Hydrofractals'03

An international conference on fractals in hydrosciences

Monte Verità, Ascona, Switzerland, 24th-29th August, 2003

A stochastic methodology for generation of seasonal time series reproducing overyear scaling

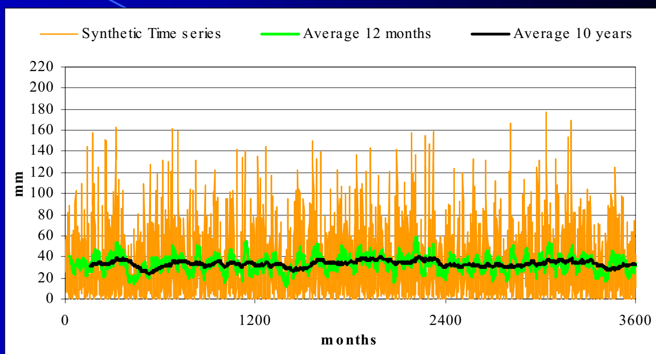
Andreas Lagoussis & Demetris Koutsoyiannis

Department of Water Resources, School of Civil Engineering

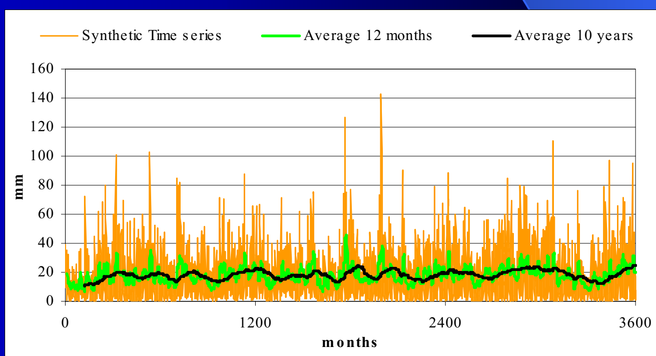
National Technical University of Athens

MAIN PURPOSE (1)

Existing seasonal models of direct sequential stochastic simulation



Developed seasonal models of direct sequential stochastic simulation



MAIN PURPOSE (2)

Our main purpose is to develop stochastic hydrological models that:

- are parsimonious (in parameters) and easy to use,
- directly generate time series at seasonal time scale,
- reproduce cyclostationarity of the process and short-term memory at seasonal (lower-level) time scale,
- preserve the marginal distributions and long-term persistence at annual (higher-level) time scale.

STATISTICAL PROPERTIES OF INTEREST AND EXISTING STOCHASTIC MODELS (1)

Sub-annual (seasonal) time scale:

- ✓ seasonal expected values (of each location),
- ✓ seasonal variances (of each location),
- ✓ seasonal skewness (of each location),
- ✓ lag one autocovariances among seasons of the same location,
- ✓ cross-covariances among locations of the same season.

Multivariate Cyclostationary Stochastic models [e.g. MPAR(1)]

Annual and overyear time scale:

- ✓ annual expected value (of each location),
- ✓ annual variance (of each location),
- ✓ annual skewness (of each location),
- ✓ overyear scaling behaviour of each location (long-term persistence).

Multivariate Stationary Stochastic models [e.g. BMA, SMA]

STATISTICAL PROPERTIES OF INTEREST AND EXISTING STOCHASTIC MODELS (2)

Disaggregation Techniques

Till now, disaggregation techniques are the only way to produce time series which are consistent with hydrological data at more than one time scales (seasonal annual and overyear).

These techniques involve two or more steps:

- where in the first step annual series are generated,
- which are subsequently disaggregated to finer scales.

However, disaggregation:

- involves several difficulties (e.g. in parameter estimation),
- inaccuracies (e.g. skewness reproduction),
- and is a slow procedure.

MOTIVATION

There are no multivariate seasonal stochastic models of direct sequential simulation, that simultaneously preserve:

long-term persistence

+

short-term memory

+

cyclostationarity

+

statistical properties in more than one time scales

MPARSMAF MODEL (1)

(Multivariate Periodic Autoregressive model with Symmetric Moving Average Filter)

The MPASMAF model is based on:

- **the implementation of an MPAR(1) model in addition to an SMA filter,**
- **a property of scaling stochastic processes,**
- **the short memory of an MPAR(1) model.**

