Tues, 2-19-19

Autoregressive-Moving-Average (ARMA) Models

1. Lightning talk
2. Self assessment on A4
3. Purpose and context of ARMA modeling
4. Equations for some simple models
5. Steps in modeling
6. Simulation

Read notes_5.pdf

A4 Self Assessment

1. Download A4x.pdf from D2L
2. Automatic points, for running assignment and having uploaded by due time, is already marked in parentheses at top of first page
3. Each assignment has maximum possible 10 points; if you make no deductions, score is 10/10
4. A4x is color coded for points; purple=1; yellow=0.5; blue=0.5
5. Open your copy of the same assignment pdf you uploaded
6. In Acrobat Reader, using “Add text box,” mark in right margin for deductions only, with deduction and segment reference : (eg., -0.5 A)
7. At top of your pdf, mark grade like this : 9.5/10
8. If necessary, put any comments at top near the grade
9. Upload your self-graded pdf to folder A4_graded in D2L

Purpose and context of ARMA modeling

ARMA Modeling

Autoregressive

Moving average

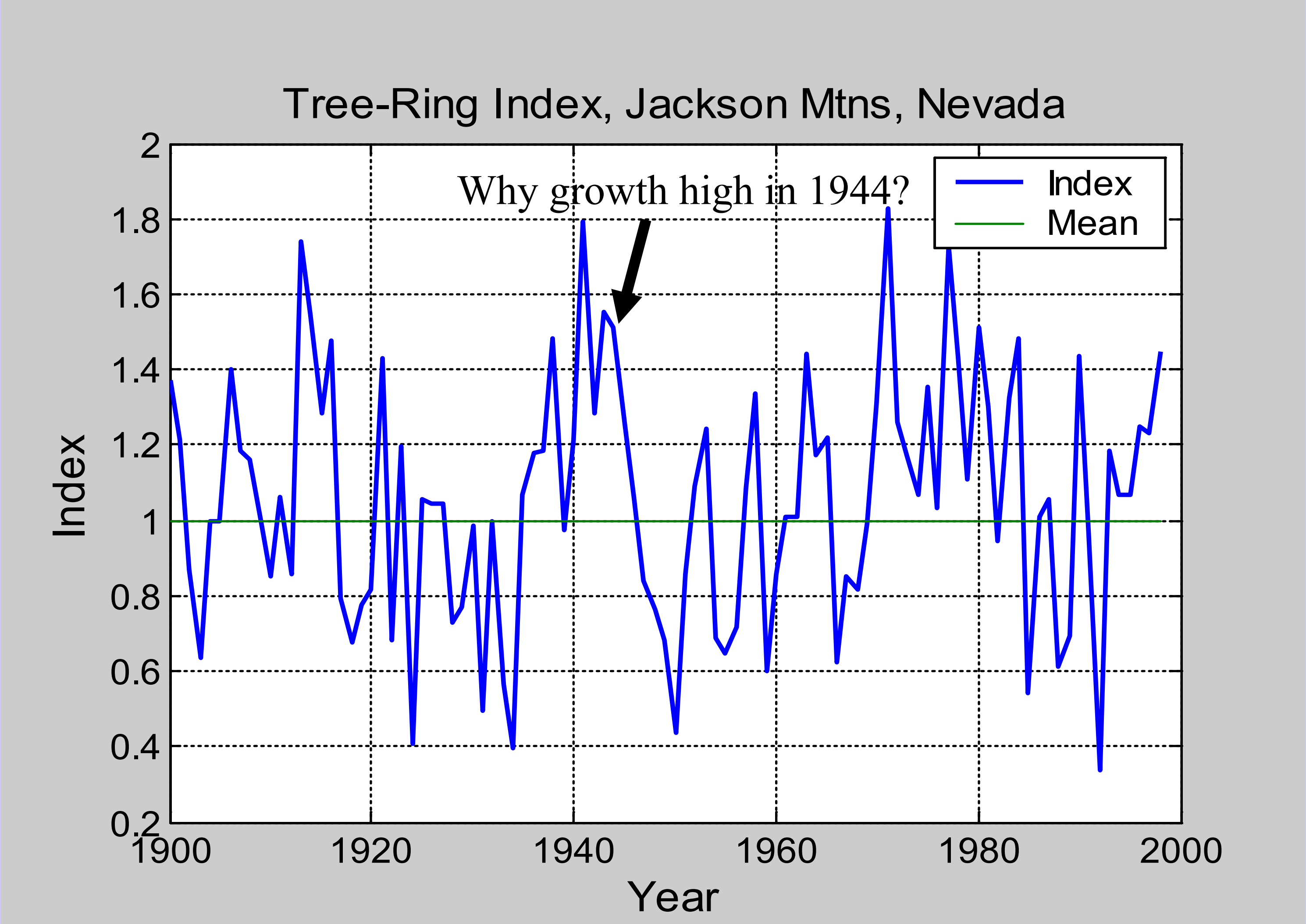
Purposes

- Get clues to processes
- Predict future behavior
- Remove persistence
- Simulation

