ENV 790.30 - Time Series Analysis for Energy Data | Spring 2022

Assignment 3 - Due date 02/08/22

Yu Hai

## Directions

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 project open the first thing you will do is change “Student Name” on line 3 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).

Please keep this R code chunk options for the report. It is easier for us to grade when we can see code and output together. And the tidy.opts will make sure that line breaks on your code chunks are automatically added for better visualization.

When you have completed the assignment, **Knit** the text and code into a single PDF file. Rename the pdf file such that it includes your first and last name (e.g., “LuanaLima\_TSA\_A03\_Sp22.Rmd”). Submit this pdf using Sakai.

## Questions

Consider the same data you used for A2 from the spreadsheet “Table\_10.1\_Renewable\_Energy\_Production\_and\_Consumption\_by\_Source.xlsx”. The data comes from the US Energy Information and Administration and corresponds to the January 2022 **Monthly** Energy Review. Once again 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.

R packages needed for this assignment:“forecast”,“tseries”, and “Kendall”. 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.\

#Load/install required package here  
library(forecast)

## Warning: 程辑包'forecast'是用R版本4.1.2 来建造的

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

library(tseries)

## Warning: 程辑包'tseries'是用R版本4.1.2 来建造的

library(Kendall)

## Warning: 程辑包'Kendall'是用R版本4.1.2 来建造的

library(lubridate)

##   
## 载入程辑包：'lubridate'

## The following objects are masked from 'package:base':  
##   
## date, intersect, setdiff, union

library(ggplot2)

library(readxl)

## Warning: 程辑包'readxl'是用R版本4.1.2 来建造的

raw\_RE\_data<-read\_excel(path="./Data/Table\_10.1\_Renewable\_Energy\_Production\_and\_Consumption\_by\_Source.xlsx",skip=10)  
raw\_RE\_data

## # A tibble: 586 x 14  
## Month `Wood Energy Prod~ `Biofuels Product~ `Total Biomass Ene~  
## <dttm> <chr> <chr> <chr>   
## 1 NA (Trillion Btu) (Trillion Btu) (Trillion Btu)   
## 2 1973-01-01 00:00:00 129.63 Not Available 129.787   
## 3 1973-02-01 00:00:00 117.194 Not Available 117.338   
## 4 1973-03-01 00:00:00 129.763 Not Available 129.938   
## 5 1973-04-01 00:00:00 125.462 Not Available 125.636   
## 6 1973-05-01 00:00:00 129.624 Not Available 129.834   
## 7 1973-06-01 00:00:00 125.435 Not Available 125.611   
## 8 1973-07-01 00:00:00 129.616 Not Available 129.787   
## 9 1973-08-01 00:00:00 129.734 Not Available 129.918   
## 10 1973-09-01 00:00:00 125.603 Not Available 125.782   
## # ... with 576 more rows, and 10 more variables:  
## # Total Renewable Energy Production <chr>,  
## # Hydroelectric Power Consumption <chr>, Geothermal Energy Consumption <chr>,  
## # Solar Energy Consumption <chr>, Wind Energy Consumption <chr>,  
## # Wood Energy Consumption <chr>, Waste Energy Consumption <chr>,  
## # Biofuels Consumption <chr>, Total Biomass Energy Consumption <chr>,  
## # Total Renewable Energy Consumption <chr>

##Trend Component

### Q1

Create a plot window that has one row and three columns. And then for each object on your data frame, fill the plot window with time series plot, ACF and PACF. You may use the some code form A2, but I want all three plots on the same window this time. (Hint: use par() function)

sub\_RE\_data <- raw\_RE\_data[-c(1),4:6]  
head(sub\_RE\_data)

## # A tibble: 6 x 3  
## `Total Biomass Energy Production` `Total Renewable Ener~ `Hydroelectric Power~  
## <chr> <chr> <chr>   
## 1 129.787 403.981 272.703   
## 2 117.338 360.9 242.199   
## 3 129.938 400.161 268.81   
## 4 125.636 380.47 253.185   
## 5 129.834 392.141 260.77   
## 6 125.611 377.232 249.859

