Mathematical derivations for seasonal models

reproducing overyear scaling behaviour

Andreas Langousis & Demetris Koutsoyiannis

Department of Water Resources, School of Civil Engineering, National Technical University, Athens
Heroon Polytechneiou 5, GR-157 80 Zographou, Greece

Introduction

This report is an accompanying material of the paper "*A stochastic methodology for generation of seasonal time series reproducing over year scaling*". Particularly, the mathematical derivations for the two stochastic hydrological models developed (i.e. MPAR-SMAF and Split model) are presented in detail in this paper. In section A, a property of scaling stochastic processes, which forms the basis of the functioning structure of the MPAR-SMAF model, is presented and proved. Section B includes the objective functions proposed for the nonlinear optimizations required by the Split model, the analytical expressions of their derivatives, as well as two algorithms developed particularly for Split model.

A. MPAR-SMAF model

We will derive a property of scaling stochastic processes, which is the basis of the MPAR-SMAF model. Particularly, we will prove that the sum of two or more stationary stochastic processes with the same Hurst coefficient H is a stationary stochastic process with Hurst coefficient equal to the initial one.

If Q_i is a stationary stochastic process where the subscript i denotes time, then the long-term persistence of the stationary stochastic process Q_i can be described by the equation

$$(Q_i^{(k)} - k\mu_q) =_{\mathrm{d}} \left(\frac{k}{\lambda}\right)^H (Q_j^{(\lambda)} - \lambda\mu_q) \tag{A.1}$$

where $=_d$ denotes equality in distribution, H is the Hurst coefficient of the stochastic process Q_i, $\mu_q := E[Q_i]$ is the expected value of the stochastic process and $Q_i^{(k)}$ is the stochastic process defined as the sum of k sequential time terms of the stochastic process Q_i,

$$Q_i^{(k)} = \sum_{j=(i-1)k+1}^{ik} Q_j \tag{A.2}$$

Let the stochastic processes P_i and $P_i^{(k)}$ be defined identically to the stochastic processes Q_i and $Q_i^{(k)}$and have the same Hurst coefficient H. In this case it is evident that

$$(P_i^{(k)} - k\mu_p) =_d \left(\frac{k}{\lambda}\right)^H (P_j^{(\lambda)} - \lambda\mu_p) \tag{A.3}$$

holds. Adding, now, equations (A.1) and (A.3) we obtain

$$(Q_i^{(k)} + P_i^{(k)} - k\mu_q - k\mu_p) =_d \left(\frac{k}{\lambda}\right)^H (Q_j^{(\lambda)} + P_j^{(\lambda)} - \lambda\mu_q - \lambda\mu_p) \tag{A.4}$$

which denotes that if two stationary stochastic processes Q_i, P_i with the same Hurst coefficient H are added, then the resulting stochastic process $Q_i + P_i$ is stationary with Hurst coefficient H equal to the initial one. It is evident, that the former proof can be generalised for more than two stationary stochastic processes.

B. Multivariate Split model

B.1 Additional material for section 3.2

B.1.1 Objective function

We group all the unknown parameters of each location l into the vector ζ^l which has $k(n+4)$ elements, i.e.

$$\zeta^l := [e^l_1, \ldots, e^l_k, \delta^l_{1,0}, \ldots, \delta^l_{k,0}, \delta^l_{1,1}, \ldots, \delta^l_{k,1}, \beta^l_0, \beta^l_1, \ldots, \beta^l_{k(n+1)-1}]^T, \; n \geq 1; l = 1, \ldots, v \tag{B.1}$$

We also group the statistical properties of the stochastic processes X_i^l and Z_i^l into the following vectors

$$\boldsymbol{\theta}^l_1 := [\mathrm{Var}[X^l_1], \mathrm{Var}[X^l_2], \ldots, \mathrm{Var}[X^l_k]]^T, \; l = 1, \ldots, v \tag{B.2}$$

$$\boldsymbol{\theta}^l_2 := [\mathrm{Cov}[X^l_k, X^l_1], \mathrm{Cov}[X^l_1, X^l_2], \ldots, \mathrm{Cov}[X^l_{k\text{-}1}, X^l_k]]^T, \; l = 1, \ldots, v \tag{B.3}$$

$$\boldsymbol{\theta}^l_3 := [\mathrm{Cov}[Z^l_i, Z^l_i], \mathrm{Cov}[Z^l_i, Z^l_{i\text{-}1}], \ldots, \mathrm{Cov}[Z^l_i, Z^l_{i\text{-}n}]]^T, \; l = 1, \ldots, v \tag{B.4}$$

and define the vectors $\tilde{\boldsymbol{\theta}}_1(\boldsymbol{\zeta}^l)$, $\tilde{\boldsymbol{\theta}}_2(\boldsymbol{\zeta}^l)$ and $\tilde{\boldsymbol{\theta}}_3(\boldsymbol{\zeta}^l)$ identically to vectors $\boldsymbol{\theta}^l_1$, $\boldsymbol{\theta}^l_2$, $\boldsymbol{\theta}^l_3$. $\tilde{\boldsymbol{\theta}}_1(\boldsymbol{\zeta}^l)$, $\tilde{\boldsymbol{\theta}}_2(\boldsymbol{\zeta}^l)$ and $\tilde{\boldsymbol{\theta}}_3(\boldsymbol{\zeta}^l)$ can be estimated from equations (15), (17), (18) and (19). In this case, the objective function that we propose for each location l is,

