A random walk on water

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/896/
A walk on water is safer on water’s solid phase...

... although somewhat cold ...

... but my family offers me a very warm solid ground.
Acheloos River: a watery origin

Acheloos @ Mesounta (my village)
ITIA.NTUA.GR: A water loving plant with deep roots

ITIA = Greek for willow tree.
NTUA = National Technical University of Athens
GR = Greece
The Henry Darcy Medal... and Tyche (Τύχη)

My personal feelings (keywords)  Assessment by friends

happiness  ευτυχία  misery  δυστυχία
success  επιτυχία  failure  αποτυχία
blessing  ευτύχημα  accident  ατύχημα
luck  τύχη  crash  δυστύχημα

Balanced approach

incident  τυχαίο συμβάν
any one  τυχών
at random  στην τύχη
by chance  κατά τύχη

“The limits of my language mean the limits of my world”.

(Ludwig Wittgenstein, 1921, §5.6)
Randomness & scientific discovery

Isaac Newton hit on the head by the famous apple, which causes the discovery of the law of universal gravitation.

Comic strip by Marcel Gotlib, Rubrique-à-Brac (1970-1974)
What is randomness? (Common reply)

- The common reply is based on a dichotomous logic (+ reductionism):
  - on ontological grounds: there exist two mutually exclusive types of events or processes—deterministic and random (or stochastic);
  - on epistemological grounds: we separate the events into these two types—the random we do not understand nor explain.

- Extended dichotomization: natural process are composed of two different, usually additive, parts or components—deterministic and random;
  - each part may be further subdivided into subparts (e.g., deterministic part = periodic + aperiodic/trend).

- The dichotomous logic is typically combined with a manichean perception:
  - the deterministic part supposedly represents cause-effect relationships and, thus, is physics and science (the “good”);
  - the random part is noise, and has little relationship with science and no relationship with understanding (the “evil”).

Naïve and incorrect view of randomness—false dichotomy.
Does “random” mean “noise”? Does Nature produce noises?

A rainfall event in Iowa measured with a resolution of 10 seconds, transformed to sound. Credit: Simon Papalexiou
What is randomness? (Alternative reply)

- Randomness = unpredictability (in deterministic terms): we may understand, we may explain, but we cannot predict.
- Randomness and determinism:
  - coexist in the same process;
  - are not separable or additive components; and
  - it is a matter of specifying the time horizon of prediction to decide which of the two dominates.
- Unpredictability (in deterministic terms) = high uncertainty for the future.
- Uncertainty is quantified by Probability:
  - Andrey Kolmogorov (1933) system: Probability is a normalized measure, i.e., a function that maps sets (areas where unknown quantities lie) to real numbers (in the interval \([0, 1]\)).
  - Random variable: a variable associated with a probability distribution (or density) function.
  - Probabilization of uncertainty: axiomatic reduction from the notion of unknown to the notion of a random variable (typical in Bayesian statistics).

"Prediction is difficult, especially of the future" (Niels Bohr).
Historical references

- **Pierre Simon Laplace:**
  - perhaps the most famous proponent of determinism in the history of philosophy of science (cf. Laplace’s demon);
  - at the same time, one of the founders of probability theory:
  - “la théorie des probabilités n'est, au fond, que le bon sens réduit au calcul”
    - “Probability theory is, au fond, nothing but common sense reduced to calculus” (Laplace, 1812).

- **James Clerk Maxwell:**
  - “the true logic for this world is the calculus of Probabilities” (Maxwell, 1850, in a letter to Lewis Campbell).

- **Edwin Thompson Jaynes’s** recent book
Emergence of randomness from determinism:
A toy model of a caricature hydrological system

- The toy model is designed intentionally simple.
- Only infiltration, transpiration and water storage are considered.
- The rates of infiltration $\varphi$ and potential transpiration $\tau_p$ are constant.

Nothing in the model is random.