Property of scaling processes

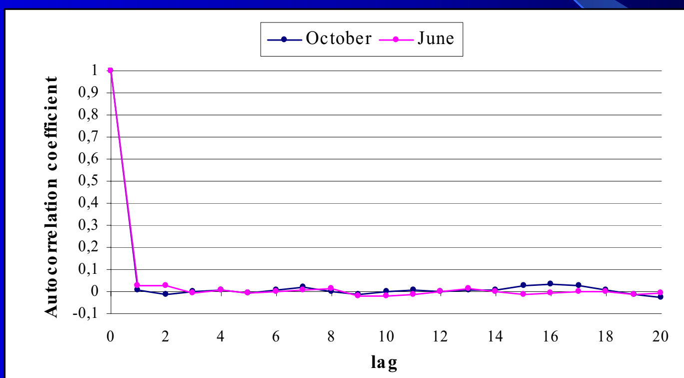
The sum of two (or more) stationary stochastic processes that have the same Hurst coefficient, is a stationary stochastic process with Hurst coefficient equal to the initial one.

Short memory of an MPAR(1) model

If W_j is an MPAR(1) cyclostationary stochastic process with period k, then for a certain s ($s = 1, \ldots, k$) the stochastic process $W_{(j-1)k+s}$ ($j = 1, \ldots$) is stationary with correlation that tends to zero if k tends to infinity.

MPARSMAF MODEL (2)

But even if k is finite (e.g. k =12 □ monthly scale), the synthetic series of the seasons (months) of an MPAR(1) model are, with high accuracy, white noise series.



Empirical autocorrelogram of the synthetic monthly series for the months of October and June, generated by an MPAR(1) model

MPARSMAF MODEL (3)

Model implementation:

1. We produce synthetic time series in seasonal (monthly) time scale using an MPAR(1) model

$$\mathbf{W}_j=\mathbf{a}_j\,\mathbf{W}_{j-1}+\mathbf{b}_j\,\mathbf{V}_j$$

$\mathbf{W}_j = [W_j^1, W_j^2, \ldots, W_j^n]^T$

cyclostationary stochastic process of location n, with period k

$n \times n$ periodically changing parameter matrices with period k

$\mathbf{V}_j = [V_j^1, V_j^2, \ldots, V_j^n]^T$

cyclostationary stochastic process of location n with zero correlation in time j and among the locations, with unit variance.

2. We use the synthetic time series of each month and location, as uncorrelated in time white noise for an SMA model.

$$X^l_{(i-1)k+s} = \sum_{j=-q}^{q} \alpha^l_{|j|}\, W^l_{(i+j-1)k+s}\ , \quad s = 1, \ldots, k \ \text{ and } \ l=1, 2, \ldots, n$$

persistent stationary stochastic process of season s and location l

number of SMA parameters

SMA parameters of location l

MPARSMAF MODEL (4)

Model implementation (continued):

- The periodically changing parameter matrices $\mathbf{a}_j$ and $\mathbf{b}_j$ (j = 1, …, k) of the MPAR(1) model can be estimated analytically.
- The SMA parameters of each location l (l = 1, …, n) can be analytically estimated using the power spectrum of a theoretical autocorrelogram (e.g. FGN autocorrelogram).

Given that there is no need for optimisation of parameters, the model is:

direct

fast

easy to use

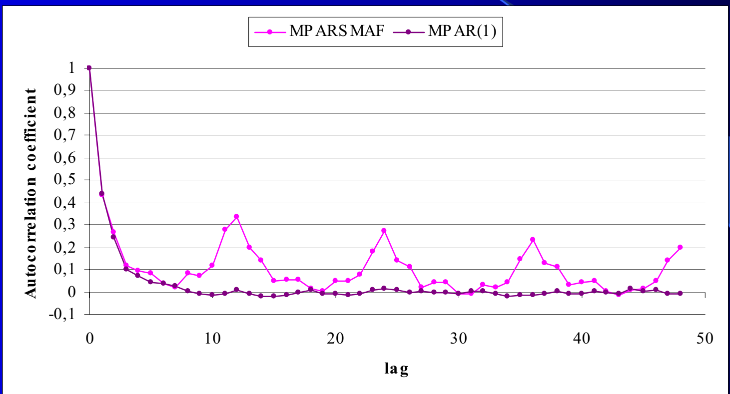
MPARSMAF MODEL (5)

Statistical properties preserved by MPARSMAF model:

- ✓ seasonal (k =12 ⇒ monthly) expected values of each location,
- ✓ seasonal (k =12 ⇒ monthly) variances of each location,
- ✓ seasonal (k =12 ⇒ monthly) skewness coefficients of each location,
- ✓ lag one autocovariances of each location,
- ✓ lag zero and lag one cross-covariances among locations,
- ✓ annual expected values (of all locations),
- ✓ **overyear scaling behaviour (long term persistence) of each location.**

MPARSMAF MODEL (6)

How does MPASMAF model work?