$$J(\boldsymbol{\zeta}^l_*) = \min[J(\boldsymbol{\zeta}^l)] = \lambda^l_1 \left\| \tilde{\boldsymbol{\theta}}_1(\boldsymbol{\zeta}^l) - \boldsymbol{\theta}^l_1 \right\|^2 + \lambda^l_2 \left\| \tilde{\boldsymbol{\theta}}_2(\boldsymbol{\zeta}^l) - \boldsymbol{\theta}^l_2 \right\|^2 + \lambda^l_3 \left\| \tilde{\boldsymbol{\theta}}_3(\boldsymbol{\zeta}^l) - \boldsymbol{\theta}^l_3 \right\|^2 +$$

$$h_1(\boldsymbol{\zeta}^l) + h_2(\boldsymbol{\zeta}^l) + h_3(\boldsymbol{\zeta}^l) + h_4(\boldsymbol{\zeta}^l) \tag{B.5}$$

where λ^l_1, λ^l_2 and λ^l_3 are positive weighting factors, $\|.\|$ denotes the Euclidean norm of a vector and $h_1(\boldsymbol{\zeta}^l)$, $h_2(\boldsymbol{\zeta}^l)$, $h_3(\boldsymbol{\zeta}^l)$, $h_4(\boldsymbol{\zeta}^l)$ are penalty terms positively valued if constraints (20)-(22) respectively are not satisfied. The value of (B.5) will be zero in case that equations (15), (17), (18) and (19) and constraints (20)-(22) are satisfied simultaneously. Otherwise the objective function will be positively valued. The term $h_1(\boldsymbol{\zeta}^l)$ which ensures the maintenance of constraint (20), is given by the equation,

$$h_1(\boldsymbol{\zeta}^l) = \kappa^l_1 \left[\sum_{s=1}^{k} \left(U(\varepsilon - \delta^l_{s,0})\, (\delta^l_{s,0} - \varepsilon)^2 \right) + U(\varepsilon - \beta^l_0)\, (\beta^l_0 - \varepsilon)^2 \right] \tag{B.6}$$

where κ^l_1 is a positively valued penalty factor and $U(x)$ is Heaviside's step function with $U(x) = 1$ if $x \geq 0$ and $U(x) = 0$ otherwise. The term $h_2(\boldsymbol{\zeta}^l)$ ensures the maintenance of constraint (21) and it is given by the equation,

$$h_2(\boldsymbol{\zeta}^l) = \kappa^l_2 \left\| \tilde{\boldsymbol{\theta}}_4(\boldsymbol{\zeta}^l) \right\|^2 \tag{B.7}$$

where κ^l_2 is a positively valued penalty factor and $\tilde{\boldsymbol{\theta}}_4(\boldsymbol{\zeta}^l)$ is a vector with k elements estimated by the expression,

$$^{(l)}\theta_4{}^j = U\{(\delta^l_{j,1})^2 - (1\text{-}\varepsilon)^2\, \delta^l_{j,0}\, \delta^l_{j\text{-}1,0}\}\ \{(\delta^l_{j,1})^2 - (1\text{-}\varepsilon)^2\, \delta^l_{j,0}\, \delta^l_{j\text{-}1,0}\}\ ,\ j = 1, \ldots, k \qquad \text{(B.8)}$$

The terms $h_3(\boldsymbol{\zeta}^l)$ and $h_4(\boldsymbol{\zeta}^l)$ which ensure the maintenance of constraint (22), are given by the equations,

$$h_3(\boldsymbol{\zeta}^l) = \kappa^l_3 \sum_{j=1}^{k(n+1)\text{-}1} \left(U(\varepsilon - \beta^l_j)\, (\beta^l_j - \varepsilon)^2\right) \qquad \text{(B.9)}$$

$$h_4(\boldsymbol{\zeta}^l) = \kappa^l_4 \sum_{j=1}^{k(n+1)\text{-}1} \left(U(\beta^l_j - (1\text{-}\varepsilon)\, \beta^l_0)\, (\beta^l_j - (1\text{-}\varepsilon)\, \beta^l_0)^2\right) \qquad \text{(B.10)}$$

where κ^l_3 and κ^l_4 are positively valued penalty factors.

B.1.2 Derivative of the objective function

The minimization of function $J(\boldsymbol{\zeta}^l)$ is facilitated by knowing the analytical expression of its derivative with respect to vector $\boldsymbol{\zeta}^l$. After algebraic manipulations it can be shown that,

$$\frac{dJ(\boldsymbol{\zeta}^l)}{d\boldsymbol{\zeta}^l} = 2\, \lambda^l_1 \{\tilde{\boldsymbol{\theta}}_1(\boldsymbol{\zeta}^l)\text{-}\boldsymbol{\theta}^l_1\}^T\, \mathbf{p}^l_1 + 2\, \lambda^l_2 \{\tilde{\boldsymbol{\theta}}_2(\boldsymbol{\zeta}^l)\text{-}\boldsymbol{\theta}^l_2\}^T\, \mathbf{p}^l_2 + 2\, \lambda^l_3 \{\tilde{\boldsymbol{\theta}}_3(\boldsymbol{\zeta}^l)\text{-}\boldsymbol{\theta}^l_3\}^T\, \mathbf{p}^l_3 + \boldsymbol{\varphi}^{l\,T}_1 +$$

$$2\, \kappa^l_2\, \tilde{\boldsymbol{\theta}}_4(\boldsymbol{\zeta}^l)^T\, \mathbf{p}^l_4 + \boldsymbol{\varphi}^{l\,T}_3 + \boldsymbol{\varphi}^{l\,T}_4 \qquad \text{(B.11)}$$

where $\mathbf{p}^l_1$, $\mathbf{p}^l_2$ and $\mathbf{p}^l_4$ are matrices with dimensions $k \times k(n+4)$, $\mathbf{p}^l_3$ is a matrix with dimensions $n+1 \times k(n+4)$ and $\boldsymbol{\varphi}^l_1$, $\boldsymbol{\varphi}^l_3$, $\boldsymbol{\varphi}^l_4$ are vectors with dimension $k(n+4)$.

