European Geosciences Union

General Assembly 2015, Vienna, Austria, 12 – 17 April 2015

Session HS7.7/NP3.8: Hydroclimatic and hydrometeorologic stochastics: Extremes, scales, probabilities

Parsimonious entropy-based stochastic modelling for changing hydroclimatic processes

Demetris Koutsoyiannis

Department of Water Resources and Environmental Engineering, School of Civil Engineering

National Technical University of Athens, Greece

(dk@itia.ntua.gr, http://www.itia.ntua.gr/dk/)

Presentation available online: www.itia.ntua.gr/1533/

Premises

- Most things are uncertain.
- Stochastics is the language of uncertainty.
- Entropy is the quantified measure of uncertainty.
- The *principle of maximum entropy*, which reflects entropy maximization in nature, can help to construct parsimonious probabilistic representations of natural phenomena.
- When time matters, the concept of *maximum entropy production* can help to construct parsimonious stochastic representations of natural processes.
- Hydrological and, more generally, geophysical processes, exhibit some peculiarities, such as:
  (a) their modelling relies very much on observational data (geophysical systems are too complex to be studied using deduction, and theories are often inadequate);
  (b) the distinction “signal vs. noise” is meaningless;
  (c) the samples are small;
  (d) the processes are often characterized by long term persistence, which makes classical statistics inappropriate.
- Long term persistence is long term change; thus modelling change in a stochastic framework is about the same as modelling long-term persistence.

Things to avoid in stochastic modelling

- Avoid stylized families of models like AR, ARMA, ARIMA, ARFIMA, etc., which (excepting their most elementary versions) are too artificial, not parsimonious and unnecessary.
- Avoid formulating a stochastic model in discrete time; rather formulate it in continuous time and infer its discrete-time properties analytically.
- In model identification and fitting, avoid the common practice of using the empirical autocorrelogram, as it is associated with high bias and uncertainty.
- Avoid applying mathematical tools from the fractals literature whose statistical / stochastic properties have not been studied, are too complicated to study or are known to involve high bias and uncertainty.
- In particular, avoid treating random variables as if they were deterministic and cursorily using uncontrollable quantities (e.g. high order moments) whose estimates are characterized by extraordinarily high bias and uncertainty.
- More generally, avoid interpreting stochastics as recipes, algorithms and series of numerical calculations that are easily performed by popular computer programs.

Proposed framework

Example of a stochastic process $\underline{x}_i$ with maximal entropy production at times $t \to 0$ and ∞:

$$\gamma(\Delta) = \lambda(1 + (\Delta/\alpha)^{2\kappa})^{\frac{H-1}{\kappa}} \quad \text{[Hybrid Hurst-Kolmogorov process, HHK]}$$

γ: variance; Δ: Time scale; $\gamma(\Delta)$: climacogram; λ: state-scale parameter; α: time-scale parameter; H: Hurst (scaling) parameter ($0 < H < 1$); κ: fractal (scaling) parameter ($0 < \kappa < 1$)

Model fitting:

Minimize the error (e.g. MSE) between the empirical climacogram $\hat{\gamma}(\Delta)$ and the theoretical expectation $\mathrm{E}\left[\underline{\hat{\gamma}}(\Delta)\right]$ (not $\gamma(\Delta)$ itself).

Generalized simulation method (for the generation of the process $\underline{x}_i$):

$$\underline{x}_i = \sum_{l=-q}^{q} a_{|l|} \underline{v}_{i+l} \quad \text{[Symmetric Moving Average technique, SMA]}$$

where q is a large integer, $\underline{v}_i$ is white noise and a_l are coefficients calculated from

$$s_{\mathrm{d}}^{a}(\omega) = \sqrt{2 s_{\mathrm{d}}(\omega)}$$

whereas $s_{\mathrm{d}}^{a}(\omega)$ is the Fourier transform of the a_l series, $s_{\mathrm{d}}(\omega)$ is the power spectrum of the discrete time process, determined from the climacogram $\gamma(\Delta)$, and ω is frequency.

Further information

Details about the methodology

Koutsoyiannis, D., Generic and parsimonious stochastic modelling for hydrology and beyond, *Hydrological Sciences Journal*, doi:10.1080/02626667.2015.1016950, 2015.

