# ENV 790.30 - Time Series Analysis for Energy Data | Spring 2021

Assignment 2 - Due date 02/05/21

Abi Vanover

```
pacman::p_load(forecast, tseries, dplyr, lubridate)
```

## Data set information

Consider the data provided in the spreadsheet "Table_10.1_Renewable_Energy_Production_and_Consumption_by_Source on our **Data** folder. The data comes from the US Energy Information and Administration and corresponds to the January 2021 Monthly Energy Review. The spreadsheet is ready to be used. Use the command *read.table*() to import the data in R or *panda.read_excel*() in Python (note that you will need to import pandas package). }

```
#Importing data set

library(readxl)
Df1 <- read_excel("../Data/Table_10.1_Renewable_Energy_Production_and_Consumption_by_Source.xlsx",skip =
View(Df1)
```

## Question 1

You will work only with the following columns: Total Biomass Energy Production, Total Renewable Energy Production, Hydroelectric Power Consumption. Create a data frame structure with these three time series only. Use the command head() to verify your data.

```
dates <- Df1[2:575, 1]
variables <- Df1[2:575, 4:6]

workingdf <- cbind(dates,variables)

head(workingdf)
```

```
##        Month Total Biomass Energy Production Total Renewable Energy Production
## 1 1973-01-01                         129.787                           403.981
## 2 1973-02-01                         117.338                             360.9
## 3 1973-03-01                         129.938                           400.161
## 4 1973-04-01                         125.636                            380.47
## 5 1973-05-01                         129.834                           392.141
## 6 1973-06-01                         125.611                           377.232
##   Hydroelectric Power Consumption
## 1                         272.703
## 2                         242.199
## 3                          268.81
```

```
## 4                          253.185
## 5                           260.77
## 6                          249.859
```

```r
biomass <- as.numeric(variables$`Total Biomass Energy Production`)
renewable <- as.numeric(variables$`Total Renewable Energy Production`)
hydro <- as.numeric(variables$`Hydroelectric Power Consumption`)
```

## Question 2

```r
#creating each variable as an individual ts to use for plotting later
tsbiomass <- ts(workingdf$`Total Biomass Energy Production`, start = 1973, end = 2020, frequency = 12)

tsrenew <- ts(variables$`Total Renewable Energy Production`, start = 1973, end = 2020, frequency = 12)

tshydro <- ts(variables$`Hydroelectric Power Consumption`, start = 1973, end = 2020, frequency = 12)


#transforming my dataframe into one time series object
tsdf <- ts(workingdf[2:575, 2:4], start = 1973, end = 2020, frequency= 12)
```

## Question 3

Compute mean and standard deviation for these three series.

```r
#Trying again for mean & Sd, converting to numeric

x <- mean(as.numeric(tsbiomass))
sd(as.numeric(tsbiomass))
```

```
## [1] 86.71481
```

```r
y <- mean(as.numeric(tsrenew))
sd(as.numeric(tsrenew))
```

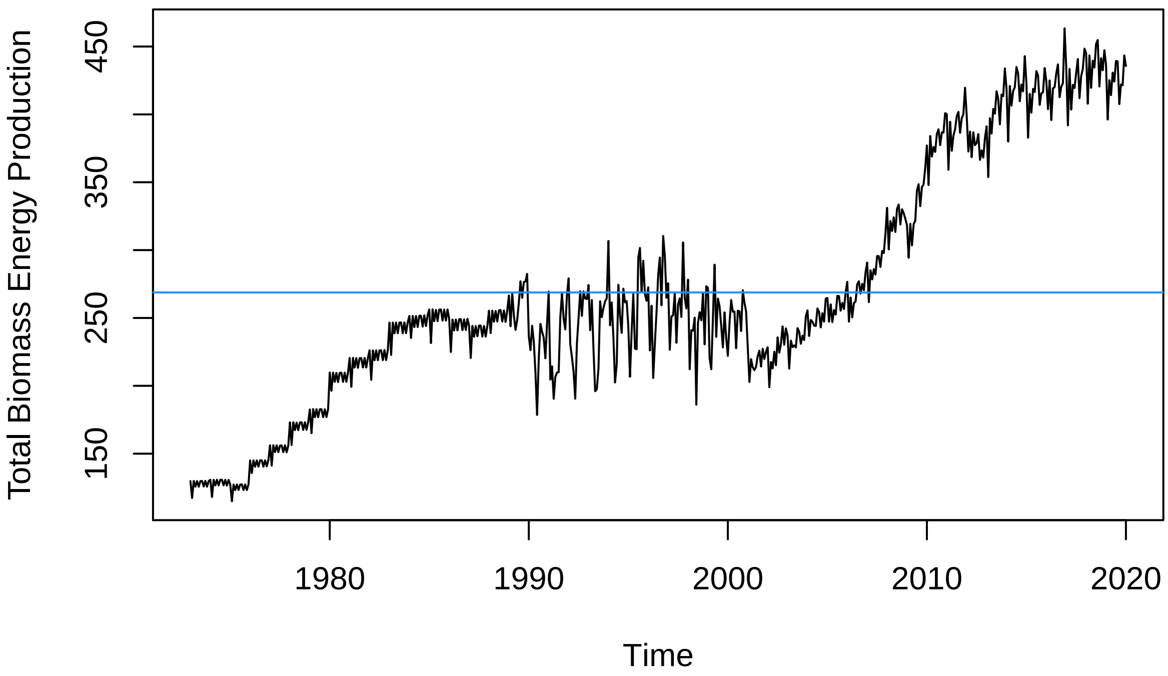
```
## [1] 161.601
```

```r
z <- mean(as.numeric(tshydro))
sd(as.numeric(tshydro))
```