The non-zero elements of matrix $\mathbf{p}^l_1$ are,

$$^{(l)}p_1{}^{i,\,i} = 2\, e^l_i\, \beta^l_0\ ,\ \ i = 1, \ldots, k \qquad \text{(B.12)}$$

$$^{(l)}p_1{}^{i,\,i+k} = 1\ ,\ \ i = 1, \ldots, k \qquad \text{(B.13)}$$

$$^{(l)}p_1{}^{i,\,3k+1} = (e^l_i)^2\ ,\ \ i = 1, \ldots, k \qquad \text{(B.14)}$$

The non-zero elements of matrix $\mathbf{p}^l_2$ are,

$$ {}^{(l)}p_2^{\,1,\,1} = e^l_k\,\beta^l_1 \tag{B.15} $$

$$ {}^{(l)}p_2^{\,1,\,k} = e^l_1\,\beta^l_1 \tag{B.16} $$

$$ {}^{(l)}p_2^{\,1,\,2k+1} = 1 \tag{B.17} $$

$$ {}^{(l)}p_2^{\,1,\,3k+2} = e^l_1\,e^l_k \tag{B.18} $$

$$ {}^{(l)}p_2^{\,i,\,i} = e^l_{i\text{-}1}\,\beta^l_1\ ,\ \ i = 2, \ldots, k \tag{B.19} $$

$$ {}^{(l)}p_2^{\,i,\,i\text{-}1} = e^l_i\,\beta^l_1\ ,\ \ i = 2, \ldots, k \tag{B.20} $$

$$ {}^{(l)}p_2^{\,i,\,i+2k} = 1\ ,\ \ i = 2, \ldots, k \tag{B.21} $$

$$ {}^{(l)}p_2^{\,i,\,3k+2} = e^l_{i\text{-}1}\,e^l_i\ ,\ \ i = 2, \ldots, k \tag{B.22} $$

The non-zero elements of matrix $\mathbf{p}^l_3$ are,

$$ {}^{(l)}p_3^{\,1,j} = 2\left[\ e^l_j\,\beta^l_0 + \left(\sum_{s=j+1}^{k}\ \ e^l_s\,\beta^l_{s\text{-}j}\right) + \left(\sum_{s=1}^{j\text{-}1}\ \ e^l_s\,\beta^l_{j\text{-}s}\right)\right]\ ,\ j = 1, \ldots, k \tag{B.23} $$

$$ {}^{(l)}p_3^{\,1,j} = 1\ ,\ j = k{+}1, \ldots, 2k \tag{B.24} $$

$$ {}^{(l)}p_3^{\,1,j} = 2\ ,\ j = 2k{+}2, \ldots, 3k \tag{B.25} $$

$$ {}^{(l)}p_3^{\,1,\,3k+1} = \sum_{s=1}^{k}\ \ (e^l_s)^2 \tag{B.26} $$

$$ {}^{(l)}p_3^{\,1,j} = \sum_{s=1}^{4k\text{-}j+1}\ \ e^l_s\,e^l_{s+j\text{-}3k\text{-}1}\ ,\ j = 3k{+}2, \ldots, 4k \tag{B.27} $$

$$ {}^{(l)}p_3^{\,2,\,2k+1} = 1 \tag{B.28} $$

$$ {}^{(l)}p_3^{\,i,j} = \sum_{s=1}^{k}\ \ \{e^l_s\,(\beta^l_{k(i\text{-}1)+j\text{-}s} + \beta^l_{k(i\text{-}1)\text{-}j+s})\}\ ,\ \ i = 2, \ldots, n{+}1;\ j = 1, \ldots, k \tag{B.29} $$

$$ {}^{(l)}p_3^{\,i,\,k(i+1)+r+1} = U(k\text{-}1\text{-}r)\left[\ \sum_{s=1}^{r}\ \ e^l_s\,e^l_{k+s\text{-}r}\right] + U(\text{-}|k\text{-}r|)\left[\ \sum_{s=1}^{k}\ \ (e^l_s)^2\right] + U(r\text{-}k\text{-}1) $$

$$\left[\sum_{s=r-k+1}^{k} e^l_s \, e^l_{s-r+k} \right], \; i = 2, \ldots, n+1; \; r = 1, \ldots, 2k-1 \tag{B.30}$$

The non-zero elements of matrix $\mathbf{p}^l_4$ are,

$${}^{(l)}p_4^{1,\,k+1} = -(1-\varepsilon)^2 \, \delta^l_{k,\,0} \tag{B.31}$$

$${}^{(l)}p_4^{i,\,k+i} = -(1-\varepsilon)^2 \, \delta^l_{i-1,\,0}, \; i = 2, \ldots, k \tag{B.32}$$

$${}^{(l)}p_4^{1,\,2k} = -(1-\varepsilon)^2 \, \delta^l_{1,\,0} \tag{B.33}$$

$${}^{(l)}p_4^{i,\,k+i-1} = -(1-\varepsilon)^2 \, \delta^l_{i,\,0}, \; i = 2, \ldots, k \tag{B.34}$$

$${}^{(l)}p_4^{i,\,2k+i} = 2 \, \delta^l_{i,\,1}, \; i = 1, \ldots, k \tag{B.35}$$

