European Geosciences Union General Assembly 2009
Vienna, Austria, 19–24 April 2009

Session HS5.3: Hydrological modelling—Adapting model complexity to the available data: approaches to model parsimony

Seeking parsimony in hydrology and water resources technology

Demetris Koutsoyiannis
Department of Water Resources and Environmental Engineering
Faculty of Civil Engineering
National Technical University of Athens, Greece (dk@itia.ntua.gr)

Presentation available online: http://www.itia.ntua.gr/en/docinfo/906/

Part 1: A stochastic* introduction

(at an elementary level)

* *Stochastic* is used here with the literal modern Greek meaning, something in between sceptical and philosophical.
(στοχάζεσθαι = to think in depth; σκέπτεσθαι = to think; φιλοσοφείν = to love wisdom)

What is the principle of parsimony?

- A principle that advises scientists to prefer the simplest theory among those that fit the data equally well.
- Alternative names: **principle of parsimony**, **principle of simplicity**, **principle of economy**, **Ockham's razor**.
- Example of a parsimonious natural law:
  - Dogs bark.
- Examples of non-parsimonious laws:
  - Black, white and spotted dogs bark.
  - Dogs bark on Mondays, Wednesdays and Fridays.
- Intuitively, the above example law should not exclude the case that a particular dog does not bark (he is mute); it should be mistaken to say "all dogs bark".
- In other words, laws of complex systems (e.g. the biological system "dog") are necessarily probabilistic in nature:
  - "Dogs bark" means "it is very likely that any dog may bark".

Failure to recognize the probabilistic character of parsimony in complex systems may create confusion (see e.g. Courtney and Courtney, 2008, and the "all crows are black" example).

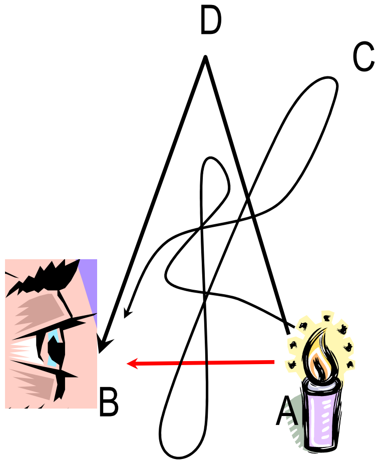
Historical reference

- Aristotle (384–322 BC), *Αναλυτικά Ύστερα* (*Posterior Analytics*): "*We may assume the superiority* ceteris paribus [other things being equal] *of the demonstration which derives from fewer postulates or hypotheses*".
  Also: "[the principles] *should, in fact, be as few as possible, consistently with proving what has to be proved*".
- Claudius Ptolemy (100-178 AD), *Μαθηματική Σύνταξις* (*Mathematical Treatise* or "*Almagest*"): parsimony helps decide between theories about planetary motions.
- Medieval philosophers: Robert Grosseteste (c. 1168-1253), Thomas Aquinas (c. 1225-1274), William of Ockham (c. 1285-1347; "*Plurality is not to be posited without necessity*").
- Nicolaus Copernicus (1473-1543), Galileo Galilei (1564-1642), Isaac Newton (1642-1727)—all used parsimony in developing their theories.
- Albert Einstein's formulation of parsimony: "*Everything should be made as simple as possible, but not simpler*".

Read more about the history and philosophy of parsimony and the scientific method in the compelling book by Gauch (2003).

Is the principle of parsimony epistemological or ontological?

- Ockham insisted that parsimony was just an epistemological principle for choosing the best theory.
- However, earlier philosophers, from Aristotle to Grosseteste had interpreted parsimony also as an ontological principle, thus expecting Nature to be simple.



- Light follows the simplest path from A to B (the red line) and not other more complex ones (e.g. the black lines ACB, ADB)?
- But what does "**simplest**" mean?

Were Nature **not parsimonious** (e.g. were paths ACB, ADB materialized) it would be difficult to understand her and life would be hard.

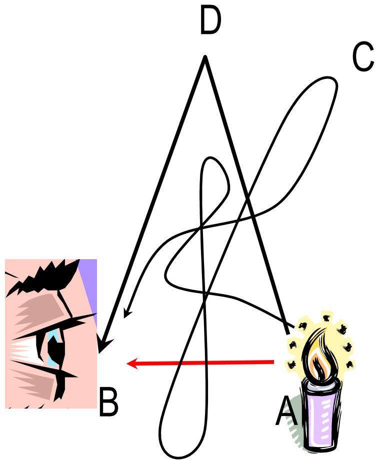
Quantification of simplicity

- The traditional approach to physics is based on writing equations.
  - Equations generally express conservation laws.
  - Such laws govern the following quantities:
    - Mass (scalar equation);
    - Linear momentum (vector equation);
    - Angular momentum (vector equation);
    - Energy (scalar equation);
    - Electric charge (scalar equation).
  - Other quantities (e.g. acceleration) are not conserved.
  - These laws refer to isolated systems that do not exchange heat and mass with the environment (in open systems there is no conservation).
- However, to find states or paths which are "**as simple as possible**", it seems more natural to formulate the problem in terms of optimization rather that using equations.
- A single optimization of a scalar function $f(\boldsymbol{s})$ of a vector quantity $\boldsymbol{s}$, i.e. a mathematical expression "optimize $f(\mathbf{s})$" is more powerful than an **equation** form, i.e. "$\boldsymbol{g}(\boldsymbol{s}) = \mathbf{0}$" (the "**optimation**" form is equivalent to as many equations as required).

Nature seems to be an **optimizer—not** an **equalizer**.

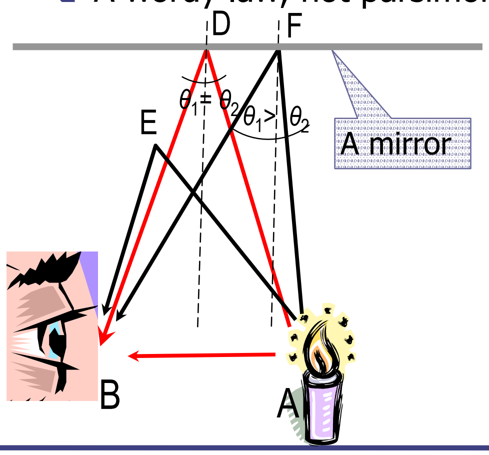
Quantification of simplicity in light trajectory Attempt 1

- Light follows the shortest possible path from A to B.
  - A parsimonious law for a parsimonious natural behaviour.
  - Further investigation will show that it is not correct (formulation simpler than "as simple as possible").