Discrete time: $i$ ("years").
- Constants (per "year")
  - Input: $\varphi = 250$ mm;
  - Potential output: $\tau_p = 1000$ mm.
- State variables (a 2D dynamical system):
  - Vegetation cover, $\nu_i$ ($0 \leq \nu_i \leq 1$);
  - Soil water (no distinction from groundwater): $x_i$ ($-\infty \leq x_i \leq a = 750$ mm).
- Actual output: $\tau_i = \nu_i \tau_p$
- Water balance
  $$x_i = x_{i-1} + \varphi - \nu_{i-1} \tau_p$$
The toy model at equilibrium

- If at some time $i - 1$:
  \[ v_{i-1} = \frac{\phi}{\tau_p} = \frac{250}{1000} = 0.25 \]
  then the water balance results in
  \[ x_i = x_{i-1} + \phi - v_{i-1} \tau_p = x_{i-1} \]
- Continuity of system dynamics demands that for some $x_{i-1}$, $v_i = v_{i-1}$. Without loss of generality we set this value $x_{i-1} = 0$ (this defines a datum for soil water).

  \[ \phi = 250 \text{ mm: Infiltration} \]
  \[ \tau = 250 \text{ mm: Transpiration} \]
  \[ v = 0.25: \text{Vegetation cover} \]

  \[ x = 0: \text{Soil water at datum} \]

  Datum

- Thus the system state:
  \[ v_i = v_{i-1} = 0.25 \]
  \[ x_i = x_{i-1} = 0 \]
represents the equilibrium of the system.

- If the system arrives at equilibrium it will stay there for ever.
Non-equilibrium state – conceptual dynamics of vegetation

The graph is described by the following equation (with $\beta = 100$ mm — a standardizing constant):

$$ V_i = \frac{\max(1 + (s_{i-1} / \beta)^3, 1) V_{i-1}}{\max(1 - (s_{i-1} / \beta)^3, 1) + (s_{i-1} / \beta)^3 V_{i-1}} $$

- **Increased vegetation**
- **Unchanged vegetation**
- **Decreased vegetation**

High soil water (above datum)
Normal soil water (at datum)
Low soil water (below datum)
System dynamics

Water balance + Vegetation cover dynamics

\[ x_i = \min(x_{i-1} + \varphi - v_{i-1} \tau_p, \alpha) \]

for finite storage \( \leq \alpha \)

\[ v_i = \frac{\max(1 + (s_{i-1} / \beta)^3, 1) \cdot v_{i-1}}{\max(1 - (s_{i-1} / \beta)^3, 1) + (s_{i-1} / \beta)^3 \cdot v_{i-1}} \]

\( \varphi = 250 \text{ mm}, \ \tau_p = 1000 \text{ mm}, \ \alpha = 750 \text{ mm}, \ \beta = 100 \text{ mm}. \)

Easy to program in a hand calculator or a spreadsheet.
Interesting trajectories produced by simple deterministic dynamics

- These trajectories of $s$ and $v$, for time $i = 1$ to 100 were produced assuming initial conditions $x_0 = 100$ mm ($\neq 0$) and $v_0 = 0.30$ ($\neq 0.25$); they can be easily reproduced using a spreadsheet (or even a hand calculator).

- The system state does not converge to the equilibrium.

- The trajectories seem periodic.

- Iterative application of the simple dynamics allows prediction for arbitrarily long time horizons (e.g., $x_{100} = -244.55$ mm; $v_{100} = 0.7423$).
Understanding of mechanisms and system dynamics

- System understanding—causative relationships:
  - There is water balance (conservation of mass);
  - Excessive soil water causes increase of vegetation;
  - Deficient soil water causes decrease of vegetation;
  - Excessive vegetation causes decrease of soil water;
  - Deficient vegetation causes increase of soil water.

- System dynamics are:
  - Fully consistent with this understanding;
  - Very simple, fully deterministic;
  - Nonlinear, chaotic.
Science vs. understanding

- Science < Latin Scientia < translation of Greek Episteme (Επιστήμη) < Epistasthai (Επιστάσθαι) = to know how to do < [epi (ἐπί) = over] + [histasthai (Ἢστασθαι) = to stand] = to overstand.

- Understanding is not identical, nor a prerequisite, to overstanding.

“I think I can safely say that nobody understands quantum mechanics”.

Richard Feynman (1965)

Credit for sketches: Demetris Jr.
Does deterministic dynamics allow a reliable prediction at an arbitrarily long time horizon?

- Axiomatic premise: A continuous (real) variable that varies in time cannot be ever known with full (infinite) precision.
- It is reasonable then to assume that there is some small uncertainty in the initial conditions (initial values of state variables).
- Sensitivity analysis allows to see that a tiny uncertainty in initial conditions may get amplified.