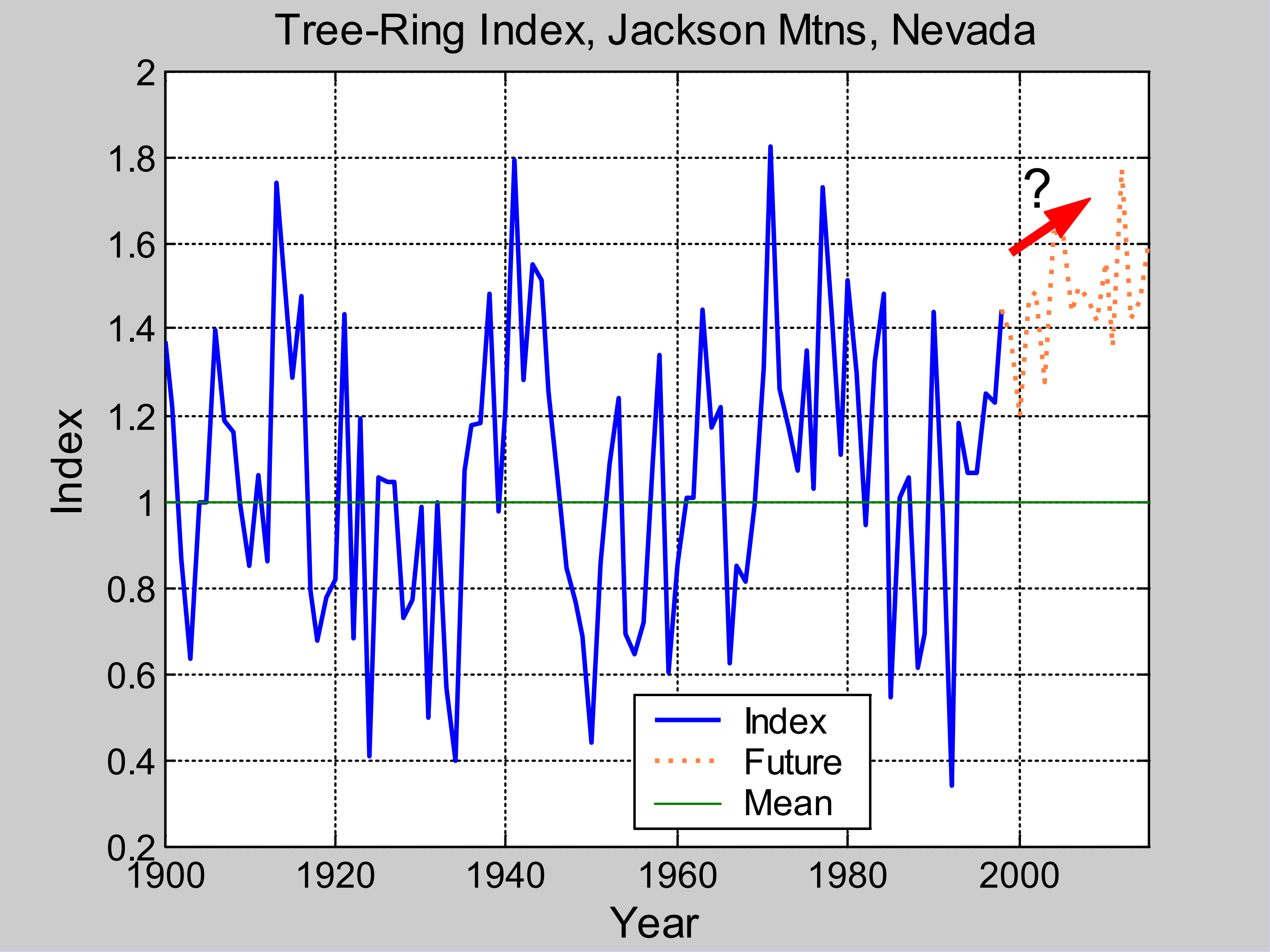
Some fields of application

- Industrial process control
- Hydrology
- Economics
- Dendrochronology

Explanation

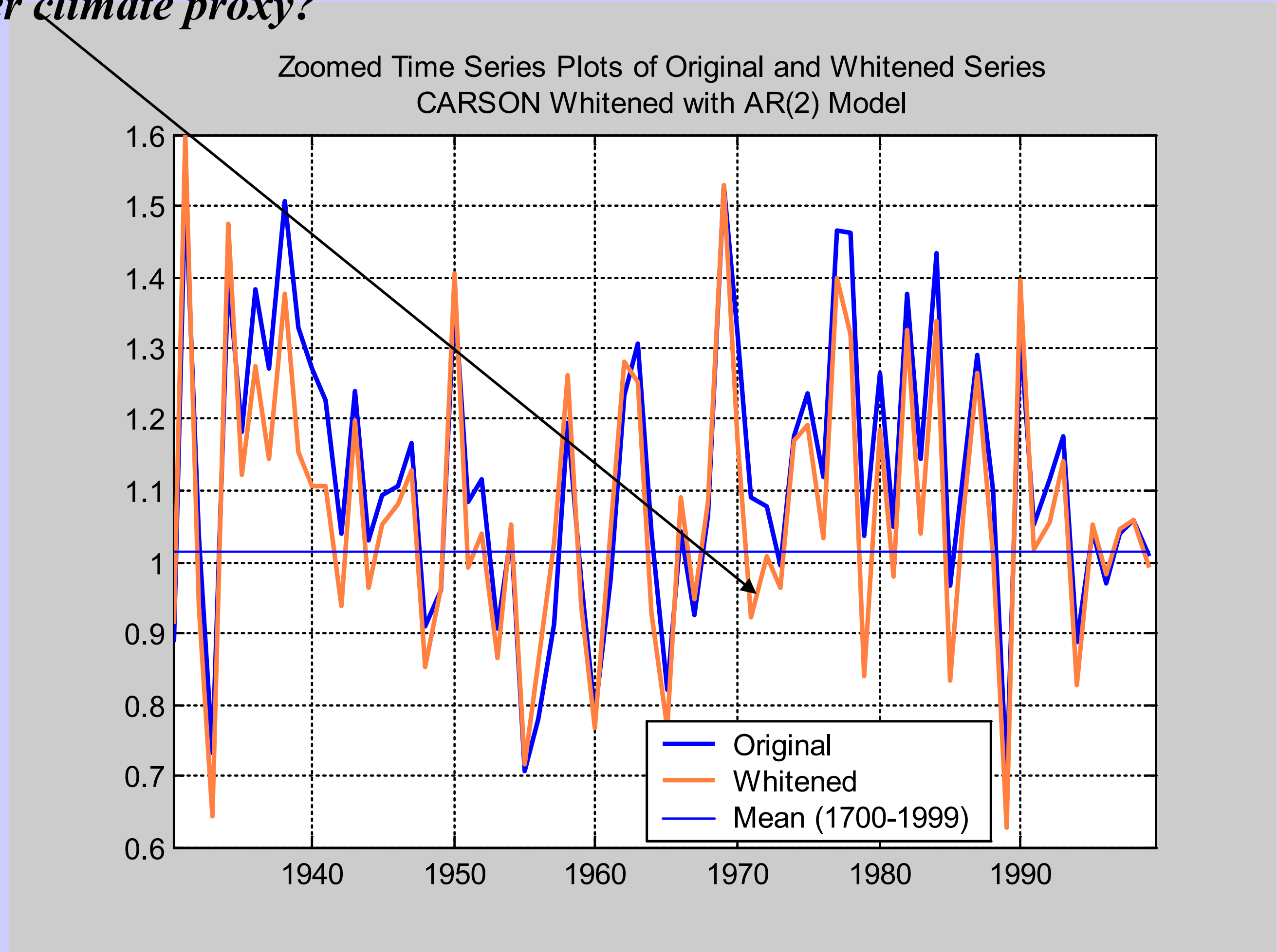


Prediction


Tree-Ring Index, Jackson Mtns, Nevada
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
Index
1900
1920
1940
1960
1980
2000
Year
?
Index
Future
Mean

Prewhitening

Better climate proxy?



