European Geosciences Union General Assembly 2010
Vienna, Austria, 2-7 May 2010
Session HS5.5: Stochastics in hydrometeorological processes: from point to global spatial scales and from minute to climatic time scales

Performance evaluation and interdependence of parameter estimators of the Hurst-Kolmogorov stochastic process

H. Tyralis and D. Koutsoyiannis
Department of Water Resources and Environmental Engineering
National Technical University of Athens
(www.itia.ntua.gr/dk)

1. Abstract

We investigate three methods for simultaneous estimation of the Hurst parameter (H) and the standard deviation (σ) for a Hurst-Kolmogorov stochastic process, namely the maximum likelihood method and two methods based on the variation of the standard deviation or of the variance with time scale. We show that the simultaneous estimation of the two parameters is important, albeit not given appropriate attention in the literature, because of the interdependence of the two parameter estimators. In addition, we test the performance of the three methods for a range of sample sizes and H values, through a simulation study and we compare it with other known results for other estimators of the literature.

2. Definition of Hurst-Kolmogorov process

- Let X_i be a Gaussian stationary stochastic process with $i = 1, 2, \ldots$ denoting discrete time. We form the vector of identically distributed variables $\boldsymbol{X}_n := (X_1, \ldots, X_n)$.
- The mean of the process is $\mu := \mathrm{E}[X_i]$ and the variance of the process is $\sigma^2 := \mathrm{Var}[X_i]$.
- Let k be an integer that represents a timescale larger than 1, the original time scale of the process X_i. The mean aggregated stochastic process on that timescale is:

$$X_i^{(k)} := (1/k) \sum_{l=(i-1)k+1}^{ik} X_l$$

- The following equation defines the Hurst-Kolmogorov process (HKp)

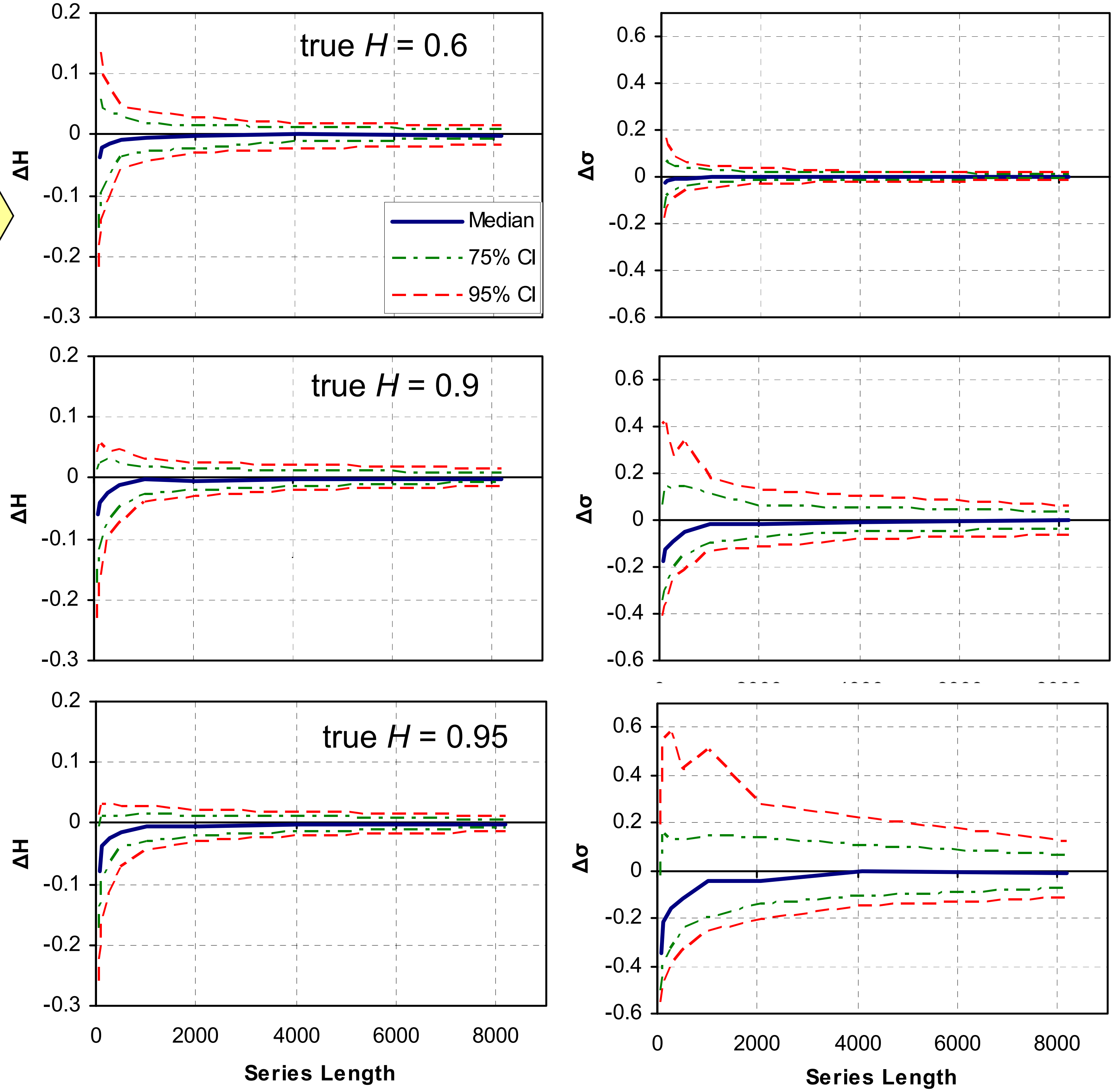
$$(X_i^{(k)} - \mu) \stackrel{\mathrm{d}}{=} \left(\frac{k}{l}\right)^H (X_j^{(l)} - \mu),\ 0 < H < 1$$

