Time Series Analysis

ARMA Models: Data Examples

Nicoleta Serban, Ph.D.
Professor
Stewart School of Industrial and Systems Engineering

U.S. Fuel Consumption:
ARMA Modeling

Georgia Tech

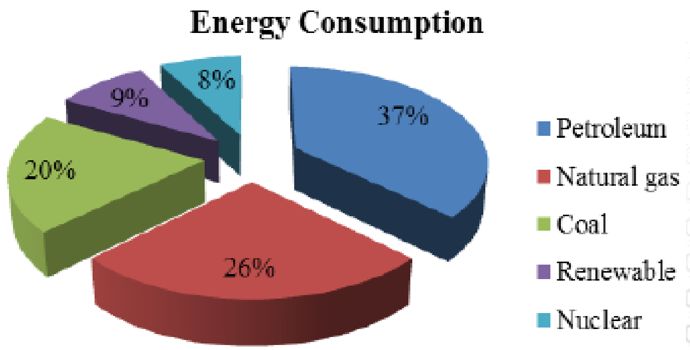
About This Lesson

Energy Consumption

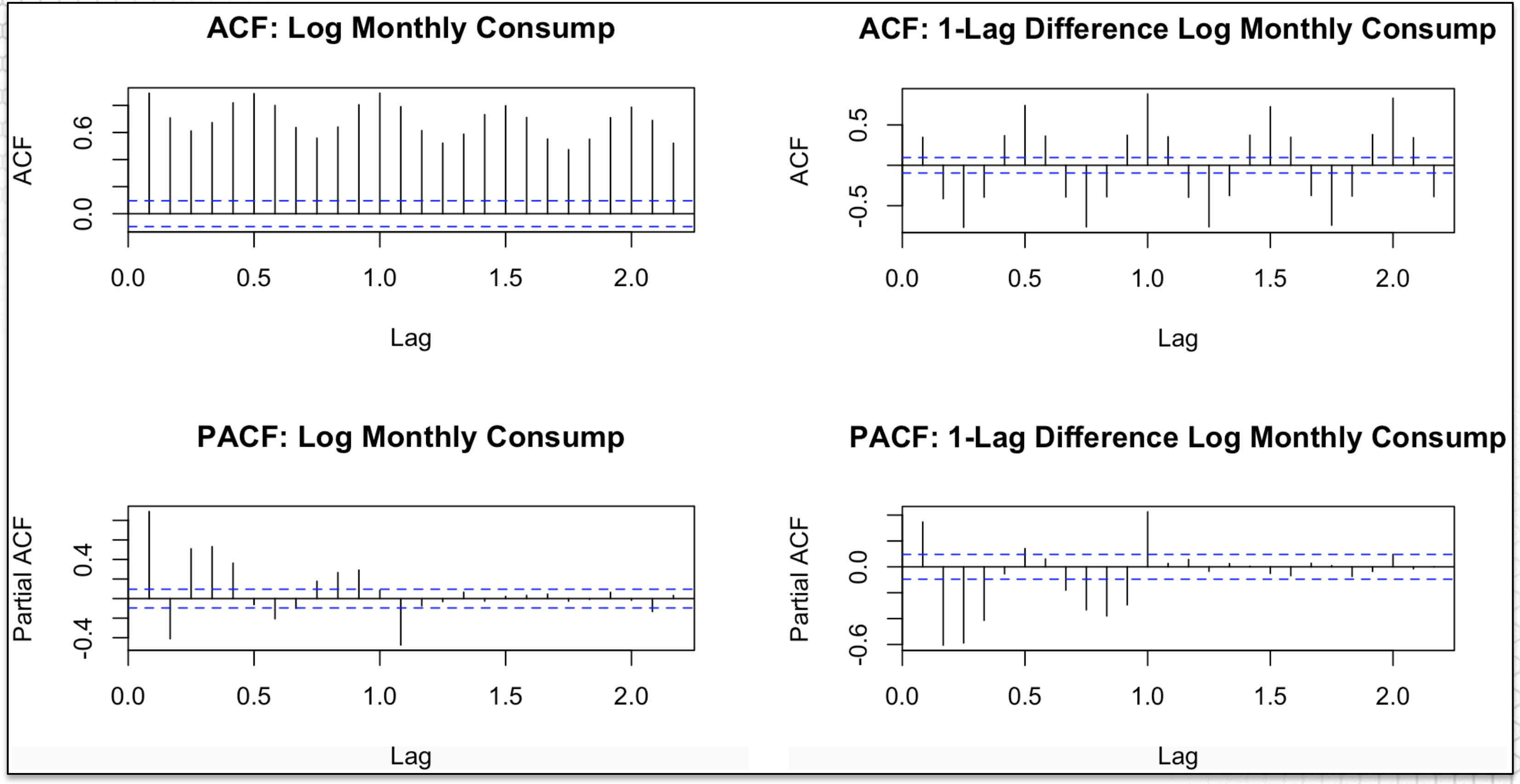
- **Data Source:** U.S. Department of Energy, The Energy Information Administration (EIA): Monthly Electricity Consumption in million megawattshours (United States Lower 48 region) over 1985-2019