Simulation

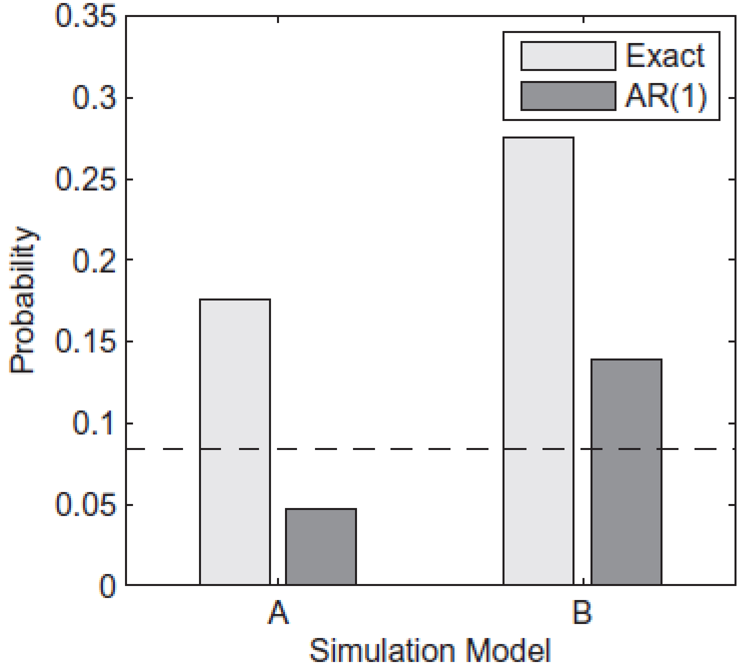


Fig. 4. Simulation-based probabilities of an 1100s low-flow in any 104-yr period. Simulations by exact simulation and first-order autoregressive (AR(1)) modeling of (A) 1906–1999 observed flows, and (B) 1906–2009 observed flows. Dashed line marks empirical probability from 1244-yr tree-ring reconstruction. See Section 4 for definition of 1100s low-flow.

- *Simulation to check how likely a "mid-1100s style" drought is from the time series properties of the observed Colorado River flows*
- *Meko et al, 2012, "Dendrochronology and links to streamflow"*

Journal of Hydrology 412–413 (2012) 200–209

Equations and terminology for some simple models

Structure

order

First-order autoregressive

AR(1) Model

Also called "random shock" or "noise" term

$$y_t + a_1 y_{t-1} = e_t$$

Residual

Mean-adjusted Time series

(mean has been subtracted)

Coefficient (autoregressive parameter)

First-order moving average

MA(1) Model

$$y_t = e_t + c_1 e_{t-1}$$

Mean-adjusted series

Residual (random)

coefficient

First-order autoregressive moving-average

ARMA(1,1) Model

AR order

MA order

Order p,q model, with p=1 and q=1

$$y_t + a_1 y_{t-1} = e_t + c_1 e_{t-1}$$

First-order autoregressive coefficient

First-order moving average coefficient

Steps in Modeling

1. **Identification** → *Structure and order?*
2. **Estimation** → *Values for parameters?*
3. **Checking** → *Are the residuals effectively without autocorrelation and is the model that successfully removes the autocorrelation as simple as possible?*

Identification

1. From diagnostic patterns of the ACF and **PACF**
2. Automatic methods based on minimizing variance of residuals

Partial Autocorrelation Function (pacf)

The pacf at lag k, or pacf(k), is equivalent to the autocorrelation at lag k of the residuals from fitting an AR(k-1) model to the time series

