Analysis

Elizabeth Thompson

11/18/2023

```
library(forecast)
```

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

```
library(tseries)
library(fUnitRoots)
library(astsa)
```

```
##
## Attaching package: 'astsa'
```

```
## The following object is masked from 'package:forecast':
##
##     gas
```

```
library(readr)
library(readxl)
library(TSA)
```

```
## Registered S3 methods overwritten by 'TSA':
##   method       from
##   fitted.Arima forecast
##   plot.Arima   forecast
```

```
##
## Attaching package: 'TSA'
```

```
## The following object is masked from 'package:readr':
##
##     spec
```

```
## The following objects are masked from 'package:stats':
##
##     acf, arima
```

```
## The following object is masked from 'package:utils':
##
##     tar
```

```
data <- read.csv("energy_final.csv", header = TRUE)
tempdata <- read_xlsx("Temperature Data.xlsx")
```

```
data <- as.vector(data$energy_used)
head(data)
```

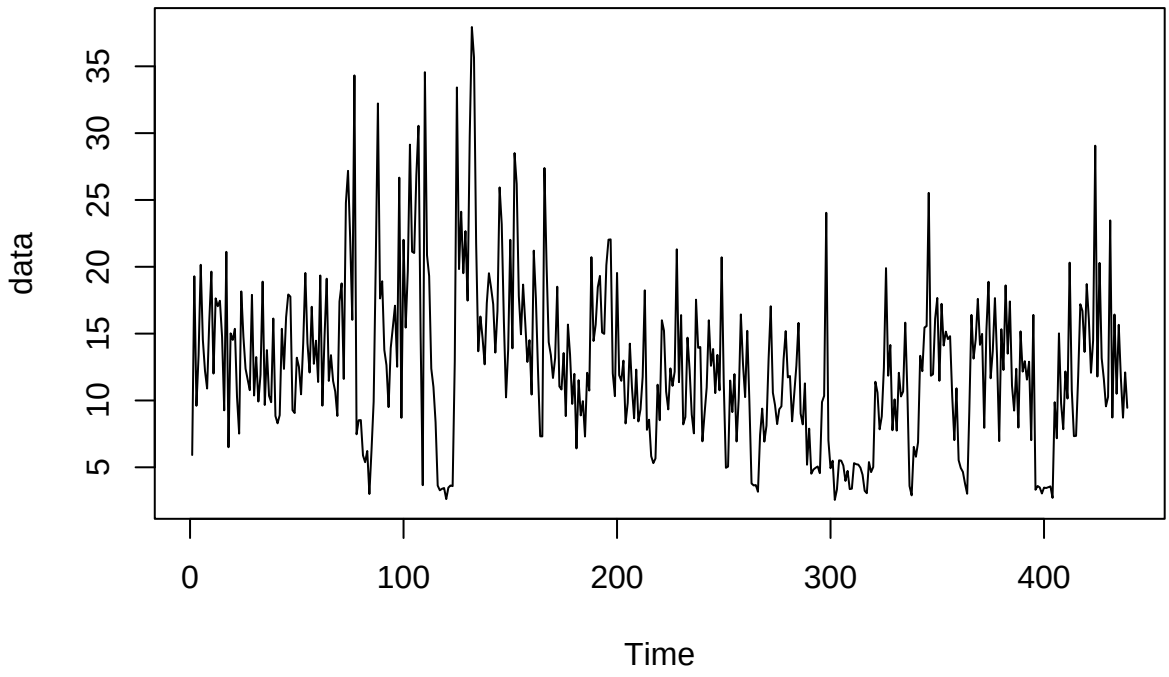
```
## [1]  5.92 19.29  9.60 13.39 20.15 14.54
```

```
write.table(data, file = "energy.txt")

tempdata <- as.vector(tempdata$Avg)
head(tempdata)
```