The non-zero elements of vector $\boldsymbol{\varphi}^l_1$ are,

$${}^{(l)}\varphi_1^{i+k} = 2 \, \kappa^l_1 \, U(\varepsilon - \delta^l_{i,\,0}) \, (\delta^l_{i,\,0} - \varepsilon), \; i = 1, \ldots, k \tag{B.36}$$

$${}^{(l)}\varphi_1^{3k+1} = 2 \, \kappa^l_1 \, U(\varepsilon - \beta^l_0) \, (\beta^l_0 - \varepsilon) \tag{B.37}$$

The non-zero elements of vector $\boldsymbol{\varphi}^l_3$ are,

$${}^{(l)}\varphi_3^{3k+1+i} = 2 \, \kappa^l_3 \, U(\varepsilon - \beta^l_i) \, (\beta^l_i - \varepsilon), \; i = 1, \ldots, k(n+1)-1 \tag{B.38}$$

The non-zero elements of vector $\boldsymbol{\varphi}^l_4$ are,

$${}^{(l)}\varphi_4^{3k+1} = -2 \, \kappa^l_4 \sum_{s=1}^{k(n+1)-1} \left[(1-\varepsilon) \, U(\beta^l_s - (1-\varepsilon) \, \beta^l_0) \, (\beta^l_s - (1-\varepsilon) \, \beta^l_0) \right] \tag{B.39}$$

$${}^{(l)}\varphi_4^{3k+1+i} = 2 \, \kappa^l_4 \, U(\beta^l_i - (1-\varepsilon) \, \beta^l_0) \, (\beta^l_i - (1-\varepsilon) \, \beta^l_0), \; i = 1, \ldots, k(n+1)-1 \tag{B.40}$$

B.2 Additional material for section 3.3

Estimating the autocovariance sequence of the stationary component process without optimisation

As we show in sub-section 3.2 of the paper, the vector ζ^l has $k(n+4)$ elements. This means, that the number of the unknown parameters of the problem increases linearly with the number of elements of the autocovariance sequence of the stochastic process Z^l_i that we want to preserve using nonlinear optimisation. This fact leads us to the conclusion that for large values of n the optimisation problem would become extremely demanding due to the computational time needed for the optimisation. For this particular case we have developed a fast algorithm based on generalised inversion (Marlow, 1993, p. 263). After having estimated the $k(n+1)-1$ ($n \geq 1$) elements of the autocovariance sequence (i.e. β^l_0, β^l_1, …, $\beta^l_{k(n+1)-1}$) of the stochastic process Y^l_i for each location l using nonlinear optimisation, we can estimate the next elements of the sequence (i.e. $\beta^l_{k(n+1)}$, …, $\beta^l_{k(q+1)-1}$, $q \geq n$) by solving a linear system of $q-n$ equations with $k(q-n)$ unknown parameters. Here, $q+1$ is the total number of elements of the autocovariance sequence of the stochastic process Z^l_i for each location l that we wish to preserve (i.e. Cov[Z^l_i, Z^l_i], Cov[Z^l_i, Z^l_{i-1}], …, Cov[Z^l_i, Z^l_{i-q}]), and $n+1$ is the number of Z^l_i autocovariances preserved using optimisation (i.e. Cov[Z^l_i, Z^l_i], …, Cov[Z^l_i, Z^l_{i-n}]). The algorithm that we developed gives us the advantage to minimize the number of the unknown parameters that need to be optimised by defining the minimum n (n = 1). The linear system that needs to be solved for each location l is given by the expression,

$$\mathbf{c}^l \, \boldsymbol{\beta}^l = \mathbf{d}^l \tag{B.41}$$

where $\mathbf{c}^l$ is a matrix with dimensions $(q-n) \times k(q-n)$, $\boldsymbol{\beta}^l = [\beta^l_{k(n+1)}, \beta^l_{k(n+1)+1}, \ldots, \beta^l_{k(q+1)-1}]^T$ is the vector of the unknown autocovariances of the stochastic process Y^l_i and $\mathbf{d}^l$ is a vector with $q-n$ known elements. The non-zero elements of matrix $\mathbf{c}^l$ can be obtained by equations,

$${}^{(l)}c^{1,j} = U(1-j)\left[\sum_{s=1}^{k} (e^l_s)^2\right] + U(j-2)\,U(k-j)\left[\sum_{s=1}^{k-j+1} e^l_s\, e^l_{s+j-1}\right], \; j = 1, \ldots, k(q-n) \tag{B.42}$$

$${}^{(l)}c^{i,(i-2)k+r+1} = U(k-1-r)\left[\sum_{s=1}^{r} e^l_s\, e^l_{s+k-r}\right] + U(-|k-r|)\left[\sum_{s=1}^{k} (e^l_s)^2\right] + U(r-k-1)$$

$$\left[\sum_{s=1}^{2k-r} e^l_s\, e^l_{s+r-k}\right], \; i = 2, \ldots, q-n; \; r = 1, \ldots, 2k-1 \tag{B.43}$$

and the elements of vector $\mathbf{d}^l$ can be obtained by equation,

$$^{(l)}d^{\,j} = \mathrm{Cov}[Z^l_i, Z^l_{i-(n+j)}] - U(1-j) \sum_{r=1}^{k-1} \left\{ \left[\sum_{s=1}^{k-r} e^l_s\, e^l_{s+r} \right] \beta_{kn+r} \right\} \;,\; j = 1, \ldots, q-n \qquad \text{(B.44)}$$

Following the algorithm of generalised inversion (Marlow, 1993, p. 263), at first we solve the linear system,

$$\mathbf{c}^l\, (\mathbf{c}^l)^T\, \boldsymbol{\lambda}^l = \mathbf{d}^l \qquad \text{(B.45)}$$

in order to obtain the vector $\boldsymbol{\lambda}^l$ for each location l and then estimate the vector $\boldsymbol{\beta}^l$ using the equation

$$\boldsymbol{\beta}^l = (\mathbf{c}^l)^T\, \boldsymbol{\lambda}^l \qquad \text{(B.46)}$$

It can be easily shown that the matrix $\mathbf{c}^l(\mathbf{c}^l)^T$ is tridiagonal. In this case the linear system (B.45) can be solved using the Thomas algorithm (Chapra *et al.*, 2002, p. 286-287), which is a very fast and easy algorithm for solving linear tridiagonal systems.

