Model Appropriateness

Introduction

Check for Model Appropriateness

Purpose of this unit:

- In previous units we have learned about a variety of useful time series models, how to fit these models to data, etc.
- In this unit, we will step back and take a closer look at the model(s) we have fit to a set of data to ascertain whether the models appear to be ***appropriate***
  - And are not simply the best we've been able to do so far
- Things to consider include
  - Whether basic assumptions of the models are satisfied
  - Ramifications of selecting certain models
  - Forecasting performance (we've already looked at this)

DataScience@SMU

White Noise Review

Check for Model Appropriateness

A. Does the model "whiten" the residuals?

Recall:

- ARMA(*p*,*q*) models are based on the assumption that the noise, a_t, is white noise
- AIC and its variations select a model ***based on reducing the white noise variance,*** $\hat{\sigma}_a^2$, while controlling the number of parameters required to do so.

Checking Residuals for White Noise

Key point:

`$avar` is actually a "residual" variance and does not measure whether or not the residuals are ***white noise***.

Notes:

- If the residuals are not white noise, this suggests that further modeling may be necessary to better explain the behavior in the data.
- The residuals are calculated within the functions **`est.ar.wge`** and **`est.arma.wge`** and found in the output variable **`$res`**.
- We briefly discuss calculation of the residuals next.

DataScience@SMU

Calculating Residuals

Calculating Residuals

Suppose we have fit the following ARMA(*p,q*) model to a realization.

$$X_t - \hat{\varphi}_1 X_{t-1} - \cdots - \hat{\varphi}_p X_{t-p} = \hat{a}_t - \hat{\theta}_1 \hat{a}_{t-1} - \cdots - \hat{\theta}_q \hat{a}_{t-q} - \bar{X}(1 - \hat{\varphi}_1 - \cdots - \hat{\varphi}_p)$$

solving for $\hat{a}_t$ we get

Light Board

$$\hat{a}_t = X_t - \hat{\varphi}_1 X_{t-1} - \cdots - \hat{\varphi}_p X_{t-p} + \hat{\theta}_1 \hat{a}_{t-1} + \cdots + \hat{\theta}_q \hat{a}_{t-q} + \bar{X}(1 - \hat{\varphi}_1 - \cdots - \hat{\varphi}_p)$$

Assuming for simplicity that $p=2$ and $q=1$, we have

$$\hat{a}_t = X_t - \hat{\varphi}_1 X_{t-1} - \hat{\varphi}_2 X_{t-2} + \hat{\theta}_1 \hat{a}_{t-1} + \bar{X}(1 - \hat{\varphi}_1 - \hat{\varphi}_2)$$

So, $\hat{a}_1 = X_t - \hat{\varphi}_1 X_0 - \hat{\varphi}_2 X_{-1} + \hat{\theta}_1 \hat{a}_0 + \bar{X}(1 - \hat{\varphi}_1 - \hat{\varphi}_2)$

$$\hat{a}_2 = X_2 - \hat{\varphi}_1 X_1 - \hat{\varphi}_2 X_0 + \hat{\theta}_1 \hat{a}_1 + \bar{X}(1 - \hat{\varphi}_1 - \hat{\varphi}_2)$$

We don't know these values: set $\hat{a}_1 = \hat{a}_2 = 0$

$$\hat{a}_3 = X_3 - \hat{\varphi}_1 X_2 - \hat{\varphi}_2 X_1 + \hat{\theta}_1 \hat{a}_2 + \bar{X}(1 - \hat{\varphi}_1 - \hat{\varphi}_2)$$

$=0$

$$\hat{a}_4 = X_4 - \hat{\varphi}_1 X_3 - \hat{\varphi}_2 X_2 + \hat{\theta}_1 \hat{a}_3 + \bar{X}(1 - \hat{\varphi}_1 - \hat{\varphi}_2)$$

etc. to $\hat{a}_n$

DataScience@SMU

Comments on Calculating the Residuals

Comments on Calculating Residuals

The residuals that we just calculated are called:

conditional residuals.

- Note that there are n-p conditional residuals instead of n
- Woodward, et al (2017) discuss the use of "backcasting" to calculate $\boldsymbol{n}$ ***unconditional residuals***
- The residuals based on backcasting are given in the output variable **`$res`** in **`est.ar.wge`** and **`est.arma.wge`**.
- "Backcasting" allows us to to estimate $\hat{a}_0$- $\hat{a}_{p-1}$through looking the series in reverse order. This allows for the estimation of all n residuals. Therefore **`est.ar.wge`** and **`est.arma.wge`** provide all n residuals .

DataScience@SMU

Testing Residuals for White Noise

Visual

Testing Residuals for White Noise

We will focus on two methods for checking the residuals for white noise.

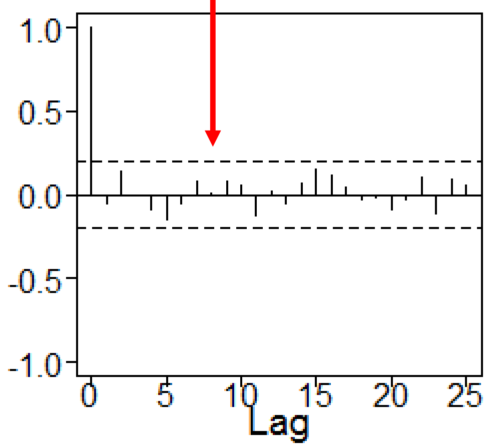
Check 1: Visually inspect plots of the residuals and their sample autocorrelations

- The residuals should look like white noise (random)
- About 95% of the sample autocorrelations of the residuals should stay within the limit lines.

Notes:

- This is the check for white noise that we recommended as the first step in analyzing a set of data.
- The limit lines provide a 5% level test for $H_0\!: \rho_k = 0$ vs $H_0\!: \rho_k \neq 0$ separately for each k.

They don't do it here, but with white noise we expect about 1 in 20 observations to randomly fall outside the limits. (5%)



DataScience@SMU

Testing Residuals for White Noise

Ljung-Box

Testing Residuals for White Noise

Check 2: *Ljung-Box test*

Whereas checking the limit lines apply separately to each lag *k*, Ljung-Box tests the hypothesis

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_K = 0$$

$$H_a : \text{ at least one } \rho_k \neq 0, \text{ for } 1 \leq k \leq K$$

The Ljung-Box test is referred to as a ***portmanteau*** [port man tō] test

- "Portmanteau" is a seldom-used word that can mean "embodying several uses or qualities"
- Ljung-Box tests the autocorrelations ***as a group***

Ljung-Box Test

The Ljung-Box test statistic for testing the null hypothesis:

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_K = 0$$

is

$$L = n(n+2)\sum_{k=1}^{K} \frac{\hat{\rho}_k^2}{n-k}$$

Under H_0, L is approximately χ^2 with $K - p - q$ d.f.

Reject H_0 if $L > \chi^2_{1-\alpha}(K - p - q)$

It is advisable to check more than one value of K

- Box and Jenkins use K=24 and 48

tswge function ljung.wge

```
ljung.wge(res,p,q,K)
# res residual file
# after ARMA(p,q) fit to data
# K is capital K above (default=24)
```

DataScience@SMU

Example 1

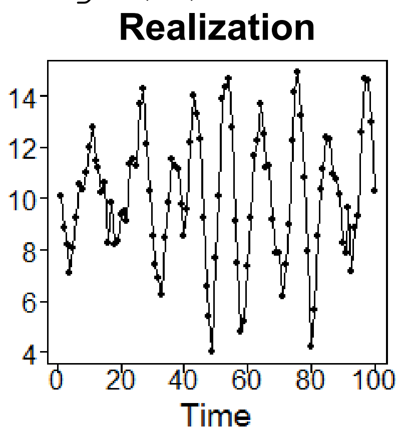
ARMA(2,1)

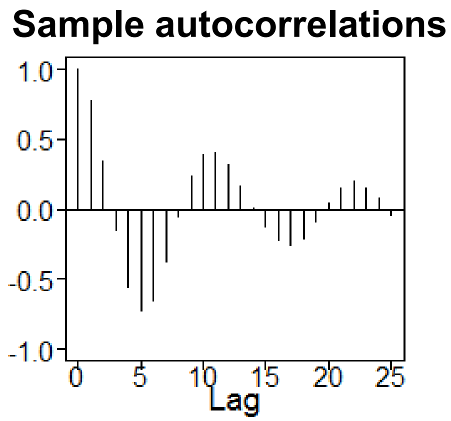
Example: ARMA(2,1)

Recall: Realization from ARMA(2,1):

$$(1-1.6B+.9B^2)(X_t-10)=(1-.8B)a_t$$

```
# generate the realization from ARMA(2,1)examined earlier
x=gen.arma.wge(n=100,phi=c(1.6,-.9),theta=.8,sn=67)
x=x+10
plotts.sample.wge(x)
```





- These seem to indicate stationarity
- Sample autocorrelations damp quickly

Example: ARMA(2,1)

```
aic.wge(x,p=0:8,q=0:4)
# AIC picks ARMA(2,1)
x21=est.arma.wge(x,p=2,q=1)
# x21$phi: 1.6194830 -0.9131788
# x21$theta: 0.868127
# x21$vara: 1.076196
mean(x)  # 10.07557
```

Final model we found in estimation unit

$$(1-1.62B+.91B^2)(X_t-10.08)=(1-.87B)a_t \quad \hat{\sigma}_a^2=1.08$$

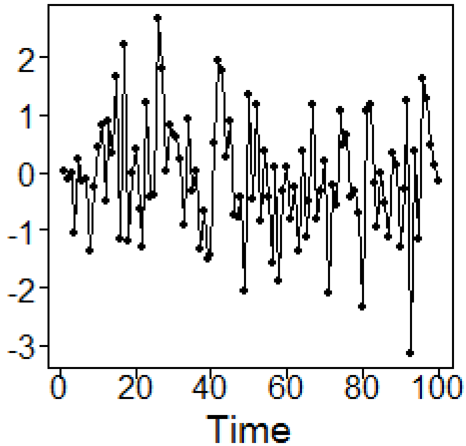
Next we examine the residuals

```
# x21$res: Contains residuals from the ARMA(2,1) fit
```

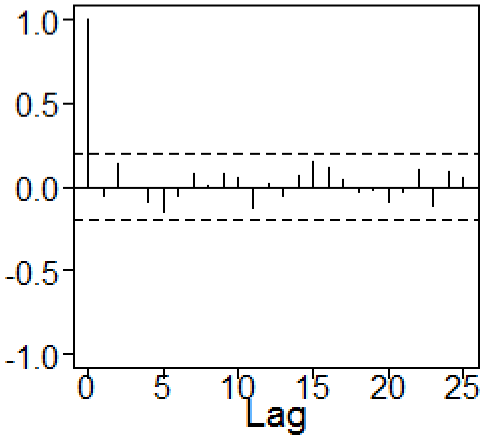
Residual Analysis: ARMA(2,1) Example

Check 1: Examine plots of residuals and their sample autocorrelations

Residuals (x21$res)



Residual sample autocorrelations with 95% limit lines



- Residuals look "white"
- Residual sample autocorrelations within 95% limit lines

Residual Analysis: ARMA(2,1) Example

Check 2: Ljung-Box test

```
ljung.wge(x21$res,p=2,q=1)
# $K: 24  (default)
# $chi.square: 20.92251
# $df: 21
# $pval: 0.4636851
```

```
ljung.wge(x21$res,p=2,q=1,K=48)
# $K: 48
# $chi.square: 44.93891
# $df: 45
# $pval: 0.4636851
```

For both K=24 and 48 we fail to reject white noise

Based on Checks 1 and 2 the residuals from the fitted ARMA(2,1) model seem to be white.

