20 Years of Nonlinear Dynamics in Geosciences
Rhodes, Greece 11-16 June 2006

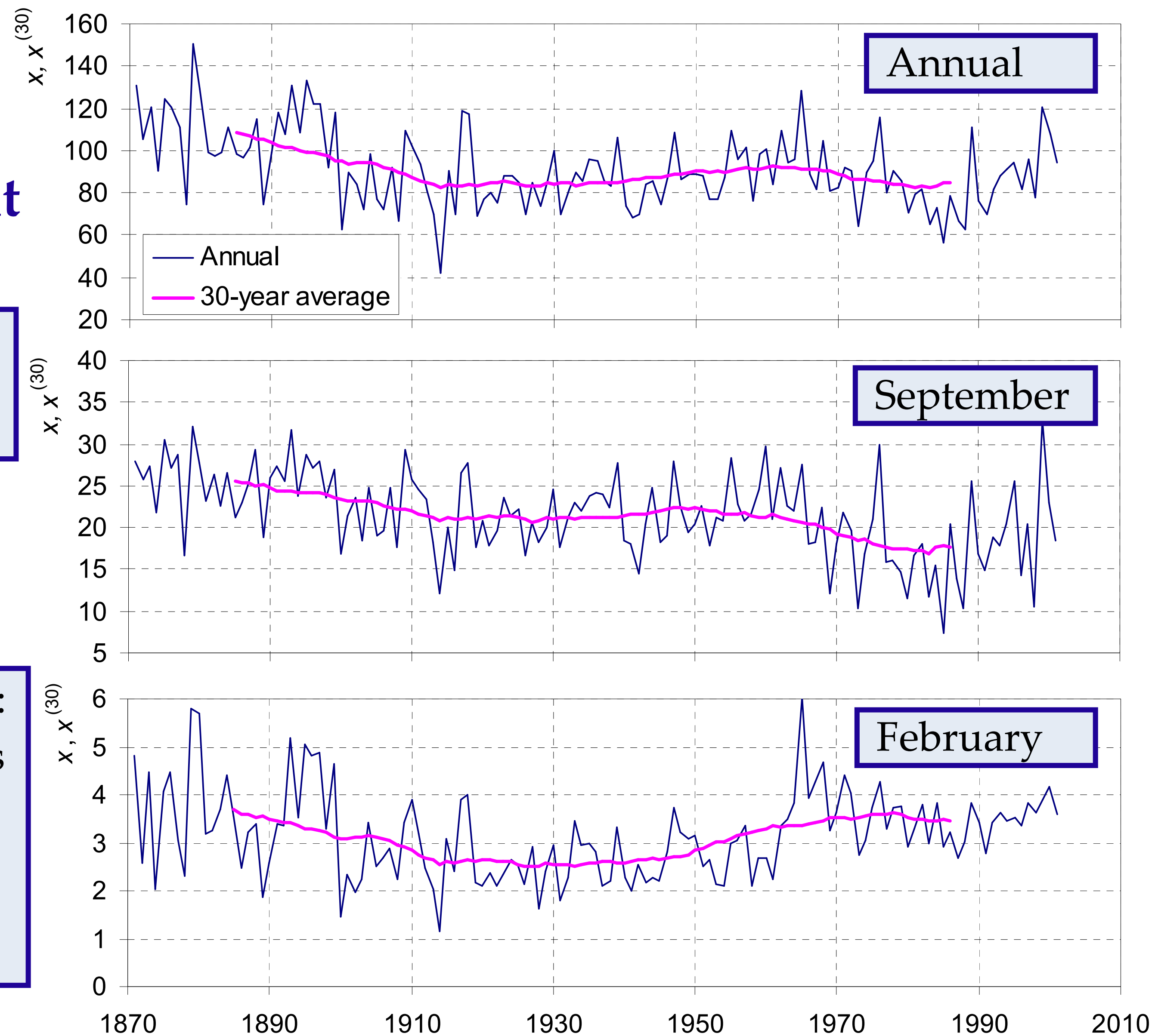
Lessons from the long flow records of the Nile: Determinism vs. indeterminism and maximum entropy

Demetris Koutsoyiannis
Department of Water Resources, School of Civil Engineering, National Technical University of Athens, Greece

Aris Georgakakos
Georgia Water Resources Institute, School of Civil and Environmental Engineering, Georgia Institute of Technology, USA

Data set 1: Contemporary record of monthly flows at Aswan

Data period 1870-2001 (131 years)
Units: km^3



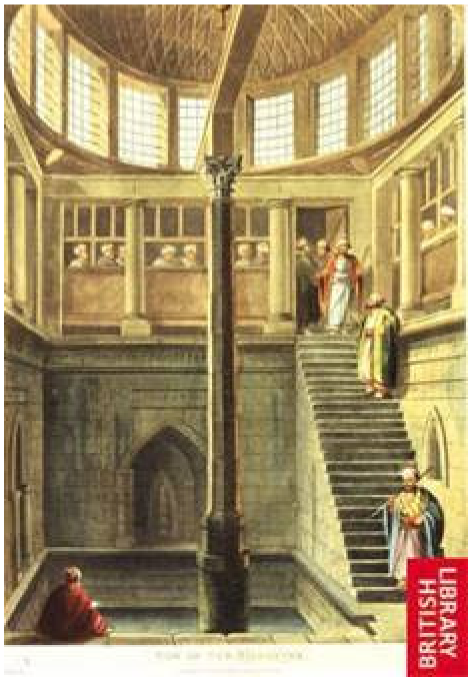
General observations:

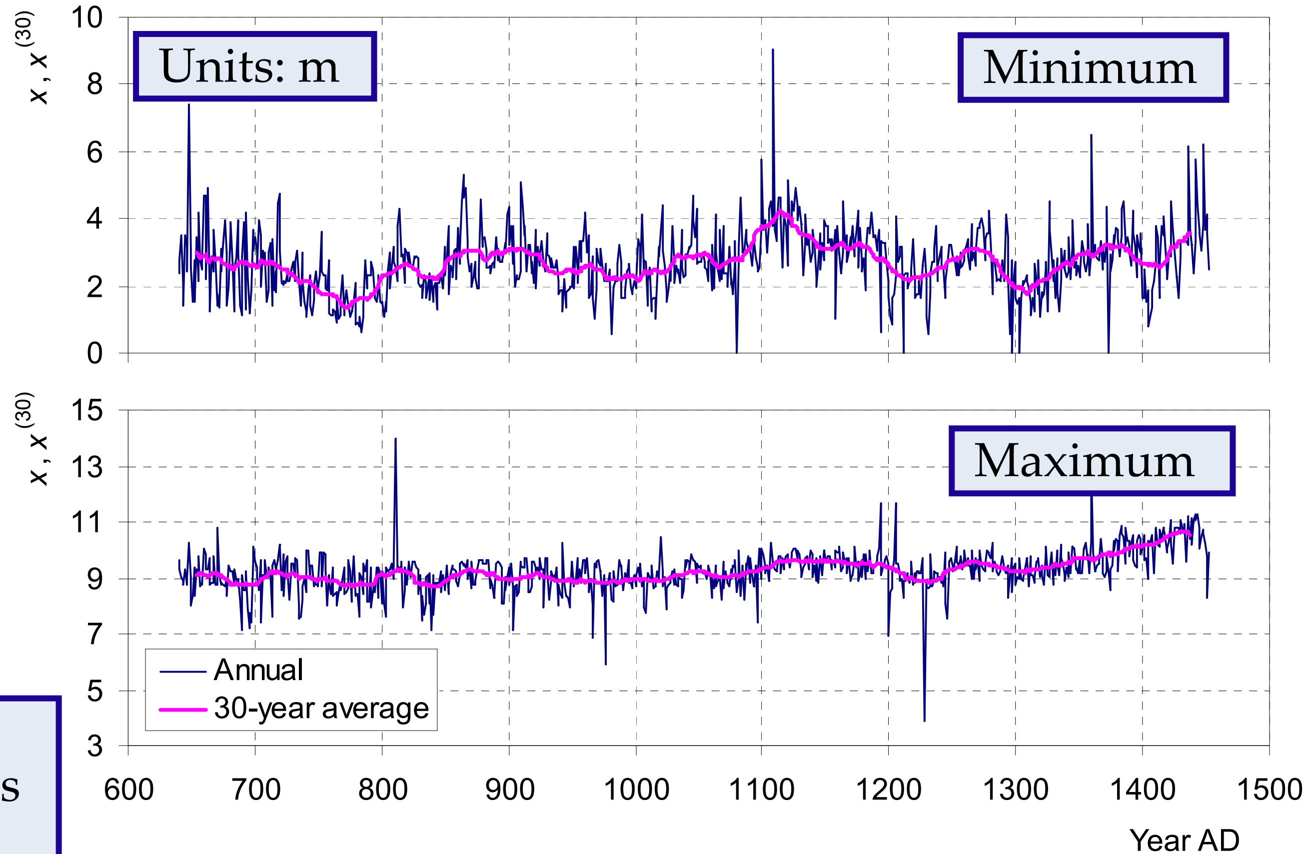
Aperiodic fluctuations on large time scales