B.3 Additional material for section 3.4

Here we will study an algorithm for modifying a non-feasible autocovariance matrix to feasible, which is applied to analytical estimation of SMA parameters. The SMA coefficients a^l_j (j = 0, 1, …, $k(q+1)$-1) of each location l (l = 1, …, v) can be estimated from the autocovariance sequence (i.e. β^l_0, β^l_1, …, $\beta^l_{k(q+1)-1}$) of the stochastic process Y^l_i. If the autocovariance matrix of Y^l_i is feasible, then the SMA coefficients a^l_j of each location l can be analytically estimated using equations (25), (26) and (27). If the autocovariance matrix of the stochastic process Y^l_i is not feasible then the constraint,

$$s^l_\beta(\omega) \geq 0 \;,\; \forall\, \omega \in [0, ½];\; l = 1, \ldots, v \qquad \text{(B.47)}$$

needed by the expression (26), is not satisfied. In order to avoid nonlinear optimisation, which is a time demanding procedure, we developed a simple algorithm that modifies the autocovariance sequence β^l_p (p = 0, 1, …, $k(q+1)-1$) of each location l to another sequence $\tilde{\beta}^l_p$ (p = 0, 1, …, $k(q+1)-1$) that is slightly different from the sequence β^l_p, obeys the constraint (B.47) and eliminates the deviation of autocovariances $\tilde{\beta}^l_0$ and $\tilde{\beta}^l_1$ from the autocovariances β^l_0 and β^l_1 respectively. The later property of the algorithm seems to be extremely useful for the Split model, as long as the autocovariances β^l_0 and β^l_1 of each location l affect not only the annual variance and annual autocovariances of location l (equations (18) and (19)), but also the seasonal variances and lag one seasonal autocovariances of each location l (equations (15) and (17)). The algorithm has the following steps:

1) We estimate the power spectrum $s^l_\beta(\omega)$ of the autocovariance sequence β^l_p (p= 0, 1, …, $k(q+1)-1$) of each location l.
2) We estimate the modified power spectrum $s^l_{\tilde{\beta}}(\omega)$ of the autocovariance sequence $\tilde{\beta}^l_p$ (p = 0, 1, …, $k(q+1)-1$) of each location l using the expression

$$s^l_{\tilde{\beta}}(\omega) = \max\{s^l_\beta(\omega), \varepsilon_1\}\ ,\ \ \forall\ \omega \in [0, ½];\ \ \varepsilon_1 \geq 0 \tag{B.48}$$

where ε_1 is a small positive number (e.g. ε_1 = 0.001).

3) We estimate the modified autocovariance sequence $\tilde{\beta}^l_p$ (p = 0, 1, …, $k(q+1)-1$) of each location l (l = 1, …, v) using Fourier transform,

$$\tilde{\beta}^l_j = \int_0^{1/2} s^l_{\tilde{\beta}}(\omega) \cos(2\pi j\omega)\mathrm{d}\omega \quad ,\ j = 0, 1, 2, \ldots, k(q+1)-1 \tag{B.49}$$

4) We replace the elements of the initial sequence β^l_p (p = 2, …, $k(q+1)-1$) with the corresponding elements of the obtained sequence $\tilde{\beta}^l_p$. The first two elements β^l_0 and β^l_1 are not changed in order not to affect the results of the first step of the estimation algorithm.

5) We repeat steps 1-4 until the deviation of autocovariances $\tilde{\beta}^l_0$ and $\tilde{\beta}^l_1$ from the autocovariances β^l_0 and β^l_1 respectively is negligible (usually this requires up to three iterations).

As we can observe at Figure B.1, the algorithm works well enough and the modifications performed are very small.

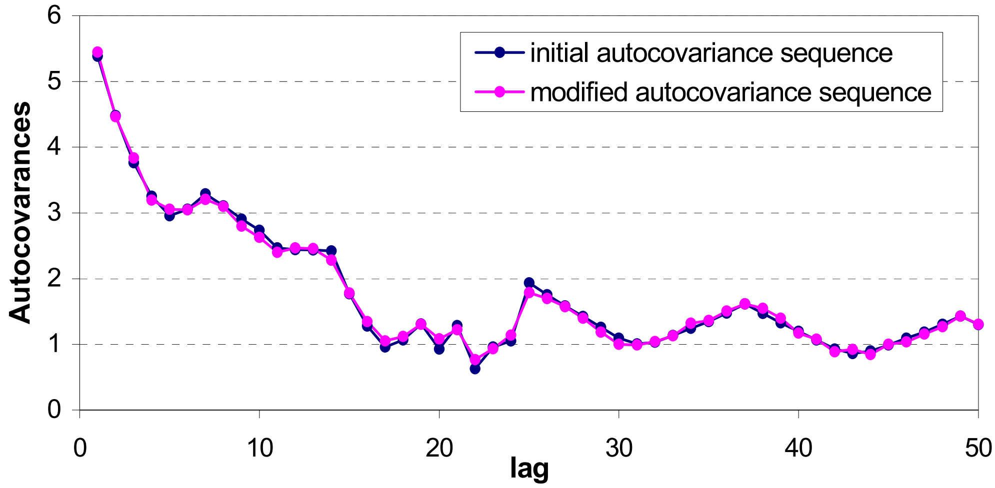

Figure B.1 Initial and modified (by the algorithm) autocovariance sequence for the application of Split model to the Boetikos Kephisos river flow (see paper)

