Time Series Analysis 3

Differencing and seasonal ARIMA prediction

Time Series Analysis
Zhe Zheng

```
rm(list=ls())
library(ggplot2)
library(mgcv)
library(FinTS)
options(digits=3)
edvoldata = read.csv("EGDailyVolume.csv",header=T)
```

1. DATA EXPLORATION AND PROCESSING

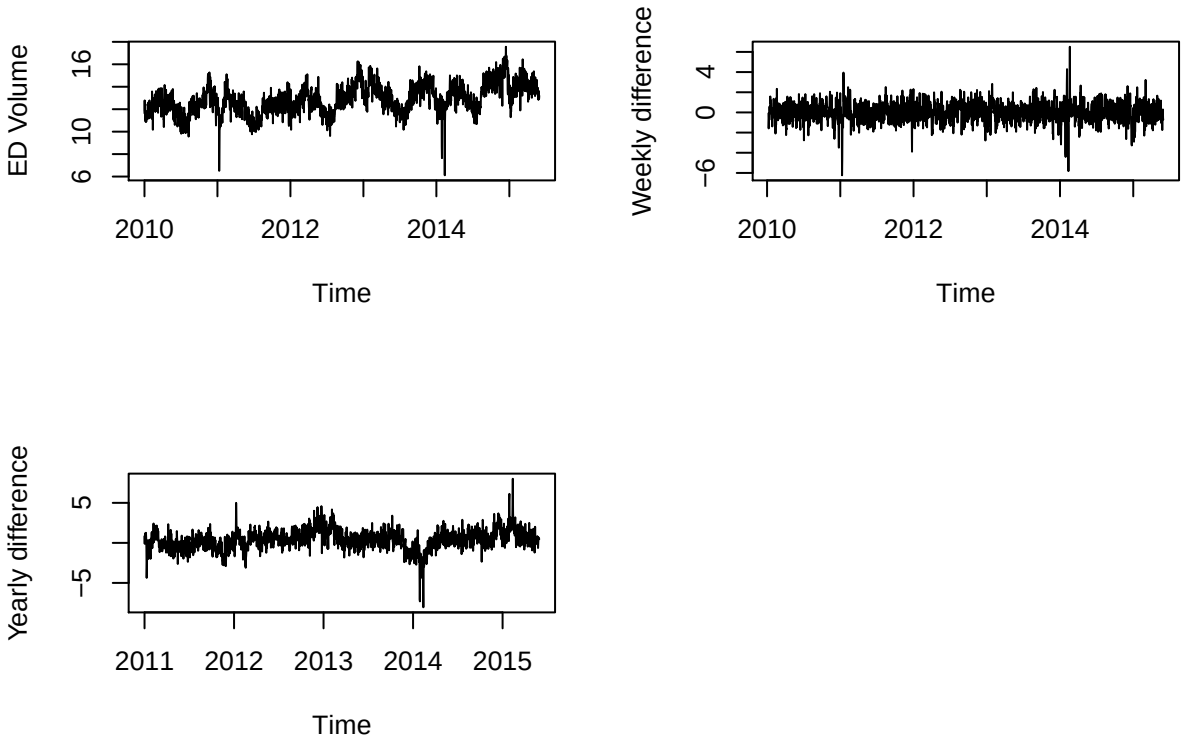
```
## process dates
year = edvoldata$Year
month = edvoldata$Month
day = edvoldata$Day
datemat = cbind(as.character(day),as.character(month),as.character(year))
paste.dates = function(date){
    day = date[1]; month=date[2]; year = date[3]
    return(paste(day,month,year,sep="/"))
}
#apply(x,margin,fun) for a matrix, margin=1 indicates rows, 2 indicates column
dates = apply(datemat,1,paste.dates)
dates = as.Date(dates, format="%d/%m/%Y")
edvoldata = cbind(dates,edvoldata)
attach(edvoldata)
Volume.tr = sqrt(Volume+3/8)
```