Questions of interest:

- What are common characteristics in energy consumption by source?
- Can we predict energy consumption over the course of a year?



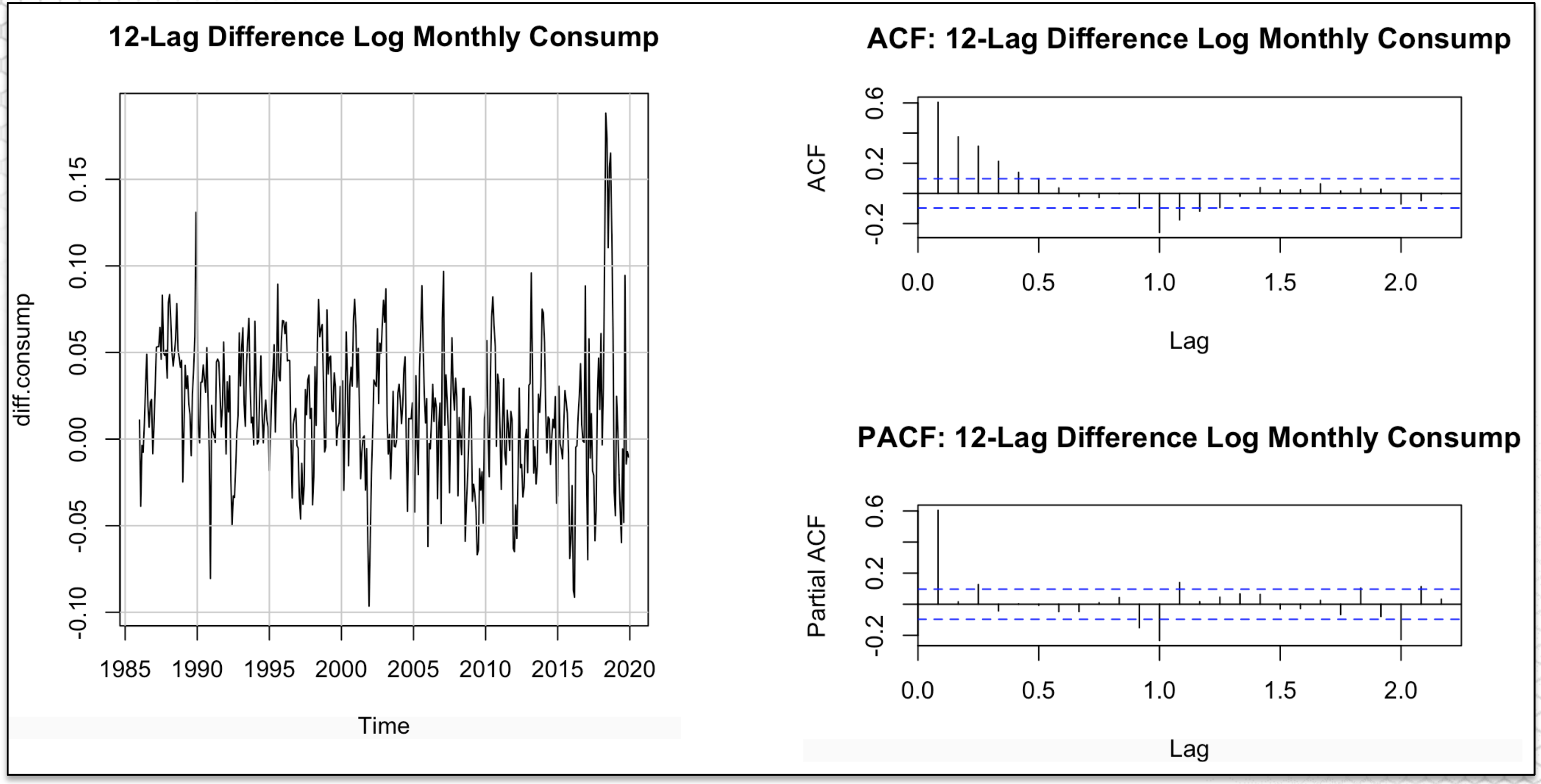
Assessing Stationarity



Assessing Stationarity: Trend & Seasonality

```
## Take the difference of lag 12 (monthly seasonality)
diff.consump = diff(log.consumption.ts,12)
diff.consump = na.omit(diff.consump)
plot(diff.consump, main="12-Lag Difference Log Monthly Consump")
grid(lty=1, col=gray(.8))
lines(dates[-c(1:12)],diff.consump,lwd=5)
par(mfrow=c(2,1))
acf(diff.consump, main="ACF: 12-Lag Difference Log Monthly Consump")
pacf(diff.consump, main="PACF: 12-Lag Difference Log Monthly
Consump")
```

Assessing Stationarity: Trend & Seasonality



Georgia Tech

Difference Time Series: ARMA Model

```
## Fit an ARMA model to the 12-lag difference log time series
## Use a maximum order of 12 due to monthly seasonality
norder = 13
p = c(1:norder)-1; q = c(1:norder)-1
n_forward=12; nfit = length(diff.consump)-n_forward
diff.consump.train = diff.consump[1:nfit]
aic = matrix(0,norder,norder)
for(i in 1:norder){
  for(j in 1:norder){
    modij = stats::arima(diff.consump.train,order = c(p[i],1,q[j]), method='ML')
    aic[i,j] = modij$aic-2*(p[i]+q[j]+1)+2*(p[i]+q[j]+1)*n/(n-p[i]-q[j]-2)
  }
}
```