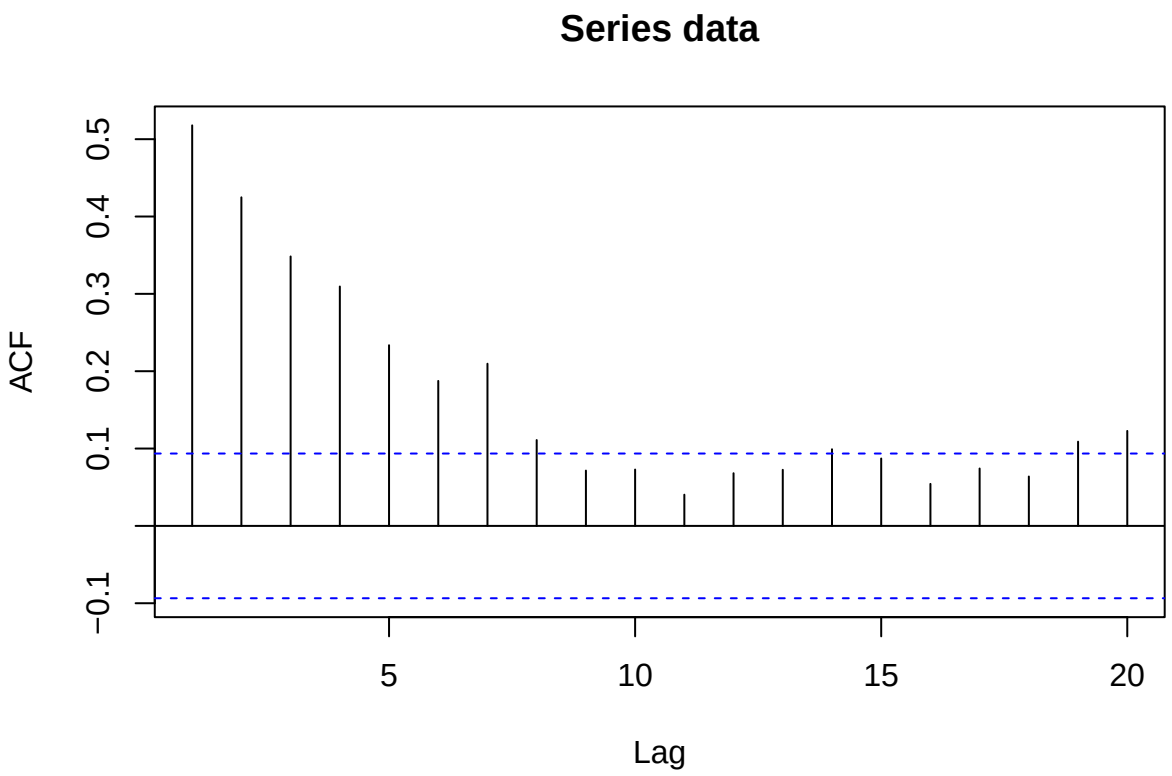
```
## [1] 79.7 75.9 72.9 72.3 74.1 78.6
```

```
write.table(tempdata, file = "avgtemp.txt")
```