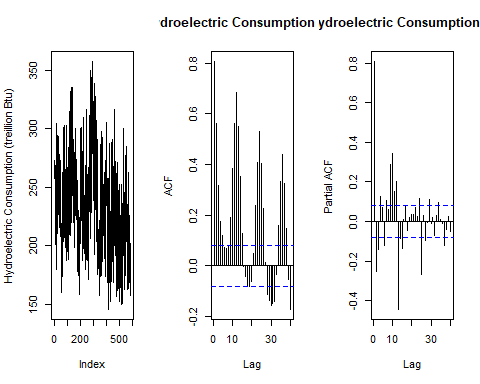
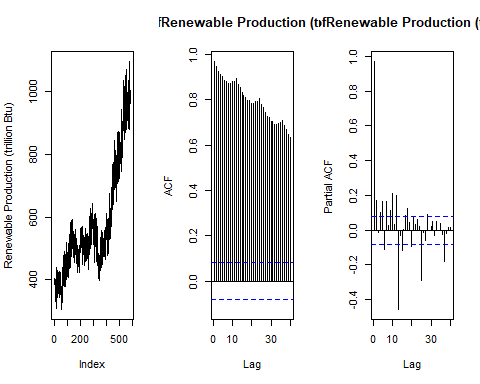
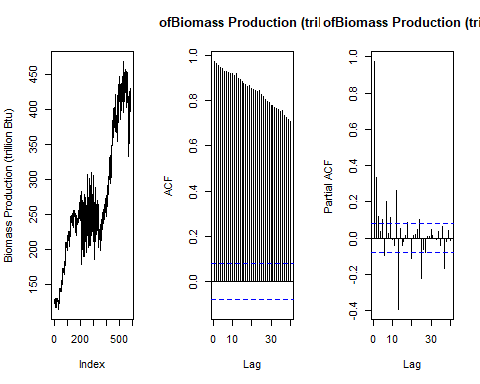
RE\_data <- cbind(raw\_RE\_data[-c(1),1],sub\_RE\_data[,])  
RE\_data$`Total Biomass Energy Production`<- as.numeric(RE\_data$`Total Biomass Energy Production`)  
RE\_data$`Total Renewable Energy Production`<- as.numeric(RE\_data$`Total Renewable Energy Production`)  
RE\_data$`Hydroelectric Power Consumption`<- as.numeric(RE\_data$`Hydroelectric Power Consumption`)  
head(RE\_data)

## Month Total Biomass Energy Production Total Renewable Energy Production  
## 1 1973-01-01 129.787 403.981  
## 2 1973-02-01 117.338 360.900  
## 3 1973-03-01 129.938 400.161  
## 4 1973-04-01 125.636 380.470  
## 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  
## 3 268.810  
## 4 253.185  
## 5 260.770  
## 6 249.859

#nhydro <- ncol(raw\_inflow\_data)-2  
#nobs <- nrow(raw\_inflow\_data)

colnames<-c("Biomass Production (trillion Btu)","Renewable Production (trillion Btu)","Hydroelectric Consumption (treillion Btu)")

#How to put column names as the plot titles?  
par(mfrow=c(1,3))  
for(i in 2:4){  
 plot(RE\_data[,i],type="l",ylab=colnames[i-1])  
 Acf(RE\_data[,i],lag.max=40,main=paste("ACF of",colnames[i-1],sep=""))   
 Pacf(RE\_data[,i],lag.max=40,main=paste("PACF of",colnames[i-1],sep=""))   
}



### Q2

From the plot in Q1, do the series Total Biomass Energy Production, Total Renewable Energy Production, Hydroelectric Power Consumption appear to have a trend? If yes, what kind of trend?

Biomass and renewable production show a linear, increasing trend, while the hydroelectric power consumption shows a linear, decreasing trend. All three series seems to have seasonality as there is consistent fluctuation pattern between observations within each year.