Behaviours in different months may be like or unlike

Data sets 2 and 3: records of annual maximum and minimum water levels at the Roda Nilometer

Data period 640-1452 (813 years): the longest instrumental data set available worldwide – high importance in understanding and modelling hydroclimatic behaviours





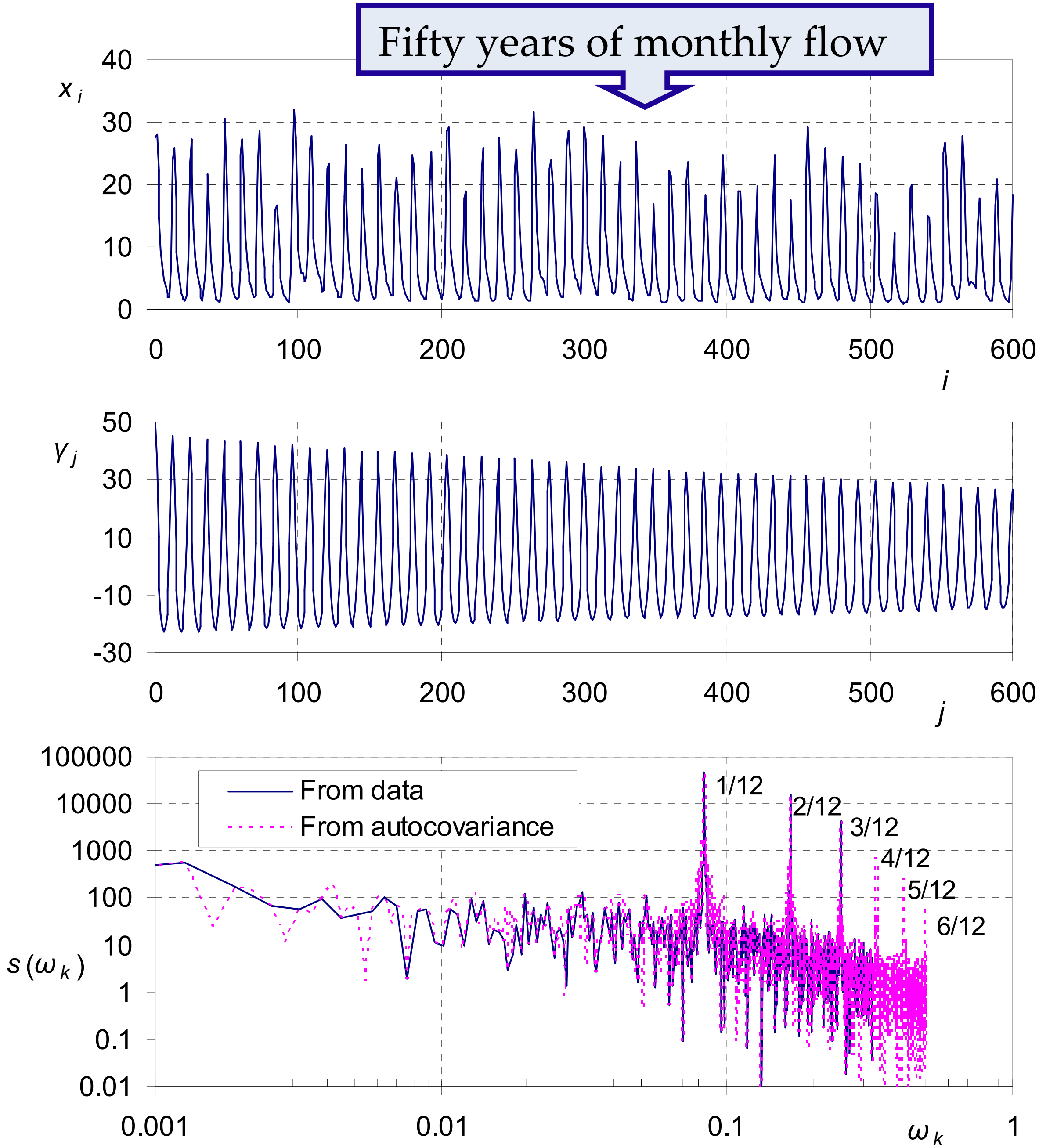
General observation: Aperiodic fluctuations on large time scales

Part 1
Seeking determinism

Analysis of the monthly flows

600 terms of autocovariance γ_j vs. lag j

Power spectrum

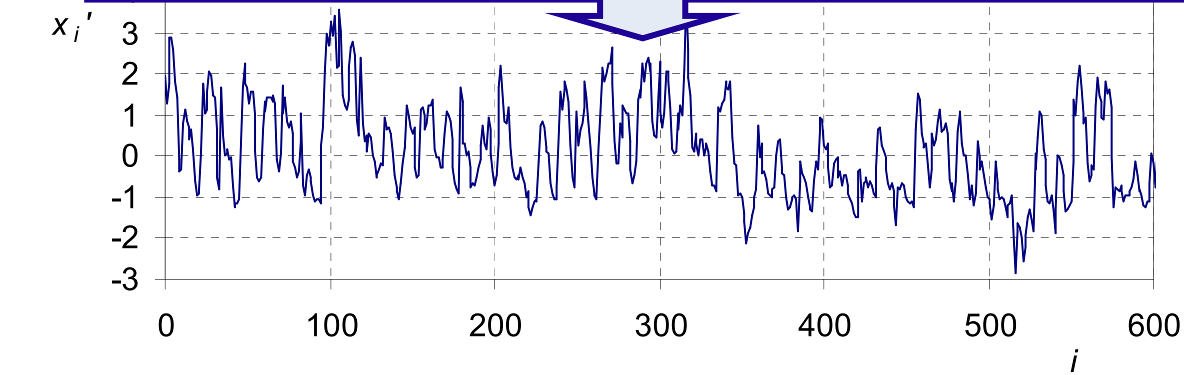


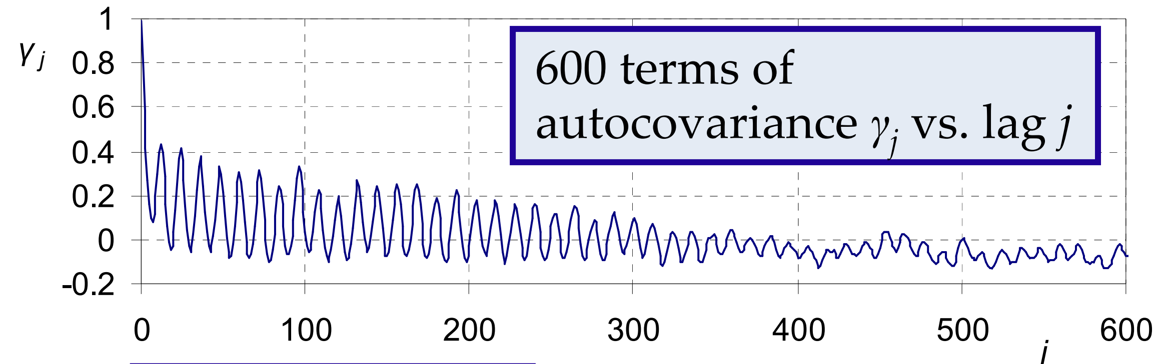
General observations:

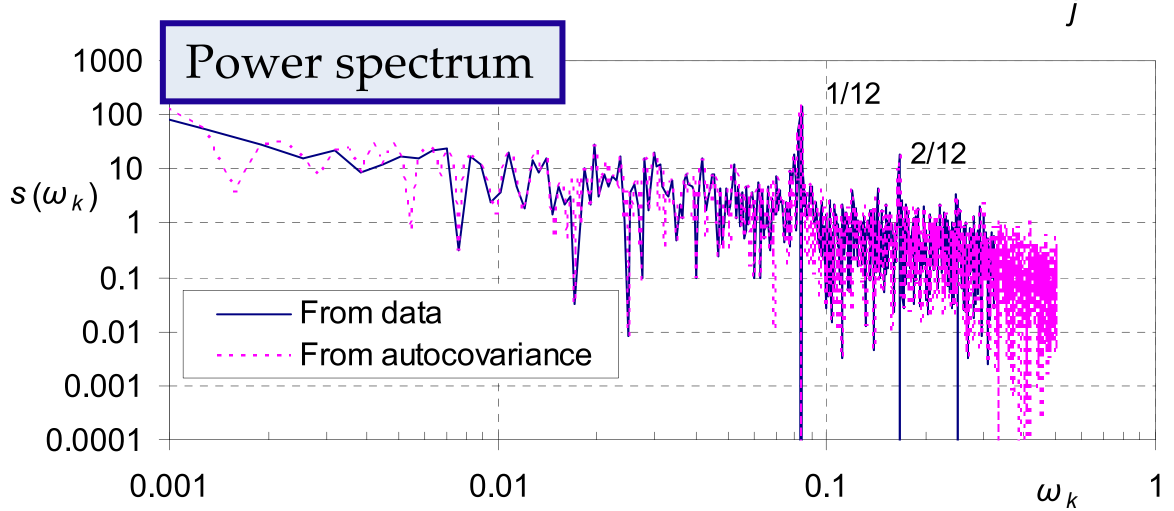
1. All three plots highlight the annual cycle
2. No apparent over-annual cycle

Analysis of the standardized monthly flows

Fifty years of monthly flow cyclically standardized





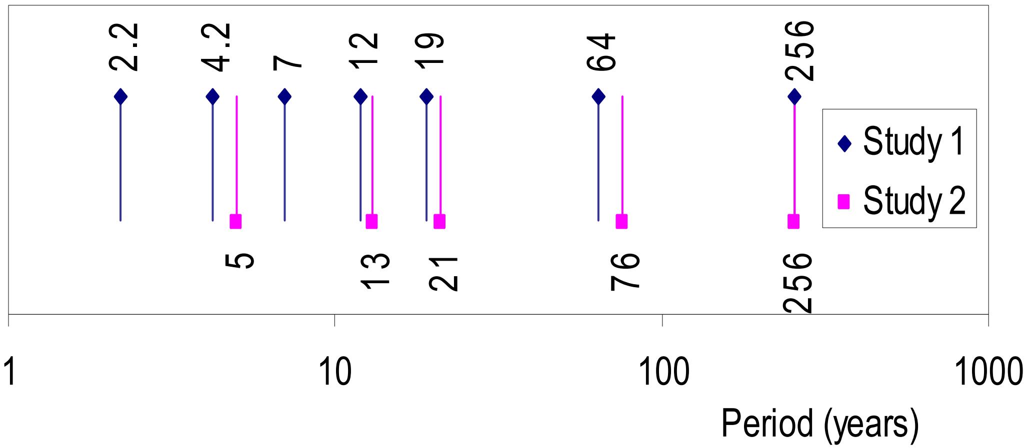


General observations:

1. The annual cycle cannot be 'removed' by standardizing the monthly time series
2. 'Linear' logics implying additivity of the type 'periodic + aperiodic' or 'signal + noise' may be inappropriate for natural processes
3. No apparent over-annual cycle

Seeking over-annual cycles

Several studies have claimed to have detected several over-annual periods from the Nilometer time series Here are two recent examples



Authors' opinion:

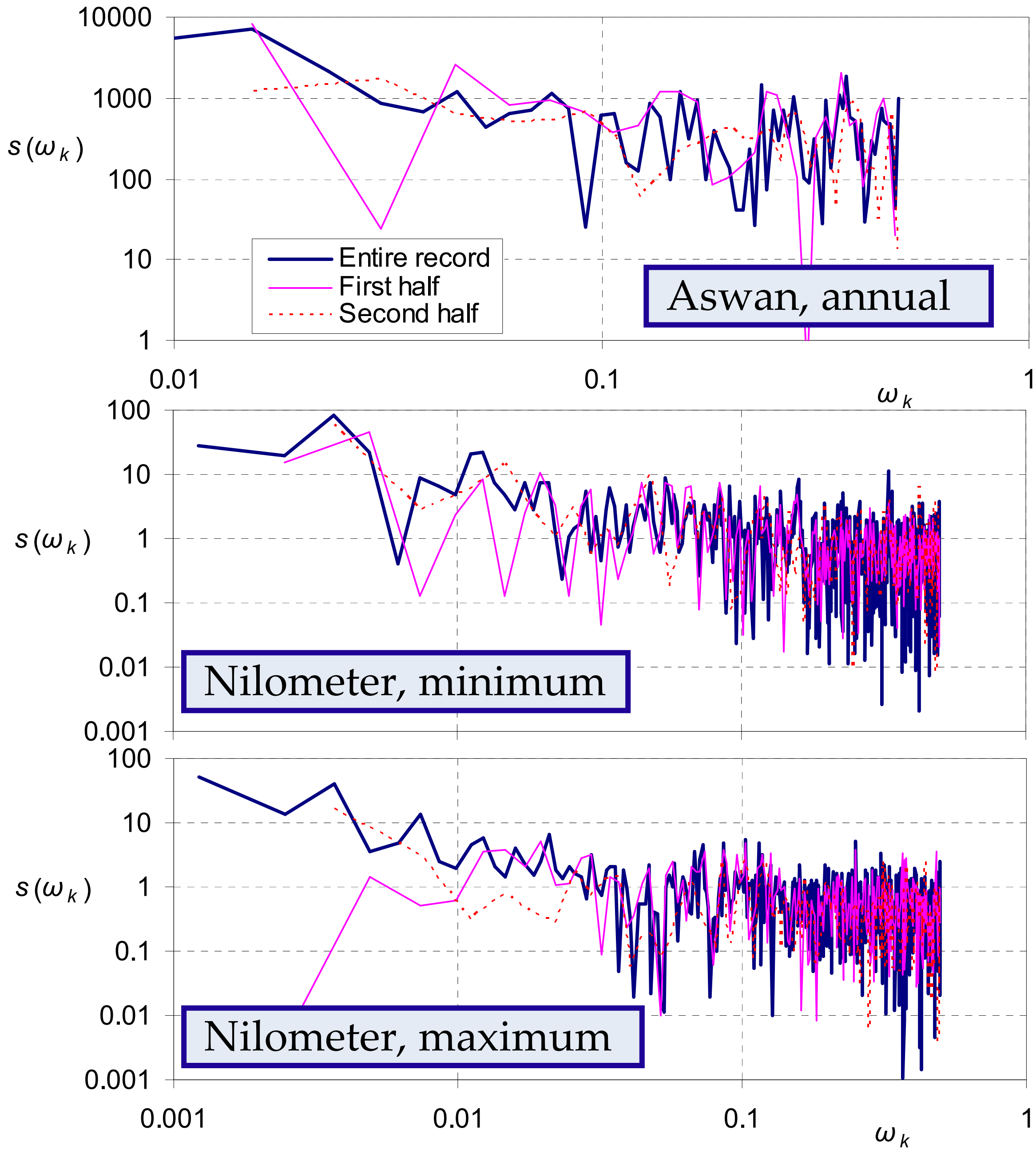
Such claims may be suspicious for several reasons

1. The plethora of periods, which are numbers asymmetric to each other, may rather indicate a stochastic behaviour (all frequencies are significant)
2. Some studies test the significance of detected periods against white noise; apparently, a white noise hypothesis is totally inconsistent with the Nile behaviour
3. Some studies may have undervalued the estimation uncertainty in stochastic processes with high autocorrelation