B.4 Additional material for section 3.5

B.4.1 Objective function

We define the vector $\boldsymbol{\psi}$ containing $(3k+1)v^2+(2k+1)v$ unknown elements, which are

$$\psi^j = b^{l,r}, \quad l, r = 1, \ldots, v; \quad j = (l-1)v+r \tag{B.50}$$

$$\psi^j = {}^0f_s^{l,r}, \quad l, r = 1, \ldots, v; \quad s = 1, \ldots, k; \quad j = v^2(2s-1)+(l-1)v+r \tag{B.51}$$

$$\psi^j = {}^1f_s^{l,r}, \quad l, r = 1, \ldots, v; \quad s = 1, \ldots, k; \quad j = 2sv^2+(l-1)v+r \tag{B.52}$$

$$\psi^j = \mu_3[R_s^l], \quad l = 1, \ldots, v; \quad s = 1, \ldots, k; \quad j = (2k+1)v^2+(s-1)v+l \tag{B.53}$$

$$\psi^j = \mu_3[H^l], \quad l = 1, \ldots, v; \quad j = (2k+1)v^2+kv+l \tag{B.54}$$

$$\psi^j = g_s^{l,r}, \quad l, r = 1, \ldots, v; \quad s = 1, \ldots, k; \quad j = (2k+s)v^2+(k+l)v+r \tag{B.55}$$

$$\psi^j = \mu_3[G_s^l], \quad l = 1, \ldots, v; \quad s = 1, \ldots, k; \quad j = (3k+1)v^2+(k+s)v+l \tag{B.56}$$

We also define the vector $\boldsymbol{\theta}_4$ with dimension $kv(v+1)/2$ and the vectors $\boldsymbol{\theta}_5$, $\boldsymbol{\theta}_6$ with dimension kv containing the statistical properties of the stochastic processes X_i^l, with elements

$$\theta_4^j = \mathrm{Cov}[X_s^l, X_s^r] \ , \ l, r = 1, \ldots, v; \ s = 1, \ldots, k \ \text{ and}$$

$$j = (s\text{-}1)\frac{v(v+1)}{2} + \left[\sum_{i=1}^{l-1} (v\text{-}i+1)\right] + r - l + 1 \tag{B.57}$$

$$\theta_5^j = \mathrm{Cov}[X^l_s, X^l_{s\text{-}1}] \ , \ l = 1, \ldots, v; \ s = 1, \ldots, k; \ j = (s\text{-}1)v+l \tag{B.58}$$

$$\theta_6^j = \mu_3[X^l_s] \ , \ l = 1, \ldots, v; \ s = 1, \ldots, k; \ j = (s\text{-}1)v+l \tag{B.59}$$

In addition, we define the vectors $\tilde{\boldsymbol{\theta}}_4(\boldsymbol{\psi})$, $\tilde{\boldsymbol{\theta}}_5(\boldsymbol{\psi})$ and $\tilde{\boldsymbol{\theta}}_6(\boldsymbol{\psi})$ identically to vectors $\boldsymbol{\theta}_4$, $\boldsymbol{\theta}_5$, $\boldsymbol{\theta}_6$. $\tilde{\boldsymbol{\theta}}_4(\boldsymbol{\psi})$, $\tilde{\boldsymbol{\theta}}_5(\boldsymbol{\psi})$ and $\tilde{\boldsymbol{\theta}}_6(\boldsymbol{\psi})$ can be estimated from equations (29)-(35). In this case, the objective function that we propose is,

$$\Phi(\boldsymbol{\psi}_*) = \min[\Phi(\boldsymbol{\psi})] = \lambda_4 \left\| \tilde{\boldsymbol{\theta}}_4(\boldsymbol{\psi}) - \boldsymbol{\theta}_4 \right\|^2 + \lambda_5 \left\| \tilde{\boldsymbol{\theta}}_5(\boldsymbol{\psi}) - \boldsymbol{\theta}_5 \right\|^2 + \lambda_6 \left\| \tilde{\boldsymbol{\theta}}_6(\boldsymbol{\psi}) - \boldsymbol{\theta}_6 \right\|^2 +$$

$$h_5(\boldsymbol{\psi}) + h_6(\boldsymbol{\psi}) \tag{B.60}$$

where λ_4, λ_5 and λ_6 are positive weighting factors and $h_5(\boldsymbol{\psi})$, $h_6(\boldsymbol{\psi})$ are penalty terms positively valued if constraints (34), (37) and (38) are not satisfied. The value of the objective function (B.60) will be zero in case that equations (29)-(35) and constraints (34), (37) and (38) are satisfied simultaneously. Otherwise, the objective function will be positively valued. The term $h_5(\boldsymbol{\psi})$ which ensures the holding of constraint (34) is,

$$h_5(\boldsymbol{\psi}) = \kappa_5 \sum_{i=1}^{v} (\tilde{u}^{i,i}\text{-}1)^2 \tag{B.61}$$

where κ_5 is a positively valued penalty factor. The term $h_6(\boldsymbol{\psi})$ which ensures the holding of constraints (37) and (38) is,

$$h_6(\boldsymbol{\psi}) = \kappa_6 \left[\sum_{j=(2k+1)v^2+1}^{(2k+1)v^2+(k+1)v} \eta_j + \sum_{j=(3k+1)v^2+(k+1)v+1}^{(3k+1)v^2+(2k+1)v} \eta_j \right] \tag{B.62}$$

$$\eta_j = U[(\psi^j)^2-(\xi_{max})^2]\,[(\psi^j)^2-(\xi_{max})^2] \tag{B.63}$$

where κ_6 is a positively valued penalty factor.

