MATH 8090: Regression with Time Series Errors

Whitney Huang, Clemson University

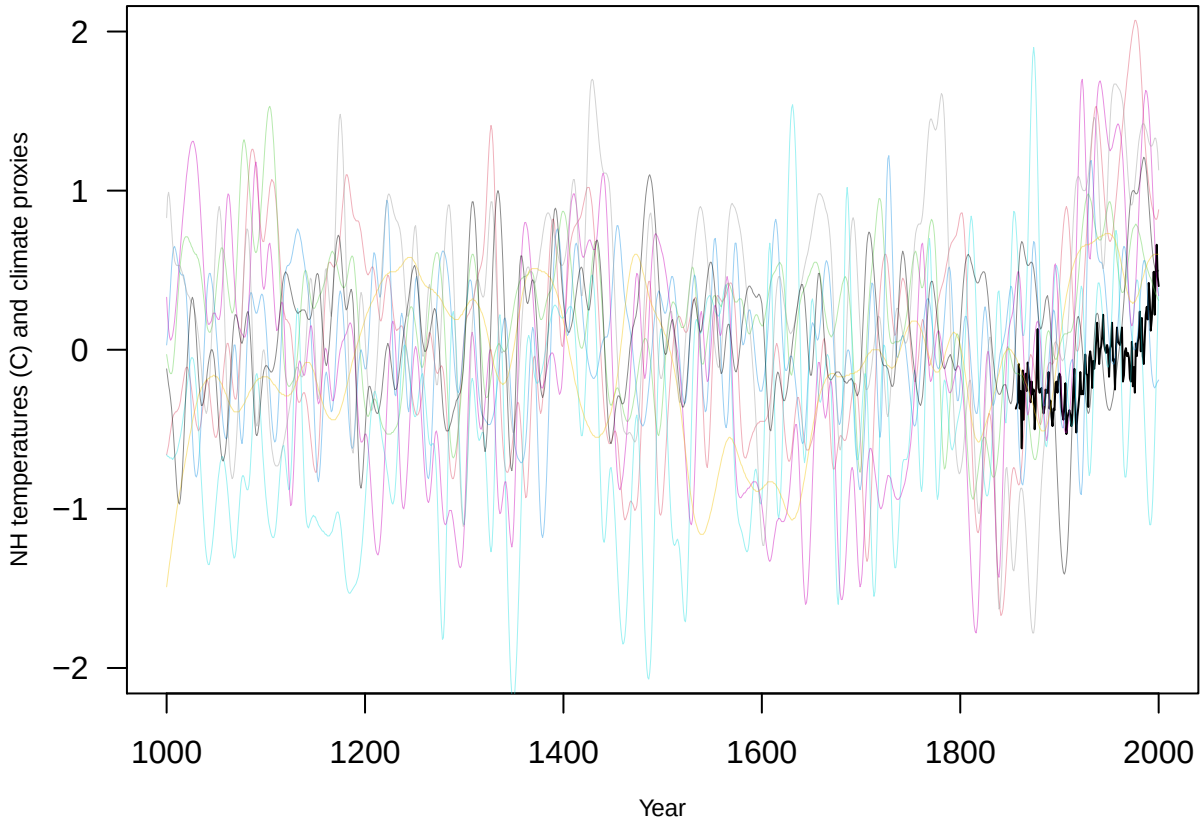
10/14/2021

Contents

Northern Hemisphere temperatures and climate proxies Jones and Mann (2004)

Plot the data

```
library(faraway)
data(globwarm)
par(las = 1, mgp = c(2.4, 1, 0), mar = c(3.6, 4, 1, 0.6))
plot(globwarm$year, globwarm$nhtemp, type = "l", ylim = c(-2, 2),
     xlab = "Year", ylab = "NH temperatures (C) and climate proxies",
     cex.lab = 0.75)
library(scales)
for (i in 2:9) lines(globwarm$year, globwarm[, i], col = alpha(i, 0.45),
                     lwd = 0.5)
```