ARIMA: p= 8; d=1; q = 12
AIC = -1625.579

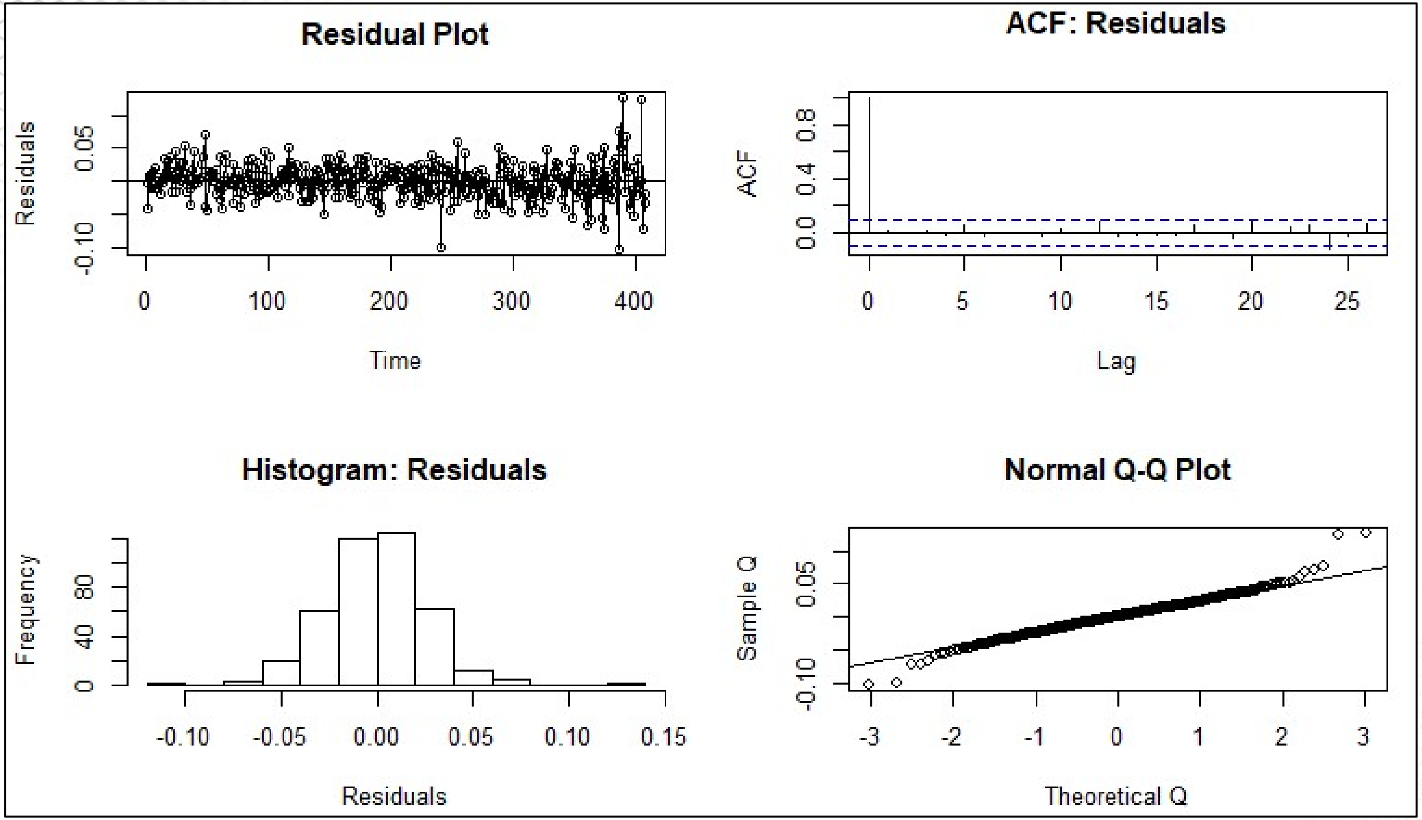
Difference Time Series: ARMA Model

```
## Fit an ARMA model to the 12-lag difference log time series
## Use a maximum order of 12 due to monthly seasonality
norder = 13
p = c(1:norder)-1; q = c(1:norder)-1
aic = matrix(0,norder,norder)
for(i in 1:norder){
  for(j in 1:norder){
    modij = stats::arima(diff.consump.train,order = c(p[i],0,q[j]), method='ML')
    aic[i,j] = modij$aic-2*(p[i]+q[j]+1)+2*(p[i]+q[j]+1)*n/(n-p[i]-q[j]-2)
  }
}
```

