STAT 8020 R Lab 22: Time Series Analysis

Whitney

November 30, 2020

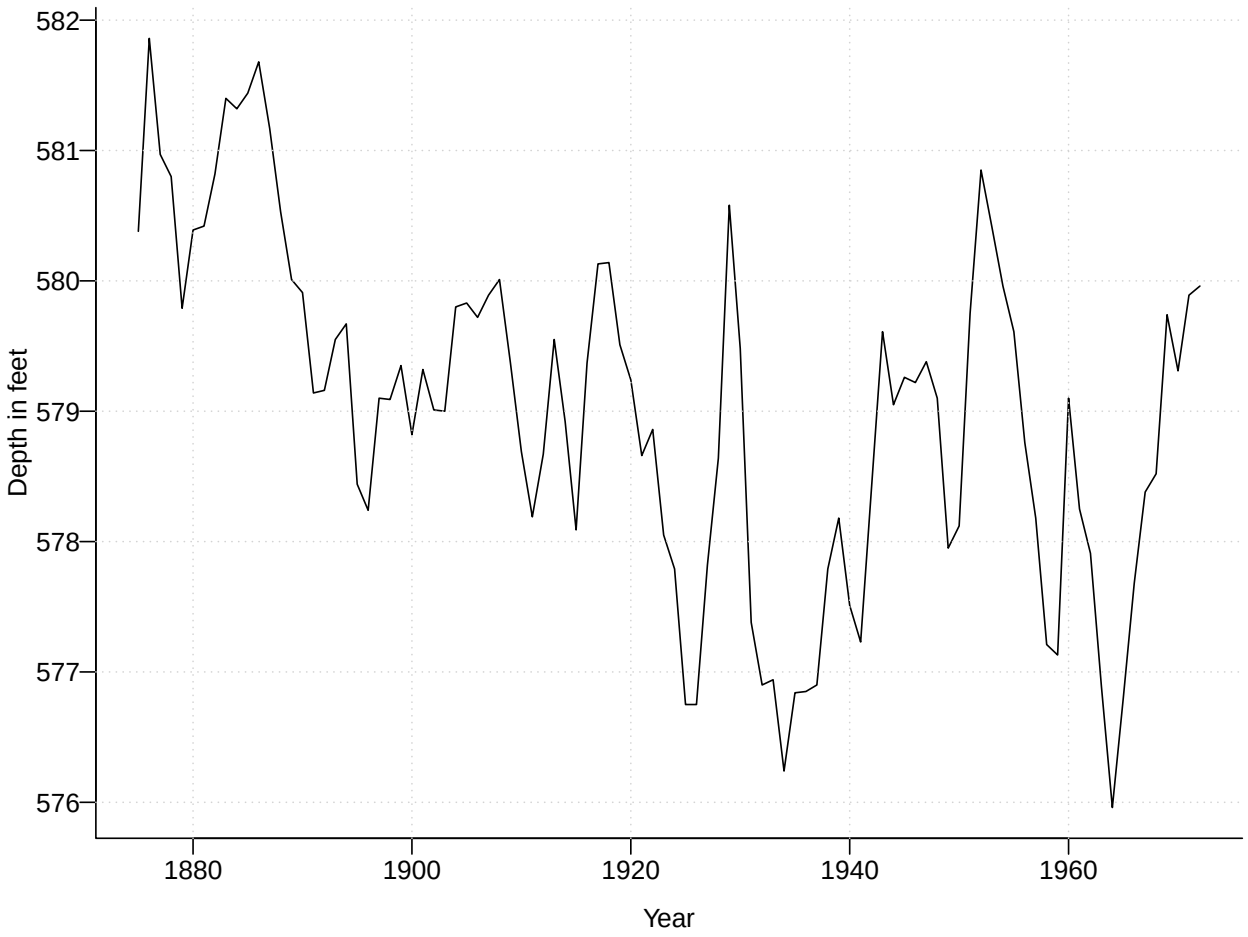
Contents

Time Series Data

Lake Huron Time Series

Annual measurements of the level of Lake Huron in feet

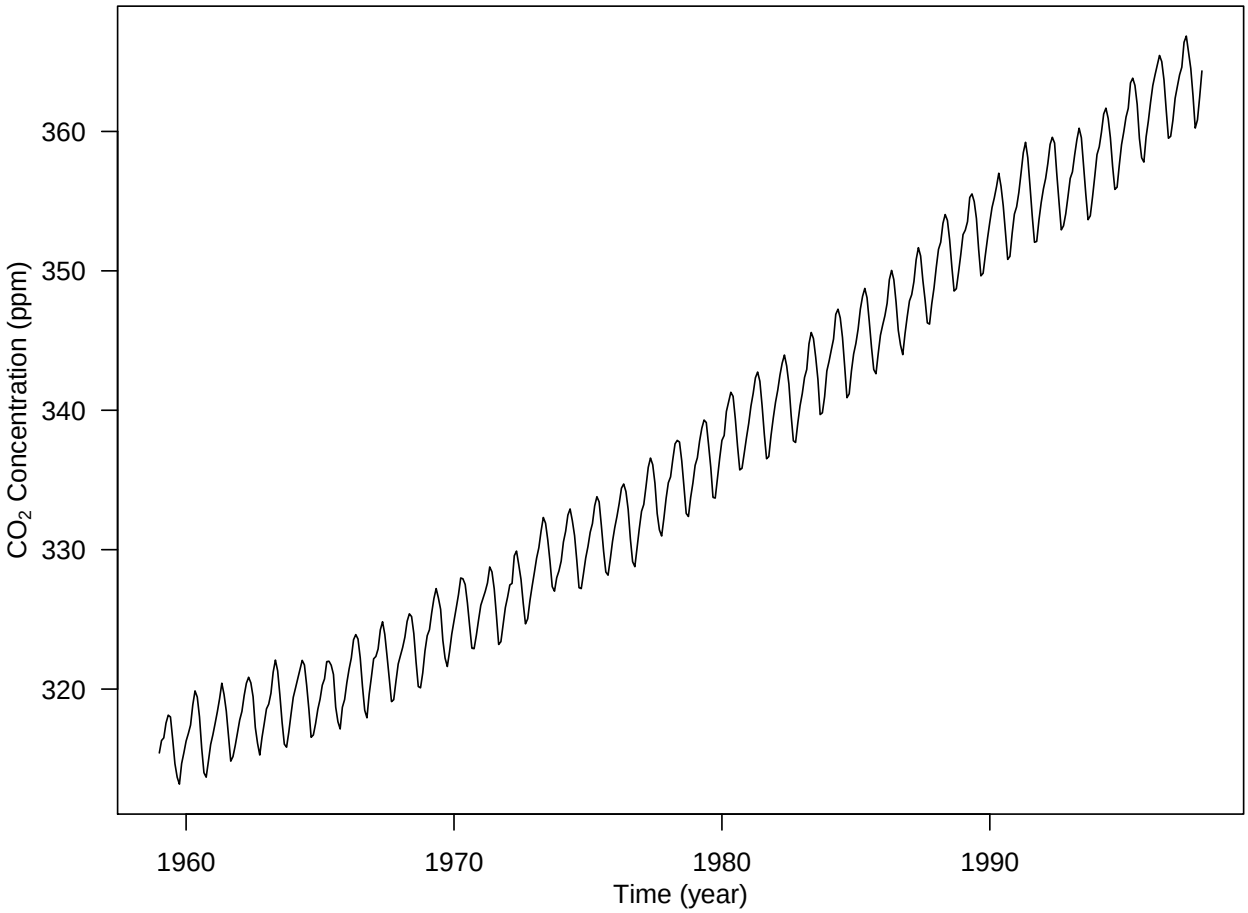
```
par(mar = c(3.2, 3.2, 0.5, 0.5), mgp = c(2, 0.5, 0), bty = "L")
data(LakeHuron)
plot(LakeHuron, ylab = "Depth in feet", xlab = "Year", las = 1)
grid()
```