- The autocorrelation function of the mean aggregated stochastic process for any aggregated timescale k is independent of k and given by the following equation

$$\rho_j^{(k)} = \rho_j = |j+1|^{2H}/2 + |j-1|^{2H}/2 - |j|^{2H}$$

3. The maximum likelihood (ML) estimator

Mean of estimated H's or σ's minus the true value and corresponding Monte Carlo confidence intervals



ΔH = mean of estimated H's − true H
$\Delta\sigma$ = mean of estimated σ's − true σ

The following equations are the ML estimates of the HKp parameters

$$\hat{\mu} = \frac{\boldsymbol{x}_n^{\mathrm{T}} \hat{\boldsymbol{R}}^{-1} \boldsymbol{e}}{\boldsymbol{e}^{\mathrm{T}} \hat{\boldsymbol{R}}^{-1} \boldsymbol{e}}$$

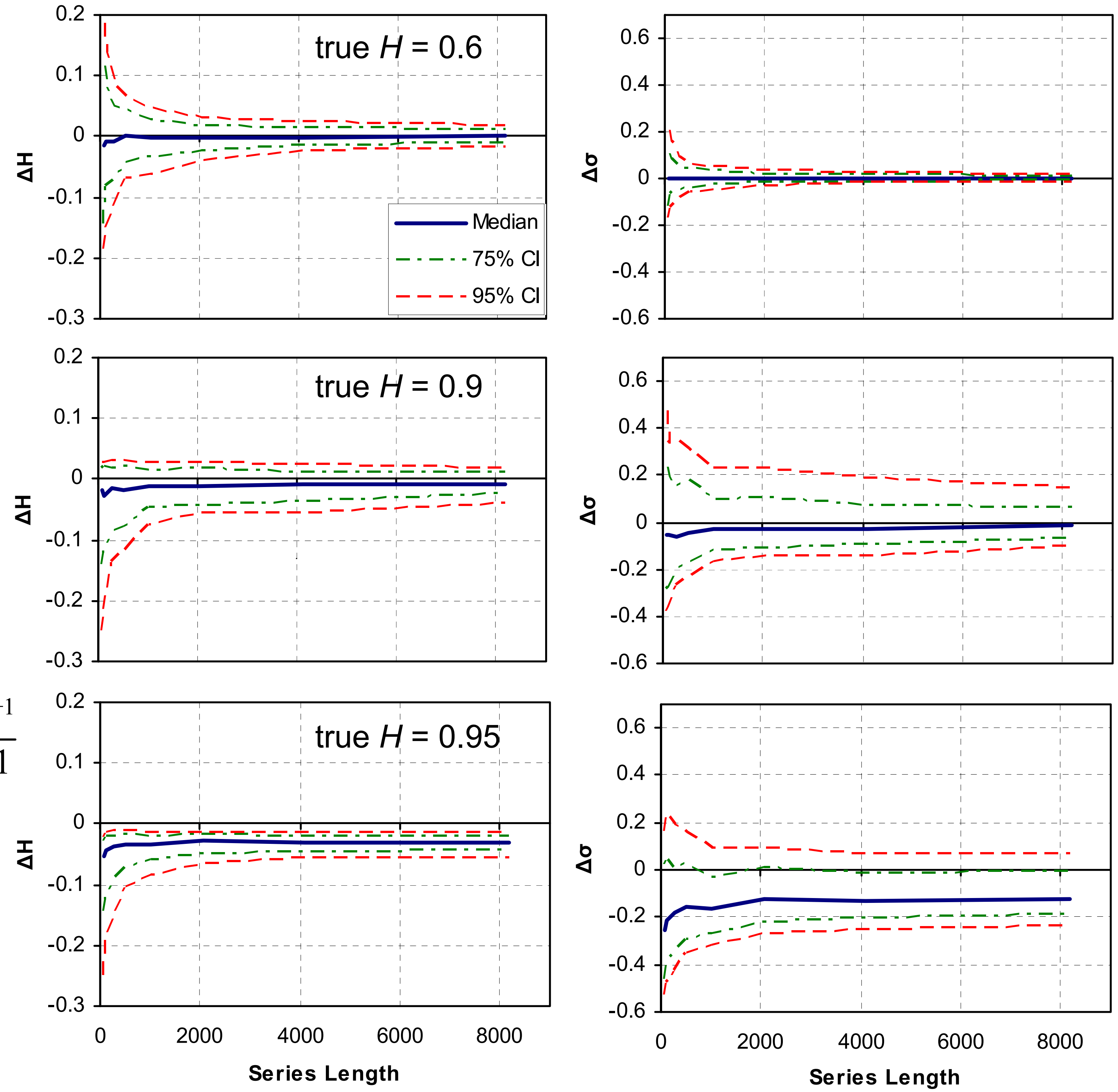
$$\hat{\sigma} = \sqrt{\frac{(\boldsymbol{x}_n - \hat{\mu}\, \boldsymbol{e})^{\mathrm{T}} \hat{\boldsymbol{R}}^{-1} (\boldsymbol{x}_n - \hat{\mu}\, \boldsymbol{e})}{n}}$$