Software implementation

Castalia: A computer system for stochastic simulation and forecasting of hydrologic processes, http://www.itia.ntua.gr/en/softinfo/2/

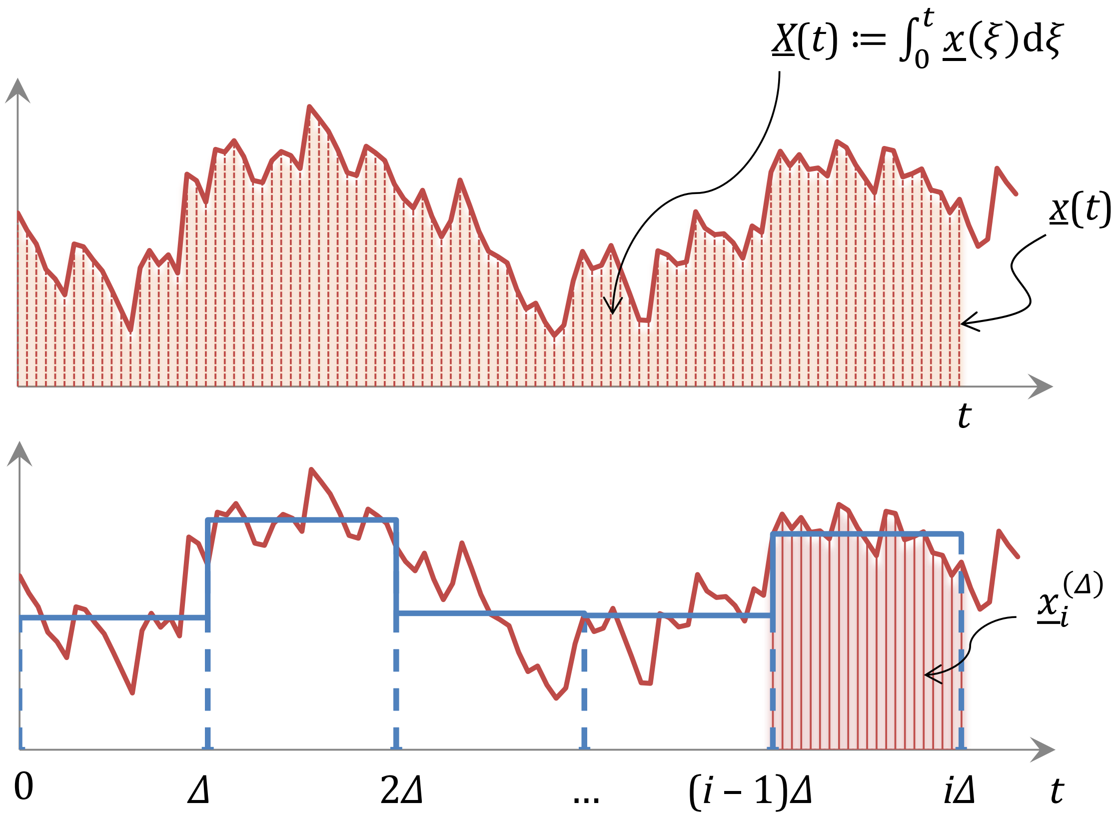
Efstratiadis, A., Y. Dialynas, S. Kozanis, and D. Koutsoyiannis, A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence, *Environmental Modelling and Software*, 62, 139–152, doi:10.1016/j.envsoft.2014.08.017, 2014.

Funding information

This research has been financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) – Research Funding Program: ARISTEIA: Reinforcement of the interdisciplinary and/ or inter-institutional research and innovation. (research project "Combined REnewable Systems for Sustainable Energy DevelOpment"—CRESSENDO; grant number 5145).

