STAT 6555 Lanqin Zhao FinalTH

Generation of Electricity Forecasting

I utilize the dataset: the-total-generation-of-electric.csv, the monthly data of total generation of electricity by the U.S. electric industry for the period January 1985 to October 1996, to build a SARIMA model and then use it to forecast the generation of electricity in next 14 months.

Exploring and preparing the data

Observing the time series plot of monthly total generation of electricity plotted in Figure 1, there is a strong trend and a strong yearly seasonality existing, which suggests a first difference and a seasonal (12nd) difference. The seasonal difference of differenced monthly total generation of electricity is shown in Figure 2 and appears to be a stable process.

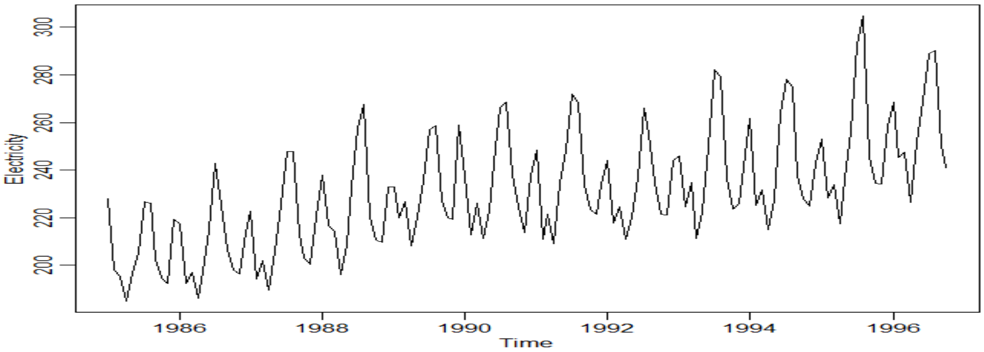

Fig. 1 monthly total generation of electricity

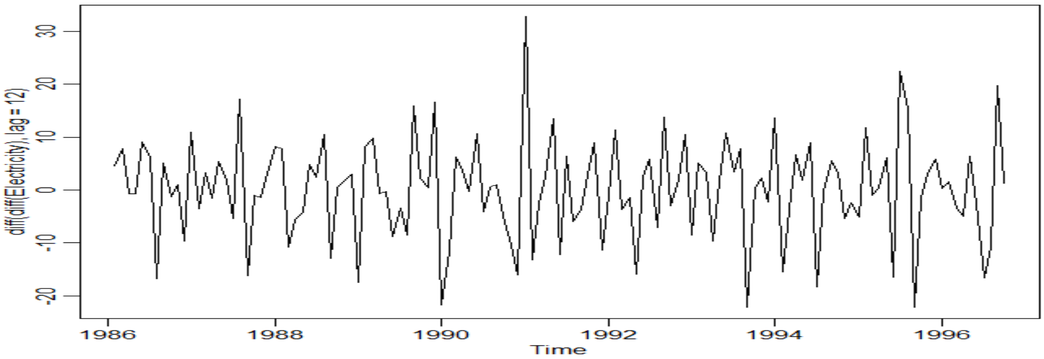

Fig. 2 seasonal difference of differenced monthly total generation of electricity.

Figure 3 shows the ACF and PACF of seasonal difference of differenced monthly total generation of electricity. Inspecting the ACF and the PACF at the seasonal (s=12) lags, there is a strong peak at h=1s in the ACF and two peaks at h=1s, 2s in the PACF. It

appears that either (i) the ACF is cutting off after lag 1s and the PACF is tailing off at the seasonal lags, or (ii) the PACF is cutting off after lag 2s and the ACF is tailing off at the seasonal lags. Using Table 3.3, this suggests either (i) P=0, Q=1 or (ii) P = 2, Q=0.
Then, inspecting the ACF and the PACF at the within seasonal lags, h=1,2, ... 11, there are peaks at h=1, 2 in the ACF and peaks at h=1, 2, 3,4 in the PACF. It appears that either (i) the ACF is cutting off after lag 2 and the PACF is tailing off, or (ii) the both are tailing off. Based on table 3.1, it suggests either (i) p=0 and q=2 or (ii) p=2 and q=1.

