ENV 790.30 - Time Series Analysis for Energy Data | Spring 2026

Assignment 2 - Due date 01/27/26

Leo Zhang

Submission Instructions

You should open the .rmd file corresponding to this assignment on RStudio. The file is available on our class repository on Github.

Once you have the file open on your local machine the first thing you will do is rename the file such that it includes your first and last name (e.g., 'LuanaLima_TSA_A02_Sp24.Rmd'). Then change 'Student Name' on line 4 with your name.

Then you will start working through the assignment by **creating code and output** that answer each question. Be sure to use this assignment document. Your report should contain the answer to each question and any plots/tables you obtained (when applicable).

When you have completed the assignment, **Knit** the text and code into a single PDF file. Submit this pdf using Sakai.

R packages

R packages needed for this assignment:'forecast', 'tseries', and 'dplyr'. Install these packages, if you haven't done yet. Do not forget to load them before running your script, since they are NOT default packages.\

```
library(forecast)
library(gt)
library(readxl)
library(tidyverse)
library(tseries)
```

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 December 2023 Monthly Energy Review. The spreadsheet is ready to be used. You will also find a *.csv* version of the data install.packages(forecast) install.packages(tseries). You may use the function *read.table*() to import the *.csv* data in R. Or refer to the file 'M2_ImportingData_CSV_XLSX.Rmd' in our Lessons folder for functions that are better suited for importing the *.xlsx*.

```
outlook <- read_excel('C:/Users/24169/Desktop/e797/797_spring26/Data/Table_10.1_Renewable_Energy_Produc
```

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.

```
outlook <- outlook[-c(1), ]
outlook <- outlook %>% select('Total Biomass Energy Production', 'Total Renewable Energy Production', 'H
outlook <- outlook %>% mutate(across(c('Total Biomass Energy Production', 'Total Renewable Energy Produc
head(outlook)
```

```
## # A tibble: 6 x 3
##   Total Biomass Energy Productio~1 Total Renewable Ener~2 Hydroelectric Power ~3
##                              <dbl>                  <dbl>                  <dbl>
## 1                            1549.                  1873.                   323.
## 2                            1562.                  1907.                   344.
## 3                            1535.                  1891.                   356.
## 4                            1474.                  1849.                   374.
## 5                            1419.                  1793.                   374.
## 6                            1394.                  1775.                   381.
## # i abbreviated names: 1: ‘Total Biomass Energy Production‘,
## #   2: ‘Total Renewable Energy Production‘,
## #   3: ‘Hydroelectric Power Consumption‘
```

Question 2

Transform your data frame in a time series object and specify the starting point and frequency of the time series using the function ts().

```
history <- ts(outlook, start = 1949, end = 2023, frequency = 1)
history
```

```
## Time Series:
## Start = 1949
## End = 2023
## Frequency = 1
##      Total Biomass Energy Production Total Renewable Energy Production
## 1949                        1549.262                          1872.627
## 1950                        1562.307                          1906.525
## 1951                        1534.669                          1890.800
## 1952                        1474.369                          1848.694
## 1953                        1418.601                          1792.616
## 1954                        1394.327                          1775.242
## 1955                        1424.143                          1820.740
## 1956                        1415.871                          1843.178
## 1957                        1333.581                          1788.598
## 1958                        1323.123                          1813.136
## 1959                        1352.874                          1834.493
## 1960                        1319.870                          1829.873
## 1961                        1294.762                          1825.773
## 1962                        1300.242                          1887.502
## 1963                        1323.316                          1900.483
```

```
## 1964                            1336.802                            1952.686
## 1965                            1334.761                            2007.517
## 1966                            1368.985                            2044.989
## 1967                            1340.249                            2108.853
## 1968                            1419.495                            2191.661
## 1969                            1440.487                            2307.418
## 1970                            1430.962                            2289.021
## 1971                            1432.323                            2353.833
## 1972                            1503.065                            2449.491
## 1973                            1529.068                            2475.544
## 1974                            1539.662                            2586.001
## 1975                            1498.734                            2544.167
## 1976                            1713.373                            2704.698
## 1977                            1838.332                            2613.474
## 1978                            2037.605                            3014.948
## 1979                            2151.906                            3131.031
## 1980                            2475.500                            3445.378
## 1981                            2596.283                            3515.922
## 1982                            2663.452                            3745.796
## 1983                            2904.414                            4069.164
## 1984                            2971.120                            4104.121
## 1985                            3016.233                            4018.174
## 1986                            2932.095                            3970.473
## 1987                            2874.884                            3774.443
## 1988                            3016.049                            3822.682
## 1989                            3159.357                            4206.136
## 1990                            2735.112                            3862.708
## 1991                            2781.797                            3901.431
## 1992                            2931.678                            3931.748
## 1993                            2908.446                            4007.312
## 1994                            3027.534                            4056.461
## 1995                            3099.082                            4294.583
## 1996                            3155.301                            4479.994
## 1997                            3107.908                            4466.063
## 1998                            2928.930                            4174.139
## 1999                            2965.132                            4202.601
## 2000                            3005.661                            4093.033
## 2001                            2624.162                            3513.751
## 2002                            2705.406                            3770.941
## 2003                            2804.778                            3914.186
## 2004                            2996.017                            4093.471
## 2005                            3101.186                            4220.160
## 2006                            3211.515                            4429.647
## 2007                            3472.080                            4582.116
## 2008                            3868.250                            5084.563
## 2009                            3956.617                            5309.146
## 2010                            4552.532                            5942.580
## 2011                            4712.436                            6404.003
## 2012                            4553.559                            6186.970
## 2013                            4834.795                            6560.783
## 2014                            5049.215                            6832.704
## 2015                            5024.958                            6839.536
## 2016                            5121.688                            7178.348
## 2017                            5155.644                            7494.525
```

```
## 2018                           5304.078                          7734.125
## 2019                           5204.933                          7742.749
## 2020                           4699.703                          7455.133
## 2021                           4903.546                          7797.231
## 2022                           5063.341                          8297.453
## 2023                           5164.910                          8437.463
##      Hydroelectric Power Consumption
## 1949                          323.365
## 1950                          344.218
## 1951                          356.131
## 1952                          374.325
## 1953                          374.015
## 1954                          380.915
## 1955                          396.597
## 1956                          427.307
## 1957                          455.017
## 1958                          490.013
## 1959                          481.619
## 1960                          509.889
## 1961                          530.690
## 1962                          586.917
## 1963                          576.594
## 1964                          615.189
## 1965                          672.111
## 1966                          675.363
## 1967                          767.525
## 1968                          770.679
## 1969                          864.834
## 1970                          856.267
## 1971                          919.641
## 1972                          941.469
## 1973                          939.769
## 1974                         1037.971
## 1975                         1034.357
## 1976                          978.986
## 1977                          762.919
## 1978                          967.183
## 1979                          965.855
## 1980                          952.569
## 1981                          900.238
## 1982                         1065.820
## 1983                         1144.012
## 1984                         1106.550
## 1985                          970.068
## 1986                         1003.146
## 1987                          862.745
## 1988                          771.456
## 1989                          927.985
## 1990                          999.258
## 1991                          986.048
## 1992                          863.536
## 1993                          957.046
## 1994                          887.549
## 1995                         1060.561
```

```
## 1996                          1184.517
## 1997                          1216.219
## 1998                          1103.221
## 1999                          1090.257
## 2000                           940.254
## 2001                           740.271
## 2002                           901.890
## 2003                           941.051
## 2004                           915.840
## 2005                           922.336
## 2006                           986.909
## 2007                           844.504
## 2008                           869.485
## 2009                           932.995
## 2010                           887.813
## 2011                          1089.639
## 2012                           942.532
## 2013                           916.345
## 2014                           884.959
## 2015                           849.861
## 2016                           913.775
## 2017                          1024.736
## 2018                           998.092
## 2019                           982.225
## 2020                           973.355
## 2021                           858.407
## 2022                           869.339
## 2023                           835.948
```

