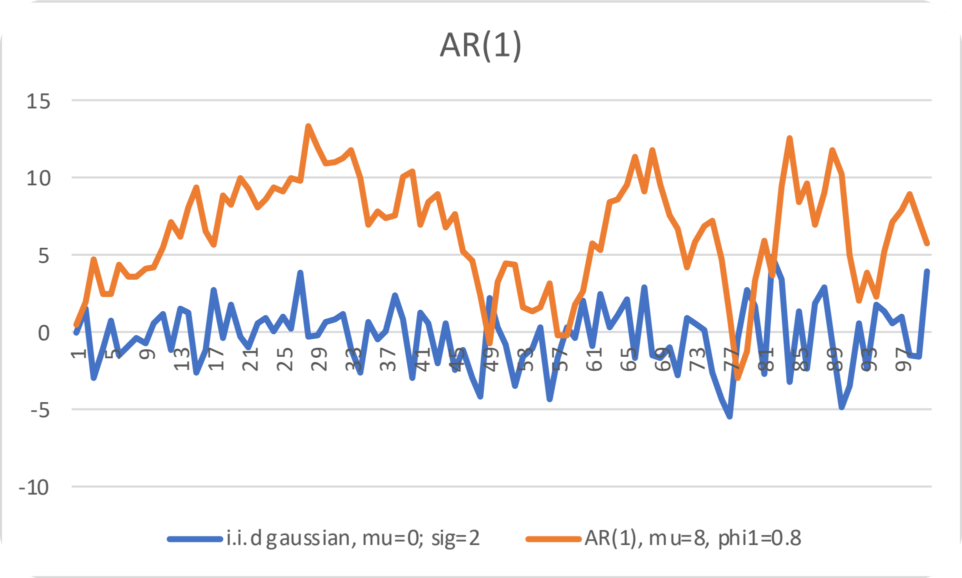
Water Resources Engineering and Management

(CIVIL-466, A.Y. 2024-2025)

5 ETCS, Master course

Prof. P. Perona
Platform of hydraulic constructions



Lecture 7-1: Univariate linear Models, periodic models, Thomas-Fiering model

Models identification

The model identification phase requires determining the orders of the p-autoregressive and the q-moving average components.

- ACF $\quad \rho_k = \dfrac{E[x_{t+k}, x_t]}{E[x_t^2]} \quad$ VS $\quad r_k = \dfrac{\sum_{t=1}^{N-k} (x_{t+k} - x_t)}{\sum_{t=1}^{N} x_t^2}$

  $\rho_m = 0$, $k > m$ $\quad$ VS $\quad r_j \simeq 0 \quad k > j$

  BECAUSE FOR A MA(q) PROCESS $\rho_k = 0$ FOR $k > q$, j CORRESPONDING TO $r_j \simeq 0$ INDICATES THE **max number of moving average components**

- PAF THE VECTOR OF PARAMETERS, ϕ_i, OF THE AUTOREGRESSIVE COMPONENT DEFINES THE **partial autocorrelation function**

  FROM THE YULE-WALKER EQUATION

$$\begin{bmatrix} \rho_1 \\ \vdots \\ \vdots \\ \rho_k \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \rho_{k-1} & \cdots & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \vdots \\ \vdots \\ \phi_k \end{bmatrix}$$

  BY SUBSTITUTING SAMPLE ACF, $r_1, \ldots, r_k$ ONE CAN ESTIMATE THE EMPIRICAL PAF AND COMPARE IT WITH PAF OF MODELS.

  FOR AR(p) $\phi_k = 0 \ \forall\, k > p$ $\quad \boxed{\hat{\phi}_j \simeq 0 \Rightarrow j = \max p}$

Courtesy of P. Burlando

Summary table

Given a time series, compute and plot the sample autocorrelation function and check if

- Exp decrease → AR(1)
- Mixed damping from lag 1 asymptotically to zero→AR(p)
- ACF lag 1 <0.5 and then drops to zero then MA(1)
- ACF becomes zero after lag q, then MA(q)

Table 4.1 Identification of Box–Jenkins models using autocorrelation and partial autocorrelation functions