E.g., for pacf(3), the partial autocorrelation at lag 3

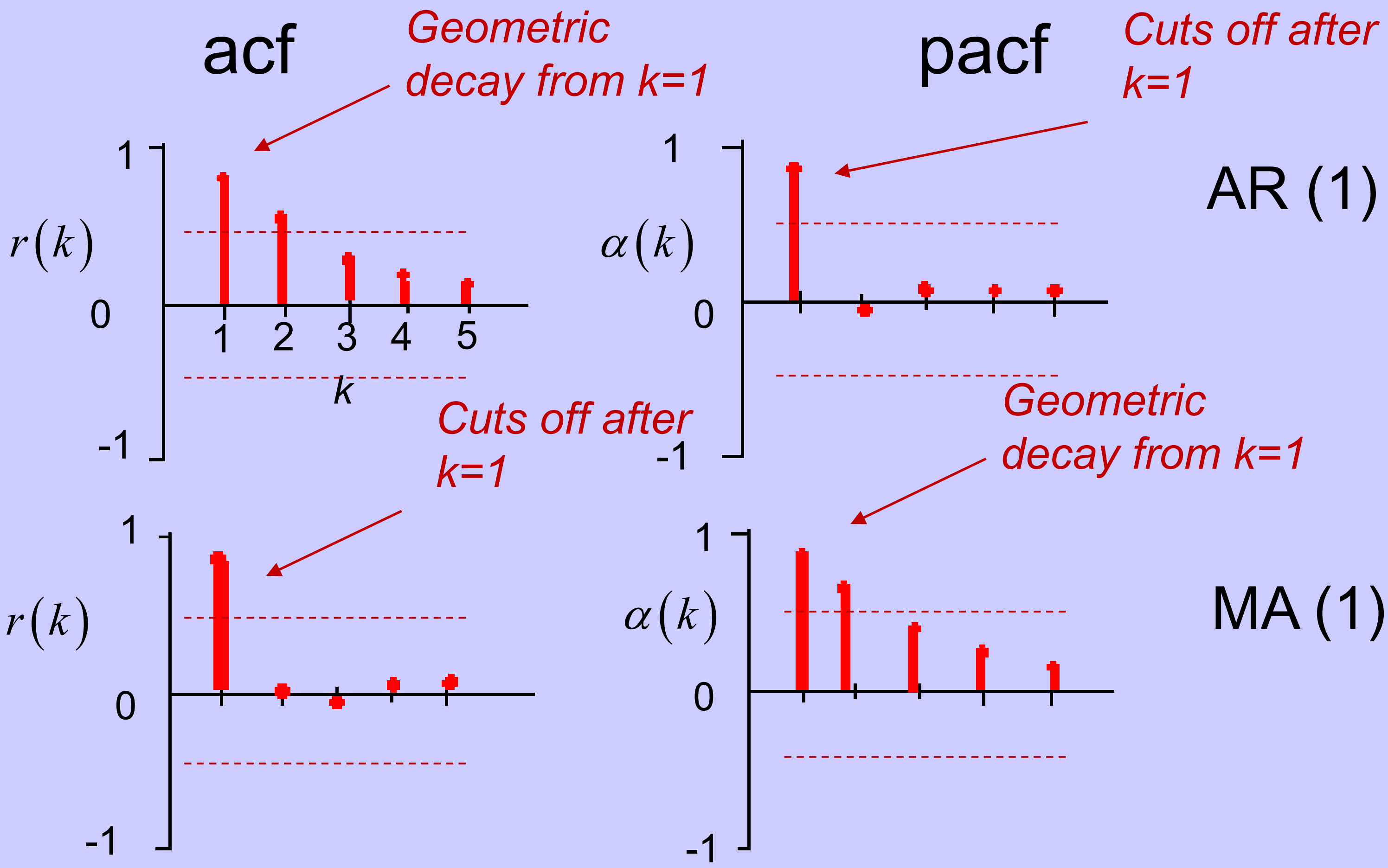
1. Fit AR(2) model to time series x(t)
2. Compute the acf of the residuals from that model → $r_e(k)$
3. Pacf(3) of x(t) is $r_e(3)$

Identification by acf and pacf patterns

Acf and pacf have characteristic decay patterns for specific-order ARMA models

1. <u>AR(1)</u>
   - Acf decreases geometrically
   - Pacf cuts off after lag 1
2. <u>MA(1)</u>
   - Pacf decreases geometrically
   - Acf cuts off after lag 1

Classic decay patterns of AR(1) and MA(1)