### Q3

Use the *lm()* function to fit a linear trend to the three time series. Ask R to print the summary of the regression. Interpret the regression output, i.e., slope and intercept. Save the regression coefficients for further analysis.

ts\_RE\_data <- ts(RE\_data[,2:4], start=c(1973, 1), end=c(2021, 09), frequency=12)  
head(ts\_RE\_data)

## Total Biomass Energy Production Total Renewable Energy Production  
## Jan 1973 129.787 403.981  
## Feb 1973 117.338 360.900  
## Mar 1973 129.938 400.161  
## Apr 1973 125.636 380.470  
## May 1973 129.834 392.141  
## Jun 1973 125.611 377.232  
## Hydroelectric Power Consumption  
## Jan 1973 272.703  
## Feb 1973 242.199  
## Mar 1973 268.810  
## Apr 1973 253.185  
## May 1973 260.770  
## Jun 1973 249.859

#nhydro <- ncol(RE\_data)-2  
nobs <- nrow(RE\_data)   
t <- c(1:nobs)  
lm1=lm(RE\_data[,2]~t)   
lm2=lm(RE\_data[,3]~t)   
lm3=lm(RE\_data[,4]~t)   
  
print("Results of Biomass Time Series")

## [1] "Results of Biomass Time Series"

print(summary(lm1))

##   
## Call:  
## lm(formula = RE\_data[, 2] ~ t)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -101.892 -24.306 4.932 33.103 82.292   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.348e+02 3.282e+00 41.07 <2e-16 \*\*\*  
## t 4.744e-01 9.705e-03 48.88 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 39.64 on 583 degrees of freedom  
## Multiple R-squared: 0.8039, Adjusted R-squared: 0.8035   
## F-statistic: 2389 on 1 and 583 DF, p-value: < 2.2e-16

For total biomass energy production, the p-value of the linear regression is less than 0.05, so there is a significant trend in biomass energy production overtime. The value of the intercept indicates that the initial biomass energy production at time t=0 (January 1973) is 134.8 trillion Btu, and after each month, it is expected that the total biomass production will increase by 0.4744 trillion Btu.

print("Results of Renewable Time Series")

## [1] "Results of Renewable Time Series"

print(summary(lm2))

##   
## Call:  
## lm(formula = RE\_data[, 3] ~ t)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -230.488 -57.869 5.595 62.090 261.349   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 323.18243 8.02555 40.27 <2e-16 \*\*\*  
## t 0.88051 0.02373 37.10 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 96.93 on 583 degrees of freedom  
## Multiple R-squared: 0.7025, Adjusted R-squared: 0.702   
## F-statistic: 1377 on 1 and 583 DF, p-value: < 2.2e-16

For total renewable energy production, the p-value of the linear regression is less than 0.05, so there is a significant trend in renewable energy production overtime. The value of the intercept indicates the initial renewable energy production at time t=0 (January 1973), that the biomass production is 323 trillion Btu, and after each month, it is expected that the total renewable energy production will increase by 0.88 trillion Btu.

print("Results of Hydroelectric Time Series")

## [1] "Results of Hydroelectric Time Series"

print(summary(lm3))

##   
## Call:  
## lm(formula = RE\_data[, 4] ~ t)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -94.892 -31.300 -2.414 27.876 121.263   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 259.18303 3.47464 74.593 < 2e-16 \*\*\*  
## t -0.07924 0.01027 -7.712 5.36e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 41.97 on 583 degrees of freedom  
## Multiple R-squared: 0.09258, Adjusted R-squared: 0.09103   
## F-statistic: 59.48 on 1 and 583 DF, p-value: 5.364e-14

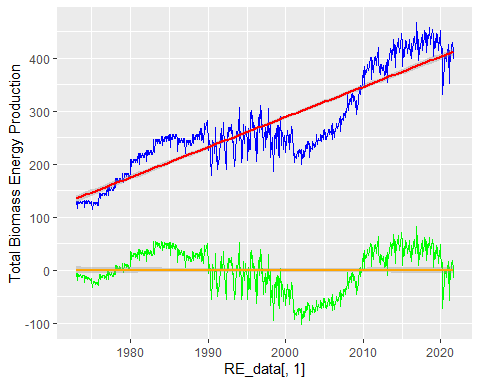
For total hydroelectric power consumption, the p-value of the linear regression is less than 0.05, so there is a significant trend in the consumption overtime. The value of the intercept indicates the initial consumption at time t=0 (January 1973) is 259.18 trillion Btu, and after each month, it is expected that the consumption will decrease by 0.079 trillion Btu.

### Q4

Use the regression coefficients from Q3 to detrend the series. Plot the detrended series and compare with the plots from Q1. What happened? Did anything change?