Seasonal autocorrelations of month October with previous months, for 2 independent synthetic series produced using MPAR(1) and MPARSMAF models.

The SMA filter assigns a periodical (twelvemonth) shift in the seasonal autocorrelogram produced by an MPAR(1) model.

SPLITMODEL (1)

Splitmodel is a multivariate model reproducing cyclostationarity, short-term memory and long-term persistence, as a weighted sum of a stationary stochastic process with long memory and a cyclostationary stochastic process with short memory.

$$X_i^l = e_i^l\, Y_i^l + W_i^l$$

cyclostationary stochastic process of location l with period k (where k is the number of seasons)

periodically changing parameter of location l with period k

stationary stochastic process of location l, with long memory

cyclostationary stochastic process of location l with period k preserving only 1st order autocovariances (short memory)

The stochastic processes Y_i^l ($l = 1, \ldots, n$) can be described by an SMA model.

$$Y_i^l = \sum_{j=-q}^{q} \alpha_{|j|}^l\, V_{i+j}^l$$

stationary stochastic process of location l

number of SMA parameters

SMA parameters of location l

stationary stochastic process uncorrelated in time $i+j$, but correlated among the locations l ($l = 1, \ldots, n$), with unit variance

SPLITMODEL (2)

- **The SMA parameters of each location l ($l = 1, \ldots, n$) can be estimated analytically using the power spectrum of the stochastic process Y_i^l ($l = 1, \ldots, n$).**
- **The vector $\mathbf{V}_i=[V_i^1, \ldots, V_i^n]^T$ (n is the number of locations) can be generated by the simple multivariate model**

$$\mathbf{V}_i = \mathbf{b}\, \mathbf{H}_i$$

$n \times n$ parameter matrix

$\mathbf{H}_i=[H_i^1, \ldots, H_i^n]^T$
vector of n stationary stochastic processes, uncorrelated in time i and among the locations l, with unit variance.

The stochastic processes W_i^l ($l = 1, \ldots, n$) can be described by a Multivariate Periodic Forward Moving Average (MPFMA) model

$$\mathbf{W}_i = {}^0\mathbf{f}_i\, \mathbf{R}_i + {}^1\mathbf{f}_i\, \mathbf{R}_{i+1} + \mathbf{g}_i\, \mathbf{G}_i$$

$\mathbf{W}_i = [W_i^1, W_i^2, \ldots, W_i^n]^T$

cyclostationary stochastic process of location n, with period k

$n \times n$ periodically changing parameter matrices with period k

$\mathbf{G}_i=[G_i^1, \ldots, G_i^n]^T$

$\mathbf{R}_i=[R_i^1, \ldots, R_i^n]^T$

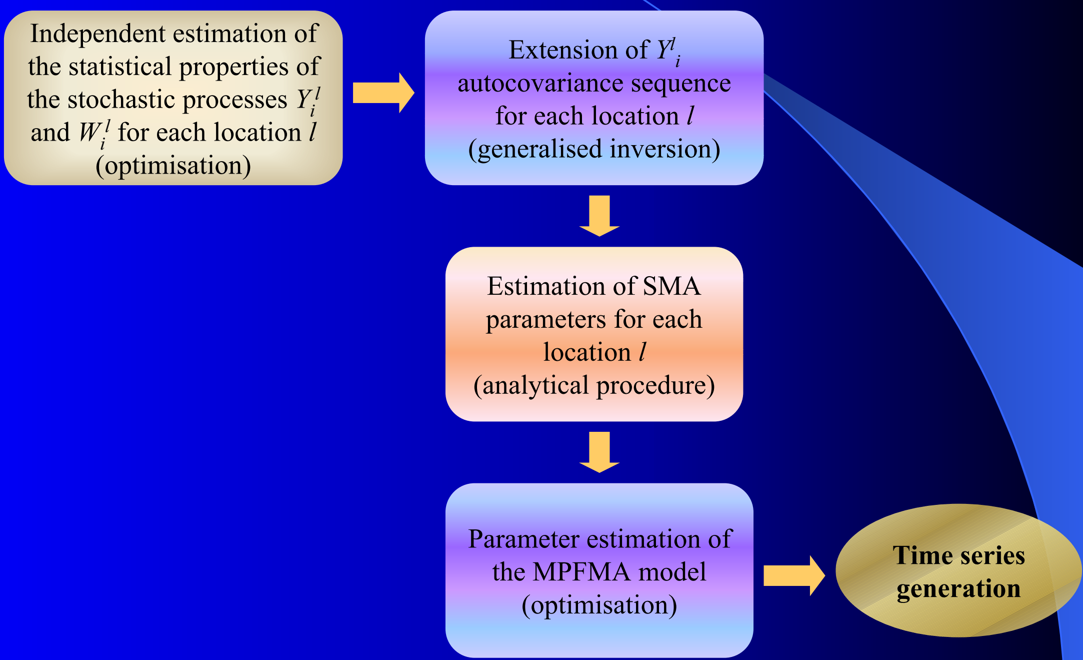
vectors of n cyclostationary stochastic processes, uncorrelated in time i and among the locations l, with unit variance.