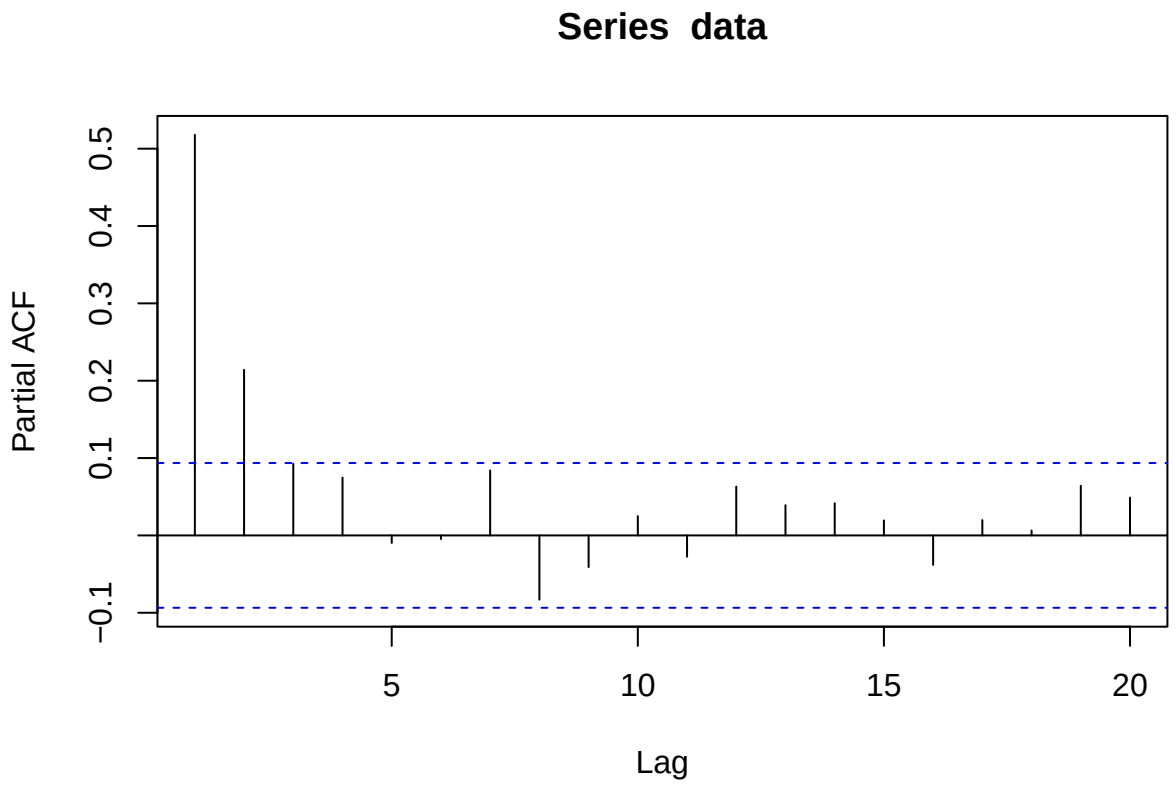
```
plot.ts(data)
```



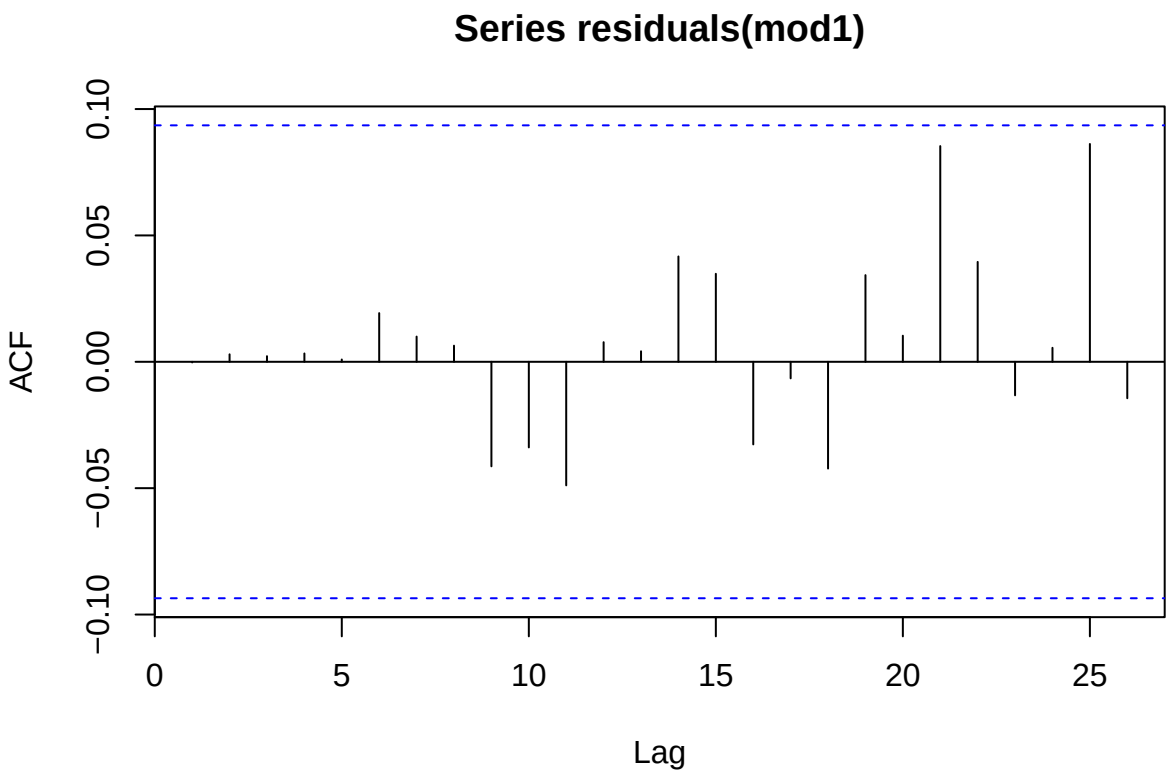

```
acf(data, lag.max = 20)
```



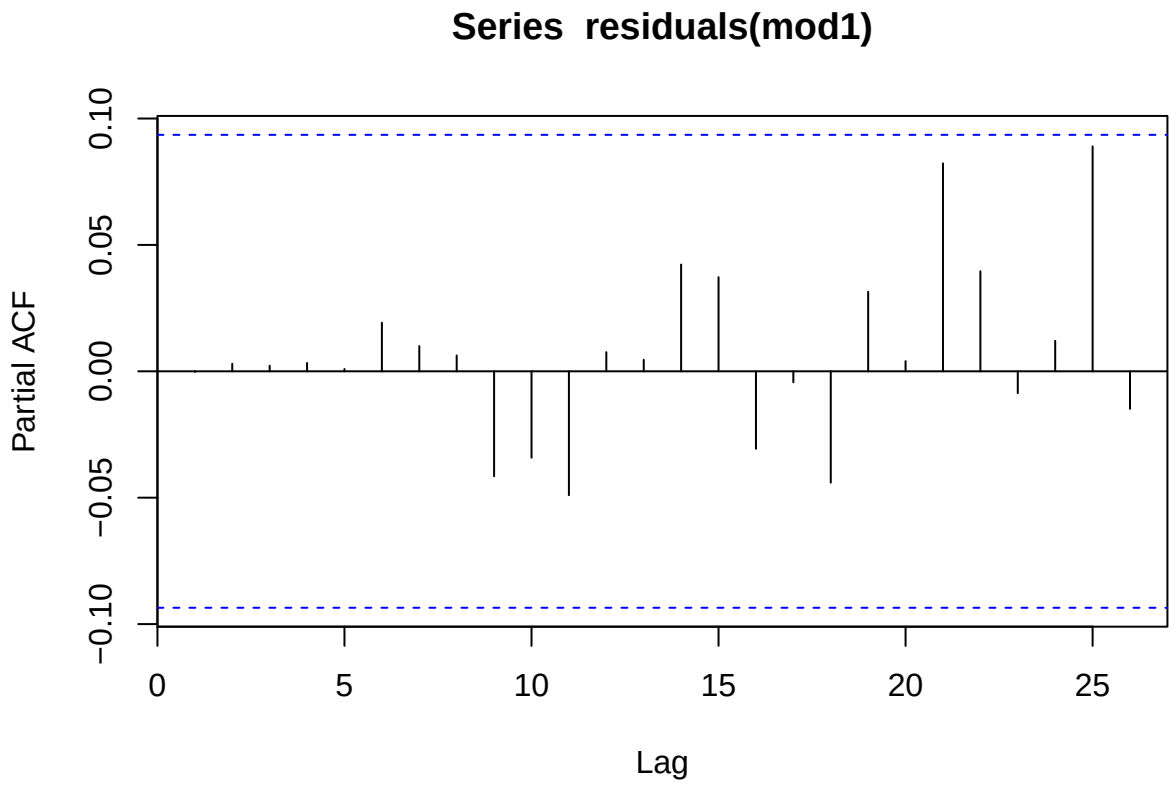

```
pacf(data, lag.max = 20)
```




```
mod1= arima(data, order = c(7, 0, 2))
acf(residuals(mod1))
```




```
pacf(residuals(mod1))
```




```
print(AIC(mod1))
```

```
## [1] 2689.39
```