B.3.2 Derivative of the objective function

In order to facilitate the minimization of function $\Phi(\boldsymbol{\psi})$, we derive the analytical expression of its derivative with respect to vector $\boldsymbol{\psi}$,

$$\frac{d\Phi(\boldsymbol{\psi})}{d\boldsymbol{\psi}} = 2\,\lambda_4\{\tilde{\boldsymbol{\theta}}_4(\boldsymbol{\psi})-\boldsymbol{\theta}_4\}^T\,\mathbf{p}_4 + 2\,\lambda_5\{\tilde{\boldsymbol{\theta}}_5(\boldsymbol{\psi})-\boldsymbol{\theta}_5\}^T\,\mathbf{p}_5 + 2\,\lambda_6\{\tilde{\boldsymbol{\theta}}_6(\boldsymbol{\psi})-\boldsymbol{\theta}_6\}^T\,\mathbf{p}_6 + \boldsymbol{\varphi}_5^{\,T} + \boldsymbol{\varphi}_6^{\,T} \tag{B.64}$$

where $\mathbf{p}_4$ is a matrix with dimensions $\{kv(v+1)/2\}\times\{(3k+1)v^2+(2k+1)v\}$, $\mathbf{p}_5$ and $\mathbf{p}_6$ are matrices with dimensions $kv\times\{(3k+1)v^2+(2k+1)v\}$ and $\boldsymbol{\varphi}_5$, $\boldsymbol{\varphi}_6$ are vectors with dimension $(3k+1)v^2+(2k+1)v$. The non-zero elements of matrix $\mathbf{p}_4$ are,

$$p_4^{\,i,j} = 2\,b^{l,m}\left[\sum_{t=-k(m+1)+1}^{k(m+1)-1}(a^l_{|t|})^2\right](e^l_s)^2\,,\;\; l=r;\;\; m=1,\ldots,v;\;\; j=(l-1)v+m \tag{B.65}$$

$$p_4^{\,i,j} = 2\;{}^0\!f_s^{\,l,m}\,,\;\; l=r;\;\; m=1,\ldots,v;\;\; j=(2s-1)v^2+(l-1)v+m \tag{B.66}$$

$$p_4^{\,i,j} = 2\;{}^1\!f_s^{\,l,m}\,,\;\; l=r;\;\; m=1,\ldots,v;\;\; j=2sv^2+(l-1)v+m \tag{B.67}$$

$$p_4^{\,i,j} = 2\,g_s^{\,l,m}\,,\;\; l=r;\;\; m=1,\ldots,v;\;\; j=(2k+s)v^2+(k+l)v+m \tag{B.68}$$

$$p_4^{\,i,j} = b^{r,m}\left[\sum_{t=-k(m+1)+1}^{k(m+1)-1}a^l_{|t|}\,a^r_{|t|}\right]e^l_s\,e^r_s\,,\;\; l\neq r;\;\; m=1,\ldots,v;\;\; j=(l-1)v+m \tag{B.69}$$

$$p_4^{\,i,j} = b^{l,m}\left[\sum_{t=-k(m+1)+1}^{k(m+1)-1}a^l_{|t|}\,a^r_{|t|}\right]e^l_s\,e^r_s\,,\;\; l\neq r;\;\; m=1,\ldots,v;\;\; j=(r-1)v+m \tag{B.70}$$

$$p_4^{\,i,j} = {}^0\!f_s^{\,r,m}\,,\;\; l\neq r;\;\; m=1,\ldots,v;\;\; j=(2s-1)v^2+(l-1)v+m \tag{B.71}$$

$$p_4^{\,i,j} = {}^0\!f_s^{\,l,m}\,,\;\; l\neq r;\;\; m=1,\ldots,v;\;\; j=(2s-1)v^2+(r-1)v+m \tag{B.72}$$

$$p_4^{\,i,j} = {}^1\!f_s^{\,r,m}\,,\;\; l\neq r;\;\; m=1,\ldots,v;\;\; j=2sv^2+(l-1)v+m \tag{B.73}$$

$$p_4^{i,j} = {}^1f_s^{l,m}\ ,\ l \neq r;\ m = 1, \ldots, v;\ j = 2sv^2+(r\text{-}1)v+m \tag{B.74}$$

$$p_4^{i,j} = g_s^{r,m}\ ,\ l \neq r;\ m = 1, \ldots, v;\ j = (2k+s)v^2+(k+l)v+m \tag{B.75}$$

$$p_4^{i,j} = g_s^{l,m}\ ,\ l \neq r;\ m = 1, \ldots, v;\ j = (2k+s)v^2+(k+r)v+m \tag{B.76}$$

where,

$$i = (s\text{-}1)\,\frac{v(v+1)}{2} + \left[\sum_{t=1}^{l\text{-}1} (v\text{-}t+1)\right] + r - l + 1\ ,\ l, r = 1, \ldots, v;\ s = 1, \ldots, k \tag{B.77}$$