→

ARIMA: p= 4; d=0; q = 12
AIC = -1722.701

ARMA Model: Residual Analysis



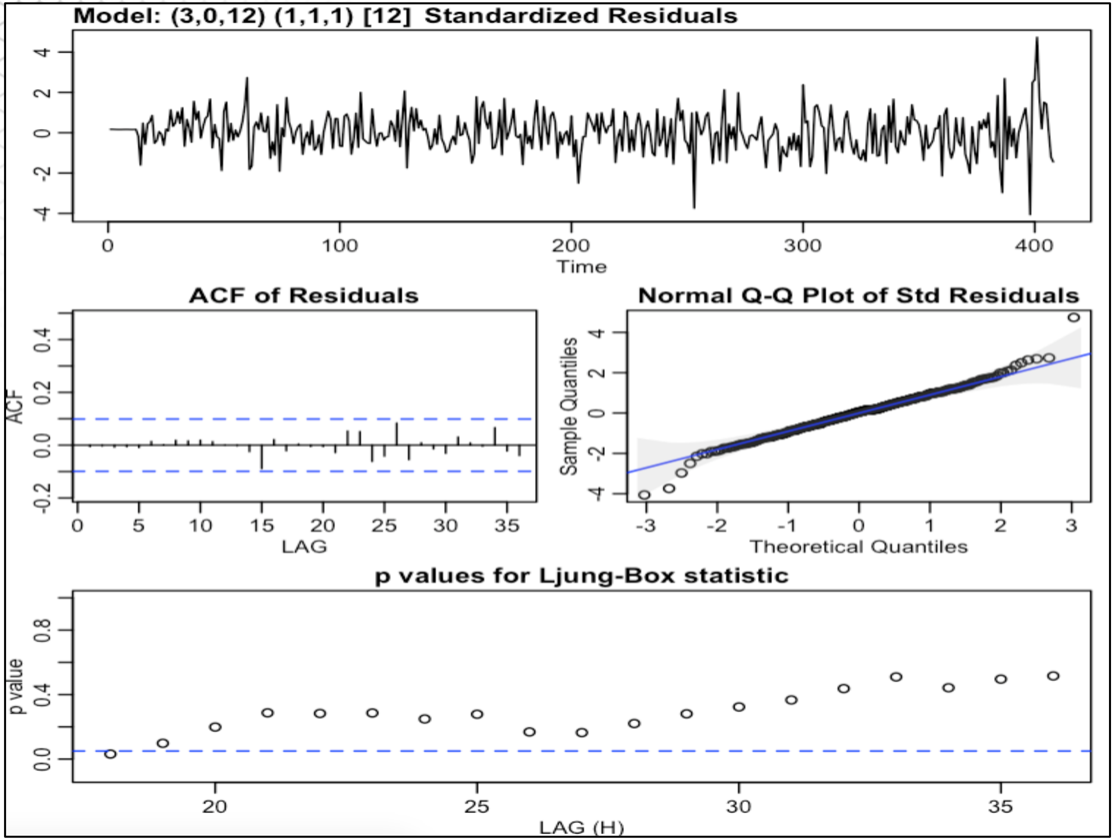
Time Series: SARIMA Model

```
## Fit an SARIMA model to the log time series
norder = 12; sorder = 2
p = c(1:norder)-1; q = c(1:norder)-1
sp = c(1:sorder)-1; sq = c(1:sorder)-1
nfit.sarima = length(log.consumption.ts)-12
log.consump.train = log.consumption.ts[1:nfit.sarima]

## hgher AR orders than 9 result in lack of convergence hence run the code for i in 1:10
## Select order for SARIMA:  p = 0,.., 9; q = 0,..,12; *sp=0; qp=1*
## For some combinations of (p,q) orders, the SARIMA model does not converge
## you will need to skip those combinations
aic_sarima_01 = matrix(0,10,norder) # # sp=1, sq=0
for(i in 1:5{
  for(j in 1:norder){
    sarima_model_select2 = sarima(log.consump.train, p[i],0,q[j], sp[2],1,sq[1], 12)
    aic_sarima_10[i,j] = sarima_model_select2$AIC
  } }
```

ARIMA: p= 3; d=0; q = 12
SARIMA: sp =1; sd= 1; sq =1

Time Series: SARIMA Model



ARMA &Differencing vs SARIMA Model

Coefficients ARMA

Coefficients:

	ar1	ar2	ar3	ar4	ar5	ar6	ma1	ma2	ma3	ma4	ma5	ma6	ma7	ma8	ma9	ma10	ma11	ma12
	0.5653	-0.1662	0.1965	-0.2301	0.1300	-0.1414	0.1314	0.2100	0.0132	0.2198	0.0787	0.2334	0.1180	0.1361	0.0916	0.2191	0.1852	-0.6910
s.e.	0.1095	0.0825	0.0768	0.0825	0.0952	0.0771	0.1045	0.0973	0.0938	0.0962	0.1129	0.1240	0.1043	0.0898	0.0785	0.0755	0.0962	0.1013

	intercept
	0.0174
s.e.	0.0038

sigma^2 estimated as 0.0006395: log likelihood = 882.36, aic = **-1726.72**

The difference in AIC is small.
BUT
The estimated coefficients are different.