where $\boldsymbol{e}$ is a nx1 vector with all its elements equal to 1, $\boldsymbol{R}$ is the autocorrelation matrix which is a function of H ("^" denotes estimate); the ML estimate of H is obtained by maximizing the one-variable function

$$g_1(H) := -\frac{n}{2} \ln\Big[\Big(\boldsymbol{x}_n - \frac{\boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{e}}{\boldsymbol{e}^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{e}}\boldsymbol{e}\Big)^{\mathrm{T}} \boldsymbol{R}^{-1} \Big(\boldsymbol{x}_n - \frac{\boldsymbol{x}_n^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{e}}{\boldsymbol{e}^{\mathrm{T}} \boldsymbol{R}^{-1} \boldsymbol{e}}\boldsymbol{e}\Big)\Big] - \frac{1}{2} \ln[\det(\boldsymbol{R})]$$

4. The Least Squares based on Standard Deviation (LSSD) estimator

Mean of estimated H's or σ's minus the true value and corresponding Monte Carlo confidence intervals



LSSD estimates of the HKp parameters σ and H are given by minimizing the following two variables function

$$e^2(\sigma, H) := \sum_{k=1}^{k'} \frac{[\ln\sigma + H \cdot \ln k + \ln c_k(H) - \ln s^{(k)}]^2}{k^p} + \frac{H^{q+1}}{q+1}$$

