Analysis of temporal phenomena

Djemel Ziou

Computer Science

Plan

- Temporal phenomena
- Univariate time series
- Estimation of the stationary autoregressive model
- Application
- Hierarchy of the autoregressive model
- Estimation of MA and ARMA models
- References

Temporal phenomena

- Dynamic phenomena are characterized by:
  - past, present, and future chronological order
  - change.

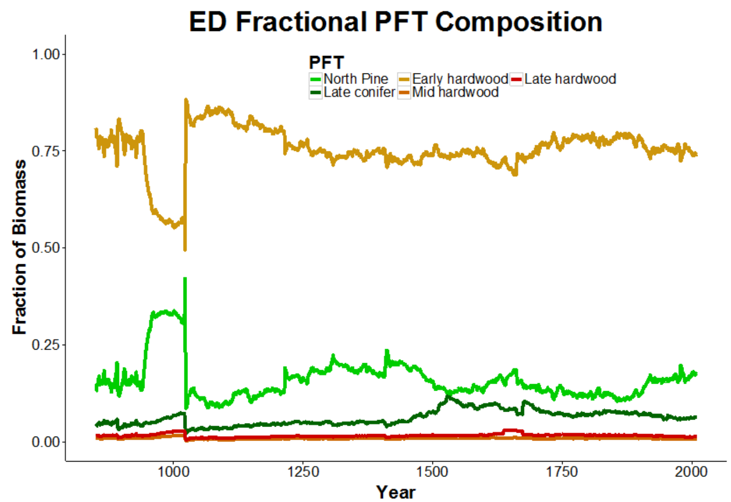
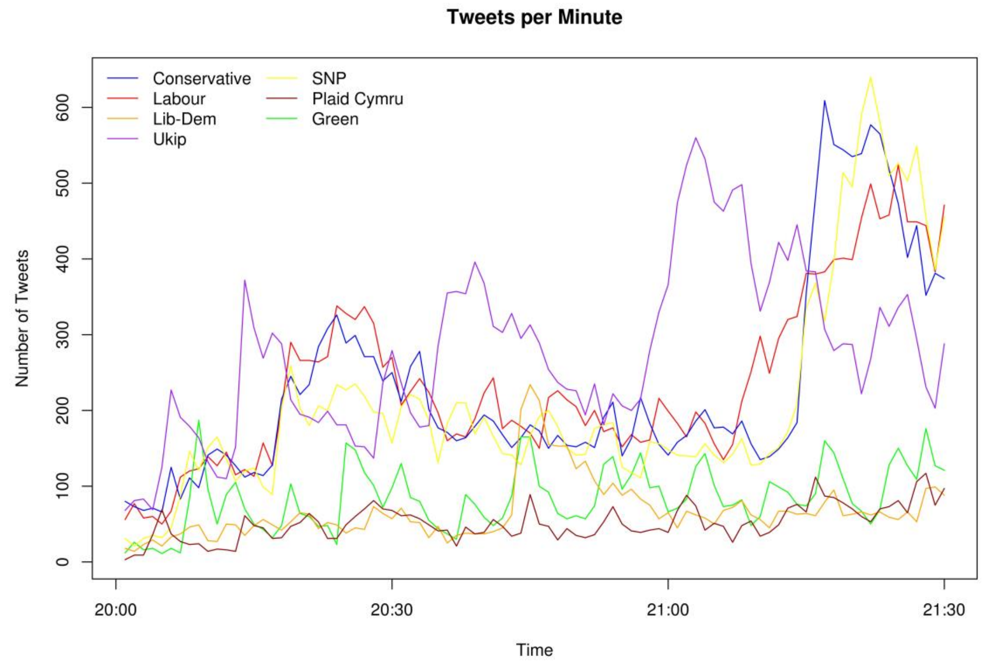


Figure 3: An example of forest composition through time in a single ecosystem model (ED).



- Origin of the uncertainty:
  - the variability of these phenomena over time can be interpreted as being random
  - measurement errors.

Temporal phenomena

Examples

- It is proposed to predict the number of calls received on the phone by u_b from u_a)
- Time is subdivided into intervals $d_1=]0, t_1]$, $d_2=]t_1,t_2]$...
- Assumptions
  - the number of calls in disjointed time intervals are independent
  - the probability of a call in a time interval is proportional to the length of that interval, the proportionality coefficient being λ
  - the probability of more than one call in a time interval is negligible.
- We show that N(t) is Poisson random variable having an intensity equals to λt, thus $P(N(t)=k)=(\lambda t)^k e^{-\lambda t}/k!$ (Poisson distribution).

Temporal phenomena

- Let the random process $\{N(t):t= t_1, t_2 \ldots\}$ representing the number of calls during the time interval $]0, t_1],]t_1, t_2], \ldots$
- Let $\bar{d}_{ab}$ the mean period between two calls from u_a to u_b, so the average of calls is $\bar{\lambda}_{ab} = 1/\bar{d}_{ab}$
- The probability $P_k^{ab}(t_n)$ that u_a receive k calls from u_b within the time interval $]t_n, t_{n+1}]$ $(d^{ab}_n = t_{n+1} - t_n)$ is:

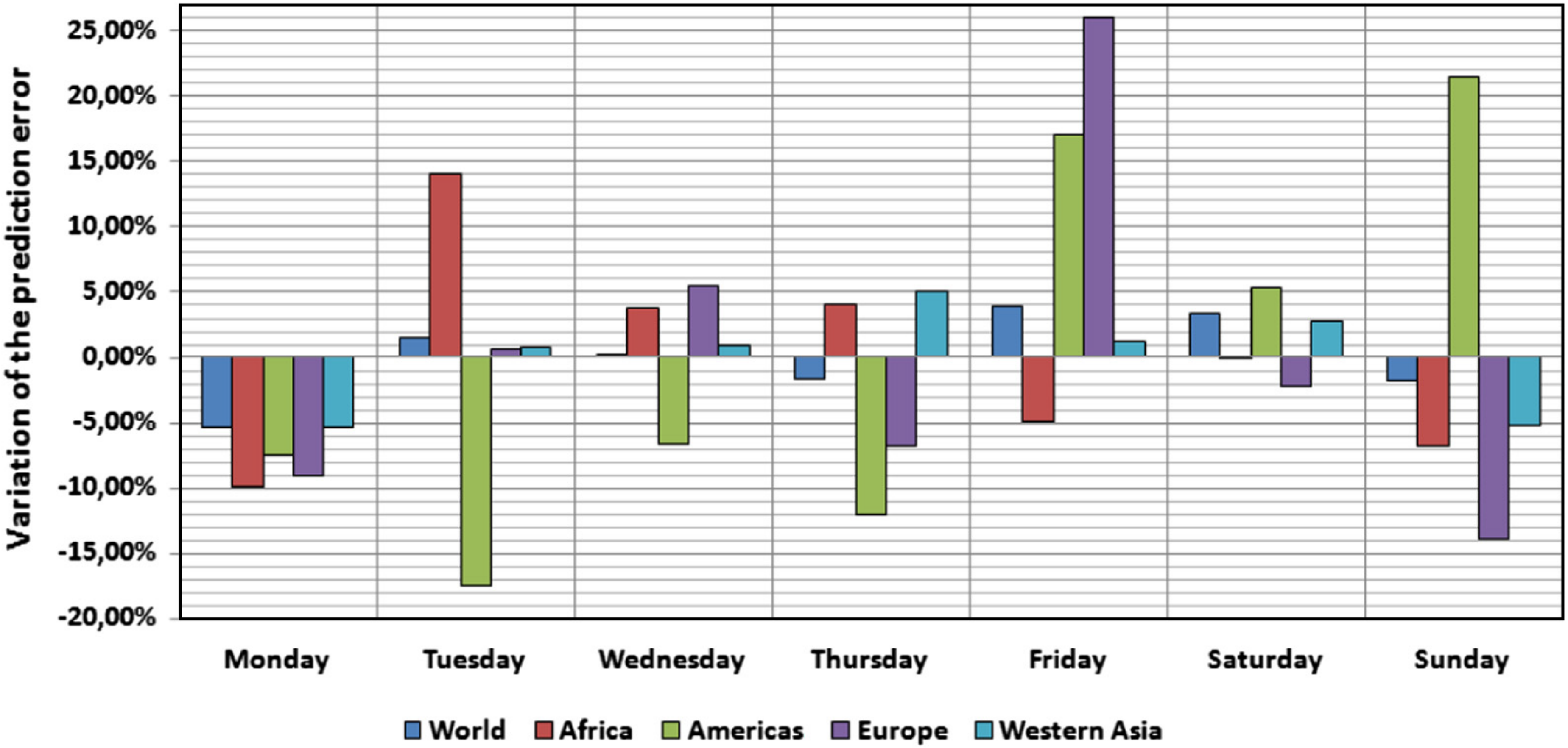
$$P_k^{ab}(t_{n+1}) = P(N^{ab}(t_{n+1}) - N^{ab}(t_n) = k) = (\bar{\lambda}_{ab} d_n^{ab})^k e^{-\bar{\lambda}_{ab} d_n^{ab}} / k!$$

- Let $\bar{\lambda}_{ab}(t_n)$ the average of calls from u_a to u_b within $]t_{n-1}, t_n]$ and

$\lambda_{ab}(t_n) = 1/d_{ab}^{(n)}$ the number of calls during the same interval. The average of class which will be received at t_{n+1} is

$$\bar{\lambda}_{ab}(t_{n+1}) = \alpha \bar{\lambda}_{ab}(t_n) + (1-\alpha)\lambda_{ab}(t_n), \quad \alpha \in [0,1]$$

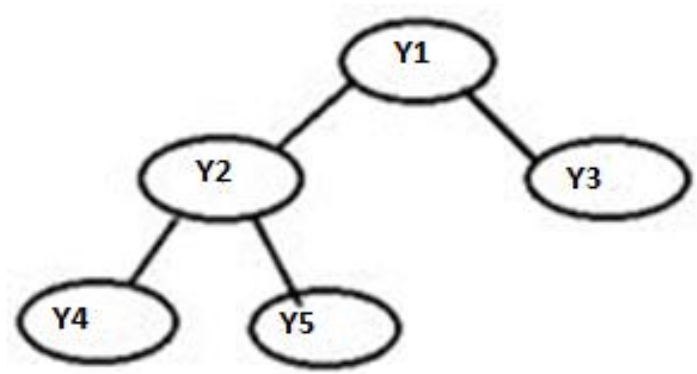
Temporal phenomena



Prediction mean error by continent and day.