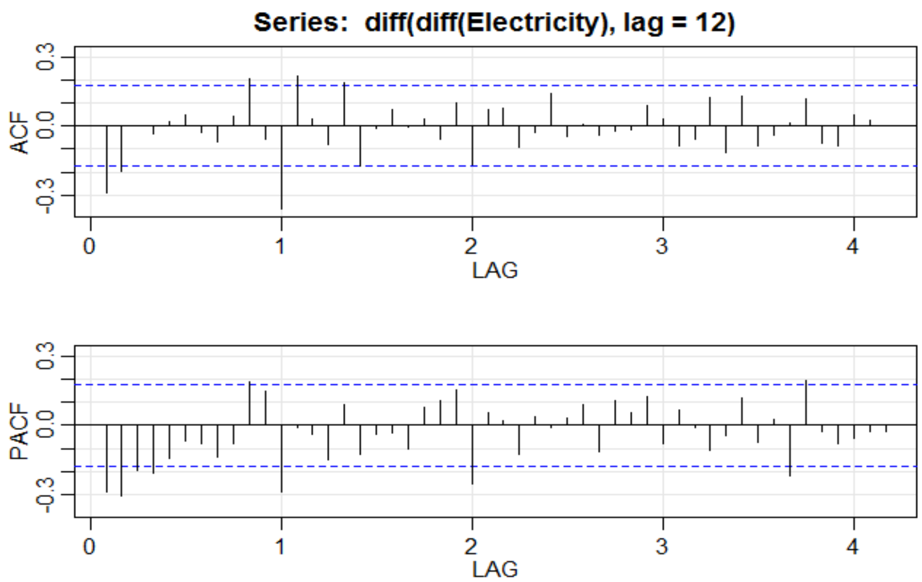

Fig. 3 ACF and PACF of seasonal difference of differenced monthly total generation of electricity

Training models on the data

Fitting the four models suggested by the above observations , we obtain:
(i) ARIMA(0; 1; 2) * $(0; 1; 1)_{12}$:
AIC= 4.751, AICc= 4.767, BIC=3.813
(ii) ARIMA(0; 1; 2) * $(2; 1; 0)_{12}$:
AIC= 4.898, AICc= 4.915, BIC=3.981
(iii) ARIMA(2; 1; 1) * $(0; 1; 1)_{12}$:
AIC= 4.763, AICc= 4.780, BIC=3.846
(iv) ARIMA(2; 1; 1) * $(2; 1; 0)_{12}$:
AIC= 4.909, AICc= 4.927, BIC=4.013
The ARIMA(0; 1; 2) * $(0; 1; 1)_{12}$ is the preferred model, and the fitted model in this case is

$$\nabla_{12}\nabla Electricity = (1-0.8053B^{12})(1-0.5755B-2298B^{2})\hat{w}_t \quad \text{with} \quad \hat{\sigma}_w^2 = 0.001367$$

all coefficients are significant.

Diagnosing

The diagnostics of residuals from the model ARIMA(0; 1; 2) * (0; 1; 1)$_{12}$ are displayed in Figure 4. Inspection of the time plot of the standardized residuals shows no obvious patterns. There are outliers, however, with a few values exceeding 3 standard deviations in magnitude. The ACF of the standardized residuals shows no apparent departure from the model assumptions, and the Q-statistic is never significant at the lags shown. The normal Q-Q plot of the residuals shows acceptable normality. There is departure at the tails due to the outliers. Overall, the model fits the data well.