where

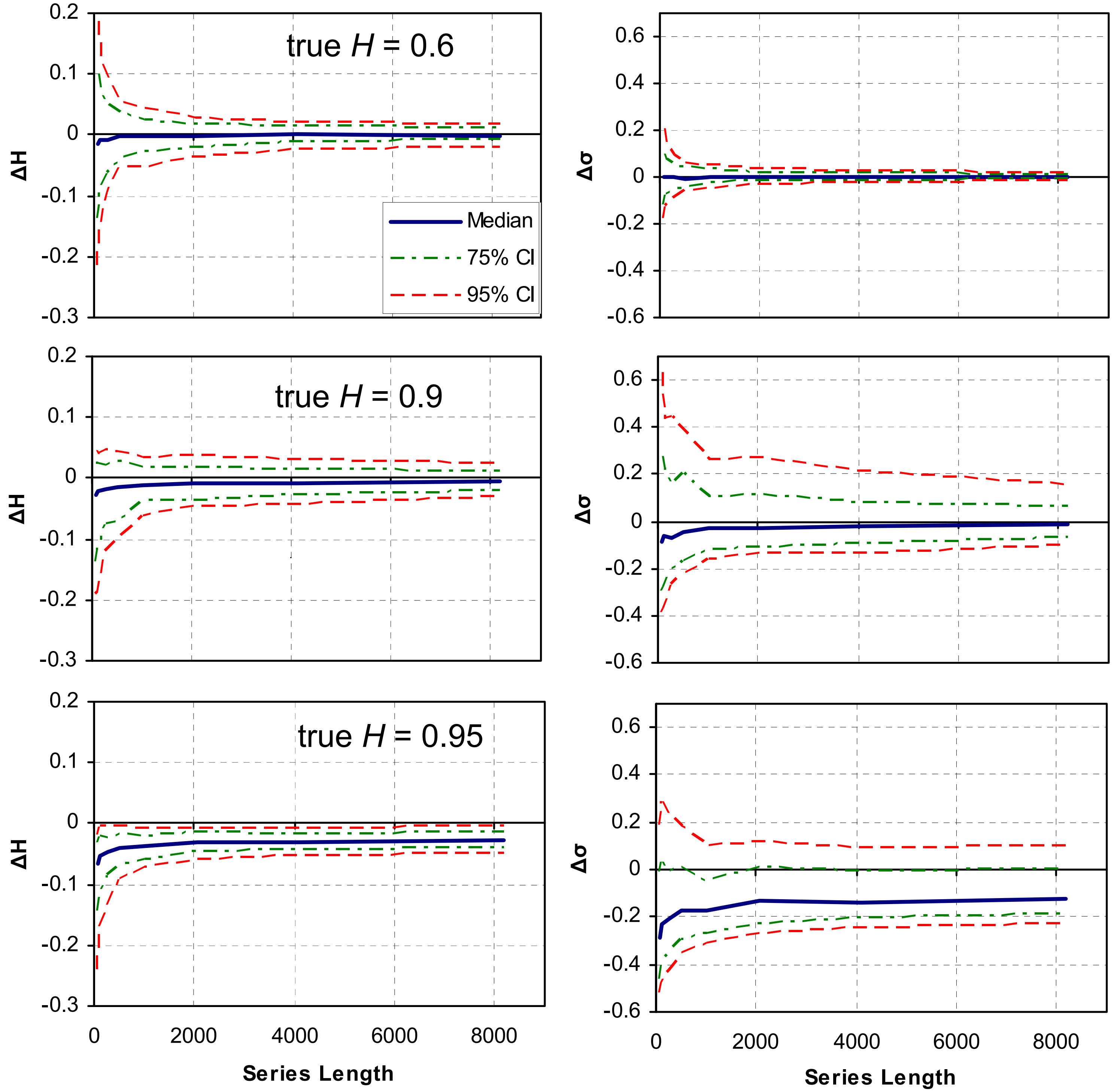
$$c_k(H) := \sqrt{\frac{n/k - (n/k)^{2H-1}}{n/k - 1/2}}$$

ΔH = mean of estimated H's − true H
$\Delta\sigma$ = mean of estimated σ's − true σ

See details about the LSSD method in Koutsoyiannis, D. (2003) Climate change, the Hurst phenomenon, and hydrological statistics. *Hydrol. Sci. J.* **48**(1), 3-24.

5. The Least Squares based on Variance (LSV) estimator

Mean of estimated H's or σ's minus the true value and corresponding Monte Carlo confidence intervals



ΔH = mean of estimated H's − true H
$\Delta\sigma$ = mean of estimated σ's − true σ

The LSV σ estimate of the HKp is given by the following equations

$$\hat{\sigma} = \sqrt{\alpha_{12}(\hat{H})/\alpha_{11}(\hat{H})}$$

where

$$\alpha_{11}(H) := \sum_{k=1}^{k'} \frac{c_k^2(H)\, k^{4H}}{k^p} \quad \alpha_{12}(H) := \sum_{k=1}^{k'} \frac{c_k(H)\, k^{2H}\, s^{2(k)}}{k^p}$$