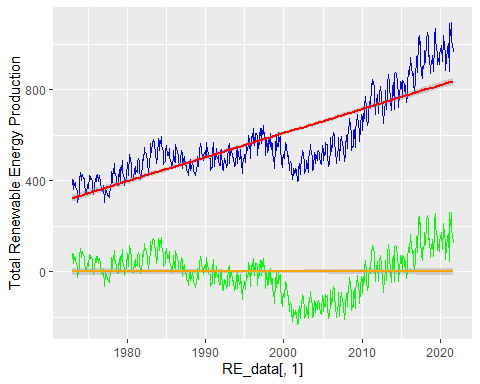
#Biomass  
beta0\_bio=as.numeric(lm1$coefficients[1])   
beta1\_bio=as.numeric(lm1$coefficients[2])  
  
detrend\_bio <- RE\_data[,2]-(beta0\_bio+beta1\_bio\*t)  
 ggplot(RE\_data, aes(x=RE\_data[,1], y=RE\_data[,2])) +  
 geom\_line(color="blue") +  
 ylab("Total Biomass Energy Production") +  
 geom\_smooth(color="red",method="lm") +  
 geom\_line(aes(y=detrend\_bio), col="green")+  
 geom\_smooth(aes(y=detrend\_bio),color="orange",method="lm")

## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'



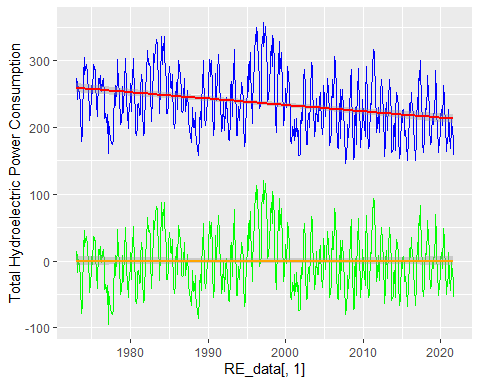
#renewable  
beta0\_renew=as.numeric(lm2$coefficients[1])   
beta1\_renew=as.numeric(lm2$coefficients[2])  
  
detrend\_renew <- RE\_data[,3]-(beta0\_renew+beta1\_renew\*t)  
 ggplot(RE\_data, aes(x=RE\_data[,1], y=RE\_data[,3])) +  
 geom\_line(color="blue") +  
 ylab("Total Renewable Energy Production") +  
 geom\_smooth(color="red",method="lm") +  
 geom\_line(aes(y=detrend\_renew), col="green")+  
 geom\_smooth(aes(y=detrend\_renew),color="orange",method="lm")

## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'



#hydro  
beta0\_hydro=as.numeric(lm3$coefficients[1])   
beta1\_hydro=as.numeric(lm3$coefficients[2])  
  
detrend\_hydro <- RE\_data[,4]-(beta0\_hydro+beta1\_hydro\*t)  
 ggplot(RE\_data, aes(x=RE\_data[,1], y=RE\_data[,4])) +  
 geom\_line(color="blue") +  
 ylab("Total Hydroelectric Power Consumption") +  
 geom\_smooth(color="red",method="lm") +  
 geom\_line(aes(y=detrend\_hydro), col="green")+  
 geom\_smooth(aes(y=detrend\_hydro),color="orange",method="lm")

## `geom\_smooth()` using formula 'y ~ x'  
## `geom\_smooth()` using formula 'y ~ x'

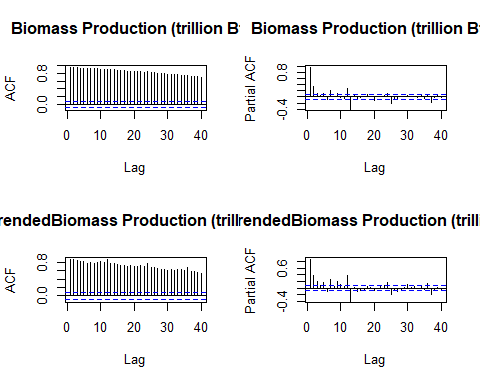


The linear trend after detrending become horizontal, so both beta 1 and beta 0 becomes 0. The detrending process removed the effects of trend from the original data but keep other components (e.g. seasonality and clynicality) the same, so the detrending process allows those components to be identified.