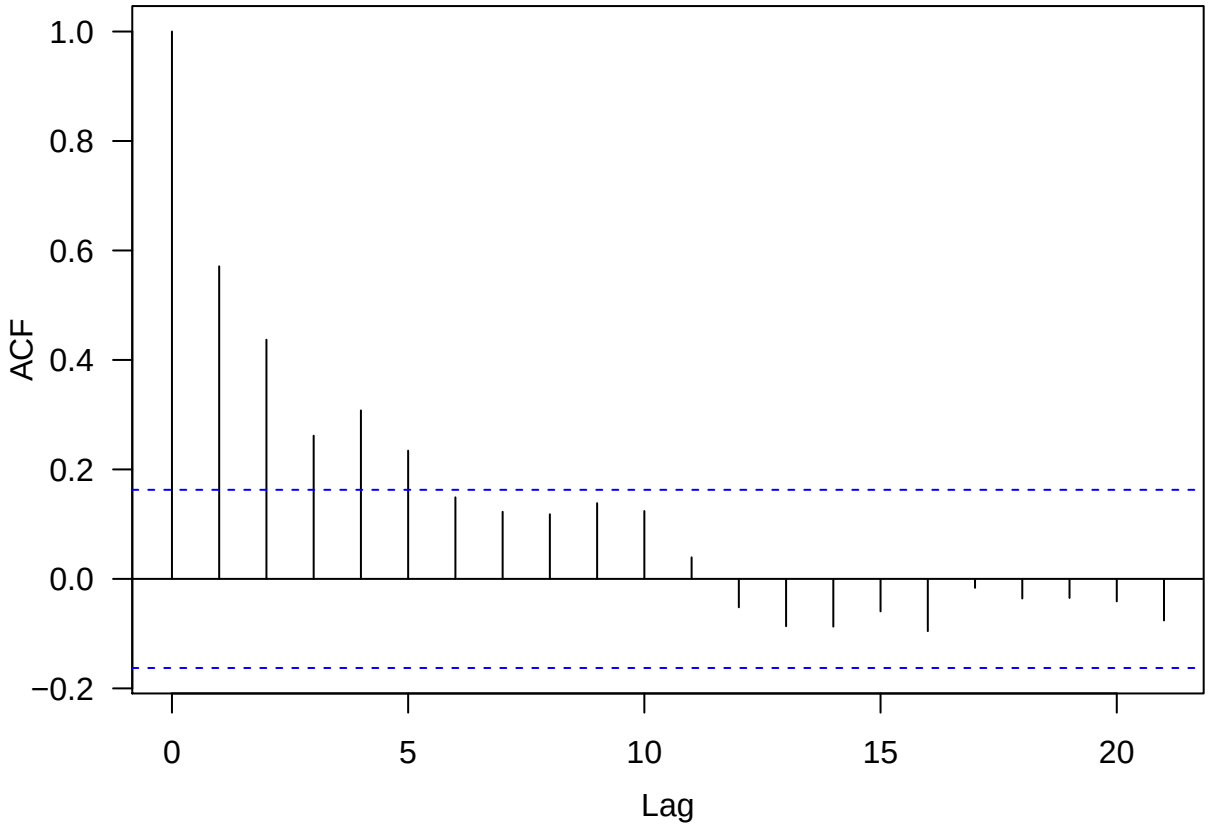
Fit an OLS an examine the residuals

```
lmod <- lm(nhtemp ~ wusa + jasper + westgreen + chesapeake
           + tornetrask + urals
           + mongolia + tasman, globwarm)
summary(lmod)
```

```
##
## Call:
## lm(formula = nhtemp ~ wusa + jasper + westgreen + chesapeake +
##     tornetrask + urals + mongolia + tasman, data = globwarm)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.43668 -0.11170  0.00698  0.10176  0.65352
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.242555   0.027011  -8.980 1.97e-15 ***
## wusa         0.077384   0.042927   1.803 0.073647 .
## jasper      -0.228795   0.078107  -2.929 0.003986 **
## westgreen    0.009584   0.041840   0.229 0.819168
## chesapeake  -0.032112   0.034052  -0.943 0.347346
## tornetrask   0.092668   0.045053   2.057 0.041611 *
## urals        0.185369   0.091428   2.027 0.044567 *
```

```
## mongolia     0.041973   0.045794   0.917 0.360996
## tasman       0.115453   0.030111   3.834 0.000192 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1758 on 136 degrees of freedom
##   (856 observations deleted due to missingness)
## Multiple R-squared:  0.4764, Adjusted R-squared:  0.4456
## F-statistic: 15.47 on 8 and 136 DF,  p-value: 5.028e-16
```

```
par(las = 1, mgp = c(2.4, 1, 0), mar = c(3.6, 4, 1, 0.6))
acf(lmod$residuals)
```