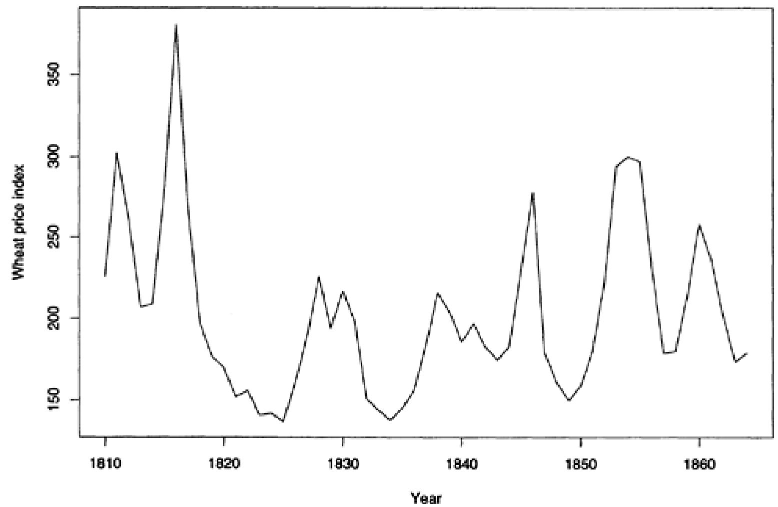
Temporal phenomena

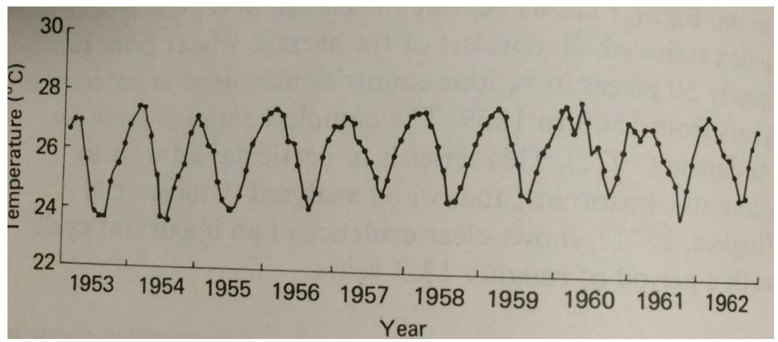
- Distributions of temporal phenomena
  - Let a discrete-time random process $\{Y_t\!:t=0..T\}$, the joint distribution is $[Y_0,Y_1\ldots Y_T]=[Y_0][Y_1|Y_0][Y_2|Y_1,Y_0]\ldots[Y_T|Y_{T-1}\ldots Y_0]$
  - it is difficult to specify all the interactions between the variables $Y_0,Y_1\ldots Y_T$. Thus, we make assumptions about the links between variables,
    $[Y_0,Y_1\ldots Y_T]=[Y_0]\prod_{t=1}^{T}[Y_t|Y_{t-1}]$

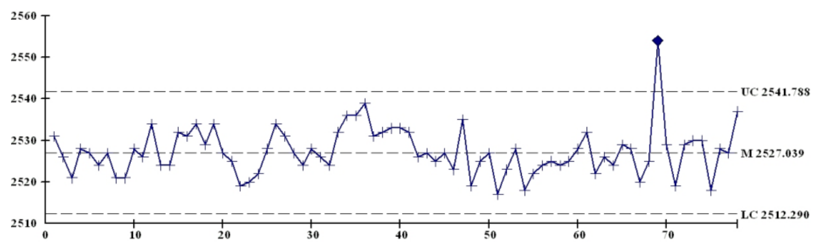


Univariate time series

- Time series is a discrete time RP process $\{Y_t : t=0..T\}$,
  - economy : Beveridge wheat price index between 1810 and 1864. Granularity = year.
  - physiques : temperature between 1953 and 1962 in Recife (Brazil). Granularity: month.
  - quality control: measuring the conformity of a product with the specifications. Granularity = the time required to perform the measurement

