AGU FALL MEETING

San Francisco | 15–19 December 2014

Session H54F

Theories and Methods for Nonstationary Hydrological Frequency Analysis II

Risks from dismissing stationarity

Demetris Koutsoyiannis

National Technical University of Athens, Greece (dk@itia.ntua.gr)

Alberto Montanari

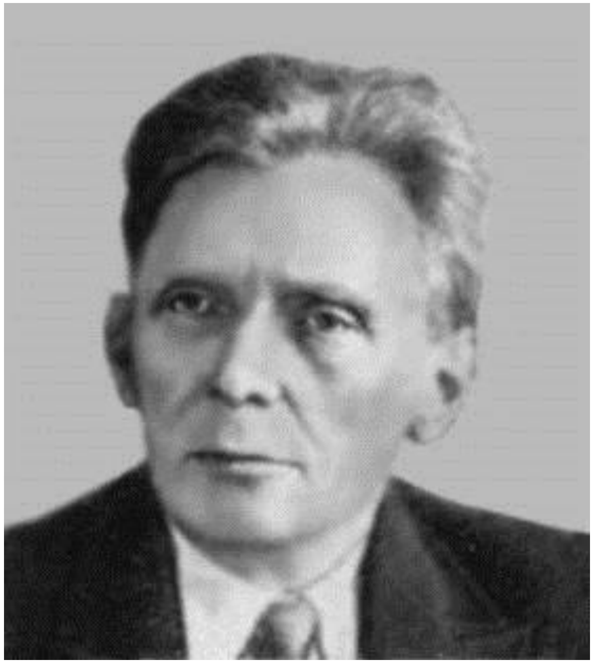
University of Bologna, Italy (alberto.montanari@unibo.it)

Presentation available online: www.itia.ntua.gr/1499/

The meaning of stationarity —A historical note

Andrey Kolmogorov
(1903–1987)

- Kolmogorov (1931)
  - clarified that the term *process* means *change of a certain system*;
  - introduced the term *stochastic process*;
  - used the term *stationary* to describe a probability density function that is unchanged in time.



Aleksandr Khinchin
(1894 – 1959)

- Khinchin (1934) gave more formal definitions of a *stochastic process* and of *stationarity*.

Definition of stationarity

- Kolmogorov (1938) gave a concise presentation of the definition as follows:

  a stationary stochastic process [...] is a set of random variables x_t depending on the parameter t, $-\infty < t < +\infty$, such that the distributions of the systems

  $$(x_{t_1}, x_{t_2}, \ldots, x_{t_n}) \text{ and } (x_{t_1+\tau}, x_{t_2+\tau}, \ldots, x_{t_n+\tau})$$

  coincide for any n, t_1, t_2, ... , t_n, and τ.
- Processes that are not stationary are called **nonstationary**; their statistical properties (at least some of them) change in time being **deterministic** functions of time.
- As far as we know:
  - This definition of stationarity has never been disputed.
  - There has never been a decent alternative definition of stationarity.
  - The terms stationary and stationarity are often misused.

Theoretical concepts that help avoid misuse of stationarity and nonstationarity

- Stationarity and nonstationarity refer to stochastic processes.
- Stochastic processes are families of random variables usually indexed by time.
- Random variables are variables associated with a probability distribution or density function.
- Attempts to conceptualize stationarity without reference to a stochastic process are inconsistent with the theory.

Model type	Studied properties	Stationary case	Nonstationary case
Deterministic	Evolution in time of system states without reference to statistical properties	Not applicable	Not applicable
Non-deterministic (statistical/stochastic)	Evolution in time of probabilities of systems states or statistical properties thereof	Probabilities and statistical properties are assumed constant in time	Probabilities and statistical properties are changing according to deterministic functions of time

In which world do stationarity and nonstationarity belong?

Real world

Physical system

Unique evolution

Time series (observed)

Many different models can be constructed

Mental copies depend on model constructed

The observed time series is unique; the simulated can be arbitrarily many

Abstract representation

Model (Stochastic process)

Ensemble (Gibbs's idea): **mental copies** of natural system

Time series (simulated)