Bold blue line corresponds to initial conditions $s_0 = 100$ mm, $v_0 = 0.30$.
All other lines represent initial conditions slightly (< 1%) different.

Short time horizons: good predictions.
Long time horizons: extremely inaccurate and useless predictions.
From determinism to stochastics

- Probabilization of uncertainty: axiomatic reduction from the notion of an uncertain quantity to the notion of a random variable.
- Any value \( x_i \) is a realization of a random variable \( x_i \) and is associated with a probability density function \( f(x_i) \).
- **Stochastics** (modern meaning): probability + statistics + stochastic processes.
- **Stochastics** (first use and definition) is the *Science of Prediction*, i.e., the science of measuring as exactly as possible the probabilities of events (Jacob Bernoulli, 1713—Ars Conjectandi, written 1684-1689).
- **Stochastics** (etymology): < Greek Stochastikos (Στοχαστικός) < Stochazesthai (Στοχάζεσθαι) < Stochos (Στόχος)
  - Stochos = target
  - Stochazesthai = (1) to aim, point, or shoot (an arrow) at a target; (2) to guess or conjecture (the target) (3) to imagine, think deeply, bethink, contemplate, cogitate, meditate.

If one 'stochazetai' (thinks deeply), eventually he goes 'stochastic' (with the probability-theoretical meaning) and he will hit 'stochos' (the target).

**Stochastics** does not necessarily mean ARMA models.
The stochastic formulation of system evolution

- We fully utilize the deterministic dynamics: \( x_i = S(x_{i-1}) \), where \( x_i := (x_i, \nu_i) \) is the vector of the system state and \( S \) is the vector function representing the known deterministic dynamics of the system.

- We assume that \( f(x_0) \) is known, e.g. a uniform distribution extending 1\% around the value \( x_0 = (100 \text{ mm}, 0.30) \).

- Given the probability density function at time \( i - 1, f(x_{i-1}) \), that of next time \( i, f(x_i) \), is given by the **Frobenius-Perron operator** \( FP \), i.e. \( f(x_i) = FP f(x_{i-1}) \), uniquely defined by an integral equation (e.g. Lasota and Mackey, 1991), which in our case takes the following form, where \( A := \{ x, x \leq (x, \nu) \} \) and \( S^{-1}(A) \) is the **counterimage** of \( A \):

\[
FPf(x) = \frac{\partial^2}{\partial x \partial \nu} \int_{S^{-1}(A)} f(u) \, du
\]

- Iterative application of the equation can determine the density \( f(x_i) \) for any time \( i \) — but we may need to calculate a high-dimensional integral.

Stochastics does not disregard the deterministic dynamics: it is included in the counterimage \( S^{-1}(A) \).
Difficulties in applying the stochastic framework and their overcoming using stochastic tools

- The stochastic representation has potentially an analytical solution that behaves like a deterministic solution, but refers to the evolution in time of admissible sets and densities, rather than to trajectories of points.
  - From $x_i = S(x_{i-1})$ to $f_i(x) = \frac{\partial^2}{\partial x \partial v} \int S^{-1}(A) f_{i-1}(u) \, du$

- In the iterative application of the stochastic description of system evolution we encounter two difficulties:
  - Despite being simple, the dynamics is not invertible and the counterimage $S^{-1}(A)$ needs to be evaluated numerically → numerical integration.
  - The stochastic formulation is more meaningful for long time horizons → high dimensional numerical integration.

- For a number of dimensions $d > 4$, a stochastic (Monte Carlo) integration method (evaluation points taken at random) is more accurate than classical numerical integration, based on a grid representation of the integration space (e.g., Metropolis and Ulam, 1949; Niederreiter, 1992).
  - In our case the Monte Carlo method bypasses the calculation of $S^{-1}(A)$.

Monte Carlo integration is very powerful, yet so easy that we may elude that we are doing numerical integration.
Results of Monte Carlo integration: Time 100

- We assume $f(x_0)$ to be a uniform density extending 1% around the value $x_0 = (100 \text{ mm}, 0.30)$.
- From 1000 simulations we are able to numerically evaluate $f(x_{100})$.
- The figure shows the density of the soil water, $f(x)$.
- Moving from time $i = 0$ to $i = 100$, the density changes:
  - from concentrated to broad;
  - from uniform to Gaussian.