$$c_k(H) := \frac{(n/k) - (n/k)^{2H-1}}{(n/k) - 1}$$

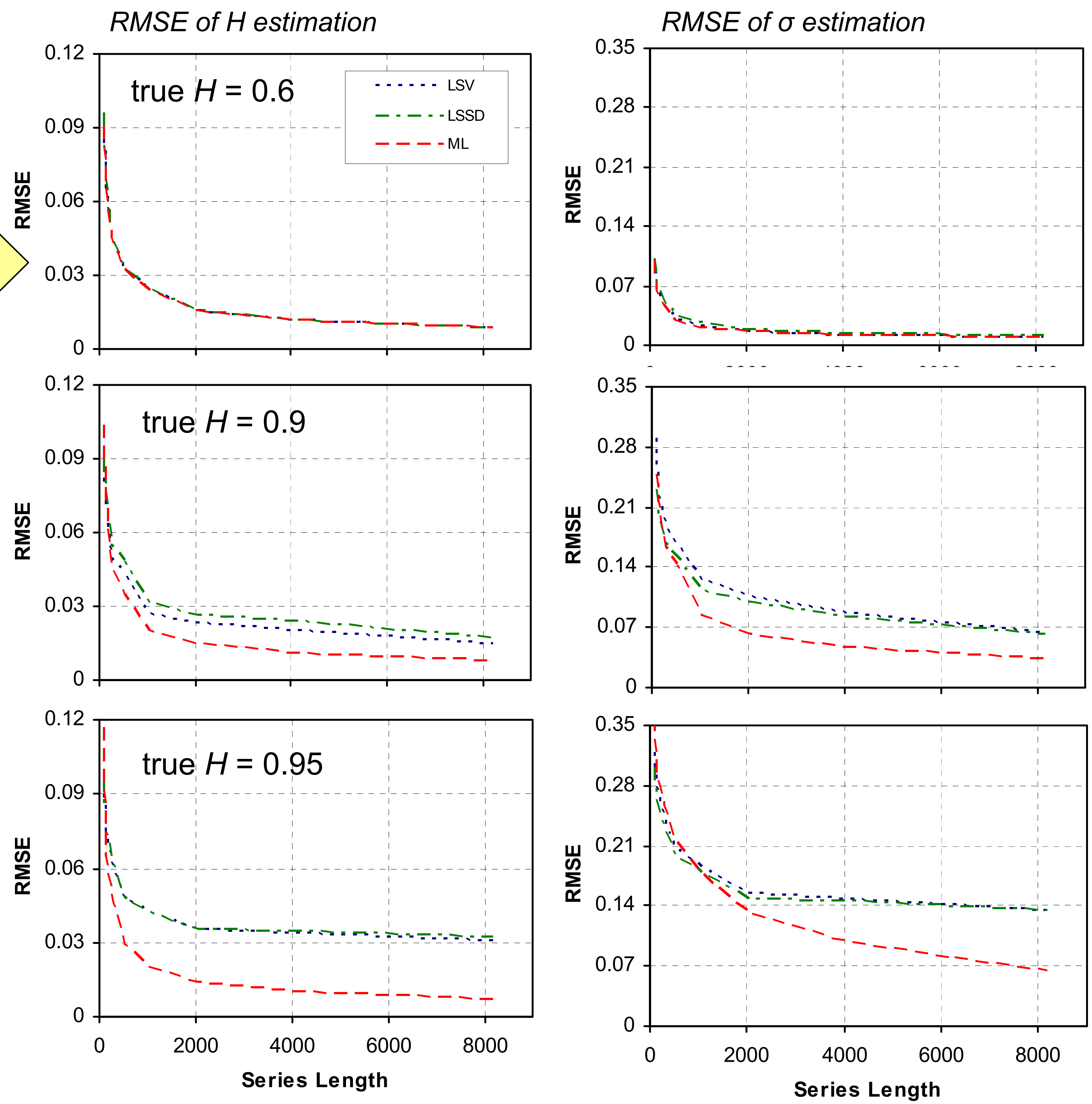
The H estimate is given by minimizing the function

$$g_2(H) := \sum_{k=1}^{k'} \frac{s^{4(k)}}{k^p} - \frac{\alpha_{12}^2(H)}{\alpha_{11}(H)} + \frac{H^{q+1}}{q+1},\ 0 < H < 1$$

See details about the LSV method in Tyralis, H. and Koutsoyiannis D. (2010) Simultaneous estimation of the Hurst-Kolmogorov parameters. *Stoch. Environ. Res. Risk Assess.* In review.