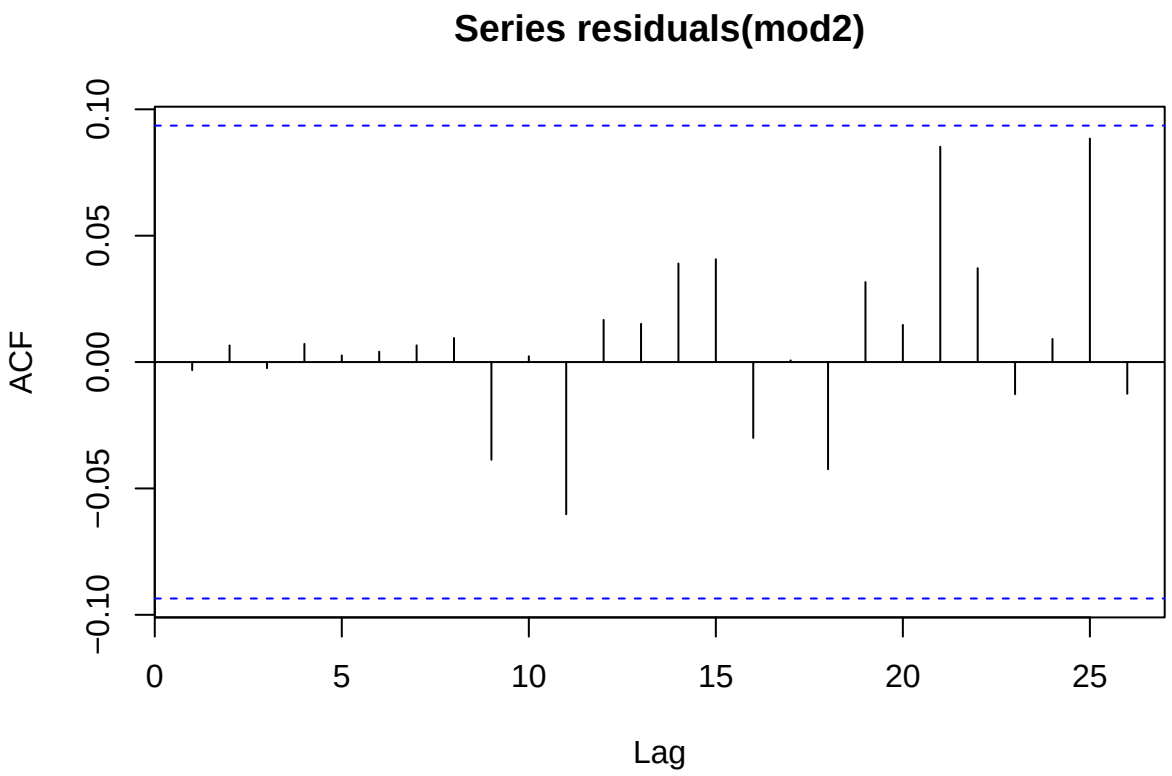
```r
print(BIC(mod1))
```

```
## [1] 2734.32
```