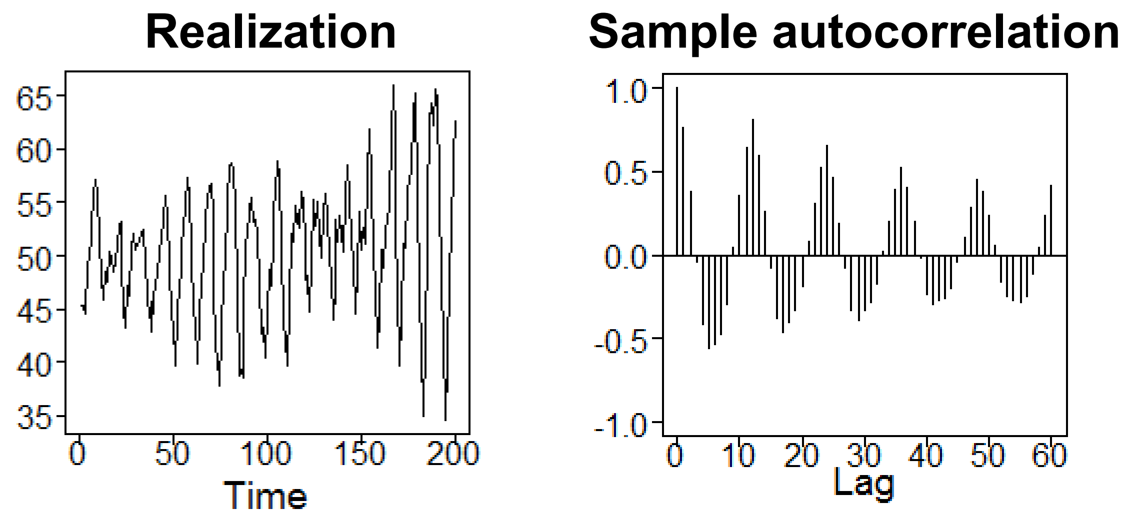
DataScience@SMU

Example 2

Seasonal $(1-B^{12})$

Simulated Seasonal Data Analyzed Earlier

$$(1 - B^{12})(1 - 1.5B + .8B^2)(X_t - 50) = a_t$$



tswge code to simulate and plot the data above

```
x=gen.aruma.wge(n=200,s=12,phi=c(1.5,-.8),sn=87)
x=x+50
plotts.sample.wge(x,lag.max=60)
```

Simulated Seasonal Data Analyzed Earlier

$$(1 - B^{12})(1 - 1.5B + .8B^2)(X_t - 50) = a_t$$

Modeling the data:

- Overfit tables suggested a factor of $(1 - B^{12})$
- We transformed the data by $(1 - B^{12})$

```
y=artrans.wge(x,phi.tr=c(0,0,0,0,0,0,0,0,0,0,0,1))
```

- Transformed data appeared to be stationary
- After transforming the data we used BIC which selected an AR(2) model

```
est.y=est.ar.wge(y,p=2)
```

- And obtained the fitted model

$$(1 - B^{12})(1 - 1.47B + .76B^2)(X_t - 49.78) = a_t \qquad \hat{\sigma}_a^2 = 1.04$$

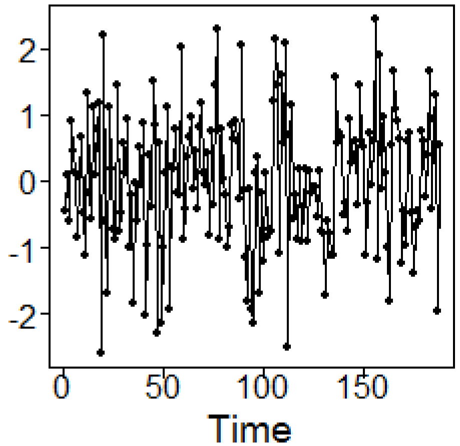
How about the residuals?

Simulated Seasonal Data Analyzed Earlier

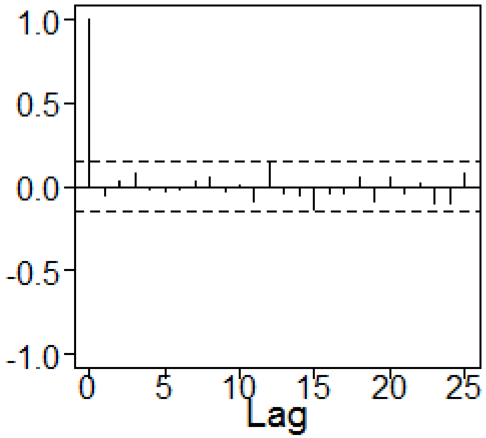
The residuals are in **`est.y$res`**

Check 1: Examine plots of residuals and their sample autocorrelations

Residuals (est.y$res)



Residual sample autocorrelations with 95% limit lines



- Residuals look “white”
- Residual sample autocorrelations within 95% limit lines

Simulated Seasonal Data Analyzed Earlier

Check 2: Ljung-Box test

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_K = 0$$

H_a : at least one $\rho_k \neq 0$, for $1 \leq k \leq K$

Recall: When we fit an ARIMA or Seasonal Model, we transform the data to stationarity. The *p* and *q* in the Ljung-Box call statement are for the estimation of the stationary component. (in this case *p*=2)

```
ljung.wge(est.y$res,p=2)
# $K: 24  (default)
# $chi.square: 20.95537
# $df: 22
# $pval: 0.5234977
```

```
ljung.wge(est.y$res,p=2,K=48)
# $K: 48
# $chi.square: 47.62509
# $df: 46
# $pval: 0.4063877
```

For both K=24 and 48 we fail to reject white noise

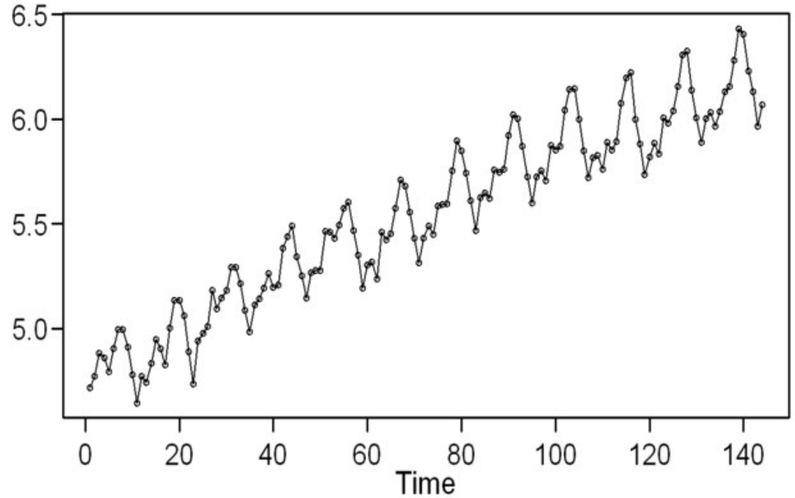
Based on Checks 1 and 2 the residuals from the fitted seasonal model seem to be white.

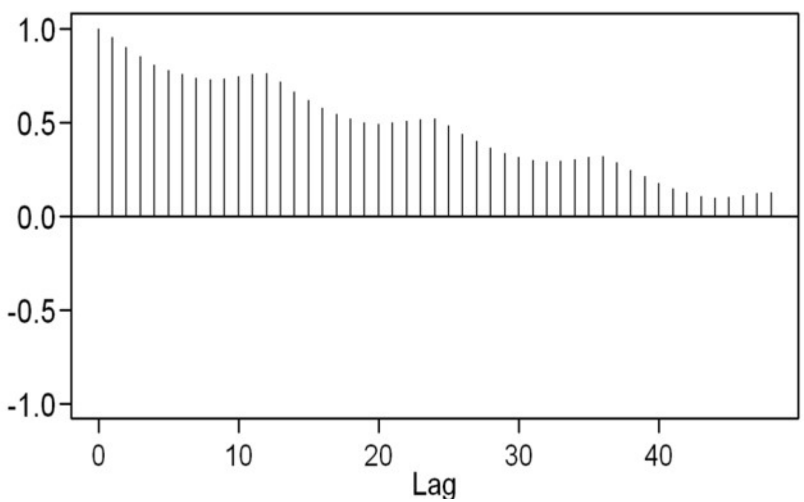
DataScience@SMU

Analysis of Residuals

Log Airline Data

Log Airline Data





Recall: We previously fit the model

$$(1-B)(1-B^{12})\varphi_{12}(B)(X_t - 5.54) = (1 - .45B)a_t \ \hat{\sigma}_a^2 = .0013$$

$$\varphi_{12}(B) = 1 - .008B - .080B^2 + .107B^3 + .021B^4 - .080B^5 - .041B^6$$
$$+.055B^7 - .036B^8 - .133B^9 + .053B^{10} + .0123B^{11} + .403B^{12}$$

Log Airline Data

In order to analyze the residuals, we recreate the steps necessary to retrieve them in **`tswge`**.