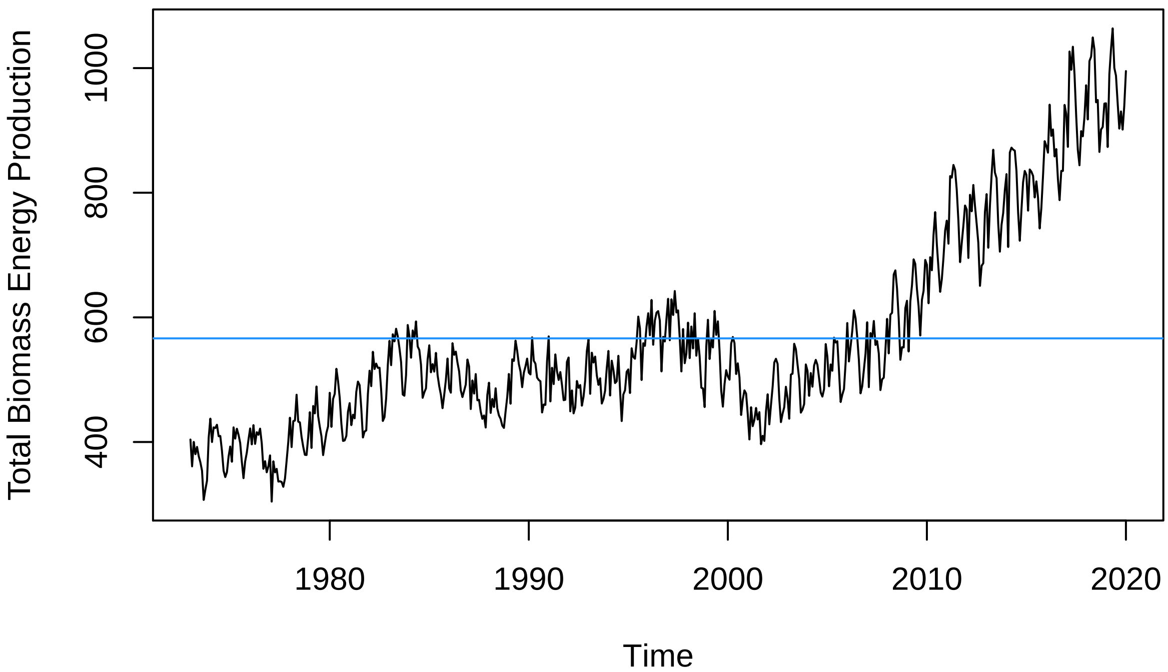
```
## [1] 43.9582
```

## Question 4

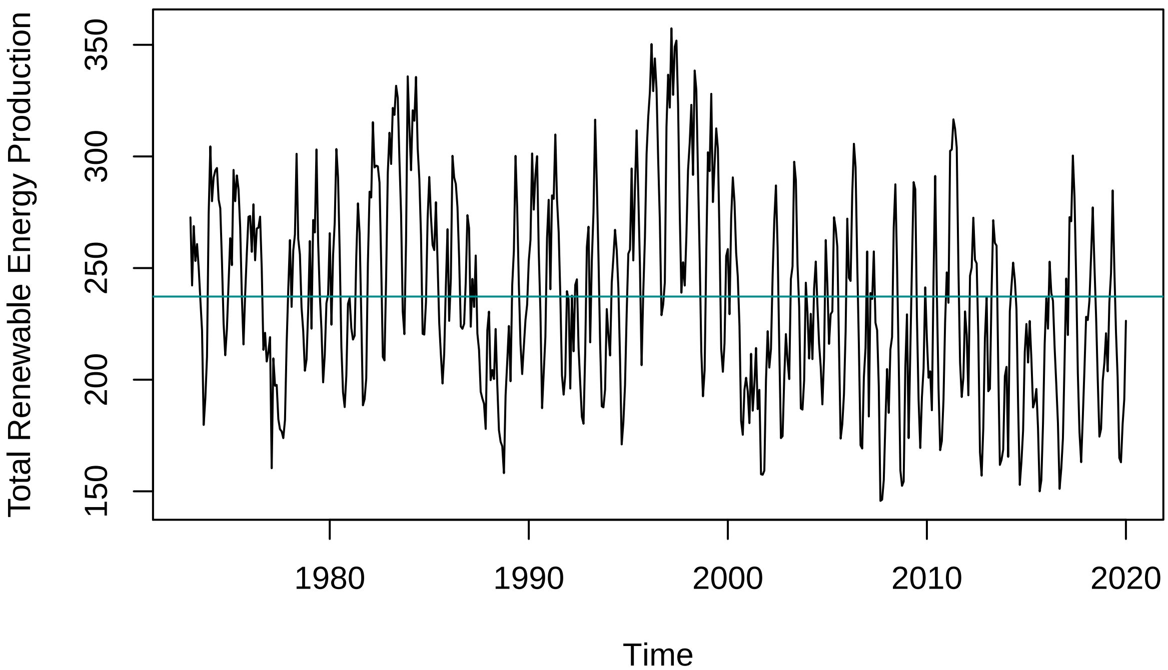
```r
#This is working now

plot(tsbiomass, xlab = "Time", ylab = "Total Biomass Energy Production")
abline(h=x, col = "dodgerblue")
```




```r
#This plot is fine and just needs a line
plot(tsrenew, xlab = "Time", ylab = "Total Biomass Energy Production")
abline(h=y, col = "dodgerblue")
```




```r
#This plot is fine and just needs a line to get added
plot(tshydro, xlab = "Time", ylab = "Total Renewable Energy Production")
abline(h=z, col = "darkcyan")
```




### Question 5

Total biomass and total renewables are significantly positively correlated, with a value of ~.923. This makes sense as total biomass would be a component of total renewables. Renewables and Hydro consumption are not correlated; the value is essentially zero. Biomass production and Hydroelectric power consumption are slightly negatively