Coefficients SARIMA

Coefficients:

	ar1	ar2	ar3	ma1	ma2	ma3	ma4	ma5	ma6	ma7	ma8	ma9	ma10	ma11	ma12	sar1	sma1
	0.5718	-0.0399	0.0913	0.0918	0.1062	-0.0401	0.1001	-0.0119	0.0874	-0.0194	0.0053	-0.0051	0.1490	0.1443	-0.7617	-0.2695	0.4411
s.e.	0.0734	0.0707	0.0672	0.0586	0.0613	0.0606	0.0554	0.0507	0.0498	0.0436	0.0463	0.0414	0.0402	0.0494	0.0586	0.2038	0.1791

	constant
	0.0014
s.e.	0.0003

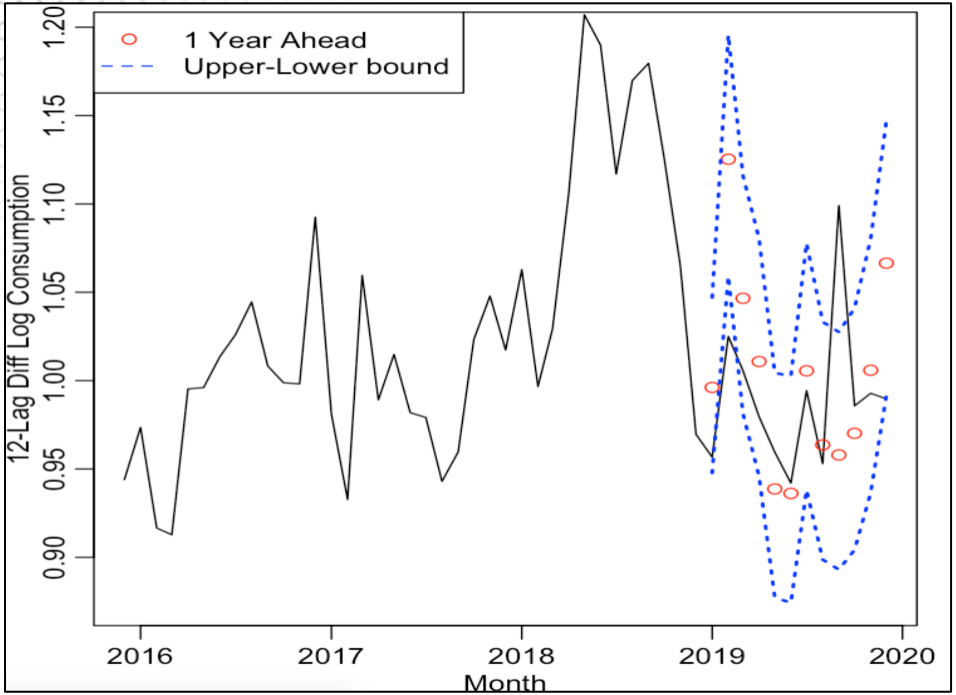
sigma^2 estimated as 0.0006416: log likelihood = 884.37, aic = **-1745.47**