The non-zero elements of matrix $\mathbf{p}_5$ are,

$$p_5^{i,j} = 2\ b^{l,m}\left[\sum_{t=-k(m+1)+1}^{k(m+1)\text{-}2} a^l_{|t|}\,a^l_{|t+1|}\right] e^l_k\, e^l_1\ ,\ s = 1\ ,\ m = 1, \ldots, v \text{ and } j = (l\text{-}1)v+m \tag{B.78}$$

$$p_5^{i,j} = {}^1f_k^{l,m}\ ,\ s = 1\ ,\ m = 1, \ldots, v \text{ and } j = v^2+(l\text{-}1)v+m \tag{B.79}$$

$$p_5^{i,j} = {}^0f_1^{l,m}\ ,\ s = 1\ ,\ m = 1, \ldots, v \text{ and } j = 2kv^2+(l\text{-}1)v+m \tag{B.80}$$

$$p_5^{i,j} = 2\ b^{l,m}\left[\sum_{t=-k(m+1)+1}^{k(m+1)\text{-}2} a^l_{|t|}\,a^l_{|t+1|}\right] e^l_{s\text{-}1}\, e^l_s\ ,\ s \neq 1\ ,\ m = 1, \ldots, v \text{ and } j = (l\text{-}1)v+m \tag{B.81}$$

$$p_5^{i,j} = {}^1f_{s\text{-}1}^{l,m}\ ,\ s \neq 1\ ,\ m = 1, \ldots, v \text{ and } j = (2s\text{-}1)v^2+(l\text{-}1)v+m \tag{B.82}$$

$$p_5^{i,j} = {}^0f_s^{l,m}\ ,\ s \neq 1\ ,\ m = 1, \ldots, v \text{ and } j = 2(s\text{-}1)v^2+(l\text{-}1)v+m \tag{B.83}$$

where,

$$i = (s\text{-}1)v+l\ ,\ l = 1, \ldots, v\ ,\ s = 1, \ldots, k \tag{B.84}$$

The non-zero elements of matrix $\mathbf{p}_6$ are,

$$p_6^{i,j} = 3\ (b^{l,m})^2\,\mu_3[H^m]\left[\sum_{t=-k(m+1)+1}^{k(m+1)\text{-}1} (a^l_{|t|})^3\right](e^l_s)^3\ ,\ m = 1, \ldots, v;\ j = (l\text{-}1)v+m \tag{B.85}$$

$$p_6^{i,j} = (b^{l,m})^3\left[\sum_{t=-k(m+1)+1}^{k(m+1)\text{-}1} (a^l_{|t|})^3\right](e^l_s)^3\ ,\ m = 1, \ldots, v;\ j = (2k+1)v^2+kv+m \tag{B.86}$$

$$p_6^{i,j} = 3\,(g_s^{l,m})^2\,\mu_3[G_s^m]\;,\;\; m = 1, \ldots, v;\;\; j = (2k+s)v^2+(k+l)v+m \tag{B.87}$$

$$p_6^{i,j} = (g_s^{l,m})^3\;,\;\; m = 1, \ldots, v;\;\; j = (3k+1)v^2+(k+s)v+m \tag{B.88}$$

$$p_6^{i,j} = 3\,({}^0f_s^{l,m})^2\,\mu_3[R_s^m]\;\;,\;\; m = 1, \ldots, v \text{ and } j = (2s\text{-}1)v^2+(l\text{-}1)v+m \tag{89}$$

$$p_6^{i,j} = ({}^0f_s^{l,m})^3\;,\;\; m = 1, \ldots, v;\;\; j = (2k+1)v^2+(s\text{-}1)v+m \tag{B.90}$$

$$p_6^{i,j} = 3\,({}^1f_s^{l,m})^2\,\mu_3[R^m{}_{s+1}]\;,\;\; s \neq k\;;\;\; m = 1, \ldots, v;\;\; j = 2sv^2+(l\text{-}1)v+m \tag{B.91}$$

$$p_6^{i,j} = ({}^1f_s^{l,m})^3\;,\;\; s \neq k\;;\;\; m = 1, \ldots, v;\;\; j = (2k+1)v^2+sv+m \tag{B.92}$$

$$p_6^{i,j} = 3\,({}^1f_k^{l,m})^2\,\mu_3[R_1^m]\;,\;\; s = k\;;\;\; m = 1, \ldots, v;\;\; j = 2kv^2+(l\text{-}1)v+m \tag{B.93}$$

$$p_6^{i,j} = ({}^1f_k^{l,m})^3\;,\;\; s = k\;;\;\; m = 1, \ldots, v;\;\; j = (2k+1)v^2+m \tag{B.94}$$

where,

$$i = (s\text{-}1)v+l\;,\;\; l = 1, \ldots, v;\;\; s = 1, \ldots, k \tag{B.95}$$

The non-zero elements of vector $\boldsymbol{\varphi}_5$ are,

$$\varphi_5{}^j = 4\,\kappa_5\,(\tilde{u}^{l,l}\text{-}1)\,b^{l,m}\;,\;\; l, m = 1, \ldots, v;\;\; j = (l\text{-}1)v+m \tag{B.96}$$

The non-zero elements of vector $\boldsymbol{\varphi}_6$ are,

$$\varphi_6{}^j = 2\,\kappa_6\,\psi^j\,U[(\psi^j)^2\text{-}(\xi_{max})^2] \tag{B.97}$$

where,

$$j = (2k+1)v^2+1, \ldots, (2k+1)v^2+(k+1)v, (3k+1)v^2+(k+1)v+1, \ldots, (3k+1)v^2+(2k+1)v \tag{B.98}$$

References

Chapra, S.C. & Canale, R.P. (2002) *Numerical Methods for Engineers*, Fourth Edition, McGraw-Hill, New York, U.S.A.

Marlow, W.H. (1993) *Mathematics for Operations Research*. Dover Publications, Mineola, New York, U.S.A.