correlated, which is reasonable as well: as biomass energy production increases, it can be inferred that it would replace some hydroelectric power.

```
##
##  Pearson's product-moment correlation
##
## data:  biomass and renewable
## t = 57.562, df = 572, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.9104276 0.9346626
## sample estimates:
##       cor
## 0.9234609
```

```
##
##  Pearson's product-moment correlation
##
## data:  renewable and hydro
## t = -0.065935, df = 572, p-value = 0.9475
```

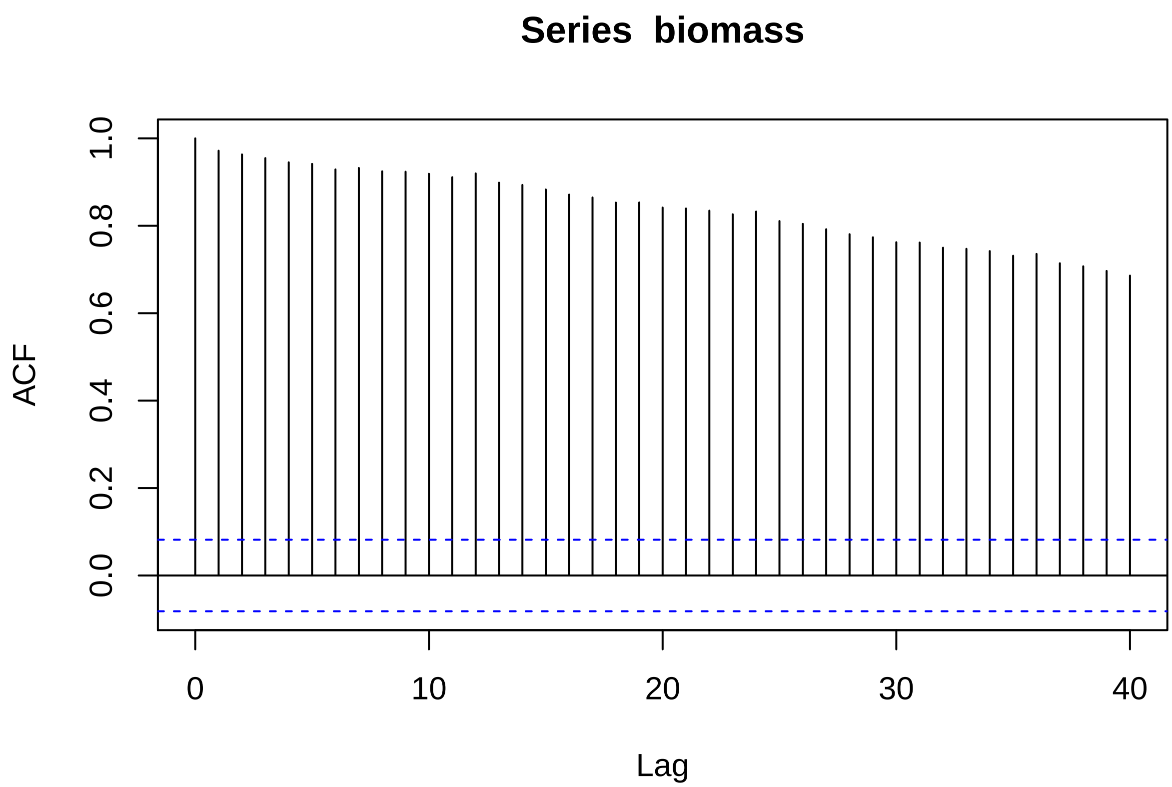
```
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.08457627  0.07909949
## sample estimates:
##          cor
## -0.002756852
```

```
##
##  Pearson's product-moment correlation
##
## data:  biomass and hydro
## t = -6.3222, df = 572, p-value = 5.195e-10
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.3304936 -0.1774402
## sample estimates:
##        cor
## -0.2555675
```