Question 3

Compute mean and standard deviation for these three series.

```
statement <- c('Mean', 'Standard Deviation')
biomass <- history[, 'Total Biomass Energy Production']
first_statistic <- c(mean(biomass), sd(biomass))
renewable <- history[, 'Total Renewable Energy Production']
second_statistic <- c(mean(renewable), sd(renewable))
hydro <- history[, 'Hydroelectric Power Consumption']
third_statistic <- c(mean(hydro), sd(hydro))
general_statistic <- data.frame(statement, first_statistic, second_statistic, third_statistic)
p1 <- general_statistic %>% gt() %>% tab_options(table.width = pct(100)) %>% tab_header(title = 'Figure
p1
```

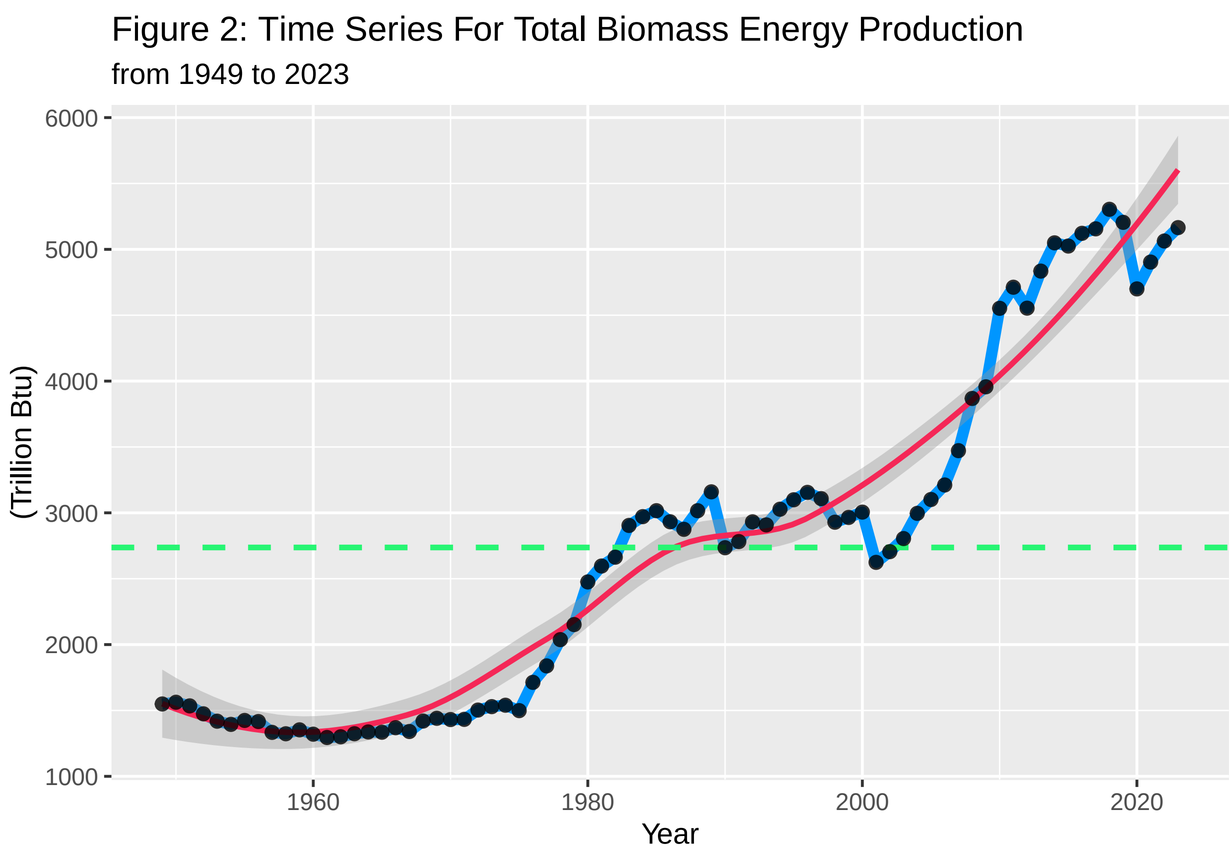
Question 4

Display and interpret the time series plot for each of these variables. Try to make your plot as informative as possible by writing titles, labels, etc. For each plot add a horizontal line at the mean of each series in a different color.

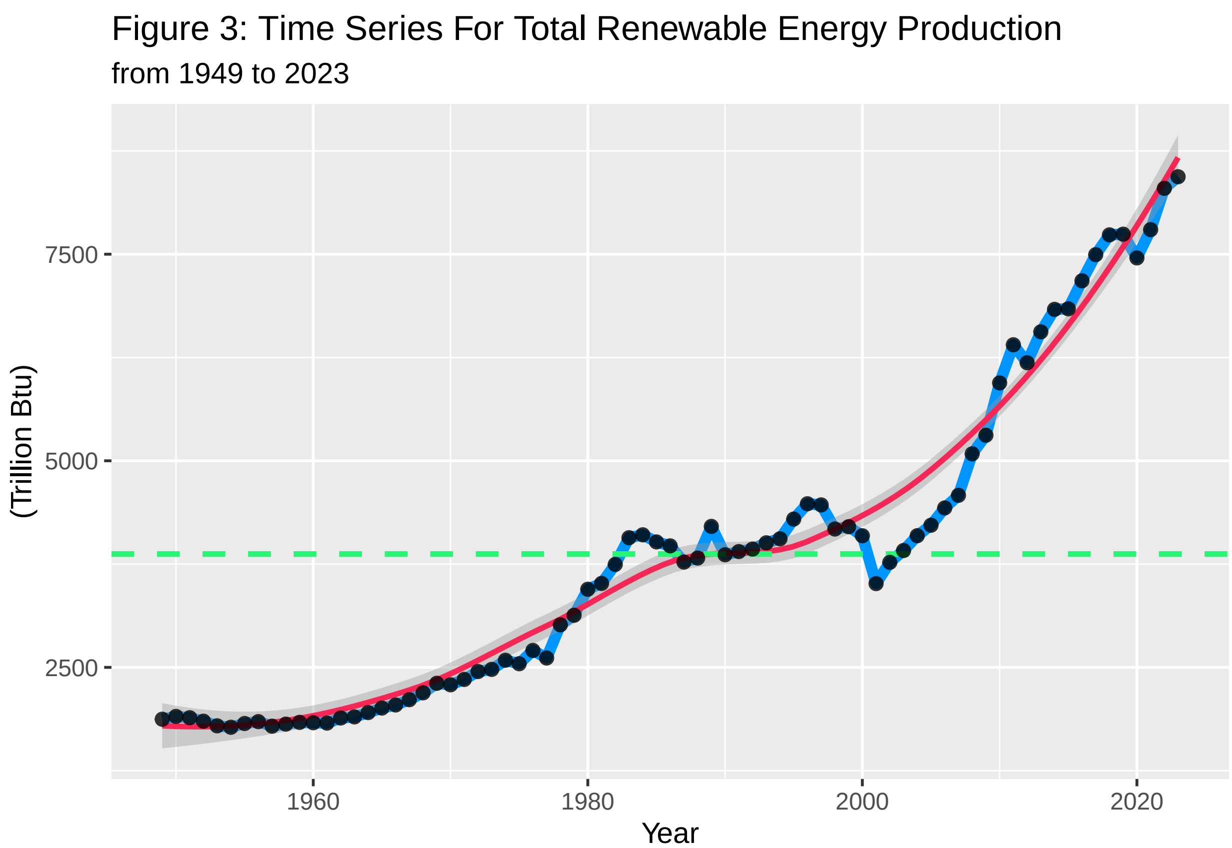
Figure 1: Summary Statistics
time series from 1949 to 2023

	Total Biomass Energy Production	Total Renewable Energy Production	Hydroele
Mean	2737.046	3871.594	
Standard Deviation	1298.192	1899.766	