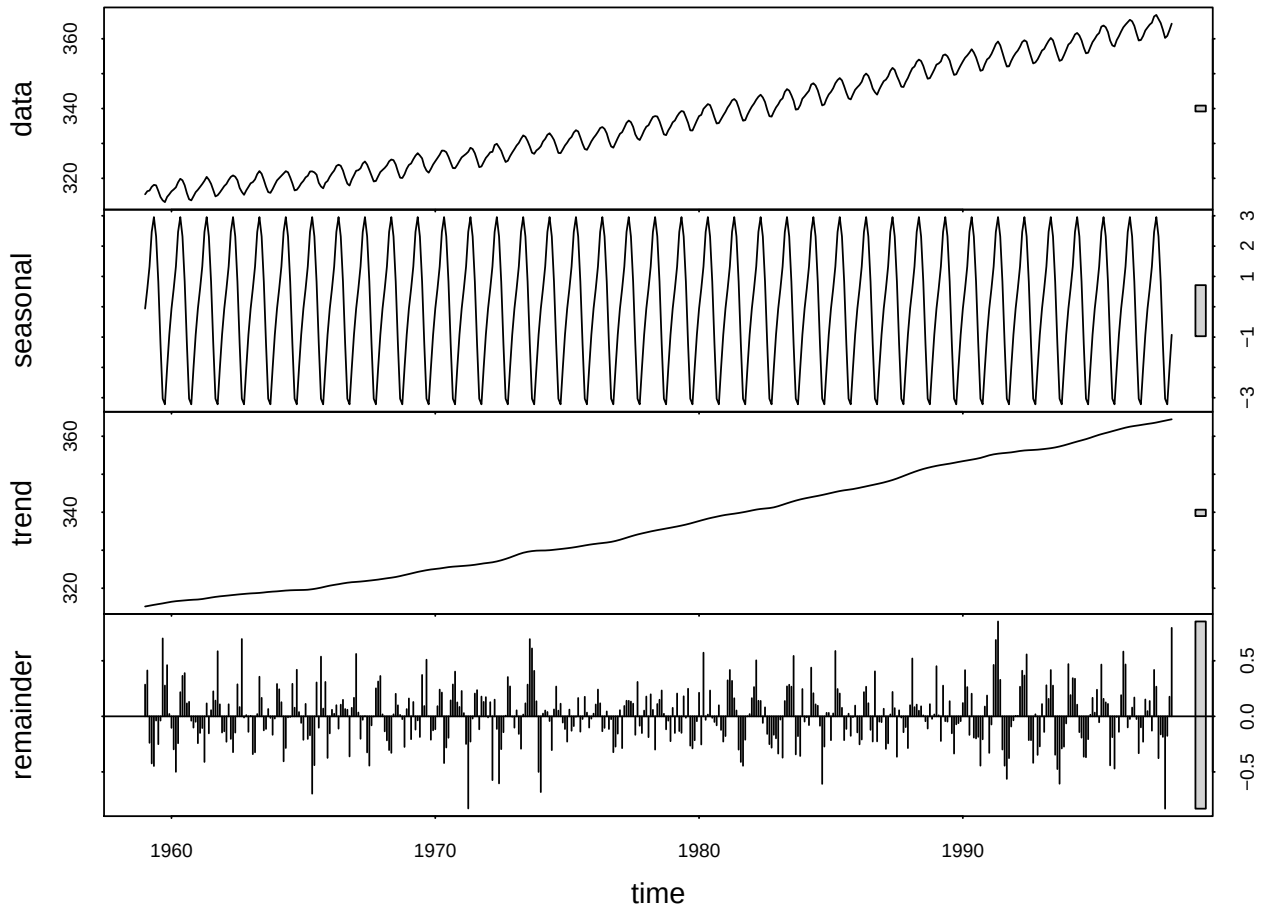
CO_2 Concentration

Atmospheric concentrations of CO2 are expressed in parts per million (ppm) and reported in the preliminary 1997 SIO manometric mole fraction scale.

```
data(co2)

par(mar = c(3.8, 4, 0.8, 0.6))
plot(co2, las = 1, xlab = "", ylab = "")
mtext("Time (year)", side = 1, line = 2)
mtext(expression(paste("CO"[2], " Concentration (ppm)")), side = 2, line = 2.5)
```




```
# Seasonal and Trend decomposition using Loess (STL)
par(mar = c(4, 3.6, 0.8, 0.6))
stl <- stl(co2, s.window = "periodic")
plot(stl, las = 1)
```



U.S. monthly unemployment rates

```
library(quantmod)
```

```
## Loading required package: xts
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Registered S3 method overwritten by 'xts':
##   method     from
##   as.zoo.xts zoo
```

```
## Loading required package: TTR
```

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

```
## Version 0.4-0 included new data defaults. See ?getSymbols.
```

```r
getSymbols("UNRATE", src = "FRED")
```

```
## 'getSymbols' currently uses auto.assign=TRUE by default, but will
## use auto.assign=FALSE in 0.5-0. You will still be able to use
## 'loadSymbols' to automatically load data. getOption("getSymbols.env")
## and getOption("getSymbols.auto.assign") will still be checked for
## alternate defaults.
##
## This message is shown once per session and may be disabled by setting
## options("getSymbols.warning4.0"=FALSE). See ?getSymbols for details.
```

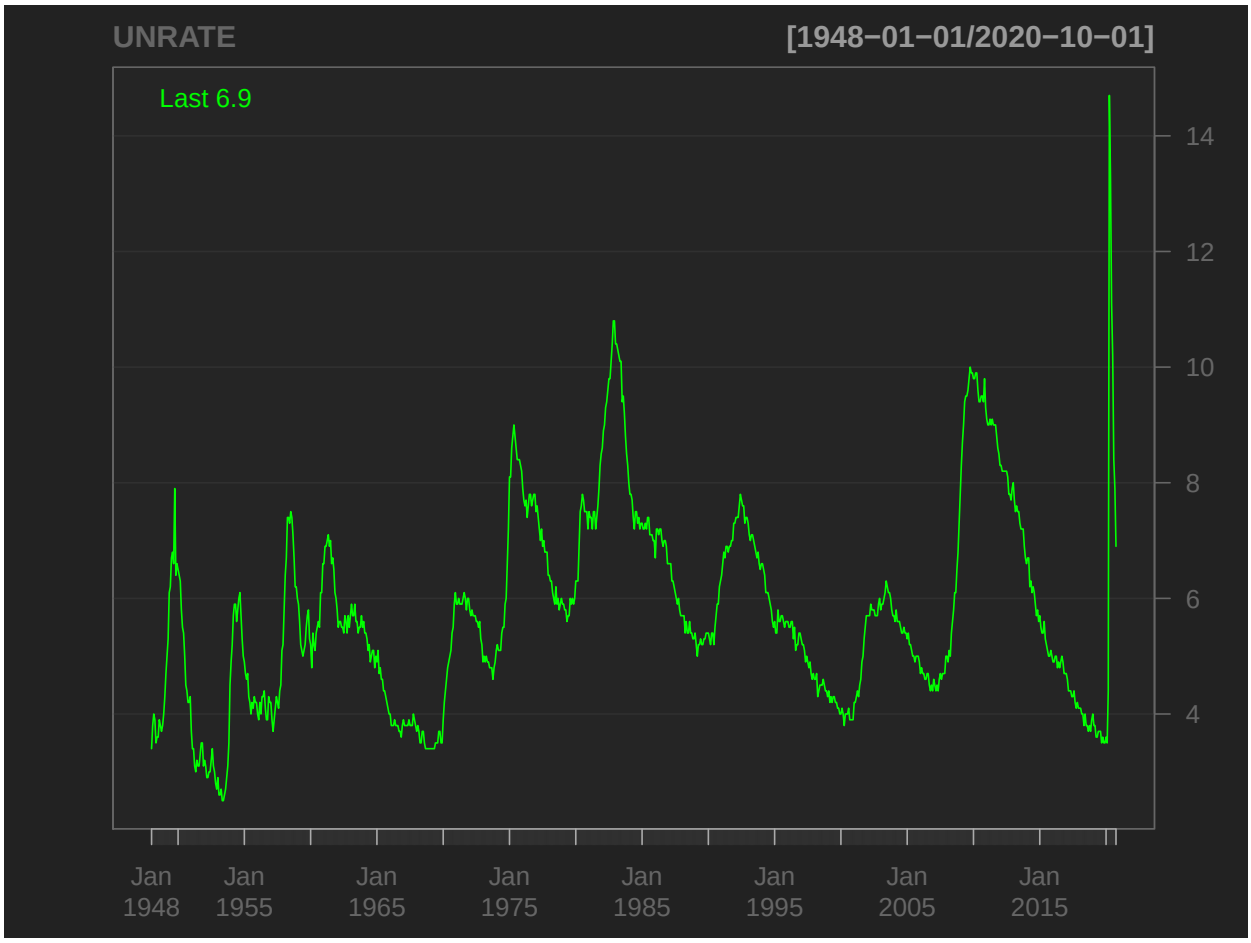
```
## [1] "UNRATE"
```

```r
head(UNRATE); tail(UNRATE)
```

```
##            UNRATE
## 1948-01-01    3.4
## 1948-02-01    3.8
## 1948-03-01    4.0
## 1948-04-01    3.9
## 1948-05-01    3.5
## 1948-06-01    3.6
```

```
##            UNRATE
## 2020-05-01   13.3
## 2020-06-01   11.1
## 2020-07-01   10.2
## 2020-08-01    8.4
## 2020-09-01    7.9
## 2020-10-01    6.9
```