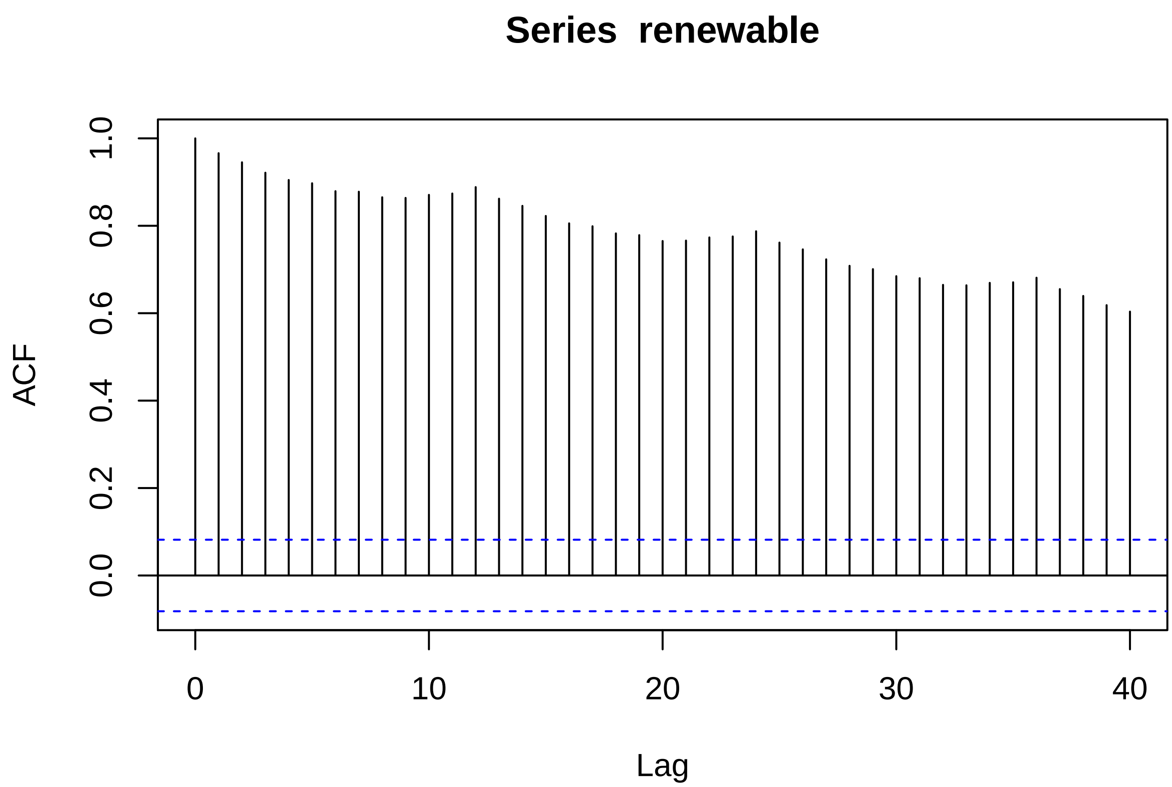
## Question 6

The plots for Biomass and Total Renewables are very similar, with high initial values of autocorrelation that decrease over time. However, the plot for Hydroelectric Power Consumption is significantly different. The autocorrelation plot for Hydro shows several peaks and valleys, with maximums decreasing over time. Several of these values are also negative, which were not present on either of the two previous plots.