6. Comparison of all three estimators

Monte Carlo RMSE for the three different methods versus the sample size



Performance of the three estimators in terms of RMSE where

$$\text{RMSE} := \sqrt{\frac{1}{K}\sum_{k=1}^{K}(H_k - H)^2}$$

The ML estimator has the lowest RMSE

ƛ. Comparison between ML, LSSD, LSV and alternative methods

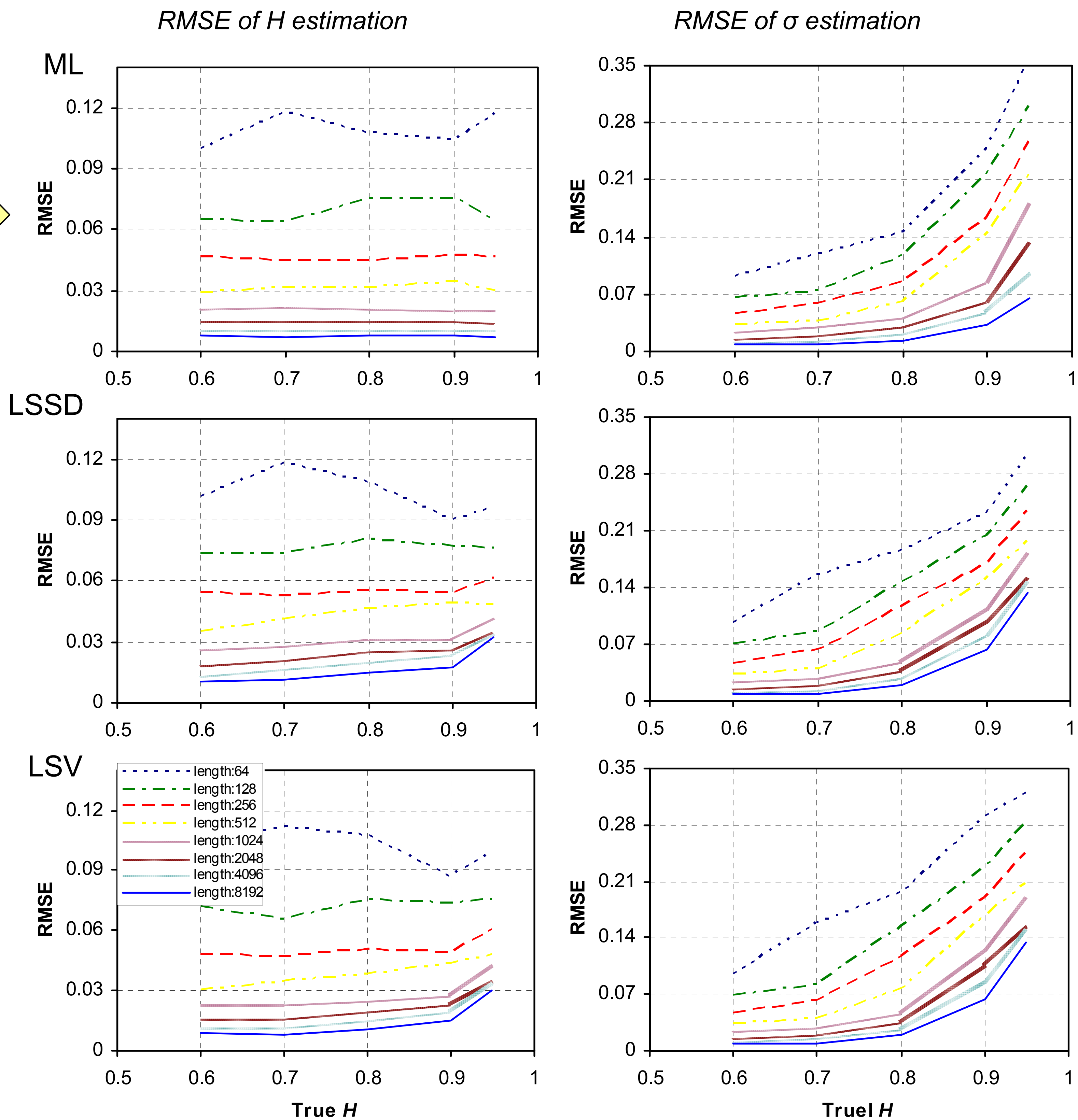
Estimation method		True H				Estimation method	True H				
		0.6	0.7	0.8	0.9		0.6	0.7	0.8	0.9	
Variance	H	.595	.687	.775	.850	Periodogram	.604	.708	.809	.912	
	RMSE	.027	.030	.036	.057		.024	.024	.026	.027	
DiffVar	H	.567	.667	.771	.864	Modified Periodogram	.565	.661	.752	.847	Worst/
	RMSE	.080	.076	.073	.070		.051	.054	.060	.063	2nd worst
Absolute	H	.594	.686	.775	.849	Local Whittle	.601	.700	.804	.902	
	RMSE	.029	.031	.038	.059		.023	.023	.023	.021	
Higuchi	H	.599	.696	.798	.888	Maximum Likelihood	.599	.700	.799	.899	
	RMSE	.028	.029	.040	.046		.008	.007	.008	.007	Best
Var. of Residuals	H	.600	.702	.801	.896	Least Squares Standard Deviation	.599	.699	.799	.892	
	RMSE	.024	.028	.030	.027		.011	.012	.015	.017	
R/S	H	.619	.706	.784	.854	Least Squares Variance	.599	.700	.800	.895	
	RMSE	.036	.033	.035	.055		.009	.008	.011	.015	2nd best

- In this table the mean of the estimated *H*'s and the corresponding RMSE for different estimation methods are shown. They were produced by a simulation study using 200 independent realizations 8 192 long.
- The three methods (ML, LSSD, LSV) seem to be unbiased for this size.
- They also seem to surpass the other methods in terms of RMSE.

Details about the above methods (excluding ML, LSSD, LSV) in Taqqu, M., Teverovsky, V. & Willinger, W. (1995) Estimators for long-range dependence: an empirical study. *Fractals* **3**(4), 785-798.