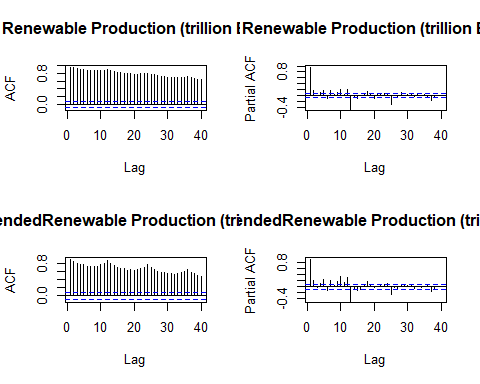
### Q5

Plot ACF and PACF for the detrended series and compare with the plots from Q1. Did the plots change? How?

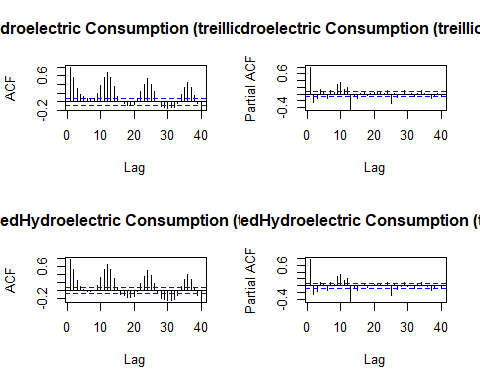
#biomass  
par(mfrow=c(2,2))  
 Acf(RE\_data[,2],lag.max=40,main=colnames[1])   
 Pacf(RE\_data[,2],lag.max=40,main=colnames[1])   
 Acf(detrend\_bio,lag.max=40,main=paste("Detrended",colnames[1],sep="") )  
 Pacf(detrend\_bio,lag.max=40,main=paste("Detrended",colnames[1],sep="") )



#renewable  
par(mfrow=c(2,2))  
 Acf(RE\_data[,3],lag.max=40,main=colnames[2])   
 Pacf(RE\_data[,3],lag.max=40,main=colnames[2])   
 Acf(detrend\_renew,lag.max=40,main=paste("Detrended",colnames[2],sep="") )  
 Pacf(detrend\_renew,lag.max=40,main=paste("Detrended",colnames[2],sep="") )



#hydro  
par(mfrow=c(2,2))  
 Acf(RE\_data[,4],lag.max=40,main=colnames[3])   
 Pacf(RE\_data[,4],lag.max=40,main=colnames[3])   
 Acf(detrend\_hydro,lag.max=40,main=paste("Detrended",colnames[3],sep="") )  
 Pacf(detrend\_hydro,lag.max=40,main=paste("Detrended",colnames[3],sep="") )

 In ACF plots of all three variables, there are greater fluctuation as ACF is decaying (i.e. the seasonality becomes more obvious), the changes in PACF plots are relatively small.

## Seasonal Component

Set aside the detrended series and consider the original series again from Q1 to answer Q6 to Q8.

### Q6

Do the series seem to have a seasonal trend? Which serie/series? Use function *lm()* to fit a seasonal means model (i.e. using the seasonal dummies) to this/these time series. Ask R to print the summary of the regression. Interpret the regression output. Save the regression coefficients for further analysis.

#Biomass  
dummies\_bio <- seasonaldummy(ts\_RE\_data[,1])   
seas\_means\_model\_bio=lm(RE\_data[,(2)]~dummies\_bio)  
summary(seas\_means\_model\_bio)

##   
## Call:  
## lm(formula = RE\_data[, (2)] ~ dummies\_bio)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -156.96 -51.40 -22.15 60.65 183.31   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 284.241 12.962 21.928 <2e-16 \*\*\*  
## dummies\_bioJan -1.498 18.238 -0.082 0.9346   
## dummies\_bioFeb -30.582 18.238 -1.677 0.0941 .   
## dummies\_bioMar -8.873 18.238 -0.486 0.6268   
## dummies\_bioApr -21.009 18.238 -1.152 0.2498   
## dummies\_bioMay -14.065 18.238 -0.771 0.4409   
## dummies\_bioJun -19.601 18.238 -1.075 0.2829   
## dummies\_bioJul -3.499 18.238 -0.192 0.8479   
## dummies\_bioAug -0.252 18.238 -0.014 0.9890   
## dummies\_bioSep -12.518 18.238 -0.686 0.4928   
## dummies\_bioOct -3.629 18.331 -0.198 0.8432   
## dummies\_bioNov -9.592 18.331 -0.523 0.6010   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 89.81 on 573 degrees of freedom  
## Multiple R-squared: 0.01056, Adjusted R-squared: -0.008439   
## F-statistic: 0.5557 on 11 and 573 DF, p-value: 0.8647