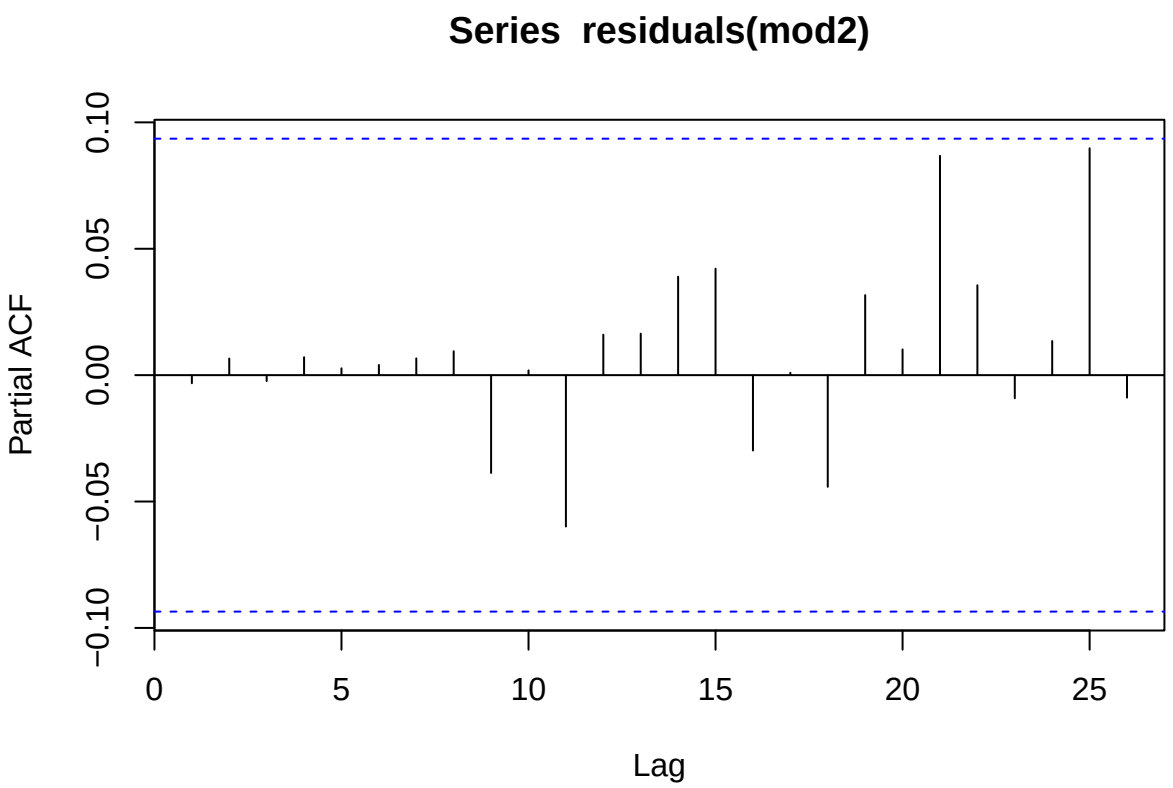
```r
residuals_mod1 <- residuals(mod1)
lags <- c(6, 12, 18, 24)
test_results <- lapply(lags, function(lag) {
  Box.test(residuals_mod1, lag=lag, type="Ljung-Box")
})

p_values <- sapply(test_results, function(x) x$p.value)
diag_table <- data.frame(Lag = lags, P_Value = p_values)

print(diag_table, digits = 16)
```

```
##   Lag            P_Value
## 1   6 0.9998913941113833
## 2  12 0.9976106749924105
## 3  18 0.9982709653174110
## 4  24 0.9941040949439239
```

```r
mod2= arima(data, order = c(8, 0, 0))
acf(residuals(mod2))
```




```r
pacf(residuals(mod2))
```




```
print(AIC(mod2))
```

```
## [1] 2687.681
```

```
print(BIC(mod2))
```

```
## [1] 2728.526
```

```
residuals_mod2 <- residuals(mod2)
lags <- c(6, 12, 18, 24)
test_results <- lapply(lags, function(lag) {
  Box.test(residuals_mod2, lag=lag, type="Ljung-Box")
})

p_values <- sapply(test_results, function(x) x$p.value)
diag_table <- data.frame(Lag = lags, P_Value = p_values)

print(diag_table, digits = 16)
```

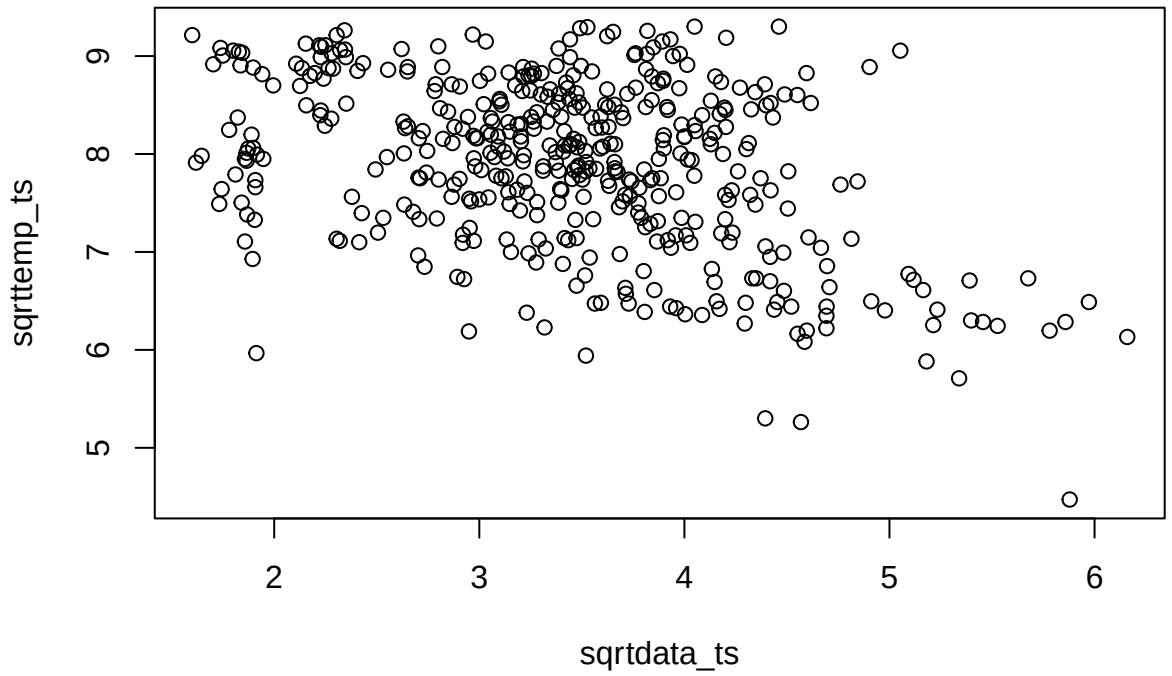
```
##   Lag            P_Value
## 1   6 0.9999956322534601
## 2  12 0.9979280963800142
## 3  18 0.9981941653470499
## 4  24 0.9944209179988882
```

```
sqrtdata = sqrt(data)
sqrttemp = sqrt(tempdata)
```