SPLITMODEL (3)

- **The statistical properties of the stochastic processes** Y_i^l ($l = 1, \ldots, n$) **και** W_i^l ($i = 1, \ldots, k$ and $l = 1, \ldots, n$)**, can be estimated using constrained gradient based nonlinear optimisation methods of real valued functions of vector variable (e.g. conjugate gradient method in combination with the method of penalties).**
- **The parameter matrix b, the periodically changing parameter matrices** ${}^j\mathbf{f}_i$ ($j = 0, 1$ and $i = 1, \ldots, k$) **and** $\mathbf{g}_i$ ($i = 1, \ldots, k$)**, as long as the statistical properties of the stochastic processes** H_i^l ($l = 1, \ldots, n$) **and** R_i^l, G_i^l, ($i = 1, \ldots, k$ and $l = 1, \ldots, n$) **can be estimated using, once more, constrained gradient based nonlinear optimisation methods of real valued functions of vector variable.**
- **To keep the number of autocovariances of the stochastic processes** Y_i^l ($l = 1, \ldots, n$) **that need to be optimised as low as possible, we have developed a fast and easy algorithm based on *generalised inversion* (the resultant system needed to be solved, is linear and tridiagonal □ Thomas algorithm).**
- **The objective functions needed for the optimisation procedures, as well as the expressions of their derivatives, have been determined analytically.**

SPLITMODEL (4)

Splitmodel implementation flow chart:



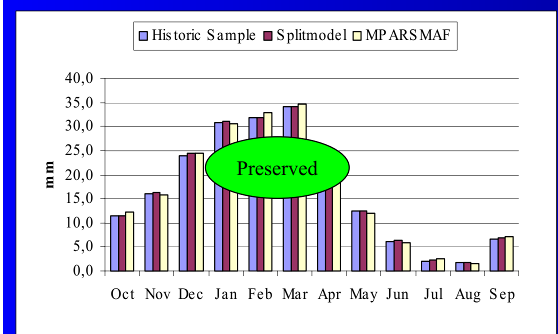
SPLITMODEL (5)

Statistical properties preserved by Splitmodel:

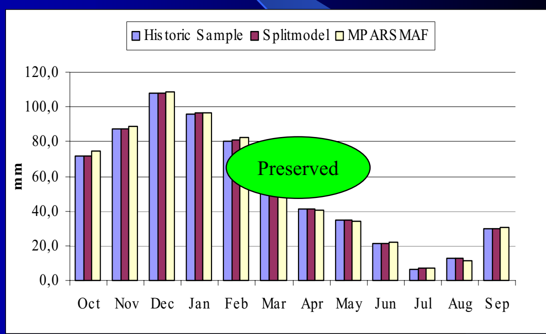
- ✓ seasonal (k =12 ⇒ monthly) expected values of each location,
- ✓ seasonal (k =12 ⇒ monthly) variances of each location,
- ✓ seasonal (k =12 ⇒ monthly) skewness coefficients of each location,
- ✓ lag one autocovariances of each location,
- ✓ lag zero cross-covariances among locations,
- ✓ annual expected values (of all locations),
- ✓ **annual variances (of all locations),**
- ✓ **overyear scaling behaviour (long term persistence) of each location.**

MODEL APPLICATION (1)

Application of Splitmodel and MPARSMAF models to the reproduction of the statistical properties of two correlated monthly time series: (1) monthly discharge of the river Viotikos Kifisos and (2) monthly rainfall at Aliartos pluviometrical station (which belongs to Viotikos Kifisos hydrological basin).