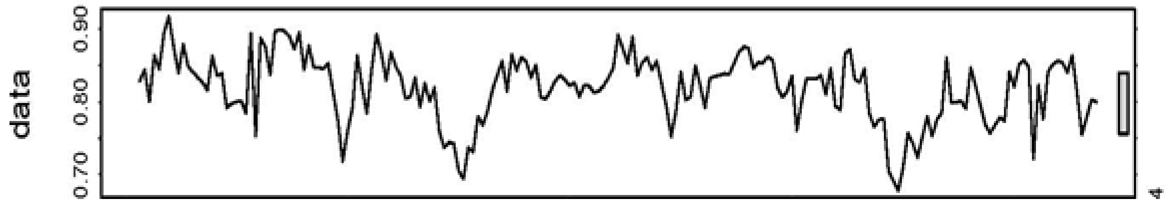




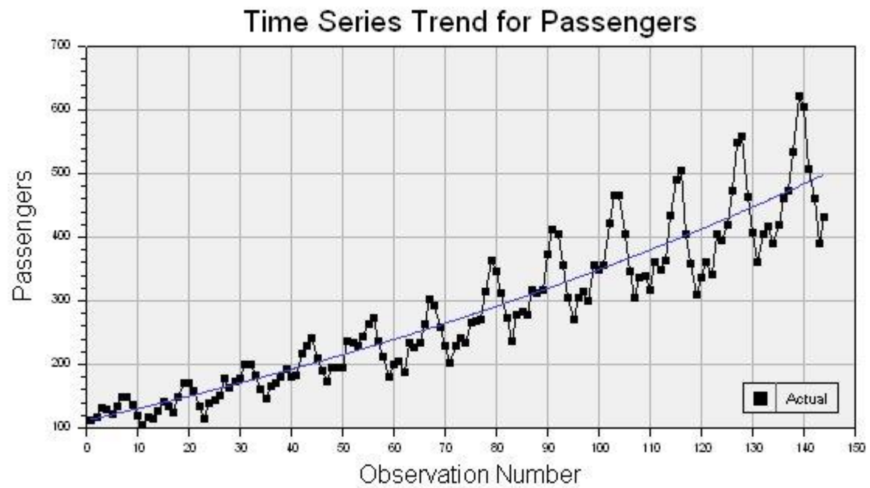


Univariate time series

- forest: pine plantation per year



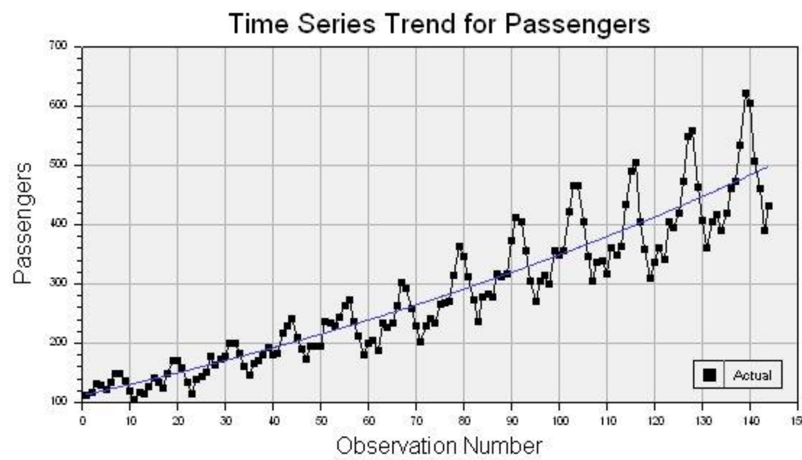
- transport: number of passengers



Univariate time series

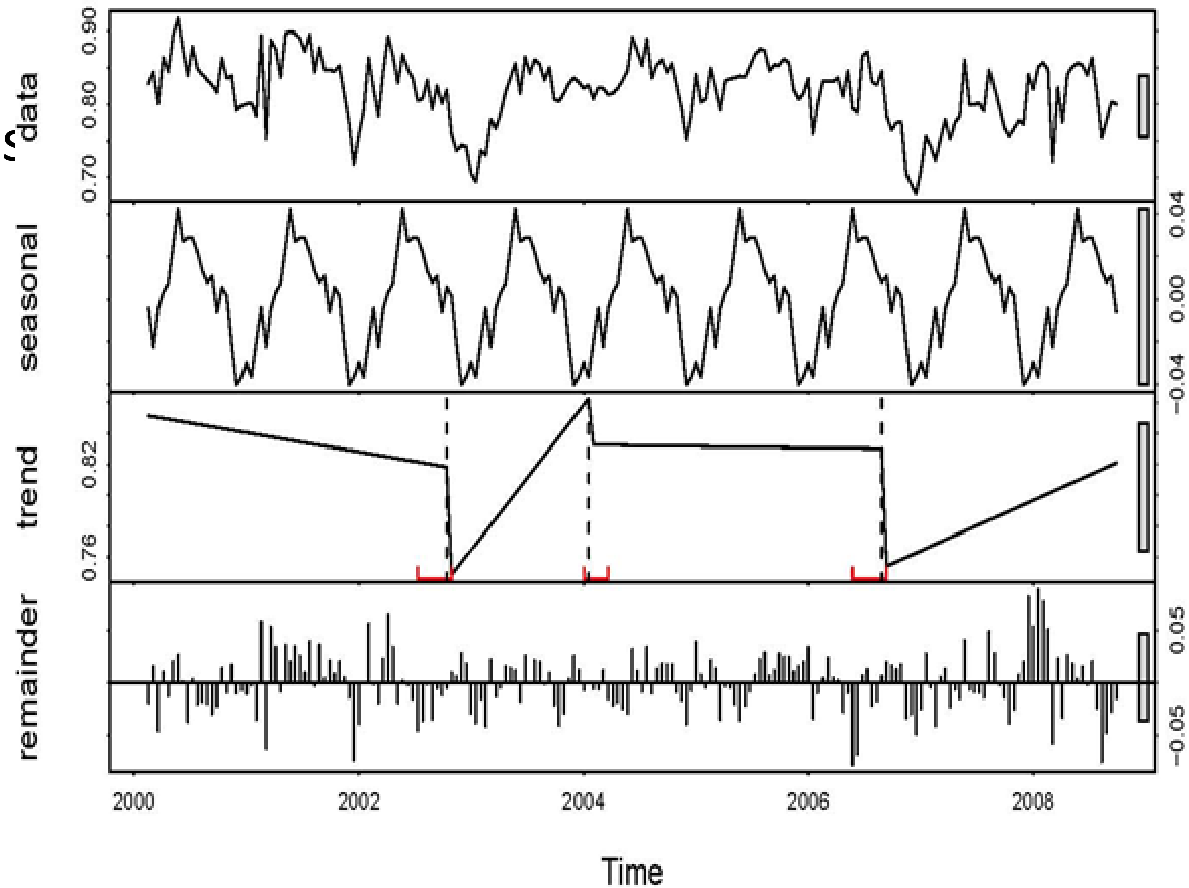
- Discreet: observations at specific times (the step is often constant).

- Classes of time series
  - univariate
  - stationary
  - linear
  - ...



Univariate time series

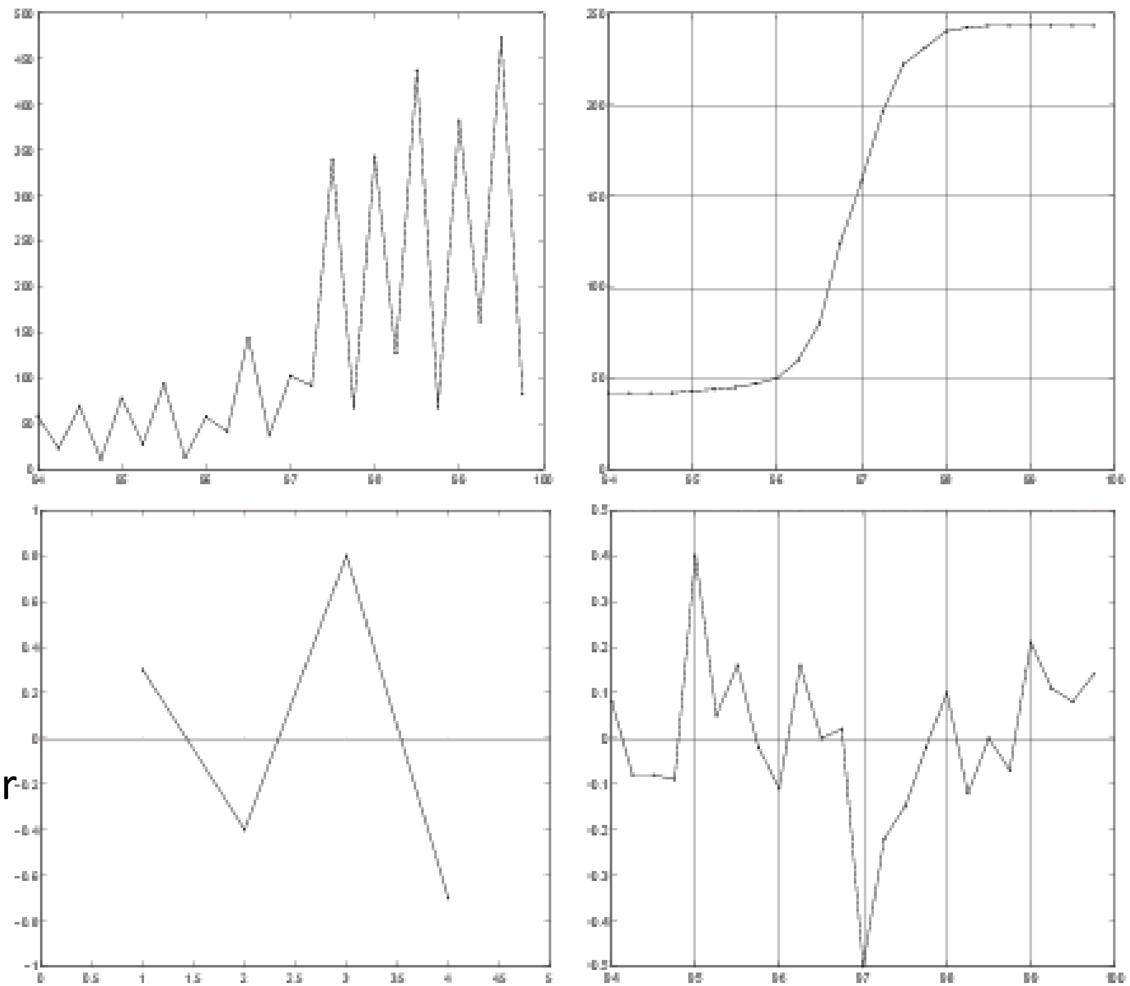
- Decomposition of a time series
  - seasonal effect
  - trend
  - random component.



- Approach to study TS
  - identify the components of the TS
  - isolate the random component
  - model the random component
  - after the prediction using the random component, add the other components.

Univariate time series

- Example: quarterly sale of sunscreen in France between 1994 and 2000



Trend
The increase in the sale is due to an advertisement. From 1998 the advertising had no effect.

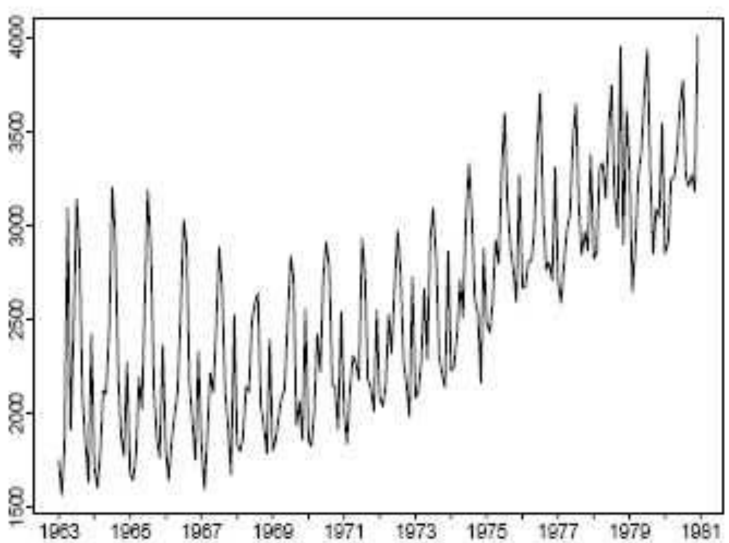
Seasonal factor
Increase of 30% in the 1st quarter (winters) and 80% 3rd quarter (summer holidays).

Random fluctuation
30% increase in sales in 1995 due to a promotion. 50% drop in sales due to a strike.

Univariate time series

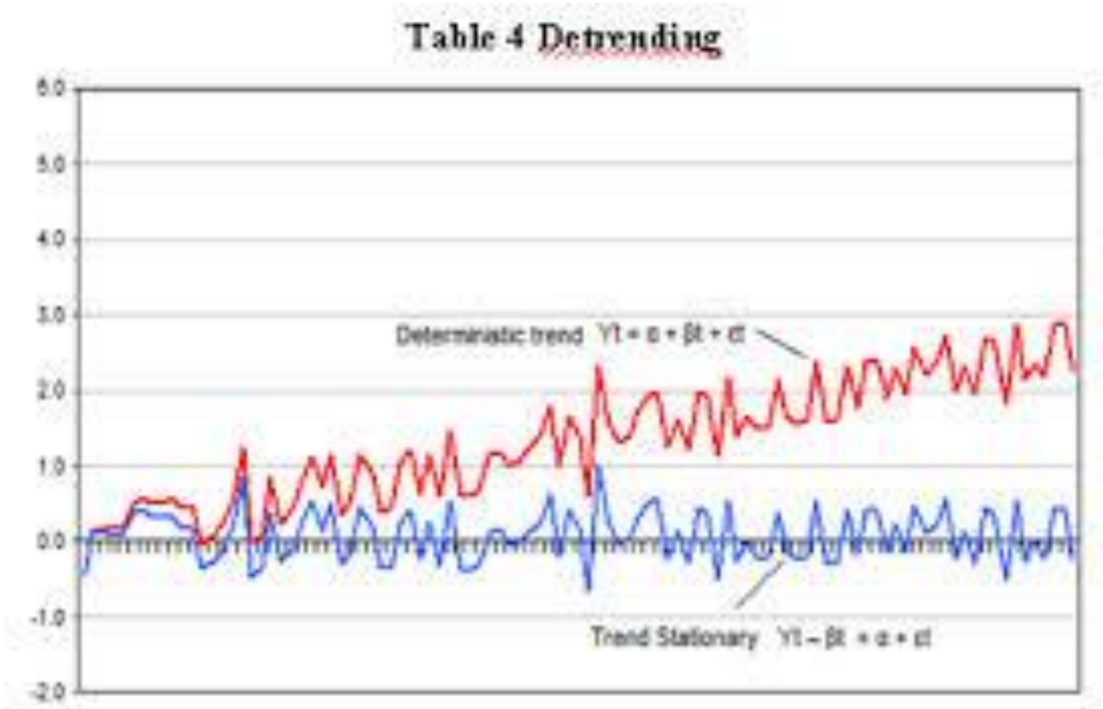
- correcting the data: in practice, you have to look at the SC data to:
- evaluate the missing data as, for example, there was no data acquisition during a day or a week ...
- bring back the data at intervals of the same length, for example, the data was collected on Monday, Wednesday, Thursday ... We have daily data and we want to analyze on the scale of the month ...
- perform data transformations when needed, for example `` stationarising '' the data, eliminate a seasonal effect ...