Additional explanations

Definitions and notation



Type	Continuous time	Discrete time, time scale Δ
Stochastic processes	$\underline{x}(t)$: instantaneous, stationary $\underline{X}(t) \coloneqq \int_0^t \underline{x}(\xi)\mathrm{d}\xi$: cumulative, nonstationary	$\underline{X}_i^{(\Delta)} \coloneqq \underline{X}(i\Delta) - \underline{X}((i-1)\Delta)$: aggregated, stationary intervals of $\underline{X}(t)$ $\underline{x}_i^{(\Delta)} \coloneqq \underline{X}_i^{(\Delta)} / \Delta$: averaged
Characteristic variances	$\gamma_0 \coloneqq \mathrm{Var}[\underline{x}(t)]$ $\Gamma(t) \coloneqq \mathrm{Var}[\underline{X}(t)]$ $\gamma(t) \coloneqq \mathrm{Var}[\underline{X}(t)/t] = \Gamma(t)/t^2$ Note: $\Gamma(0) = 0;\ \gamma(0) = \gamma_0$	$\mathrm{Var}[\underline{X}_i^{(\Delta)}] = \Gamma(\Delta)$ $\mathrm{Var}[\underline{x}_i^{(\Delta)}] = \gamma(\Delta)$
Autocovariance function	$c(\tau) \coloneqq \mathrm{Cov}[\underline{x}(t), \underline{x}(t+\tau)]$ Note: $c(0) \equiv \gamma_0 = \gamma(0)$	$c_j^{(\Delta)} \coloneqq \mathrm{Cov}[\underline{x}_i^{(\Delta)}, \underline{x}_{i+j}^{(\Delta)}]$ Note: $c_0^{(\Delta)} \equiv \gamma(\Delta)$
Power spectrum (spectral density)	$s(w) \coloneqq 2\int_{-\infty}^{\infty} c(\tau)\cos(2\pi w\tau)\,\mathrm{d}\tau$	$s_\mathrm{d}^{(\Delta)}(\omega) \coloneqq 2\sum_{j=-\infty}^{\infty} c_j^{(\Delta)} \cos(2\pi\omega j)$ $s^{(\Delta)}(w) = \Delta s_\mathrm{d}^{(\Delta)}(w\Delta)$ Note: $w = \omega/\Delta$

Properties of the Hybrid Hurst-Kolmogorov process (HHK)

Climacogram:

$$\gamma(\Delta) = \lambda(1 + (\Delta/\alpha)^{2\kappa})^{\frac{H-1}{\kappa}}$$

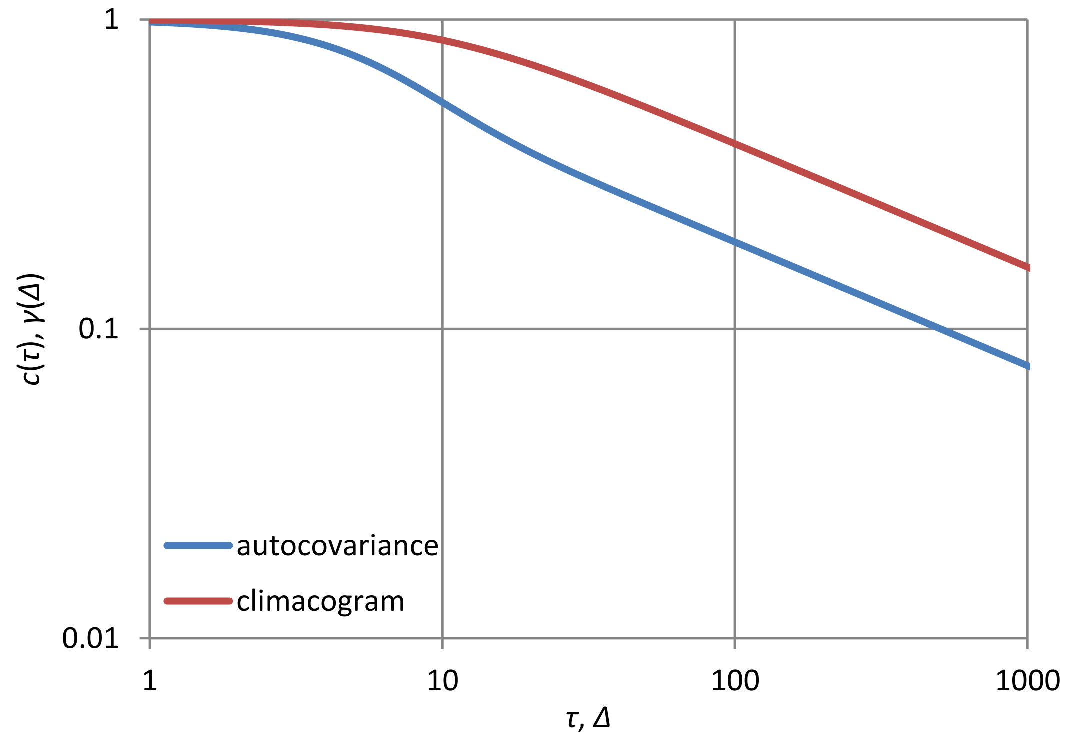
Autocovariance in continuous time (for lag τ):