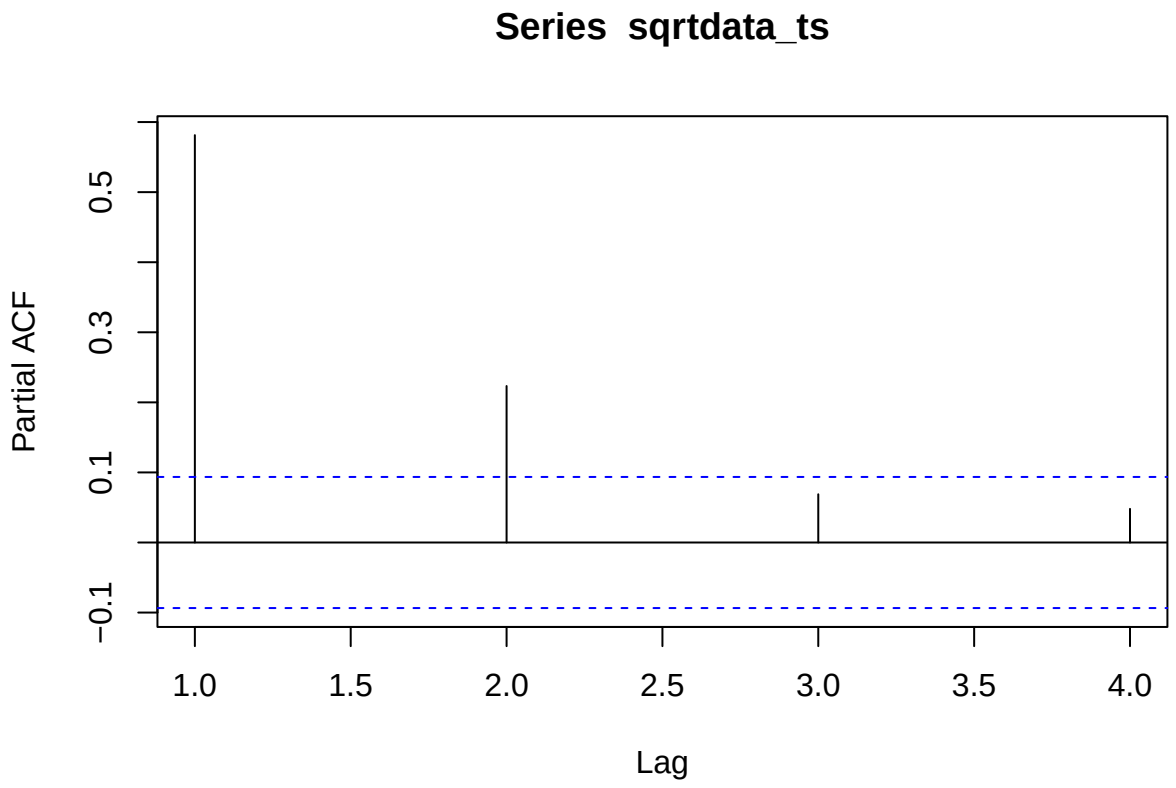
Try Transfer Function Model Yt = Energy data Xt = Temperature data

```
sqrtdata_ts = ts(sqrtdata)
sqrttemp_ts = ts(sqrttemp)
```

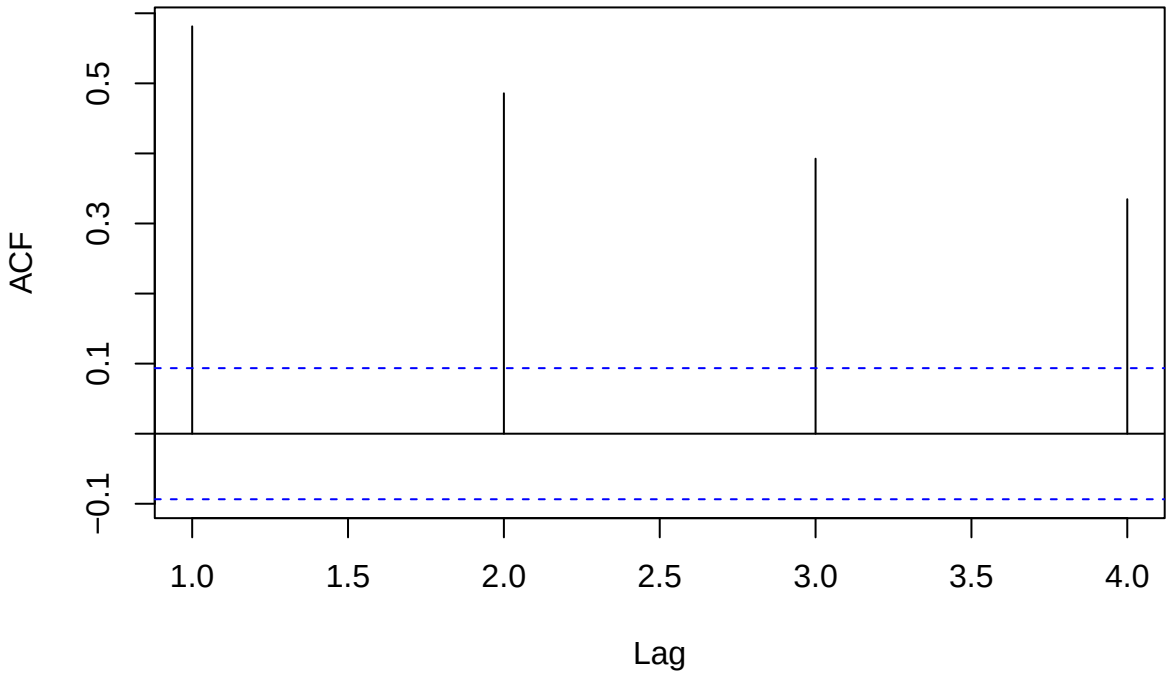
```
plot.ts(sqrtdata_ts, sqrttemp_ts)
```