The figure shows the density of the soil water, $x$. Moving from time $i = 0$ to $i = 100$, the density changes:
- from concentrated to broad;
- from uniform to Gaussian.
Why is the distribution of soil water, after a long time, Gaussian?

- There are a number of theoretical reasons resulting in Gaussian distribution; see Jaynes (2003).
- Among them the most widely known is the Central Limit Theorem, which does not apply here (there are no sums of variables).
- Here applies the Principle of Maximum Entropy: for fixed mean and variance the distribution that maximizes entropy is the normal distribution (or the truncated normal, if the domain of the variable is an interval in the real line).
- Entropy [\(< \text{Greek } \varepsilonντροπία < \text{entrepomai (εντρέπομαι)} = \text{to turn into}\) is a probabilistic concept, which for a continuous random variable \(x\) is defined as
  \[
  \varphi[x] := E[-\ln f(x)] = -\int_{-\infty}^{\infty} f(x) \ln f(x) \, dx
  \]
- Entropy is a typical measure of uncertainty, so its maximization indicates that the uncertainty spontaneously becomes as high as possible (this is the basis of the Second Law of thermodynamics).
The propagation of uncertainty is completely determined using stochastics.

In summary, the stochastic representation:
- incorporates the deterministic dynamics—yet describes uncertainty;
- has a rigorous analytical expression (Frobenius-Perron);
- provides and utilizes a powerful numerical integration method (Monte Carlo).

The so-called ensemble forecasting in weather and flood prediction does not differ from this stochastic framework.
Do we really need the deterministic dynamics to make a long-term prediction?

- Working hypothesis: A set of observations contains enough information, which for long horizons renders knowledge of dynamics unnecessary.
- Here we use 100 “years” of “past observations”, for times $i = -100$ to -1.
- Initial conditions: $x_{-100} = 73.99$ and $v_{-100} = 0.904$.
- At time $i = 0$, the resulting state is $x_0 = 99.5034 \approx 100$; $v_0 = 0.3019 \approx 0.30$.
- Interpreting “observations” as a statistical sample, we estimate: mean = -2.52; standard deviation = 209.13.

In further investigations, we will refer to the state $x_0 = 99.5034; v_0 = 0.3019$ as the exact initial state and $x_0 = 100; v_0 = 0.30$ as the rounded off initial state.
A naïve statistical prediction vs. deterministic prediction

- We compare two different predictions:
  - That derived by immediate application of the system dynamics;
  - A naïve prediction: the future equals the average of past data.

- For long prediction times, the naïve prediction is more skilful.

- Its error $e_i$ is smaller than that of deterministic prediction by a factor of $\sqrt{2}$.

- This result is obtained both by Monte Carlo simulation and by probability-theoretic reasoning (assuming independence among different trajectories).

For long horizons use of deterministic dynamics gives misleading results. Unless a stochastic framework is used, neglecting deterministic dynamics is preferable.
Past data and ergodicity

- **Ergodicity** (εργοδικός < [ἔργον = work] + [οδός = path]) is an important concept in dynamical systems and stochastics.
- By definition (e.g. Lasota and Mackey, 1994, p. 59), a transformation is ergodic if all its invariant sets are trivial (have zero probability [= measure]).
- In other words, in an ergodic transformation starting from any point, a trajectory will visit all other points, without being trapped to a certain subset. (In contrast, in non ergodic transformations there are invariant subsets, such that a trajectory starting from within a subset will never depart from it).
- An important theorem by George David Birkhoff (1931) says that for an ergodic transformation $S$ and for any integrable function $g$ the following property holds true:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} g(S^i(x)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$

- For instance, for $g(x) = x$, setting $x_0$ the initial system state, observing that the sequence $x_0$, $x_1 = S(x_0)$, $x_2 = S^2(x_0)$, ..., $n$ represents a trajectory of the system and taking the equality in the limit as in approximation with finite terms, we obtain that the time average equals the true (ensemble) average:

$$\frac{1}{n} \sum_{i=0}^{n-1} x_i \approx \int_{-\infty}^{\infty} x f(x)dx$$

Ergodicity allows estimation of the system properties using past data only.
A more informative prediction

- Reduction of uncertainty for long time horizons: **No way!**
  - No margin for better knowledge of dynamics (full knowledge already).
  - Indifference of improved knowledge of initial conditions (e.g. reduction of initial uncertainty from 1% to $10^{-6}$ results in no reduction of final uncertainty at $i = 100$ (try it!).