Evolution of passenger traffic
SNCF between 1960 and 1980



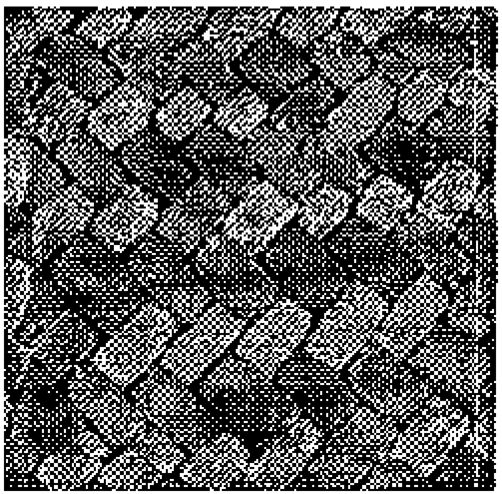
Univariate time series

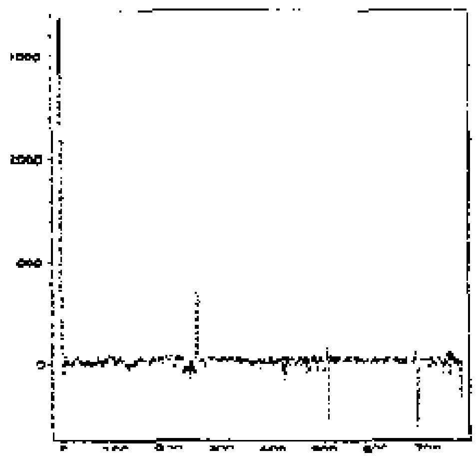
- Example of transformation
  - Take the logarithm; $X_t = m_t s_t e_t$
  - Reduction of fluctuations (bruit, e seasonal effect) by using a smoothing
  - Use the derivative of time series $Y_t = DX_t$, where D is a differentiation operator..



Univariate time series

- Statistics of the TS
  - A discrete time RP $\{Y_t{:}t=0..T\}$,
  - The expectation $\mu_t = E(Yt)$
  - the covariance $C_Y(t,r)=Cov(Y_t,Y_r)=E((Y_t-\mu_t)(Y_r-\mu_r))$
- Covariance
  - It describes how the RP co-varies as a function the lag between two variables
  - symmetry $C_Y(r,t)=C_Y(t,r)$
  - variance $C_Y(t,t)$
  - autocorrelation $R_Y(t,r)=C_Y(t,r)/(C_Y(t,t)C_Y(r,r))^{0.5} \in [-1,1]$





First raw of covariance matrix

Univariate time series

- Stationarity
  - strong: $[Y_{t1} \dots Y_{tm}]$ and $[Y_{t1+\tau} \dots Y_{tm+\tau}]$ have the same distribution for $\tau = \pm 1, \pm 2 \dots$

  - weak : moments 1 and 2 exist and do not dependent on time.

    Let N realisations $\{y_t^{(1)}\} \dots \{y_t^{(N)}\}$ of RP $\{Y_t : t=0..T\}$, we have

    $E(Y_t) = \int_{\Omega} y_t \, f(y_t) \, dy_t = \operatorname{plim}_{N \to \infty} \sum_{n=1}^{N} y_t^{(n)}/N$

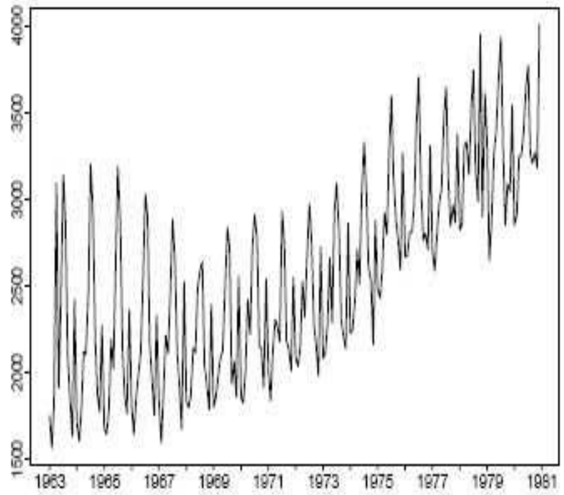
    $E((Y_t - \mu_t)(Y_{t-k} - \mu_t)) = \int_{\Omega} \dots \int_{\Omega} (y_t - \mu_t)(y_{t-k} - \mu_t) f(y_t \dots, y_{t-k}) \, dy_t \dots dy_{t-k} = \operatorname{plim}_{N \to \infty} \sum_{n=1}^{N} (yt^{(n)} - \mu_t)(yt_{-k}^{(n)} - \mu_t)/N$

    RP are weak stationary means

    $E(Y_t) = \mu < \infty$

    $E((Y_t - \mu_t)(Y_{t-k} - \mu_t)) = C(k) < \infty$

    Autocorrelation of a stationary RP tend to 0 when k become large.



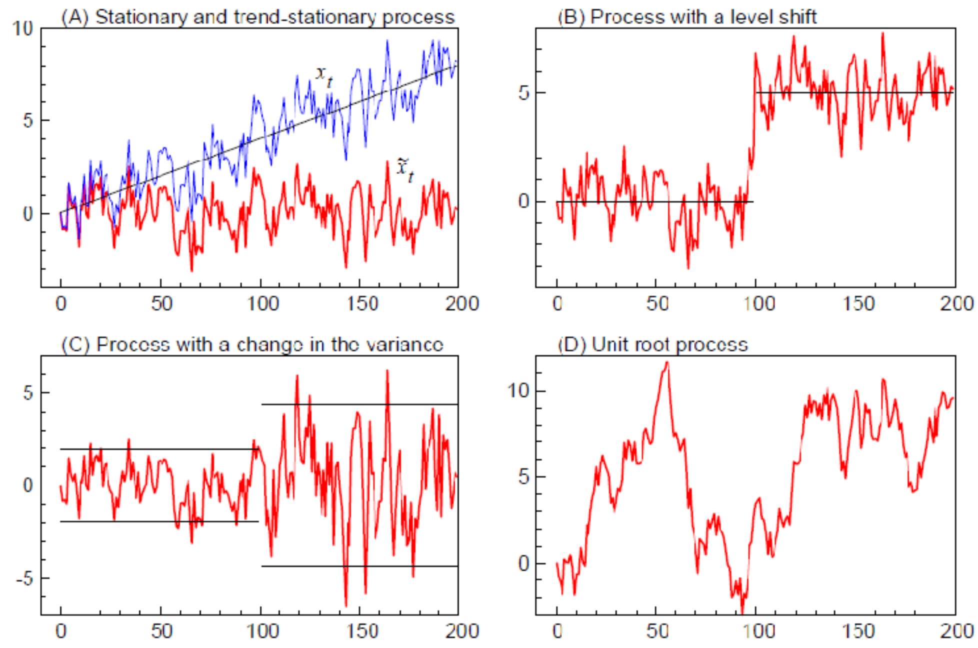


Figure 1: Simulated examples of non-stationary time series.

Univariate time series

- Examples
  - $Y_t = \mu + \varepsilon_t$ is a weak stationnairy because $E(Y_t) = \mu + E(\varepsilon_t) = \mu$, $E((Y_t - \mu)(Y_{t-k} - \mu)) = \sigma^2$ if k=0 and 0 otherwise, with $\varepsilon_t \sim N(0, \sigma^2)$ and the ε_t are independents
  - $Y_t = \mu t + \varepsilon_t$ *is not weak stationnary because* $E(Y_t) = \mu t + E(\varepsilon_t) = \mu t$, E((Y$_t$- *are* independents
- Strong stationarity and $C_Y(t,t)$ is finite implies the weak stationarity. The opposite is not true.