- We overfit the data with *p*=14 and 16 (steps not shown) and determined the need to transform to obtain $(1-B)(1-B^{12})X_t$=**`d1.12`** in the code below

```
data(airlog)
# transform data
# Difference the data
d1=artrans.wge(airlog,phi.tr=1)
# Transform differenced data by 1-B^12
s12=c(0,0,0,0,0,0,0,0,0,0,0,1)
d1.12=artrans.wge(d1,phi.tr=s12)
aic.wge(d1.12,p=0:15,q=0:3)
# aic and aicc pick ARMA(12,1)
# estimate parameters of stationary part
est.12.1=est.arma.wge(d1.12,p=12,q=1)
```

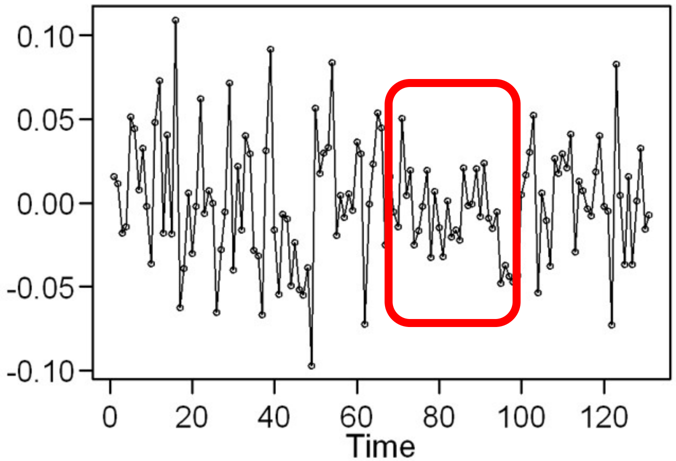
Log Airline Model: Analysis of Residuals

The code on the previous slide produces the fitted model obtained earlier.

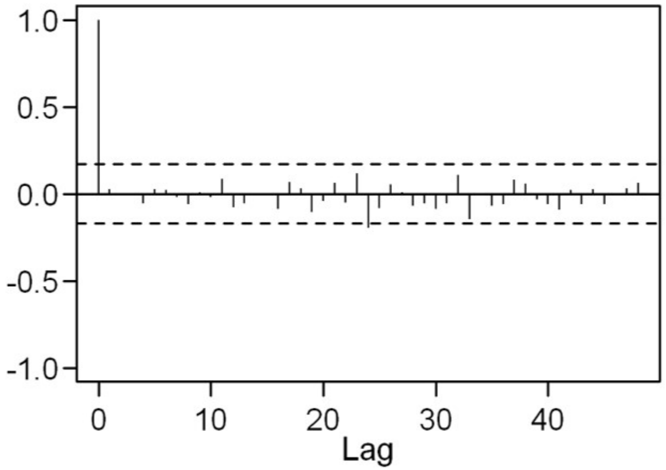
- The residuals are in **`est.12.1$res`**

Check 1: Examine residuals and their sample autocorrelations

Residuals



Residual sample autocorrelations



- Residuals look “fairly white” (unusual behavior between 65-100)
- Residual sample autocorrelations within 95% limit lines

Log Airline Model: Analysis of Residuals

Check 2: Ljung-Box test

$H_0 : \rho_1 = \rho_2 = \cdots = \rho_K = 0$

$H_a :$ at least one $\rho_k \neq 0,$ for $1 \leq k \leq K$

As with the previous example, *p* and *q* for the Ljung-Box test are those obtained when fitting a stationary model to the transformed data.
(in this case *p*=12, *q*=1)

```
ljung.wge(est.12.1$res,p=12,q=1)
# $K: 24  (default)
# $chi.square: 17.30648
# $df: 11
# $pval: 0.09913114

ljung.wge(est.12.1$res,p=12,q=1,K=48)
# $K: 48
# $chi.square: 35.93309
# $df: 35
# $pval: 0.4245906
```

Log Airline Model: Analysis of Residuals

Conclusions and comments:

- The residuals "pass" both checks for white noise
  - We noted some behavior that was somewhat worrisome in the residual plot
  - For $K=24$, we did not reject H_0 but we would have if testing at $\alpha = .10$
- The first two examples were simulated data from ARMA and seasonal models
  - The residuals were "nice and white"
- For the log airline data, the seasonal model we fit is our "best guess" at a model that describes the behavior of the data
  - In practice, we often see residual analyses that aren't as definitive as in the simulated examples
  - In fact, the airline results are quite good

DataScience@SMU

Next Steps

Next Step:
Answer Questions of Interest

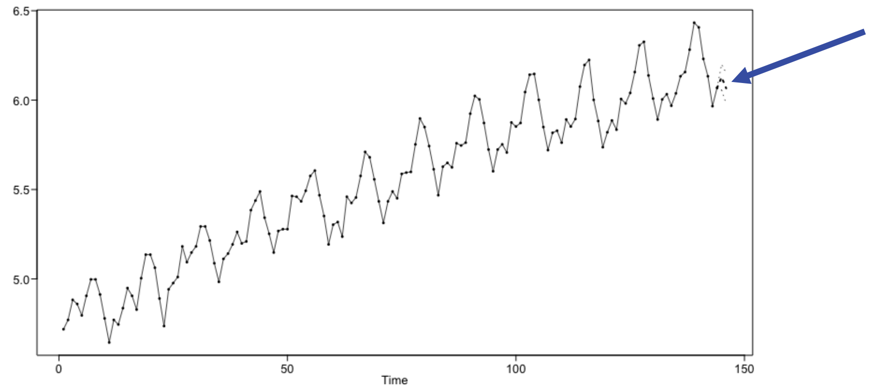
The question of interest may have been to forecast the number of airline passengers two months later and to quantify our uncertainty.

TwoMonthFore = fore.aruma.wge(airlog,d = 1, s = 12, phi = est.12.1$phi, theta = est.12.1$theta,n.ahead = 2, limits = TRUE)
TwoMonthFore

```
$f
[1] 6.124662 6.070614

$ll
[1] 6.053338 5.989049

$ul
[1] 6.195986 6.152180
```



Conclusion: In two months we are 95% confident that the number of airline passengers will be between 399,415 ($e^{5.99}$ * *1000)* and 468,717 ($e^{6.15}$ * *1000) passengers*. Our best estimate is 432,681 ($e^{6.07}$ * *1000) passengers*.

DataScience@SMU

More Checks for Model Appropriateness

Does the model make sense?

Another important check for model appropriateness

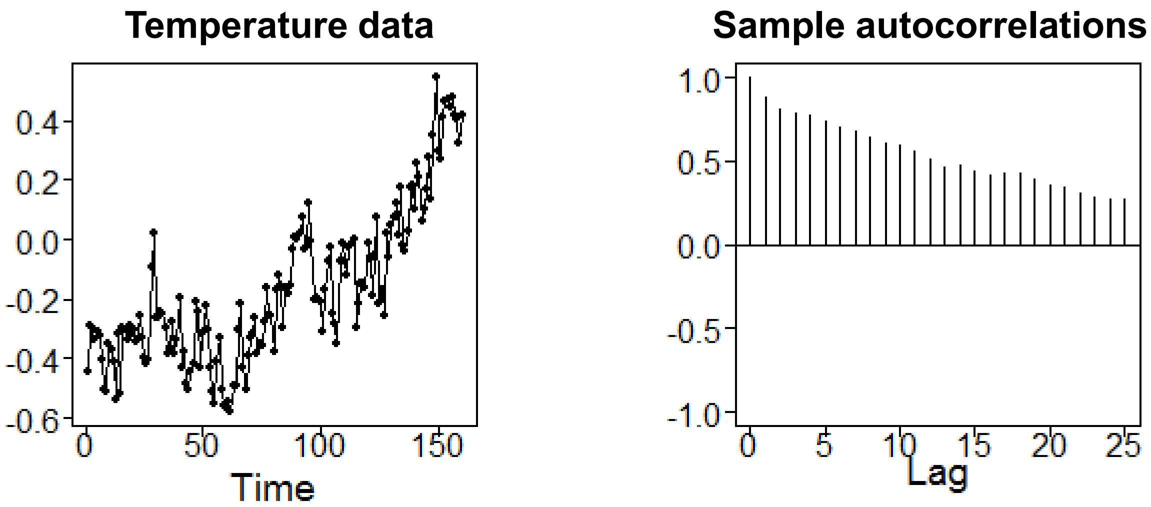
- Stationary vs. nonstationary
- Seasonal vs. non-seasonal
- Correlation-based vs. signal-plus-noise model
- Are characteristics of fitted model consistent with those of the data
  - Forecasts and spectral estimates make sense?
  - Do realizations and their characteristics behave like the data?

DataScience@SMU

Stationary Model

Global Temperature Data

Modeling Global Temperature Data



a) Fitting a stationary model to the data:

```
data(hadley)
mean(hadley)  # -0.1684937
plotts.sample.wge(hadley)
aic5.wge(hadley,p=0:6,q=0:1)
# AIC picks an ARMA(3,1) stationary model
had.est=est.arma.wge(hadley,p=3,q=1)
# $phi: 1.2700171 -0.4685313  0.1911988
# $theta: 0.6322319
# $avar: 0.01074178
```

Stationary model:

Fitted ARMA(3,1) model

$$(1 - 1.27B + .47B^2 - .19B^3)(X_t + .17) = (1 - .63B)a_t$$

where $\hat{\sigma}_a^2 = .0107$

or in factored form

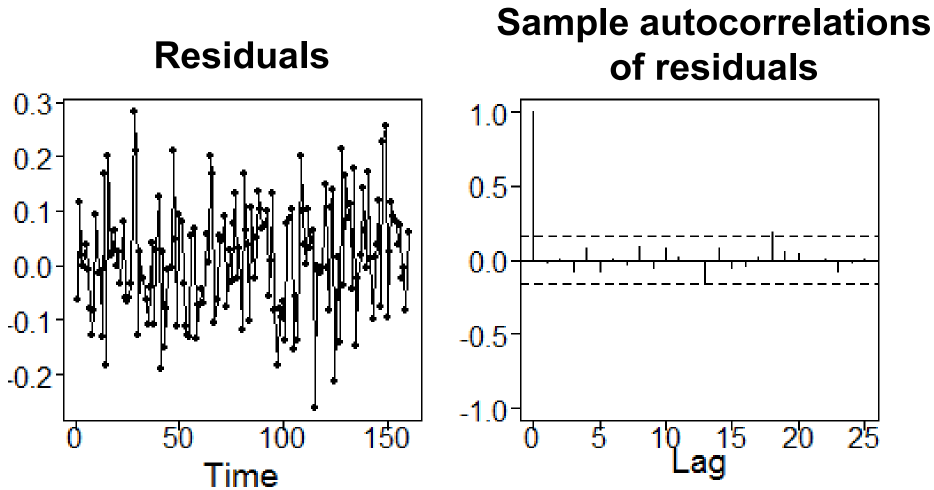
$$(1 - .99B)(1 - .28B + .19B^2))(X_t + .17) = (1 - .63B)a_t$$

(this is a “nearly nonstationary” model)

Stationary model:

Check residuals

```
plotts.sample.wge(had.est$res,arlimits=TRUE)
ljung.wge(had.est$res,p=3,q=1)
ljung.wge(had.est$res,p=3,q=1,K=48)
```



Residuals look "white" and residual sample autocorrelations stay sufficiently within 95% limit lines