Inverse patterns of acf and pacf for these two simple models

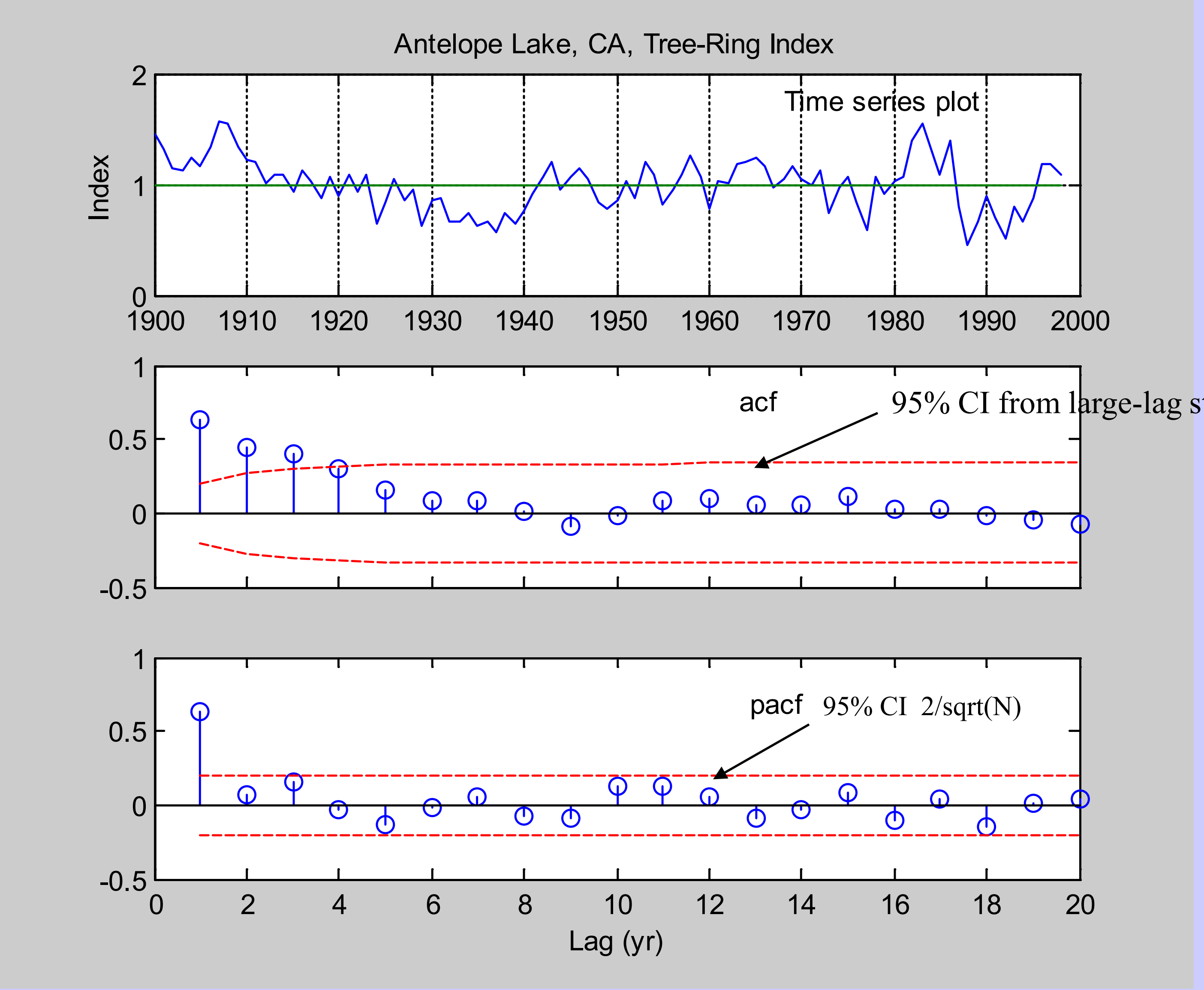
Classic decay patterns
(for higher-order AR(*p*) and MA(*q*) models)

Patterns can be difficult to identify, especially if noise component of model is large. But, some broad guidelines:

$$\mathrm{AR}\left(p\right): \text{ pacf "cuts off" after lag } k = p$$
$$\text{acf decays, possibly irregularly}$$

$$\mathrm{MA}\left(q\right): \text{ acf "cuts off" after lag } k = q$$
$$\text{pacf decays, possibly irregularly}$$