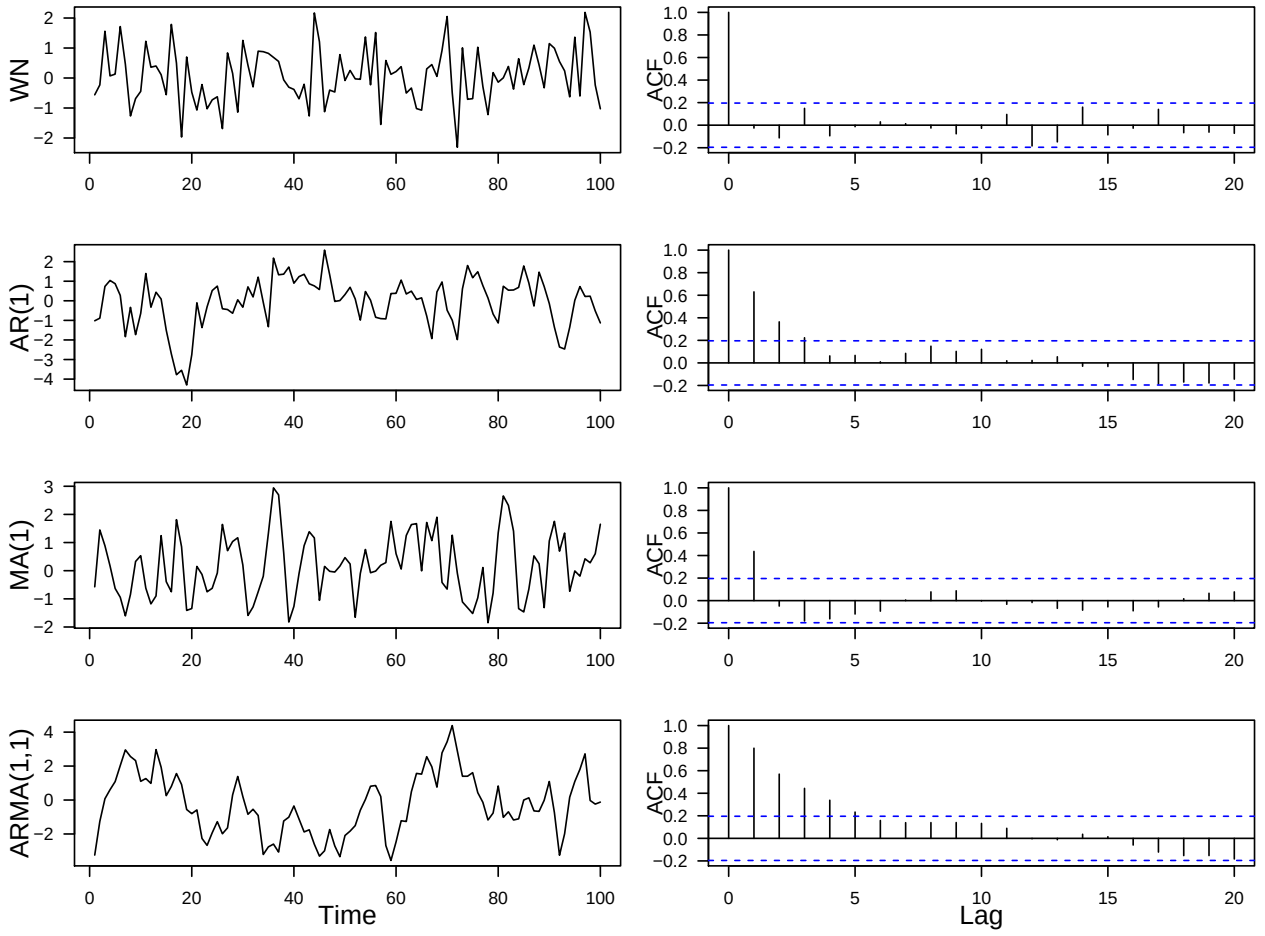
```r
chartSeries(UNRATE)
```



ARMA and ACF

```
set.seed(123)
n = 100
WN <- rnorm(n)

par(mfrow = c(4, 2), mar = c(3.6, 3.6, 0.8, 0.6))
plot(1:n, WN, type = "l", las = 1, xlab = "", ylab = "")
mtext("WN", side = 2, line = 2)
acf(WN, xlab = "", ylab = "", main = "", las = 1)
mtext("ACF", side = 2, line = 2, cex = 0.8)
# AR(1) phi = 0.8
AR <- arima.sim(n = n, model = list(ar = 0.8))
plot(1:n, AR, type = "l", las = 1, xlab = "", ylab = "")
mtext("AR(1)", side = 2, line = 2)
acf(AR, xlab = "", ylab = "", main = "", las = 1)
mtext("ACF", side = 2, line = 2, cex = 0.8)
# MA(1) theta = 0.5
MA <- arima.sim(n = n, model = list(ma = 0.5))
plot(1:n, MA, type = "l", las = 1, xlab = "", ylab = "")
mtext("MA(1)", side = 2, line = 2)
acf(MA, xlab = "", ylab = "", main = "", las = 1)
mtext("ACF", side = 2, line = 2, cex = 0.8)
```

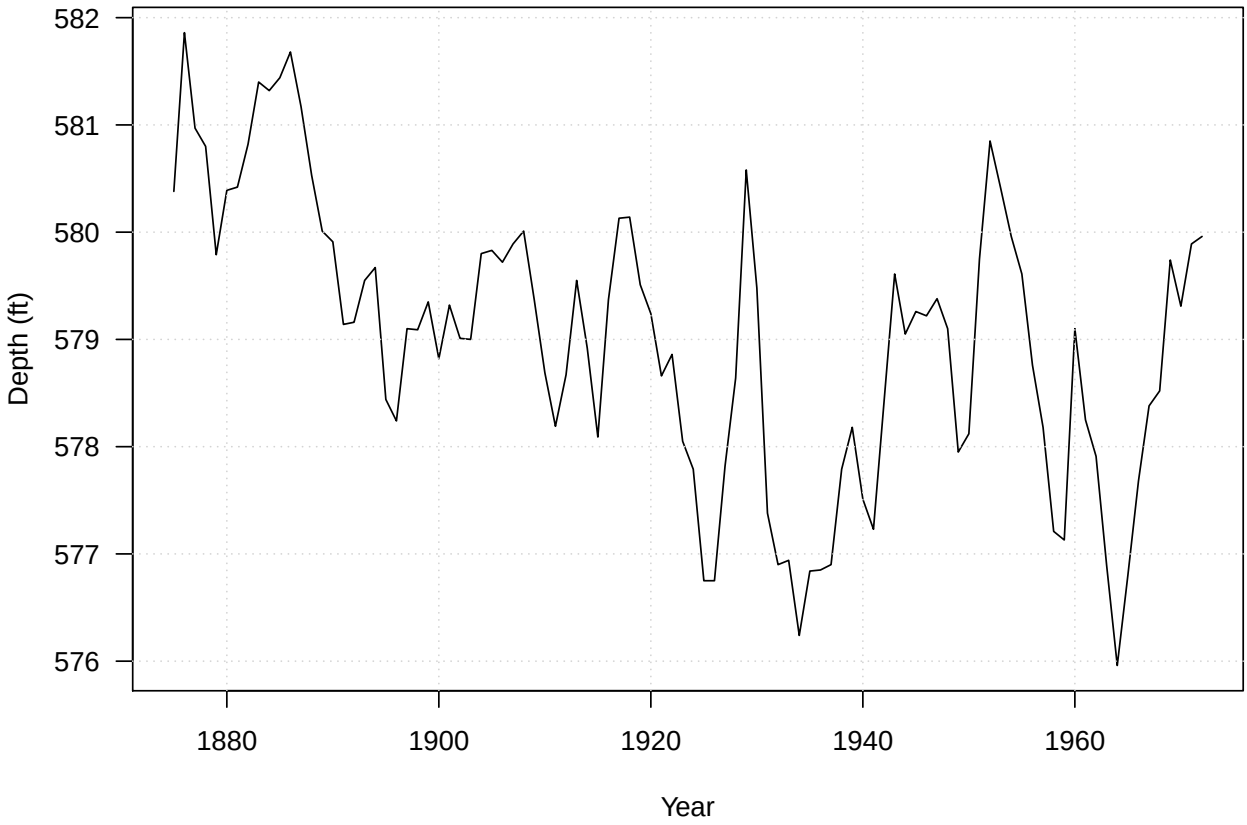
```r
# ARMA(1, 1) phi = 0.8, theta = 0.5
ARMA <- arima.sim(n = n, model = list(ar = 0.8, ma = 0.5))
plot(1:n, ARMA, type = "l", las = 1, xlab = "", ylab = "")
mtext("ARMA(1,1)", side = 2, line = 2)
mtext("Time", side = 1, line = 2)
acf(ARMA, xlab = "", ylab = "", main = "", las = 1)
mtext("ACF", side = 2, line = 2, cex = 0.8)
mtext("Lag", side = 1, line = 2)
```



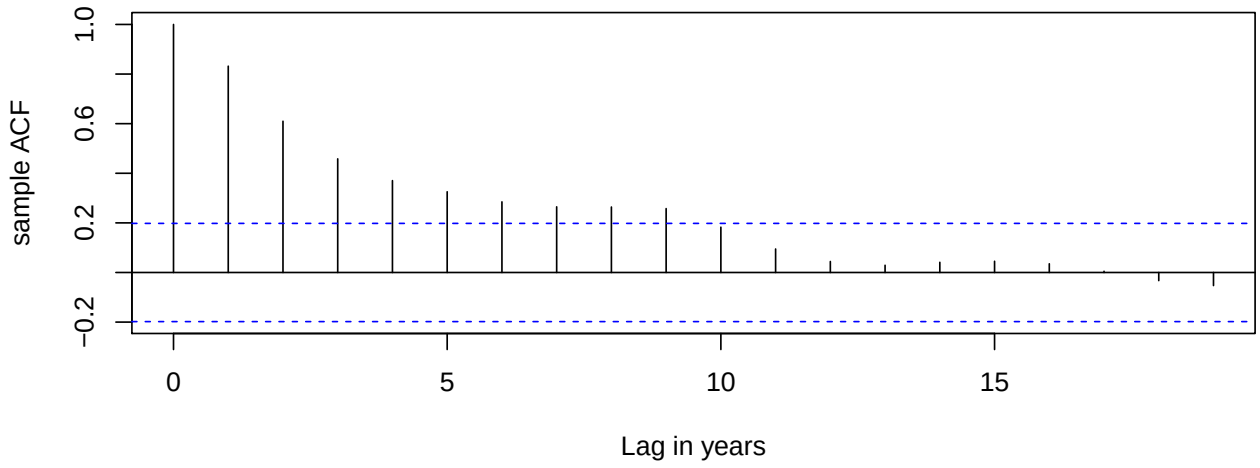
Lake Huron Case Study

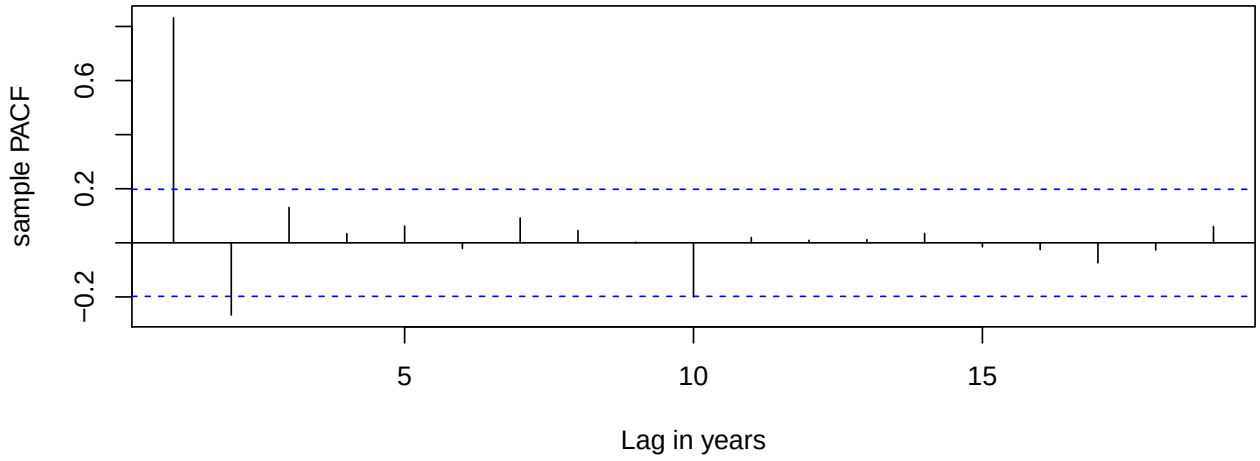
Detrend

```r
## Let us create a 'years' variable.
years <- time(LakeHuron)

plot(LakeHuron, ylab = "Depth (ft)", xlab = "Year", las = 1)
grid()
```