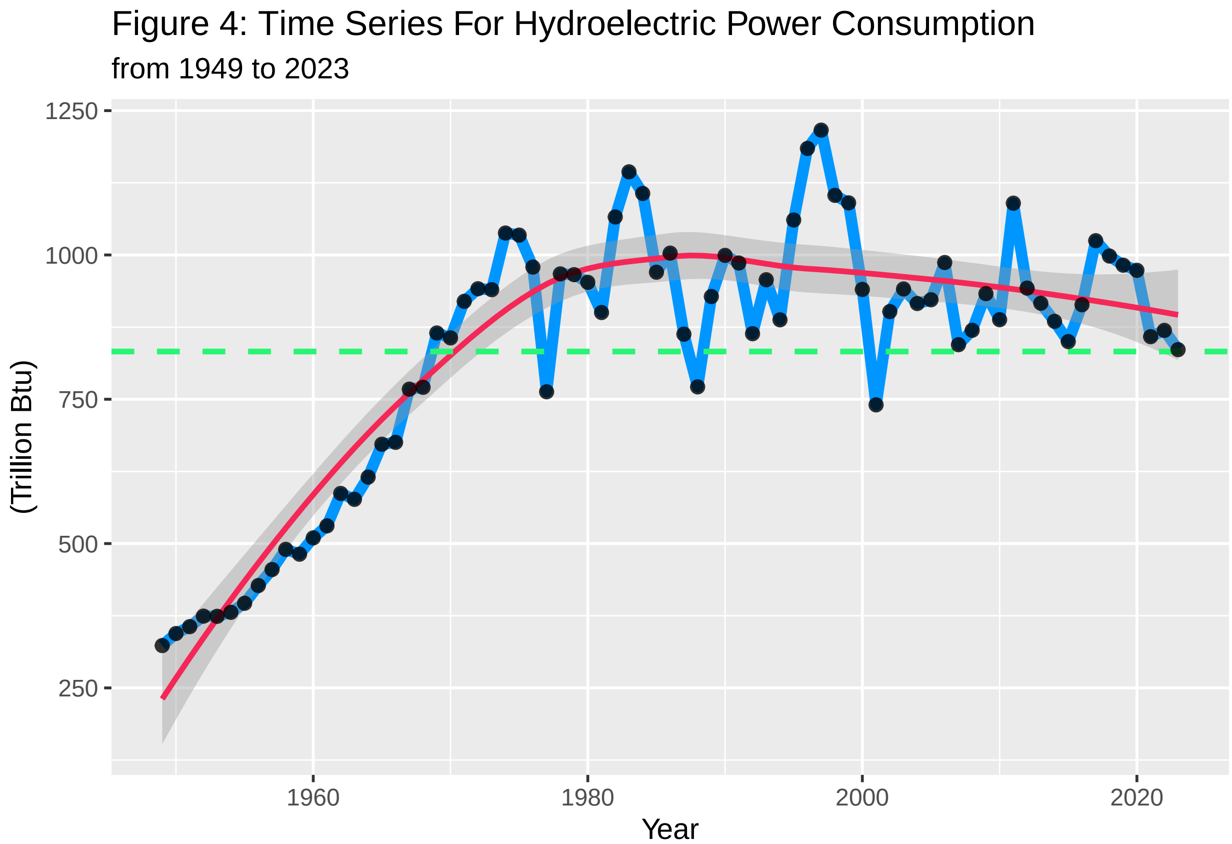
```
p2 <- autoplot(biomass) + geom_line(color = '#0096FF', linewidth = 2) + geom_smooth(color = '#F52757', 
p2
```




```
p3 <- autoplot(renewable) + geom_line(color = '#0096FF', linewidth = 2) + geom_smooth(color = '#F52757'
p3
```




```
p4 <- autoplot(hydro) + geom_line(color = '#0096FF', linewidth = 2) + geom_smooth(color = '#F52757', lin
p4
```



For 75 years, the total biomass energy production has been increasing. Particularly, between 1970 and 1980, the total biomass energy production increased rapidly. The growth was more considerable after 2000. For 75 years, the total renewable energy production has also been increasing. The growth was also more considerable after 2000. Before 1980, the hydroelectric power consumption seemed to increase steadily. However, after 1980, with apparent fluctuations, the hydroelectric power consumption had a declining tendency.

Question 5

Compute the correlation between these three series. Are they significantly correlated? Explain your answer.

```
correlate <- cor(history)
correlate
```

```
##                                  Total Biomass Energy Production
## Total Biomass Energy Production                        1.0000000
## Total Renewable Energy Production                      0.9861982
## Hydroelectric Power Consumption                        0.5276925
##                                  Total Renewable Energy Production
## Total Biomass Energy Production                          0.9861982
## Total Renewable Energy Production                        1.0000000
## Hydroelectric Power Consumption                          0.5457158
##                                  Hydroelectric Power Consumption
## Total Biomass Energy Production                        0.5276925
```

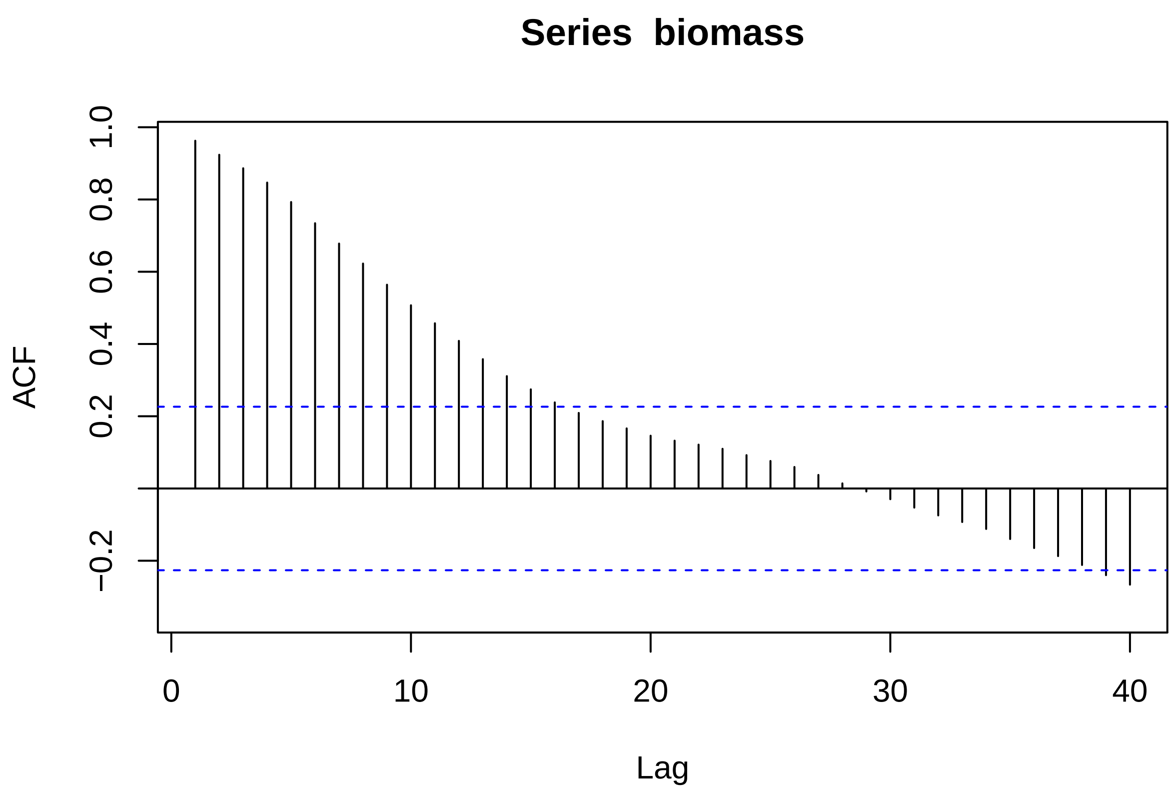
```
## Total Renewable Energy Production                        0.5457158
## Hydroelectric Power Consumption                          1.0000000
```