Perpetual change

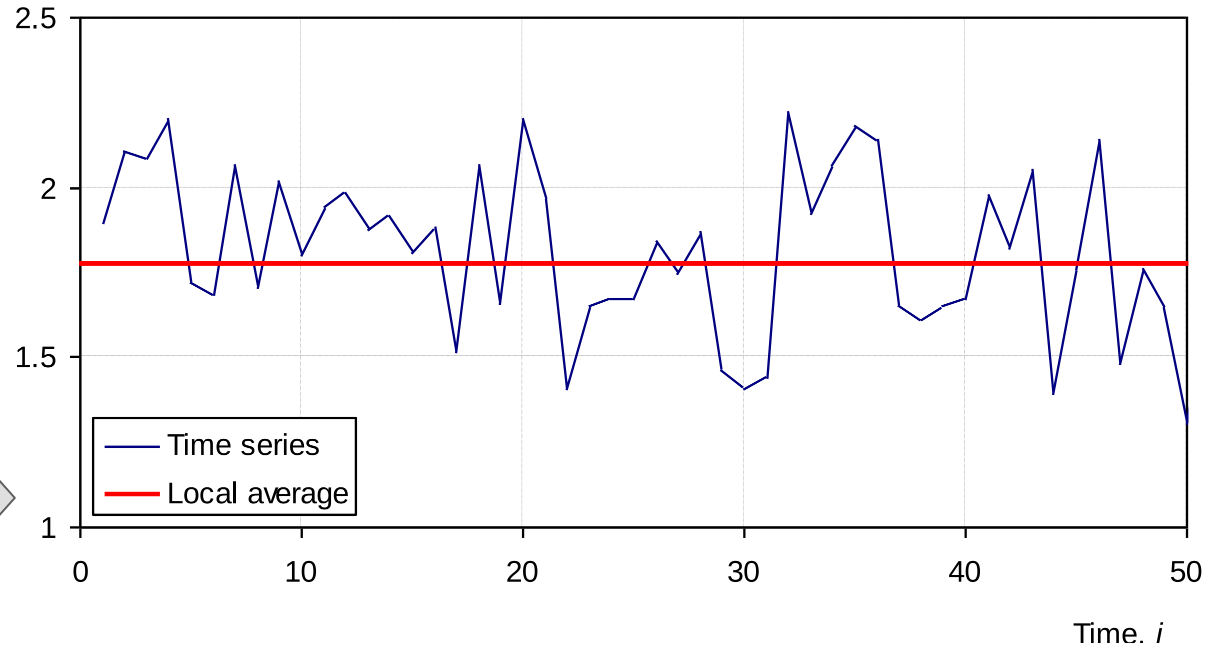
An important consequence:
Stationarity is immortal

Both stationarity and nonstationarity apply here (not in the real world)

Does a time series tell us if it is stationary or nonstationary?

- Not actually.
- Actually, a time series is neither stationary nor nonstationary.
- These are properties of the stochastic process that generated the time series.

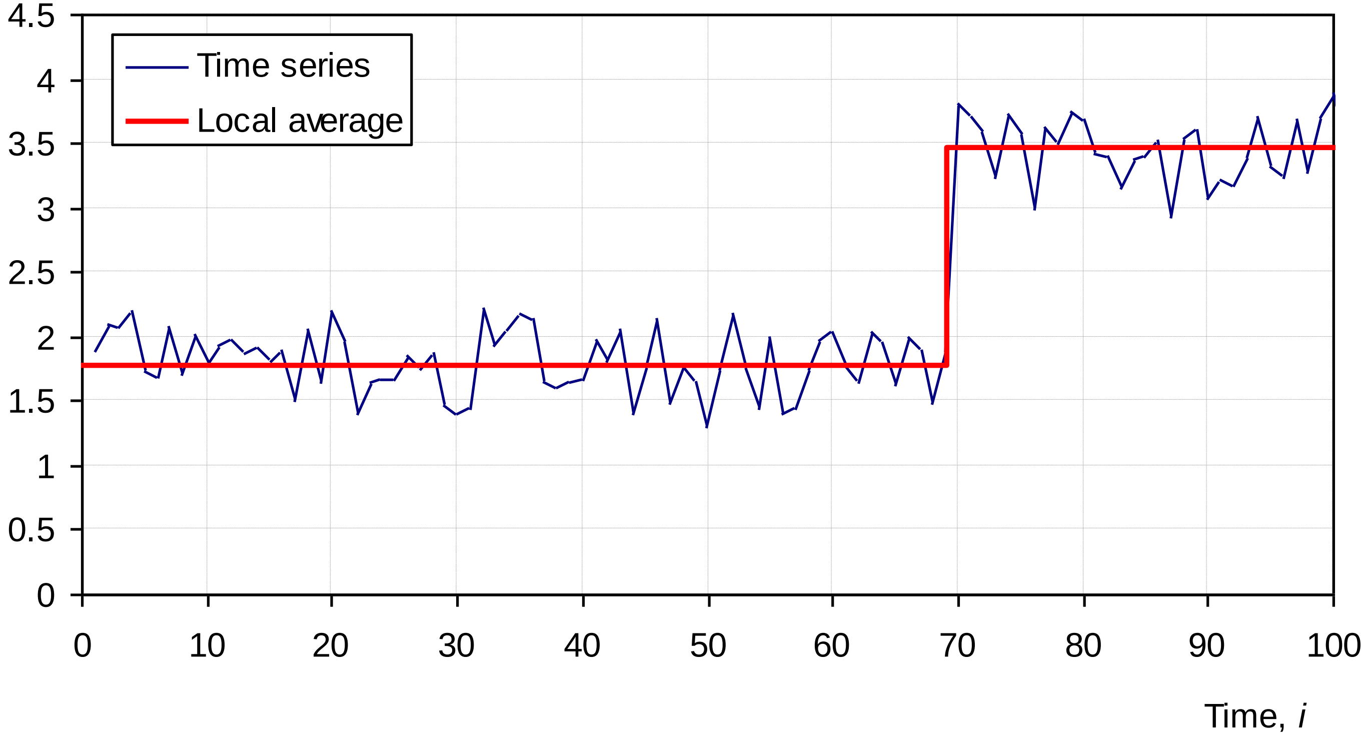
Example:
50 terms of a synthetic time series