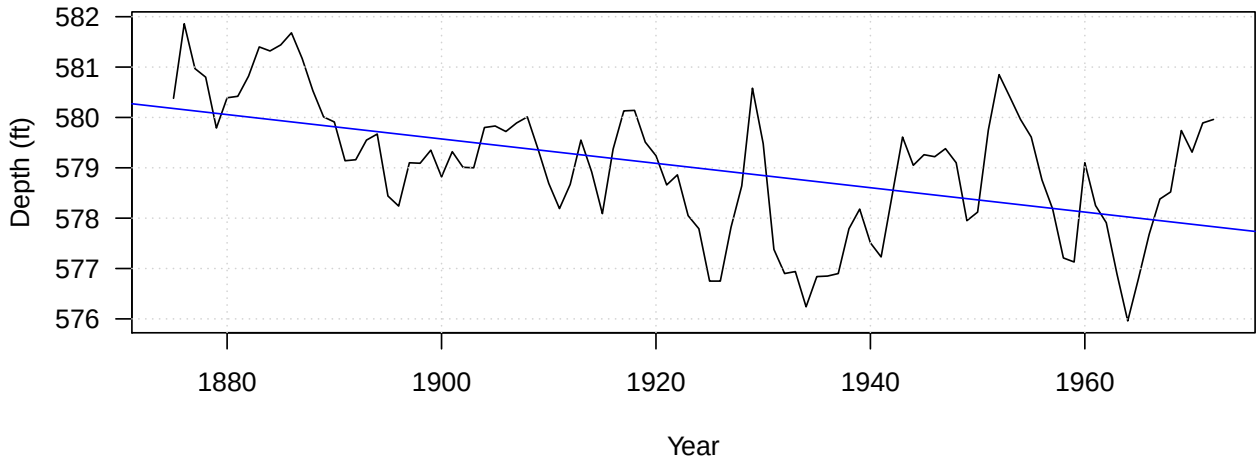



```
par(mfrow = c(2, 1), mar = c(4, 4, 1, 1))
acf(LakeHuron, xlab="Lag in years", ylab = "sample ACF", main = "")
pacf(LakeHuron, xlab="Lag in years", ylab = "sample PACF", main = "")
```

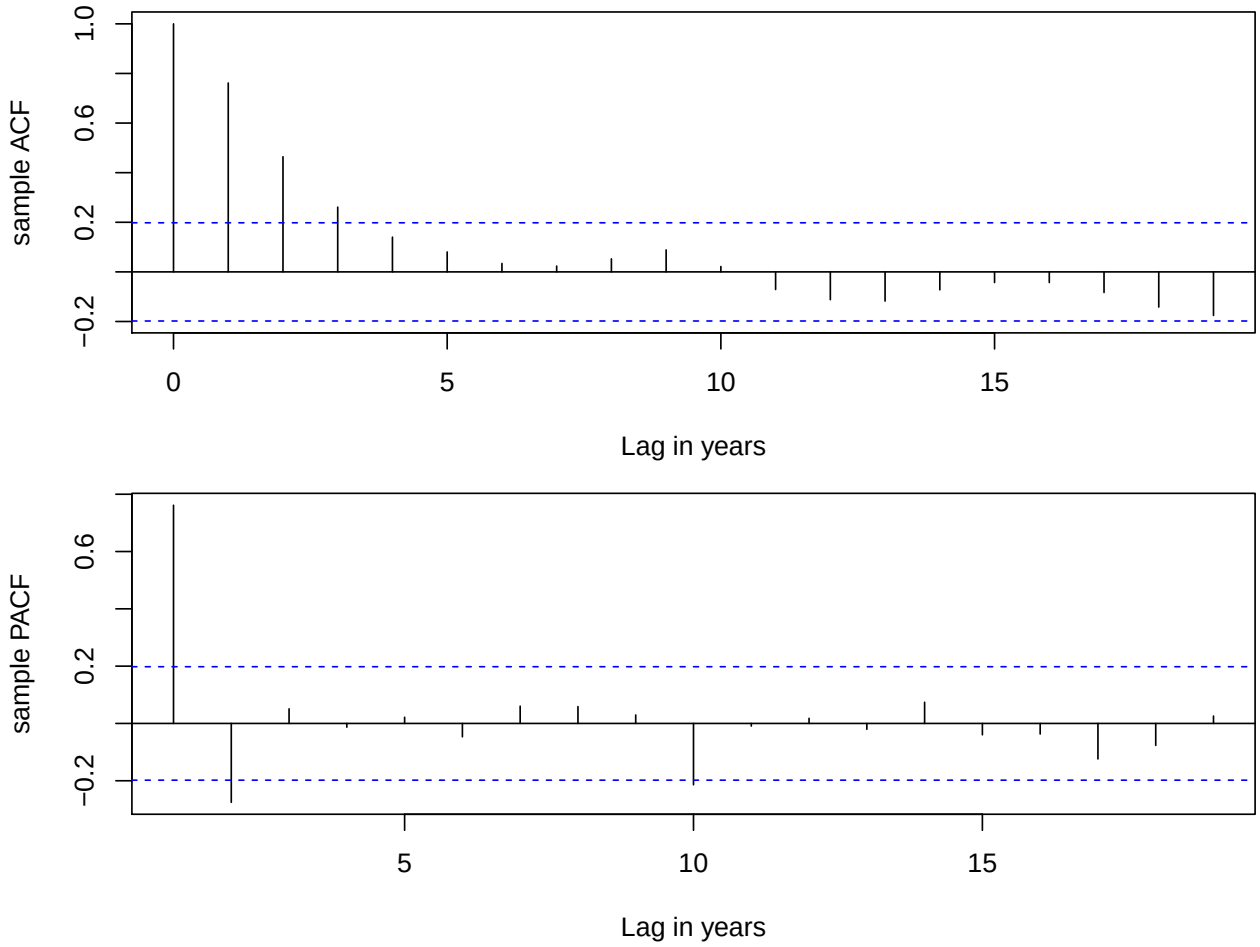




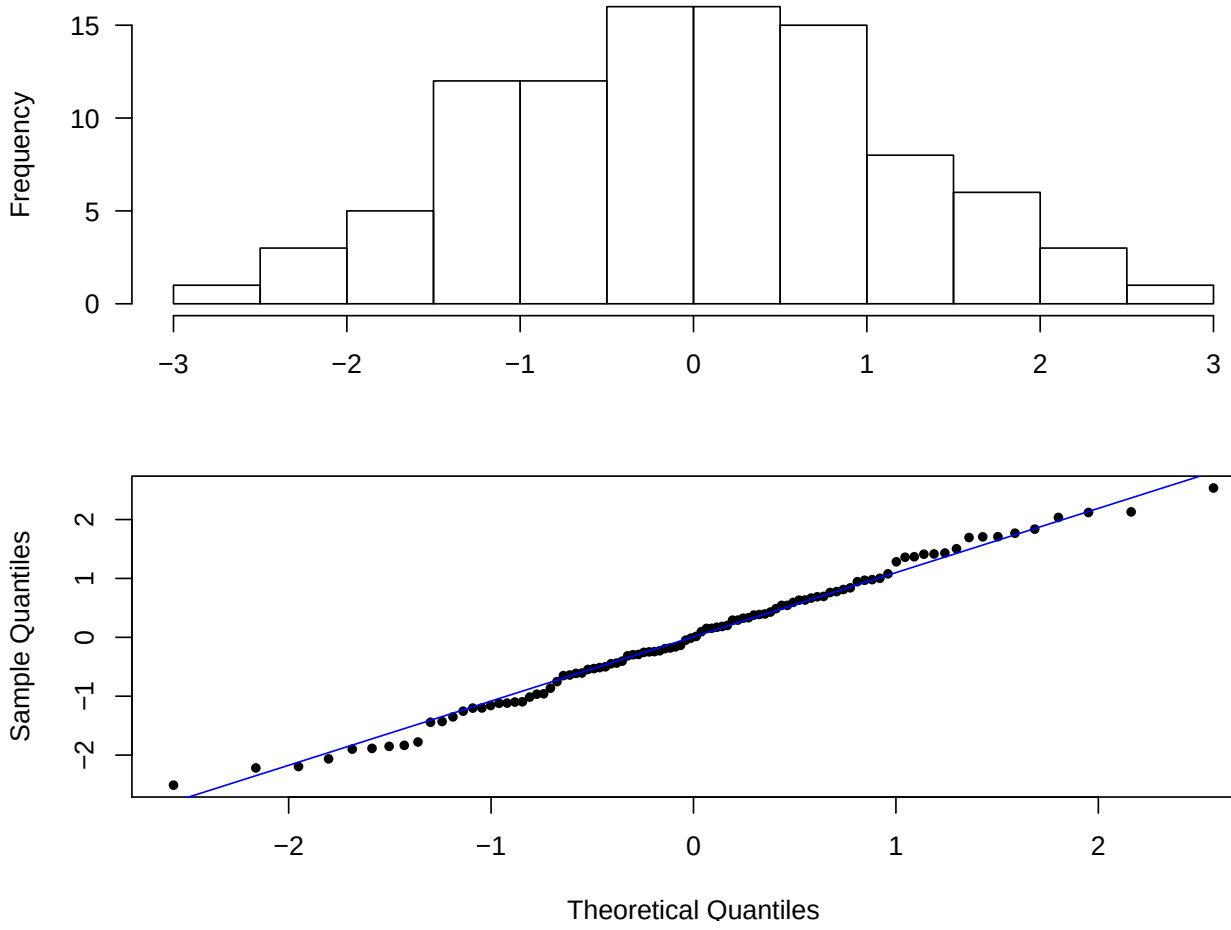

```r
# Estimate the linear trend
lm <- lm(LakeHuron ~ years)
plot(LakeHuron, ylab = "Depth (ft)", xlab = "Year", las = 1); grid()
abline(lm, col = "blue")

deTrend <- resid(lm)

par(mfrow = c(2, 1), mar = c(4, 4, 1, 1))
```



```r
acf(deTrend, xlab="Lag in years", ylab = "sample ACF", main = "")
pacf(deTrend, xlab="Lag in years", ylab = "sample PACF", main = "")
```




```r
hist(deTrend, main = "", xlab = "", las = 1)
qqnorm(deTrend, main = "", pch = 16, cex = 0.8); qqline(deTrend, col = "blue")
```