The correlation coefficient between total biomass energy production and total renewable energy production is 0.9861982, which may indicate a very strong positive correlation. There is a moderate positive correlation between hydroelectric power consumption and the rest.

Question 6

Compute the autocorrelation function from lag 1 up to lag 40 for these three variables. What can you say about these plots? Do the three of them have the same behavior?

```
first_acf <- Acf(biomass, lag.max = 40, type = 'correlation', plot = TRUE)
```

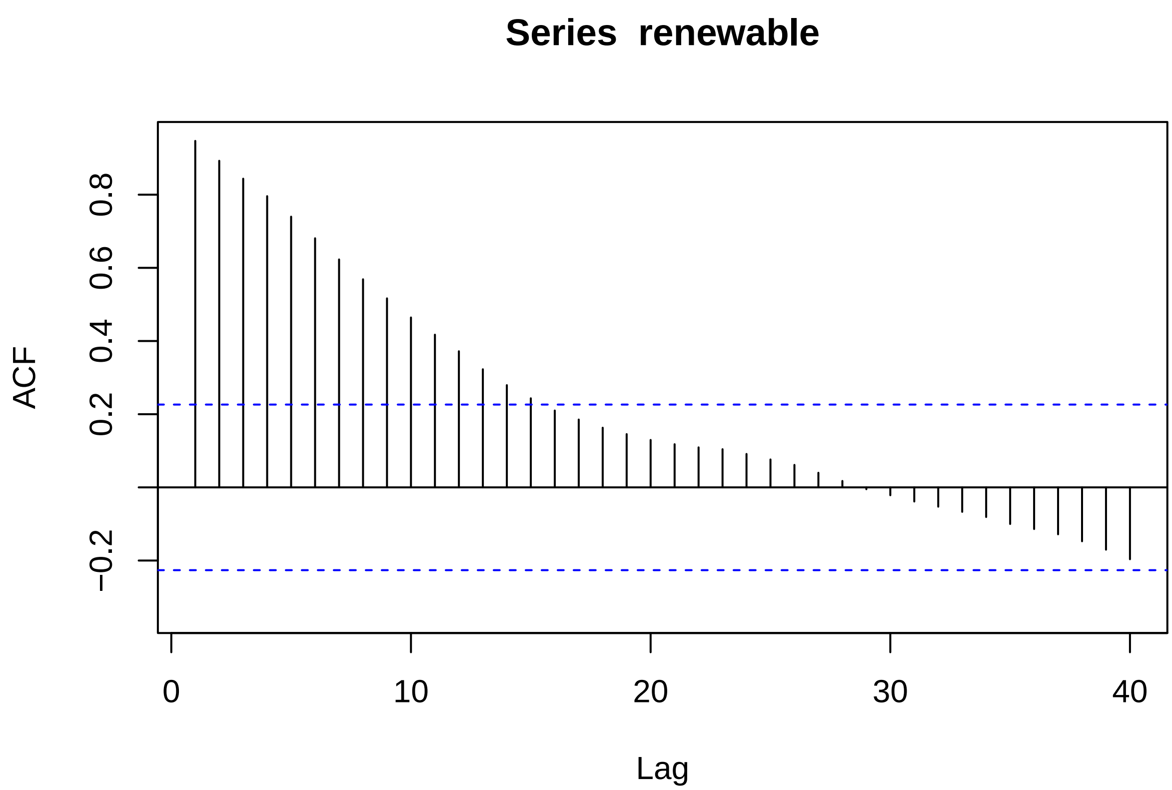



```
first_acf
```

```
##
## Autocorrelations of series 'biomass', by lag
##
##     0     1     2     3     4     5     6     7     8     9    10
## 1.000 0.963 0.924 0.887 0.847 0.793 0.734 0.678 0.623 0.564 0.507
##    11    12    13    14    15    16    17    18    19    20    21
## 0.457 0.409 0.358 0.311 0.274 0.238 0.209 0.186 0.166 0.146 0.132
##    22    23    24    25    26    27    28    29    30    31    32
```

```
##  0.121  0.110  0.092  0.076  0.060  0.038  0.014 -0.008 -0.030 -0.053 -0.075
##     33     34     35     36     37     38     39     40
## -0.093 -0.112 -0.140 -0.165 -0.187 -0.212 -0.240 -0.266
```

