticity, a GARCH or ARCH model could be utilised. Auto Regressive Conditional Heteroscedastic (ARCH) models are often used for AR models and generally assume the variance of residuals to be a function of the previous error terms: often the variance is related to the squares of the residuals. Whereas Generalised Auto Regressive Conditional Heteroscedastic (GARCH) models are often used for ARMA models and generally assume the same as ARCH models, but that the residuals are also a function of the time-dependent variance of the previous error terms. In the ARCH(q) model, it is assumed that each $\varepsilon_t = v_t\sigma_t$ where v_t is a white noise process and σ_t is a time-dependent standard deviation which is modelled by the following equation: $\sigma_t^2 = \hat{\alpha}_0 + \sum_{i=1}^{q} \hat{\alpha}_i\hat{\varepsilon}_t^2$. To decide the order of the ARCH model, the ACF and PACF of the squared residuals should be examined. If the autocorrelations of the squared residuals decay exponentially and the partial autocorrelations of the squared residuals drops below the line of significance after lag k, then it would expected that the optimal ARCH model would have order k.

The ACF and PACF of the squared residuals are given in Figure 17 and Figure 18. After lag = 0, the autocorrelations are lie below the line of significance. Similarly, all of the partial correlations of the squared residuals lie below the line of significance. This indicates that the squared residuals do not depend on each other, and therefore an ARCH or GARCH model is not a preferred model.

Figure 17

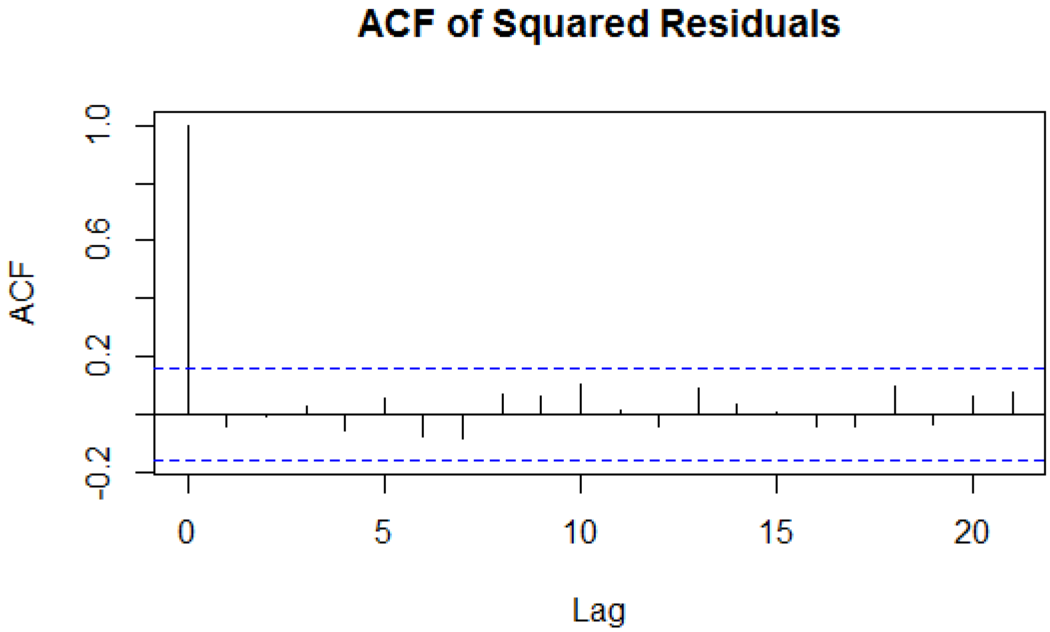
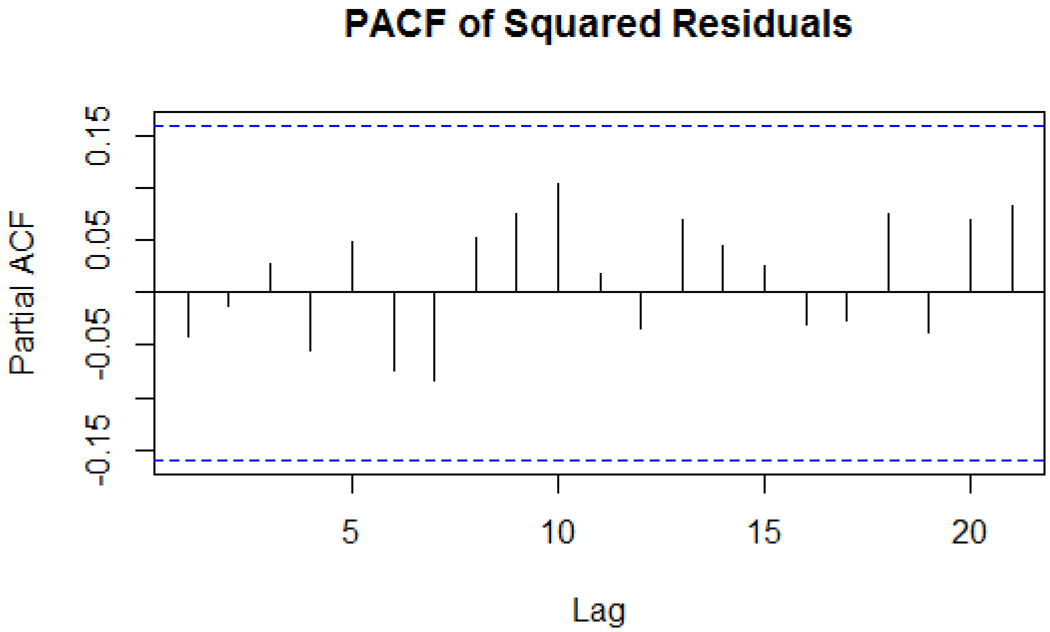