get weekly and montly differenced data

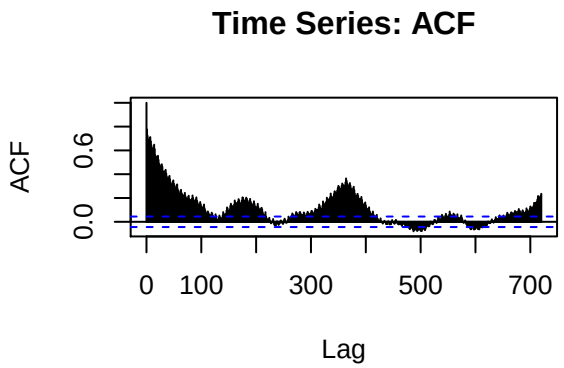
```
volume.ts = ts(Volume.tr,start=c(2010,1,1),frequency=365.25) #daily data
dvolume7=diff(volume.ts,7) # weekly
dvolume12=diff(volume.ts,365) # monthly (annual)
# test1<-data.frame(volume.ts)
# test2<-data.frame(dvolume7)
# test3<-data.frame(dvolume12)
# test_all<-cbind(test1,test2,test3)
```

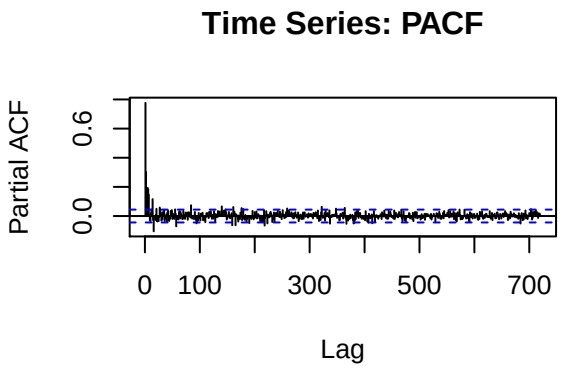
plot time series and difference processes

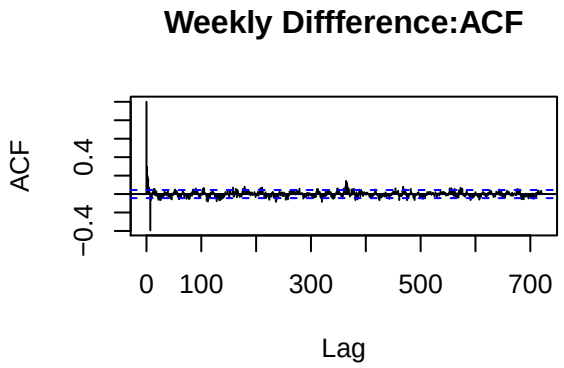
```
par(mfrow=c(2,2))
ts.plot(volume.ts,ylab="ED Volume")
ts.plot(dvolume7,ylab="Weekly difference")
ts.plot(dvolume12,ylab="Yearly difference")

par(mfrow=c(2,2))
```

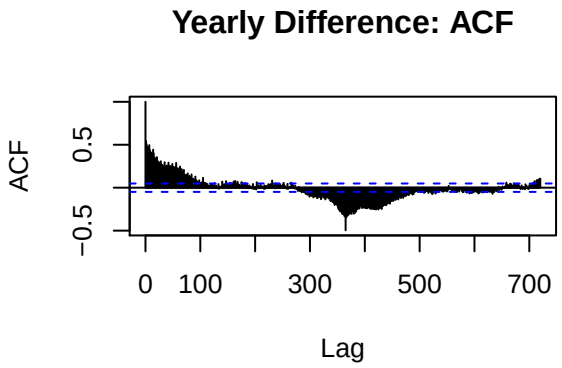



```
acf(as.vector(volume.ts), main='Time Series: ACF',lag.max=360*2)
acf(as.vector(volume.ts),type="partial", main='Time Series: PACF',lag.max=360*2)
acf(as.vector(dvolume7) , main='Weekly Diffference:ACF',lag.max=360*2)
acf(as.vector(dvolume12), main='Yearly Difference: ACF',lag.max=360*2)
```