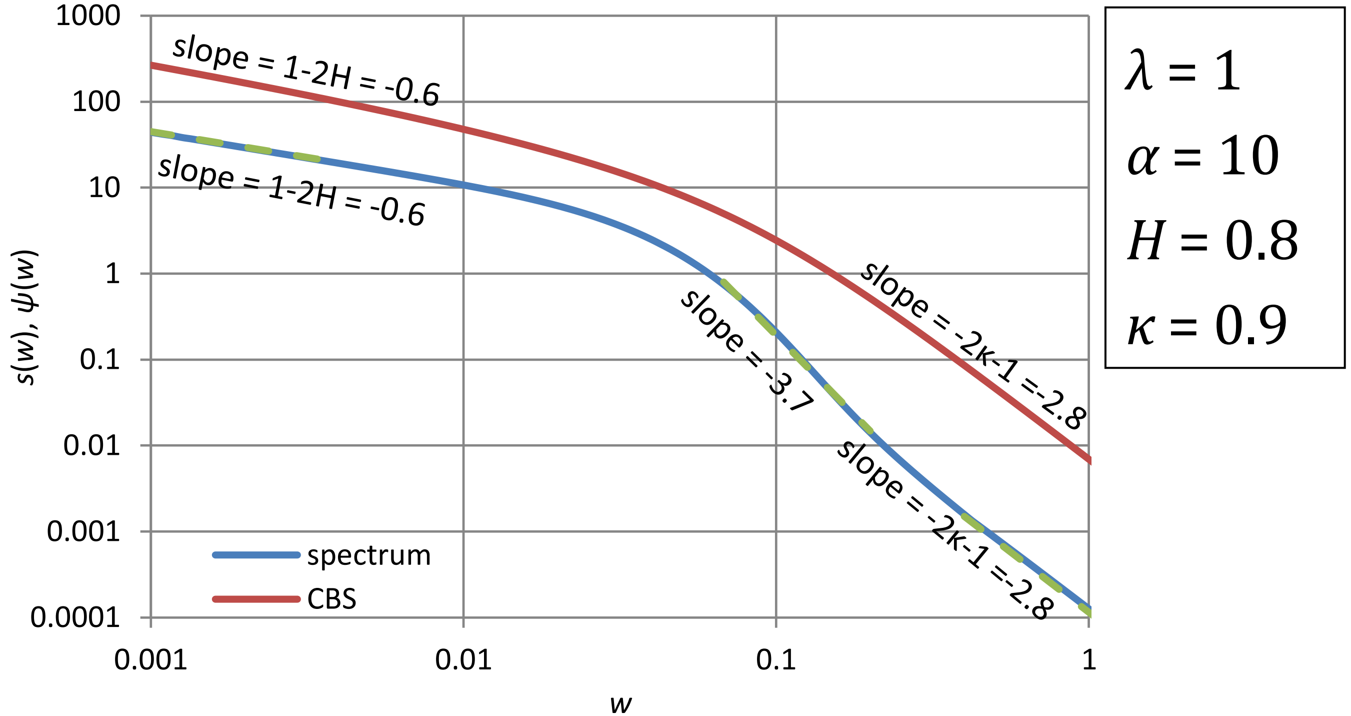
$$c(\tau) = \gamma(\tau)\,\frac{1+(3H+2\kappa H-2\kappa-1)(\tau/a)^{2\kappa}+H(2H-1)(\tau/a)^{4\kappa}}{1+2(\tau/a)^{2\kappa}+(\tau/a)^{4\kappa}}$$

The expressions of other properties are complex, but can be easily be evaluated numerically. Furthermore, the asymptotic properties are easily derived analytically (mostly in terms of log-log derivatives, e.g. $\gamma^{\#}(x)$) and given below (see more details in Koutsoyiannis, 2013b, 2015):