Ljung-Box results

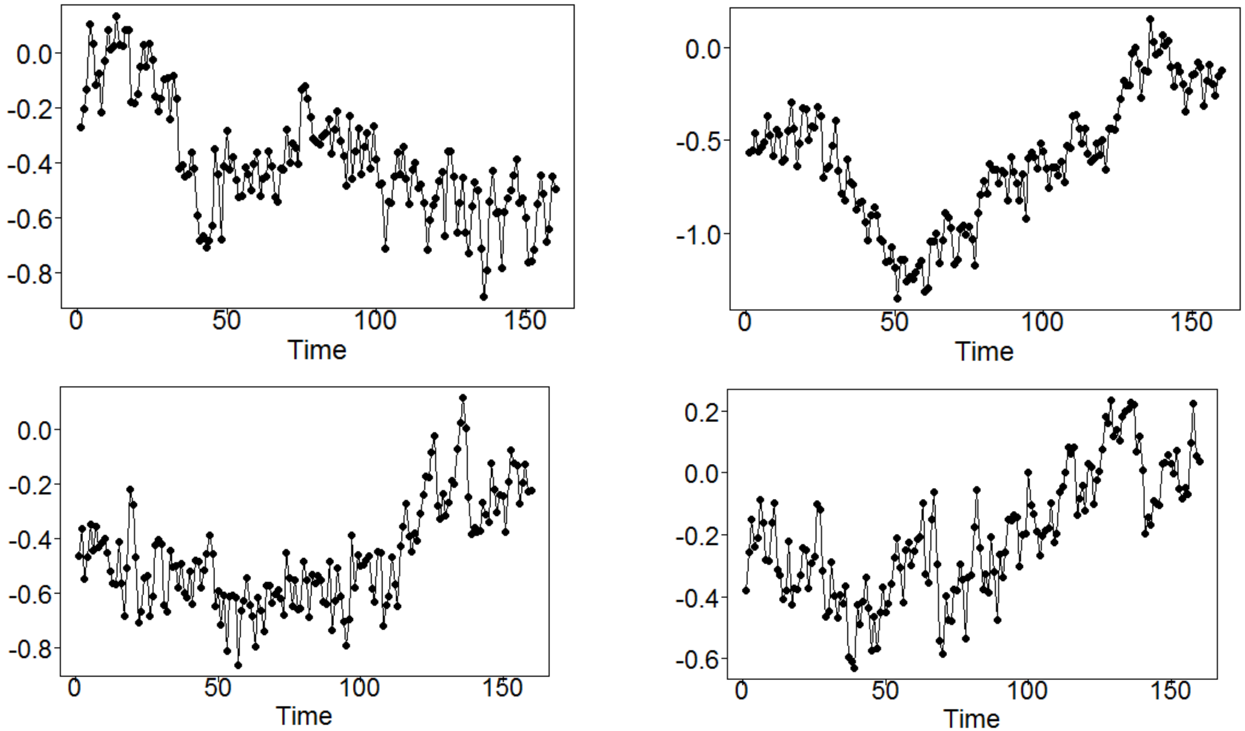
P-values for $K=24$ and $K=48$ are 0.42 and 0.41, respectively.

Conclusion: Residuals for stationary ARMA(3,1) fit appear to be white.

Stationary model:

Realizations

- Does the ARMA(3,1) model produce realizations that “look like” the temperature data?
- Realizations below were generated from the ARMA(3,1) model
- They have the same general behavior as temperature data



DataScience@SMU

Nonstationary Model

Global Temperature Data

Nonstationary model fit to temperature data:

Indications of a unit root of +1

There are several indications that an ARIMA model might be appropriate for the temperature data.

- The stationary model has a factor of $(1 - .99B)$
- The wandering behavior and fairly slowly damping sample autocorrelations
- The overfit tables with p=8 and p=12 (not shown) suggest the possibility of a single unit root of +1
- The Dickey-Fuller test of H_0: the model has a unit root, is not rejected (p-value=.5611)

b) Nonstationary model fit to temperature data:

Suppose that based on the evidence on the previous slide ***we make the decision to fit an ARIMA model*** to the temperature data

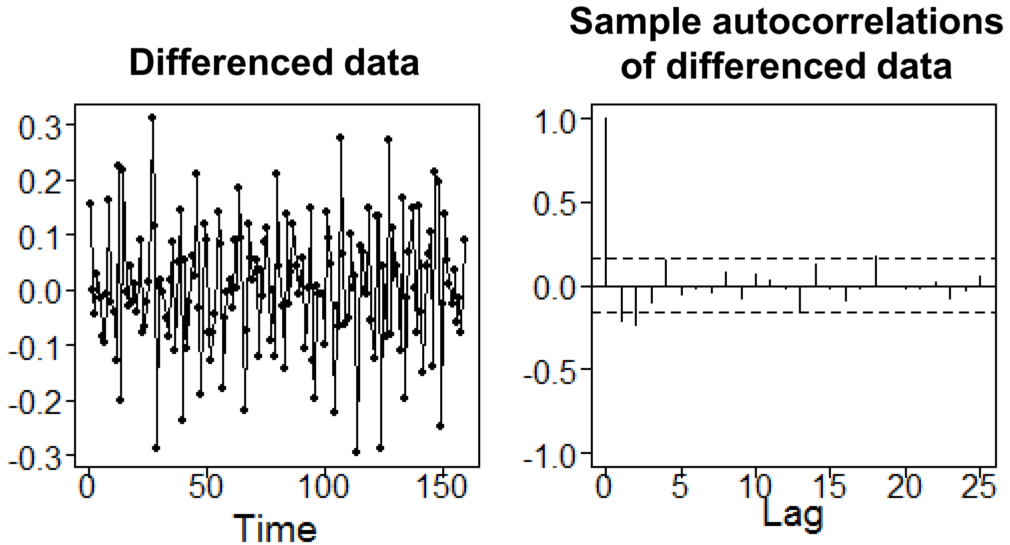
- (Even though the model checks: white noise residuals, realizations that have the appearance of the data, etc. were good for the ARMA model)

In this case we proceed by differencing the **`hadley`** data.

Nonstationary model fit to temperature data:

`tswge` code to difference the data

```
d1.temp=artrans.wge(hadley,phi.tr=1)
plotts.sample.wge(d1.temp,arlimits=TRUE)
```



- The differenced data appear to be stationary
  - And nearly “white”
- However, the fact that the first two sample autocorrelations are outside the limits lines suggests that we continue to model

Nonstationary model fit to temperature data:

`tswge` code to model the differenced data

```
aic5.wge(d1.temp,p=0:6,q=0:1)
# AIC selects an ARMA(2,1)
d1.temp.est=est.arma.wge(d1.temp,p=2,q=1)
# $phi: 0.3274341 -0.1786827
# $theta: 0.704618
$avar: 0.01058826
```

Fitted ARIMA(2,1,1) model

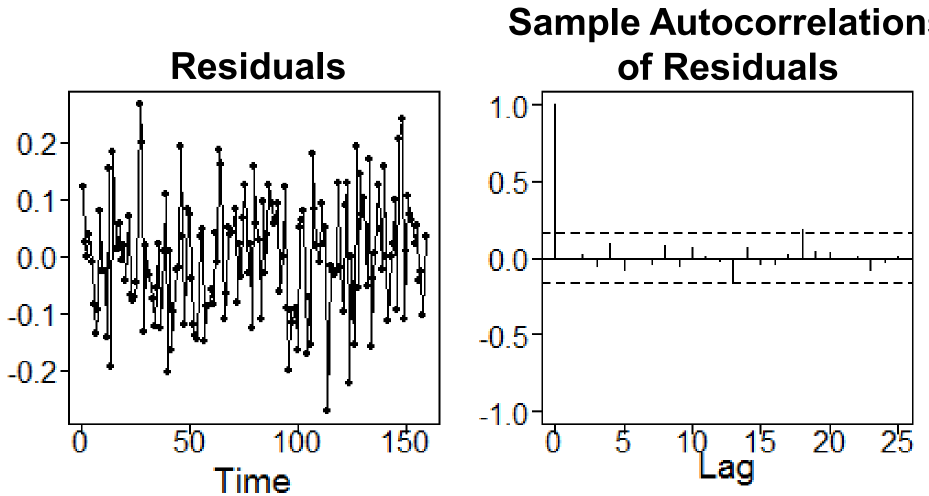
$$(1-B)(1-.33B+.18B^2))(X_t+.17)=(1-.70B)a_t$$

where $\hat{\sigma}_a^2=.0106$

Nonstationary model fit to temperature data:

Check residuals

```
plotts.sample.wge(d1.temp.est$res,arlimits=TRUE)
ljung.wge(d1.temp.est$res,p=2,q=1)
ljung.wge(d1.temp.est$res,p=2,q=1,K=48)
```



Residuals look “white” and residual sample autocorrelations stay sufficiently within 95% limit lines

Ljung-Box results

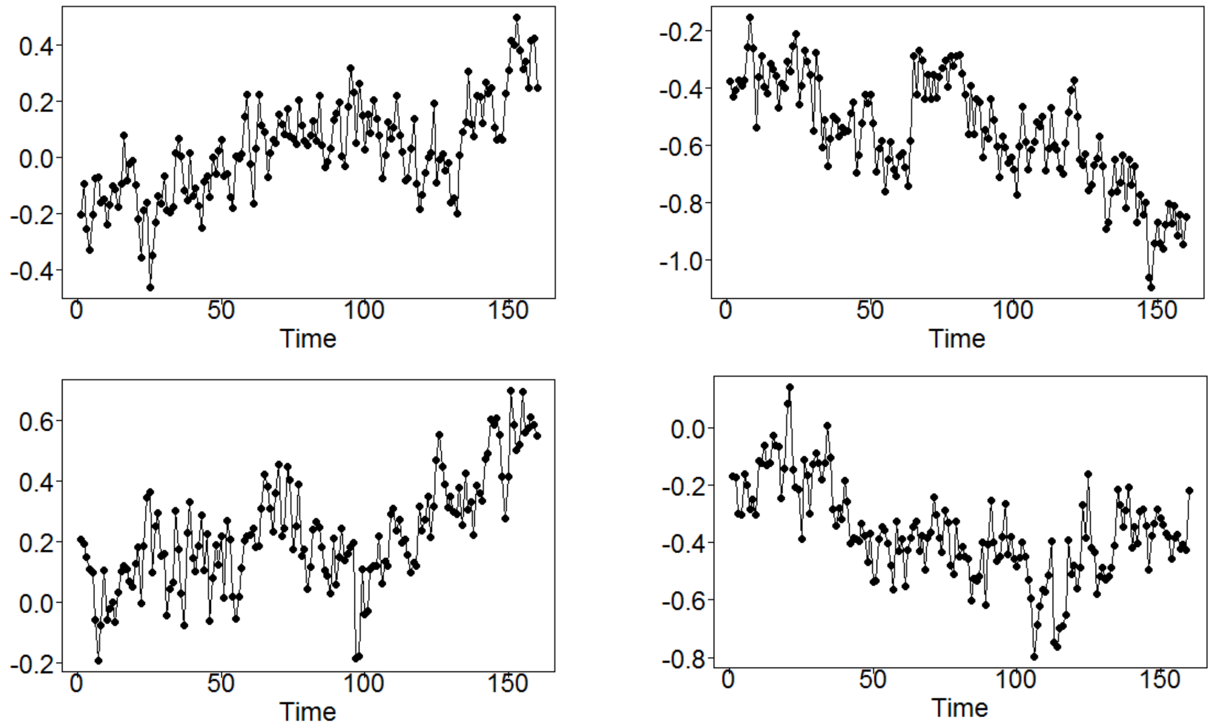
P-values for $K=24$ and $K=48$ are 0.47 and 0.58, respectively.