Model type	*Autocorrelation function*	*Partial autocorrelation function*
AR(1), first-order autoregressive (Markov)	Exponential decrease	$\phi_{1,1} \neq 0$ $\phi_{i,i} = 0$ for $i = 2, 3, 4, \ldots$
AR(p), pth-order autoregressive	Mixed type of damping from lag 1	$\phi_{i,i} \neq 0$ for $i \leqslant p$ $\phi_{i,i} = 0$ for $i > p$
MA(1), first-order moving-average	$\rho_1 \neq 0$ $\rho_i = 0$ for $i = 2, 3, 4, \ldots$	Exponential decrease
MA(q), qth-order moving-average	$\rho_i = 0$ for $i > q$ $\rho_i \neq 0$ for $i \leqslant q$	Mixed type of damping from lag 1
ARMA(1, 1), autoregressive moving-average	Exponential decrease after lag 1	Exponential decrease after lag 1
ARMA(p, q), autoregressive moving-average	Mixed type of damping after lag $q+1$	Mixed type of damping after lag $p-q$

Model estimators based on the method of moments

PLATEFORME DE CONSTRUCTIONS HYDRAULIQUES

AR(1) model

$$\hat{\mu} = \bar{y}$$

$$\hat{\sigma}_\varepsilon^2 = s^2 (1 - r_1^2)$$

$$\hat{\phi}_1 = r_1$$

MA(1) model

Obvious given the type of ACF and the mean

ARMA(1,1) model

$$\hat{\sigma}_\varepsilon^2 = \frac{s^2 (1 - \hat{\phi}_1^2)}{(1 - 2\,\hat{\phi}_1 \hat{\theta}_1 + \hat{\theta}_1^2)}$$

$$\hat{\phi}_1 = \frac{r_2}{r_1}$$

$$\hat{\theta}_1 = \frac{-b \pm \sqrt{b^2 - 4\,(r_1 - \hat{\phi}_1)^2}}{2\,(r_1 - \hat{\phi}_1)}$$

Model testing and diagnostic

PLATEFORME DE CONSTRUCTIONS HYDRAULIQUES

After model estimate, then :

1) Check if model correctly represent process statistics (obviously yes if calibration was done with method of the moment;

2) Use the model to decompose noise from the observed time series.
Revert the equation, e.g., for an AR(2)

$$\varepsilon_t = y_t - \mu - \phi_1(y_{t-1} - \mu) - \phi_2\,(y_{t-2} - \mu)$$

<u>Here now y is the data</u>

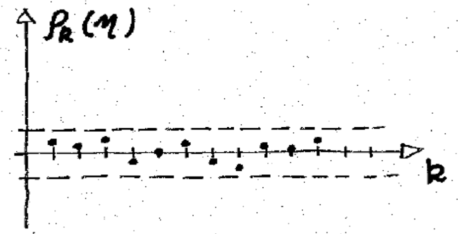
Reconstruct the series of the noise ε_t, and check if its serial correlation is zero

$\rho_k(\eta) \approx 0 \quad \forall\, k > 0$

VISUAL CHECK!

$\downarrow$

TESTS!



YES, serial correlation of reconstructed noise is zero for all lags

FINISHED, model good to be used

NO, serial correlation of reconstructed noise is non zero for all lags

Model is not able to remove all serial correlation → change model

Always use models with as lowest order as possible (parsimoniousness)

Modeling of univariate periodic time series

Periodic Autoregressive Models

PLATEFORME DE CONSTRUCTIONS HYDRAULIQUES

PAR(p) & PARMA(p,q)

seasonal hydrologic series

$$\boxed{x_{\nu,\tau} = \mu_\tau + \sum_{1}^{p} {}_j\, \phi_{j,\tau} (x_{\nu,\tau-j} - \mu_{\tau-j}) + \varepsilon_{\nu,\tau}}$$

ν = year $\quad$ τ = season $\quad$ $\tau = 1, \dots, \omega$

MODEL PARAMETERS

$\varepsilon_{\nu,\tau}$ = NORMAL VARIABLE, UNCORRELATED, ZERO MEAN, $\sigma_\tau^2(\varepsilon)$

PAR(1)

$$\boxed{x_{\nu,\tau} = \mu_\tau + \phi_{1,\tau} (x_{\nu,\tau-1} - \mu_{\tau-1}) + \varepsilon_{\nu,\tau}}$$

e.g. THOMAS-FIERING MODEL

Courtesy of P. Burlando

PARMA (1,1)

$$\boxed{x_{\nu,\tau} = \mu_\tau + \phi_{1,\tau}(x_{\nu,\tau-1} - \mu_{\tau-1}) + \varepsilon_{\nu,\tau} - \theta_1 \varepsilon_{\nu,\tau-1}}$$

PRESERVATION OF BOTH SEASONAL & ANNUAL STATISTICS

PARAMETER ESTIMATION

see SALAS, J. in Handbook of Hydrology, Ch. 19, McGraw-Hill, 19..