Seeking over-annual cycles: Aswan and Nilometer annual series

A simple methodology:
A 'real' peak in the power spectrum (manifesting determinism rather than a random effect) will appear also at **the same frequency** if we **split** the sample in two halves or in three tertiaries

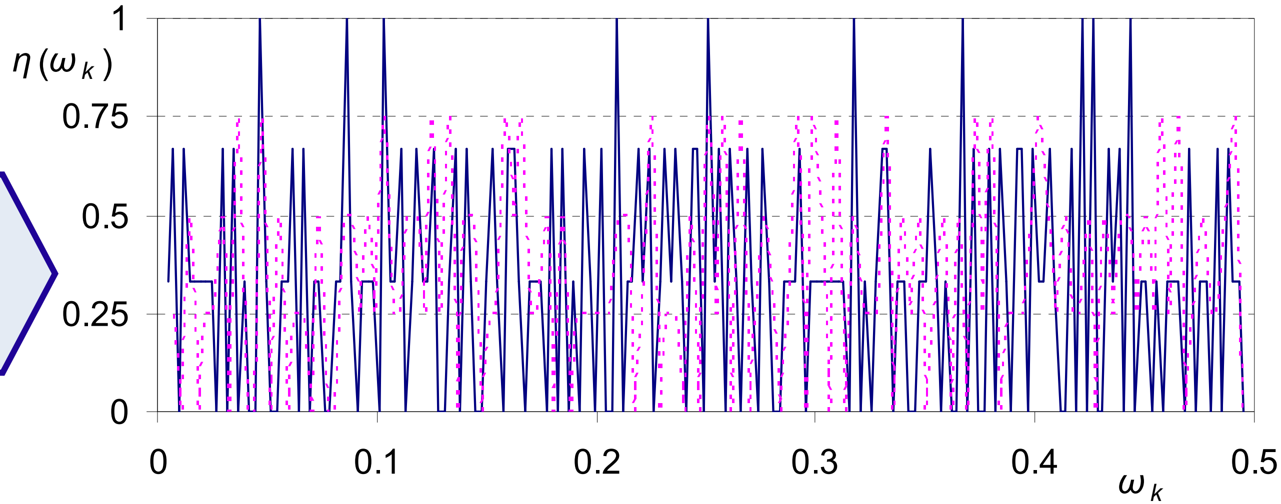
Conclusion:
No over-annual cycle appears in any of the three time series



Algorithmic details in seeking over-annual cycles

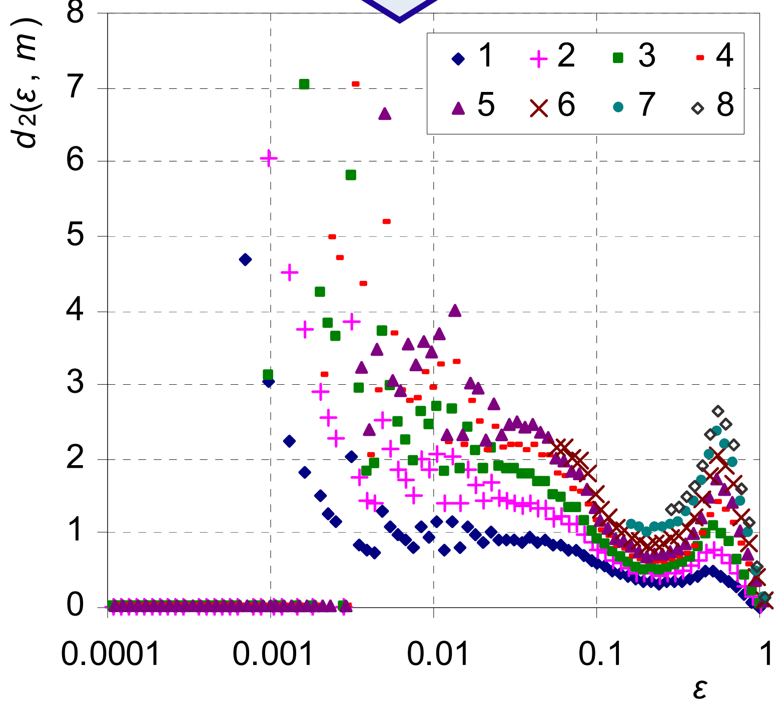
1. We split the sample into two halves; if the sample size n is not even we omit one value so that the half-sample size $v = n/2$ is integral
2. We estimate the periodograms of the full time series and its two halves at frequencies $\omega_k = k/v$ for integral k such that $1/v \le \omega_k \le 1/2$
3. For each ω_k we calculate a likelihood measure of a 'real' peak $\eta(\omega_k)$ as the number of peaks in the three periodograms divided by 3 (a real peak should have $\eta(\omega_k) = 1$)
4. We repeat (analogously) steps 1-3 splitting the sample into three tertiaries (here $v = n/3$)
5. We plot $\eta(\omega_k)$ vs. ω_k for the two cases (halves, tertiaries) and locate frequencies ω in whose neighborhood $\eta(\omega) = 1$ in both cases
6. For these frequencies, we inspect in the periodogram of the full series whether or not the peak is higher than neighbouring peaks; if yes we can say that the identified ω manifest determinism rather than a random effect

Power spectrum peaks plot, Nilometer, maximum

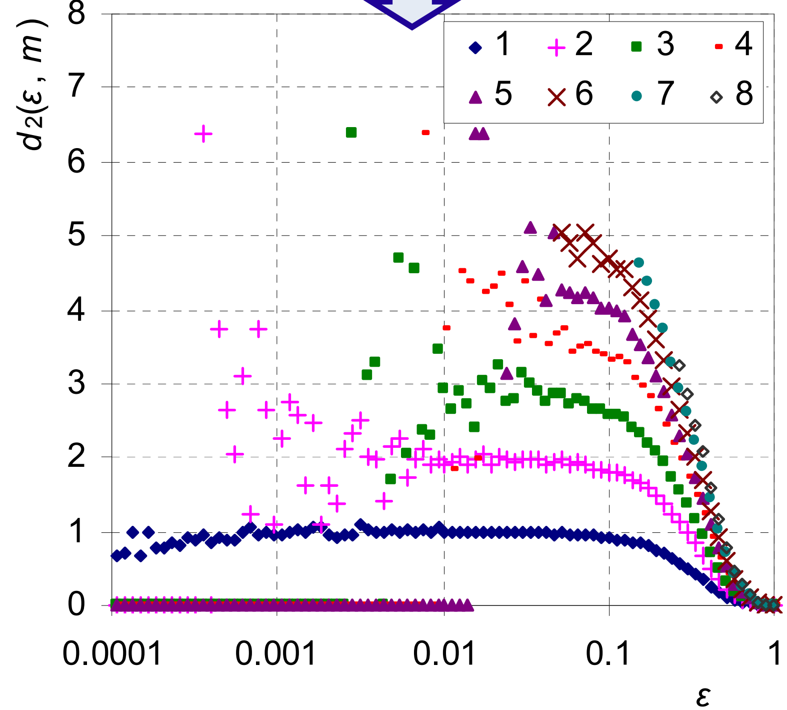


Seeking a chaotic attractor

Natural monthly flows of Nile, Time lag 1



Standardized monthly flows of Nile, Time lag 6



Notation: ε = scale length; m = embedding dimensions (from 1 to 8 as indicated in the legends); $d_2(\varepsilon, m)$ = local slope of correlation sums