Conclusion: Residuals for stationary ARMA(2,1) fit appear to be white.

b) Nonstationary model fit to temperature data:

Realizations

- Does the ARIMA(2,1,1) model produce realizations that "look like" the temperature data?
- Realizations below were generated from the ARMA(3,1) model
- They have same general behavior as temperature data



Stationary ARMA(3,1) vs. ARIMA(2,1,1) Fit to Temperature Data

Clearly: The two models are quite similar to each other.

Stationary model

$(1-.99B)(1-.28B+.19B^2))(X_t+.17)=(1-.63B)a_t$

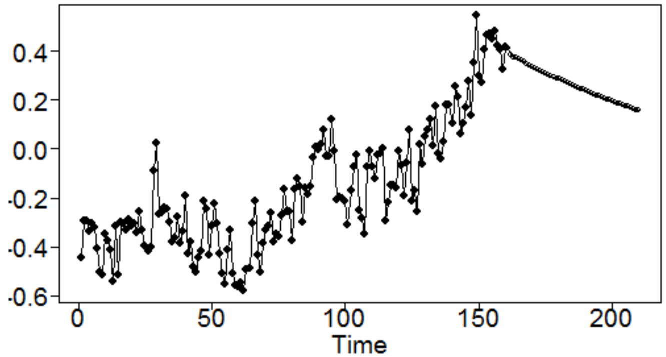
Nonstationary model

$(1-B)(1-.33B+.18B^2))(X_t+.17)=(1-.70B)a_t$

- The main difference is the stationary factor $(1-.99B)$ vs. the nonstationary factor $(1-B)$
- Residuals appear to be white for both models
- Realizations from the two models are similar
- ***How about forecasts?***

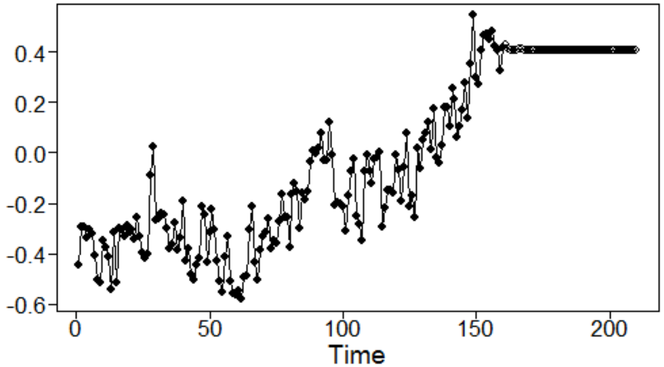
Forecasts using stationary model

```
data(hadley)
fore.arma.wge(hadley,phi=c(1.27,-.47,.19),
theta=.63,n.ahead=50,limits=FALSE)
```



Forecasts using nonstationary model

```
fore.aruma.wge(hadley,d=1,phi=c(.33,-.18),
theta=.7,n.ahead=50,limits=FALSE)
```



Notes:

- The two “similar” models produce very different forecasts
- It is important to understand the properties of the selected model
  - The selection of a **stationary** model will *automatically* produce forecasts that eventually *tend toward the mean* of the observed data
  - Forecasts from the ARIMA(2,1,1) model will *tend to a horizontal line* at about the level of the last data value.
- **That is:** It was the ***decision*** to use an ARMA or ARIMA model that resulted in the forecasts produced

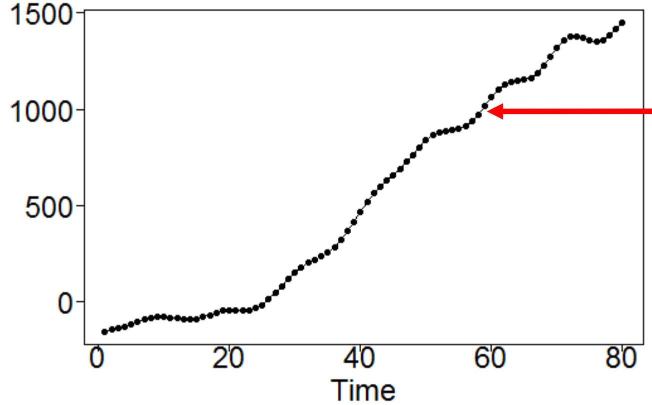
Beware of results by analysts who choose a model in order to produce desired results

Note: For the temperature data, the debate centers around whether we should ***predict the trend to continue***.

For "correlation-based" models:

In order for a trend to be predicted to continue we need 2 unit roots.

- We did see this in the airline data
- In the non-seasonal case, realizations from models with 2 unit roots often look like



Highly correlated wandering pattern – much more so than the temperature data

- The choice of a correlation based model makes it difficult to conclude that we should predict a trend to continue

Caution: The decision concerning whether the observed warming trend should be predicted to continue is one that involves a variety of climatological issues we are not discussing here.

We simply ask the question:

Given reasonable models fit to the historical data, would these models predict the current trend to continue?

- The answer is "No" based on the standard ARMA and ARIMA fits to the temperature data
  - Which seem like reasonable models and easily passed the checks for white noise residuals, etc.
  - ***So, is there any statistical argument for claiming the trend should continue?***

DataScience@SMU

Signal-Plus-Noise Model

Global Temperature Data

Let's consider the signal-plus-noise model:

$$X_t = a + bt + Z_t$$

We have already discussed the facts that:

- It is often difficult to distinguish between realizations from the above model and those from ARMA/ARIMA models.
- Such models are appealing since they ***simultaneously consider correlation based vs. deterministic trending behavior.***
- If the null hypothesis $H_0 : b = 0$ is rejected then the existing trend would be predicted to continue into the future
  - With the usual precautions about extrapolating into the future where conditions might change.

Fitting Signal-Plus-Noise Models

When testing $H_0\colon b = 0$ we noted that:

- A simple linear regression analysis is inappropriate because it fails to take into account the correlation in the residuals (Z_t).
- Methods such as Cochrane-Orcutt and MLE have been proposed for dealing with the correlated residuals issue.
  - But they have an issue with rejected the null hypothesis when it is true more than 5% of the time.

However: For purposes of this example, we will use the Cochrane-Orcutt procedure for testing for trend in the temperature data set.

Signal-plus-noise model fit to temperature data: testing b (slope) with Cochrane-Orcutt

Model: $X_t = a + bt + Z_t$

Testing: $H_0: b = 0$ vs $H_a: b \neq 0$

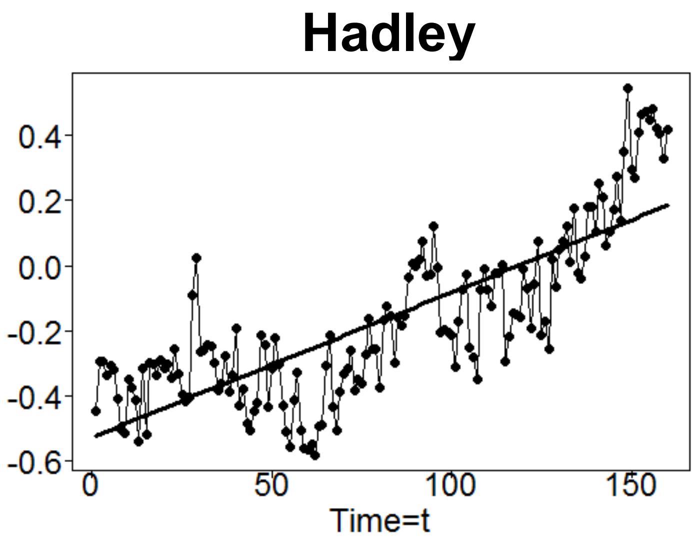
The following R code provides a Cochrane-Orcutt test for the temperature data

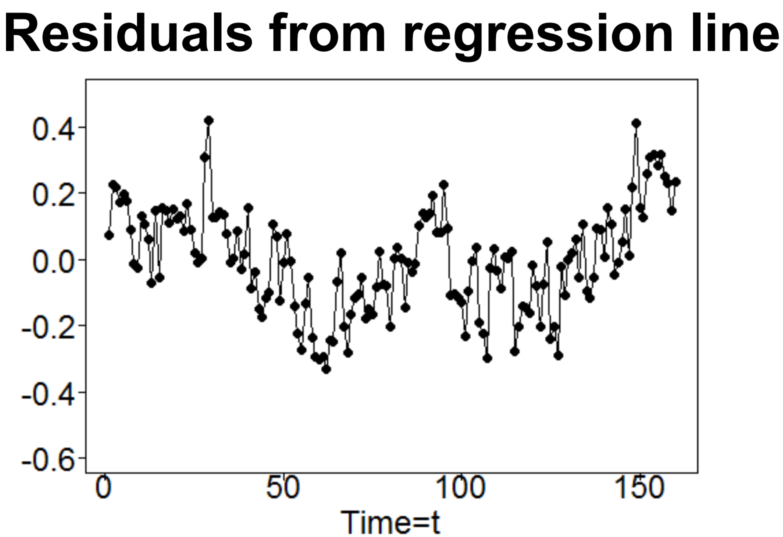
1. Fit a regression line to the data and find the residuals from the line

```
x=hadley
n=length(x)
t=1:n
d=lm(x~t)
x.z=x-d$coefficients[1]-d$coefficients[2]*t
#x.z are the residuals from the regression line
```

Signal-plus-noise model fit to temperature data: using Cochrane-Orcutt

Below are plotted the Hadley temperature data and residuals from the regression line





The trending behavior is mostly removed in the residuals but there is still substantial autocorrelation