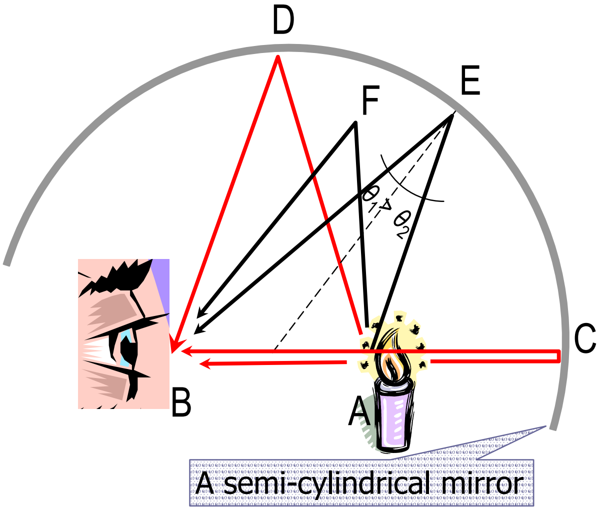
Quantification of simplicity in light trajectory Attempt 2

- In the presence of a mirror, light follows both red paths from A to B (AB, ADB)—but not other (the black) ones (e.g. AEB, AFB).
- The previous formulation of the law is not valid.
- Replacement: Light follows the shortest path, but when there is a mirror, it also follows a second path with a reflection by the mirror such that the angle of incidence equals the angle of reflection.
  - A wordy law, not parsimonious ("equalizer" thinking…).



- We observe that the mirror has imposed an inequality constraint to possible paths (by disallowing light to go through it) and thus "generated" a second minimum in the "shortest path" problem.
- The paths followed by light have minimum length (either global or local minimum).
- Parsimonious law—Principle of Hero of Alexandria (~1st cent. BC)

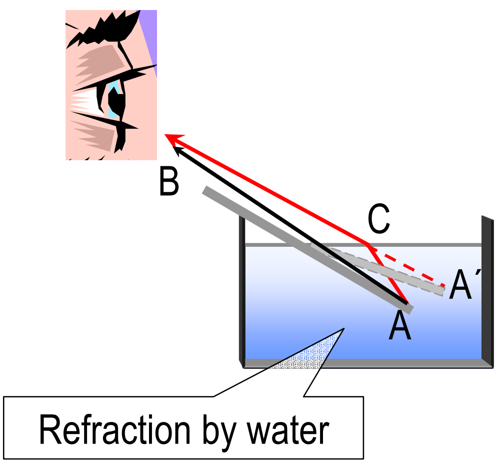
Quantification of simplicity in light trajectory Attempt 3



- If we replace the flat mirror with a cylindrical mirror, the light follows three paths from A to B (red lines AB, ACB, ADB).
- Of these, AB is the global minimum, ACB is a local minimum and ADB is a local maximum.
- The paths followed by light have extremal length (either global or local minimum or maximum).

Nature is a skilful optimizer, as she finds **all** local minima and maxima (put many mirrors to see lots of paths materializing).
Failure to observe this makes things difficult to explain, as indicated for instance in the debate by Gaertner (2003) and Schoemaker (2003).

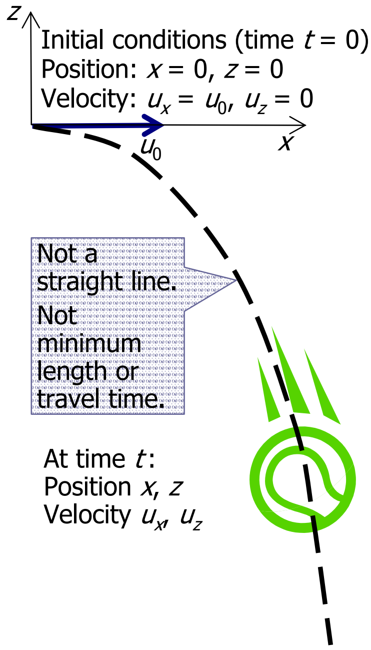
Quantification of simplicity in light trajectory Attempt 4



- Refraction makes clear that light does not always follow the shortest (straight line) path; here it follows the broken line path ACB.
- This is related to the fact that the light speed in water is smaller than in air.
- It is easily proved that the path ACB has the least travel time.
- Final law (Fermat’s principle, corrected for extremum—instead of minimum):
Light follows paths that have **extremal travel time**.

Nature **is indeed parsimonious** (ontological parsimony).
The final law is parsimonious (epistemological parsimony), reflecting the parsimony of Nature.

Generalization to the trajectory of a weight
The principle of extremal (stationary) action



Quantities involved

- Potential energy: $V = m g z$
- Kinetic energy: $T = (1/2) m u^2 = (1/2) m (u_x^2 + u_z^2)$
- Lagrangian: $L = T - V = (1/2) m (u_x^2 + u_z^2) - m g z$
- Action: $S = \int_\Pi L \, dt$ along the path Π

Principle of extremal action (Hamilton)

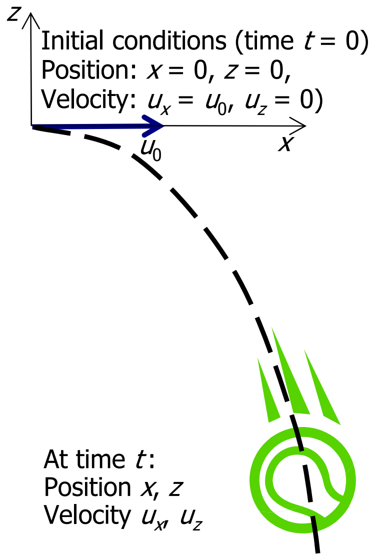
- From all possible motions between two points, the true motion has extremal (stationary) action.
- Credit for the principle is given to Pierre-Louis Moreau de Maupertuis, who wrote about it in 1744; Leonhard Euler discussed it in 1744, whereas Gottfried Leibniz preceded both by 39 years.
- The principle is applicable both in classical and in quantum physics).

Solution

- Extremization of action results in Euler-Lagrange equation:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial u_x}\right) - \frac{\partial L}{\partial x} = 0, \frac{d}{dt}\left(\frac{\partial L}{\partial u_z}\right) - \frac{\partial L}{\partial z} = 0$$

The trajectory of a weight
Application of the principle of extremal action



- The Euler-Lagrange equation results in a single (global) minimum (**least action**):

  $$u_x = u_0 \text{ (= constant)}, \; u_z = -g t$$

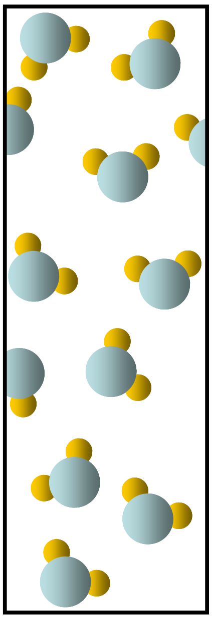
  from which we obtain:

  $$x = u_0 t, \; z = -g t^2/2 \text{ or } z = -(g / 2u_0^2) \, x^2$$

  (parabola; going down).
- In the above formulation, we have not used Newton's laws.
- We did not even use preservation of energy.
- Rather, the preservation of total energy $E = T + V$ results from the least action solution.
- The solution gives not only the geometry (parabola) and direction (down) of the trajectory but the full description of the movement of the weight.

A single principle (Hamilton's with Fermat's as a special case) describes diverse phenomena in optics and classical mechanics. (**Real parsimony!**)

From simple to complex systems



- When we have a system of many "bodies" (e.g. many molecules of water in solid, liquid or gaseous phase), we are not interested on the properties (position, momentum) of each particular body.
- Even if we were interested, it would be difficult (and extremely unparsimonious) to know them; e.g. 1 m^3 of a gas in standard conditions contains 2.7×10^{25} molecules.
- Only macroscopic/statistical (or thermodynamical) properties of the system are of interest.
- Macroscopic properties are state variables such as pressure, internal energy, entropy, temperature, and characteristic constants such as specific heat and latent heat.
- Inevitably—albeit often not stated explicitly—macroscopic descriptions are probabilistic descriptions and involve uncertainty.
- However, when the system components are very many and identical, due to the applicability of the laws of large numbers, uncertainty becomes almost certainty.