See details of this example in Koutsoyiannis (2011)

Does this example suggest stationarity or nonstationarity?

Example time series extended up to time 100



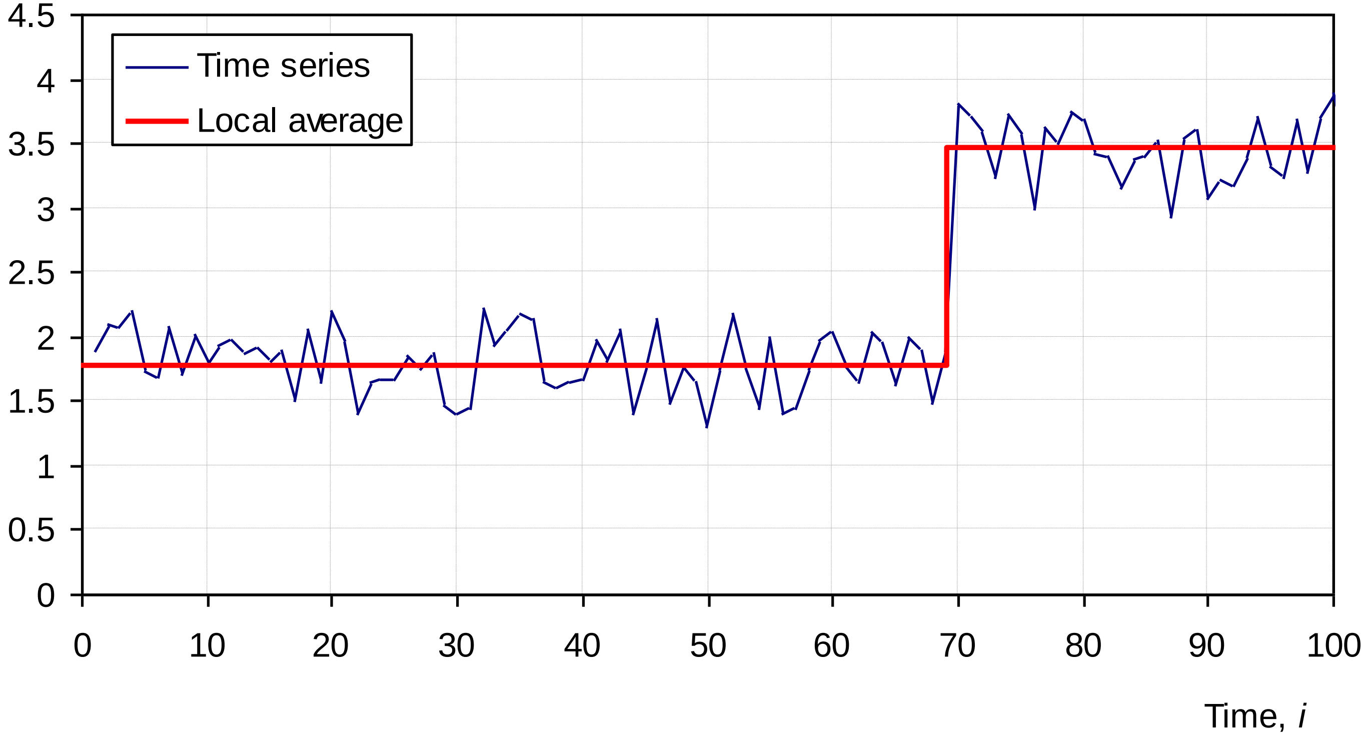
Mean m (red line) of time series (blue line) is:

m = 1.8 for i < 70

m = 3.5 for $i \geq 70$

Reformulation of question: Does the red line reflect a **deterministic** function?