Example, a jeffrey pine tree-ring chronology from N. California



Identification by automated model selection

1. Fit various candidate models
2. Select “best” model by some objective criterion (e.g. FPE)
3. Follow up by checking model

Akaike's Final Prediction Error (FPE)

$$FPE = \frac{1 + n/N}{1 - n/N} * V$$

Number of estimated parameters

Sample size (series length)

Variance of ARMA residuals

Estimation and checking of model

1. Identification
2. Estimation
3. Checking

To get values of parameters; this process Is largely transparent to the user

Will cover in next lecture

Simulation

in context of simple AR & MA

Building a fake time series
that has specified statistical properties

1. White noise
2. MA(1)
3. AR(1)

$\Big\}$ . . .

White noise

mean

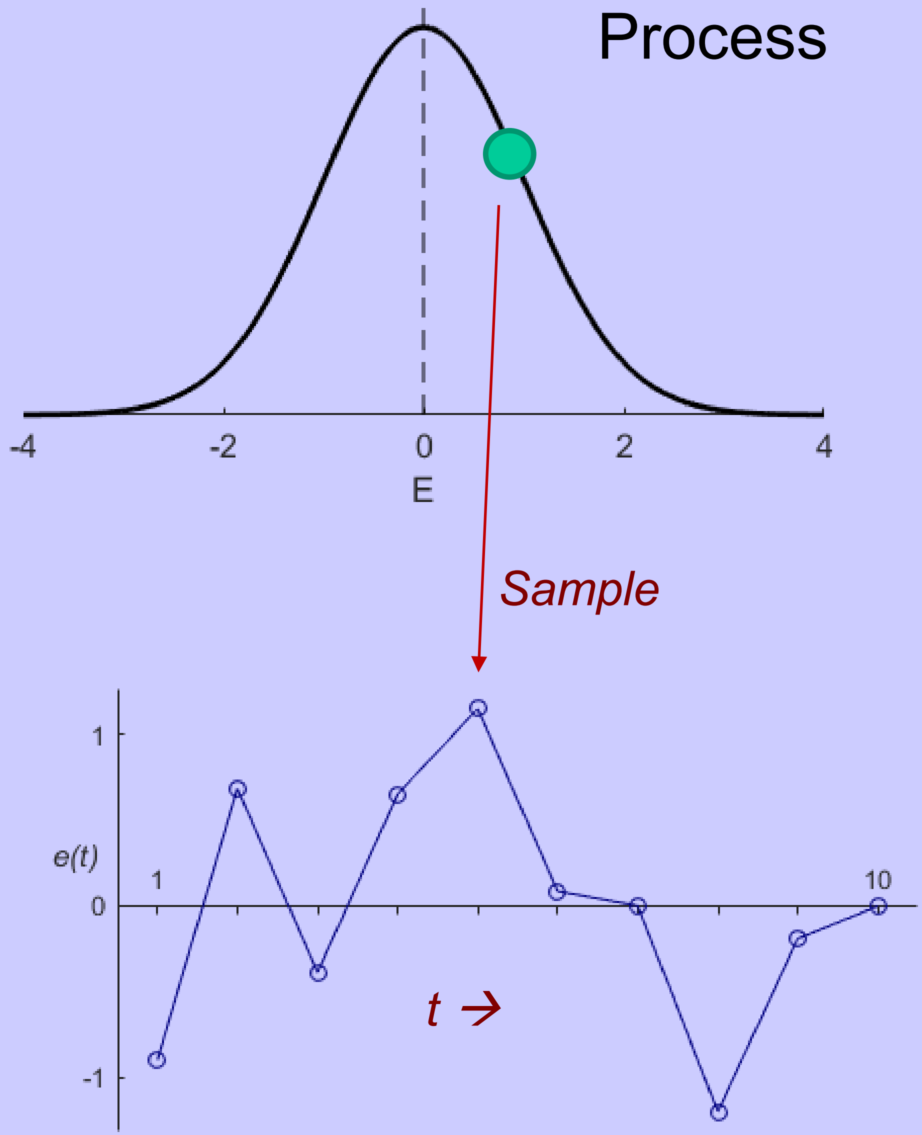
vari

$$E_t \sim I\,N\left(0, \sigma^2\right)$$

Process

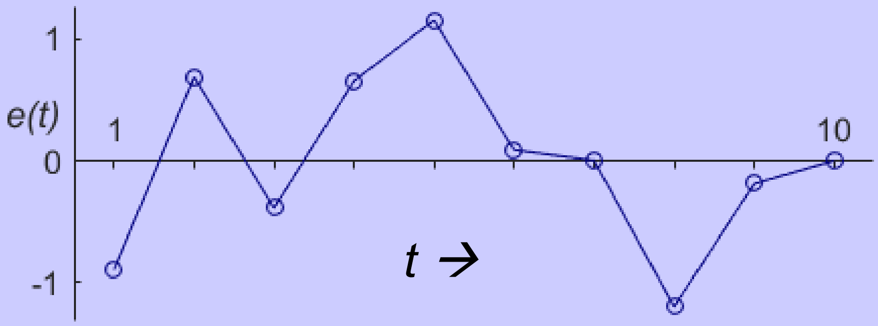
Normal