Ⅲ. Influence of *H* to the estimators' performance

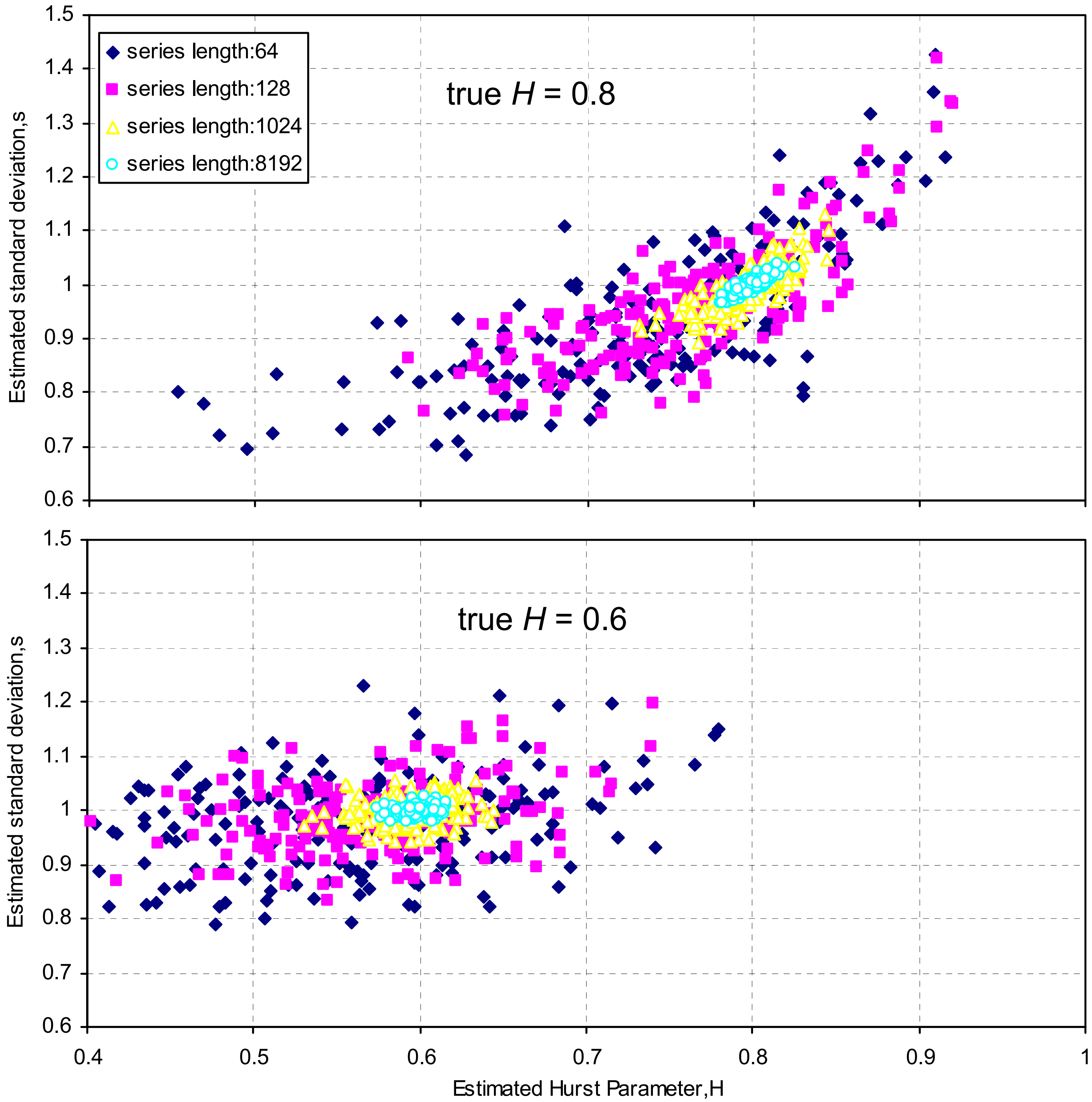
Monte Carlo RMSE for the three different methods versus the true *H* value



Performance in terms of RMSE for different *H* values. An increase of *H* results in higher errors when estimating σ.

