FACETS OF UNCERTAINTY
Kos Island, Greece | 17-19 October 2013
European Geosciences Union · International Association of Hydrological Sciences · International Union of Geodesy and Geophysics
5th EGU LEONARDO CONFERENCE • HYDROFRACTALS '13 • STATISTICAL HYDROLOGY—STAHY '13

How to parsimoniously disaggregate rainfall in time

ROMA TRE UNIVERSITÀ DEGLI STUDI

Federico Lombardo[(1)], Elena Volpi[(1)] and Demetris Koutsoyiannis[(2)]

[(1)] Dipartimento di Ingegneria, Università Roma Tre, Italy – federico.lombardo@uniroma3.it
[(2)] Department of Water Resources and Environmental Engineering, National Technical University of Athens, Greece

ITIA RESEARCH TEAM ΕΡΕΥΝΗΤΙΚΗ ΟΜΑΔΑ

ABSTRACT

Generating finer scale time series of rainfall that are fully consistent with any given coarse-scale totals is still an important and open issue in hydrology. This is commonly tackled by disaggregation models. We focus on a simple and parsimonious model based on a particular nonlinear transformation of the variables obtained by a stepwise disaggregation approach, which generates time series with Hurst-Kolmogorov dependence structure (*Lombardo et al.*, 2012). Unfortunately, nonlinear transformations of the variables do not preserve the additive property, which is one of the main attributes of the original disaggregation scheme.

To overcome this problem, an empirical adjusting procedure is suggested in order to restore consistency, but such a procedure may, in turn, introduce bias in all statistics that are to be preserved. We modify the time series generated by our model in a way to be consistent with a given higher-level time series, without affecting the stochastic structure implied by our model.

Mathematical framework

Rainfall is a natural process $R(t)$ defined in continuous time t; we observe or study it in discrete time as $R_j(\delta)$, which is the average of $R(t)$ over a fixed time scale δ in discrete time steps j (with $j = 1, 2, \ldots$)

$$R_j^{(\delta)} \coloneqq \frac{1}{\delta}\int_{(j-1)\delta}^{j\delta} R(t)dt \quad (1)$$

The aggregated and mean aggregated processes at the $f\delta$ (f is a positive integer and for convenience δ will be omitted) are

$$Z_j^{(f)} \coloneqq \sum_{l=(j-1)f+1}^{jf} R_l = f R_j^{(f)} \quad (2)$$

Hurst-Kolmogorov (HK) process

Stochastic stationary process which, for any integers i and j and any time scales f and l, has the property

$$\left(\tilde{R}_j^{(f)} - \tilde{\mu}\right) \stackrel{d}{=} \left(\frac{f}{l}\right)^{H-1} \left(\tilde{R}_i^{(l)} - \tilde{\mu}\right) \quad (3)$$

where $0 < H < 1$ is the Hurst coefficient. Typically, $\tilde{R}_j$ is Gaussian and $\tilde{\mu} = \langle \tilde{R} \rangle$.

For the relevant $\tilde{Z}_j^{(f)}$ process the following scaling law holds

$$\mathrm{var}\left[\tilde{Z}_j^{(f)}\right] = f^2 \mathrm{var}\left[\tilde{R}_j^{(f)}\right] = f^{2H}\tilde{\sigma}^2 \quad (4)$$

where $\tilde{\sigma}^2 = \mathrm{var}[\tilde{R}]$ and the correlation coefficient is given by

$$\tilde{\rho}^{(f)}(t) = \tilde{\rho}(t) = \frac{|t+1|^{2H}}{2} + \frac{|t-1|^{2H}}{2} - |t|^{2H} \quad (5)$$

HK downscaling model (*Lombardo et al.*, 2012)

Let $Z_1^{(f)}$ be the cumulative rainfall depth at the time origin ($j = 1$) aggregated on the largest time scale f which is to be downscaled to a certain scale of interest: it is assumed to be a log-normally distributed random variable with mean μ_0 and variance σ_0^2.

Let us now introduce an auxiliary Gaussian random variable of the aggregated HK process on the time scale f, $\tilde{Z}_{1,0} = \tilde{Z}_1^{(f)} = \ln Z_1^{(f)}$ with $\tilde{\mu}_0 = \langle \tilde{Z}_{1,0} \rangle$ and $\tilde{\sigma}_0^2 = \mathrm{var}[\tilde{Z}_{1,0}]$. $\tilde{Z}_{1,0}$ is to be disaggregated by a dyadic additive cascade, following the linear generation scheme

$$\begin{cases} \tilde{Z}_{2j-1,k} = \boldsymbol{\theta}^T \mathbf{Y} + V \\ \tilde{Z}_{2j,k} = \tilde{Z}_{j,k-1} - \tilde{Z}_{2j-1,k} \end{cases} \quad (6)$$

where $\mathbf{Y} = \left[\tilde{Z}_{2j-5,k}, \tilde{Z}_{2j-4,k}, \tilde{Z}_{2j-3,k}, \tilde{Z}_{2j-2,k}, \tilde{Z}_{j,k-1}, \tilde{Z}_{j+1,k-1}, \tilde{Z}_{j+2,k-1}\right]^T$, $\boldsymbol{\theta}$ is the vector of parameters and V the variance of the innovation term (which are estimated in terms of the correlation coefficients and of the variance of the HK process at the level k). In order to increase the model accuracy in reproducing the summary statistics of the underlying stochastic process, we expand the number of lower- and higher-level variables that are considered in the generation procedure, eq. (6), with respect to the original formulation proposed by Lombardo et al. (2012).