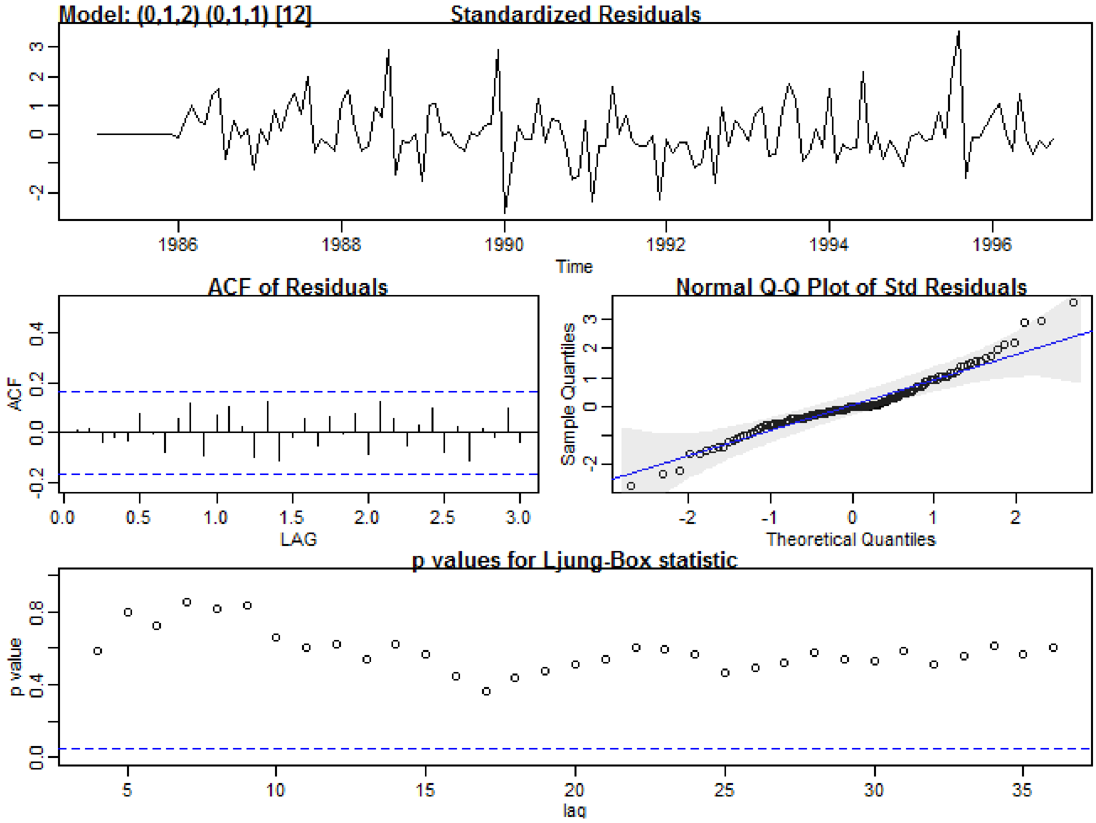

Fig. 4 Diagnostics of the residuals from the ARIMA(0; 1; 2) * (0; 1; 1)$_{12}$ fit on the monthly total generation of electricity data.

Forecasting

Using the model built to forecast the next 14 months: November and December in 1996, and 12 months in 1997, yields the results:

The forecasts and 95% confident intervals for next 14 months are:

Time	Electricity Estimate	Lower Bound	Upper Bound
Nov 1996	240.3628	227.8271	252.8986
Dec 1996	262.6877	249.0694	276.3060
Jan 1997	270.3573	256.5248	284.1897
Feb 1997	244.0965	230.0506	258.1424

```
Mar 1997       249.9720 235.7159 264.2281
Apr 1997       232.8143 218.3510 247.2776
May 1997       249.4647 234.7972 264.1322
Jun 1997       270.8394 255.9705 285.7084
Jul 1997       295.6689 280.6012 310.7366
Aug 1997       295.8175 280.5536 311.0813
Sep 1997       257.2003 241.7428 272.6578
Oct 1997       246.5068 230.8580 262.1556
Nov 1997       245.1111 228.7095 261.5127
Dec 1997       267.2620 250.4930 284.0311
```

The forecasts and limits are also shown in Figure 5.