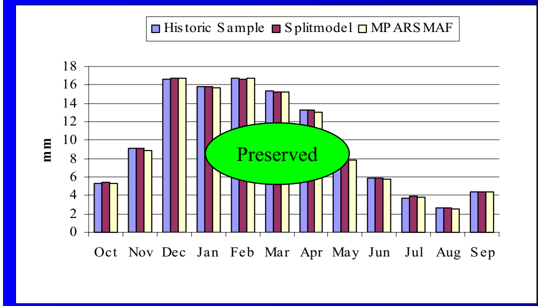


Monthly expected values of discharge time series

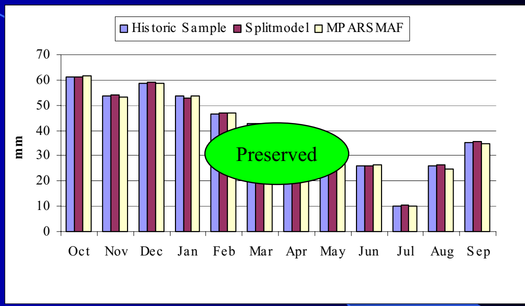


Monthly expected values of rainfall time series

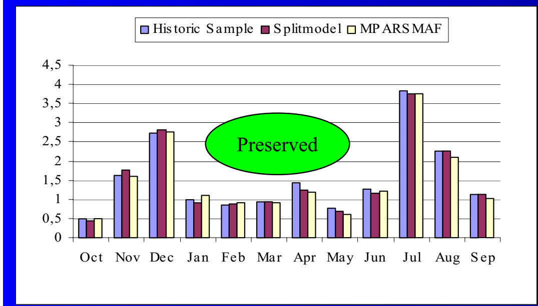
MODEL APPLICATION (2)



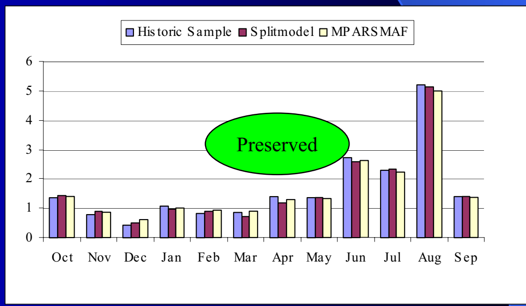
Monthly standard deviations of discharge time series



Monthly standard deviations of rainfall time series

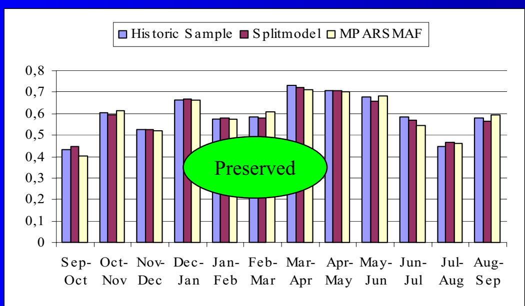


Monthly skewness coefficients of discharge time series

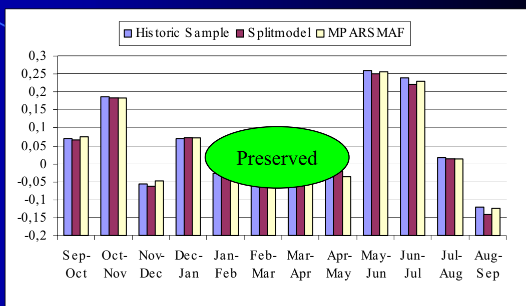


Monthly skewness coefficients of rainfall time series

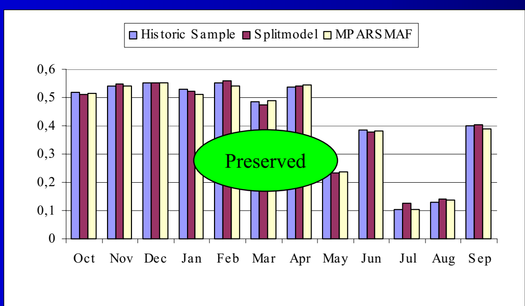
MODEL APPLICATION (3)