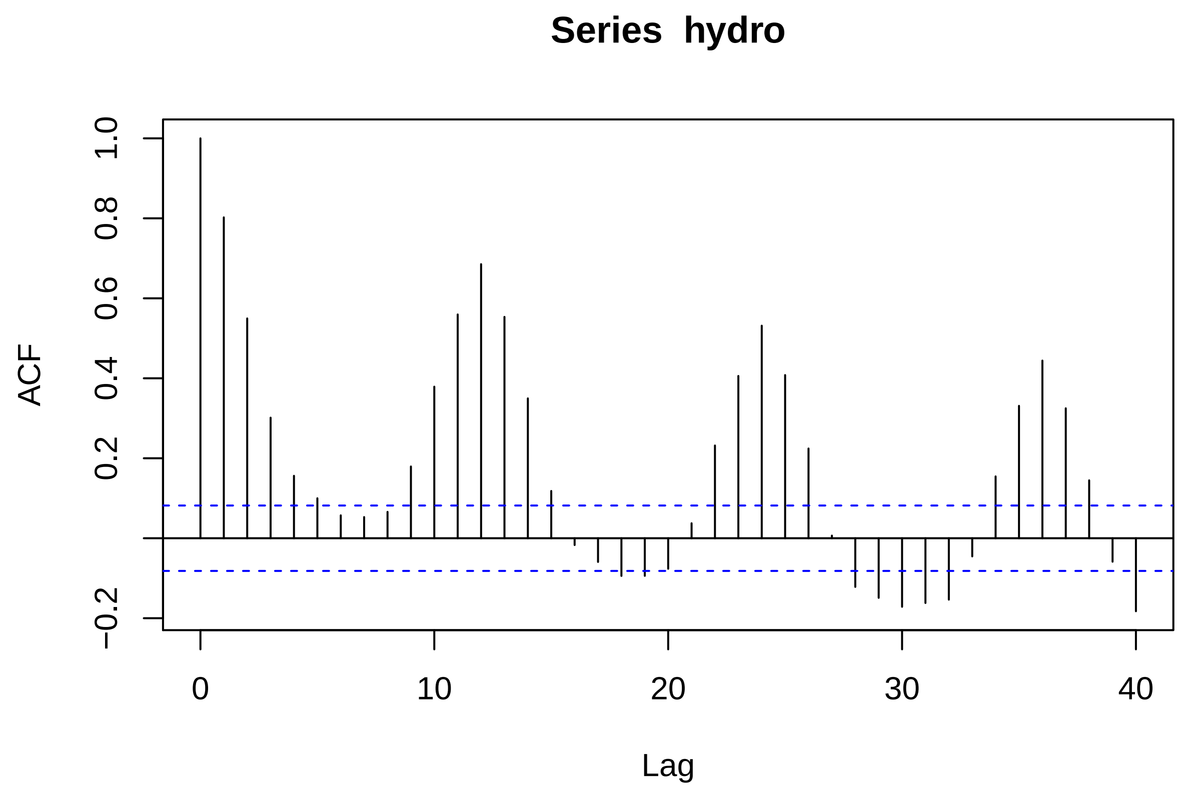
```
acfb=acf(biomass, lag.max = 40, type = "correlation") #decreasing at a shallow slope
```




```
acfr=acf(renewable, lag.max = 40, type = "correlation") #smoother decline
```




```
acfh=acf(hydro, lag.max = 40, type = "correlation") #decreasing sharper, wave pattern
```




```
acfb
```

```
##
## Autocorrelations of series 'biomass', by lag
##
##     0     1     2     3     4     5     6     7     8     9    10    11    12
## 1.000 0.972 0.963 0.955 0.945 0.942 0.929 0.932 0.925 0.924 0.919 0.911 0.920
##    13    14    15    16    17    18    19    20    21    22    23    24    25
## 0.899 0.893 0.883 0.871 0.865 0.853 0.853 0.842 0.839 0.835 0.826 0.833 0.811
##    26    27    28    29    30    31    32    33    34    35    36    37    38
## 0.804 0.792 0.781 0.774 0.762 0.762 0.750 0.747 0.742 0.731 0.736 0.714 0.707
##    39    40
## 0.697 0.686
```

```
acfr
```

```
##
## Autocorrelations of series 'renewable', by lag
##
##     0     1     2     3     4     5     6     7     8     9    10    11    12
## 1.000 0.966 0.945 0.921 0.905 0.897 0.879 0.878 0.865 0.864 0.871 0.874 0.889
##    13    14    15    16    17    18    19    20    21    22    23    24    25
## 0.862 0.846 0.823 0.806 0.799 0.783 0.779 0.765 0.766 0.773 0.776 0.788 0.762
##    26    27    28    29    30    31    32    33    34    35    36    37    38
## 0.746 0.723 0.708 0.701 0.685 0.680 0.665 0.664 0.669 0.671 0.681 0.655 0.640
```

```
##    39    40
## 0.618 0.604
```