Conclusion: No low-dimensional determinism

Part 2
Stochastic description

Multi-scale setting of stochastic analysis

A process at continuous time t	$X(t)$
The averaged process at a scale k	$X_i^{(k)} := \frac{1}{k} \int_{(i-1)\,k}^{i\,k} X(t)\,dt$
Properties of the process at an arbitrary observation scale $k = 1$ (e.g. annual)	
Standard deviation	$\sigma \equiv \sigma^{(1)}$
Autocorrelation function (for lag j)	$\rho_j \equiv \rho_j^{(1)}$
Power spectrum (for frequency ω)	$s(\omega) \equiv s^{(1)}(\omega)$
Properties of the process at any other scale	Can be derived analytically from those at scale $k = 1$ – depend on ρ_j
Specific properties at any scale of a **simple scaling stochastic** process (SSS - fractional Gaussian noise) with Hurst exponent H	
Standard deviation	$\sigma^{(k)} = k^{H-1}\,\sigma \quad (0.5 < H < 1)$
Autocorrelation function (for lag j)	$\rho_j^{(k)} = \rho_j \approx H\,(2\,H - 1)\,\lvert j \rvert^{2H-2}$
Power spectrum (for frequency ω)	$s^{(k)}(\omega) \approx 4\,(1 - H)\,\sigma^2\,k^{2H-2}\,(2\,\omega)^{1-2\,H}$

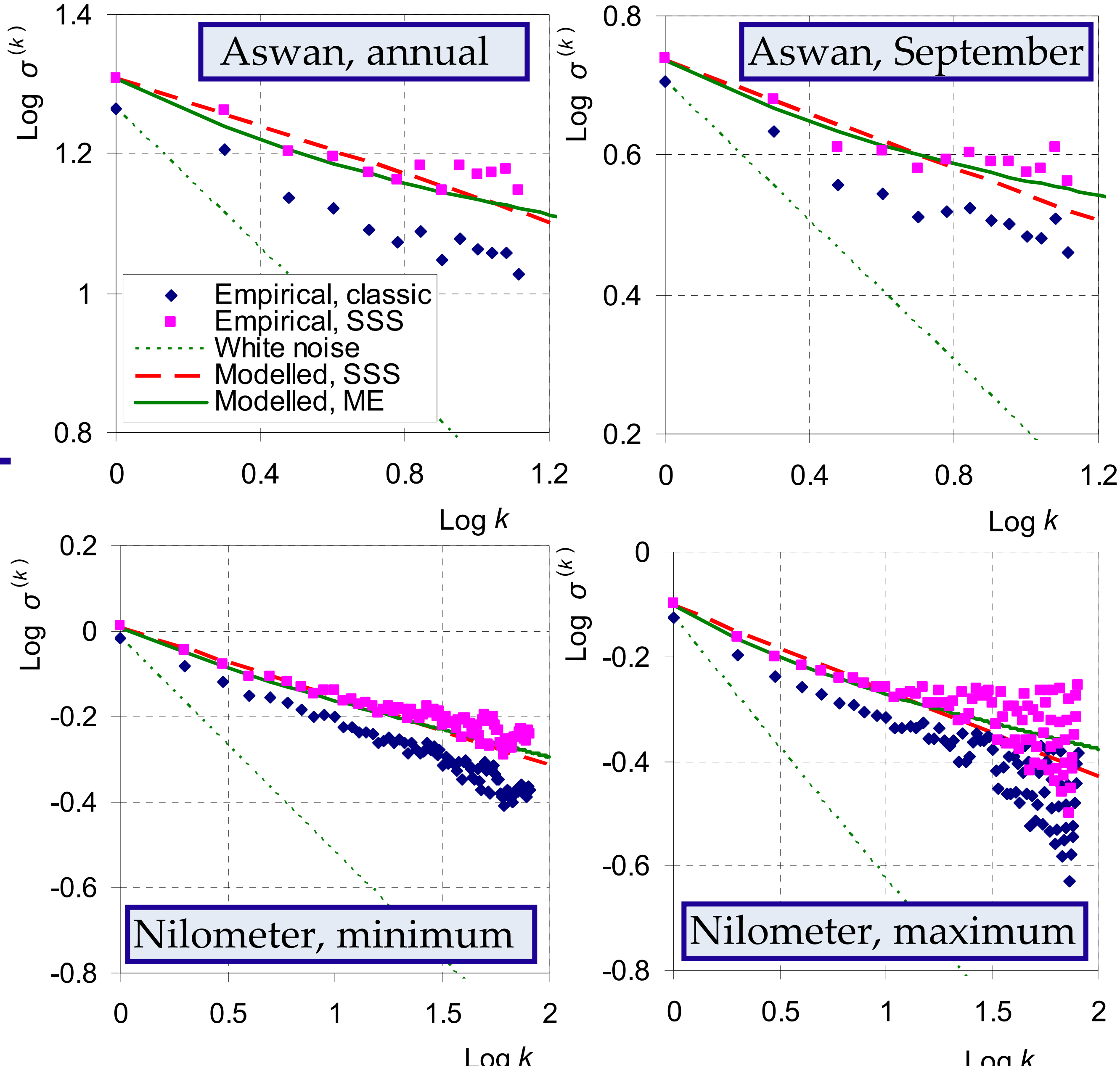
Manifestation of aperiodic fluctuations into the statistics of the time series

Month	μ (km^3)	σ (km^3)	C_s	C_k	τ_3	τ_4	H	ρ_{FGN1}	ρ_1	ρ_2	ρ_{12}
Aug	19.03	4.38	-0.06	-0.28	-0.01	0.11	0.74	0.40	0.58	0.06	0.17
Sep	21.55	5.07	-0.21	-0.18	-0.03	0.11	0.81	0.53	0.74	0.24	0.25
Oct	14.58	4.27	0.17	-0.24	0.01	0.11	0.91	0.77	0.85	0.61	0.50
Nov	8.16	2.31	0.64	-0.14	0.14	0.09	0.85	0.62	0.86	0.72	0.33
Dec	5.57	1.40	1.21	1.91	0.23	0.20	0.90	0.74	0.90	0.75	0.46
Jan	4.33	1.03	0.74	1.05	0.14	0.19	0.87	0.68	0.90	0.75	0.44
Feb	3.17	0.91	0.64	0.53	0.11	0.13	0.82	0.56	0.92	0.76	0.40
Mar	2.76	0.87	0.79	1.25	0.10	0.14	0.83	0.59	0.90	0.75	0.41
Apr	2.52	1.05	0.43	-0.87	0.12	0.01	0.93	0.82	0.80	0.62	0.73
May	2.33	1.01	0.46	-1.11	0.14	-0.02	0.94	0.83	0.94	0.73	0.79
Jun	2.24	0.82	0.82	0.69	0.15	0.07	0.88	0.69	0.78	0.69	0.46
Jul	5.26	1.76	0.74	0.58	0.13	0.13	0.83	0.59	0.57	0.35	0.39
Average			0.53	0.27	0.10	0.10	0.86	0.65	0.81	0.59	0.44
Annual	91.51	18.38	0.48	0.23	0.11	0.13	0.83	0.58	0.32	0.32	

Series	μ (m)	σ (m)	C_s	C_k	τ_3	τ_4	H	ρ_{FGN1}	ρ_{FGN2}	ρ_1	ρ_2
Minima	2.70	0.96	0.70	3.20	0.05	0.16	0.84	0.60	0.46	0.46	0.38
Maxima	9.27	0.75	-0.13	6.24	-0.01	0.22	0.84	0.59	0.45	0.43	0.36

High autocorrelation (ρ) – High Hurst coefficient ($H > 0.5$)