- Ergodicity
  - Expectation $E(Y_t) = \mathrm{plim}_{T \to \infty} \sum_{t=1}^{T} y_t^{(n)} / T$
  - covariance $E((Y_t - \mu_t)(Y_{t-k} - \mu_t)) = \mathrm{plim}_{T \to \infty} \sum_{t=k}^{T} (yt^{(n)} - \mu_t)(yt_{-k}^{(n)} - \mu_t) / N$
  - ergodicity means that the RP moments can be estimated from only one realization (n).
  - …

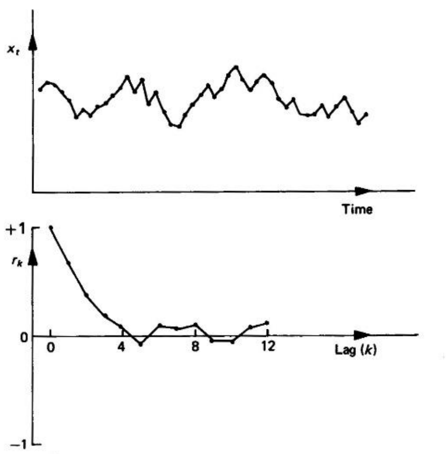
Univariate time series

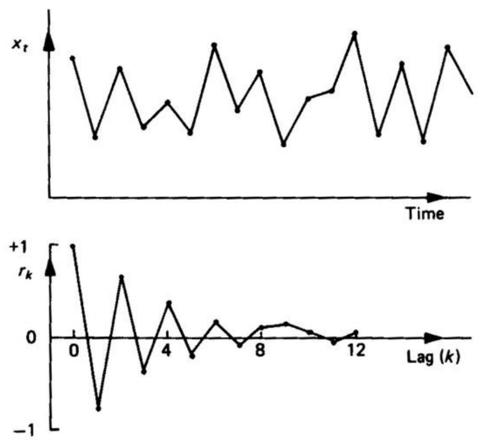
- Example
  - $Y_t = \mu_1 + \varepsilon_t$ if z=0 et $Y_t = \mu_2 + \varepsilon_t$ si z=1 is not ergodic. Indeed, $p(Y_t)=p(Y_t, z=0)+p(Y_t, z=1)=p(Y_t| z=0)p(z=0)+p(Y_t| z=1)p(z=1)$. Thus, $E(Y_t)=E(Y_t| z=0)p(z=0)+E(Y_t| z=1)p(z=1)= \mu_1 p(z=0)+ \mu_2 p(z=1)$

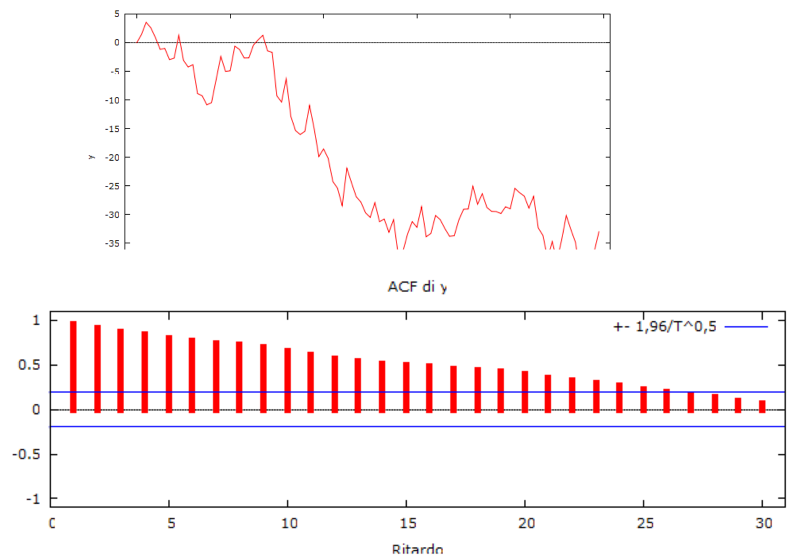
The mean $\sum_{t=1}^{T} y_t^{(n)}/T= \mu_1 + \sum_{t=1}^{T} \varepsilon_t /T$ if z=0 and $\mu_2 + \sum_{t=1}^{T} \varepsilon_t /T$ if z=1. The mean converge to one of the two mu. The RP is not ergodic.

Univariate time series

- Moments estimation in the case of stationary TS
  - let the observation $y_1 \ldots y_N$
  - mean: $\mu = \sum_{t=1}^{N} y_t / N$
  - covariance: we form pairs $(y_1, y_{1+k}), \ldots (y_n, y_{n+k})$
    $C_Y(k) = E((Y_t - \mu)(Y_{t-k} - \mu)) = \sum_{t=1}^{n} (y_t - \mu)(y_{t+k} - \mu)/n$
  - correlogram: graphical representation of the autocorrelation coefficients as function of the lag.







Univariate time series

- Example of correlogram use

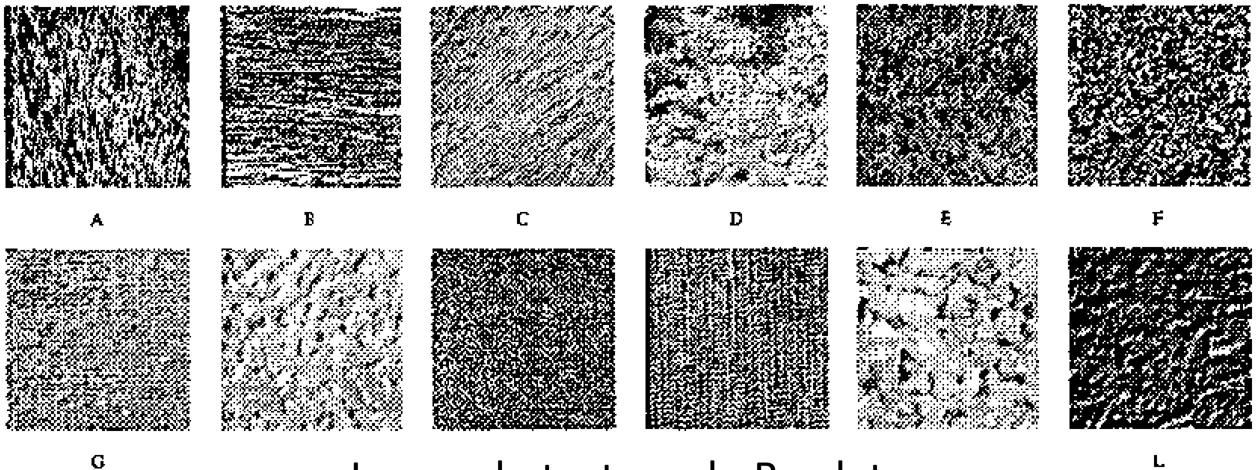
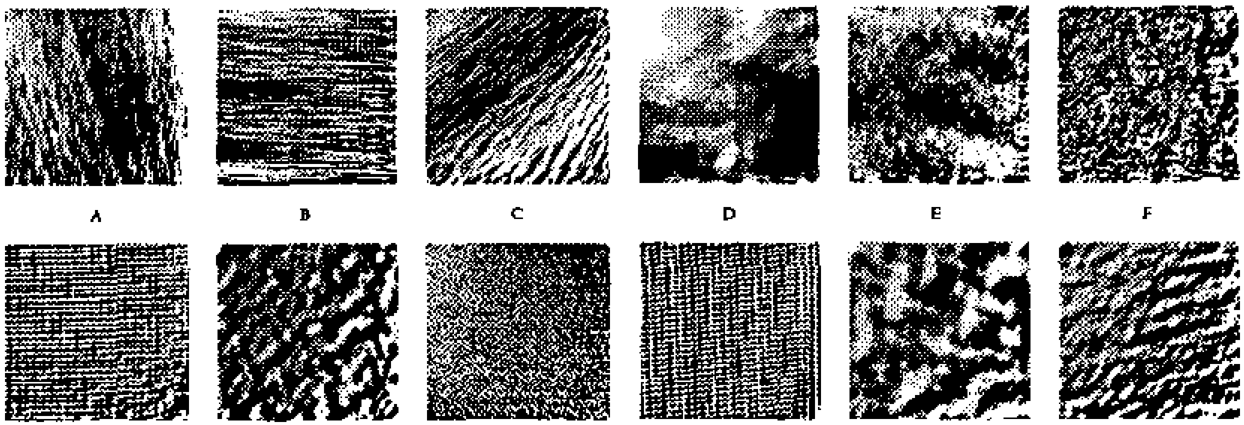


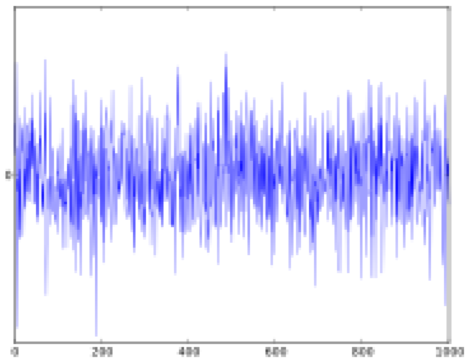
Image de texture de Brodatz



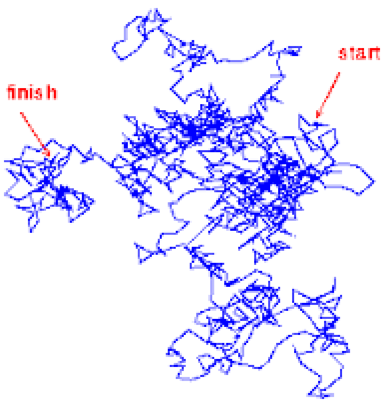
Autocorrélation

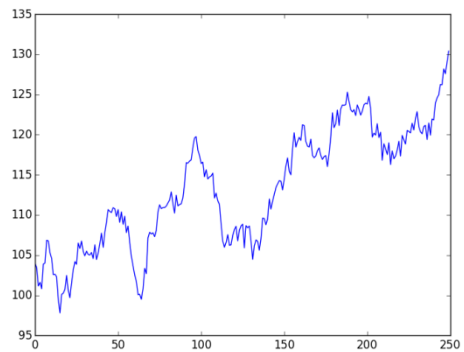
Univariate time series

- Basic TS
  - White noise $\{Y_t, t=1, 2\ldots\}$, $Y_t=e_t$, where e_t are independent, $\mu=0$, covariance $C_e(\tau) = \sigma_e{}^2\ si\ \tau = 0$ et 0 si $\tau = \pm 1, \pm 2\ldots$
    Example $e_t \sim N(0, \sigma_e{}^2)$