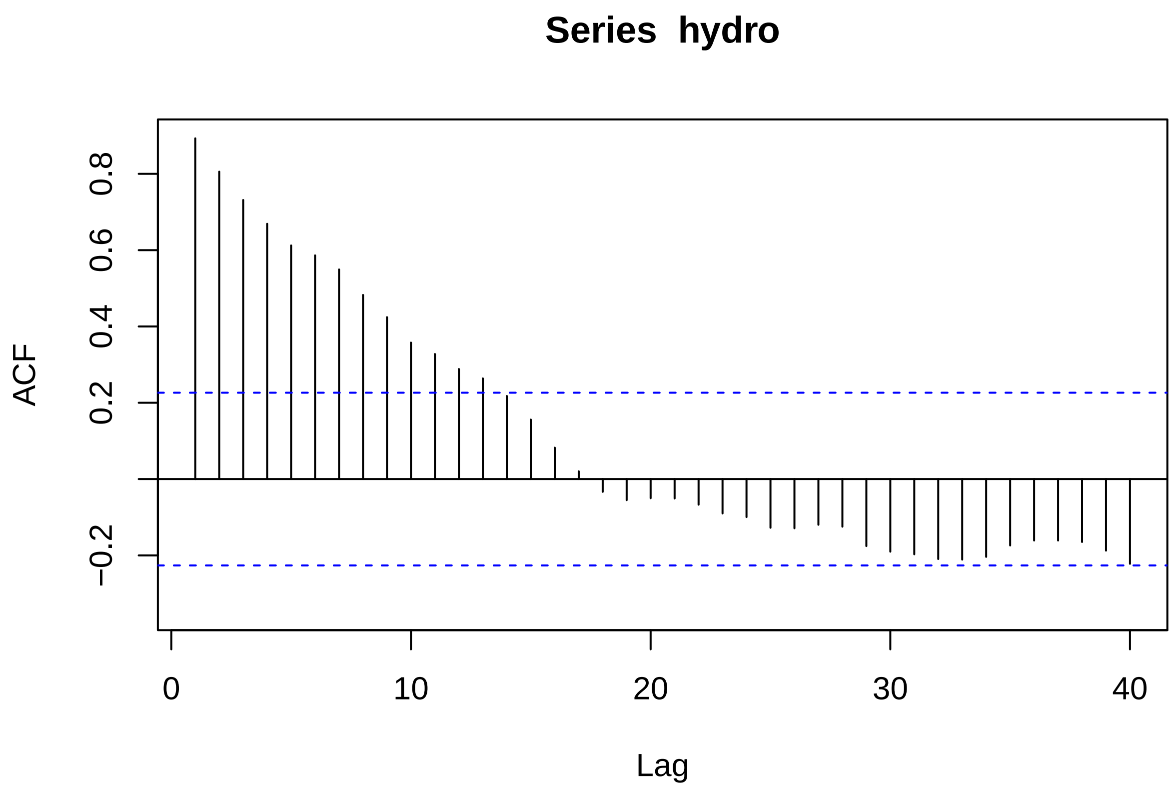
```r
second_acf <- Acf(renewable, lag.max = 40, type = 'correlation', plot = TRUE)
```




```r
second_acf
```

```
##
## Autocorrelations of series 'renewable', by lag
##
##      0      1      2      3      4      5      6      7      8      9     10
##  1.000  0.947  0.893  0.844  0.796  0.740  0.681  0.623  0.569  0.517  0.464
##     11     12     13     14     15     16     17     18     19     20     21
##  0.417  0.372  0.323  0.279  0.243  0.210  0.185  0.163  0.146  0.130  0.118
##     22     23     24     25     26     27     28     29     30     31     32
##  0.109  0.104  0.091  0.076  0.061  0.040  0.017 -0.005 -0.022 -0.039 -0.053
##     33     34     35     36     37     38     39     40
## -0.067 -0.081 -0.100 -0.114 -0.128 -0.147 -0.170 -0.196
```

```r
third_acf <- Acf(hydro, lag.max = 40, type = 'correlation', plot = TRUE)
```




```
third_acf
```