Fit a GLS with an AR(1) error structure

```
library(nlme)
glmod <- gls(nhtemp ~ wusa + jasper + westgreen + chesapeake + tornetrask + urals +
               mongolia + tasman, correlation = corAR1(form = ~ year),
             data = na.omit(globwarm))
summary(glmod)
```

```
## Generalized least squares fit by REML
##   Model: nhtemp ~ wusa + jasper + westgreen + chesapeake + tornetrask +      urals + mongolia + tasma
##   Data: na.omit(globwarm)
##         AIC       BIC   logLik
```

```
##   -108.2074 -76.16822 65.10371
##
## Correlation Structure: AR(1)
##  Formula: ~year
##  Parameter estimate(s):
##       Phi
## 0.7109922
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept) -0.23010624 0.06702406 -3.433188  0.0008
## wusa         0.06673819 0.09877211  0.675678  0.5004
## jasper      -0.20244335 0.18802773 -1.076668  0.2835
## westgreen   -0.00440299 0.08985321 -0.049002  0.9610
## chesapeake  -0.00735289 0.07349791 -0.100042  0.9205
## tornetrask   0.03835169 0.09482515  0.404446  0.6865
## urals        0.24142199 0.22871028  1.055580  0.2930
## mongolia     0.05694978 0.10489786  0.542907  0.5881
## tasman       0.12034918 0.07456983  1.613913  0.1089
##
##  Correlation:
##             (Intr) wusa   jasper wstgrn chespk trntrs urals  mongol
## wusa        -0.517
## jasper      -0.058 -0.299
## westgreen    0.330 -0.533  0.121
## chesapeake   0.090 -0.314  0.230  0.147
## tornetrask -0.430  0.499 -0.197 -0.328 -0.441
## urals       -0.110 -0.142 -0.265  0.075 -0.064 -0.346
## mongolia     0.459 -0.437 -0.205  0.217  0.449 -0.343 -0.371
## tasman       0.037 -0.322  0.065  0.134  0.116 -0.434  0.416 -0.017
##
## Standardized residuals:
##         Min          Q1         Med          Q3         Max
## -2.31122523 -0.53484054  0.02342908  0.50015642  2.97224724
##
## Residual standard error: 0.204572
## Degrees of freedom: 145 total; 136 residual
```

```
intervals(glmod, which = "var-cov")
```

```
## Approximate 95% confidence intervals
##
##  Correlation structure:
##         lower      est.     upper
## Phi 0.5099744 0.7109922 0.8383752
## attr(,"label")
## [1] "Correlation structure:"
##
##  Residual standard error:
##     lower      est.     upper
## 0.1540709 0.2045720 0.2716263
```

Lake Huron Example

Ploting function by Peter F. Craigmile at OSU

```
plot.residuals <- function (x, y = NULL, lag.max = NULL,
                            mean.line = TRUE, acf.ylim = c(-0.25, 1),
                            mfrow = c(2, 2), lags = NULL, ...){
  if (!is.null(mfrow))
    par(mfrow = mfrow)
  if (is.null(y)){
    y <- as.numeric(x); x <- seq(length(y))
  } else {
    x <- as.numeric(x); y <- as.numeric(y)
  }
  if (is.null(lag.max)) {
    lag.max <-  floor(10 * log10(length(x)))
  }
  plot(x, y, type = "l", ...)
  if (mean.line) abline(h = 0, lty = 2)
  qqnorm(y, main = ""); qqline(y)
  if (is.null(lags)) {
    acf(y, main = "", lag.max = lag.max, xlim = c(0, lag.max),
        ylim = acf.ylim, ylab = "sample ACF")
    pacf(y, main = "", lag.max = lag.max, xlim = c(0, lag.max),
         ylim = acf.ylim, ylab = "sample PACF")
  }
  else {
    acf(y, main = "", lag.max = lag.max, xlim = c(0, lag.max),
        ylim = acf.ylim, ylab = "sample ACF", xaxt = "n")
    axis(side = 1, at = lags)
    pacf(y, main = "", lag.max = lag.max, xlim = c(0, lag.max),
         ylim = acf.ylim, ylab = "sample PACF", xaxt = "n")
    axis(side = 1, at = lags)
  }
  Box.test(y, lag.max, type = "Ljung-Box")
}
```

A two-step fit

```
data(LakeHuron)
years <- time(LakeHuron)
lm <- lm(LakeHuron ~ years)
lm$coefficients
```

```
##  (Intercept)        years
## 625.55491791  -0.02420111
```

```
(MLE_est1 <- arima(lm$residuals, order = c(2, 0, 0),
                   include.mean = FALSE))
```