beta\_int\_bio=seas\_means\_model\_bio$coefficients[1]  
beta\_coeff\_bio=seas\_means\_model\_bio$coefficients[2:12]

#Renewable  
dummies\_renew <- seasonaldummy(ts\_RE\_data[,2])   
seas\_means\_model\_renew=lm(RE\_data[,(3)]~dummies\_renew)  
summary(seas\_means\_model\_renew)

##   
## Call:  
## lm(formula = RE\_data[, (3)] ~ dummies\_renew)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -272.95 -111.55 -59.35 65.68 480.41   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 589.971 25.464 23.169 <2e-16 \*\*\*  
## dummies\_renewJan 11.793 35.828 0.329 0.7422   
## dummies\_renewFeb -40.992 35.828 -1.144 0.2530   
## dummies\_renewMar 21.892 35.828 0.611 0.5414   
## dummies\_renewApr 8.908 35.828 0.249 0.8037   
## dummies\_renewMay 37.500 35.828 1.047 0.2957   
## dummies\_renewJun 19.465 35.828 0.543 0.5871   
## dummies\_renewJul 8.115 35.828 0.227 0.8209   
## dummies\_renewAug -18.359 35.828 -0.512 0.6086   
## dummies\_renewSep -62.115 35.828 -1.734 0.0835 .   
## dummies\_renewOct -51.377 36.012 -1.427 0.1542   
## dummies\_renewNov -41.789 36.012 -1.160 0.2464   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 176.4 on 573 degrees of freedom  
## Multiple R-squared: 0.03139, Adjusted R-squared: 0.0128   
## F-statistic: 1.688 on 11 and 573 DF, p-value: 0.07235

beta\_int\_renew=seas\_means\_model\_renew$coefficients[1]  
beta\_coeff\_renew=seas\_means\_model\_renew$coefficients[2:12]

#Hydroelectric  
dummies\_hydro <- seasonaldummy(ts\_RE\_data[,3])   
seas\_means\_model\_hydro=lm(RE\_data[,(4)]~dummies\_hydro)  
summary(seas\_means\_model\_hydro)

##   
## Call:  
## lm(formula = RE\_data[, (4)] ~ dummies\_hydro)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -90.253 -23.017 -3.042 21.487 99.478   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 237.841 4.892 48.616 < 2e-16 \*\*\*  
## dummies\_hydroJan 13.558 6.883 1.970 0.04936 \*   
## dummies\_hydroFeb -8.090 6.883 -1.175 0.24037   
## dummies\_hydroMar 20.067 6.883 2.915 0.00369 \*\*   
## dummies\_hydroApr 16.619 6.883 2.414 0.01607 \*   
## dummies\_hydroMay 39.961 6.883 5.805 1.06e-08 \*\*\*  
## dummies\_hydroJun 31.315 6.883 4.549 6.57e-06 \*\*\*  
## dummies\_hydroJul 10.511 6.883 1.527 0.12732   
## dummies\_hydroAug -17.853 6.883 -2.594 0.00974 \*\*   
## dummies\_hydroSep -49.852 6.883 -7.242 1.43e-12 \*\*\*  
## dummies\_hydroOct -48.086 6.919 -6.950 9.96e-12 \*\*\*  
## dummies\_hydroNov -32.187 6.919 -4.652 4.08e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 33.89 on 573 degrees of freedom  
## Multiple R-squared: 0.4182, Adjusted R-squared: 0.4071   
## F-statistic: 37.45 on 11 and 573 DF, p-value: < 2.2e-16