When we move from single to complex systems, parsimony demands replacement of microscopic with macroscopic properties and of deterministic with probabilistic descriptions.

What does Nature "extremize" in complex systems?

- The quantity that gets extremized seems to be the **entropy**.
- The word is (ancient) Greek*.
- The scientific term is due to Clausius (1850-1865).
- The entropy concept was fundamental to formulate the **second law of thermodynamics**.
- Boltzmann (1866) showed that the entropy of a macroscopic stationary state is proportional to the logarithm of the number w of possible microscopic states that correspond to this macroscopic state.
- Gibbs (1902) studied the concept further in a statistical mechanical context.
- Shannon (1948) generalized the mathematical form of entropy and also explored it further.
- Kolmogorov (1956, 1958) founded the concept on more mathematical grounds on the basis of the measure theory and introduced entropy to the theory of dynamical systems.

*Εντροπία, a feminine noun from the verb εντρέπομαι meaning: to turn into; to turn towards someone's position; to turn round and round.

What is entropy?

- The modern definition of entropy is based of probability theory.
- For a discrete random variable $\underline{x}$ taking values x_j with probability mass function $p_j \equiv p\,(x_j)$, $j = 1,\ldots,w$, the Boltzmann-Gibbs-Shannon (or extensive) entropy is defined as

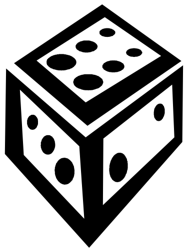
$$\varphi := E[-\ln p(\underline{x})] = -\sum_{j=1}^{w} p_j \ln p_j\,, \qquad \text{where } \sum_{j=1}^{w} p_j = 1$$

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

$$\varphi := E[-\ln f(\underline{x})] = -\int_{-\infty}^{\infty} f(x) \ln f(x)\, \mathrm{d}x, \quad \text{where } \int_{-\infty}^{\infty} f(x)\, \mathrm{d}x = 1$$

- In both cases the entropy φ is a measure of **uncertainty** about $\underline{x}$ and equals the **information** gained when $\underline{x}$ is observed.
- Entropy is also regarded as a measure of **order**/**disorder** and **complexity** (e.g. in statistical mechanics, thermodynamics, dynamical systems, fluid mechanics).
- Generalizations of the entropy definition have been introduced more recently (Renyi, Tsallis).

Entropy maximization: The die example



- What is the probability that the outcome of a toss of a die will be i? (i = 1, …, 6)
- The entropy is:

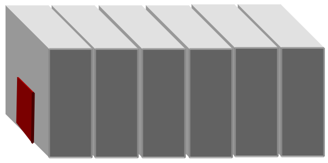
$$\varphi := E\,[-\ln p(\underline{X})] = -p_1 \ln p_1 - p_2 \ln p_2 - \ldots - p_6 \ln p_6$$

- The equality constraint (mass preservation) is

$$p_1 + p_2 + \ldots + p_6 = 1$$

- The inequality constraint is $p_i \geq 0$.
- Solution of the optimization problem (e.g. by the Lagrange method) yields a single maximum: $p_1 = p_2 = \ldots = p_6 = 1/6$.
- This method, the application of the Principle of Maximum Entropy (ME; mathematically, an "optimation" form) is equivalent to the Principle of Insufficient Reason (IR; Bernoulli-Laplace; mathematically, an "equation" form).
- However, ME is much more powerful and can perform in non-symmetric cases (e.g. in a loaded die, when it is known that the probability of a six is twice as high than that of an ace).

Entropy maximization and elementary statistical thermophysics



- Take a room with volume, say, 60 m^3 and assume a partitioning into six slices, 10 m^3 each.
- Assume standard conditions, so that the room contains 60/0.0224 moles of air or $N = 6.022\times10^{23} \times 60/0.0224 = 1.6\times10^{27}$ molecules.
- The case that all molecules are in the front sixth and all other slices are empty is not impossible; it has just very low probability: $(1/6)^N = 10^{-1\ 255\ 000\ 000\ 000\ 000\ 000\ 000\ 000\ 000}$.
- With respect to this partition to six slices, this state ($p_1 = 1$; $p_2 = \ldots = p_6 = 0$) has minimum entropy $\varphi = 1 \ln 1 = 0$.
- The ME principle results again in $p_1 = p_2 = \ldots = p_6 = 1/6$ with maximum $\varphi = \ln 6 = 1.79$.
- Accordingly, the air molecules will be uniformly distributed in the six slices, so that each has an expected number of molecules $E[M] = (1/6)\ N = 2.7\times10^{26}$.
- The standard deviation of the number of molecules is $\mathrm{StD}[M] = [N (1/6)\ (5/6)]^{0.5} = 1.5\times10^{13}$.
- Hence, the variation is $\mathrm{StD}[M]\ /\ E[M] = 5.6\times10^{-14}$.
- Thus, the high uncertainty at the microscopic scale becomes almost certainty at a macroscopic level.
- This enables description of the conditions in the entire room with a few state variables, such as density, pressure, temperature **(Parsimony again!)**.

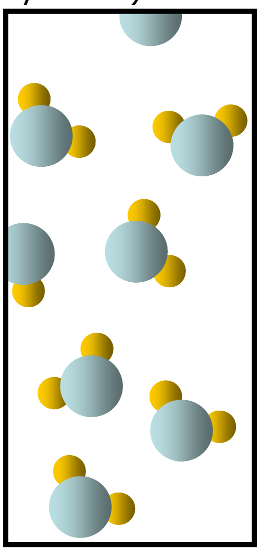
Is the principle of maximum entropy ontological or epistemological?

- In thermodynamics and statistical physics the principle of maximum entropy is clearly ontological:
  - It determines (macroscopic thermodynamical) actual states of physical systems, which can be measured.
- Jaynes (1957) introduced the principle of maximum entropy as an epistemological principle in a probabilistic context:
  - It is used to infer unknown probabilities from known information.
  - The (unknown) density function $f(x)$ of a random variable $\underline{x}$ is the one that maximizes the entropy φ, subject to any known constraints.
- Are these two different principles or one?
- If Nature aligns itself with the (ontological) principle, why not use the same principle in logic for inference about Nature?
- Some more difficult questions: Is the principle of maximum entropy related to the principle of extremal action? Is **extremal action** a special case of **extremal entropy**? Or are the two different special cases of a more general principle?

Part 2: Hydrological applications

Typical thermodynamical vs. hydrological systems

A system of water molecules (classical thermo-dynamics).



Topographical background of a hydrological system (Acheloos River @ Mesounta, Greece; image from Google Earth).

Three-dimensional detail of a hydrological system (credit: Lessovaia *et al.*, 2008).

Identical elements Deduction: possible	Different, unique elements Deduction: impossible; Induction: method of choice

Is the principle of parsimony applicable in the complex hydrological systems?