Courtesy of P. Burlando

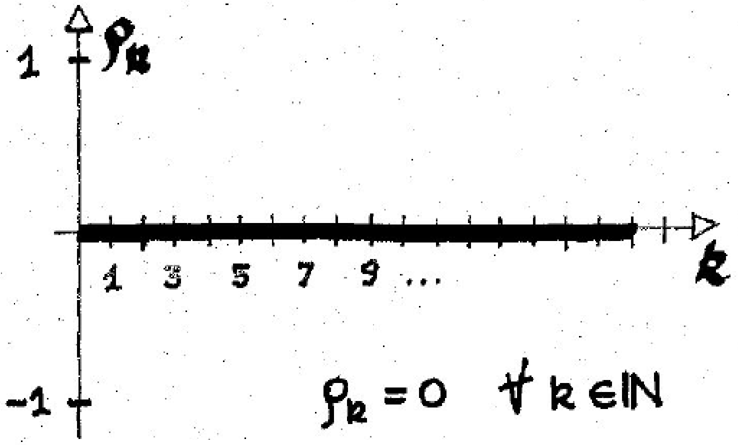
Identification of time series models

BASED ON AUTOCORRELOGRAM ANALYSIS

COMPARISON BETWEEN

γ_k PLOT and ρ_k PLOT → STATIONARY and INDEPENDENT STOCHASTIC PROCESS

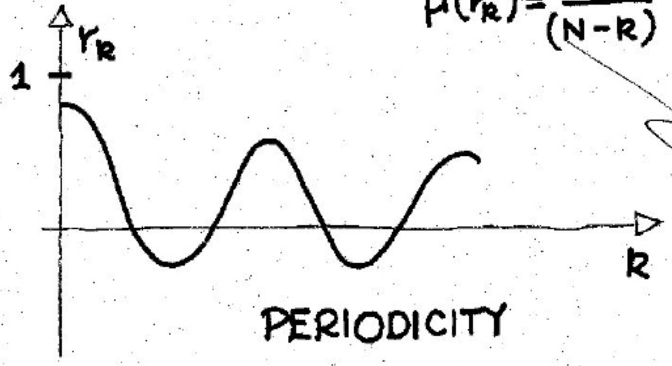
The analysis of the correlogram should allow to identify the deterministic component that must be considered to obtain a STATIONARY and INDEPENT time series