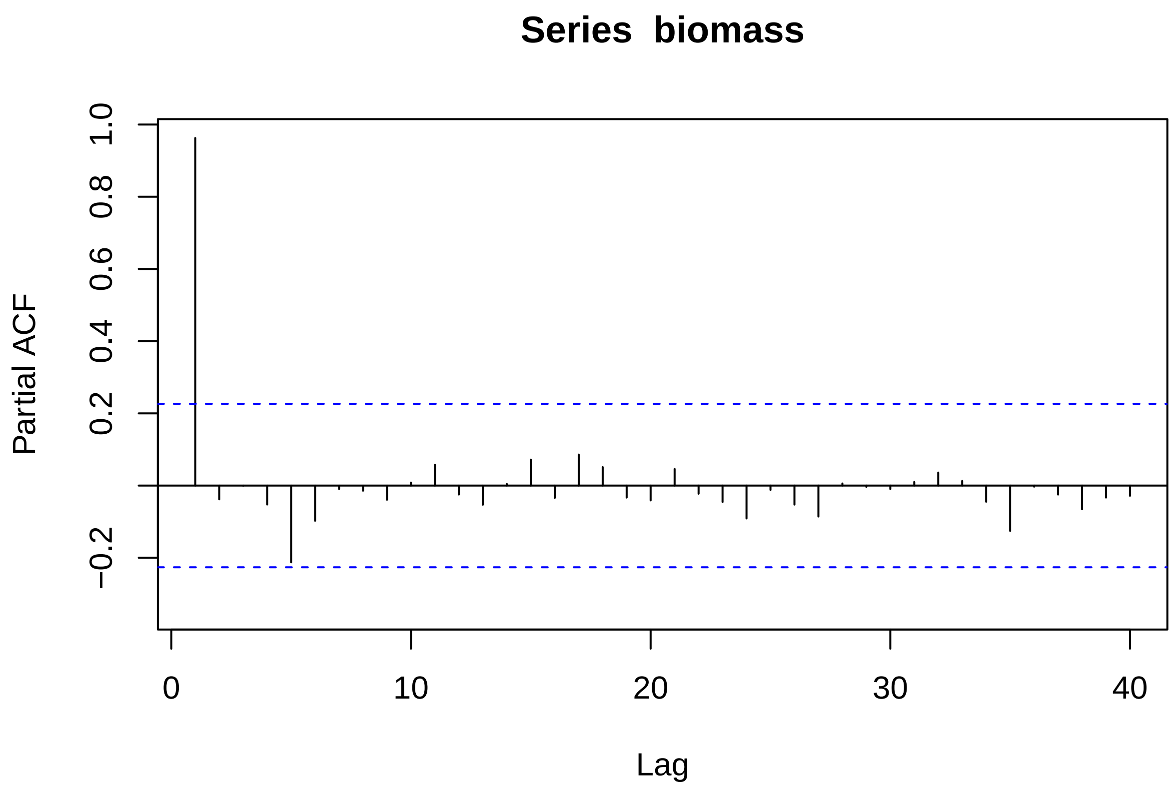
```
##
## Autocorrelations of series 'hydro', by lag
##
##      0      1      2      3      4      5      6      7      8      9     10
##  1.000  0.893  0.806  0.731  0.669  0.612  0.586  0.549  0.483  0.424  0.358
##     11     12     13     14     15     16     17     18     19     20     21
##  0.328  0.288  0.264  0.218  0.156  0.082  0.020 -0.033 -0.055 -0.050 -0.051
##     22     23     24     25     26     27     28     29     30     31     32
## -0.067 -0.090 -0.100 -0.128 -0.129 -0.120 -0.125 -0.176 -0.190 -0.197 -0.210
##     33     34     35     36     37     38     39     40
## -0.211 -0.204 -0.174 -0.161 -0.161 -0.165 -0.187 -0.222
```

The behavior of total biomass energy production is similar to that of total renewable energy production. For both total biomass energy production and total renewable energy production, the coefficient becomes negative with lag 29. Hydroelectric power consumption behave differently because its coefficient becomes negative with lag 18. This faster decline may indicate that the temporal dependence is weaker for hydroelectric power consumption.