- The answer underlying the mainstream hydrological research of the last two decades seems to be negative.
- Hopes were invested to the power of computers that would enable faithful and detailed representation of the diverse system elements and the hydrological processes, based on merely "first principles", thus resulting in "physically-based" models that tend to approach in complexity the real world systems.
- Few loud voices debated this direction, e.g.:
  - James Dooge (1997): Searching for simplicity in hydrology.
  - Keith Beven (1993): Prophecy, reality and uncertainty in distributed hydrological modelling.

What is the mathematical tool to reconcile the complexity of hydrological systems with parsimony?

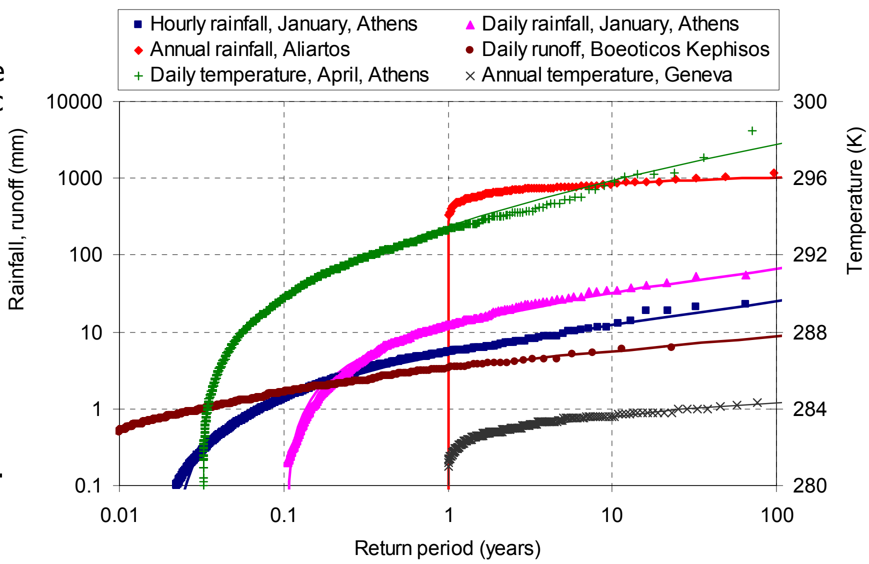
- As in thermodynamics, a consistent theory should necessarily be based on probability—but in an enhanced setting.
- The tool is **Stochastics** = **Probability theory** + **Statistics** + **Stochastic processes**.
- Probability theory provides the theoretical basis for:
  - moving from a microscopic to a macroscopic view of phenomena by mapping sets of diverse elements and events of hydrological systems to single numbers (a probability or an expected value);
  - making induction.
- Statistics provides the empirical basis for:
  - summarizing data;
  - making inference from data; and
  - supporting decision making.
- Stochastic processes and Monte Carlo simulations provide the means for:
  - probabilistic predictions;
  - uncertainty estimation;
  - design and management of complex hydrosystems.

A note on the "enhanced setting" of stochastics for hydrology

- Classical statistics is based on the prototype of independence and repeatability (the "coin-tossing" prototype).
- Hydrological and water resource systems (and more generally, real-world systems) behave differently from the classical prototype.
- Typical stochastic models (particularly the multivariate ones) are often not parsimonious themselves.
- Therefore, substantial advancement of stochastics is necessary to make models more consistent with:
  - observed natural behaviours, and
  - the principle of parsimony.

Application 1: ME and the marginal distribution of hydrometeorological variables

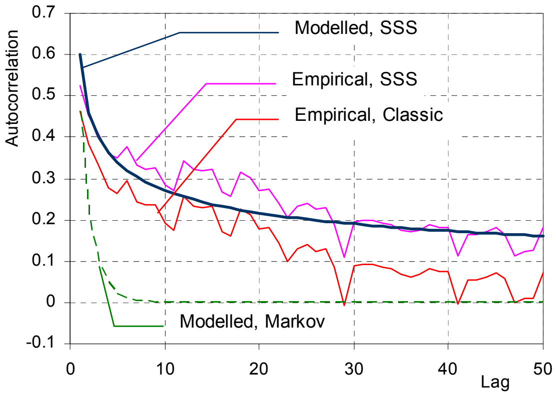
- The mean μ and the standard deviation σ are the most characteristic summary statistics of a hydrometeorological variable.
- Can μ and σ determine the distribution function completely?
- Roughly yes, if combined with the principle of maximum entropy.
- The shape of the distribution (from Pareto to Gaussian) depends on the variation σ/μ.



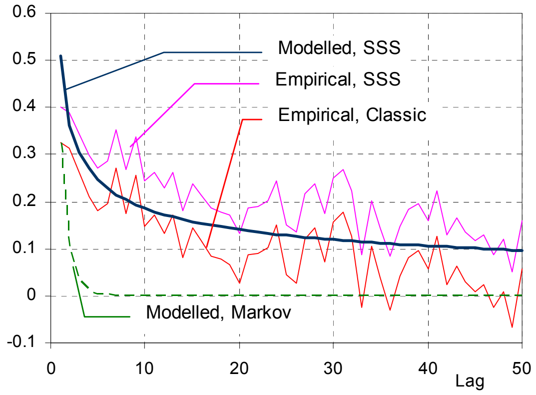
Points: empirical distributions; lines: maximum entropy distributions; see details in Koutsoyiannis (2005a).

Application 2: ME and the dependence structure of hydrometeorological processes

- How many parameters do we need to express the autocorrelation function (ACF, i.e. ρ_j for lag j) of a hydrometeorological process?
- It seems that one parameter may suffice: the ACF is given by entropy maximization at a range of scales with a single dependence constraint, a specified ρ_1.
- The resulting structure is that of the simple scaling stochastic process (SSS process), whereas maximization at a single scale would result in a Markovian process. The parameter best expressing the ACF is the Hurst coefficient *H*.



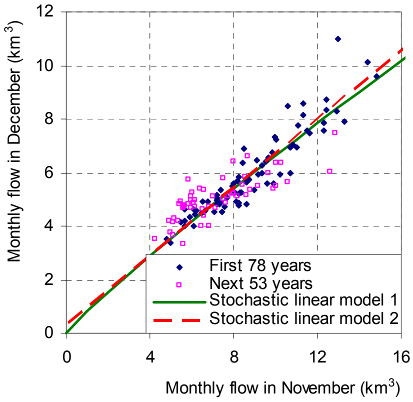
Roda Nilometer annual minimum water level, 640-1452 AD (H = 0.84).



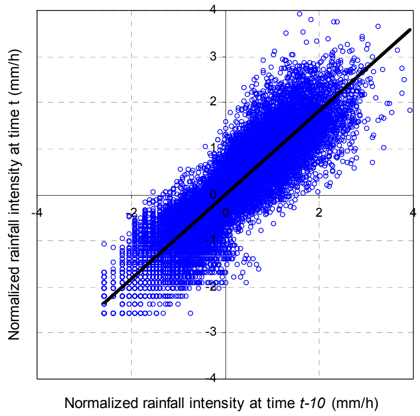
Geneva annual average temperature, 1753-1980 (H = 0.80; Koutsoyiannis, 2005b).