Independent



Sometimes "white noise" is defined as Gaussian, sometimes not. ARMA modeling usually assumes Gaussian white noise

MA(1)



- *Sample from white noise*
- *Apply model equation to get simulated time series*

$$y_t = e_t + c_1 e_{t-1}$$

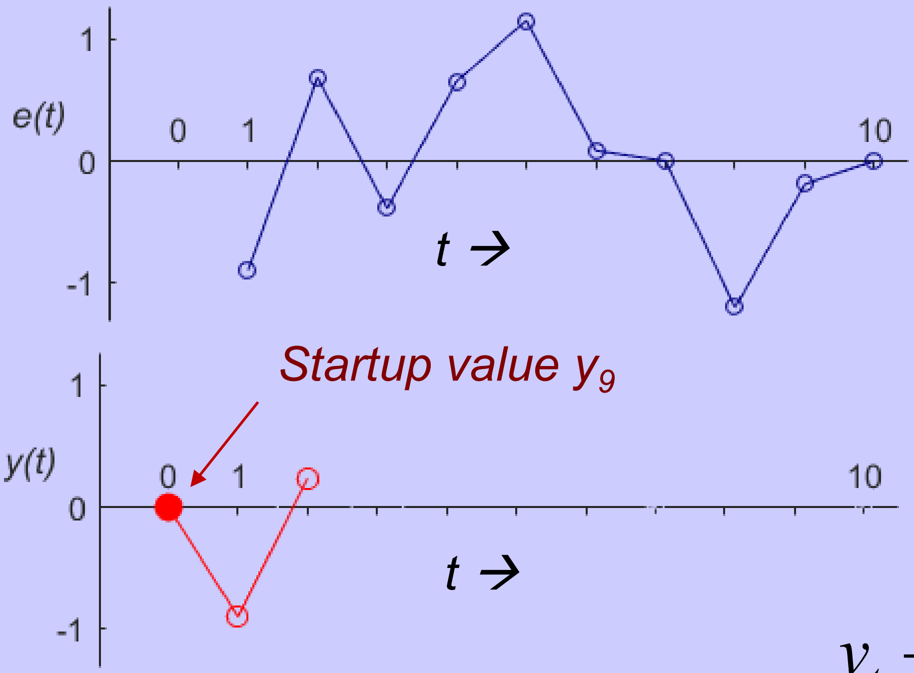
MA(1) parameter

"Memory" goes back only one previous time step

In practice, you might begin with the observed time series and use that to get the estimated moving average parameter before simulating

$-1 < c_1 < 1$ *Parameters of ARMA models are restricted to specified ranges that guarantee the model is stationary and invertible*

AR(1)



White noise, or "random shock"

y_t is built recursively from past values and noise term

$$y_t + a_1 y_{t-1} = e_t$$

$$y_t = -a_1 y_{t-1} + e_t$$

1) $y_1 = -a_1 y_0 + e_1$

2) $y_2 = -a_1 y_1 + e_2$

3) $y_3 = -a_1 y_2 + e_3$

$\vdots$

$-1 < a_1 < 1$ *Assures stationarity*