Question 7

Compute the partial autocorrelation function from lag 1 to lag 40 for these three variables. How these plots differ from the ones in Q6?

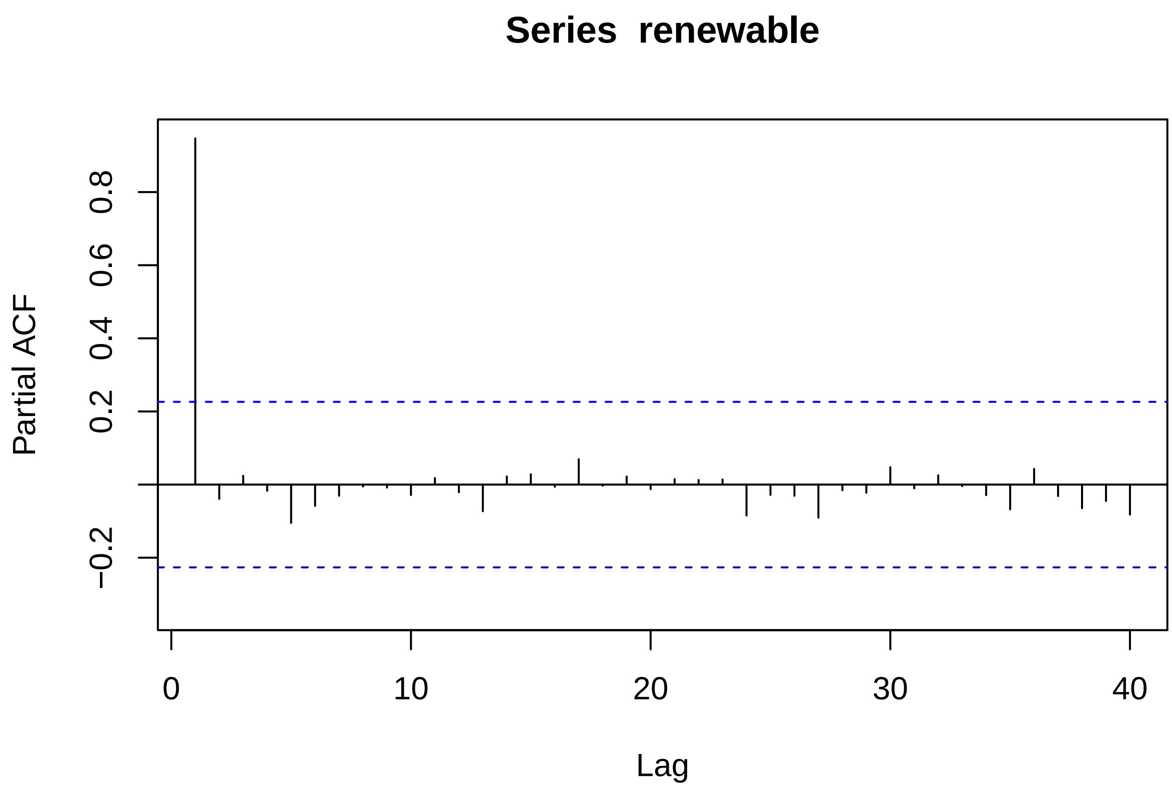
```r
first_pacf <- Pacf(biomass, lag.max = 40, plot = TRUE)
```




```r
first_pacf
```

```
##
## Partial autocorrelations of series 'biomass', by lag
##
##      1      2      3      4      5      6      7      8      9     10     11
##  0.963 -0.038  0.000 -0.053 -0.213 -0.097 -0.009 -0.014 -0.039  0.008  0.057
##     12     13     14     15     16     17     18     19     20     21     22
## -0.025 -0.053  0.004  0.072 -0.034  0.086  0.051 -0.033 -0.041  0.046 -0.023
##     23     24     25     26     27     28     29     30     31     32     33
## -0.046 -0.091 -0.012 -0.053 -0.086  0.006 -0.004 -0.010  0.010  0.036  0.013
##     34     35     36     37     38     39     40
## -0.045 -0.126 -0.003 -0.025 -0.066 -0.033 -0.028
```