Application 3: ME and linearity of stochastic representations

- How many parameters, in addition to those of the marginal distribution and autocorrelation, do we need to express the full dependence in time of a hydrometeorological process?
- None, because the dependence is linear (if the process is transformed to Gaussian).
- Linearity is a result of ME.
- Linearity in stochastic terms (very common) should not be confused with linearity in deterministic terms (very rare); the two are fundamentally different.



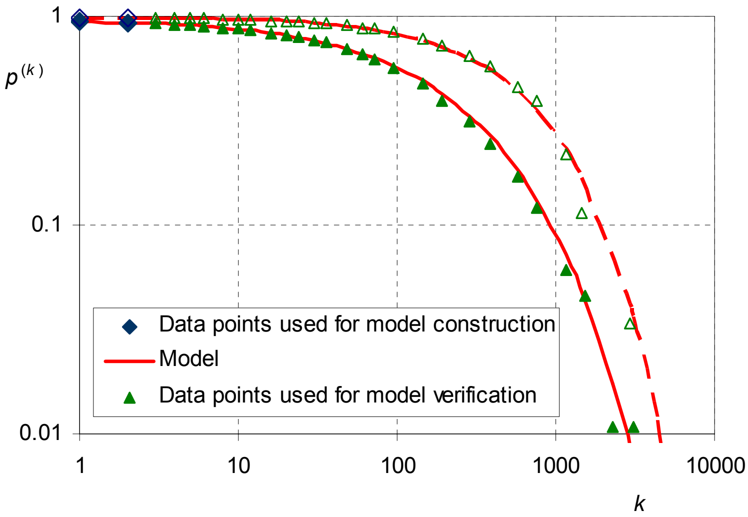
Monthly Nile flows (Koutsoyiannis *et al.*, 2008).



Iowa high temporal resolution rainfall (Koutsoyiannis *et al.*, 2007b).

Application 4: ME and clustering of rainfall occurrence

- How many parameters do we need to express the (intermittent) rainfall occurrence process at all scales, including its time dependence and its clustering behaviour (or overdispersion)?
- It seems that two parameters may suffice.
- The solution is given by maximizing, for a range of scales, the entropy of the binary-state rainfall process maximized with two constraints representing the observed probabilities dry at two specified time scales.
- The solution is not Markovian neither scaling but in between.



Probability $p^{(k)}$ that an interval of k hours is dry, as estimated from the Athens rainfall data set and predicted by the model of maximum entropy for the entire year (full triangles and full line) and the dry season (empty triangles and dashed line); see details in Koutsoyiannis (2006).

Application 5: Parameter parsimonious stochastic generation schemes

- ARMA-type models become parameter-excessive when they try to reproduce an autocorrelation function with a long tail.
- The problem is that an ARMA-type model is simultaneously a representation of an autocovariance function (γ_j for lag j) and a stochastic generating scheme for the process $\underline{x}_i$. The ARMA(p, q) process is:

$$\underline{x}_i = \sum_{j=1}^{p} a_j\, \underline{x}_{i-j} + \sum_{j=0}^{q} b_j\, \underline{v}_{i-j}$$

where a_j and b_j groups of parameters estimated from data, and $\underline{v}_i$ is white noise.
- This is a bad idea and the two could be separated. The autocovariance function γ_j could be defined using, say, one (as in the SSS case) or two parameters estimated from the data (depending on the size of the data set and the prior information).
- Then the generating scheme can be independent from the data, e.g. a simple symmetric moving average scheme:

$$\underline{x}_i = \sum_{j=-s}^{s} b_{|j|}\, \underline{v}_{i+j}$$

where b_j are now internal algorithmic coefficients estimated by

$$s_b(\omega) = \sqrt{2 s_\gamma(\omega)}$$

whereas $s_\gamma(\omega)$ and $s_b(\omega)$ are the inverse finite Fourier transforms of γ_j and b_j, respectively.

See details in Koutsoyiannis (2000).

Application 6: Parsimonious multivariate stochastic modelling

- Multivariate stochastic modelling is needed in multi-site simulation, in simulation of single-site cyclostationary processes, in prediction problems with different types of predictors, etc.
- For example, we consider the prediction $\underline{w}$ of the monthly flow one month ahead, conditional on a number s of other variables with known values that compose the vector $\underline{\boldsymbol{z}}$, using the linear model:

$$\underline{w} = \boldsymbol{a}^T \underline{\boldsymbol{z}} + \underline{v}$$

  where $\boldsymbol{a}$ is a vector of parameters (the superscript T denotes the transpose of a vector or matrix) and $\underline{v}$ is the prediction error, assumed independent of $\underline{\boldsymbol{z}}$; for simplicity, all elements of $\underline{\boldsymbol{z}}$ are assumed normalized and with zero mean and unit variance.
- For the model to take account of both long-range and short-range dependence, a possible composition of $\underline{\boldsymbol{z}}$ may include the following:
  - The flows of a few previous months of the same year.
  - All available flow measurements of the same month on previous years.
- The model parameters are estimated from (Koutsoyiannis, 2000):

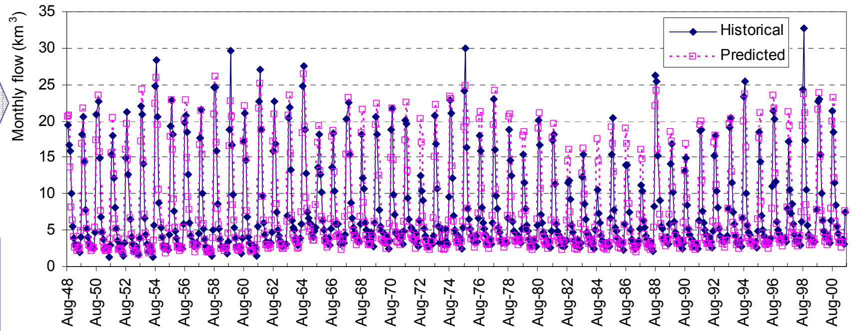
$$\boldsymbol{a}^T = \boldsymbol{\eta}^T \boldsymbol{h}^{-1}, \quad \mathrm{Var}[\underline{v}] = 1 - \boldsymbol{\eta}^T \boldsymbol{h}^{-1} \boldsymbol{\eta} = 1 - \mathbf{a}^T \boldsymbol{\eta}$$

  where $\boldsymbol{\eta} := \mathrm{Cov}[\underline{w}, \underline{\boldsymbol{z}}]$ and $\boldsymbol{h} := \mathrm{Cov}[\underline{\boldsymbol{z}}, \underline{\boldsymbol{z}}]$.

Application 6 (cont'd)

- Both the vector $\boldsymbol{\eta} := \mathrm{Cov}[\underline{w}, \underline{\boldsymbol{z}}]$ and the matrix $\boldsymbol{h} := \mathrm{Cov}[\underline{\boldsymbol{z}}, \underline{\boldsymbol{z}}]$ may contain numerous items, typically of the order of 10^3-10^4 (e.g. for a dimensionality 100, if we have 100 years of observations: 100 + 100 × 100 = 10 100 items—but reduced due to symmetry).
- Traditionally, the items of such covariance matrices and vectors have been estimated directly from data; this is totally illogical (100 years of data cannot support the statistical estimation of 1000-10 000 parameters).
- An alternative approach is to use data to estimate a couple of parameters per month and derive all other 'unestimated' parameters by maximizing entropy.
- Such entropy maximization may in fact be very simple (suggestive of a generalized Cholesky matrix decomposition).