- Informative prediction = point prediction + quantified uncertainty.
  - Past data: temporal mean & variance at times $i = -100$ to 0.
  - Ergodicity: ensemble mean & variance at time $i = 100$.
  - Principle of maximum entropy: Gaussian distribution.

Stochastic inference using (a) past data, (b) ergodicity, and (c) maximum entropy provides an informative prediction. Knowledge of dynamics does not improve this prediction.
Stochastics for ever...

The stochastic representation is good for both short and long horizons, and helps figure out when the deterministic dynamics should be considered or neglected.
Further exploration of the system properties: Is the system evolution periodic?

- A longer simulation of the system (10 000 terms) using the rounded-off initial conditions shows that the period $\delta$ between consecutive peaks is **not constant** but varies between 4 and 10 “years”.
- The period with maximum frequency $\nu$ is 6 “years”.

The trajectories of the system state do not resemble a typical periodic deterministic system—nor a purely random process.
A stochastic tool to detect periodicity: Periodogram

- The square absolute value of the Discrete Fourier Transform (a real function $p(\omega)$ where $\omega$ is frequency) of the time series (here 10,000 terms) is the periodogram of the time series.
- $p(\omega) \, d\omega$ is the fraction of variance explained by $\omega$ and thus excessive values of $p(\omega)$ indicate strong cycles with period $1/\omega$.
- Here we have large $p(\omega)$ at $1/\omega$ between 4 and 12 “years” without a clearly dominant frequency.
- Rather the shape indicates a combination of persistence (short periods) and antipersistence (long periods).

Antipersistence is often confused with periodicity—however, the two are different.
A stochastic tool to detect periodicity and dependence: Autocorrelation function

- The Finite Fourier Transform of the periodogram is the empirical autocorrelation function (autocorrelogram), which is a sequence of values $\rho_j$, where $j$ is a lag. It is more easily determined as $\rho_j = \frac{\text{Cov}[x_i, x_{i-j}]}{\text{Var}[x_i]}$.

- The positive $\rho_1$ is expected because of physical consistency.

- The existence of negative values is an indication of antipersistence.

![Graph showing autocorrelation function]
A different perspective of long-term predictability and the key consequence of antipersistence

- Arguably, when we are interested for a prediction for a long time horizon, we do not demand to know the exact value at a specified time but an average behaviour around that time (the “climate” rather than the “weather”).
- The plot of the soil water for a long period (1000 “years”) indicates:
  - High variability at a short (annual) scale—with peculiar variation patterns;
  - A flat time average at a 30-year scale (“climate”).

Antipersistence enhances **climatic-type predictability** (prediction of average).

\[ x_i \]

\[ \text{Annual storage} \quad \text{Moving average of 30 values} \]
An index of variability

To study the peculiar variability of the soil water $x_i$ we introduce the random variable $y_i := |x_i - x_{i-6}|$ where the lag 6 was chosen to be equal to the most frequent period appearing in the time series of $x_i$.

We call $y_i$ the **variability index**.

The plot of the time series of $y_i$ for a long period (1000 “years”) indicates:

- High variability at a short (“annual”) scale;
- Long excursions of the 30-“year” average (“the climate”) from the global average (of 10000 values).

The frequent and long excursions of the local average from the global average indicate **long-term persistence**. **Persistence** is often confused with **nonstationarity**—but the two are different.
The autocorrelation of the variability index

Models with all auto-correlations negative are not physically consistent.

Consistently positive autocorrelation
Persistence (not nonstationarity)

Alternating positive-negative autocorrelation
Antipersistence (not periodicity)

The consistently positive autocorrelations $\rho_j$ for high lags $j$ indicate long-term persistence.
Multi-scale stochastics and the Hurst-Kolmogorov dynamics

- A discrete-time random variable \( x_i \) refers to a specific time scale.
- A multi-scale stochastic representation defines a process at any scale \( k \geq 1 \) by:
  \[
  x_i^{(k)} := \frac{1}{k} \sum_{l=(i-1)k}^{ik} x_l
  \]
- A key multi-scale characteristic is the standard deviation \( \sigma^{(k)} \) of \( x_i^{(k)} \) which is a function of the scale \( k \), typically depicted on a log-log plot.
- The quantity \( H = 1 + \text{slope} \) in this plot is termed the Hurst coefficient.
- \( H = 0.5 \) indicates pure randomness.
- \( H \) between 0 and 0.5 indicates antipersistence.
- \( H \) between 0.5 and 1 indicates persistence.