lag one autocorrelation coefficients of discharge time series

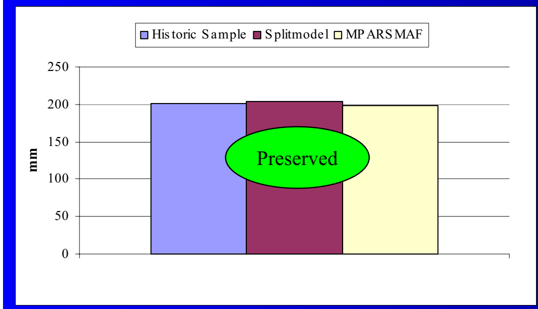


lag one autocorrelation coefficients of rainfall time series

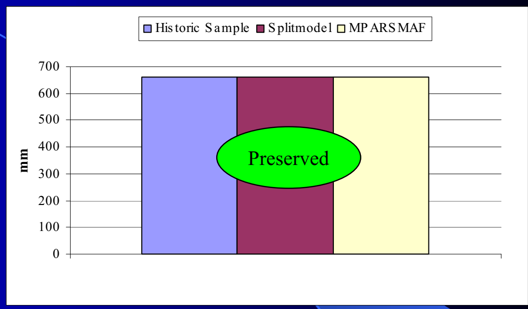


lag zero cross-correlation coefficients of the two time series

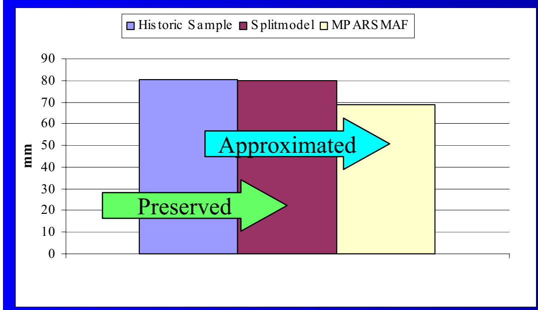
MODEL APPLICATION (4)



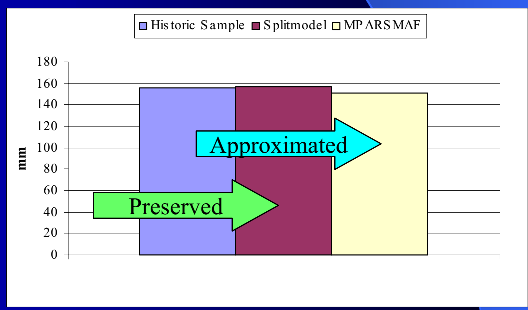
Annual expected value of discharge time series



Annual expected value of rainfall time series



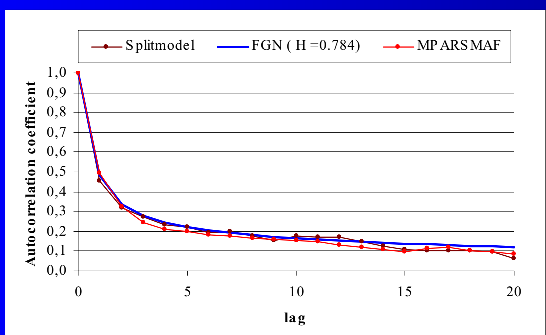
Annual standard deviation of discharge time series



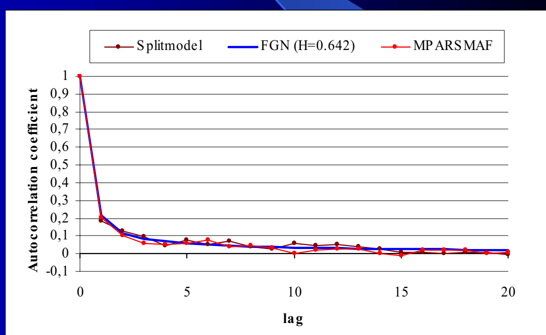
Annual standard deviation of rainfall time series

MODEL APPLICATION (5)

Hurst Phenomenon



Autocorrelogram of annual discharge time series



Autocorrelogram of annual rainfall time series

Preserved

CONCLUSIONS

Two cyclostationary stochastic models have been developed, that:

- are multivariate (very important for planning and design of hydrosystems),
- reproduce short-term memory and long-term persistence,
- preserve the statistical properties of hydrological data in more than one time scales (seasonal, annual and overyear),
- avoid the use of disaggregation.

The models have been applied to real word hydrological data with satisfactory results