Example: One month ahead predictions of Nile flow in comparison to historical values for the validation period (Efficiency = 91%).



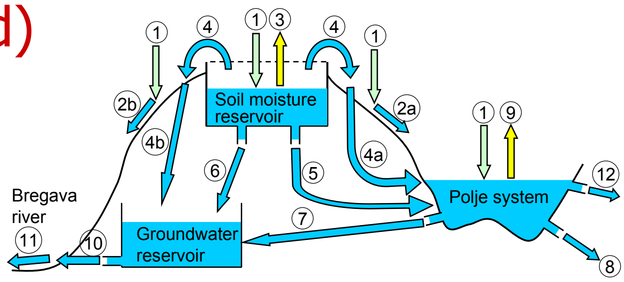
See details in Koutsoyiannis *et al.* (2008).

Application 7: Lumped vs. detailed hydrological modelling

- A simple lumped representation of a complex hydrological system may be more skilful than detailed "physically-based" representations.
- As a recent example, we consider the hydrological modelling of a karst system in Eastern Herzegovina, which involves several ground and surface transport and storage elements (e.g. poljes).
- Three different approaches were developed by three independent modelling teams:
  - A detailed quasi-physically-based model (3DNet), performing full dynamic flow simulation in a network of tunnels and reservoirs and in the unsaturated zone.
  - A detailed transfer-function-based model (Černi) that represents the dynamics of the transformation of precipitation into the karstic inflow, surface water flow and flow through porous medium using transfer functions.
  - A simplistic lumped conceptual model (K-Sim^2), with caricature dynamics.

Application 7 (cont'd)

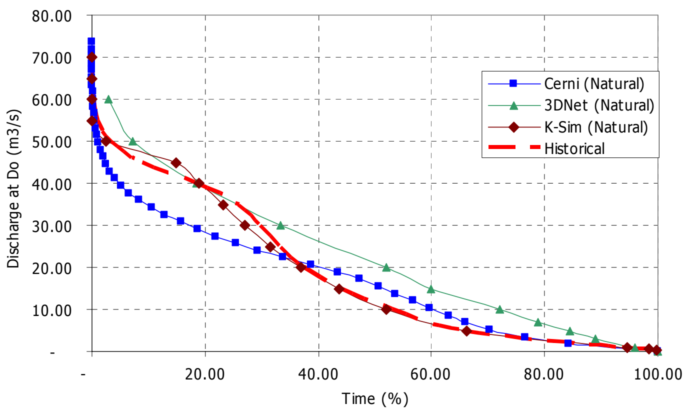
Schematic of the simplistic K-Sim^2 model. Three reservoirs represent the different storage types, surface (poljes), soil and groundwater; these are connected through transport paths.



1. Rainfall
2. Direct flow
3. Evapotranspiration
4. Quick flow

5-8. Groundwater fluxes

9. Lake evaporation
10. Spring discharge
11. River flow
12. Diversion tunnel flow



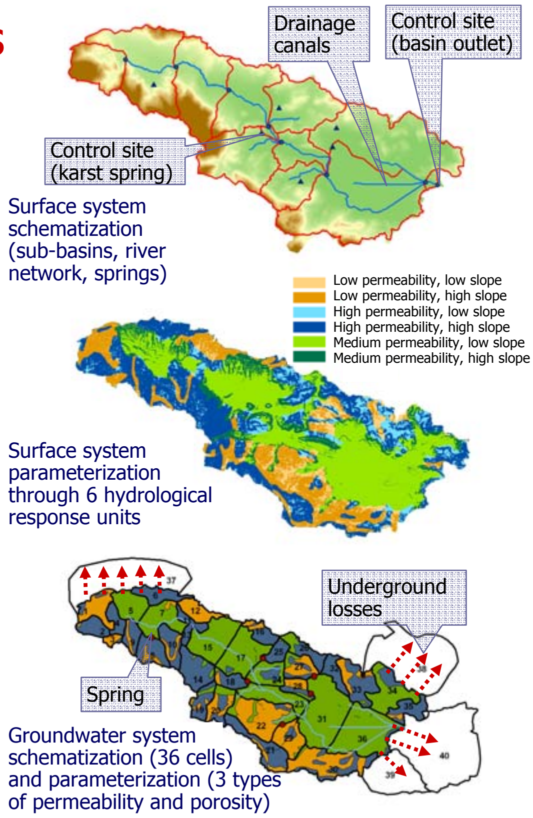
The simplest K-Sim^2 model gave substantially better proximity with reality in terms of fitting flood hydrographs and the historical flow frequency curve.

See details in Makropoulos *et al.* (2008).

Application 8: Parsimonious semi-distributed catchment modelling

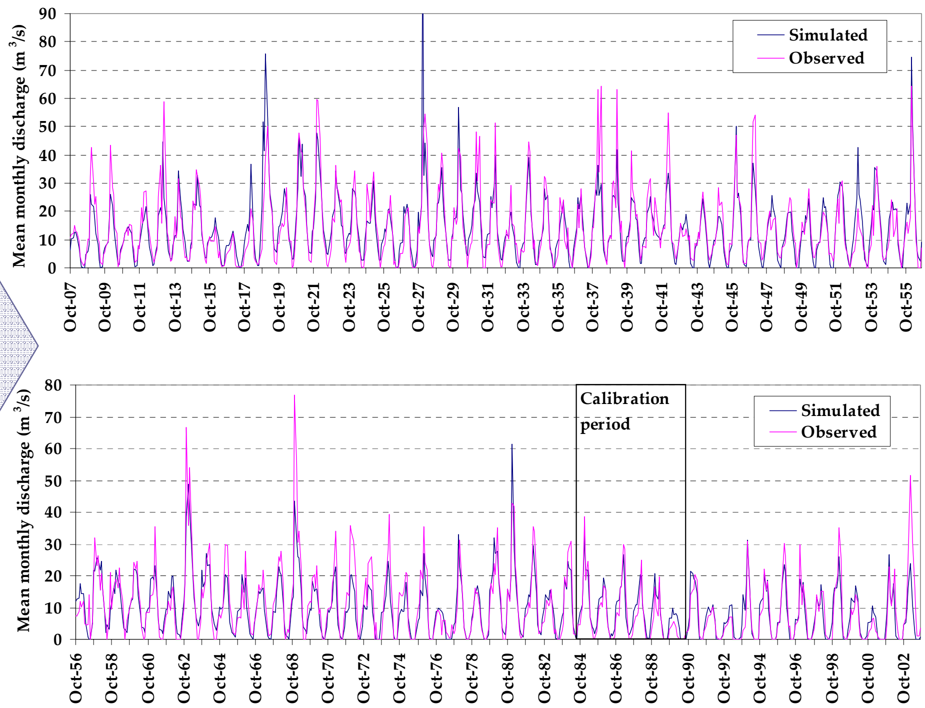
- Semi-distributed modelling may be a good parsimonious choice when the problem involves multiple processes or multiple control points.
- As a recent example, we consider the hydrological modelling of a peculiar modified catchment in Greece (Boeoticos Kephisos river basin, ≈2000 km^2) which lies mostly on a karst subsurface and has no outlet to the sea.
- The problem demands modelling of the interaction of surface- and ground-water processes, plus human intervention.
- Basic choices for parsimony:
  - Separation of surface and sub-surface discretization;
  - Separation of schematization (discretization) and parameterization.