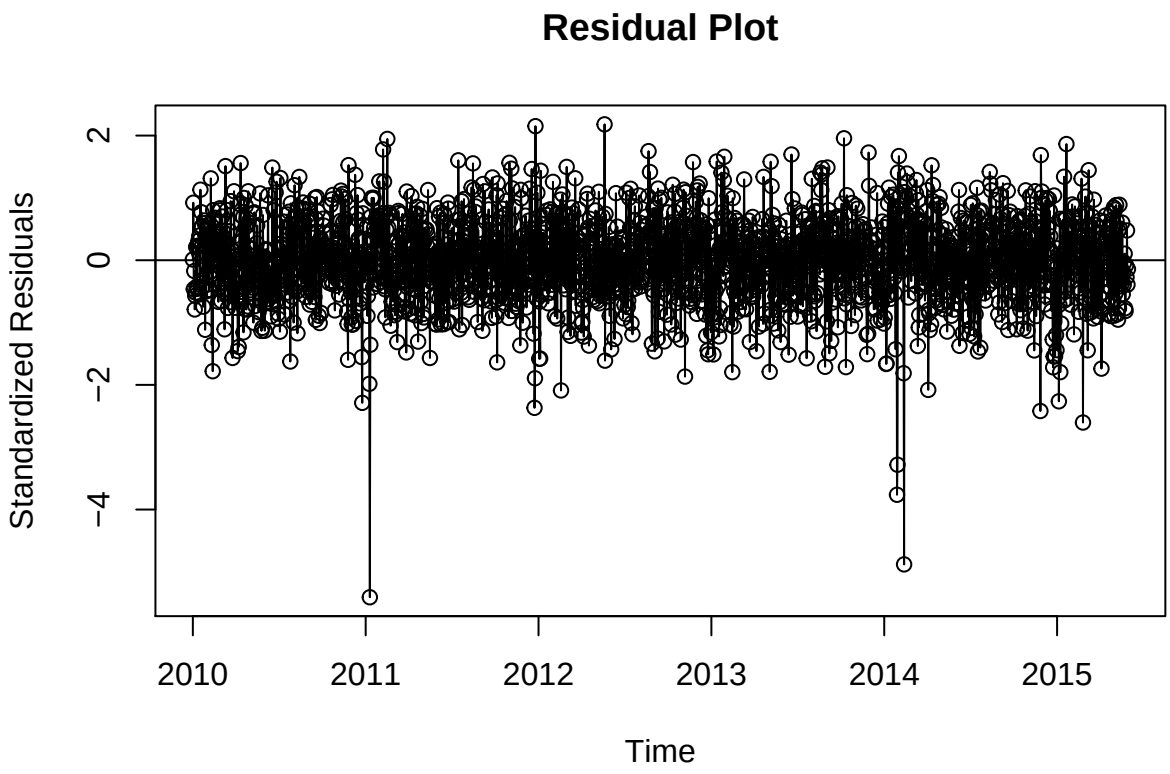




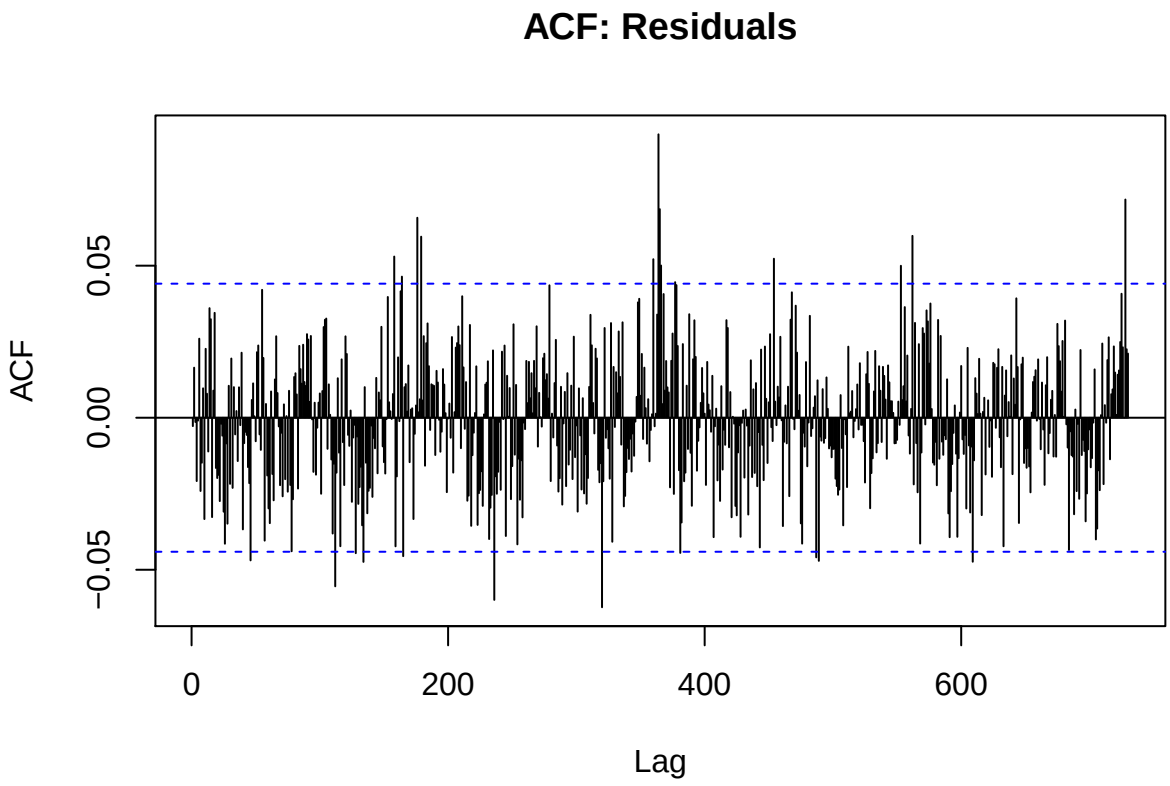


2.Model Fitting ARIMA(5,1,5)+seasonal ARMA(1,1)

```
# the acf results showed weekly difference is the best, so period = 7
mod = arima(volume.ts, order = c(5,1,5),seasonal = list(order = c(1,0,1),period=7),method = "ML")