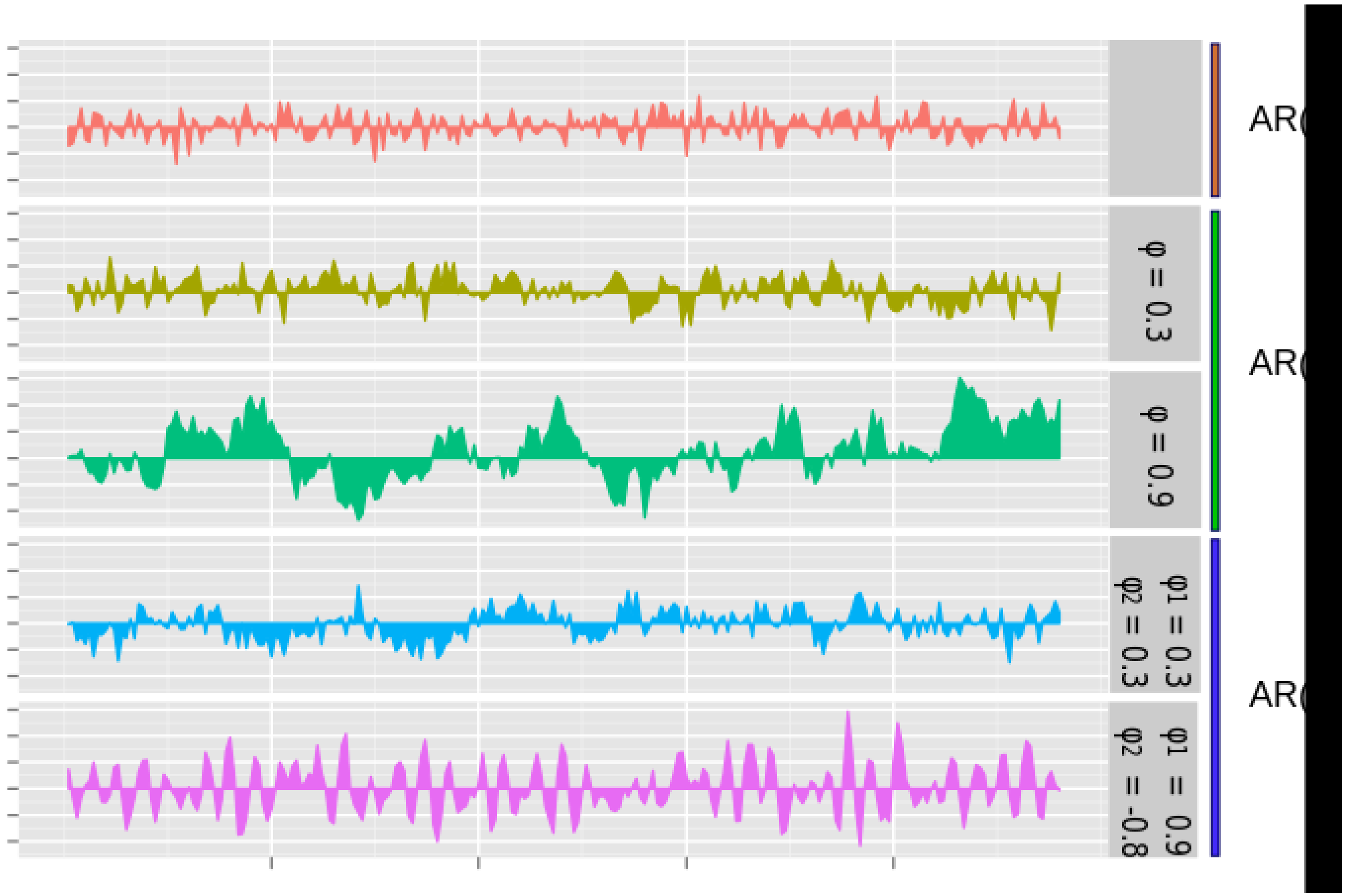
  - Random walk $\{Y_t, t=1, 2\ldots\}$, $Y_t=Y_{t-1}+e_t$, $e_t \sim N(0, \sigma_e{}^2)$, $E(Y_t)=E(Y_0)$, $Var(Y_t)=Var(Y_0)+t\sigma_e{}^2$
    non stationary RP, but $e_t=Y_t-Y_{t-1}$ it is.





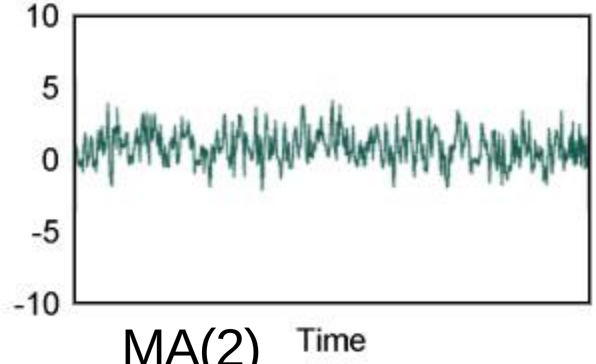
Univariate time series

- Autoregressive model (AR) $\{Y_t, t=0, 1, 2\ldots\}$
  $Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots a_n Y_{t-n} + e_t$, $e_t \sim N(0, \sigma_e^2)$
  Example



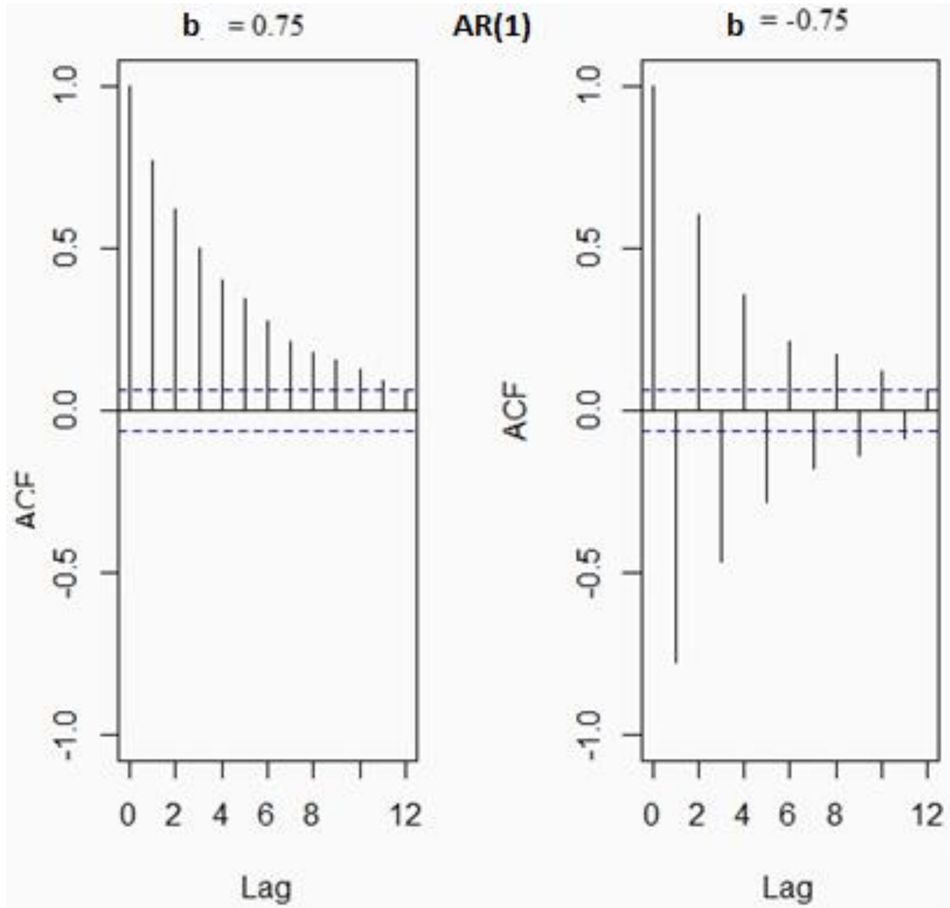
Univariate time series

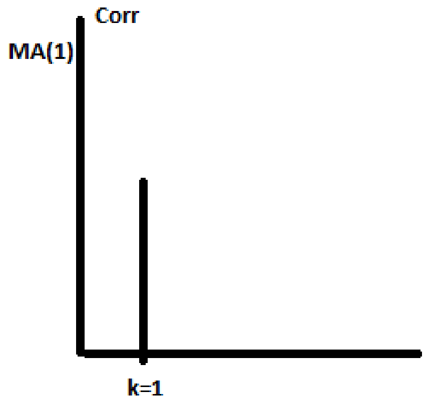
- Mobile average (MA) $\{Y_t, t=0, 1, 2\ldots\}$, $Y_t= e_t+a_1\, e_{t-1}+\ldots a_n\, e_{t-n}$ and $e_t \sim N(0, \sigma_e^2)$



MA(2)

- Comparison of AR (1) and MA (1)
  - models $Y_t=b_0+b_1\, Y_{t-1}+e_t$ and $Y_t= e_t+a\, e_{t-1}$
  - MA(1) : $\mathrm{var}(Y_t) = \sigma_e^2(1+a^2)$, $\mathrm{Cov}(Y_t, Y_{t-k})=a\,\sigma_e^2$ if k=1 and 0 if k>1 and $\mathrm{Corr}(Y_t, Y_{t-k})=a/(1+a^2)$ si k=1 …
  - AR(1) : $\mathrm{Corr}(Y_t, Y_{t-k})=b_1^k$





Univariate time series

- Link between AR and MA

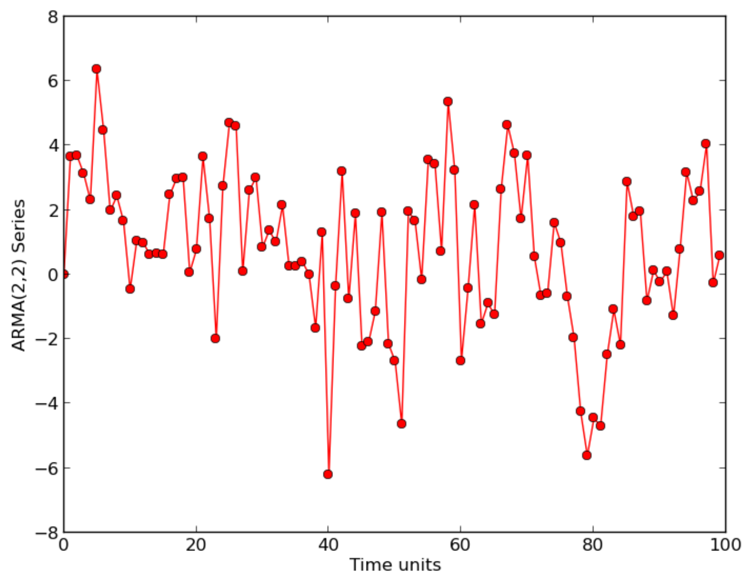
Example

$Y_t = b\,Y_{t-1} + e_t = b(Y_{t-2} + e_{t-1}) + e_t = bY_{t-2} + be_{t-1} + e_t = \ldots$