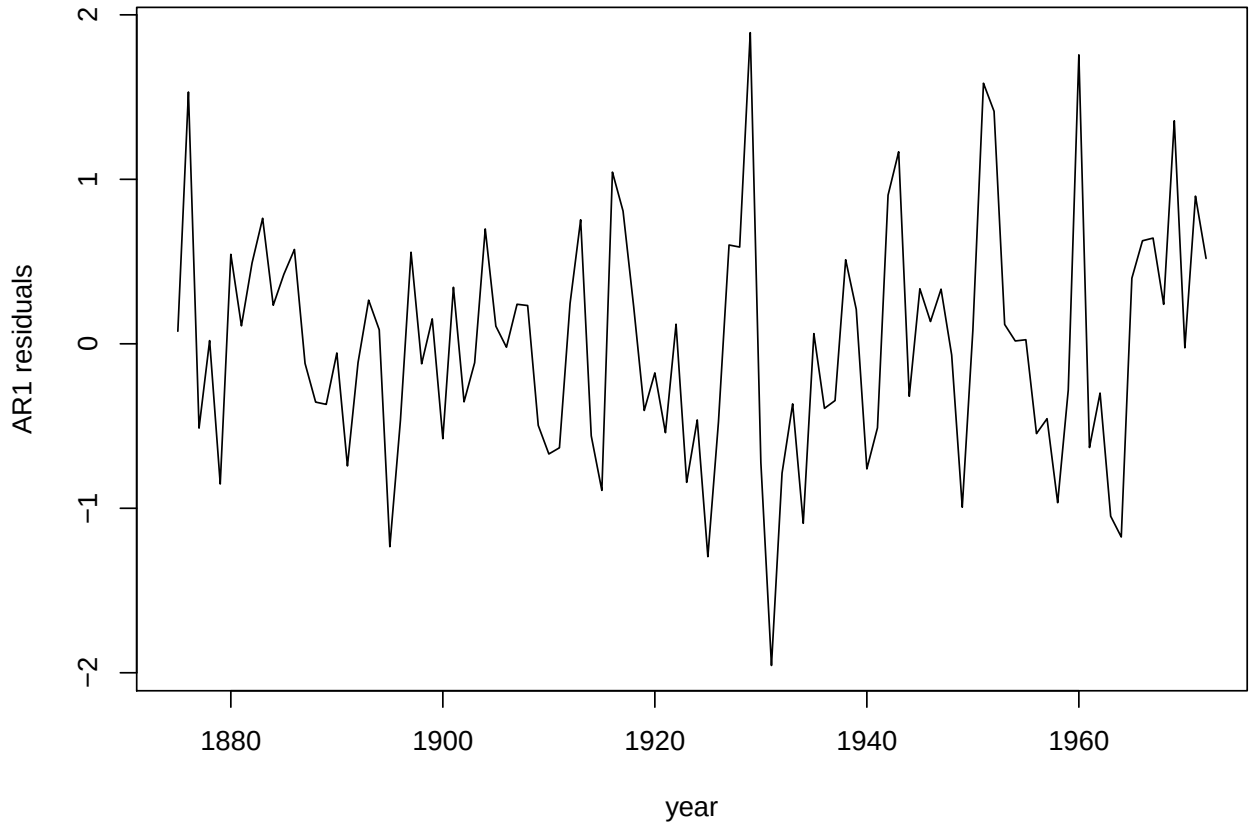


Model Selection/Fitting

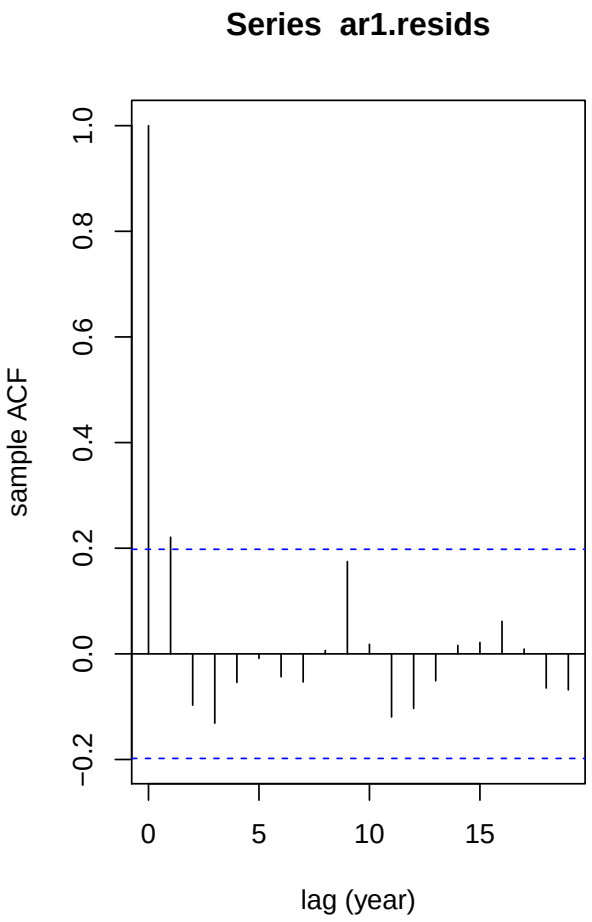
```
## AR(1)
ar1.model <- arima(deTrend, order = c(1, 0, 0))
ar1.model
```

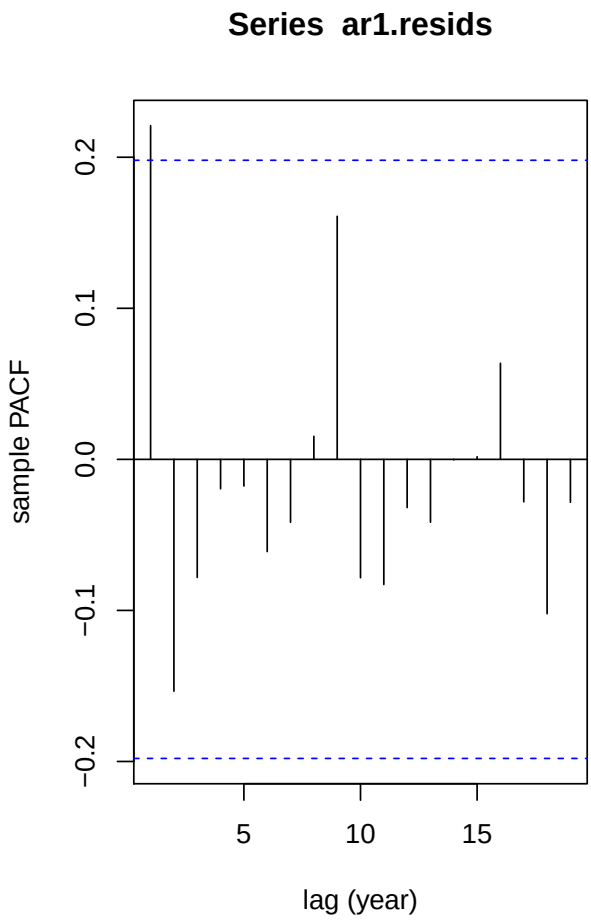
```
##
## Call:
## arima(x = deTrend, order = c(1, 0, 0))
##
## Coefficients:
##          ar1  intercept
##       0.7829     0.0797
## s.e.  0.0634     0.3178
##
## sigma^2 estimated as 0.4972:  log likelihood = -105.29,  aic = 216.58
```

```
ar1.resids <- resid(ar1.model)
plot(1875:1972, ar1.resids, type = "l", xlab = "year", ylab = "AR1 residuals")
```




```
## Sample ACF and PACF of the residuals
par(mfrow = c(1, 2))
acf(ar1.resids, ylab = "sample ACF", xlab = "lag (year)")
pacf(ar1.resids, ylab = "sample PACF", xlab = "lag (year)")
```





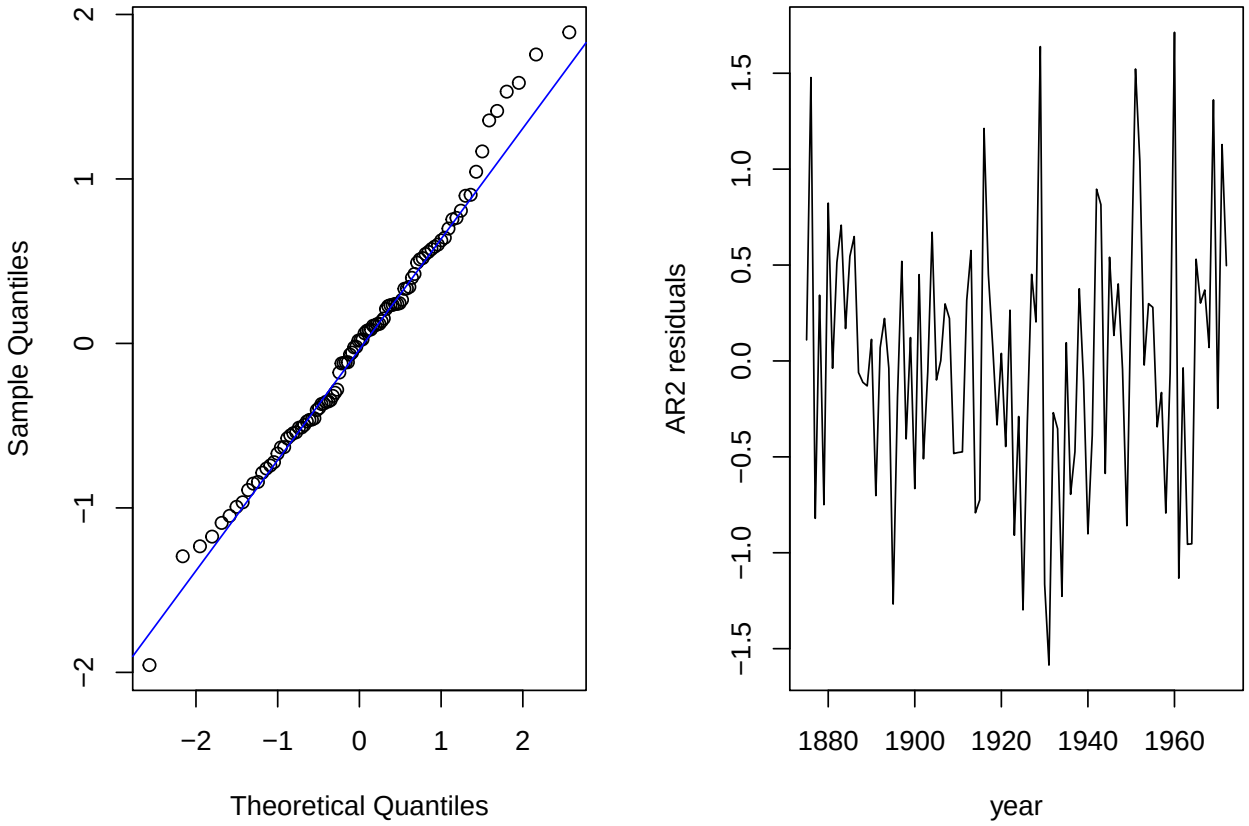

```
## Normal Q-Q plot for the residuals
qqnorm(ar1.resids, main = ""); qqline(ar1.resids, col = "blue")
## Test for time dependence for the residuals
Box.test(ar1.resids, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  ar1.resids
## X-squared = 4.93, df = 1, p-value = 0.02639
```