We apply the following specific exponentiation to the HK process (auxiliary process) to make it log-normal (actual process) but preserving its scaling properties (4)

$$Z_{j,k} = \exp\left[\alpha(k)\tilde{Z}_{j,k} + \beta(k)\right] \quad (7)$$

with $\alpha(k) = \frac{2^{Hk}}{\tilde{\sigma}_0}\sqrt{\ln[2^{2k(1-H)}(\exp(\tilde{\sigma}_0^2) - 1) + 1]}$, and

$\beta(k) = -k\ln 2 - \tilde{\mu}_0\left[\frac{\alpha(k)}{2^k} - 1\right] - \frac{\tilde{\sigma}_0^2}{2}\left[\frac{\alpha^2(k)}{2^{2Hk}} - 1\right]$.

The log-normality hypothesis and our specific exponential transformation enable the analytical formulation of the main statistics of the actual rainfall field, $\mu_k = \langle Z_{j,k} \rangle = \frac{\mu_0}{2^k}$, $\sigma_k^2 = \mathrm{var}[Z_{j,k}] = \frac{\sigma_0^2}{2^{2Hk}}$ and $\rho_k(t) = \frac{\exp\left(\tilde{\sigma}_k^2\tilde{\rho}(t)\right)-1}{\exp(\tilde{\sigma}_k^2)-1}$. The lower-level rainfall time series generated are only statistically consistent with the given field $\boldsymbol{Z}$ at the coarser scale.

References

- Lombardo, F., Volpi, E. and D. Koutsoyiannis, 2012. Rainfall downscaling in time: Theoretical and empirical comparison between multifractal and Hurst-Kolmogorov discrete random cascades. Hydrological Science Journal, 57, 6, 1052-1066, doi: 10.1080/02626667.2012.695872
- Koutsoyiannis, D., and A. Manetas, Simple disaggregation by accurate adjusting procedures, Water Resources Research, 32(7), 2105-2117, 1996.

Adjusting procedure

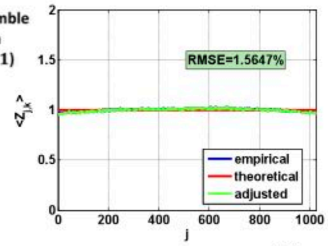
The empirical adjusting procedure is introduced in the downscaling model in order to restore consistency, i.e. to preserve exactly the rainfall mass at the higher level of the actual process $Z_1^{(f)}$. We apply herein a proportional adjusting procedure (*Koutsoyiannis and Manetas*, 1996), which modifies the generated values at the k level, $Z_{j,k}$ with $\{j = 1, 2, \ldots n\}$ to get the adjusted values $Z'_{j,k}$ according to

$$Z'_{j,k} = Z_{j,k}\left(Z_1^{(f)} / \textstyle\sum_{j=1}^{n} Z_{j,k}\right) \quad (7)$$

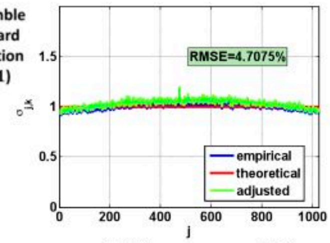
Numerical simulations

We generate 30000 time series with sample size $n = 2^{10}$ =1024, unit mean, standard deviation $\sigma \in [0.5, 2]$ and $H = 0.85$. A good adjusting procedure, in addition to reinstating the additive property, should preserve certain statistics or even the complete distribution of lower-level variables; we analyze the behaviour of the ensemble mean, standard deviation and autocorrelation function.

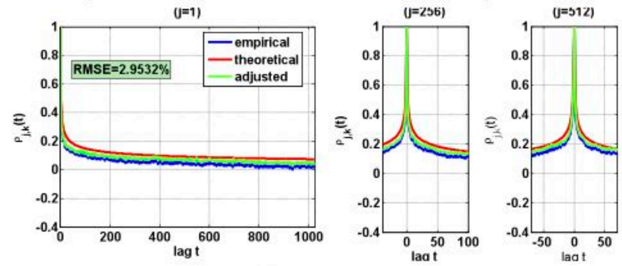
Ensemble mean ($\sigma = 1$)



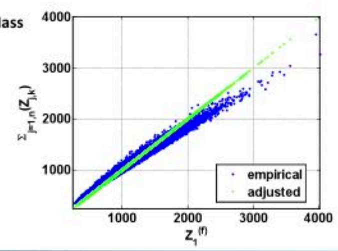
Ensemble standard deviation ($\sigma = 1$)

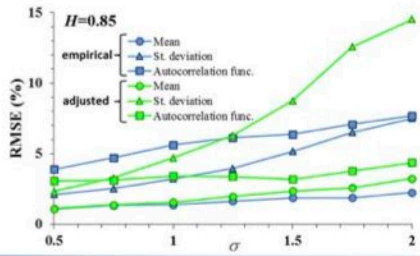


Ensemble autocorrelation function ($\sigma = 1$)



Mass





CONCLUSIONS

We show some preliminary results of a disaggregation method based on the Hurst-Kolmogorov process. We use a simple and parsimonious model based on exponentiation of the variables obtained by a stepwise disaggregation approach, in a way to generate lognormal time series with Hurst-Kolmogorov dependence structure. Then, we use an accurate proportional adjusting procedure to allocate the error in the additive property, i.e., the departure of the sum of lower-level variables within a period from the corresponding higher-level variable. It is accurate in the sense that it preserves explicitly certain statistics of lower-level variables. Note that some statistics are not explicitly preserved by our adjusting procedure, especially when increasing the variance of the synthetic series.

In future work, we should improve the approximations of these statistics by repetition, i.e., by combining conditional sampling with adjusting: Generate the lower-level variables until the error in the additive condition is small and apply an adjusting procedure to allocate this error among the different sub-periods.