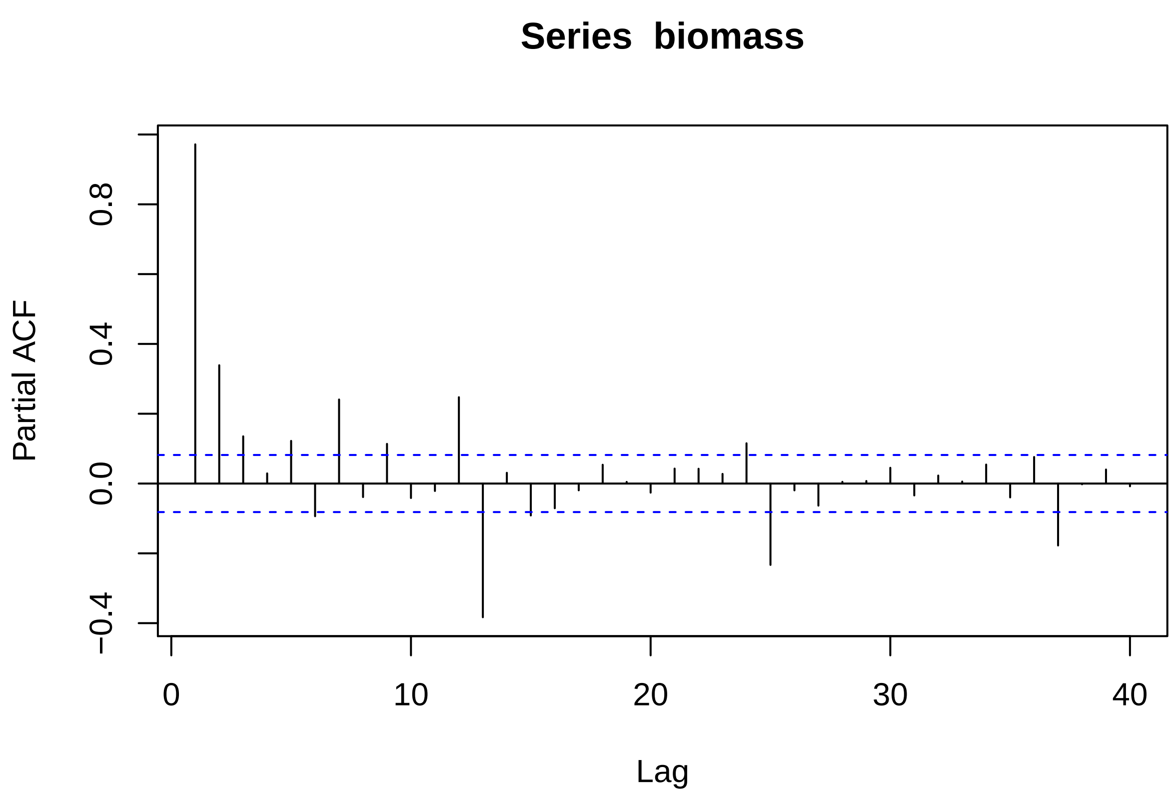
```
acfh
```

```
##
## Autocorrelations of series 'hydro', by lag
##
##      0      1      2      3      4      5      6      7      8      9     10
##  1.000  0.802  0.550  0.302  0.156  0.100  0.057  0.053  0.066  0.179  0.379
##     11     12     13     14     15     16     17     18     19     20     21
##  0.560  0.685  0.554  0.350  0.118 -0.017 -0.059 -0.094 -0.094 -0.076  0.037
##     22     23     24     25     26     27     28     29     30     31     32
##  0.232  0.406  0.532  0.408  0.224  0.007 -0.122 -0.149 -0.171 -0.162 -0.154
##     33     34     35     36     37     38     39     40
## -0.045  0.155  0.331  0.444  0.325  0.145 -0.059 -0.183
```