Property	Global behaviour	Local behaviour
Climacogram (in terms of variance $\gamma(\Delta)$ or standard deviation $\sigma(\Delta)$)	$\gamma^{\#}(\infty) = 2\sigma^{\#}(\infty) = 2H - 2$	$\gamma^{\#}(0) = \sigma^{\#}(0) = 0$
Autocovariance $c(\tau)$	$c^{\#}(\infty) = 2H - 2$	$c^{\#}(0) = 0$
Power spectrum $s(w)$	$s^{\#}(0) = 1-2H$	$s^{\#}(\infty) = -2\kappa - 1$
Climacogram-based spectrum, $\psi(w) \coloneqq \frac{2\,\gamma(1/w)}{w}\left(1 - \frac{\gamma(1/w)}{\gamma_0}\right)$	$\psi^{\#}(0) = 1-2H$	$\psi^{\#}(\infty) = -2\kappa - 1$
Climacogram-based structure function, $g(\Delta) \coloneqq \gamma_0 - \gamma(\Delta)$	$g^{\#}(\infty) = 0$	$g^{\#}(0) = 2\kappa$
Entropy production in logarithmic time (unconditional $\varphi(\Delta)$, conditional $\varphi_C(\Delta)$)	$\varphi(\infty) = \varphi_C(\infty) = H$	$\varphi(0) = 1$ $\varphi_C(0) = 1 + \kappa$