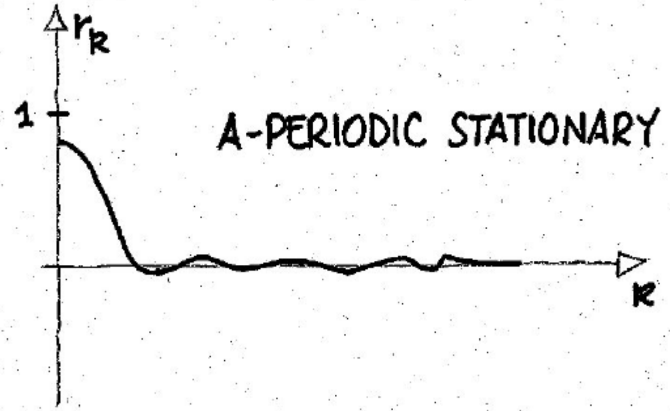


Courtesy of P. Burlando

example

$$\mu(r_k) = \frac{1}{(N-k)}$$

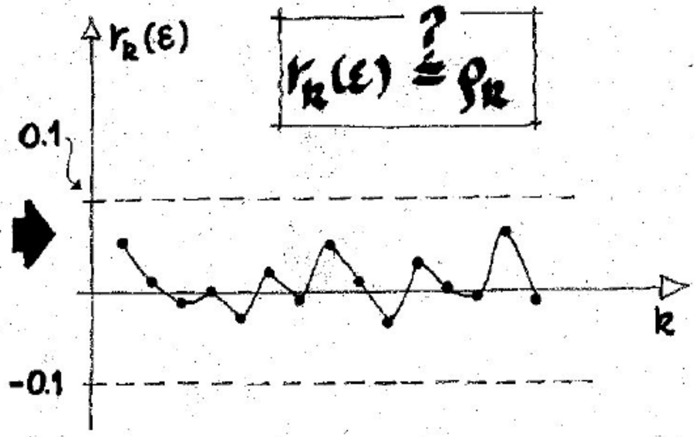




FOR THE A-PERIODIC, TREND-FREE TIME SERIES ONE CAN DEFINE A CONFIDENCE INTERVAL, BASED ON THE ASSUMPTION THAT THE RESIDUALS ARE STATIONARY, INDEPENDENT AND NORMALLY DISTR.

$$\mu(r_k) \pm K_\gamma \sigma(r_k) \qquad \sigma^2(r_k) = \frac{N-k-1}{N-k}$$

$$Pr = 0.90 \rightarrow K_\gamma = 1.645$$



$$\boxed{r_k(\varepsilon) \overset{?}{=} \rho_k}$$

Courtesy of P. Burlando

Model of univariate periodic time series

PLATEFORME DE CONSTRUCTIONS HYDRAULIQUES

NON STATIONARITY INDUCED BY **seasonality**

e.g. river flows → monthly persistence

AR(p) ⇒

$$\boxed{x_t = \sum_{1}^{p} {}_i\, \phi_i\, x_{t-1} + \eta_t}$$

BEING x_t A STATIONARY VARIABLE

$$x_t = \frac{y_t - \mu}{\sigma}$$

PERIODIC AR(p)

MEAN and STANDARD DEVIATION
FOR "SEASON" τ

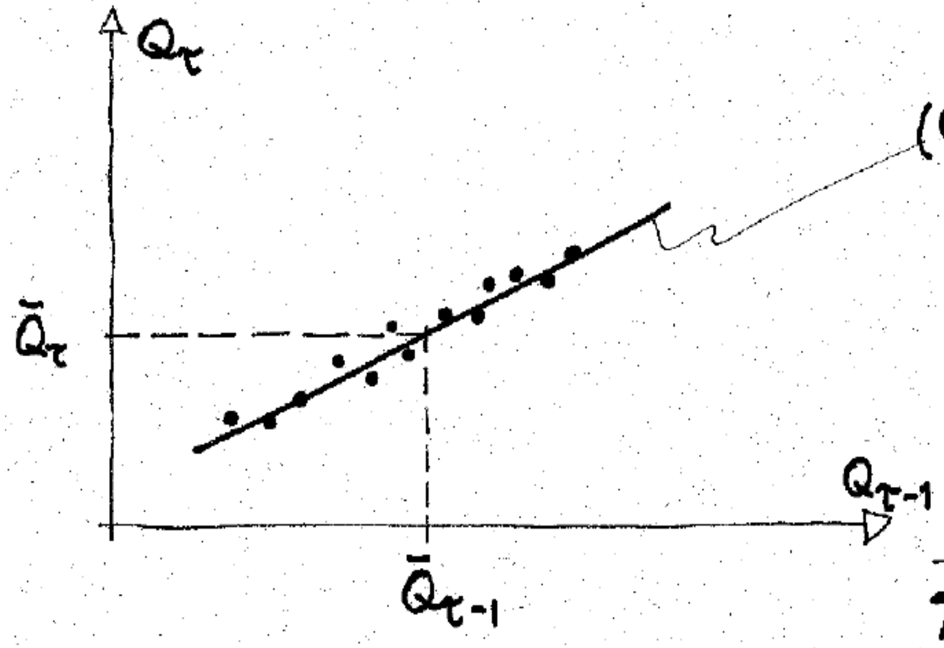
$$\boxed{y_{t,\tau} = \mu_\tau + \sigma_\tau \sum_{1}^{p} {}_i\, \phi_i\, (y_{t-i,\tau-i} - \mu_{\tau-i})/\sigma_{\tau-i} + \sigma_\tau \eta_{t,\tau}}$$