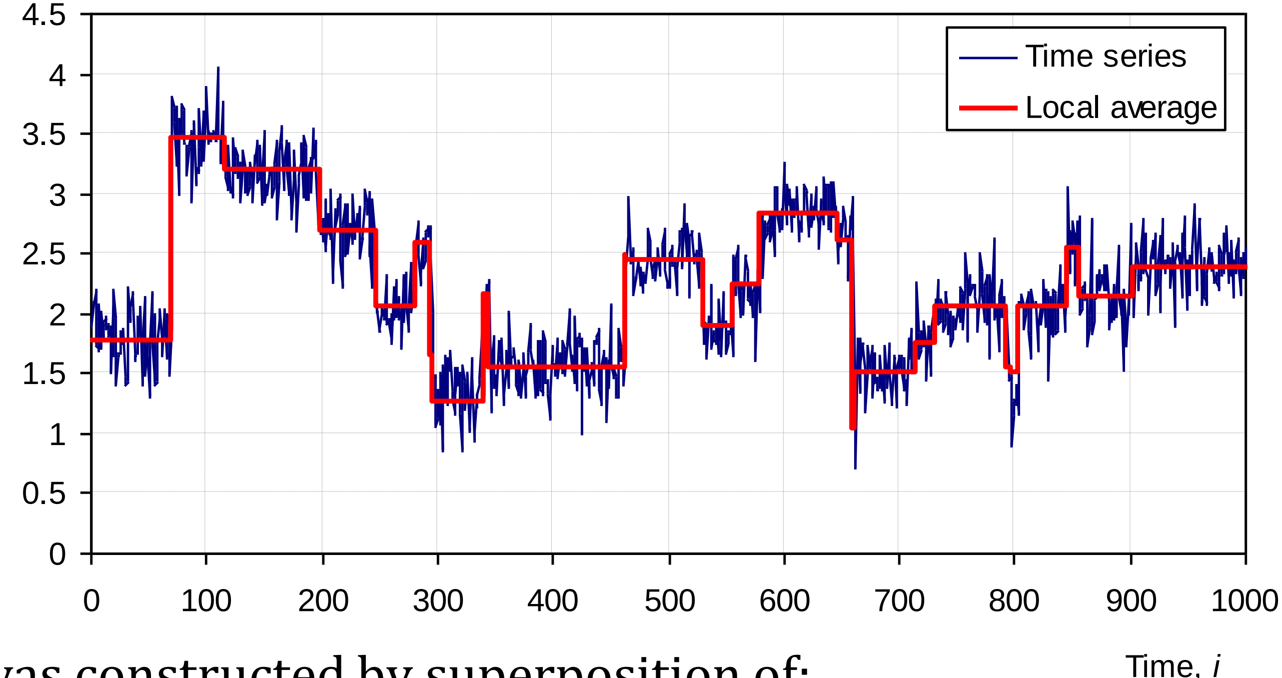
Example time series extended up to time 100



- If the red line is a deterministic function of time: → **nonstationarity**.
- If the red line is a random function (realization of a stationary stochastic process) → **stationarity**.

Answer of the last question: the red line is a realization of a stochastic process