Properties of the HHK process—an example

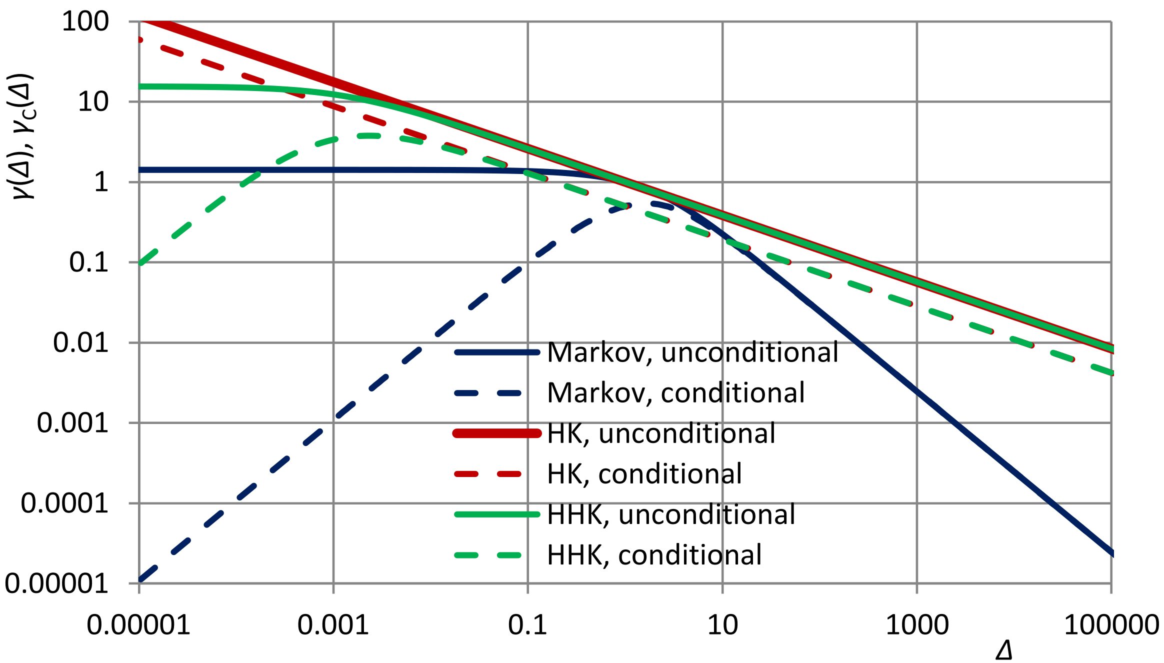


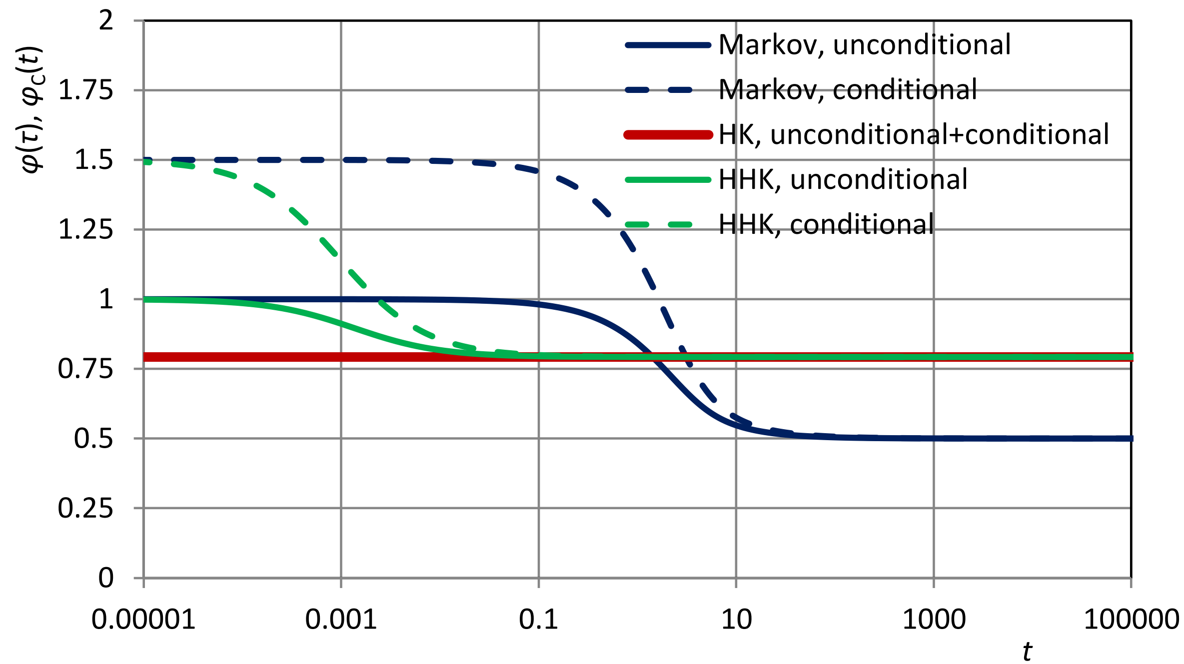


Notes

1. For large time scales HHK exhibits Hurst behaviour. The scaling behaviour at small time scales (small frequencies) is quite different.
2. The intermediate steep slope that appears in the power spectrum is artificial and does not indicate a scaling behaviour.

Comparison of the HHK process with Markov and HK processes





In all processes	HHK parameters
$\gamma(1) = 1$ $c_1^{(1)} = 0.5$	$a = 0.0013539$ $\lambda = 15.5093$ $\kappa = 0.5$ $H = 0.7925$

For $\kappa = 0.5 \neq H$ and for small scales the HHK process behaves like Markov and for large scales behaves as Hurst-Kolmogorov. (In the special case $H = \kappa = 0.5$, HHK is practically indistinguishable from a Markov process—even though not precisely identical). As $\alpha \to 0$, the process tends to a pure HK process with the same Hurst coefficient H.

Statistical estimation—Model fitting

We assume that the observation period T is an integer multiple of time scale Δ of the averaged process $\underline{x}_i^{(\Delta)}$, i.e., $n = T/\Delta$ is an integer. The (unbiased) estimator of the common mean μ of the instantaneous process $\underline{x}(t)$ as well as of the discrete process $\underline{x}_i^{(\Delta)}$ is

$$\underline{\bar{x}}^{(\Delta)} := \frac{1}{n}\sum_{i=1}^{n} \underline{x}_i^{(\Delta)} = \frac{\underline{X}(T)}{n\Delta} = \frac{\underline{X}(T)}{T} = \underline{x}_1^{(T)}$$

The standard estimator $\underline{\hat{\gamma}}(\Delta)$ of the variance $\gamma(\Delta)$ of the averaged process $\underline{x}_i^{(\Delta)}$ is

$$\underline{\hat{\gamma}}(\Delta) := \frac{1}{n-1}\sum_{i=1}^{n}\left(\underline{x}_i^{(\Delta)} - \underline{\bar{x}}^{(\Delta)}\right)^2 = \frac{1}{T/\Delta-1}\sum_{i=1}^{T/\Delta}\left(\underline{x}_i^{(\Delta)} - \underline{\bar{x}}^{(\Delta)}\right)^2$$

This is biased (except for white noise) where the bias correction coefficient η is estimated as follows (Koutsoyiannis, 2011, 2013b, 2015):

$$\mathrm{E}\left[\underline{\hat{\gamma}}(\Delta)\right] = \eta(\Delta, T)\gamma(\Delta)$$

$$\eta(\Delta, T) = \frac{1-\gamma(T)/\gamma(\Delta)}{1-\Delta/T} = \frac{1-(\Delta/T)^2\Gamma(T)/\Gamma(\Delta)}{1-\Delta/T}$$

Important note: Direct estimation of the variance $\gamma(\Delta)$ (even more so, γ_0) is not possible merely from the data: we should assume a stochastic model which evidently influences the estimation of $\gamma(\Delta)$.