ARMA & Differencing: Forecasting

```
## Forecasting with ARMA, 1 Year (12 Months) Ahead
outpred = predict(final_model,n.ahead=n_forward)
ubound = outpred$pred+1.96*outpred$se #confidenec interval
lbound = outpred$pred-1.96*outpred$se
consump_true = as.vector(exp(diff.consump[(nfit+1):n]))
consump_pred = exp(outpred$pred)

dates.diff = dates[-c(1:12)]
n = length(diff.consump)
plot((dates.diff)[(n-n_forward-20):n],exp(diff.consump[(n-n_forward-20):n]),type="l",
ylim=c(ymin,ymax), xlab="Time", ylab="12-Lag Diff Log Consumption")
points((dates.diff)[(nfit+1):n],exp(outpred$pred),col="red")
lines((dates.diff)[(nfit+1):n],exp(ubound),lty=3,lwd= 2, col="blue")
lines((dates.diff)[(nfit+1):n],exp(lbound),lty=3,lwd= 2, col="blue")
legend('topleft', legend=c("1 Year Ahead ","Upper-Lower bound"),lty = 2, col=c("red","blue"))
```

ARMA & Differencing: Forecasting



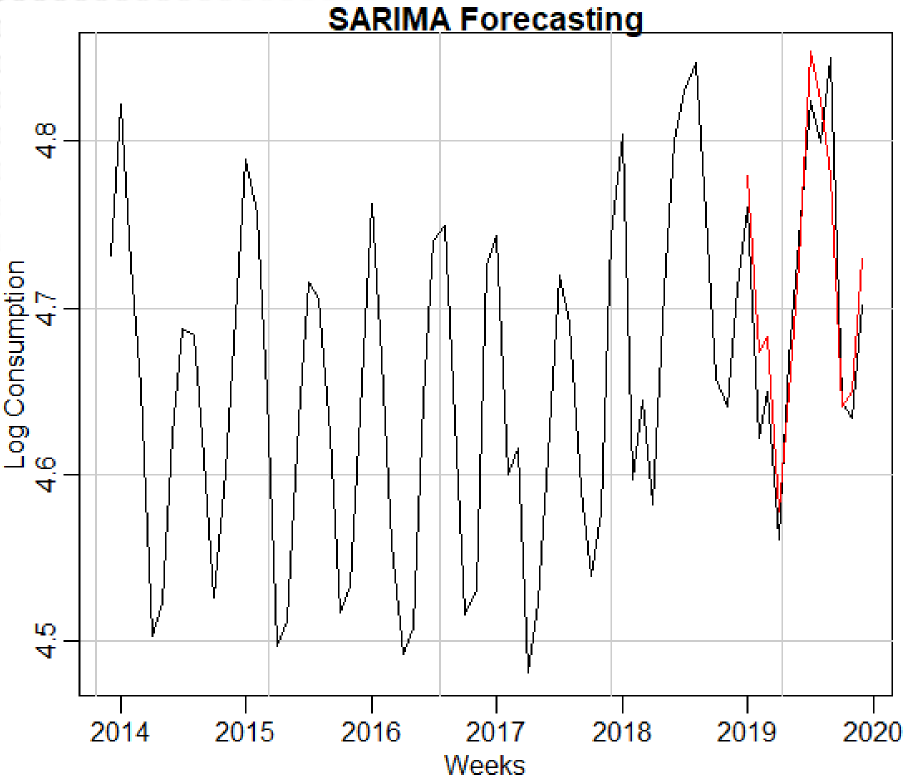
MSPE = 0.00343
MAE = 0.04224
MAPE = 0.04140
PM = 2.14127
CI = 3

SARIMA: Forecasting

```
## Forecasting with SARMA, 1 Year (12 Months) Ahead
sarima.consump = astsa::sarima(log.consumption.ts, 3,0,12, 1,1,1, 12)
sarima_1year = sarima.for(log.consump.train, 12, 3,0,12, 1,1,1, 12)
resid_SARIMA = log.consumption.ts[(nfit.sarima+1):length(log.consumption.ts)]-
sarima_1year$pred
true_SARIMA =log.consumption.ts[(nfit.sarima+1):length(log.consumption.ts)]

plot(dates[(n-n_forward-20):n],log.consumption.ts[(n-n_forward-20):n],type="l",
xlab="Weeks", ylab="Log Consumption",main="SARIMA Forecasting")
grid(lty=1, col=gray(.8))
lines(dates[(n-n_forward-20):n],log.consumption.ts[(n-n_forward-20):n])
lines(dates[(nfit.sarima+1):n],sarima_1year$pred, col=2)
```

