# Data analysis methods in weather and climate research

Dr. David B. Stephenson
Department of Meteorology
University of Reading
www.met.rdg.ac.uk/cag

9. Introduction to time series analysis & modelling

- Basic concepts
- Trend, periodic, and irregular components
- Filtering and smoothing
- Serial correlation
- Autoregressive time series models



## 9. Introductory reading on time series

- www.met.rdg.ac.uk/cag stats@reading
- C. Chatfield, The Analysis of Time-series: An Introduction, CRC press, 6th edition 2003.
- P.J. Brockwell and R.A. Davis, Introduction to Time Series and Forecasting, Springer Verlag, 1996.
- P. Bloomfield, Fourier Analysis of Time Series: An Introduction (Wiley Series in Probability & Statistics), 2000.



## 9. Some basic concepts

- Discrete time series = sequence of values $x_1, x_2, x_{3,} \ldots, x_n$ recorded at discrete times $t_1, t_2, t_{3,} \ldots, t_n$ - usually regular $t_n = n\Delta$
- Aims: to monitor, forecast, and control systems
- *Approaches: Time domain* or *frequency domain*
- Exploratory data analysis EDA (descriptive) → time series modelling (inferential)

## 9. Components of a time series

$$X_n = T_n + P_n + I_n$$

- Trend component

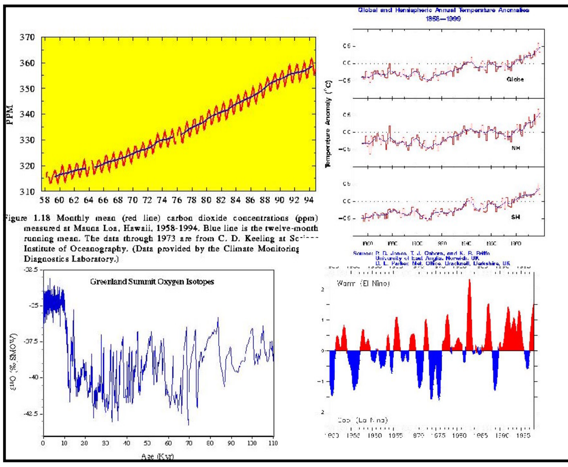
  *smooth long-term changes in mean*

- (Quasi-)Periodic component

  *cyclic terms e.g. annual cycle, diurnal cycle, etc*.

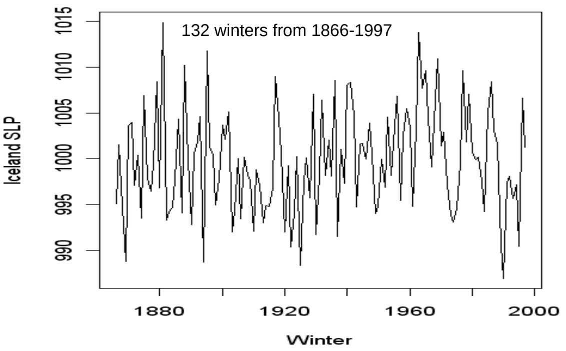
- Irregular component

  *noisy random part*




5

## 9. Time series of winter SLP at Iceland

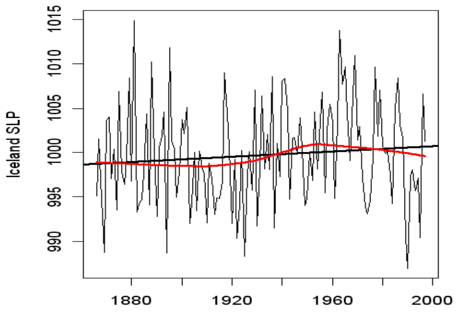



```
> plot(x,y,type="l",xlab="Winter",ylab="Iceland SLP")
```



6

## 9. Long-term time trend T




```
> abline(lsfit(x,y),lwd=3,col=1)
> lines(lowess(x,y),lwd=3,col=2)
```

## 9. Filtering and smoothing

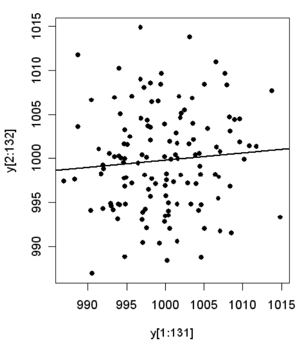
Lanczos filter with span k:

$$y'_t = \sum_{lag=-k}^{lag=+k} w_{lag} y_{t+lag}$$

- Low-pass (smoothing) e.g. MA(.), binomial,…
- High pass: y-y'
- Band-pass: y''-y'