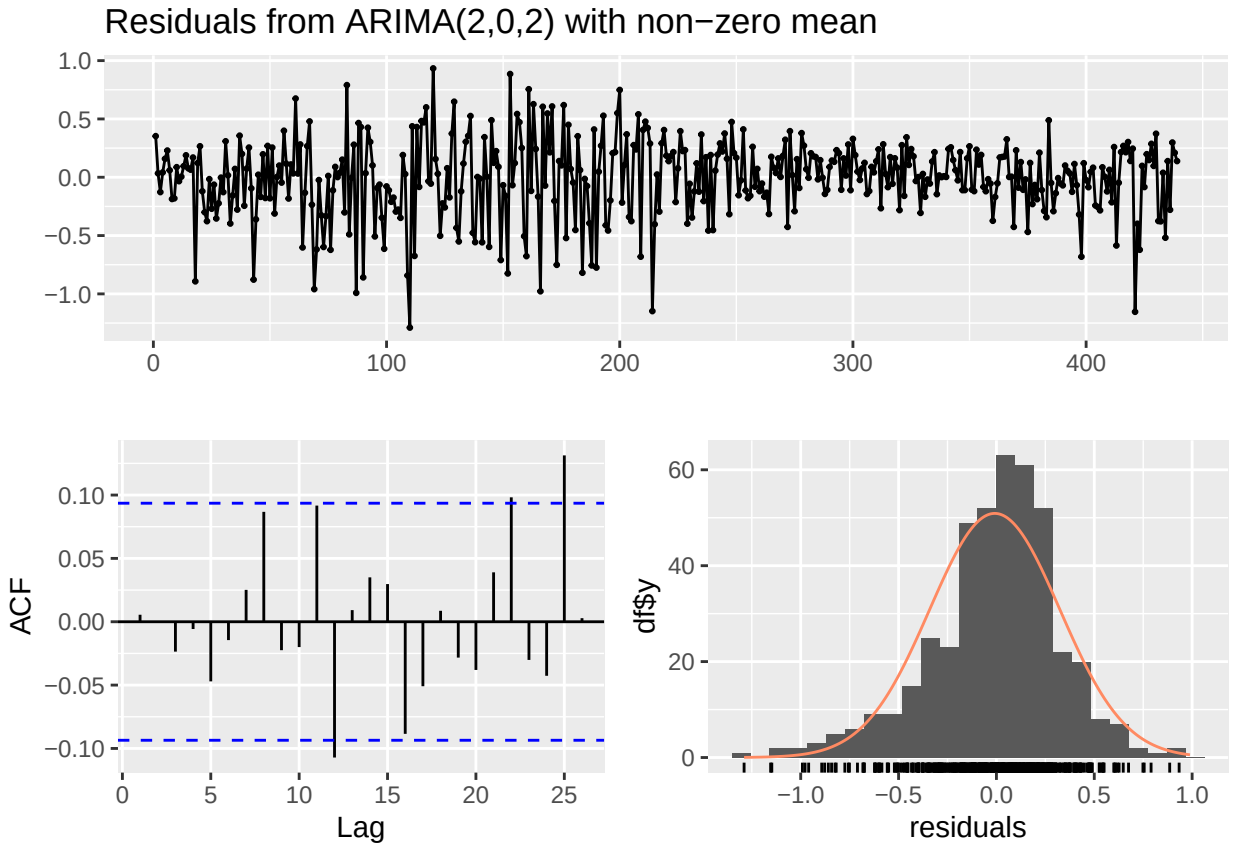

```
pacf(sqrtdata_ts, sqrttemp_ts)
```




```
acf(sqrtdata_ts, sqrttemp_ts)
```




```
tfmod = arimax(sqrttemp_ts, order=c(2,0,2), xreg=sqrtdata_ts, xtransf=sqrtdata_ts, transfer=list(c(1,0)

# Check model diagnostics
checkresiduals(tfmod)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,0,2) with non-zero mean
## Q* = 5.427, df = 6, p-value = 0.4903
##
## Model df: 4.   Total lags used: 10
```

```
# Summary of the model
summary(tfmod)
```

```
##
## Call:
## arimax(x = sqrttemp_ts, order = c(2, 0, 2), xreg = sqrtdata_ts, xtransf = sqrtdata_ts,
##     transfer = list(c(1, 0)))
##
## Coefficients:
```

```
## Warning in sqrt(diag(x$var.coef)): NaNs produced
```

```
##          ar1      ar2      ma1      ma2  intercept     xreg   T1-AR1  T1-MA0
##       1.4663  -0.4715  -0.4699  -0.3775     8.2177  -0.1367  -0.2298  0.0731
## s.e.  0.0741   0.0726   0.0714   0.0502     0.4003      NaN      NaN     NaN
##
## sigma^2 estimated as 0.1104:  log likelihood = -140.32,  aic = 296.65
##
## Training set error measures:
##                       ME      RMSE       MAE        MPE     MAPE      MASE
## Training set -0.008510468 0.3322535 0.2498361 -0.3185433 3.346807 0.9130868
```