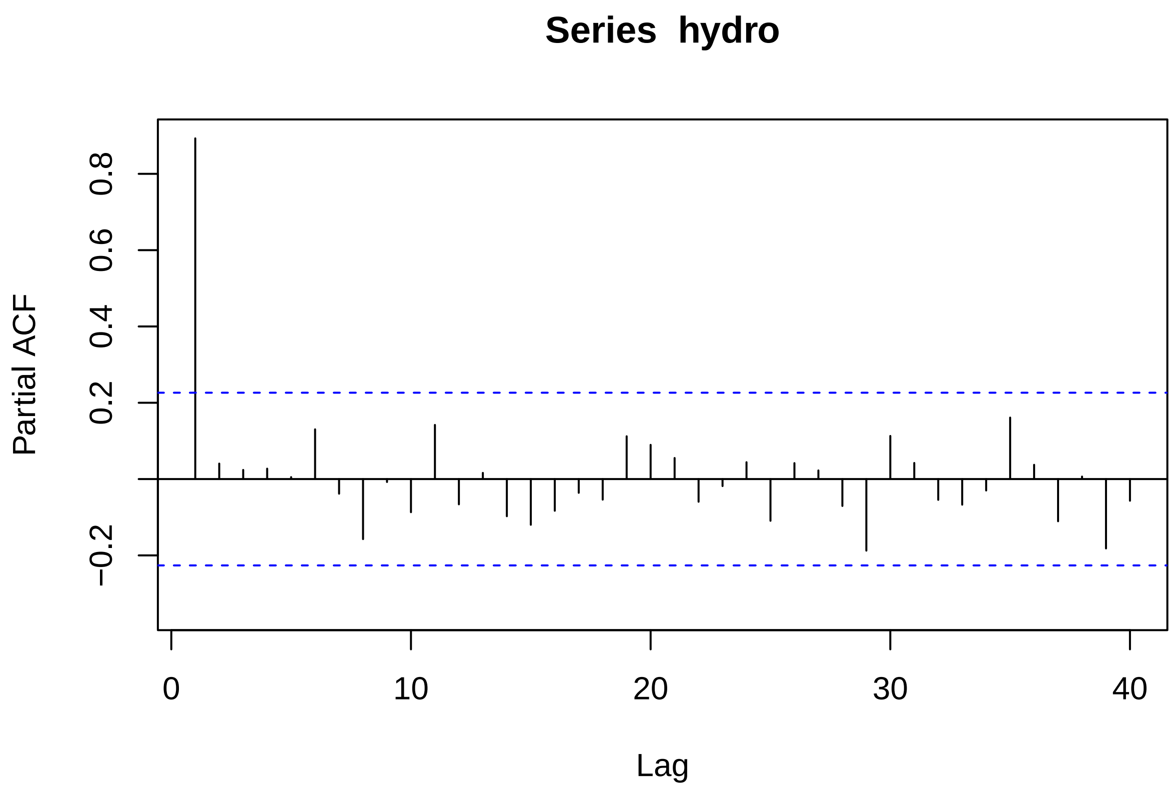
```r
second_pacf <- Pacf(renewable, lag.max = 40, plot = TRUE)
```




```
second_pacf
```

```
##
## Partial autocorrelations of series 'renewable', by lag
##
##      1      2      3      4      5      6      7      8      9     10     11
##  0.947 -0.039  0.024 -0.017 -0.105 -0.058 -0.031 -0.005 -0.008 -0.029  0.018
##     12     13     14     15     16     17     18     19     20     21     22
## -0.021 -0.073  0.022  0.028 -0.006  0.069 -0.003  0.022 -0.013  0.015  0.013
##     23     24     25     26     27     28     29     30     31     32     33
##  0.014 -0.085 -0.028 -0.031 -0.091 -0.016 -0.022  0.048 -0.010  0.025 -0.004
##     34     35     36     37     38     39     40
## -0.029 -0.068  0.043 -0.031 -0.065 -0.045 -0.082
```

```
third_pacf <- Pacf(hydro, lag.max = 40, plot = TRUE)
```




```
third_pacf
```

```
##
## Partial autocorrelations of series 'hydro', by lag
##
##      1      2      3      4      5      6      7      8      9     10     11
##  0.893  0.041  0.024  0.027  0.005  0.130 -0.038 -0.157 -0.008 -0.087  0.142
##     12     13     14     15     16     17     18     19     20     21     22
## -0.066  0.016 -0.097 -0.120 -0.083 -0.036 -0.054  0.112  0.090  0.055 -0.059
##     23     24     25     26     27     28     29     30     31     32     33
## -0.018  0.044 -0.109  0.042  0.023 -0.070 -0.188  0.113  0.042 -0.055 -0.067
##     34     35     36     37     38     39     40
## -0.030  0.161  0.037 -0.111  0.007 -0.182 -0.057
```

Coefficients no longer follow a declining trend. PACF eliminates the influence of intermediate variables.