SARIMA: Forecasting

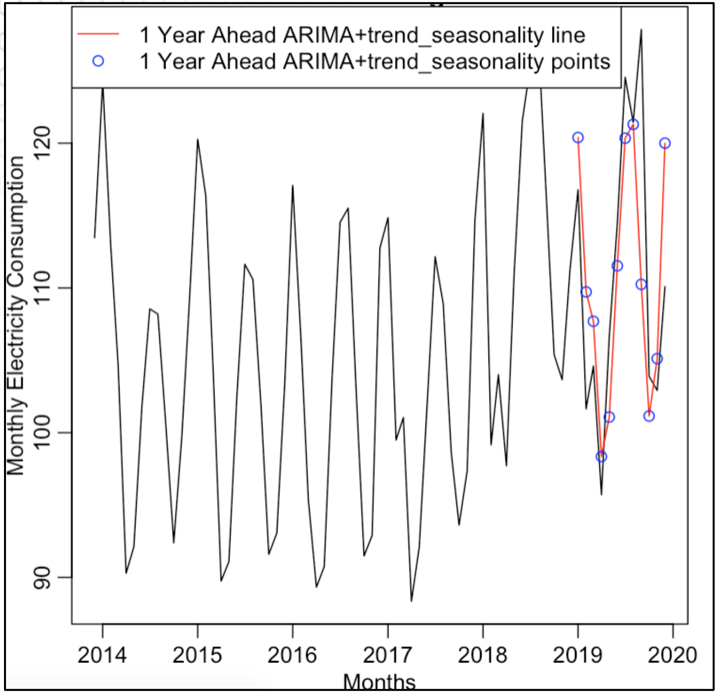


MSPE = 0.00106
MAE = 0.02736
MAPE = 0.00579
PM = 0.14295

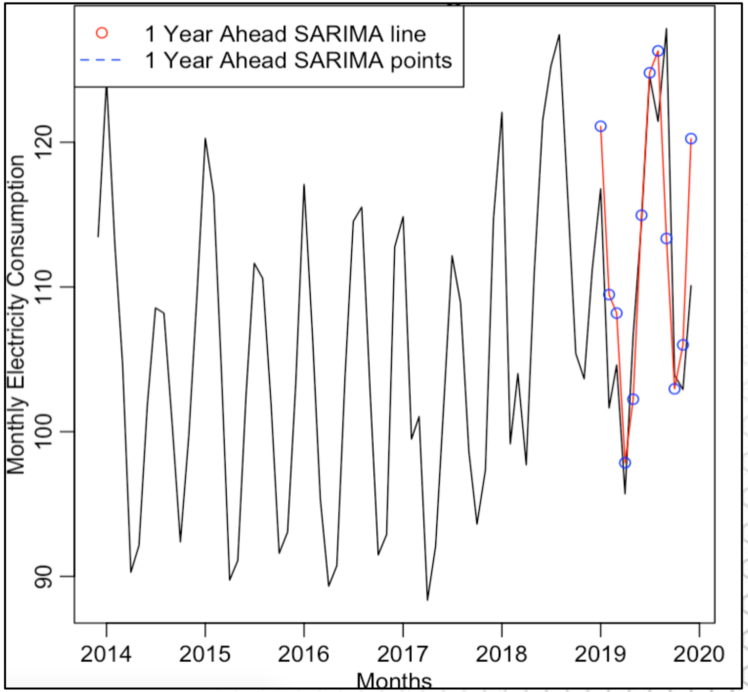
SARIMA forecast has better accuracy compared to ARIMA difference forecast.

Forecasting Comparison

ARMA: Forecasting



SARIMA: Forecasting



Georgia Tech

Summary

Georgia Tech