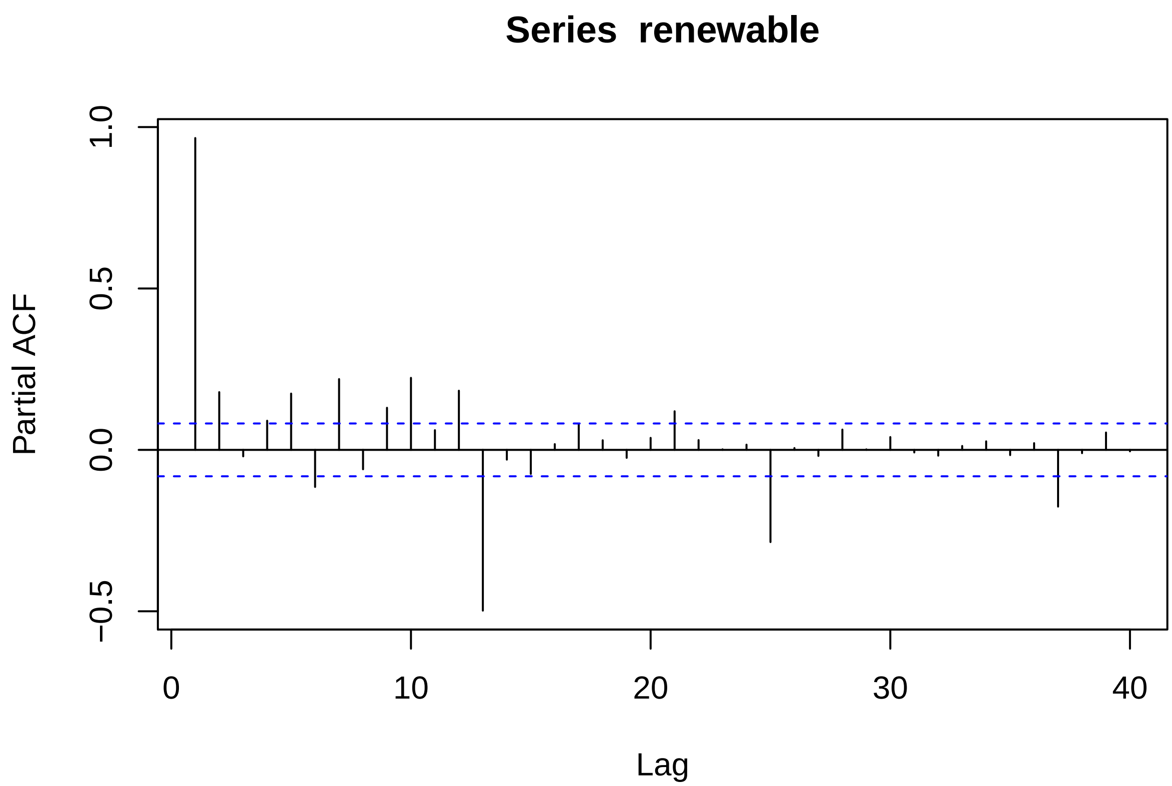
### Question 7

These plots are all very similar to each other, but differ sharply from the initial ACF plots for the Biomass & Total Renewables. For Biomass & Renewables, the ACF plots had a steadily decreasing trendline of all positive values. However, the PACF shows a more cyclical trend of values, with both negative and positive values. For all plots though, the absolute values of the PACF's are lower than the absolute values of ACF's.