$E[\eta_t]=0 \quad Var[\eta_t]=\sigma_\eta^2$
uncorrelated
normal variable

Courtesy of P. Burlando

example of MONTHLY PERSISTENCE

river flows (effects of storage and subsurface flow)



$$(Q_\tau - \bar{Q}_\tau) = b(Q_{\tau-1} - \bar{Q}_{\tau-1})$$

12 months → 12 equations

⇓

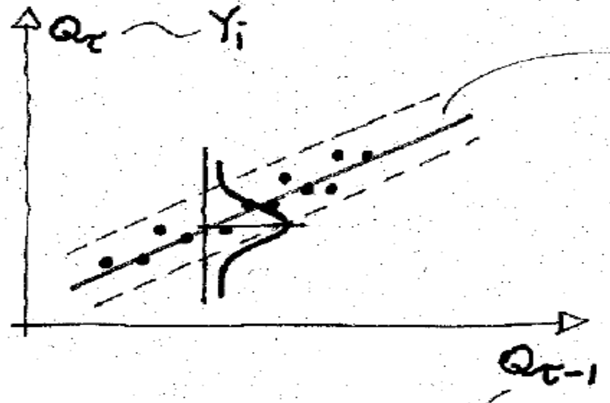
Thomas-Fiering model

Build this plot by comparing, for example, mean or total Q in February vs same Q in January for all available years

Repeat for all month pairs

Courtesy of P. Burlando

Thomas-Fiering model



$$\boxed{Q_\tau = \bar{Q}_\tau + b\,(Q_{\tau-1} - \bar{Q}_{\tau-1}) + \varepsilon_\tau} \qquad (A)$$

Linear regression

$$Y_i = \bar{Y} + b\,(X_i - \bar{X})$$

$$b_{YX} = \frac{\sum_1^n i\; X_i Y_i - n\bar{X}\bar{Y}}{\sum_1^n i\; X_i^2 - n\bar{X}^2}$$

STANDARD ERROR OF ESTIMATE $\rightarrow$ $\sigma_Y \sqrt{1-r^2}$ (B)

VARIANCE $\quad \sigma^2 = \sigma^2 r^2 + \sigma^2 (1-r^2)$

$$r = \frac{\sum_1^n i\; X_i Y_i - n\bar{X}\bar{Y}}{\sqrt{\left(\sum_1^n i\; X_i^2 - n\bar{X}\right)\left(\sum_1^n i\; Y_i^2 - n\bar{Y}^2\right)}}$$

Courtesy of P. Burlando

$$\boxed{b_{YX} = r\,\frac{\sigma_Y}{\sigma_X}}$$

A

B

(*) ⇨

$$\boxed{Q_\tau = \bar{Q}_\tau + r_\tau\,\frac{\sigma_\tau}{\sigma_{\tau-1}}\left(Q_{\tau-1} - \bar{Q}_{\tau-1}\right) + \eta_\tau\,\sigma_\tau\sqrt{1 - r_\tau^2}}$$

$p = 1$

AR(1) $\rightarrow \rho_1$!

B

$$y_{t,\tau} = \mu_\tau + \frac{\sigma_\tau}{\sigma_{\tau-1}}\,\phi_1\left(y_{t-1,\tau-1} - \mu_{\tau-1}\right) + \sigma_{\tau,\eta}\,\eta_\tau$$

36 parameters!

$12\,\bar{Q}_\tau$, $12\,\sigma_\tau$, $12\,r_\tau$

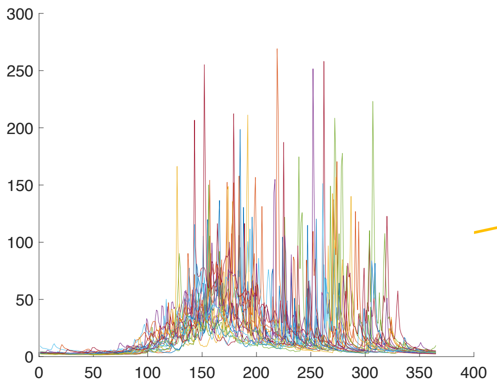
- $\bar{Q}_\tau$ AVERAGE FLOW FOR τ-th MONTH
- $\bar{Q}_{\tau-1}$ " " FOR $(\tau-1)$th MONTH
- r_τ AUTOCORR. COEFF. τ-th $\rightarrow (\tau-1)$-th MONTH
- η_τ GAUSSIAN RANDOM VARIATE
- σ_τ VARIANCE OF FLOW FOR τ-th MONTH