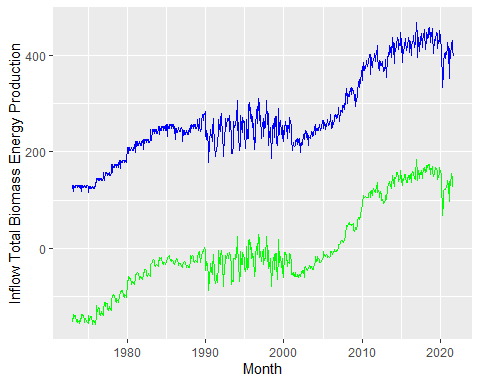
beta\_int\_hydro=seas\_means\_model\_hydro$coefficients[1]  
beta\_coeff\_hydro=seas\_means\_model\_hydro$coefficients[2:12]

The seasonal mean models show that there is a significant seasonality for hydroelectric power consumption (p<0.05), while there isn’t a significant seasonality for biomass energy production (p=0.86) or for renewable energy production (p=0.07).

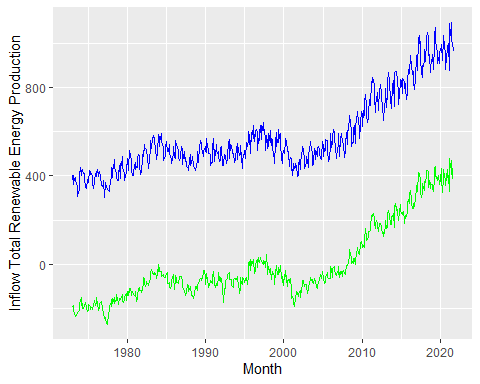
### Q7

Use the regression coefficients from Q6 to deseason the series. Plot the deseason series and compare with the plots from part Q1. Did anything change?

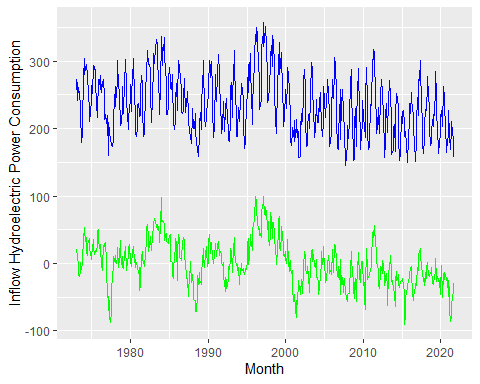
#Biomass deseason  
RE\_seas\_comp\_bio=array(0,nobs)  
for(i in 1:nobs){  
 RE\_seas\_comp\_bio[i]=(beta\_int\_bio+beta\_coeff\_bio%\*%dummies\_bio[i,])  
}  
deseason\_RE\_data\_bio <- RE\_data[,2]-RE\_seas\_comp\_bio  
ggplot(RE\_data, aes(x=Month, y=RE\_data[,2])) +  
 geom\_line(color="blue") +  
 ylab(paste0("Inflow ",colnames(RE\_data)[2],sep="")) +  
 geom\_line(aes(y=deseason\_RE\_data\_bio), col="green")



#Renewable  
RE\_seas\_comp\_renew=array(0,nobs)  
for(i in 1:nobs){  
 RE\_seas\_comp\_renew[i]=(beta\_int\_renew+beta\_coeff\_renew%\*%dummies\_renew[i,])  
}  
deseason\_RE\_data\_renew <- RE\_data[,3]-RE\_seas\_comp\_renew  
  
ggplot(RE\_data, aes(x=Month, y=RE\_data[,3])) +  
 geom\_line(color="blue") +  
 ylab(paste0("Inflow ",colnames(RE\_data)[3],sep="")) +  
 geom\_line(aes(y=deseason\_RE\_data\_renew), col="green")