Empirical statistical analysis 1: Standard deviation vs. scale

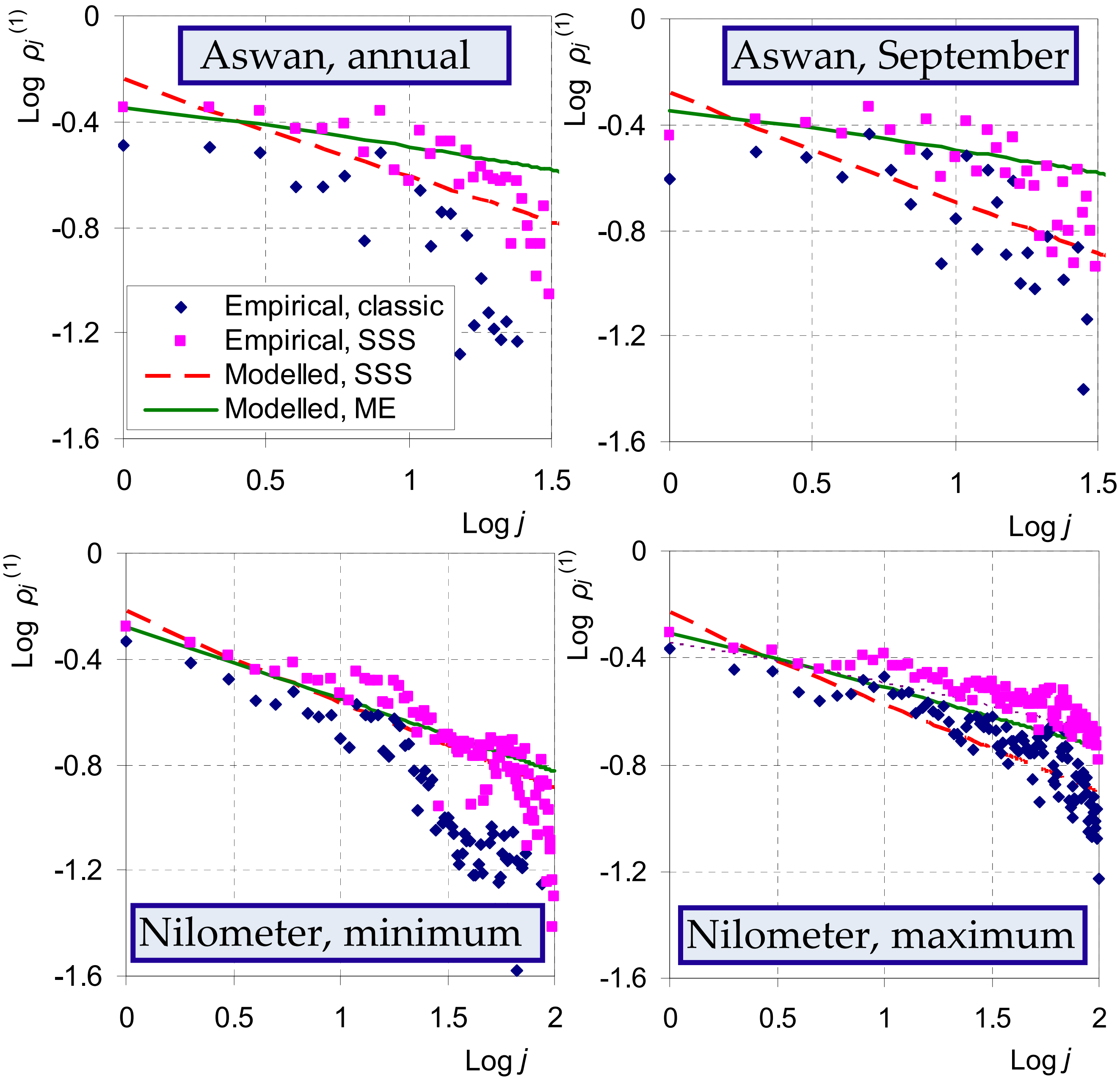


Observation:

In all cases an SSS model seems satisfactory

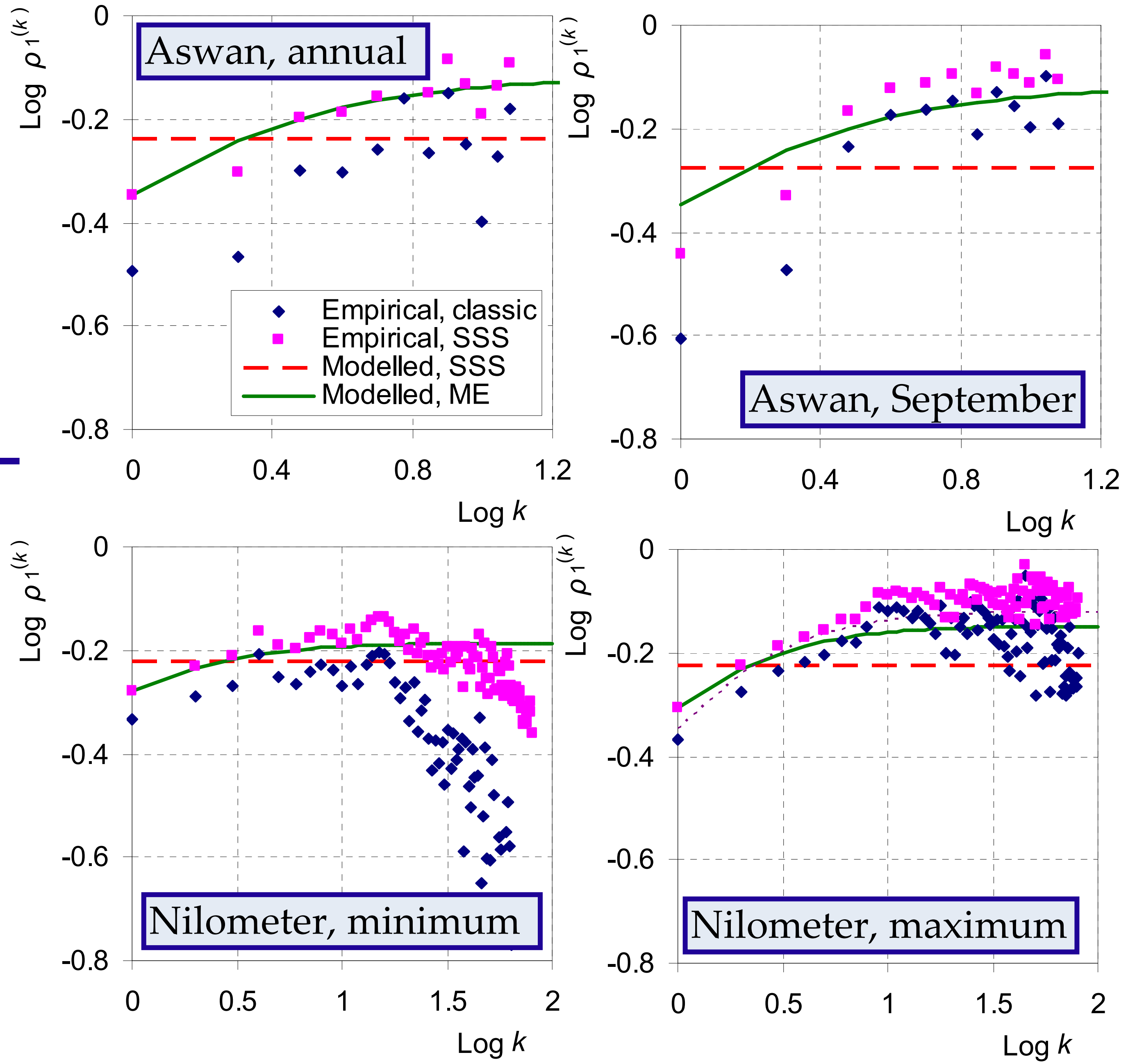
(White noise is far from satisfactory)

Empirical statistical analysis 2: Auto-correlation vs. lag at scale 1



Observation:
In all cases an SSS model seems satisfactory

Empirical statistical analysis 3: Auto-correlation vs. scale at lag 1



Observation:

In all cases even an SSS model seems not very satisfactory

Some type-"why?" questions

- Why the **temperature** of this room tends to equal that of the environment (in the absence of heating/cooling systems)
- Why the probability for each outcome of a **die** is 1/6?
- Why the **normal** (**non scaling**) distribution is so common for processes with relatively low variation?
- Why variables with high variation tend to have asymmetric inverse-J-shaped (rather than bell-shaped) distributions?
- Why variables with high variation tend to have a **scaling** behaviour **in state** (i.e. **non-normal** distribution)?
- Why **stochastic** dependence of natural quantities in consecutive time steps appears so commonly to be **linear**?
- Why the Hurst phenomenon (**scaling** behaviour **in time)** is so common in geophysical, biological, socioeconomical and technological processes?

Because this behaviour maximizes entropy (εντροπία entropy entropie Entropie entropia entropía entropi entrópia entroopia entropija энтропия ентропія)

What is entropy?