See details in Efstratiadis *et al.* (2009);
A poster paper at this session this afternoon.



Application 7 (cont'd)

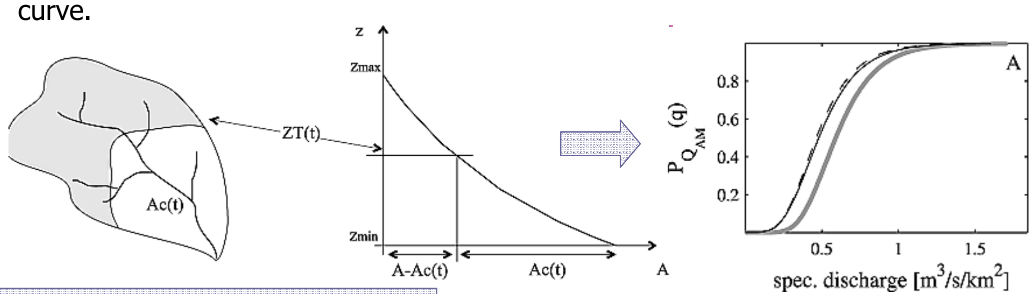
The approach seems to have good prediction skill, as indicated from almost a century long validation.



From an earlier application in Koutsoyiannis *et al.* (2007a).

Application 9: Using probabilistic representations in catchment modelling

- Several hydrological models effectively use parsimonious probabilistic representations of the geographically distributed system characteristics or state variables—instead of deterministic descriptions inflationary in detail.
- A nice recent example is an analytical model developed to determine the effects of catchment elevation on the flood frequency distribution.
- In this, the probability distribution of the annual maximum discharge is analytically derived, based on simple assumptions on the stochastic process of precipitation. The shape and the moments of this distribution explicitly relate to basin hypsometric curve (i.e. frequency distribution of elevation) and to the seasonality of temperatures.
- The model can explain the attenuation of flood quantiles in high-elevation basins by describing the freezing level elevation in the catchment, based on the hypsometric curve.



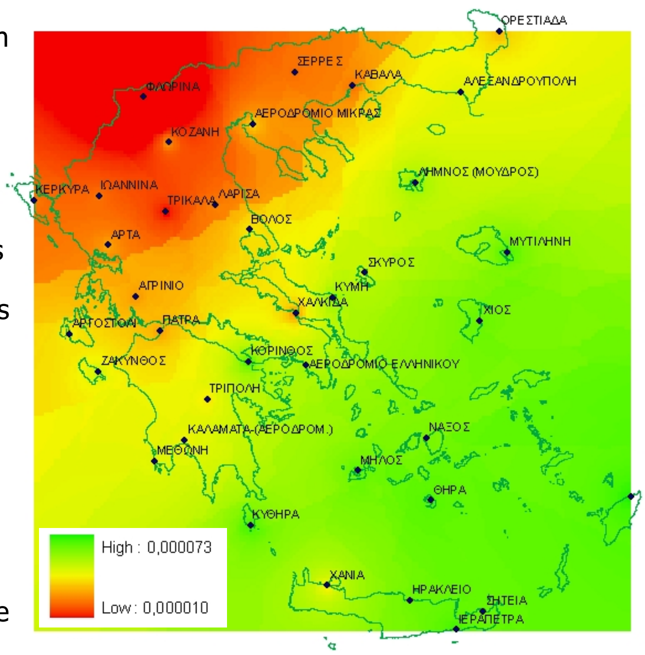
See details in Allamano *et al.* (2009).

Application 10: Regionalization

- Regionalization techniques have been very effective in summarizing geographically distributed processes and mapping their parameters, and hence in the inference in ungauged basins.
- In a recent example, a simplification of the Penman-Montieth method was developed for estimating potential evapotranspiration at a monthly basis from incomplete data.
- The simplified equation obtained is

$$\mathrm{PE} = \frac{aS_0}{1 - cT_a}$$

where PE is the potential evapotranspiration in kg m^{-2} d^{-1} (or mm/d), S_0 is the extraterrestrial shortwave radiation (calculated), T_a the air temperature (the only variable needed to measure), c = 0.0234°C^{-1} and a is a parameter varying geographically as in the map (in kg/kJ).



See details in Tegos *et al.* (2009);
A poster paper at this session this afternoon.

Application 11: From nonsense deterministic to parsimonious stochastic approaches in engineering

Problem: Given a reservoir with storage capacity k and inflows i_t for time (year) $t = 1, 2, \ldots, n$, find the release d that can be achieved on a year-to-year steady state basis.

- Classical **deterministic** formulation:

  maximize d

  s.t. $s_t = s_{t-1} + i_t - d - w_t$

  $s_t \le k, \quad s_t, w_t, d \ge 0, \quad s_n \ge s_0$

  where s_t and w_t are, respectively, the reservoir storage and spill at time t.
- The equality constraint represents the water balance in the reservoir whereas the non-equality constraints represent physical or methodological restrictions.
- This is a typical linear programming problem with only one actual control variable (the steady state release d), but with $2n$ additional control variables (s_t, w_t) and $2n + 1$ constraints (not including the non-negativity constraints).
- For example, in a simulation with n = 1000, the problem includes 2001 control variables and 2001 constraints.

- Alternative **stochastic** formulation:

  maximize $L(d) = r_{(\beta n)}$

  where β is an acceptable probability of failure and $r_{(m)}$ denotes the mth smallest value of the series of releases r_t (e.g., for n = 1000 and β = 1%, $r_{(\beta n)} \equiv r_{(10)}$ is the tenth smallest value), determined by:

  $r_t = \min(d, s_{t-1} + i_t)$

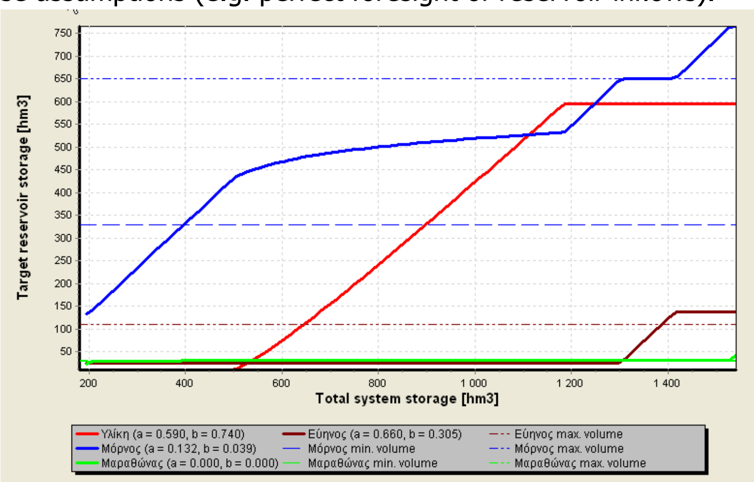
  $w_t = \max(0, s_{t-1} + i_t - d - k)$

  $s_t = s_{t-1} + i_t - r_t - w_t$
- This is a nonlinear problem with **one control** variable only and no constraints.
- The formulation is very **parsimonious**, and the solution is **reliability-based** and does not depend on the sequence of **unknown inflows** but only on their statistical characteristics.