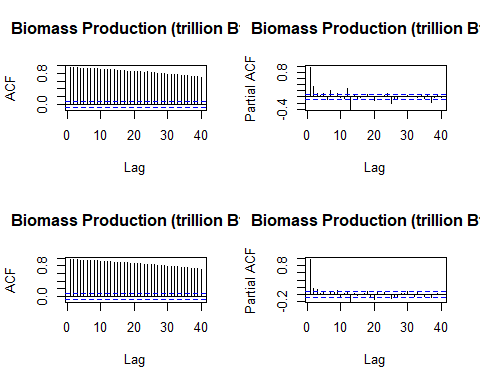
#Hydro  
RE\_seas\_comp\_hydro=array(0,nobs)  
for(i in 1:nobs){  
 RE\_seas\_comp\_hydro[i]=(beta\_int\_hydro+beta\_coeff\_hydro%\*%dummies\_hydro[i,])  
}  
deseason\_RE\_data\_hydro <- RE\_data[,4]-RE\_seas\_comp\_hydro  
ggplot(RE\_data, aes(x=Month, y=RE\_data[,4])) +  
 geom\_line(color="blue") +  
 ylab(paste0("Inflow ",colnames(RE\_data)[4],sep="")) +  
 geom\_line(aes(y=deseason\_RE\_data\_hydro), col="green")

 There are changes in all three variables: there are less variations between observations that within one year in the deseason series. As the variation caused by seasonality is removed, and the long-term trend is more obvious.

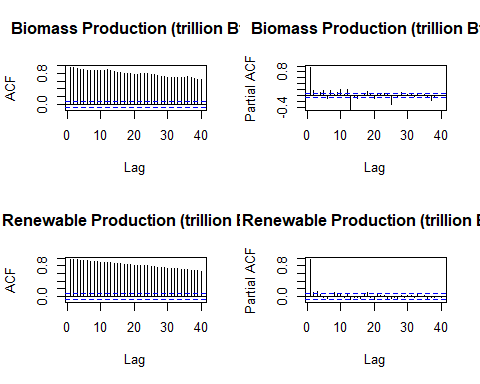
### Q8

Plot ACF and PACF for the deseason series and compare with the plots from Q1. Did the plots change? How?

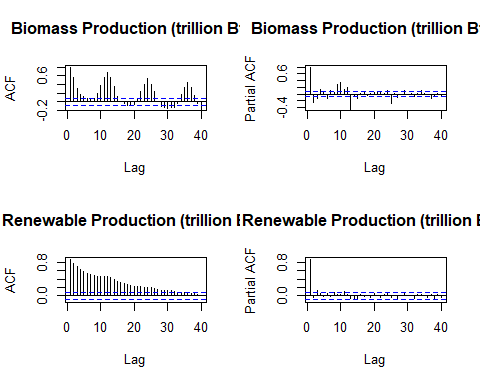
deseason\_RE\_data<-c(deseason\_RE\_data\_bio,deseason\_RE\_data\_renew,deseason\_RE\_data\_hydro)  
#biomass  
par(mfrow=c(2,2))  
 Acf(RE\_data[,2],lag.max=40,main=colnames[1])   
 Pacf(RE\_data[,2],lag.max=40,main=colnames[1])  
 Acf(deseason\_RE\_data\_bio,lag.max=40,main=colnames[1])   
 Pacf(deseason\_RE\_data\_bio,lag.max=40,main=colnames[1])



#renewable  
par(mfrow=c(2,2))  
 Acf(RE\_data[,3],lag.max=40,main=colnames[1])   
 Pacf(RE\_data[,3],lag.max=40,main=colnames[1])  
 Acf(deseason\_RE\_data\_renew,lag.max=40,main=colnames[2])   
 Pacf(deseason\_RE\_data\_renew,lag.max=40,main=colnames[2])



#hydroelectric  
par(mfrow=c(2,2))  
 Acf(RE\_data[,4],lag.max=40,main=colnames[1])   
 Pacf(RE\_data[,4],lag.max=40,main=colnames[1])  
 Acf(deseason\_RE\_data\_hydro,lag.max=40,main=colnames[2])   
 Pacf(deseason\_RE\_data\_hydro,lag.max=40,main=colnames[2])

 For biomass and renewable energy production, the seasonal mean models indicate that there is no significant seasonality, but from deseason ACF plots, we can still observe less fluctuation as ACF is decaying. For hydroelectric power consumption ACF plot, all ACF becomes positive, the ACF decays faster, and the seasonal pattern disappears. In PACF for all three variables, the significant PACF with the original data between the observations at lag=0 and at lag=12, 24, and 36 become weaker, especially for renewable production and hydroelectric consumption, those PACF values recede into the insignificant range (indicated by blue dash lines).