A process with constant slope and \( H \) between 0.5 and 1 is a **Hurst-Kolmogorov** process (after **Hurst**, 1951, and **Kolmogorov**, 1940) with long term persistence.
For an one-step ahead prediction, a purely random process \( x_i \) is the most unpredictable.

Dependence enhances one-step ahead predictability; e.g. in a Markovian process with \( \rho_1 = 0.5 \) (comparable to that of our series \( x_i \) and \( y_i \)) the conditional standard deviation is \( \sqrt{1 - \rho_1^2} \) times the unconditional, i.e. by 13% smaller.

However, in the climatic-type predictions, where we are interested on the average behaviour rather than on exact values, the situation is different.

In the example shown, at the 30-“year” climatic scale, predictability is deteriorated by a factor of 3 for the persistent process \( y_i \) (thus eliminating the 13% reduction due to conditioning on the past).

Contrary to what is believed, positive dependence/persistence substantially deteriorates predictability over long time scales—but antipersistent improves it.
Demonstration of unpredictability of processes with persistence

- The plot shows 1000 “years” of the time series $y_i$ (variability index) at the annual and the climatic, 30-“year” scale and for initial conditions
  - exact, and
  - rounded.
- The departures in the two cases are evident.

Even a **fully deterministic system** is **fully unpredictable** at a long (climatic) time scale when there is **persistence**.
Recovery of dynamics from time series

- Stochastics—the concept of entropy in particular—provides a way to recover the dynamics of a system, if the dynamics is deterministic and unknown and if a long time series is available.
- Forming time delayed vectors with trial dimensions $m$ and calculating the multidimensional entropy of vector trajectories we are able to recover the unknown dynamics (employing Takens, 1981, theorem).
- In the example we find that the dimensionality of our toy system is 2.

Recovering of unknown deterministic dynamics does not enhance long-term predictability.
From the toy model to the real world

- In comparison to our simple toy model, a natural system (e.g., the atmosphere, a river basin, etc.):
  - is extremely more complex;
  - has time-varying inputs and outputs;
  - has spatial extent, variability and dependence (in addition to temporal);
  - has greater dimensionality (virtually infinite);
  - has dynamics that to a large extent is unknown and difficult or impossible to express deterministically; and
  - has parameters that are unknown.

- Hence uncertainty and unpredictability are naturally even more prominent in a natural system.

- The role of stochastics is even more crucial:
  - to infer dynamics (laws) from past data;
  - to formulate the system equations;
  - to estimate the involved parameters;
  - to test any hypothesis about the dynamics.

- Data offer the only solid grounds for all these tasks, and failure of founding on, and testing against, evidence from data renders the hypothesized dynamics worthless.
Some questions related to the real world:
(i) Physically-based modelling in hydrology

- What is *physically-based* modelling of hydrological (and other geophysical) systems?
  - Is *physics* synonym to *determinism*?
  - Is *physically-based* synonym to *mechanistic*?
  - Are *first principles* mechanistic principles?
  - Is not *statistical physics* part of *physics*?
  - Is not *entropy maximization* a first principle?
  - Is not *stochastic modelling* part of *physical modelling*?

- Will it ever be possible to achieve such a *physically-based* modelling of hydrological systems that will not depend on data or stochastic representations?
  - Can detailed representations and reduction to first principles render hydrologic measurements unnecessary?
  - What level of detail is needed in such reductionist modelling for a catchment of, say, 1000 km²?
    - $10^3$ cells of 1 km² each?
    - $10^9$ cells of 1 m² each?
    - $10^{15}$ cells of 1 mm² each?
  - How far can current research trend toward detailed “physically-based” models advance hydrology and water resources science and technology?
(ii) Hydrological uncertainty and its reduction