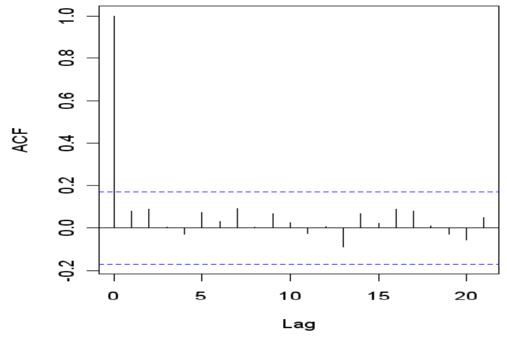

## 9. Serial correlation: SLP(t+1) versus SLP(t)




r1
= cor(y(t+1),y(t))
= 0.081

```
> plot(y[1:131],y[2:132],pch=16)
> abline(lsfit(y[1:131],y[2:132]),lwd=2)
```

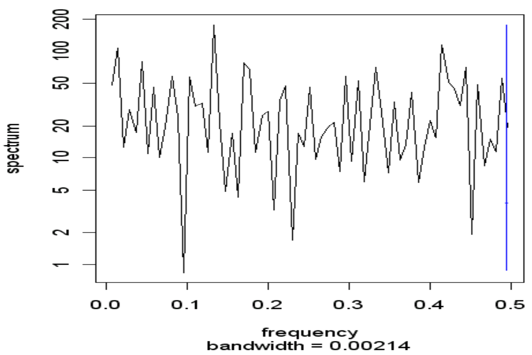
## 9. Autocorrelation function for Iceland




```
> acf(y,lag.max=20)
```



## 9. Unsmoothed power spectrum: periodogram




```
> spectrum(y)
```

## 9. Autoregressive and Moving Average models

AR(1) - Autoregressive order 1 time series model:

$$y_{t+1} = \beta y_t + \varepsilon_t$$

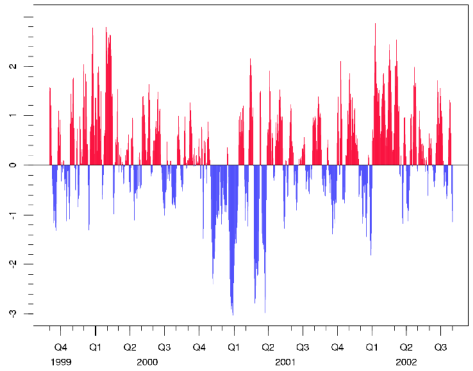
AR(3) time series model:

$$y_{t+1} = \beta_1 y_t + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \varepsilon_t$$

MA(2) time series model:

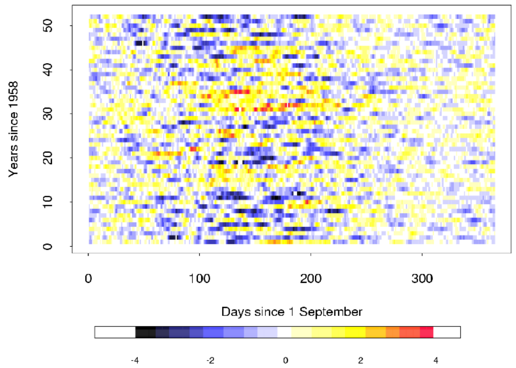
$$y_{t+1} = \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2}$$