Figure 18



To investigate this further, I conducted a Ljung-Box test on the squared residuals with lags equal to 30, 60, 90, 120, and 149. The null hypothesis is that the squared residuals are independent and the alternative hypothesis is that the squared residuals are dependent on each other. The resulting p-values were 0.539, 0.5567, 0.5161, 0.9477, and 0.9991. Since all of these p-values > $\alpha = 0.05$, H_0 is rejected in favour of H_A. Thus, there is insufficient evidence to suggest that the squared residuals are dependent on each other at the 5% level of significance.

In order to confirm that the ARCH model is not appropriate, I fitted ARCH(1), ARCH(2), and ARCH(3) models to the residuals. If the ARCH model really is inappropriate then all the parameters should be insignificant. The results of fitting these models to the residuals are given below in Figure 19, Figure 20 and Figure 21. A t-test can be used to test the null hypothesis that a particular term's coefficient is equal to zero given that all the other term are in the model. The alternative hypothesis is that that term's coefficient is not equal to zero given that all the rest of the terms are included in the model. Then, the p-values of the ARCH terms are 1, 1, 0.74, 1, 0.654, and 0.617. Since all of these p-values > $\alpha = 0.05$, H_0 is not rejected in favour of H_A. There is insufficient evidence to suggest any of these parameters are significant given the other terms are included in the model at the 5% level of significance. Therefore, the ARCH/GARCH model should not be fitted to the residuals.