Steps for the calculation of the coefficients of the SMA technique

1. A model is assumed in terms of its climacogram $\gamma(\Delta)$ or $\Gamma(\Delta) = \gamma(\Delta)\,\Delta^2$.
2. The model is fitted by minimizing the error (e.g. MSE) between the empirical climacogram $\hat{\gamma}(\Delta)$ and the theoretical expectation $\mathrm{E}\left[\underline{\hat{\gamma}}(\Delta)\right]$, as the latter is determined in the previous page.
3. The autocovariance function at scale Δ is determined from:

$$c_j^{(\Delta)} = \frac{1}{2}\,\frac{\delta_\Delta^2 \Gamma(j\Delta)}{\Delta^2} = \frac{1}{\Delta^2}\left(\frac{\Gamma\big((j+1)\Delta\big)+\Gamma\big((j-1)\Delta\big)}{2} - \Gamma(j\Delta)\right)$$

4. The power spectrum of the discrete-time process is calculated from the Fourier transform:

$$s_\mathrm{d}^{(\Delta)}(\omega) = 2c_0^{(\Delta)} + 4\sum_{j=1}^{\infty} c_j^{(\Delta)} \cos(2\pi\omega j)$$

5. The Fourier transform of the a_l series is calculated from:

$$s_\mathrm{d}^{a}(\omega) = \sqrt{2 s_\mathrm{d}^{(\Delta)}(\omega)}$$

6. The coefficients a_l are calculated from the inverse Fourier transform:

$$a_l = \int_0^{1/2} s_\mathrm{d}^{a}(\omega) \cos(2\pi\omega l)\, \mathrm{d}\omega$$

Dealing with truncation errors

The use of a finite number q of coefficients a_l in in the SMA technique (Koutsoyiannis, 2010) introduces a truncation error. To deal with this, the following methods have been proposed (Koutsoyiannis, 2015).

Method 1 (best): Replace $s_{\mathrm{d}}^{a}(\omega)$ with $s_{\mathrm{d}}^{a}(\omega)(1 - \mathrm{sinc}(2\pi\omega q))$ in the previous equation and calculate coefficients a_l' as

$$a_l' = \int_0^{1/2} s_{\mathrm{d}}^{a}(\omega)(1 - \mathrm{sinc}(2\pi\omega q)) \cos(2\pi\omega l)\, \mathrm{d}\omega$$

Then calculate the constant

$$a'' = \sqrt{\frac{\gamma(\Delta) - \Sigma\alpha'^2}{2q+1} + \left(\frac{\Sigma\alpha'}{2q+1}\right)^2} - \frac{\Sigma a'}{2q+1}$$

where $\Sigma\alpha' \coloneqq a_0' + 2(a_1' + \cdots + a_q')$ and $\Sigma\alpha'^2 \coloneqq a_0'^2 + 2(a_1'^2 + \cdots + a_q'^2)$. The final coefficients a_l to be applied for SMA are

$$a_l = a_l' + a''$$

Method 2 (easier): As Method 1 but without replacing $s_{\mathrm{d}}^{a}(\omega)$ with $s_{\mathrm{d}}^{a}(\omega)(1 - \mathrm{sinc}(2\pi\omega q))$.

Method 3 (good when, due to numerical imperfections, it happens that $\gamma(\Delta) - \Sigma\alpha'^2 < 0$): As Method 2 but with proportional adjustment, i.e.,

$$a_l = a_l' \sqrt{\check{\gamma}_0 / \Sigma\alpha'^2}$$

Appendix: Definition and importance of entropy

Historically, entropy was introduced in thermodynamics but later it was given a rigorous definition within probability theory (owing to Boltzmann, Gibbs and Shannon). Thermodynamic and probabilistic entropy are essentially the same thing (Koutsoyiannis, 2013b, 2014a; but others have different opinion).

Entropy is a dimensionless measure of uncertainty defined as follows:

For a *discrete random variable* $\underline{z}$ with probability mass function $P_j := P\{\underline{z} = z_j\}$

$$\Phi[\underline{z}] := \mathrm{E}[-\ln P(\underline{z})] = -\sum_{j=1}^{w} P_j \ln P_j$$

For a *continuous random variable* $\underline{z}$ with probability density function $f(z)$:

$$\Phi[\underline{z}] := \mathrm{E}\left[-\ln \frac{f(\underline{z})}{h(\underline{z})}\right] = -\int_{-\infty}^{\infty} \ln \frac{f(z)}{h(z)} f(z)\mathrm{d}z$$

where $h(\mathrm{z})$ is the density of a background measure (usually $h(\mathrm{z}) = 1[z^{-1}]$).