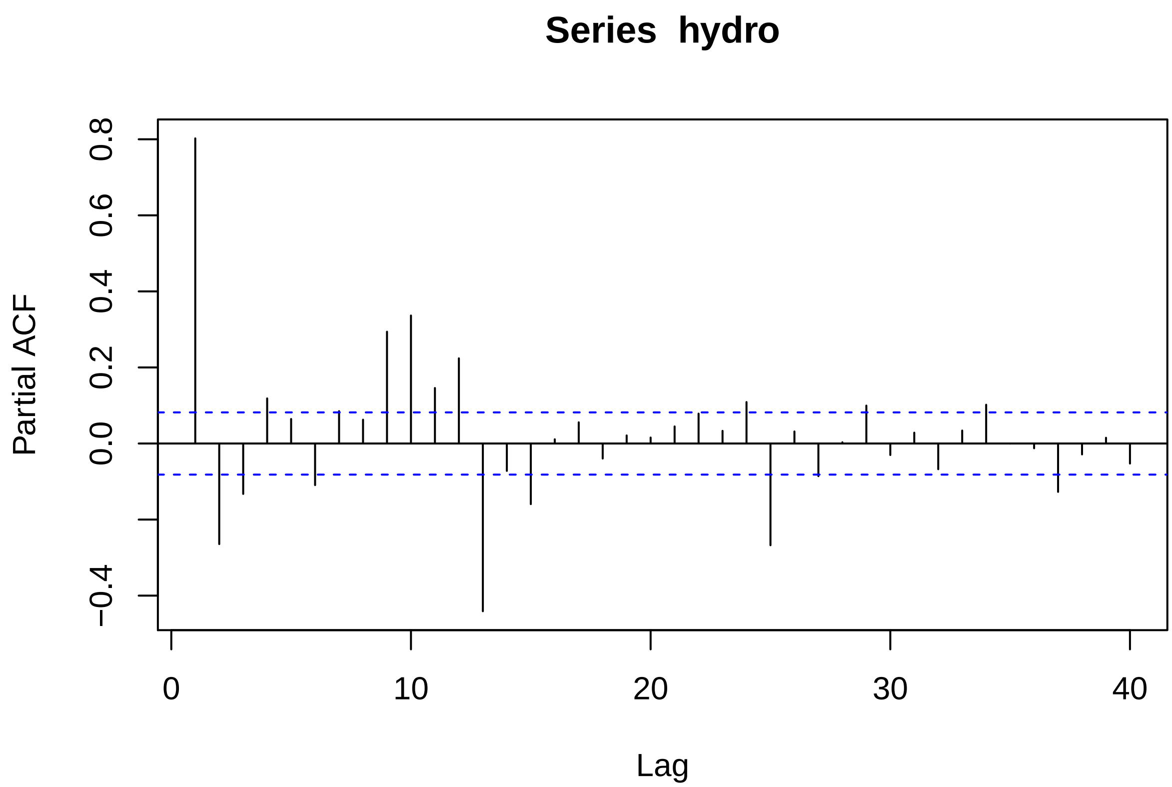
```
pacfb=pacf(biomass, lag.max = 40) #shrinking waves, decreasing over time
```



```r
pacfr=pacf(renewable, lag.max = 40) #similar pattern, but sharper
```




```r
pacfh=pacf(hydro, lag.max = 40) #similar pattern, but slightly wider range
```




```
pacfb
```

```
##
## Partial autocorrelations of series 'biomass', by lag
##
##      1      2      3      4      5      6      7      8      9     10     11
##  0.972  0.339  0.135  0.029  0.122 -0.094  0.241 -0.039  0.114 -0.042 -0.021
##     12     13     14     15     16     17     18     19     20     21     22
##  0.247 -0.383  0.031 -0.091 -0.071 -0.019  0.054  0.005 -0.026  0.043  0.042
##     23     24     25     26     27     28     29     30     31     32     33
##  0.028  0.115 -0.233 -0.020 -0.063  0.005  0.007  0.045 -0.034  0.023  0.006
##     34     35     36     37     38     39     40
##  0.054 -0.040  0.076 -0.178 -0.002  0.040 -0.008
```

```
pacfr
```

```
##
## Partial autocorrelations of series 'renewable', by lag
##
##      1      2      3      4      5      6      7      8      9     10     11
##  0.966  0.179 -0.020  0.090  0.174 -0.115  0.219 -0.060  0.130  0.223  0.061
##     12     13     14     15     16     17     18     19     20     21     22
##  0.183 -0.498 -0.030 -0.074  0.018  0.079  0.030 -0.025  0.037  0.120  0.030
##     23     24     25     26     27     28     29     30     31     32     33
##  0.002  0.016 -0.286  0.006 -0.019  0.063  0.002  0.040 -0.008 -0.018  0.012
```

```
##     34     35     36     37     38     39     40 
##  0.027 -0.016  0.021 -0.176 -0.010  0.054 -0.004
```

```
pacfh
```

```
## 
## Partial autocorrelations of series 'hydro', by lag
## 
##      1      2      3      4      5      6      7      8      9     10     11 
##  0.802 -0.265 -0.133  0.119  0.064 -0.109  0.085  0.063  0.294  0.337  0.146 
##     12     13     14     15     16     17     18     19     20     21     22 
##  0.224 -0.441 -0.072 -0.159  0.011  0.056 -0.040  0.021  0.016  0.045  0.079 
##     23     24     25     26     27     28     29     30     31     32     33 
##  0.033  0.109 -0.268  0.032 -0.086  0.003  0.100 -0.030  0.029 -0.068  0.034 
##     34     35     36     37     38     39     40 
##  0.102  0.000 -0.013 -0.127 -0.029  0.015 -0.053
```