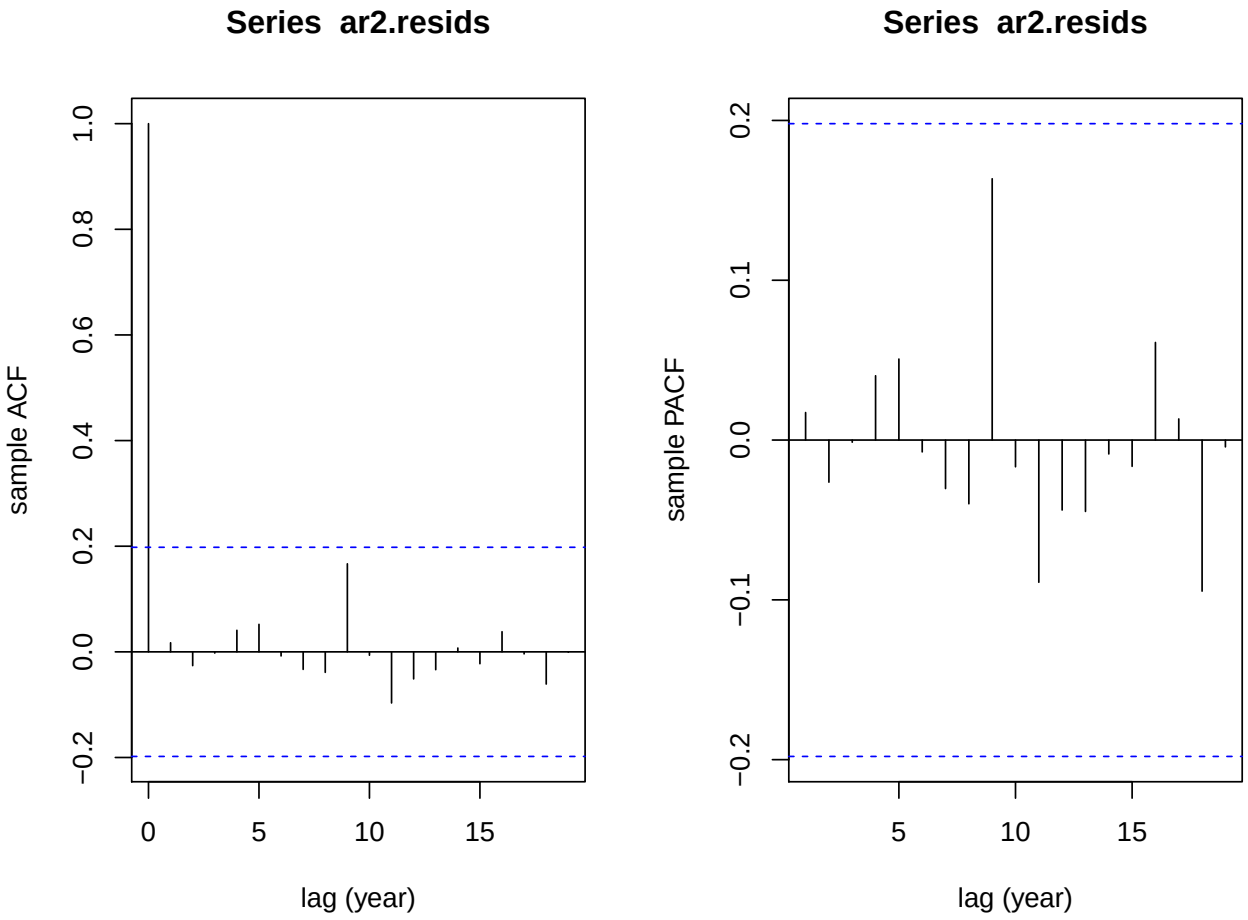
```
## AR(2)
ar2.model <- arima(deTrend, order = c(2, 0, 0))
## summarize the model
ar2.model
```

```
##
## Call:
## arima(x = deTrend, order = c(2, 0, 0))
##
## Coefficients:
##          ar1      ar2  intercept
##       1.0047  -0.2919     0.0196
## s.e.  0.0977   0.1004     0.2351
##
## sigma^2 estimated as 0.4571:  log likelihood = -101.25,  aic = 210.5
```

```r
## calculate the residuals
ar2.resids <- resid(ar2.model)
## time series plot of the residuals
plot(1875:1972, ar2.resids, type = "l", xlab = "year", ylab = "AR2 residuals")
```




```r
## Sample ACF and PACF of the residuals
par(mfrow=c(1,2))
acf(ar2.resids,  ylab = "sample ACF", xlab = "lag (year)")
pacf(ar2.resids,  ylab = "sample PACF", xlab = "lag (year)")
```




```
## Normal Q-Q plot for the residuals
qqnorm(ar2.resids, main = ""); qqline(ar2.resids, col = "blue")

## Test for time dependence for the residuals
Box.test(ar2.resids, type = "Ljung-Box")
```

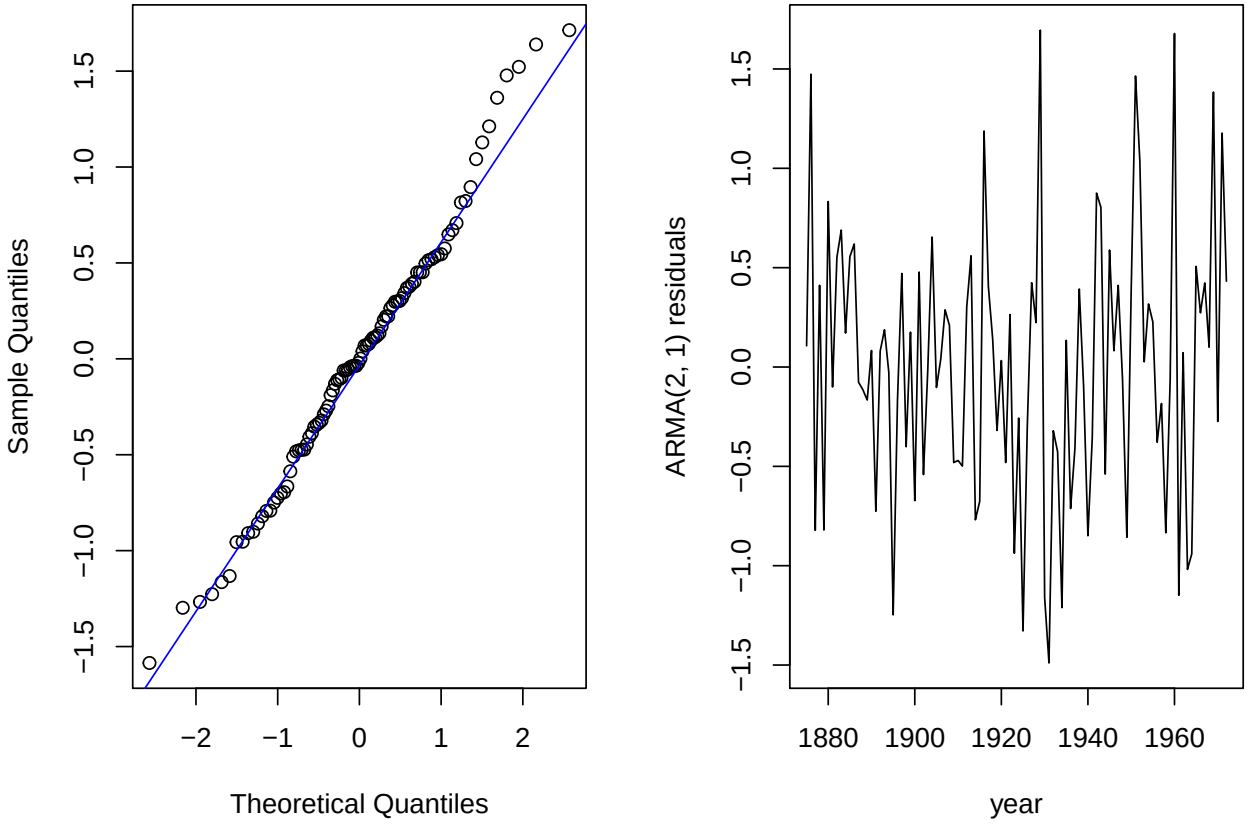
```
##
##  Box-Ljung test
##
## data:  ar2.resids
## X-squared = 0.029966, df = 1, p-value = 0.8626
```

```
## Fit the ARMA(2, 1) model
arma21.model <- arima(deTrend, order = c(2, 0, 1))
## summarize the model
arma21.model
```