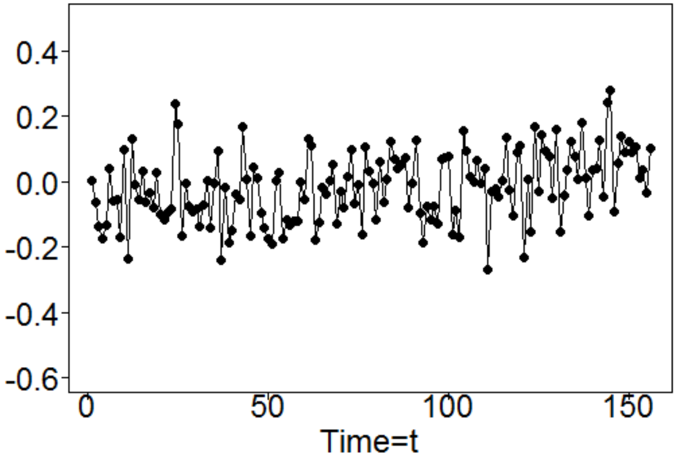
- i.e. correlated residuals

Signal-plus-noise model fit to temperature data: using Cochrane-Orcutt

2. Fit an AR(*p*) model $\widehat{\varphi}_Z(B)$ to the residuals and find $\hat{Y}_t = \widehat{\varphi}_Z(B)X_t$ $(\widehat{Y_t}$=x.trans in the code below)

```
ar.z=aic.wge(x.z,p=0:6)
# ar.z$p is the order p
#ar.z$phi is vector of ar.z$p estimated AR coefficients
x.trans=artrans.wge(hadley,phi.tr=ar.z$phi)
```

x.trans (plotted against time *t*)



Signal-plus-noise model fit to temperature data: using Cochrane-Orcutt

3. Transform the independent variable (time)

   $\hat{T}_t = \hat{\varphi}_Z(B)T_t \; T_1 = 1, T_2 = 2\; etc.$ $\hat{T}_t$= t.trans

```
#ar.z$phi is vector of ar.z$p estimated AR coefficients
t.trans=artrans.wge(t,phi.tr=ar.z$phi)
```

4. Regress $\hat{Y}_t$ on $\hat{T}_t$ using OLS.

```
fit = lm(y.trans~t.trans)
summary(fit)
```

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.094089   0.018883  -4.983 1.67e-06 ***
t.trans      0.005700   0.001248   4.566 1.01e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

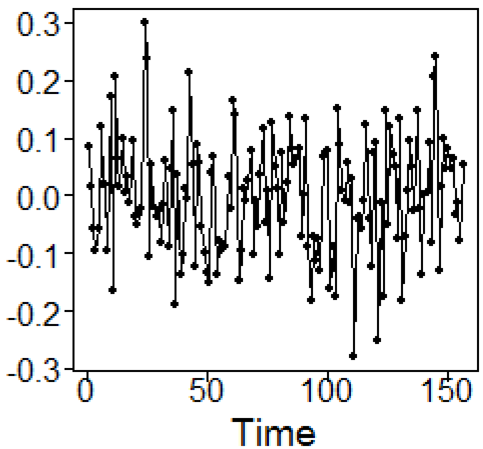
After accounting for the serial correlation (AR(4)), there is strong evidence to suggest that the slope is significantly different from zero (pvalue < .0001).

Signal-plus-noise model fit to temperature data: using Cochrane-Orcutt

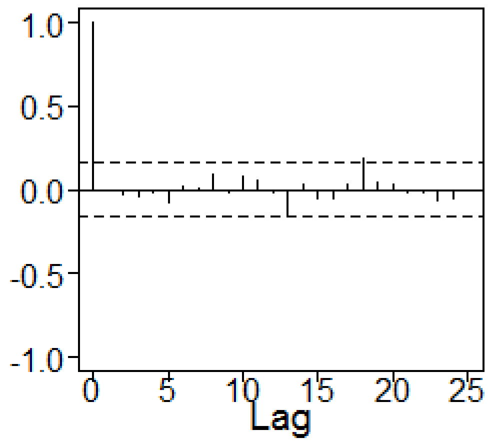
Evaluating residuals (after Cochrane–Orcutt)

```
plot(fitco$residuals)
acf(fitco$residuals)
ljung.wge(fitco$residuals)
```

Residuals from transformed fit



Sample autocorrelations of residuals



Sample autocorrelations tend to stay within limit lines and Ljung-Box test has p-values of .805 and .577 for K=24 and 48 respectively.

Signal-plus-noise model fit to temperature data: using Cochrane-Orcutt

Estimated signal-plus-noise model

$X_t = -.5257 + .0044t + Z_t$ where $\hat{\sigma}_a^2 = .0103$

$(1 - .614B + .044B^2 - .078B^3 - .026B^4)Z_t = a_t$

or $(1 - .92B)(1 - .21B + .43B^2)(1 + .52B)Z_t = a_t$

Note 1: The above ***is the signal-plus-noise fit to the temperature data.*** Cochrane-Orcutt is a procedure to assess the significance of the slope (adjusting for the correlated errors).

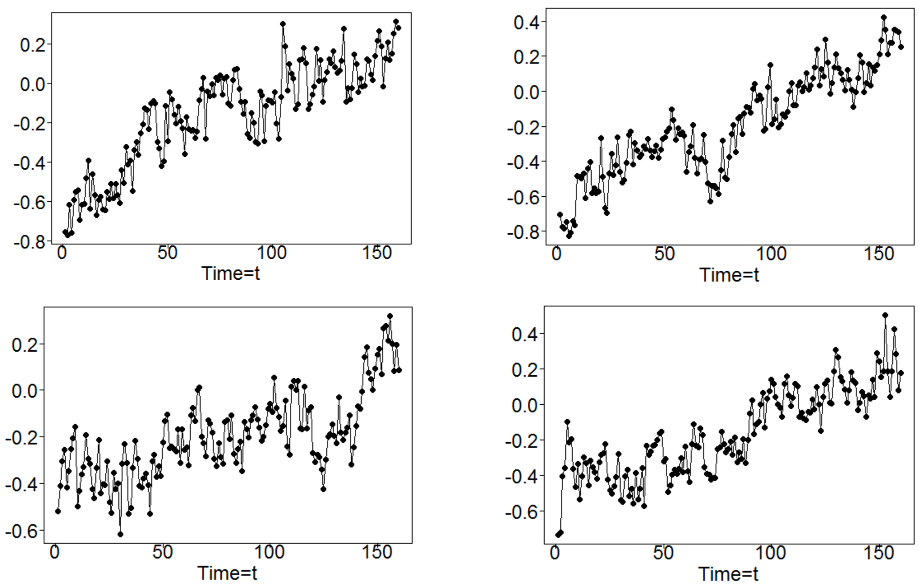
Note 2: We had to code the Cochrane-Orcutt Procedure manually since the function cochrane.orcutt() is only for AR(1) correlation.

Signal-plus-noise model fit to temperature data: using Cochrane-Orcutt

Evaluating realizations from fitted model

$X_t = -.5257 + .0044t + Z_t$, $\hat{\sigma}_a^2 = .0103$, and Z_t is AR(4)

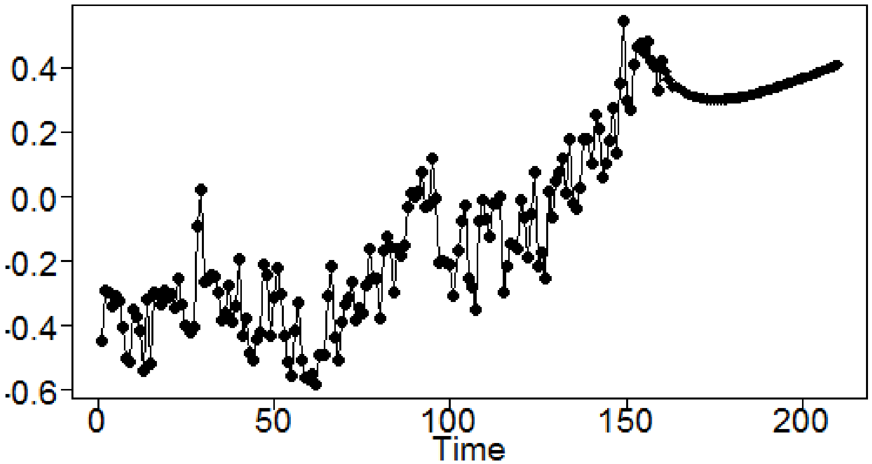
- Realizations have the appearance of temperature data
- All tend to increase because of the line with positive slope




```
gen.sigplusnoise.wge(160, b0 = -.5257, b1 = .0044, phi = ar.z$phi, vara = .0103)
```

Forecasts using Signal-Plus-Noise Model

```
data(hadley)
fore.sigplusnoise.wge(hadley,max.p=4,n.ahead
=50,limits=FALSE)
```



Interestingly, the forecasts suggest an initial decline but eventually predict the trend to continue

DataScience@SMU

Final Thoughts on Temperature Data

Important Points

- Realizations from AR (ARMA/ARUMA) models have ***random trends***
  - Unless there are 2 unit roots, these models ***will not*** forecast a trend to continue
- Realizations from $X_t = S_t + Z_t$ have deterministic trends
  - *If conditions do not change*, then these trends ***will be*** forecast to continue
- Regarding the temperature data, if there is a deterministic signal in the data, it almost assuredly is not simply a straight line

Final Thoughts for Temperature Data

- All three models (ARMA, ARIMA, and signal-plus-noise) seemed to be satisfactory models from the standpoint of
  - Residual analysis
  - Realizations generated
- However, the three produced strikingly different forecasts
- Knowledge of the physical situation can help guide you, incorrect assumptions may lead you to the wrong conclusion
- ***We can't stress enough:***

Beware of results by analysts who choose a model in order to produce desired results

DataScience@SMU

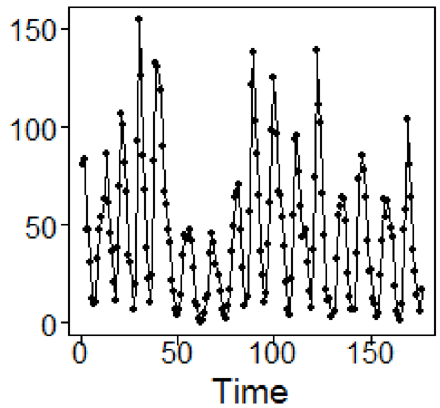
DataScience@SMU

More on Model Appropriateness

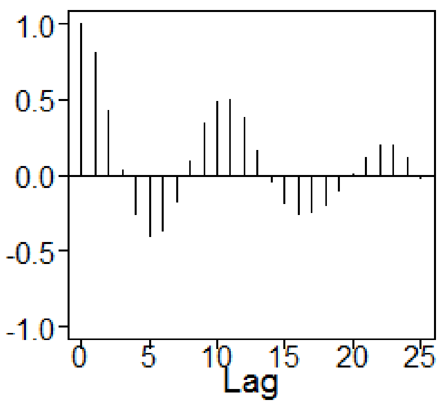
Sunspot Data

Sunspot Data:
1749-1924 (sunspot.classic)

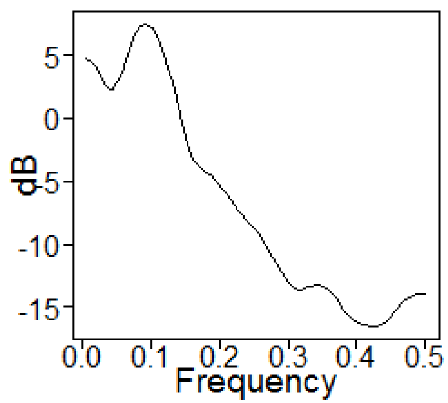
Sunspot data



Sample autocorrelations



Parzen spectral density



Note:

This is a classic and often analyzed time series data set.