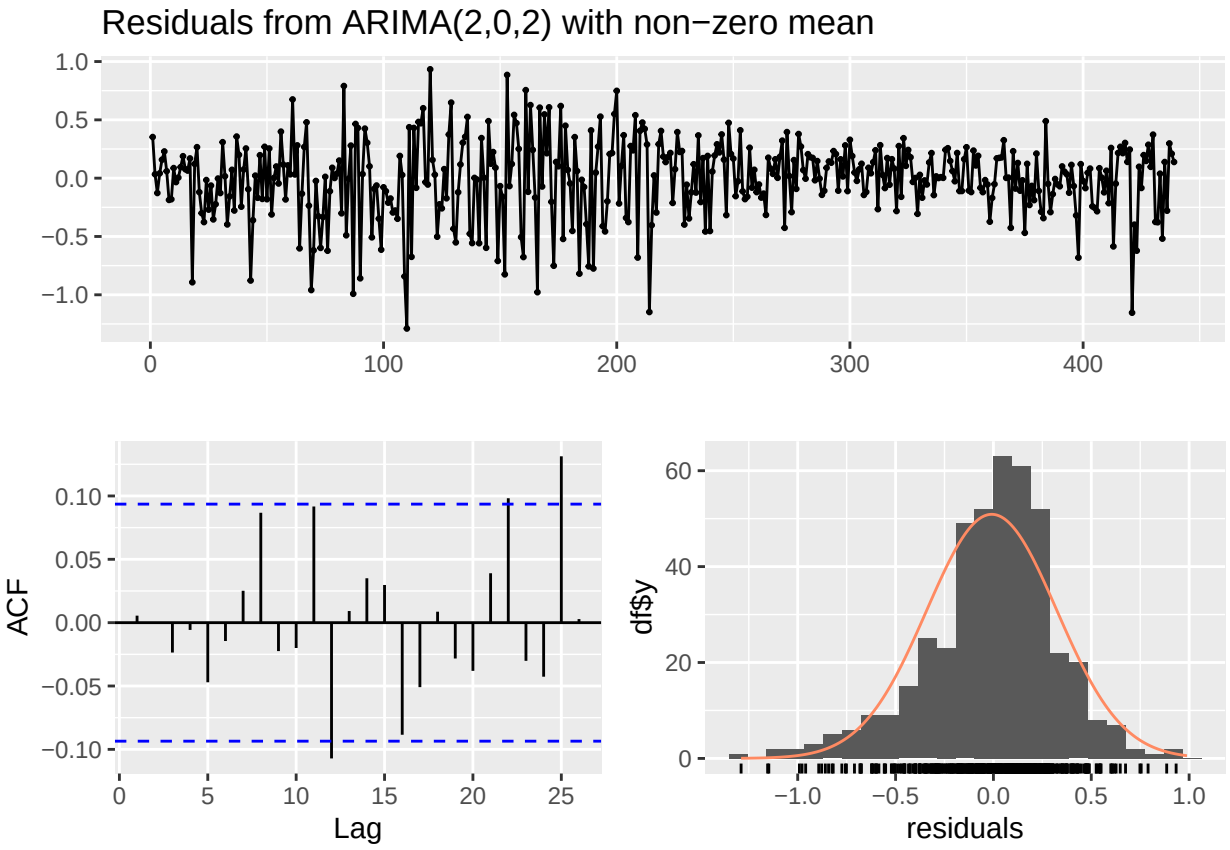
```
##                      ACF1
## Training set 0.005530523
```

```
residuals_tfmod <- residuals(tfmod)
lags <- c(6, 12, 18, 24)
test_results <- lapply(lags, function(lag) {
  Box.test(residuals_tfmod, lag=lag, type="Ljung-Box")
})

p_values <- sapply(test_results, function(x) x$p.value)
diag_table <- data.frame(Lag = lags, P_Value = p_values)

print(diag_table, digits = 16)
```

```
##   Lag            P_Value
## 1   6 0.9684245902051334
## 2  12 0.2738408751120571
## 3  18 0.3194171641710828
## 4  24 0.2716634600960421
```

There is feedback in our Transfer Function Model

```
checkresiduals(tfmod)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,0,2) with non-zero mean
## Q* = 5.427, df = 6, p-value = 0.4903
```

```
##
## Model df: 4.   Total lags used: 10
```

```
tempdata_ts = ts(tempdata)
sqrtdata_ts = ts(sqrtdata)

# Compute and plot cross-correlation
ccf_result = ccf(sqrtdata_ts, tempdata_ts, lag.max=10, plot = FALSE)

plot(ccf_result, main="Cross-Correlation between x and y", ylab="CCF")
```