# residual analysis
plot(resid(mod), ylab='Standardized Residuals',type='o',main="Residual Plot")
abline(h=0)
```

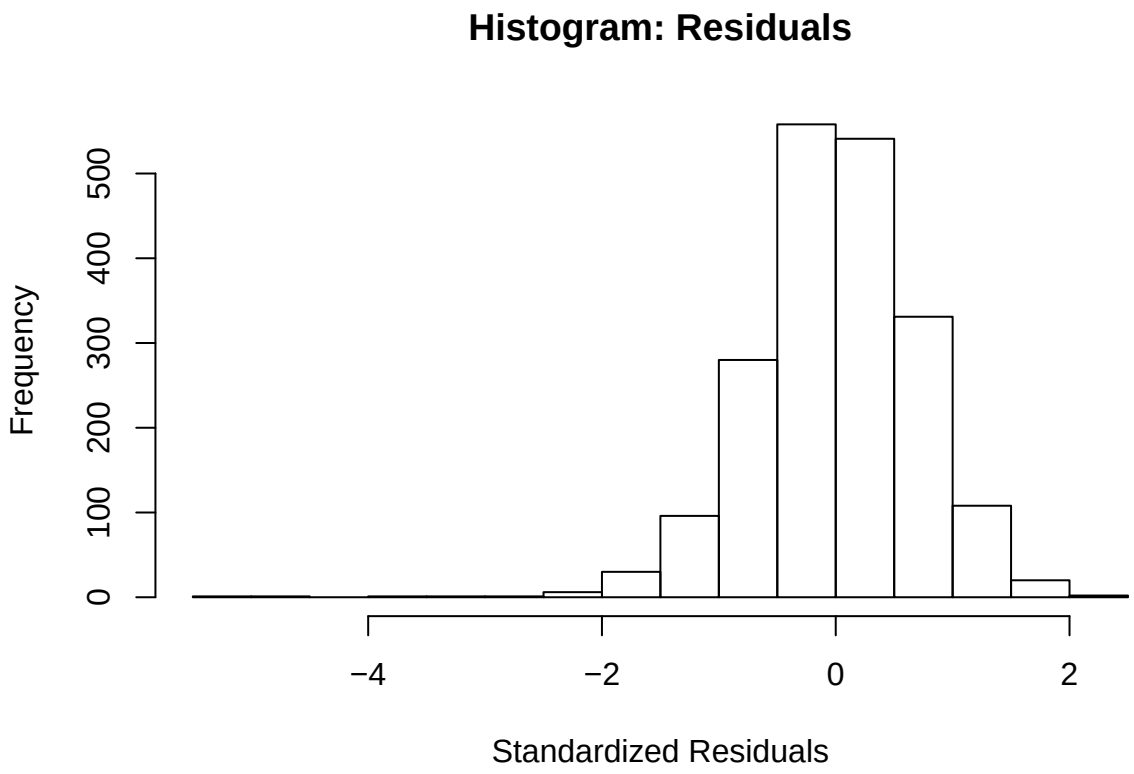



```
# Acf func in FinTs package omits zero lag
Acf(as.vector(resid(mod)),lag.max=365*2,main="ACF: Residuals")
```