- To what extent can hydrological uncertainty be reduced?
  - Can uncertainty be eliminated by uncovering the system’s deterministic dynamics?
  - Is uncertainty epistemic or structural?
- When there is potential for reduction of uncertainty, what are the most effective means for reduction?
  - Better understanding?
  - Better deterministic models?
  - More detailed discretizations?
  - Better data?
- When the limits of uncertainty reduction have been reached, what are the appropriate scientific and engineering attitude?
  - Confession of failure—no action?
  - Quantification of uncertainty and risk through stochastics—action under risk?
- Is there potential to improve current stochastic methods in hydrology?
  - Are current methods consistent with observed natural behaviours?
(iii) Uncertainty and water resources engineering and management

- Can there be risk-free hydraulic engineering and water management?
- Can deterministic methods provide solid scientific grounds for water resources engineering and management?
- Are there deterministic upper limits in extreme hydrological phenomena, such as precipitation and flood, and can they be determined with certainty?
- Are the concepts of probable maximum precipitation (PMP) and probable maximum flood (PMF) scientific?
- Are the so-called hydrometeorological methods for determining PMP and PMF deterministic or statistical?
- Are there stochastic alternatives to PMP and PMF, able to quantify the risk?
- Do PMP and PMF remove risk by implying existence of upper deterministic limits?
- Do PMP and PMF remove responsibility for a decision as to the degree of acceptable risk or protection by implying that there is no risk?

“Ignorance is preferable to error and he is less remote from the truth who believes nothing than he who believes what is wrong”.

Thomas Jefferson (1781)
(iv) Hydrology, water resources and climate

- Is the current interface between hydrology and climate satisfactory?
- Should hydrology and water resources planning rely on climate model outputs?
- Are climate models properly validated (i.e., for periods and scales not used in calibration)?
- Is the evolution of climate and its impacts on water resources deterministically predictable?
- What is Climate and Climate Science?

**A remark for the definition:** “Gotta love Climate Science ... a scientific field with no agreed-upon subject of study.

- **What is the climate, Daddy?**
- **Son, if I knew that, I wouldn’t be a Climate Scientist.**


**A definition:** Climate [is] the long-term statistics describing the conditions in the atmosphere, ocean, and ice sheets and sea ice, such as means and extremes.

Long term climate predictions are trendy...

From 2100 AD (Battisti and Naylor, Science, 2009)...

... to 3000 AD (Solomon et al., Nature Geoscience, 2009)

...to 100 000 AD (Shaffer et al., PNAS, 2009)
Is there any indication that climate is predictable in deterministic terms?

Comparison of 3 IPCC TAR and 3 IPCC AR4 climate models with historical series of more than 100 years length in 55 stations worldwide.

Comparison of 3 IPCC AR4 climate models with reality in sub-continental scale (contiguous USA).

Efficiency: -97 to -375

Source: Anagnostopoulos, et al. (2009); Currently displayed next door (XY299).
Is climate predictability in stochastic terms better or worse than in a purely random process?

Many temperature series and proxies indicate very strong persistence with $H \approx 0.94$.

Global satellite-derived temperature of lower troposphere, monthly scale, 1979-today.

Data from the US National Space and Technology Center (lt5.2 series)
[vortex.nsstc.uah.edu/; vortex.nsstc.uah.edu/public/msu/t2lt/tltglhmam_5.2]

Greenland proxy temperature during the last 50 000 years.

Reconstructed from the GISP2 Ice Core. Data from:

Source: Koutsoyiannis et al., 2009.
<table>
<thead>
<tr>
<th>Question</th>
<th>More common reply</th>
<th>Less common reply</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can natural processes be divided in deterministic and random components?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Are probabilistic approaches unnecessary in systems with known deterministic dynamics?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Is stochastics a collection of mathematical tools, unable to give physical explanations?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Are deterministic systems deterministically predictable?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Can uncertainty be eliminated (or radically reduced) by discovering a system’s deterministic dynamics?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Does positive autocorrelation (i.e. dependence), improve long term predictions?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Are deterministic predictions of climate possible?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Are the popular climate “predictions” or “projections” trustworthy and able to support decisions on water management, hydraulic engineering, or even “geoengineering” to control Earth’s climate?</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
Who is right?

*Αἰών παῖς ἐστι παίζων πεσσεύων. Παιδός ἡ βασιληί.*

*Time is a child playing, throwing dice. The ruling power is a child’s.*

(Heraclitus; ca. 540-480 BC; Fragment 52)

*I am convinced that He does not throw dice.*

(Albert Einstein, in a letter to Max Born in 1926)
References