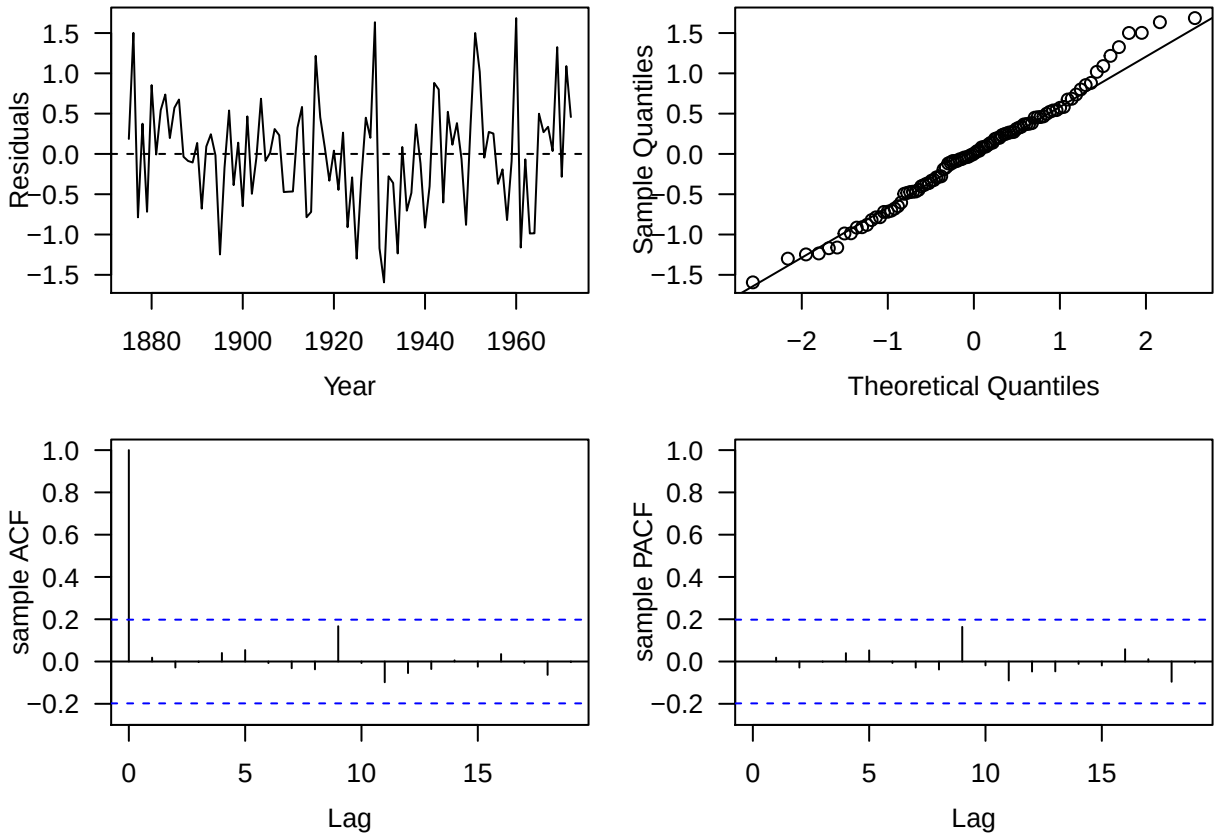
```
##
## Call:
## arima(x = lm$residuals, order = c(2, 0, 0), include.mean = FALSE)
##
## Coefficients:
##          ar1      ar2
##       1.0050  -0.2925
## s.e.  0.0976   0.1002
##
## sigma^2 estimated as 0.4572:  log likelihood = -101.26,  aic = 208.51
```

One-step MLE fit

```
mle <- arima(LakeHuron, order = c(2, 0, 0),
             xreg = cbind(rep(1,length(LakeHuron)), years),
             include.mean = FALSE)
mle
```

```
##
## 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
```

```
par(las = 1, mgp = c(2.4, 1, 0), mar = c(3.6, 4, 1, 0.6), mfrow = c(2, 2))
plot.residuals(years, resid(mle), xlab = "Year", ylab = "Residuals")
```




```
##
##  Box-Ljung test
##
## data:  y
## X-squared = 6.2088, df = 19, p-value = 0.9974
```

Comparing CIs

```
confint(lm)
```

```
##                    2.5 %      97.5 %
## (Intercept) 610.14291793 640.9669179
## years        -0.03221272  -0.0161895
```

```
confint(MLE_est1)
```

```
##          2.5 %      97.5 %
## ar1  0.8137180  1.19630830
## ar2 -0.4888881 -0.09606208
```

```
confint(mle)
```

```
##                                 2.5 %        97.5 %
```

```
## ar1                          0.81348340   1.196124084
## ar2                         -0.48806617  -0.094573470
## rep(1, length(LakeHuron)) 589.98093574 651.042054268
## years                       -0.03744268  -0.005694972
```

References

Jones, Philip D, and Michael E Mann. 2004. "Climate over Past Millennia." *Reviews of Geophysics* 42 (2).