Figure 19

```
Estimated Coefficient(s) of ARCH(1):
    Estimate  Std. Error  t value Pr(>|t|)
a0 1.833e+01   2.704e+00     6.78  1.2e-11
a1 4.901e-13   1.013e-01     0.00        1
```

Figure 20

```
Estimated Coefficient(s) of ARCH(2):
    Estimate  Std. Error  t value Pr(>|t|)
a0 1.737e+01   3.057e+00    5.682 1.33e-08
a1 3.367e-14   1.003e-01    0.000     1.00
a2 3.510e-02   1.057e-01    0.332     0.74
```

Figure 21

```
Estimated Coefficient(s) of ARCH(3):
    Estimate  Std. Error  t value Pr(>|t|)
a0 1.641e+01   3.257e+00    5.037 4.72e-07
a1 2.717e-13   9.887e-02    0.000    1.000
a2 4.931e-02   1.099e-01    0.449    0.654
a3 3.699e-02   7.397e-02    0.500    0.617
```

Conclusion

After generating a completely new time series, I had to find a model that best describes the time series. Initially, the time series appeared to be non-stationary. Despite seemingly fluctuating about a constant mean, the time series displayed signs of heteroscedasticity. I chose to fit an AR(3) model to the time series. This fit the data well, as the residuals appeared to be independent and normally distributed with zero mean. However, they exhibited signs of non-constant variance. Then, I examined the possible reasons why the time series appeared to be non-stationary. I concluded that the time series was not difference-stationary and the trend terms were not influential. As a consequence, the heteroscedasticity of the residuals was accepted as the main cause for the apparent non-stationarity. In order to model this heteroscedasticity, I tried to implement an ARCH model. However, it was quite clear that the squared residuals were independent and thus, the ARCH/GARCH models were not of use.

As a result, my conclusion is that the apparent heteroscedasticity of the residuals must have been due to random variation. Consequently, the time series itself is stationary. It is unsurprising that the residuals could have appeared to have non-constant variance when the number of observations is only 150.

Finally, the model that fits the time series the best is the AR(3) with the following equation:

$$Y_t = 2.6725 + 0.1662*(Y_{t-1} - 2.6725) + 0.2415*(Y_{t-2} - 2.6725) + 0.2868*(Y_{t-3} - 2.6725) + \varepsilon_t$$

It takes into account past values and random variation, which are the only factors I know that could possibly have an effect on this time series. This model is a good fit since the residuals are white noise variables. They are independent and follow a Normal distribution with zero mean and constant variance.

Appendix

I referenced Dr. Patrick Murphy's lecture notes along with the book 'Time Series Analysis With Applications in R' by Jonathan D. Cryer and Kung-Sik Chan throughout this report.