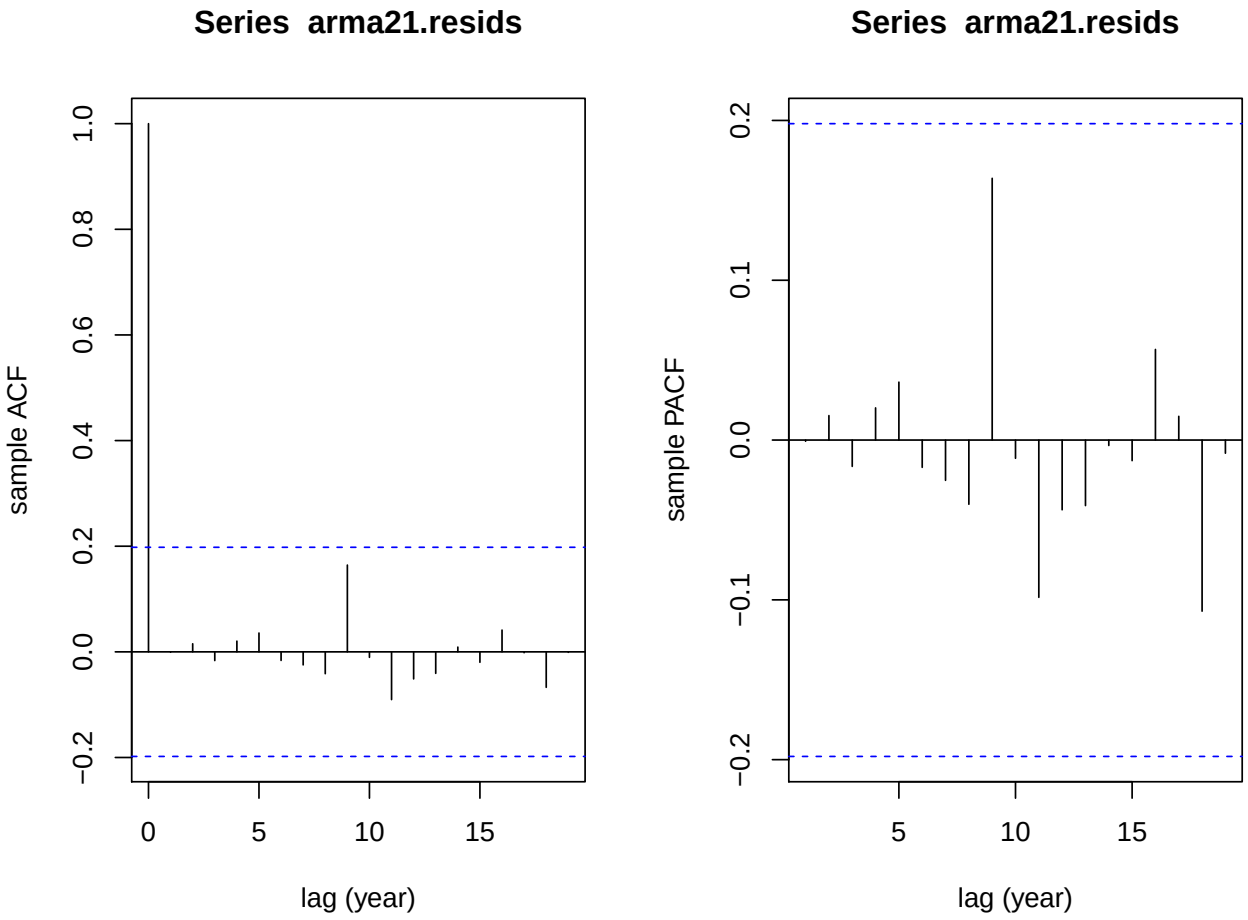
```
##
## Call:
## arima(x = deTrend, order = c(2, 0, 1))
##
## Coefficients:
##          ar1      ar2     ma1  intercept
##       0.8374  -0.1622  0.1846     0.0245
## s.e.  0.3180   0.2621  0.3180     0.2452
##
```

```
## sigma^2 estimated as 0.4556:  log likelihood = -101.09,  aic = 212.18
```

```
## calculate the residuals
arma21.resids <- resid(arma21.model)
## time series plot of the residuals
plot(1875:1972, arma21.resids, type = "l", xlab = "year", ylab = "ARMA(2, 1) residuals")
```




```
## Sample ACF and PACF of the residuals
par(mfrow=c(1,2))
acf(arma21.resids,  ylab = "sample ACF", xlab = "lag (year)")
pacf(arma21.resids,  ylab = "sample PACF", xlab = "lag (year)")
```




```
## Normal Q-Q plot for the residuals
qqnorm(arma21.resids, main = ""); qqline(arma21.resids, col = "blue")
## Test
Box.test(arma21.resids, type = "Ljung-Box")
```

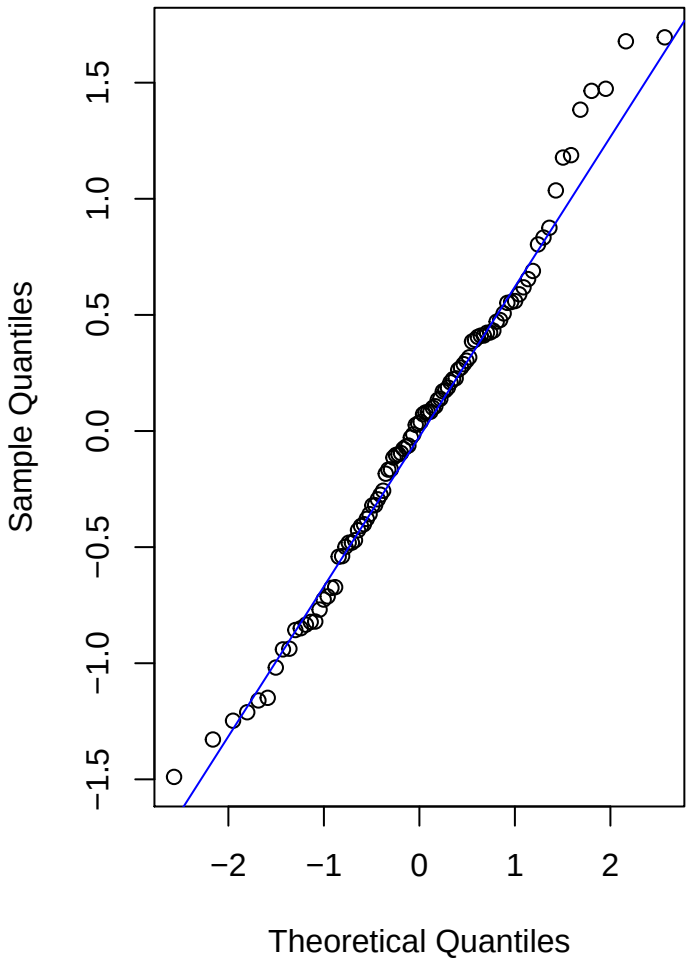
```
##
##  Box-Ljung test
##
## data:  arma21.resids
## X-squared = 5.5105e-05, df = 1, p-value = 0.9941
```

```
arma21.model <- arima(deTrend, order=c(2, 0, 1))

AIC(ar1.model); AIC(ar2.model); AIC(arma21.model)
```

```
## [1] 216.5835
```

```
## [1] 210.5032
```

```
## [1] 212.1784
```



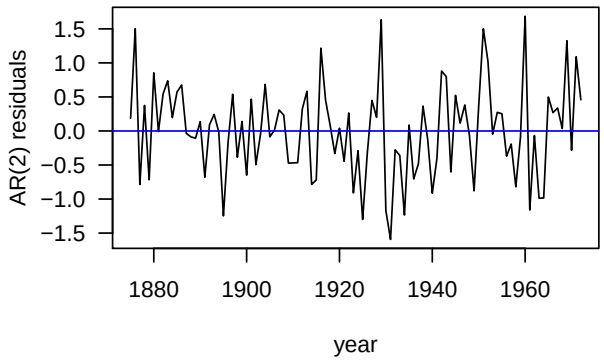
AR(2) Fitting

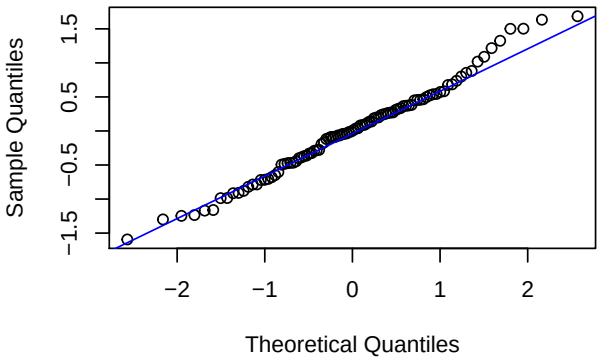
```
## Fit the full regression model plus AR(2) errors
full.model <- arima(LakeHuron, order = c(2, 0, 0),
                    xreg = cbind(rep(1, length(LakeHuron)), years), include.mean = FALSE)
full.model
```

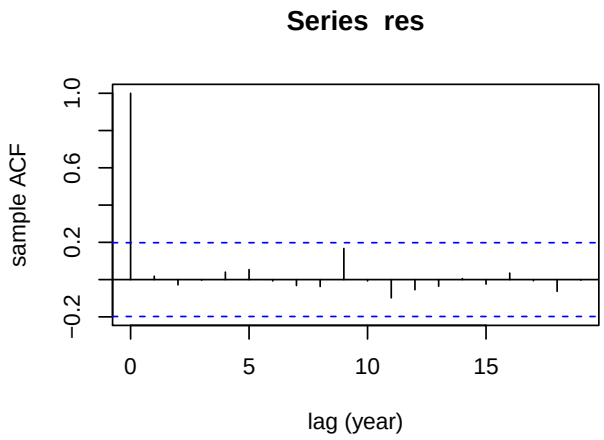
```
##
## Call:
## arima(x = LakeHuron, order = c(2, 0, 0), xreg = cbind(rep(1, length(LakeHuron)),
##     years), include.mean = FALSE)
##
## Coefficients:
##          ar1      ar2  rep(1, length(LakeHuron))    years
##       1.0048  -0.2913                   620.5115  -0.0216
## s.e.  0.0976   0.1004                    15.5771   0.0081
##
## sigma^2 estimated as 0.4566:  log likelihood = -101.2,  aic = 212.4
```