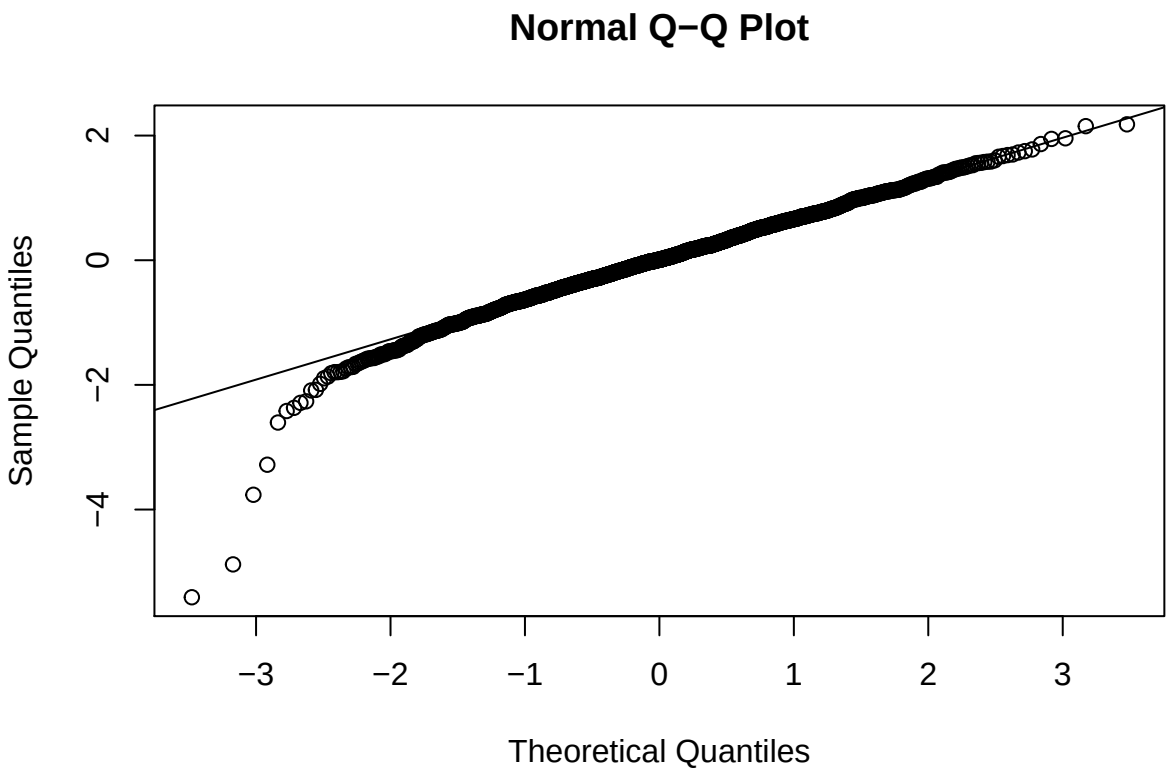


using histogram and qq plot to see wheather the residual is normal distributed

```
hist(resid(mod),xlab='Standardized Residuals',main='Histogram: Residuals')
```