Also, all output provided in this report was due to operations performed using the R software package. The code that produced the outputs in this report is given below:

```
#Simulate AR(3) with phi1=1/4, phi=1/3, phi3=1/4, mean=2 and sigma=3
set.seed(20)
ep1=rnorm(250, mean=0, sd=3)
ar1=c()
ar1[1]=2+ep1[101]
ar1[2]=2+(1/4)*(ar1[1]-2)+ep1[102]
ar1[3]=2+(1/4)*(ar1[2]-2) +(1/3)*(ar1[1]-2)+ep1[103]
for(i in 1:147) ar1[i+3]=2+(1/4)*(ar1[i+2]-2)+(1/3)*(ar1[i+1]-
2)+(1/4)*(ar1[i]-2)+ep1[i+103]

#Simulate AR(3) with phi1=1/4, phi=1/3, phi3=1/4, mean=2 and sigma=5
ep2=rnorm(250, mean=0, sd=5)
ar2=c()
ar2[1]=2+ep2[101]
ar2[2]=2+(1/4)*(ar2[1]-2)+ep2[102]
ar2[3]=2+(1/4)*(ar2[2]-2) +(1/3)*(ar2[1]-2)+ep2[103]
for(i in 1:147) ar2[i+3]=2+(1/4)*(ar2[i+2]-2)+(1/3)*(ar2[i+1]-
2)+(1/4)*(ar2[i]-2)+ep2[i+103]

#Simulate AR(3) with phi1=1/4, phi=1/3, phi3=1/4, mean=2 and sigma=4
ep3=rnorm(250, mean=0, sd=4)
ar3=c()
ar3[1]=2+ep3[101]
ar3[2]=2+(1/4)*(ar3[1]-2)+ep3[102]
ar3[3]=2+(1/4)*(ar3[2]-2) +(1/3)*(ar3[1]-2)+ep3[103]
for(i in 1:147) ar3[i+3]=2+(1/4)*(ar3[i+2]-2)+(1/3)*(ar3[i+1]-
2)+(1/4)*(ar3[i]-2)+ep3[i+103]

#Concatenate 50 observations each from the 3 AR(3) time series
AR=c()
for(i in 1:50) AR[i]=ar1[i+100]
for(i in 1:50) AR[i+50]=ar2[i+100]
for(i in 1:50) AR[i+100]=ar3[i+100]

#Plots of new time series
plot.ts(AR, main= "Plot of Revised Time Series")
acf(AR, main= "ACF of Time Series")
pacf(AR, main= "PACF of Time Series")
library("tseries")

#Fit an ARIMA model - Looks like it could be AR(3)
arima(AR, c(3,0,0))
arima(AR, c(4,0,0))
arima(AR, c(3,0,1))

#Pick Best ARIMA model as ARIMA(3,0,0)
Fit=arima(AR, c(3,0,0))

#Conduct Residual Analysis
Residuals=Fit$res
plot(Residuals, main= "Plot of Residuals")
acf(Residuals, main= "ACF of Residuals")
```

```
qqnorm(Residuals, main= "QQ Plot of Residuals");qqline(Residuals,
col="red")
hist(Residuals, main="Histogram of Residuals")
Box.test(Residuals,lag=30,type=c("Ljung-Box"), fitdf=3)
Box.test(Residuals,lag=60,type=c("Ljung-Box"), fitdf=3)
Box.test(Residuals,lag=90,type=c("Ljung-Box"), fitdf=3)
Box.test(Residuals,lag=120,type=c("Ljung-Box"), fitdf=3)
Box.test(Residuals,lag=149,type=c("Ljung-Box"), fitdf=3)
shapiro.test(Residuals)

#Check for Unit Roots
diff1=diff(AR,1)
plot.ts(diff1, main="Plot of Differenced Time Series")
diff2=diff(AR,2)
plot.ts(diff2, main="Plot of Twice Differenced Time Series")
acf(diff1, main="ACF of Differenced Time Series")
adf.test(AR)

#Check for Trend Term
t = 1:150
linear=lm(AR~t)
summary(linear)
tsq = (1:150)^2
quadratic=lm(AR~tsq)
summary(quadratic)
tcub = (1:150)^3
cubic=lm(AR~tcub)
summary(cubic)
trend = lm(AR~t+tsq+tcub)
summary(trend)
summary(trend)$r.squared
summary(trend)$adj.r.square

#ARCH/GARCH
#Examine Squared Residuals
SqResiduals=c()
for(i in 1:150) SqResiduals[i]=(Residuals[i])^2
plot.ts(SqResiduals, main= "Plot of Squared Residuals")
acf(SqResiduals, main= "ACF of Squared Residuals")
pacf(SqResiduals, main= "PACF of Squared Residuals")
Box.test(SqResiduals,lag=30,type=c("Ljung-Box"), fitdf=3)
Box.test(SqResiduals,lag=60,type=c("Ljung-Box"), fitdf=3)
Box.test(SqResiduals,lag=90,type=c("Ljung-Box"), fitdf=3)
Box.test(SqResiduals,lag=120,type=c("Ljung-Box"), fitdf=3)
Box.test(SqResiduals,lag=149,type=c("Ljung-Box"), fitdf=3)
garch(Residuals, order=c(1,1))
garch(Residuals, order=c(0,1))
for(i in 1:10)garch(Residuals, order=c(0,i))
Arch1=garch(Residuals,c(0,1))
Arch2=garch(Residuals,c(0,2))
Arch3=garch(Residuals,c(0,3))
summary(Arch1)
summary(Arch2)
summary(Arch3)
```