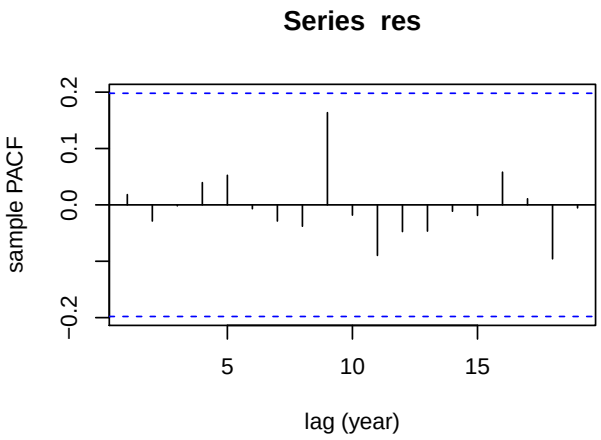
```
## Examine the residuals of the model
par(mfrow = c(2, 2))
res <- full.model$residuals
```

```r
plot(res, type = "l", xlab = "year", ylab = "AR(2) residuals", las = 1)
abline(h = 0, col = "blue")
qqnorm(res, main = ""); qqline(res, col = "blue")
acf(res,  ylab = "sample ACF", xlab = "lag (year)")
pacf(res,  ylab = "sample PACF", xlab = "lag (year)")
```









Forecasting

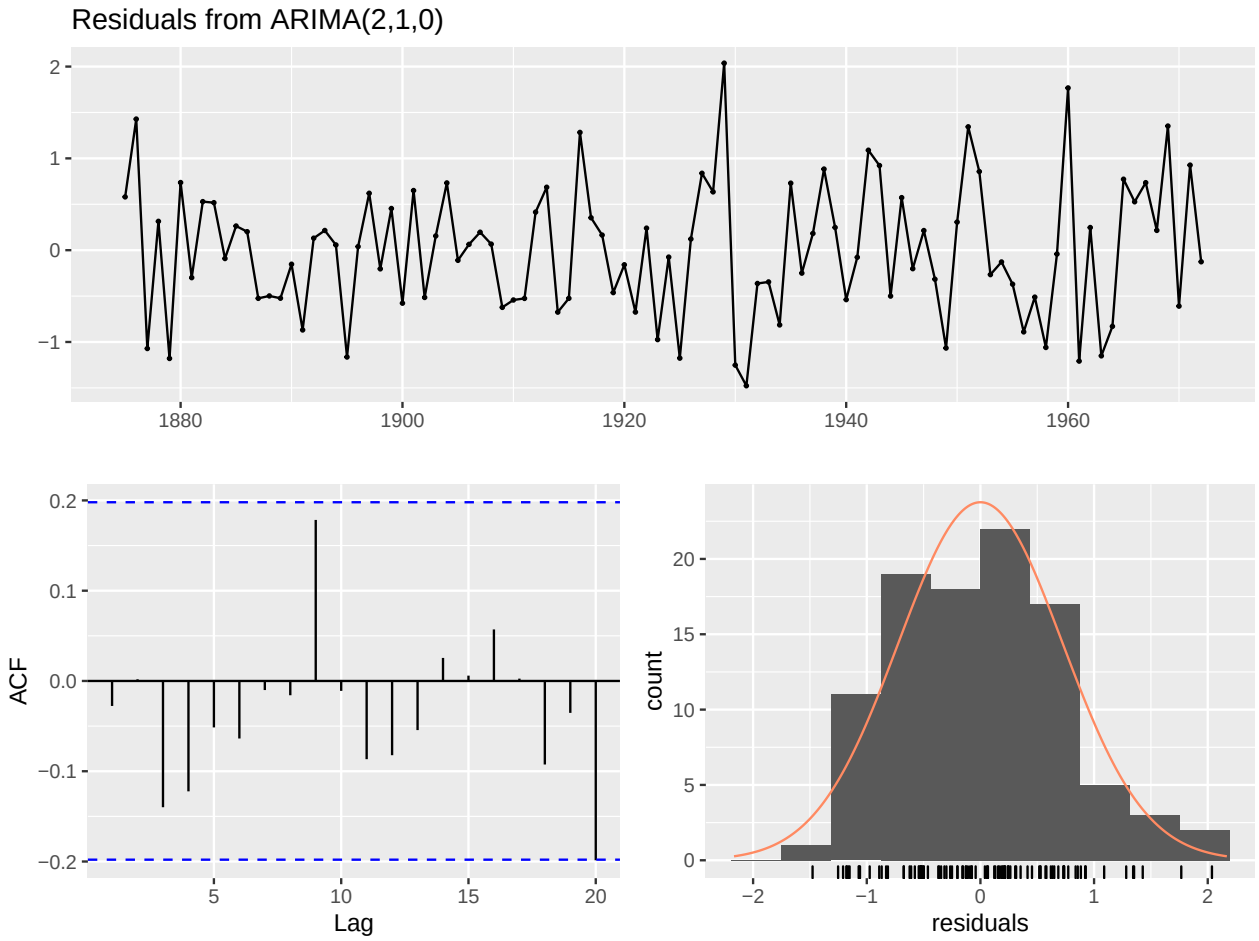
```r
library(forecast)
```

```
## Warning: package 'forecast' was built under R version 3.6.2
```

```r
(fit <- Arima(LakeHuron, order = c(2, 1, 0)))
```

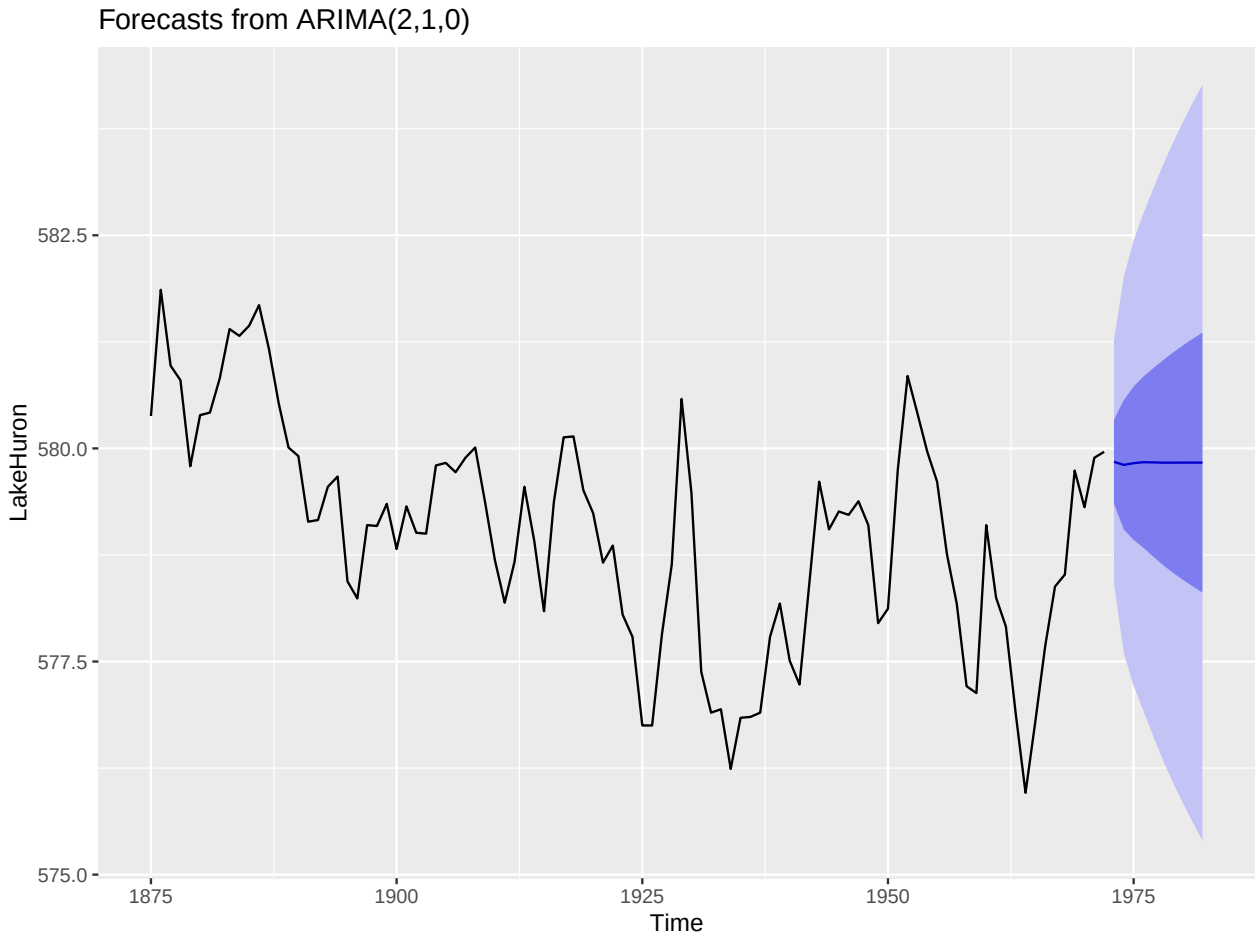
```
## Series: LakeHuron
## ARIMA(2,1,0)
##
## Coefficients:
##          ar1      ar2
##       0.1728  -0.2233
## s.e.  0.1012   0.1015
##
## sigma^2 estimated as 0.5333:  log likelihood=-105.87
## AIC=217.74   AICc=218   BIC=225.47
```

```r
checkresiduals(fit)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,1,0)
## Q* = 7.9226, df = 8, p-value = 0.4411
##
## Model df: 2.   Total lags used: 10
```

```
autoplot(forecast(fit, level = c(50, 95)))
```


Forecasts from ARIMA(2,1,0)
LakeHuron
582.5
580.0
577.5
575.0
1875
1900
1925
1950
1975
Time