$= bY_0 + be_0 + \ldots + be_{t-1} + e_t$

- Autoregressive and mobile average models (ARMA(p,q)) $\{e_t,\ t=0, 1, 2\ldots\}$, $\{Y_t,\ t=0, 1, 2\ldots\}$ and $e_t \sim N(0, \sigma_e^2)$

  $Y_t = e_t + a_1\, e_{t-1} + \ldots . a_q\, e_{t-q} + b_1\, Y_{t-1} + b_2\, Y_{t-2} + \ldots . b_p\, Y_{t-p},$



ARMA(2,2)

Estimation of stationary autoregressive model

- Moment based method
  - AR(n): $Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots a_n Y_{t-n} + e_t$. We want to estimate the model $a_0, a_1, a_2 \ldots, a_n$
  - expectation $\mu = a_0 + a_1 \mu + a_2 \mu + \ldots a_n \mu$, thus $a_0 = (1 - a_1 - a_2 - \ldots a_n)\mu$
  - We can write AR(n) : $Y_t - \mu = a1 (Y_{t-1} - \mu) + a_2 (Y_{t-2} - \mu) + \ldots a_n (Y_{t-n} - \mu) + e_t$
  - multiplying AR(p) by $Y_{t-\tau} - \mu$: $(Y_t - \mu)(Y_{t-\tau} - \mu) = a1 (Y_{t-1} - \mu)(Y_{t-\tau} - \mu) + a_2 (Y_{t-2} - \mu)(Y_{t-\tau} - \mu) + \ldots a_n (Y_{t-n} - \mu)(Y_{t-\tau} - \mu) + (Y_{t-\tau} - \mu) e_t$
  - Taking the expectation:

    $C_Y(0) = a_1 C_Y(-1) + a_2 C_Y(-2) + \ldots a_n C_Y(-n) + \sigma_e^2$

    $C_Y(\tau) = a1 C_Y(\tau - 1) + a_2 C_Y(\tau - 2) + \ldots a_n C_Y(\tau - n)$ $\tau > 0$
  - To find the model, we need to solve the linear equation system (Yule-Walker equations) of n+1 unknowns and n+1 equations.

$$C_Y(0) = a_1 C_Y(-1) + a_2 C_Y(-2) + \ldots a_n C_Y(-n) + \sigma_e^2$$

$$C_Y(1) = a_1 C_Y(0) + a_2 C_Y(-1) + \ldots a_n C_Y(-n+1)$$

$$\ldots$$

$$C_Y(n) = a_1 C_Y(n - 1) + a_2 C_Y(n - 2) + \ldots a_n C_Y(0)$$

Estimation of stationary autoregressive model

- Maximum likelihood: AR(1) case
  - Y_t=a_0+a_1 Y_{t-1}+e_t. We want to estimate the parameters. For this, we need to define gradually the likelihood function. The observations are y_T, y_{T-1},…,y_1
  - We specify the joint probability [y_T, y_{T-1},…,$y_{1;}$ θ].
  - y_1 is sampled form a Gaussian N($\mu_y, \sigma_y{}^2$), where μ=a_0/(1-a_1) and $\sigma_y{}^2$=$\sigma_e{}^2$/(1-$a_1{}^2$). Thus ,

    [y_1; θ]= N($\mu_y, \sigma_y{}^2$)

    let us consider y_2, the conditional distribution of y_2 given to Y_1=y_1, i.e.

Y_2=a_0+a_1 Y_1+e_2 (a constant plus e_2)

[y_2|y_1 ;θ]=N(a_0+a_1y_1, $\sigma_e{}^2$)

because [y_2,y_1 ;θ]= [y_2|y_1 ;θ] [y_1; θ]=N(a_0+a_1y_1, $\sigma_e{}^2$) N($\mu_y, \sigma_y{}^2$)

let us consider y_3, the conditional distribution of y_3 given Y_2=y_2 and Y_1=y_1 ; i.e. Y_3=a_0+a_1 Y_2+e_3 (constant plus e_3)

[y_3|y_2 , y_1;θ]=N(a_0+a_1y_2, $\sigma_e{}^2$)

because [y_3, y_2,y_1 ;θ]= [y_3|y_2 , y_1;θ] [y_2|y_1 ;θ] [y_1; θ]= N(a_0+a_1y_2, $\sigma_e{}^2$)N(a_0+a_1y_1, $\sigma_e{}^2$) N($\mu_y, \sigma_y{}^2$)

Estimation of stationary autoregressive model

- The same reasoning leads to $[y_t|y_{t-1}\ldots, y_3, y_2, y_1 ;\theta]$= N($a_0+a_1y_{t-1}$, $\sigma_e{}^2$)
- The likelihood (joint pdf) : $[y_T,\ldots, y_3, y_2, y_1 ;\theta]$=N($\mu_y, \sigma_y{}^2$) $\prod_{t=2}^{T}$ N($a_0+a_1y_{t-1}$,$\sigma_e{}^2$)
- The maximum likelihood max $_{\sigma y,\ a0,\ a1,\ \sigma e}$ ln(N($\mu_y, \sigma_y{}^2$) $\prod_{t=2}^{T}$ N($a_0+a_1y_{t-1}$,$\sigma_e{}^2$)). Do it as exercise.

- simplification, we consider constant Y_1; i.e. max $_{a0,\ a1,\ \sigma e}$ ln($\prod_{t=2}^{T}$ N($a_0+a_1y_t$,$\sigma_e{}^2$)). The closed form solution :

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} T-1 & \sum_{t=2}^{T} y_{t\_1} \\ \sum_{t=2}^{T} y_{t\_1} & \sum_{t=2}^{T} y_t{}^2{}_{\_1} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=2}^{T} y_t \\ \sum_{t=2}^{T} y_{t\_1}\, y_t \end{bmatrix}$$

$$\sigma_e{}^2=\sum_{t=2}^{T}(y_t-a_0-a_1y_{t-1})^2/(T-1)$$

Estimation of stationary autoregressive model

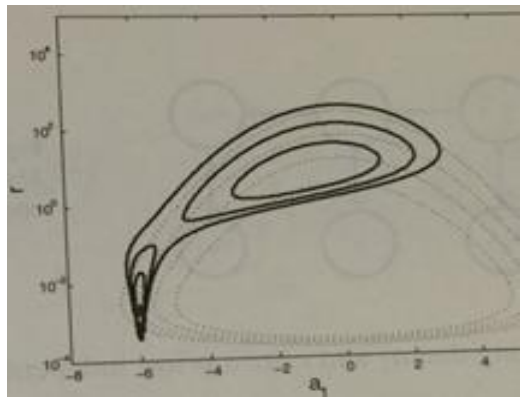
- Maximum likelihood: AR(n) case
  - $Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots a_n Y_{t-n} + e_t$.
  - observations $y_T, y_{T-1}, \ldots, y_1$
  - the joint pdf $[y_T, y_{T-1}, \ldots, y_{1;}\, \theta] = [y_n, \ldots, y_{1;}\, \theta] \prod_{t=n+1}^{T} [y_t | y_{t-1,\ldots,} y_{t-n}; \theta]$
  - the conditional pdf $[y_t | y_{t-1}, \ldots, y_{t-n;}\, \theta] = N(a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \ldots a_n y_{t-n}, \sigma_e{}^2)$.
  - The pdf of the first n values, considering le vector $(y_n, y_{n-1}, \ldots, y_1)^t$, the pdf $[y_n, y_{n-1}, \ldots, y_{1;}\, \theta] = N(\mu, \emptyset)$, where μ is the mean vector of $y_n, y_{n-1}, \ldots, y_1$ and $\emptyset$ the covariance matrix; $\emptyset_{ij} = E(Y_i - \mu_i)(Y_j - \mu_j)$ because TS is stationary $\mu_i = \mu_j$
  - Finding the model, requires maximizing the likelihood (pdf) $N(\mu, \emptyset) \prod_{t=n+1}^{T} \quad N(a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \ldots a_n y_{t-n}, \sigma_e{}^2)$

Estimation of stationary autoregressive model

- Bayesian estimation
  - AR(n) : $Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \dots a_n Y_{t-n} + e_t$.
  - $[a_0, a_1 \dots a_n, \sigma_e^2 \mid y_T, y_{T-1}, \dots, y_1] = [y_T, y_{T-1}, \dots, y_1 \mid a_0, a_1 \dots a_n, \sigma_e^2][a_0, a_1 \dots a_n, \sigma_e^2] / [y_T, y_{T-1}, \dots, y_1]$

    e.g. $[a_0, a_1 \dots a_n, \sigma_e^2] = \mathrm{Invgamma}(\beta, \beta/v) \prod_{i=0}^{n} N(m, \tau^2)$
  - example

  AR(1) : $Y_t = a_1 Y_{t-1} + e_t$. $[a_1] = N(0, \tau^2)$, $[\sigma_e^2] = \mathrm{Invgamma}(\beta, \beta/v) = \exp(-(v +$