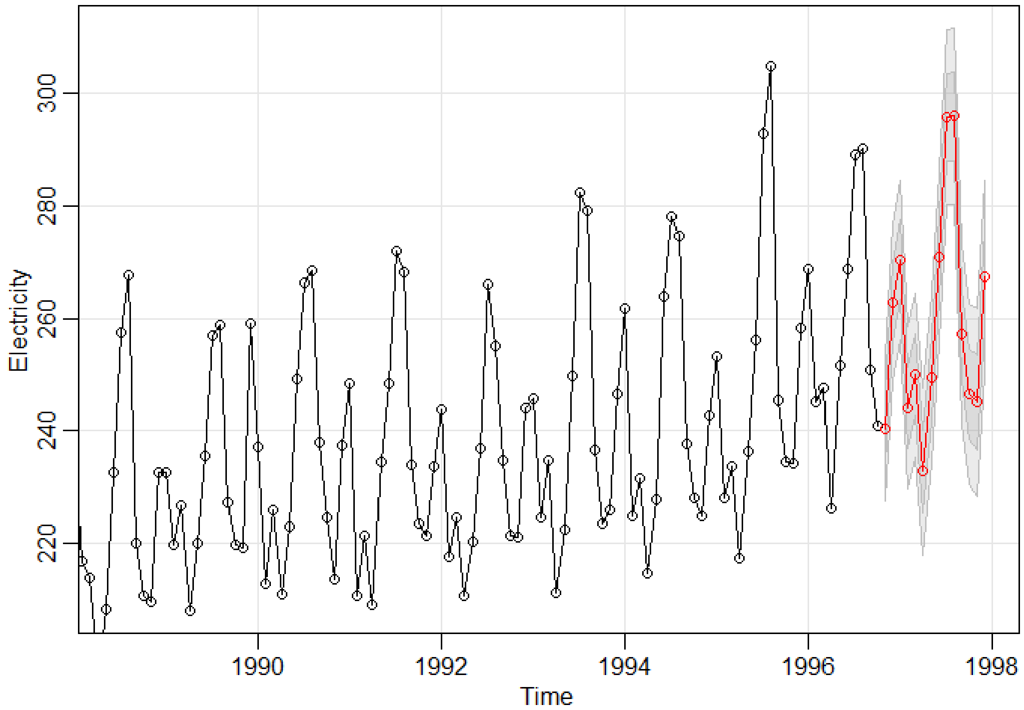

Fig. 5 14 months forecasts and limits for generation of electricity.

Appendix

R Code:

```
#data exploration and preparation
library(astsa)
electric=read.csv('the-total-generation-of-electric1.csv',head=F)
electric=ts(electric,frequency = 12,start=c(1985,1),names='Electricity')
Electricity=electric[,1]
plot(Electricity)
acf2(Electricity,max.lag = 50)
acf2(diff(Electricity),max.lag = 50)
plot(diff(diff(Electricity),lag=12))
abline(h=0,lty='dotted')
acf2(diff(diff(Electricity),lag=12),max.lag = 50)

#build models and diagnose
sarima(Electricity,0,1,2,0,1,1,12)
sarima(Electricity,0,1,2,2,1,0,12)
sarima(Electricity,2,1,1,0,1,1,12)
sarima(Electricity,2,1,1,2,1,0,12)

# forecast
pred=sarima.for(Electricity,14,0,1,2,0,1,1,12)
Electricity.hat=pred$pred
u=pred$pred+qnorm(1-0.05/2)*pred$se
l=pred$pred-qnorm(1-0.05/2)*pred$se
cbind(Electricity.hat, l, u)
```