## 9. Daily NAO index from Sep 1999 - 2002

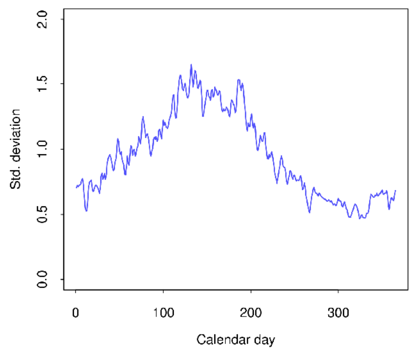



13

## 9.Persistence in the daily NAO index

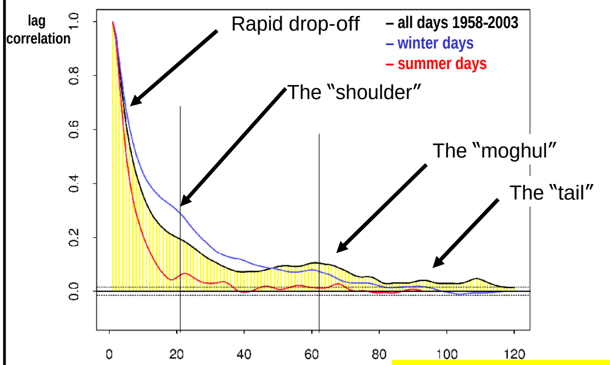



## 9.Varying variance (heteroskedasticity)



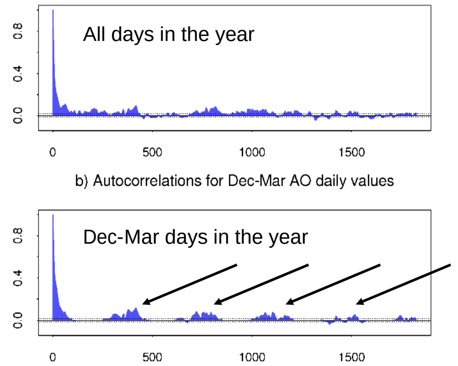

→ Almost 10 times the variance in winter than in summer

# 9. Daily NAO autocorrelation 1958-2003



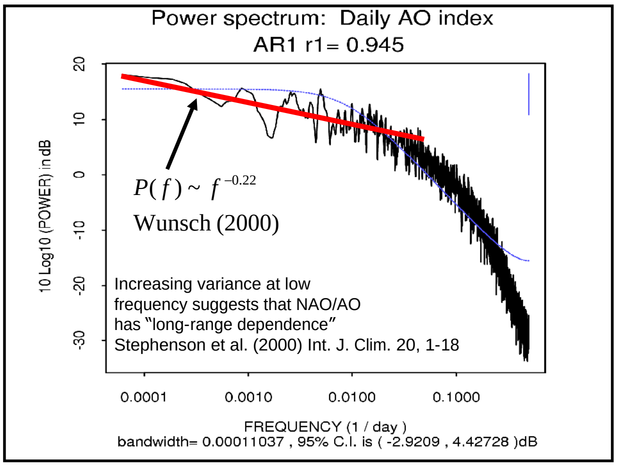

$$r_{lag} = cor(x_t, x_{t+lag})$$

# 9. Long-range correlation in daily NAO




→ Long memory from year to year

17



## 9. Covariations in sea temperatures and SLP

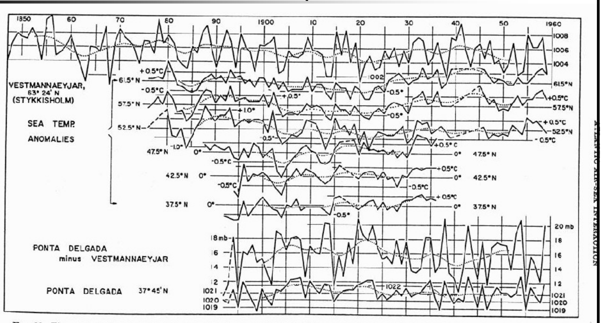



FIG. 11. Time series of annual sea temperatures in a profile Iceland-Azores. Top curve: Time series of annual sea-level pressure for Vestmannaeyar (extrapolation before 1881 based on Stykkisholm). Bottom curves: Time series of annual Ponta Delgada minus Vestmannaeyar sea-level pressure (zonal index), and time series of annual sea-level pressure for Ponta Delgada (from *Geophysica Norvegica*, **24**, No. 3, 1962).

J. Bjerknes, Geophysica Norvegica, **24**, No. 3, 1962.

## 9. Pressure and sea temperature associations

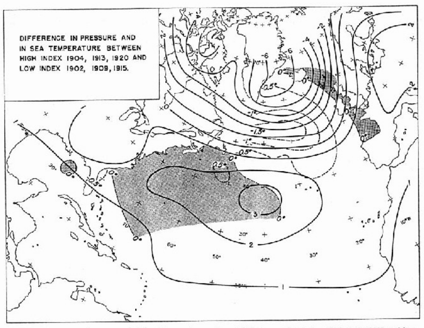
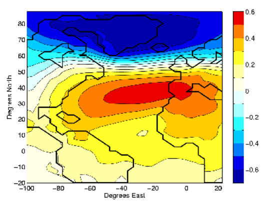



FIG. 14. Difference in pressure (mb) and in sea temperature (°C) between high-index 1904, 1913, 1920 and low-index 1902, 1909, 1915 (from *Geophysica Norvegica*, **24**, No. 3, 1962).

| Lat | r(NAO,T) |
|---|---|
| 61.5N | -0.34 |
| 57.5N | -0.82 |
| 52.5N | -0.2 |
| 47.5N | 0.18 |
| 42.5N | 0.33 |

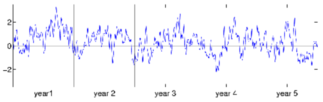
J. Bjerknes, Geophysica Norvegica, **24**, No. 3, 1962.