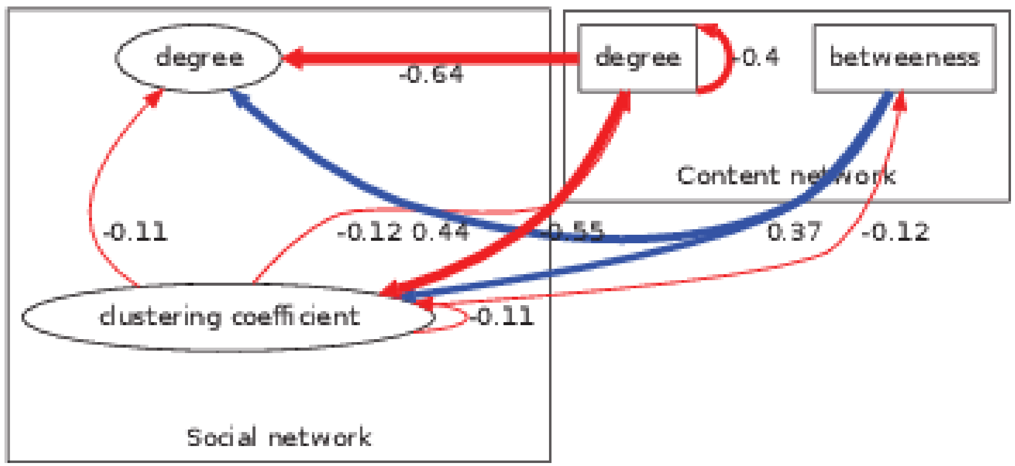
Example – influence in social networks

- It is proposed to study the evolution of co-publication and research topics in a conference that takes place each year.
- We are building
  - a labeled and undirected (social) graph of the authors. The node is an author and the arc indicates the number of articles written by two authors.
  - a labeled and undirected graph (content) of the keywords (themes). The knot is a theme and the bow is the co-occurrence of two themes in the articles.
  - the links between the two previous graphs are explicit in the articles. We know what theme is associated with which author and vice versa.
- A graph is characterized by
  - social: degree of centrality (importance for the author to write articles with others), intermediate centrality, grouping coefficient (importance for the author to write articles with the same authors)
  - content: degree of centrality (popularity of a theme), intermediate centrality (how the theme relates to other themes)
  - Let denotes $y_1 \dots y_P$ the authors, each $y_i = (y_{i1},...,y_{i5})$ the features of i^{th} author.

Example – influence in social networks

- Prediction:
  - Estimation of AR for each variable $Y^{t}_{pj}=a_{pj0}+a_{pj1}\,Y^{t-1}_{pj}+a_{pj2}\,Y^{t-2}_{pj}+\ldots a_{pjn}\,Y^{t-n}_{pj}+e^{t}_{pj}\ \forall p \text{ and } \forall j$
  - a feature of an author can be predicted
  - this solution cannot lead to measure the influence of one feature on another.

- Influence
  - Estimation of a mixed AR $(Y^{t}_{pj}-\mu)/\sigma_y=a_{pj1}\,(Y^{t-1}_{p1}-\mu)\,/\sigma_y+a_{pj2}\,(Y^{t-1}_{p2}-\mu)\,/\sigma_y+\ldots a_{pj5}\,(Y^{t-1}_{p5}-\mu)\,/\sigma_y+e^{t}_{pj}\ \forall p \text{ and } \forall j$
  - The influence of one feature on another can be measured.



Influence network

Example – influence in social networks

- Influence (generalisation) :
  - $(Y^{t}_{p}-\mu)/\sigma_y = M^{t-1}\,(Y^{t-1}_{p}-\mu)\,/\sigma_y + \ldots + M^{t-n}\,(Y^{t-n}_{p}-\mu)\,/\sigma_y + e^{t}_{p}\ \forall p$
  - The influence of one feature on another can be measured.

Hierarchical AR model

- [process, parameters|data] =[data|process, parameters][process|parameters][parameters]
- Example 1

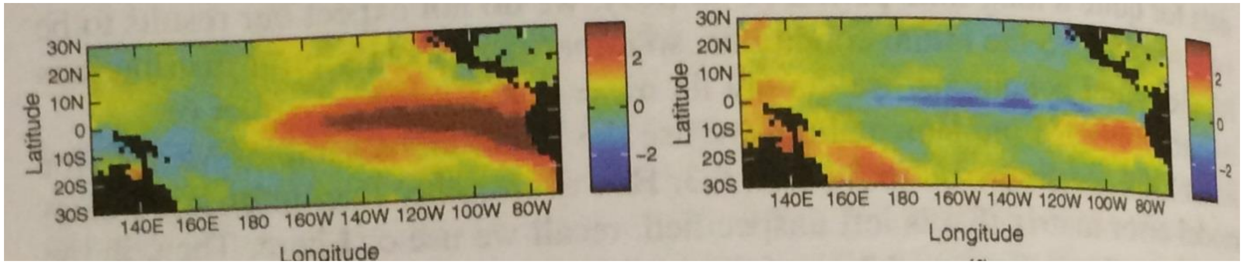
  Data: $Z_t = H_t Y_t + e^t_z$
  Model: $Y_t = M_t Y_{t-1} + e^t_y$
  Parameters: H_t, M_t, et $\sum^t{}_Z, \sum^t{}_Y$ are covariance of e^t_z and e^t_y
  let $Z_1 \ldots Z_T$ the observations, smoothing of Y_t : $E(Y_t | Z_1 \ldots Z_t)$,
  prediction: $E(Y_{t+\tau} | Z_1 \ldots Z_t)$.

- Example 2: prediction of the temperature in a area in Pacific ocean.

  Data: $Z_t = H\alpha_t + e^t_z$, where H is nxp matrix (the first p columns of the observed covariance matrix), $e^t_z \sim N(0, \sigma_z{}^2 I)$
  Model: $\alpha_{t+\tau} = M\alpha_t + e^t_\alpha$, where M is pxp matrix, $e^t_\alpha \sim N(0, \sum^t{}_\alpha)$
  Parameters: column(M)~N(0, 100I), $(\sum^t{}_\alpha)^{-1} \sim$ Wishart(100(p-1)I, p-1), $\sigma_z \sim IGamma(0.1, 100)$,
  prediction: $E(Y_{t+\tau} | Z_1 \ldots Z_t)$ where $\tau = 6\, mois$.

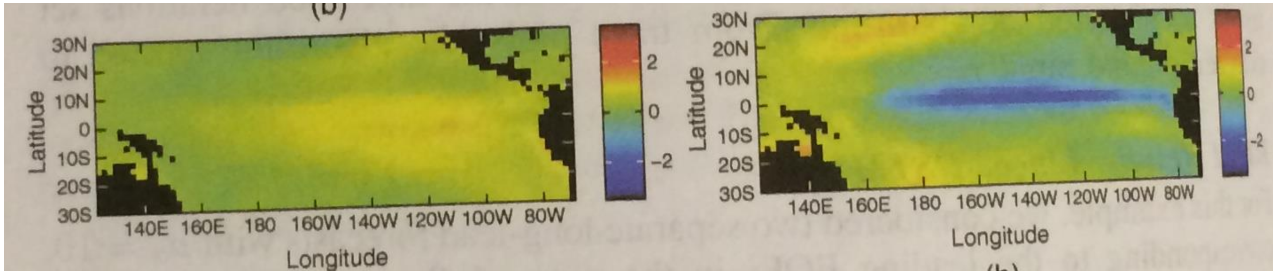
Hierarchical AR model

- p=10



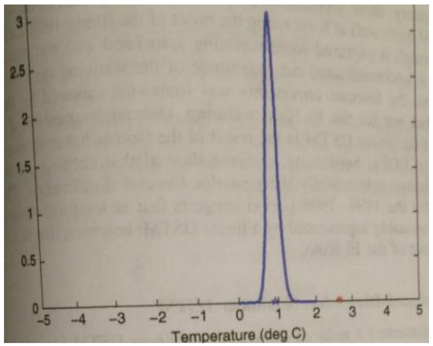
Data, October 1997 (El Nino)

Data, October 1998 (La Nina)

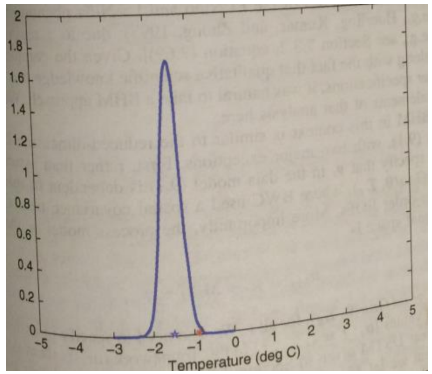


Estimated process (10/1997) using Acquired data until 04/1997

Estimated process (10/1998) using acquired data until 04/1998



Posterior $[Y_{t+\tau}| Z_1 \ldots Z_6]$, El Nino 1997



Posterior $[Y_{t+\tau}| Z_1 \ldots Z_6]$, El Nino 1998

The proposed hierarchical model is good for La Nina, but not for El Nino

Estimation of MA and ARMA models

- The method of moments is not very precise.
- Maximum likelihood estimators or Bayesian are used. The approach is similar to that used for the AR.

References

- M. R. Haddad, H. Baazaoui, D. Ziou, H. B. Ghézala. A predictive model for recurrent consumption behavior: An application on phone calls. Knowledge-Based Systems 04/2014.
- J. D. Hamilton. Time Series Analysis. Princeton University Press, 1994.
- S. Wang and P. Groth. Measuring the Dynamic Bi-directional Influence between Content and Social Networks. LNCS 6496, pp. 814–829, 2010.
- N. Cressie and C. K. Wikle. Statistics for Spatio-Temporal Data. Wiely 2011.
- D. Barber, A. Taylan Cemgil, and S. Chiappa. Bayesian Time Series Models. Cambridge University Press, 2011.