See details in Koutsoyiannis and Economou (2003).

Application 12: Extension to the management of multi-reservoir systems

- This example refers to the strategic management of the Athens water supply system comprising 4 reservoirs, assuming a 10-year control horizon and trying to maximize reliability and minimize cost.
- A classical deterministic approach would involve 4 × 12 × 10 = 480 control variables and would rely on nonsense assumptions (e.g. perfect foresight of reservoir inflows).
- The adopted stochastic approach expresses the management in terms of parameterized rules referring to an optimal distribution of total system storage to target storages in each reservoir.
- This approach assigns two control variables per reservoir (a total of 8) and its results can be easily implemented and monitored in real time.



See details in Koutsoyiannis *et al.* (2003) and Mamassis *et al.* (2009).

Conclusions

- Nature seems to be naturally parsimonious.
- It is then natural to try to build parsimonious models for natural processes.
- Simple systems can be parsimoniously modelled by deterministic approaches.
- In complex systems parsimony should necessarily be combined with stochastic approaches.
- Recently mainstream research invested hopes in detailed approaches by building complicated models.
- However, comparisons of complicated models with parsimonious ones indicate that the latter:
  - can facilitate insight and comprehension;
  - improve accuracy, efficiency and predictive capacity; and
  - require fewer data to achieve the same accuracy with the former.
- In water engineering and management, parsimonious formulations and solutions of problems are more reasonable and rational, and easier to apply and monitor in practice.

References

- Allamano, P., P. Claps, and F. Laio, An analytical model of the effects of catchment elevation on the flood frequency distribution, *Water Resour. Res.*, 45, W01402, doi:10.1029/2007WR006658, 2009.
- Beven, K.J., Prophecy, reality and uncertainty in distributed hydrological modelling, *Adv. Water Resour.*, 16(1), 41–51, 1993.
- Boltzmann, L., Über die Mechanische Bedeutung des Zweiten Hauptsatzes der Wärmetheorie. *Wiener Berichte*, 53, 195–220, 1866.
- Clausius, R., Mechanical Theory of Heat – with its Applications to the Steam Engine and to Physical Properties of Bodies. London: John van Voorst, 1850-1865.
- Courtney, A., and M. Courtney, Comments Regarding "On the Nature of Science", *Physics in Canada*, 64(3), 7-8, 2008.
- Dooge, J.C.I., Searching For Simplicity In Hydrology, Surveys in Geophysics, 18(5), 511-534, 1997.
- Efstratiadis, A., A. Mazi, A. D. Koussis, and D. Koutsoyiannis, Flood modelling in complex hydrologic systems with sparsely resolved data, *European Geosciences Union General Assembly 2009, Geophysical Research Abstracts, Vol. 11*, Vienna, 4157, European Geosciences Union, 2009.
- Gaertner, H.-M., Huygens' Principle: A case against optimality, *Behavioral and Brain Sciences*, 26, 779–781, 2003.
- Gauch, H.G., Jr., *Scientific Method in Practice*, Cambridge University Press, Cambridge, 2003.
- Gibbs, J.W., *Elementary Principles in Statistical Mechanics*, Yale University Press, New Haven, Connecticut, 1902.
- Jaynes, E. T., Information theory and statistical mechanics. *Phys. Rev.*, 106(4), 620–630, 1957.
- Kolmogorov, A.N., The theory of transmission of information, *Session of the USSR Academy of Sciences on Scientific Problems of Automating Productio*n, Plenary Meetings 66-99. Izd. Akad. Nauk SSSR, Moscow (In Russian), 1956.
- Kolmogorov, A.N., A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces, *Dokl. Akad. Nauk SSSR*, 119, 861-864 (in Russian), 1958.
- 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., Uncertainty, entropy, scaling and hydrological stochastics, 1, Marginal distributional properties of hydrological processes and state scaling, *Hydrological Sciences Journal*, 50 (3), 381–404, 2005a.
- Koutsoyiannis, D., Uncertainty, entropy, scaling and hydrological stochastics, 2, Time dependence of hydrological processes and time scaling, *Hydrological Sciences Journal*, 50 (3), 405–426, 2005b.
- Koutsoyiannis, D., An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence, *Water Resources Research*, 42 (1), W01401, doi:10.1029/2005WR004175, 2006.
- Koutsoyiannis, D., and A. Economou, Evaluation of the parameterization-simulation-optimization approach for the control of reservoir systems, *Water Resources Research*, 39 (6), 1170, doi:10.1029/2003WR002148, 2003.
- Koutsoyiannis, D., G. Karavokiros, A. Efstratiadis, N. Mamassis, A. Koukouvinos, and A. Christofides, A decision support system for the management of the water resource system of Athens, *Physics and Chemistry of the Earth*, 28 (14-15), 599–609, 2003.
- Koutsoyiannis, D., A. Efstratiadis, and K. Georgakakos, Uncertainty assessment of future hydroclimatic predictions: A comparison of probabilistic and scenario-based approaches, *Journal of Hydrometeorology*, 8 (3), 261–281, 2007a.
- Koutsoyiannis, D., S.M. Papalexiou, and A. Montanari, Can a simple stochastic model generate a plethora of rainfall patterns?, *The Ultimate Rainmap: Rainmap Achievements and the Future in Broad-Scale Rain Modelling*, Oxford, Engineering and Physical Sciences Research Council, 2007b.
- 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.
- Lessovaia, S. N., Dmitricheva, L. E., Goryachkin, S. V., Vodyanitsky Y. N., Gagarina, E. I., Semikolennykh, A. A., Shelemina, A. N., and Tuyukina, T. Y.: Soil-ecology conditions of unique taiga landscapes, Institute of Ecology Soil Science MSU, http://soilinst.msu.ru/pub 01/text paper.htm, last access: August 2008.
- Makropoulos, C., D. Koutsoyiannis, M. Stanic, S. Djordevic, D. Prodanovic, T. Dasic, S. Prohaska, C. Maksimovic, and H. S. Wheater, A multi-model approach to the simulation of large scale karst flows, *Journal of Hydrology*, 348 (3-4), 412–424, 2008.
- Mamassis, N., A. Efstratiadis, G. Karavokiros, D. Nikolopoulos, Master plan of the Athens water resource system - Year 2008-2009, *Maintenance, upgrading and extension of the Decision Support System for the management of the Athens water resource system*, Contractor: Department of Water Resources and Environmental Engineering – National Technical University of Athens, Report, Athens, 2009.
- Schoemaker, P.J.H., Huygens versus Fermat: No clear winner, *Behavioral and Brain Sciences*, 26, 781–783, 2003.
- Shannon, C.E., A mathematical theory of communication, *Bell Systems Tech. J.*, 27(379), 623-656, 1948.
- Tegos, A., N. Mamassis, and D. Koutsoyiannis, Estimation of potential evapotranspiration with minimal data dependence, *European Geosciences Union General Assembly 2009, Geophysical Research Abstracts, Vol. 11*, Vienna, 1937, European Geosciences Union, 2009.