№ Interdependence between the H and σ estimators

Dependence between the estimated σ and the estimated H



The Fisher information matrix for the parameter vector θ is defined below

$$\boldsymbol{I}_{ij}(\boldsymbol{\theta}) := -\,\mathrm{E}\left[\frac{\partial^2 \ln[p(\boldsymbol{\theta}|\boldsymbol{x}_n)]}{\partial\theta_i\,\partial\theta_j}\right]$$

$$\boldsymbol{\theta} := (\theta_1, \theta_2, \theta_3) \equiv (\mu, \sigma, H)$$

It is easily proved that

$$\boldsymbol{I}_{12}(\boldsymbol{\theta}) = \boldsymbol{I}_{13}(\boldsymbol{\theta}) = 0$$

$$\boldsymbol{I}_{23}(\boldsymbol{\theta}) = (1/\sigma)\ \mathrm{Tr}(\boldsymbol{R}^{-1}\frac{\partial \boldsymbol{R}}{\partial H}) \neq 0$$

which shows the dependence between σ and H. This dependence increases for high values of H.

10. Conclusions

- Three estimators (ML, LSSD, LSV) relying on the structure of the HK stochastic process are used to estimate its parameters.
- These estimators have the advantage to be more accurate when compared to the usual estimators of the literature.
- The finite sample properties of these estimators are explored.
- They seem to behave well for small samples but their performance declines for large values of the Hurst parameter H.
- Their main advantage is they estimate simultaneously the Hurst parameter and the standard deviation σ of the stochastic process.
- This property is essential because of the dependence of the two parameters. For example, the estimate of σ given H, is not guaranteed to vary slowly with small changes of H.
- Some theoretical results concerning the asymptotic properties of the LSV estimator with respect to H and its bracketing in [0, 1] are also given.

11. Appendix – Theoretical results (1)

Derivation of the ML estimator

The likelihood function is

$$p(\boldsymbol{\theta}|\boldsymbol{x}_n) = \frac{1}{(2\pi)^{n/2}}\,\frac{1}{\sigma^n}\,[\det(\boldsymbol{R})]^{-1/2}\exp[-\frac{1}{2\,\sigma^2}\,(\boldsymbol{e}^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}\,(\mu - \frac{\boldsymbol{x}_n^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}}{\boldsymbol{e}^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}})^2 + \frac{\boldsymbol{e}^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}\,\boldsymbol{x}_n^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{x}_n - (\boldsymbol{x}_n^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e})^2}{\boldsymbol{e}^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}})]$$

This function is maximized when $\hat{\mu} = \frac{\boldsymbol{x}_n^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}}{\boldsymbol{e}^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}}$

because $\boldsymbol{R}$ is a positive definite matrix and $\boldsymbol{e}^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e} > 0$

After substituting this value of μ the ML estimates of σ is given after equating the partial derivative of the log-likelihood function to 0.

$$\frac{\partial \ln[p(\boldsymbol{\theta}|\boldsymbol{x}_n)]}{\partial\sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\,(\boldsymbol{x}_n - \hat{\mu}\,\boldsymbol{e})^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,(\boldsymbol{x}_n - \hat{\mu}\,\boldsymbol{e})$$

So we can obtain the estimate of σ from the following equation

$$\hat{\sigma} = \sqrt{\frac{(\boldsymbol{x}_n - \hat{\mu}\,\boldsymbol{e})^{\mathrm{T}}\,\hat{\boldsymbol{R}}^{-1}\,(\boldsymbol{x}_n - \hat{\mu}\,\boldsymbol{e})}{n}}$$

and the H estimate from the minimization of the following function

$$g_1(H) := -\frac{n}{2}\,\ln[(\boldsymbol{x}_n - \frac{\boldsymbol{x}_n^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}}{\boldsymbol{e}^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}}\,\boldsymbol{e})^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,(\boldsymbol{x}_n - \frac{\boldsymbol{x}_n^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}}{\boldsymbol{e}^{\mathrm{T}}\,\boldsymbol{R}^{-1}\,\boldsymbol{e}}\,\boldsymbol{e})] - \frac{1}{2}\,\ln[\det(\boldsymbol{R})]$$

12. Appendix – Theoretical results (2)

Bracketing of *H* for the LSV estimator

Suppose that $H_2 > 1$ and $\sigma_2 > 0$ (It's easy to prove that an estimated $\hat{\sigma} > 0$ always). Now for any $H_1 \in (0, 1)$ we can always find a $\sigma_1 > 0$, such that $c_k(H_1)\, k^{2H_1}\, \sigma_1^2 - s^{2(k)} < 0$ for every k. For these values of H_1 and σ_1: $|\, c_k(H_1)\, k^{2\cdot H_1}\, \sigma_1^2 - s^{2(k)} \,| < |\, c_k(H_2)\, k^{2\cdot H_2}\, \sigma_2^2 - s^{2(k)} \,|$ for every k. This proves that $e^2(\sigma_1, H_1) < e^2(\sigma_2, H_2)$. So $e^2(\sigma, H)$ attains its minimum for $H \leq 1$.

This property is important since other estimators permit estimated values of H higher than 1, for which a HKp cannot be defined.