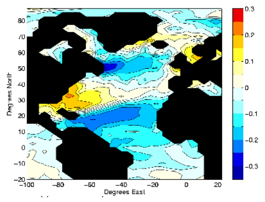
## 9. Granger causality in the N. Atlantic



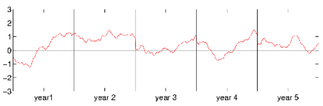
Mosedale et al. (2005)

Daily SLP NAO index



Daily SST tripole index




21

## 9. Causality analysis in macro-economics

**Granger Causality**

If the past history of one variable, X, gives extra predictive information for another variable, Y, then X "Granger causes" Y

"Testing for causality and feedback," *Econometrica*, 37, 1969, 424-438.

Clive W.J. Granger receiving a Nobel Prize for econometric time series analysis from His Majesty the King of Sweden at the Stockholm Concert Hall. 10 December 2003

http://www.nobel.se/economics/laureates/2003/

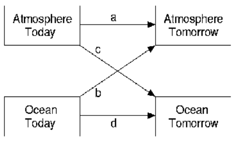


## 9. Granger causality testing

- Fit a Vector AutoRegressive VAR(p) time series model to the data:

$$N_t = \sum_{i=1}^{p} a_i N_{t-i} + \sum_{i=1}^{p} b_i S_{t-i} + \varepsilon_t$$

$$S_t = \sum_{i=1}^{p} c_i N_{t-i} + \sum_{i=1}^{p} d_i S_{t-i} + \eta_t$$



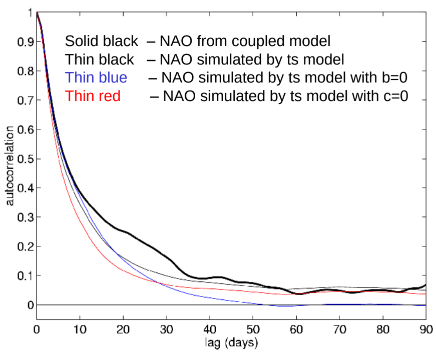

- Determine the best order p by minimising the Schwarz criterion
- Significance test for b=0 by comparing the log-likelihood ratio

$$L = N\left(\log\left|\Omega_{b=0}\right| - \log\left|\Omega\right|\right)$$

to the chi-squared distribution with p degrees of freedom



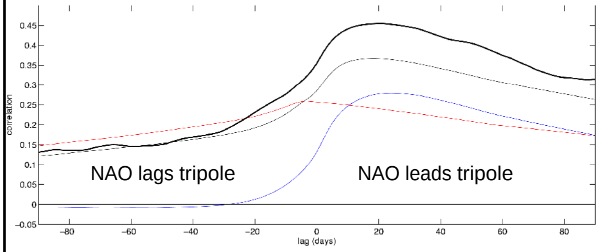
## 9. Autocorrelation of simulated NAO




Need o→a interaction to get long tail in NAO

## 9. Cross-correlation of NAO and tripole

Solid black – NAO from coupled model
Thin black – NAO simulated by ts model
Thin blue – NAO simulated by ts model with b=0
Thin red – NAO simulated by ts model with c=0

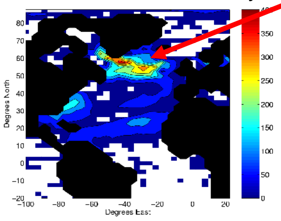



Need a→o interaction to get cross-correlations



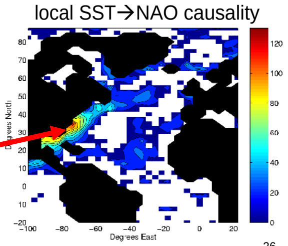

## 9. Causality for NAO and local SSTs?

NAO→local SST causality




NAO causal effect on
N. Atlantic SSTs
(as found by Bjerknes)

local SST→NAO causality




Interesting causal
effect of Gulf stream
SSTs on NAO




## 9. Summary

- Time series can have fascinatingly rich serial structure
- EDA exists for exploring the different time series components: trend+periodic+irregular
- Regression models can be used to MODEL the process that generated the time series and hence provide insight.
- Lots of good books on time series analysis that are worth reading: Chatfield, Brockwell and Davis, Bloomfield, Priestley, …

## Probability Problem Solution

Two types of simple event: G=girl B=boy

Event space: GG GB BG BB

P(GG|(GB or BG or GG))

=P(GG)/P(GB or BG or GG)

=0.25/3*0.25=1/3

Three types of simple event: B=boy R=girl rare name C=girl not rare name

P(B)=1/2 P(R)=p/2 p(C)=(1-p)/2

Event space: RR RC RB CC CR CB BR BC BB

P(RR or RC or CR|(RR or RC or CR or RB or BR))

=(2-p)/(4-p) (but note that there are many girl's names so p is often small!)