- It is in data file `sunspot.classic` in `tswge`
- Box and Jenkins modeled the data in their textbook using the “Box-Jenkins” approach mentioned earlier
- We will consider the problem of modeling these data and checking model appropriateness

Some Final Comments on Checking Models for Appropriateness

We have briefly discussed the issue of checking models to see if they produce realizations that "behave like the original data." In general, we may also want to check whether generated realizations:

- Have sample autocorrelations similar to those for the actual data?
- Have spectral densities that are similar to those for the data?
- Let's investigate further using the classic sunspot data (sunspot.classic in tswge).

DataScience@SMU

Sunspot Data

Box-Jenkins Model

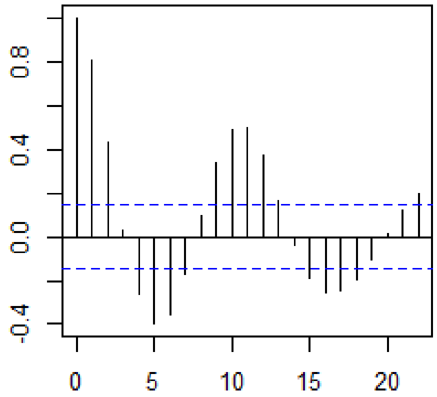
Sunspot Data: Box-Jenkins Model

The Box-Jenkins procedure involves ***plotting the sample autocorrelations and sample partial autocorrelations and looking for patterns***

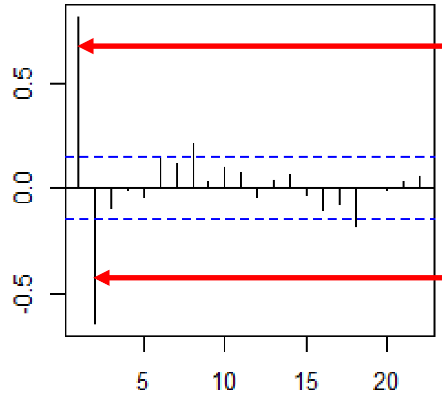
- Sample autocorrelations and partial autocorrelations can be obtained using base R functions

```
acf(sunspot.classic)
pacf(sunspot.classic)
```

Sample autocorrelations



Partial autocorrelations



The two large partial autocorrelations strongly suggest an AR(2)

Sunspot Data: Box-Jenkins Model

Box and Jenkins fit an AR(2) model to the data. Using MLE estimates for the AR(2) we obtain

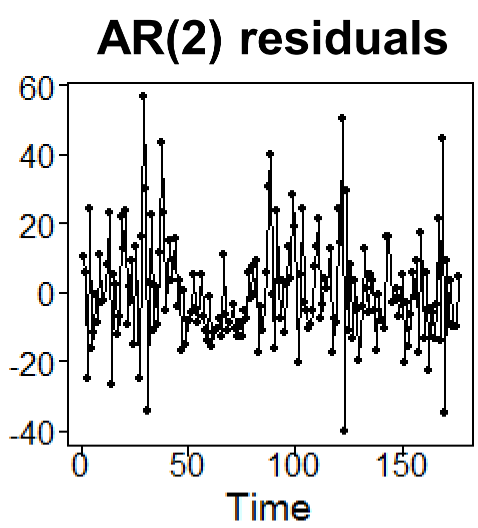
```
s2=est.ar.wge(sunspot.classic,p=2)
# s2$phi: 1.3347282 -0.6474423
# s2$avar:235.993
mean(sunspot.classic)  # 44.78409
```

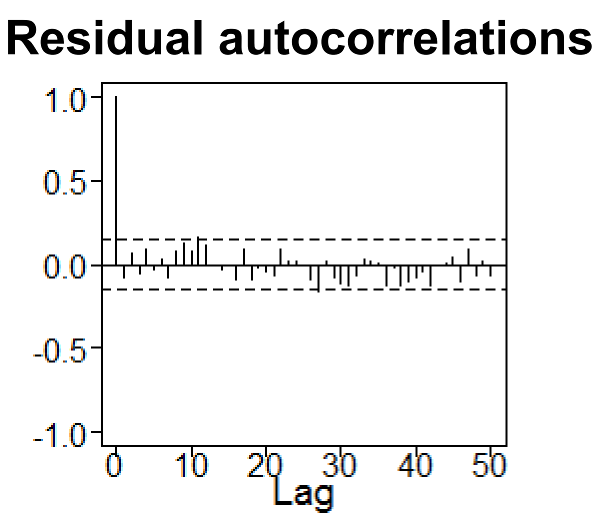
$(1-1.33B+.65B^2)(X_t-44.78)=a_t$ where $\hat{\sigma}_a^2=236$

```
Factor                     Abs recip    System freq
1-1.3347B+0.6474B^2         0.8046        0.0943
```

The factor table shows that this model is associated with a pseudo-cyclic behavior with frequency $f_0=.094$ or cycle length $1/.094=10.6$ years which is consistent with the data.

Sunspot Data: Box-Jenkins Model





- The residuals look reasonably white
- Sample autocorrelation of the residuals stay within the 95% limit lines
- Ljung-Box did not reject the null of white noise at K=48 (p-value=.21) but at K=48 (p-value=.05) the conclusion of white noise is somewhat questionable

DataScience@SMU

Sunspot Data

AIC Model

Sunspot Data: AIC Model Selection

We next let AIC select a model for the sunspot data. We have chosen to select the best AR model using the code below:

```
aic5.wge(sunspot.classic,p=0:10,q=0:0)
# AIC picks an AR(8)
# FYI BIC selects an AR(2)
s8=est.ar.wge(sunspot.classic,p=8)
# s8$phi: 1.22872595 -0.47331327 -0.13807811  0.15688938 -
0.14030802  0.07050449 -0.12841889 0.20692558
# s8$avar:212.6003
mean(sunspot.classic)  # 44.78409
```

The resulting AR(8) model selected by AIC is:

$(1 - 1.23B + .47B^2 - .14B^3 - .16B^4 + .14B^5 - .07B^6 + .13B^7 - .21B^8)(X_t - 44.78) = a_t$ where $\hat{\sigma}_a^2 = 213$

Sunspot Data: AIC Model Selection

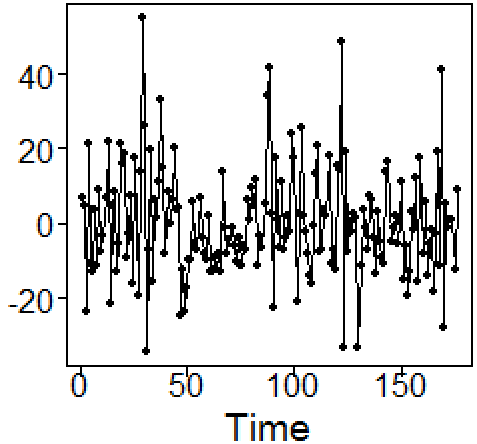
The factor table shown below also has a dominant frequency associated with the 10.5 year period.

Factor	Abs recip	System freq
$1-1.5565B+0.8970B^2$	0.9471	0.0965
$1-0.8771B$	0.8771	0.0000
$1-0.4147B+0.6550B^2$	0.8093	0.2088
$1+0.7964B$	0.7964	0.5000
$1+0.8231B+0.5043B^2$	0.7101	0.3484

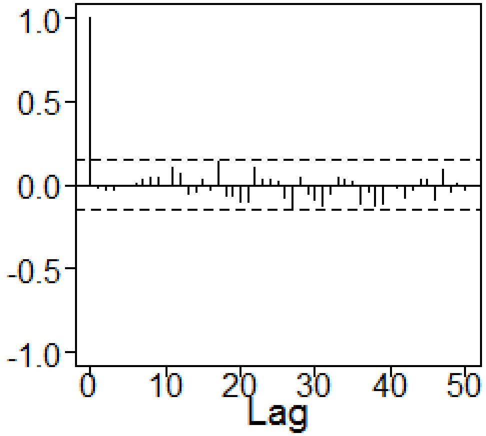
Sunspot Data: AIC Model Selection

$(1 - 1.23B + .47B^2 - .14B^3 - .16B^4 + .14B^5 - .07B^6 + .13B^7 - .21B^8)(X_t - 44.78) = a_t$ where $\hat{\sigma}_a^2 = 213$

AR(8) residuals



Residual autocorrelations



- Again, the residuals look reasonably white
- Sample autocorrelation of the residuals stay within the 95% limit lines
- Ljung-Box did not reject the null of white noise at K=24 or at K=48 with p-values of .33 and .24, respectively