- For a discrete random variable X taking values x_j with probability mass function $p_j \equiv p(x_j)$, the Boltzmann-Gibbs-Shannon (or extensive) entropy is defined as

$$\phi := E[-\ln p(X)] = -\sum_{j=1}^{w} p_j \ln p_j, \quad \text{where} \quad \sum_{j=1}^{w} p_j = 1$$

- For a continuous random variable X with probability density function $f(x)$, the entropy is defined as

$$\phi := E[-\ln f(X)] = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, dx, \quad \text{where} \quad \int_{-\infty}^{\infty} f(x)\, dx = 1$$

- Several generalizations of the entropy definition have been proposed; the most promising seems to be that of Tsallis (1988)
- In both cases the entropy ϕ is a measure of **uncertainty** about X and equals the **information** gained when X is observed
- Sometimes entropy is regarded as a measure of order or disorder and complexity

Entropic quantities of a stochastic process

- The *order 1 entropy* (or simply *entropy* or *unconditional entropy*) refers to the marginal distribution of the process X_i:

$$\phi := E[-\ln f(X_i)] = -\int f(x) \ln f(x)\, dx$$

- The *order n entropy* refers to the joint distribution of the vector of variables $\mathbf{X}_n = (X_1, \ldots, X_n)$ taking values $\mathbf{x}_n = (x_1, \ldots, x_n)$:

$$\phi_n := E[-\ln f(\mathbf{X}_n)] = -\int f(\mathbf{x}_n) \ln f(\mathbf{x}_n)\, \mathrm{d}\mathbf{x}_n$$

- The *order m conditional entropy* refers to the distribution of a future variable (for one time step ahead) conditional on known m past and present variables (Papoulis, 1991):

$$\phi_{\mathrm{c},m} := E[-\ln f(X_1 \mid X_0, \ldots, X_{-m+1})] = \phi_m - \phi_{m-1}$$

- The *conditional entropy* refers to the case where the entire past is observed:

$$\phi_\mathrm{c} := \lim_{m \to \infty} \phi_{\mathrm{c},m}$$

- The *information gain* when present and past are observed is:

$$\psi := \phi - \phi_\mathrm{c}$$

The theory is for discrete time stationary processes

What is the principle of maximum entropy (ME)?

- In a probabilistic context, the principle of ME was introduced by Jaynes (1957) as a generalization of the “principle of insufficient reason” attributed to Bernoulli (1713) or to Laplace (1829)
- In a probabilistic context, the ME principle is used to infer unknown probabilities from known information
- In a physical context, the ME principle determines thermodynamical states
- The principle 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
- Typical constraints used in a probabilistic or physical context are:

Mass; Mean/Momentum; Non-negativity

$$\int_{-\infty}^{\infty} f(x)\, dx = 1, \quad E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx = \mu \qquad x \geq 0$$

Variance/Energy; Dependence/Stress

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \sigma^2 + \mu^2, \; E[X_i X_{i+1}] = \int_{-\infty}^{\infty} x_i x_{i+1} f(x_i, x_{i+1})\, dx_i\, dx_{i+1} = \rho\, \sigma^2 + \mu^2$$

Application of the ME principle at the basic time scale

- Maximization of either ϕ_n (for any n) or ϕ_c with the mass/mean/variance constraints results in **Gaussian white noise,** with maximized entropy

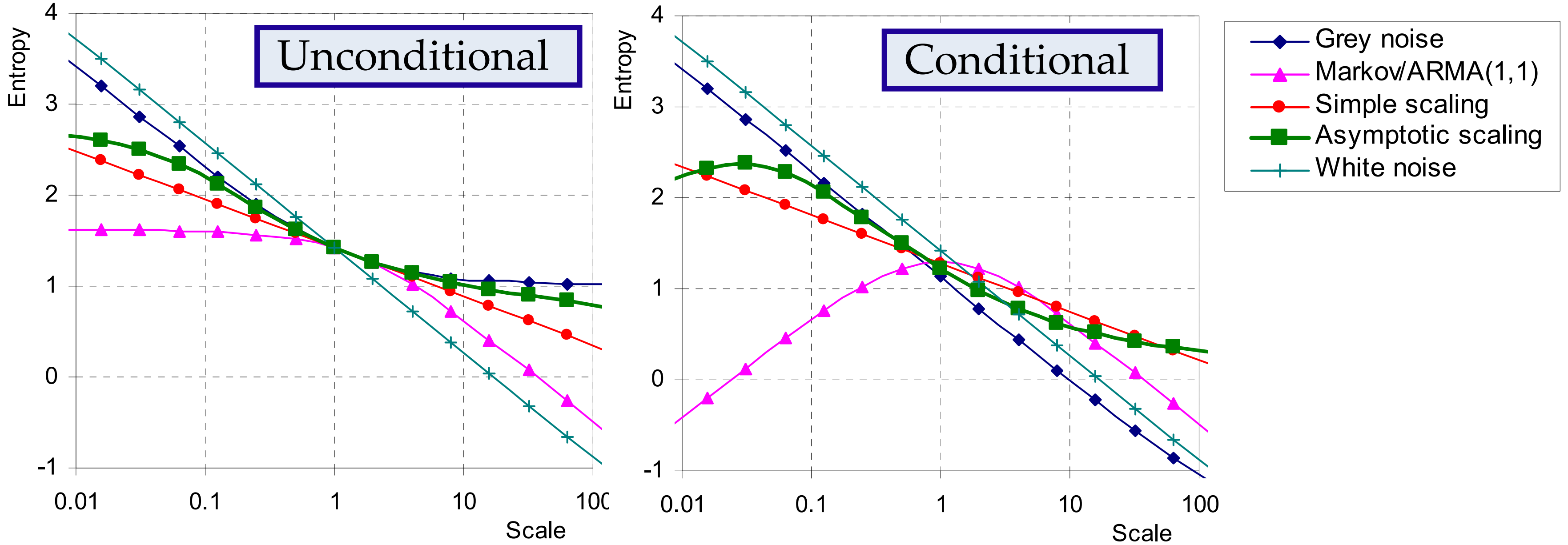
$$\phi = \phi_c = \ln(\sigma \sqrt{2\pi e}), \quad \phi_n = n\,\phi$$

  and information gain $\psi = 0$. This result remains valid even with the non-negativity constraint if variation is low ($\sigma/\mu \ll 1$)
- Maximization of either ϕ_n (for any n) or ϕ_c with the additional constraint of dependence with $\rho > 0$ (for lag one) results in a **Gaussian Markov process** with maximized entropy

$$\phi = \ln(\sigma \sqrt{2\pi e}), \quad \phi_c = \ln[\sigma \sqrt{2\,\pi\, e\,(1-\rho^2)}], \quad \phi_n = \phi + (n-1)\,\phi_c$$

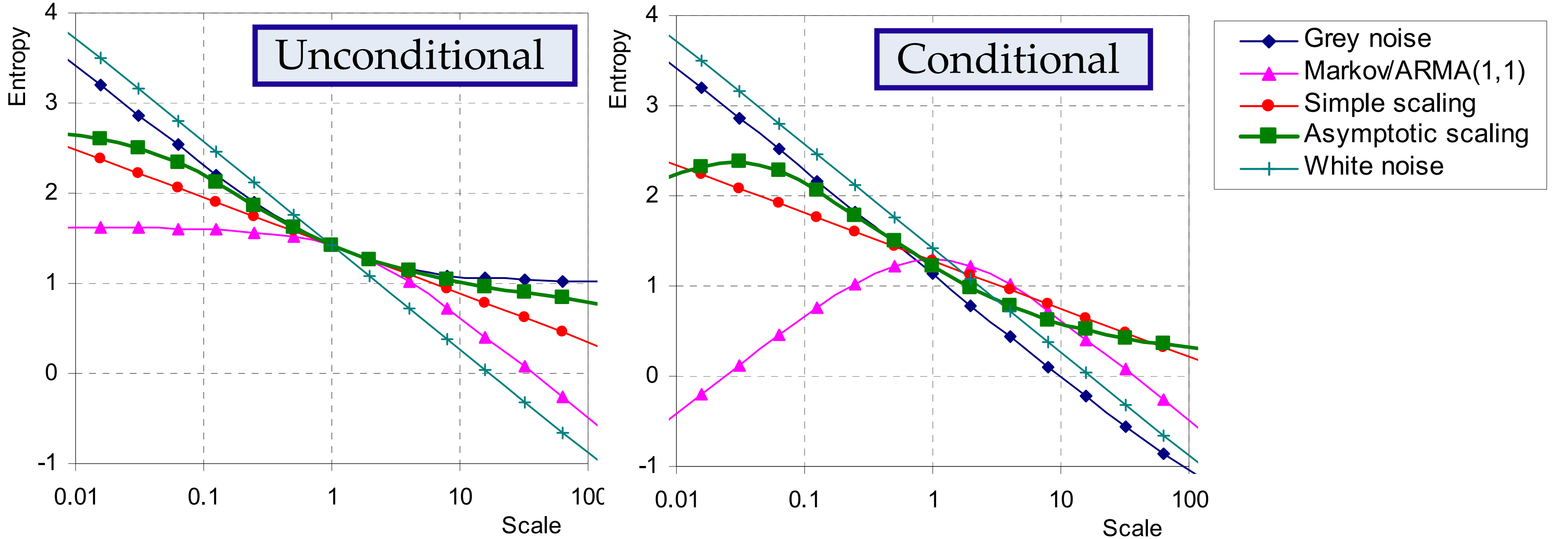
  and information gain $\psi = -\ln\sqrt{1-\rho^2}$
- The question is how we should maximize entropy in a continuous time process:
  - At the observation scale? (=> a Markov process – but not reasonable)
  - At multiple time scales simultaneously? (see Koutsoyiannis, 2005 => SSS)
  - At scale tending to zero? (a 'local' contemplation)
  - At scale tending to infinity? (a 'global' contemplation)