Courtesy of P. Burlando

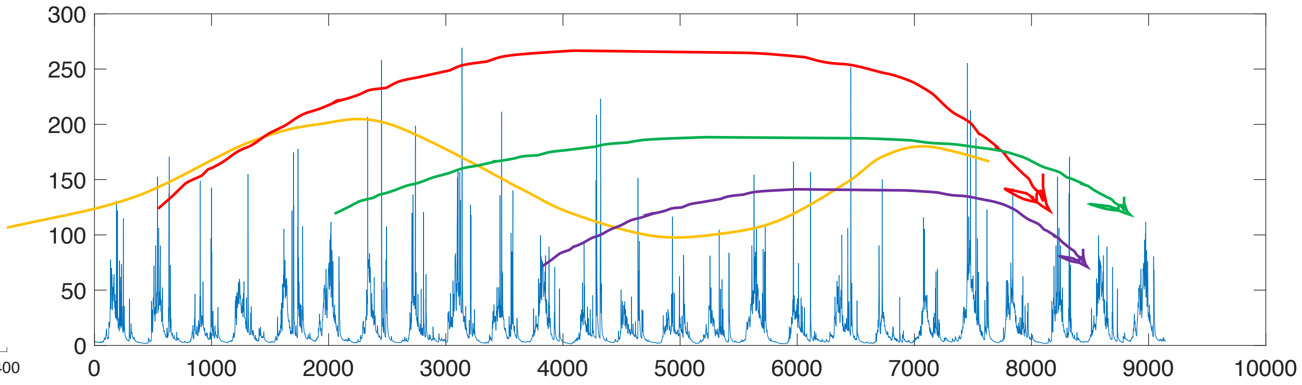
Other techniques

PLATEFORME DE CONSTRUCTIONS HYDRAULIQUES

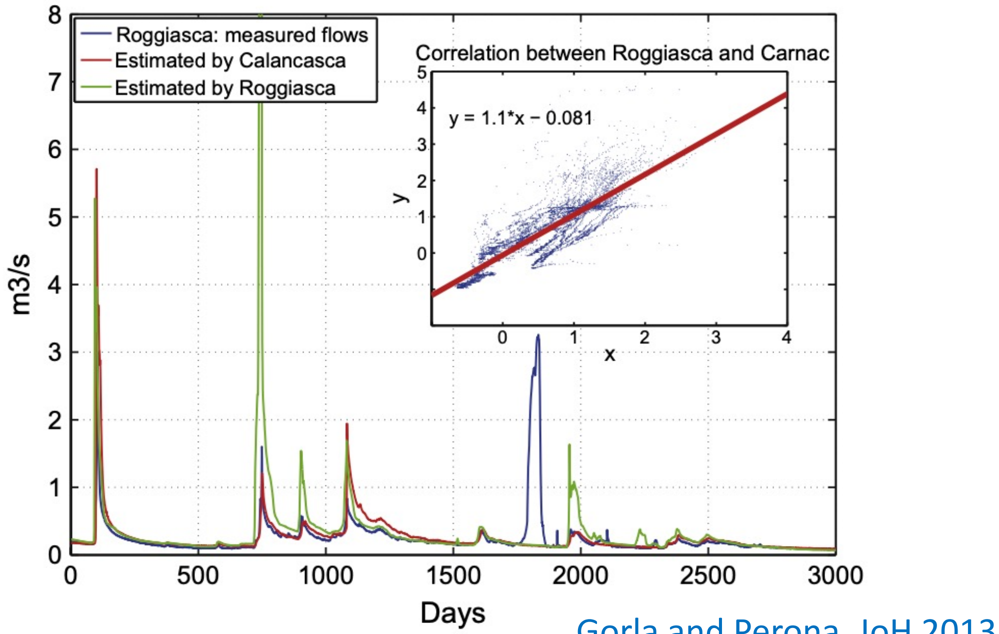
- Reshuffling by swapping data
  conditions:
  no lower frequency trends,
  long enough time series
  preserve serial correlation
  swap cycles



Check this cycle, then swap years



- Use time series from neighborhood gauging stations
  conditions:
  Catchments must have similar dynamics
  Size can be rescaled, but try maintain variance



Gorla and Perona, JoH 2013