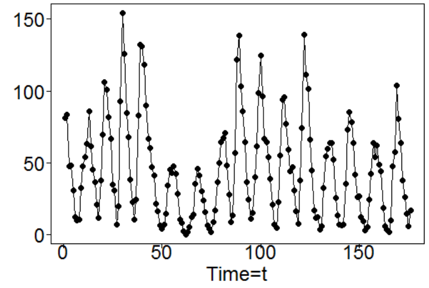
DataScience@SMU

Sunspot Models

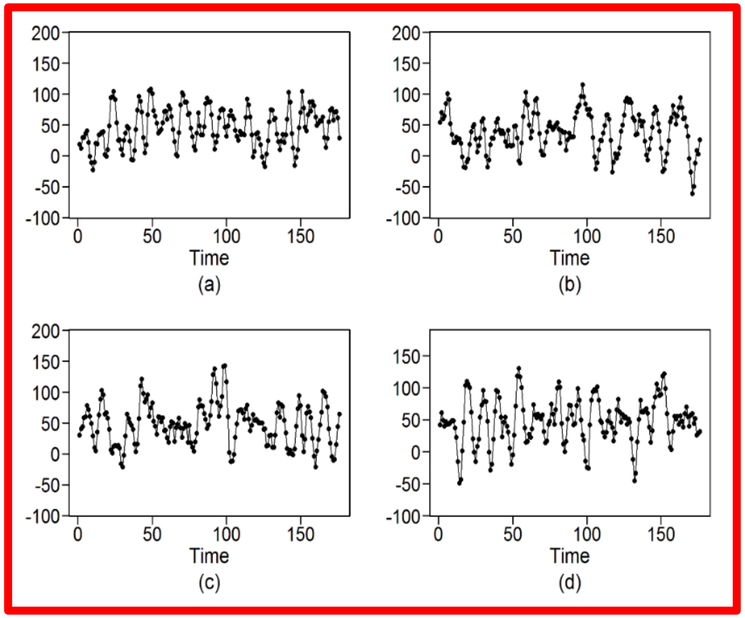
Comparing Realizations

Sunspot Models: Comparing Realizations

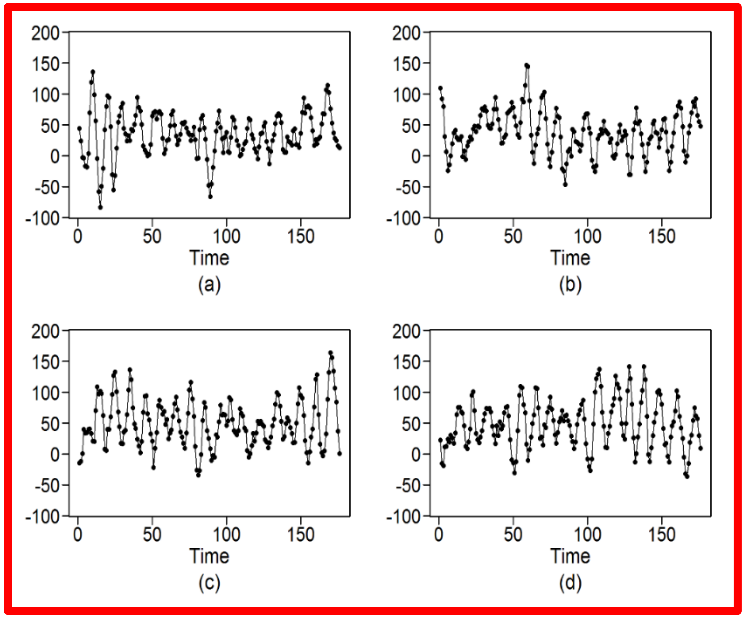
Sunspot data



Realizations from the AR(2) model



Realizations from the AR(8) model



DataScience@SMU

Sunspot Models: Comparing Realizations

Clearly neither model produced realizations that were very similar to the sunspot data

- The sunspot data has an asymmetric appearance in which the peaks are more variable than the troughs
  - The AR models produce no such asymmetry
  - Might need nonlinear models not covered here
- Also, the cyclic behavior holds together stronger in the sunspot data than in the realizations from the two models
  - Although the cycles may tend to hold up better in the AR(8) realizations

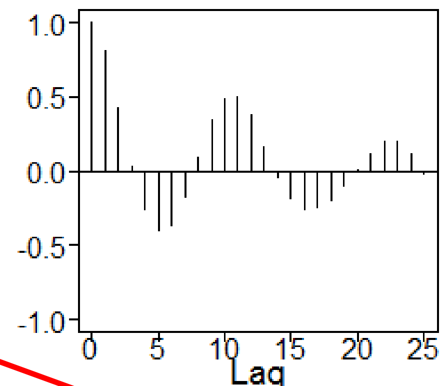
DataScience@SMU

Sunspot Models

Comparing Autocorrelations

Sunspot Models: Comparing Autocorrelations

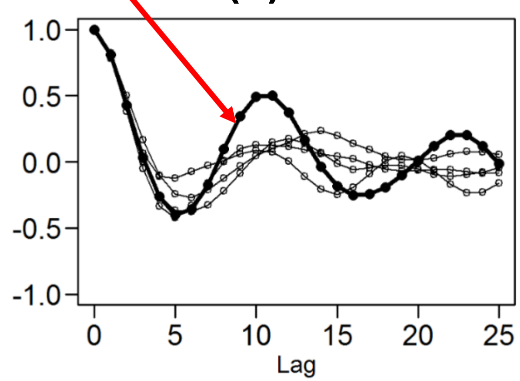
Sunspot data



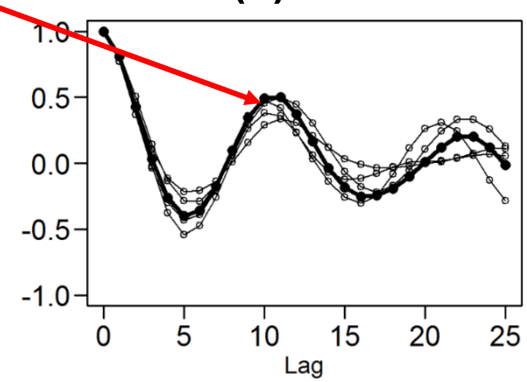
Dark curve shows sunspot autocorrelations

4 other curves in each plot are the sample autocorrelations from the realizations

AR(2) model



AR(8) model



AR(8) autocorrelations track the sunspot autocorrelations much better

DataScience@SMU

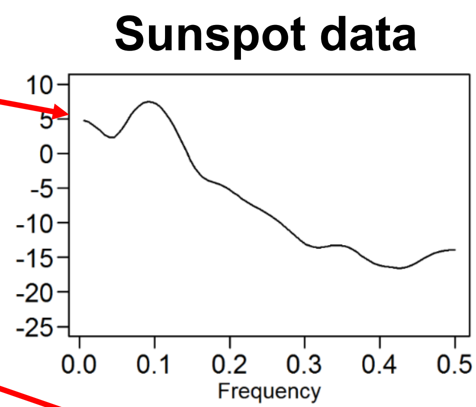
Sunspot Models

Comparing Spectral Densities

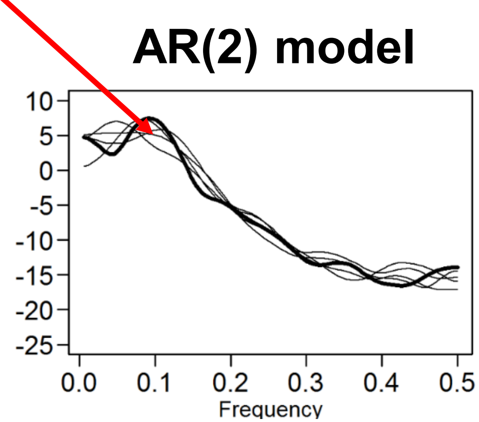
Sunspot Models: Comparing Spectral Densities

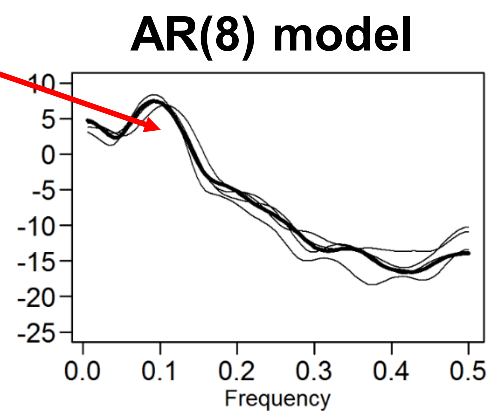
Peak at zero has to do with wandering behavior of peak sizes

Dark curve shows sunspot spectral density



4 other curves in each plot are the spectral densities from the realizations





AR(8) spectral densities do a better job of showing a peak at $f=0$ (AR(2) model only has one peak in its true spectral density).

DataScience@SMU

Sunspot Models

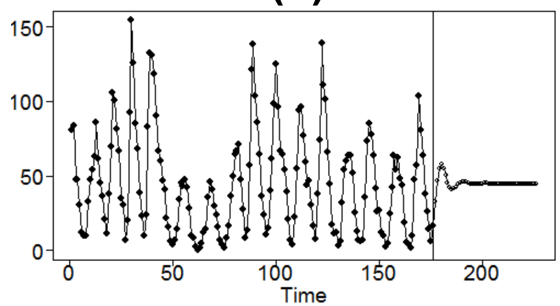
Comparing Forecasts

Sunspot Models: Comparing Forecasts

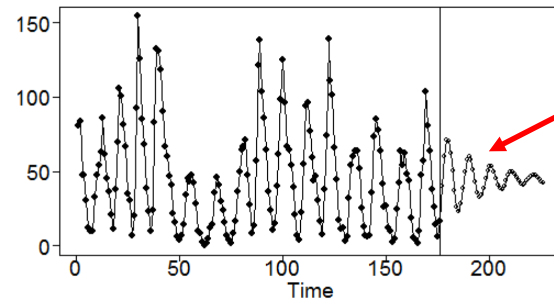
Forecast next 50

These look better than AR(2)

AR(2)



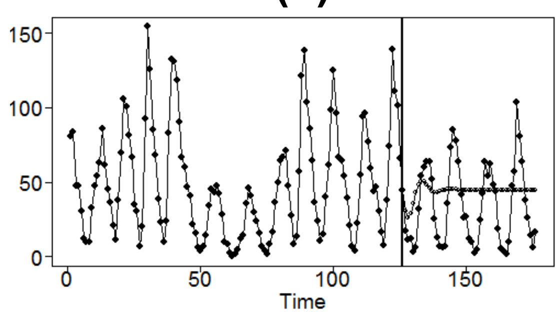
AR(8)



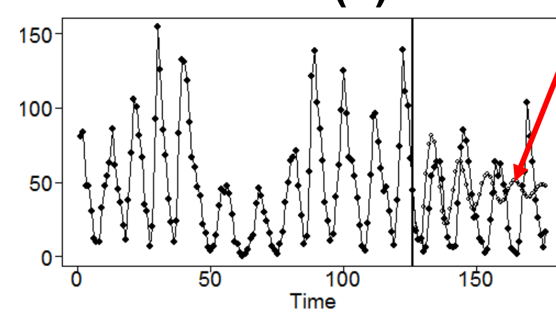
Forecast last 50

Forecast peak at a trough!

AR(2)



AR(8)



Oops!

DataScience@SMU

Sunspot Models

Summary

Sunspot Models: Summary

The AR(8) model seemed superior to the AR(2) model picked using partial autocorrelations

- White noise test more conclusive
- More similar sample autocorrelations
- More similar spectral densities that tended to show the peak at zero in the sunspot spectral density
- Forecasts pick up the cyclic behavior better
  - But forecasts of last 50 years got "off cycle"
  - Sunspot cycle length varies some from cycle to cycle (pseudo periodic)
  - ***Be careful about long term forecasts!***
- ***BONUS: Check Out: https://spaceplace.nasa.gov/solar-cycles/en/***

DataScience@SMU

DataScience@SMU