Entropy extremizing solutions - Comparison



- All five solutions are constrained by $\sigma = 1$ at the annual scale ($k = 1$);
- All solutions but white noise are also constrained by $\rho = 0.449$, as in the Nile flow record at the annual scale ($k = 1$); white noise is plotted just for comparison
- To derive the simple scale and asymptotic scaling solutions two additional inequality constraints (restrictions) were used
  1. $\psi^{(0)} \geqslant \psi^{(k)}$ for any $k > 0$ (meaning that predictability at any time scale k is lower than that instantaneously after the measurement)
  2. $\psi(\infty) < \infty$ (prohibiting an illimitable predictability at very large scales)

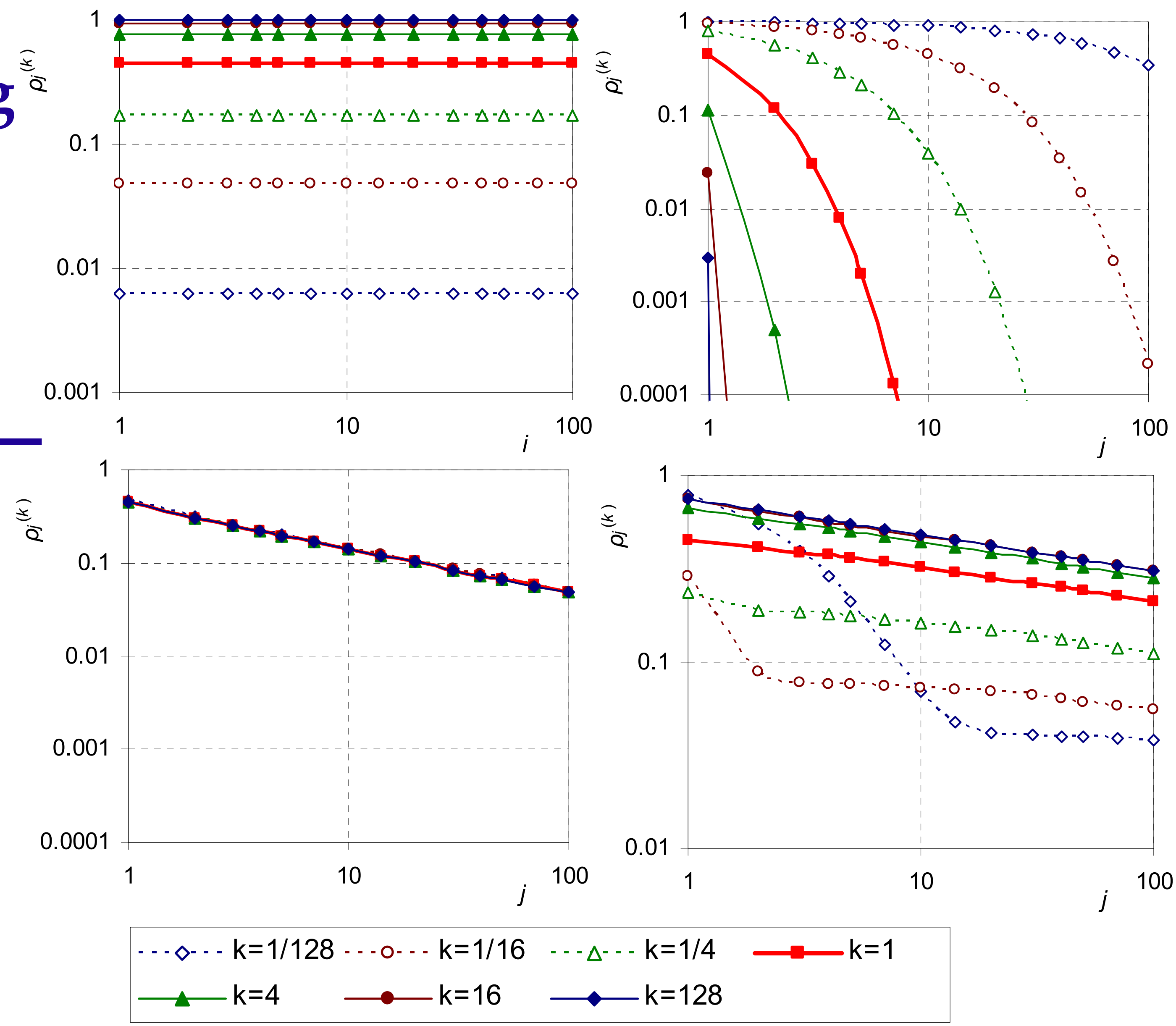
Entropy extremizing solutions - Comparison



	Small scales (→ 0)		Scale 1	Large scales (→ ∞)	
	UE	CE	CE	UE	CE
Grey noise	Max	Max		Max	Min
Markov/ARMA(1,1)	Min	Min	Max	Min	
Simple scaling	Max/R	Max/R		(Large)	(Large)
Asymptotic scaling	(Large)			Max/R	Max/R

Note R: solutions with additional restrictions

Entropy extremizing solutions – Auto-correlation functions



Conclusions

- Apart from the annual cycle, no other signatures of determinism appear in the long records of the Nile River
- The type of stochastic behaviour observed in the Nile records is very different from white noise and suggests a simple scaling or asymptotic scaling process
- Both these processes are consistent with the maximum entropy principle and suggest much higher entropy on small and large scales in comparison to Markov processes
- Thus, the observed behaviour can be interpreted as dominance of uncertainty in nature
- (… which makes the world interesting, i.e. not boring)

Both classical physics and quantum physics are indeterministic
Karl Popper (in his book "Quantum Theory and the Schism in Physics")
The future is not contained in the present or the past
W. W. Bartley III (in Editor's Foreward to the same book)

References

- Bernoulli, J. (1713), *Ars Conjectandi*, Reprinted in 1968 by Culture et Civilisation, Brussels
- Jaynes, E.T. (1957), Information Theory and Statistical Mechanics I, *Physical Review*, 106, 620-630
- Koutsoyiannis, D. (2005), Uncertainty, entropy, scaling and hydrological stochastics, 2, Time dependence of hydrological processes and time scaling, *Hydrological Sciences Journal*, 50(3), 405-426, 2005
- Laplace, P.S. de, (1829), *Essai philosophique sur les Probabilités*, H. Remy, 5th édition, Bruxelles
- Papoulis, A. (1991) *Probability, Random Variables, and Stochastic Processes* (third edn.), McGraw-Hill, New York, USA
- Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics, *J. Stat. Phys.* 52, 479-487

Both classical physics and quantum physics are indeterministic

Karl Popper (in his book "Quantum Theory and the Schism in Physics")

The future is not contained in the present or the past

W. W. Bartley III (in Editor's Foreward to the same book)