Entropy acquires its importance from the *principle of maximum entropy* (Jaynes, 1957), which postulates that the entropy of a random variable should be at maximum, under some conditions, formulated as constraints, which incorporate the information that is given about this variable.

Its physical counterpart, the tendency of entropy to become maximal (2nd Law of thermodynamics) is the driving force of natural change.

Appendix (contd.): Entropy production in stochastic processes

In a stochastic process the change of uncertainty in time can be quantified by the *entropy production*, i.e. the time derivative (Koutsoyiannis, 2011):

$$\Phi'[\underline{X}(t)] := \mathrm{d}\Phi[\underline{X}(t)]/\mathrm{d}t$$

A more convenient (and dimensionless) measure is the entropy production (i.e. the derivative) in logarithmic time (EPLT):

$$\varphi(t) \equiv \varphi[\underline{X}(t)] := \Phi'[\underline{X}(t)]\, t \equiv \mathrm{d}\Phi[\underline{X}(t)] \,/\, \mathrm{d}(\ln t)$$

For a Gaussian process, the entropy depends on its variance $\Gamma(t)$ only and is given as (Papoulis, 1991):

$$\Phi[\underline{X}(t)] = (1/2)\ln(2\pi\mathrm{e}\,\Gamma(t))$$

The EPLT of a Gaussian process is thus easily shown to be:

$$\varphi(t) = \Gamma'(t)\, t \,/\, 2\Gamma(t) = 1 + ½\, \gamma^{\#}(t)$$

When the past and the present are observed, instead of the unconditional variance $\Gamma(t)$ we should use a variance $\Gamma_{\mathrm{C}}(t)$ conditional on the known past and present. This turns out to be:

$$\Gamma_{\mathrm{C}}(t) \approx 2\Gamma(t) - \Gamma(2t)/2$$

Extremization of EPLT for asymptotic times ($t \to 0$ and ∞) with relevant constraints results in Markov, HK or HHK process (Koutsoyiannis, 2011, 2014a, b, 2015; for other uses of the principle of maximum entropy for parsimonious stochastic modelling see Koutsoyiannis et al. 2008).

References

- Efstratiadis, A., Y. Dialynas, S. Kozanis, and D. Koutsoyiannis, A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence, *Environmental Modelling and Software*, 62, 139–152, doi:10.1016/j.envsoft.2014.08.017, 2014.
- Jaynes, E.T. Information theory and statistical mechanics, *Physical Review*, 106 (4), 620-630, 1957.
- Koutsoyiannis, D., A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research*, 36 (6), 1519–1533, 2000.
- Koutsoyiannis, D., Hurst-Kolmogorov dynamics as a result of extremal entropy production, *Physica A*, 390 (8), 1424–1432, 2011.
- Koutsoyiannis, D., Physics of uncertainty, the Gibbs paradox and indistinguishable particles, *Studies in History and Philosophy of Modern Physics*, 44, 480–489, 2013a.
- Koutsoyiannis, D., Encolpion of stochastics: Fundamentals of stochastic processes, Department of Water Resources and Environmental Engineering – National Technical University of Athens, Athens, 2013b (http://www.itia.ntua.gr/en/docinfo/1317/).
- Koutsoyiannis, D., Entropy: from thermodynamics to hydrology, *Entropy*, 16 (3), 1287–1314, 2014a.
- Koutsoyiannis, D., Random musings on stochastics (Lorenz Lecture), *AGU 2014 Fall Meeting*, San Francisco, USA, American Geophysical Union, 2014b (http://www.itia.ntua.gr/en/docinfo/1500/).
- Koutsoyiannis, D., Generic and parsimonious stochastic modelling for hydrology and beyond, *Hydrological Sciences Journal*, doi: 10.1080/02626667.2015.1016950, 2015.
- Koutsoyiannis, D., H. Yao, and A. Georgakakos, Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods, *Hydrological Sciences Journal*, 53 (1), 142–164, 2008.
- Papoulis, A., *Probability, Random Variables and Stochastic Processes*, 3rd edn., McGraw-Hill, New York, 1991.