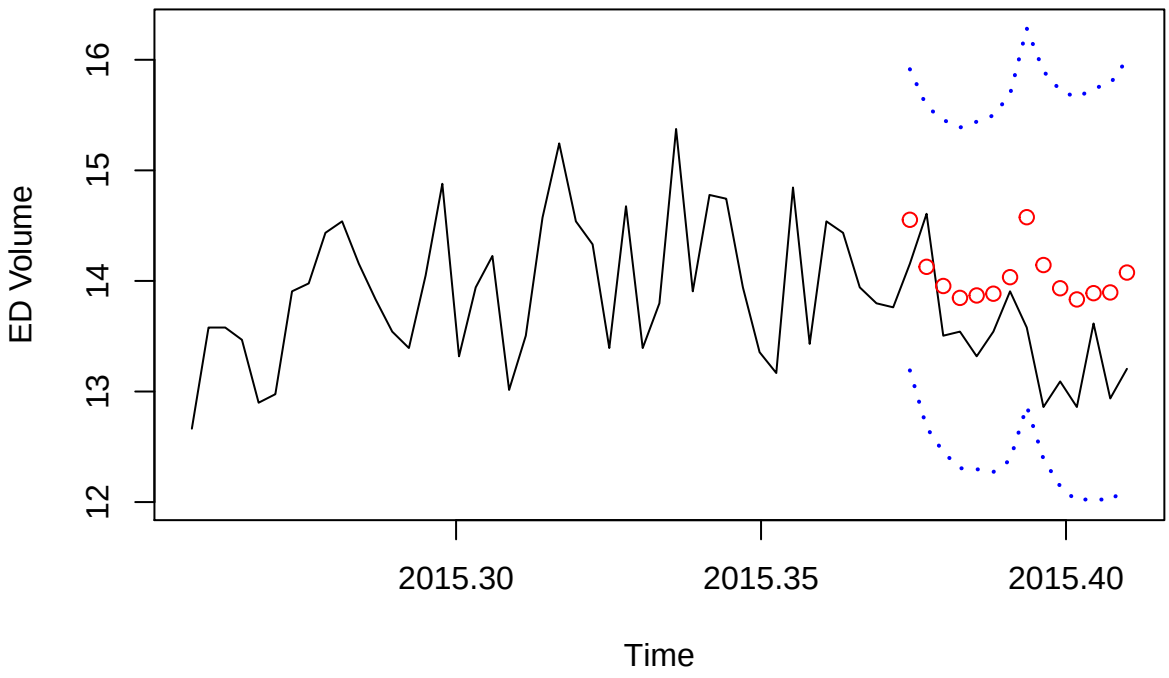



```
qqnorm(resid(mod))
qqline(resid(mod))
```



3.Forecasting with ARIMA: 2 Weeks Ahead

```
n = length(volume.ts)
nfit = n-14
outvol = arima(volume.ts[1:nfit], order = c(5,1,5),seasonal = list(order = c(1,0,1),period=7),method =
out_pred = as.vector(predict(outvol,n.ahead=14))

timevol=time(volume.ts)
ubound = out_pred$pred+1.96*out_pred$se #create 95% bound
lbound = out_pred$pred-1.96*out_pred$se
ymin = min(lbound)
ymax = max(ubound)
par(mfrow=c(1,1))
#plot the ts that is close to predicting period
plot(timevol[(n-56):n],volume.ts[(n-56):n],type="l", ylim=c(ymin,ymax), xlab="Time", ylab="ED Volume")
points(timevol[(nfit+1):n],out_pred$pred,col="red")
lines(timevol[(nfit+1):n],ubound,lty=3,lwd= 2, col="blue")
lines(timevol[(nfit+1):n],lbound,lty=3,lwd= 2, col="blue")
```


ED Volume
12
13
14
15
16
2015.30
2015.35
2015.40
Time