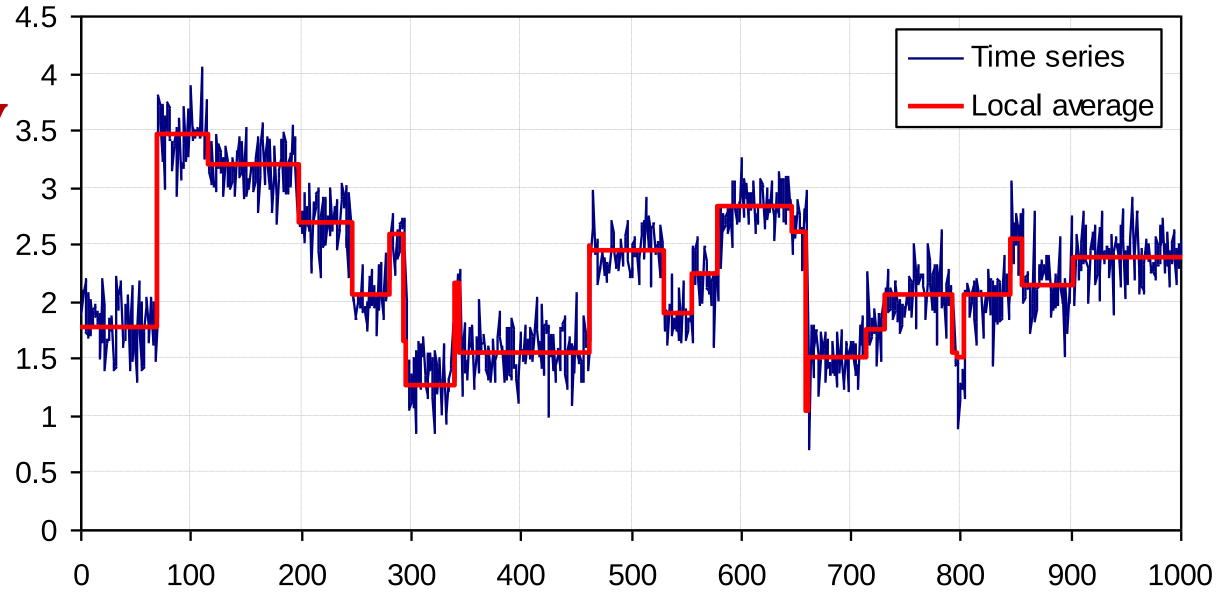
Example 1 extended up to time 1000



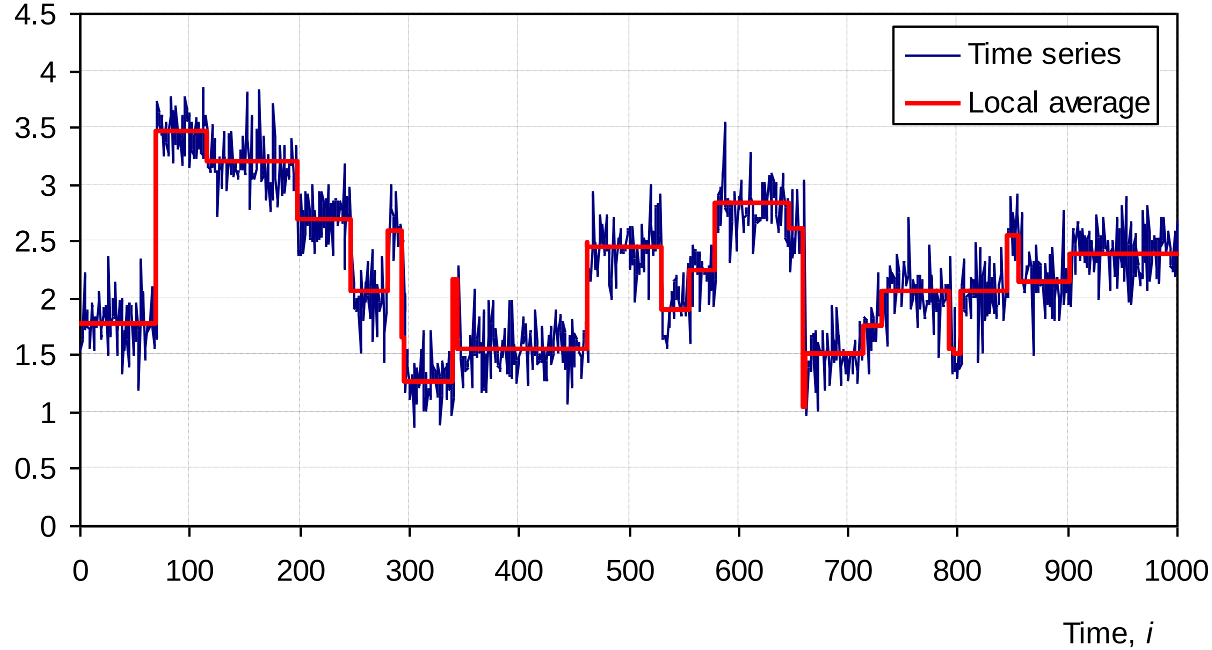
- The time series was constructed by superposition of:
  - A stochastic process with values $\underline{m}_j \sim N(2, 0.5)$ each lasting a period $\underline{\tau}_j$ exponentially distributed with $E[\underline{\tau}_j] = 50$ (red line);
  - White noise $N(0, 0.2)$.
- Nothing in the model is nonstationary.
- The process of our example is **stationary.**

The big difference of **nonstationarity** and stationarity (1)

The initial time series

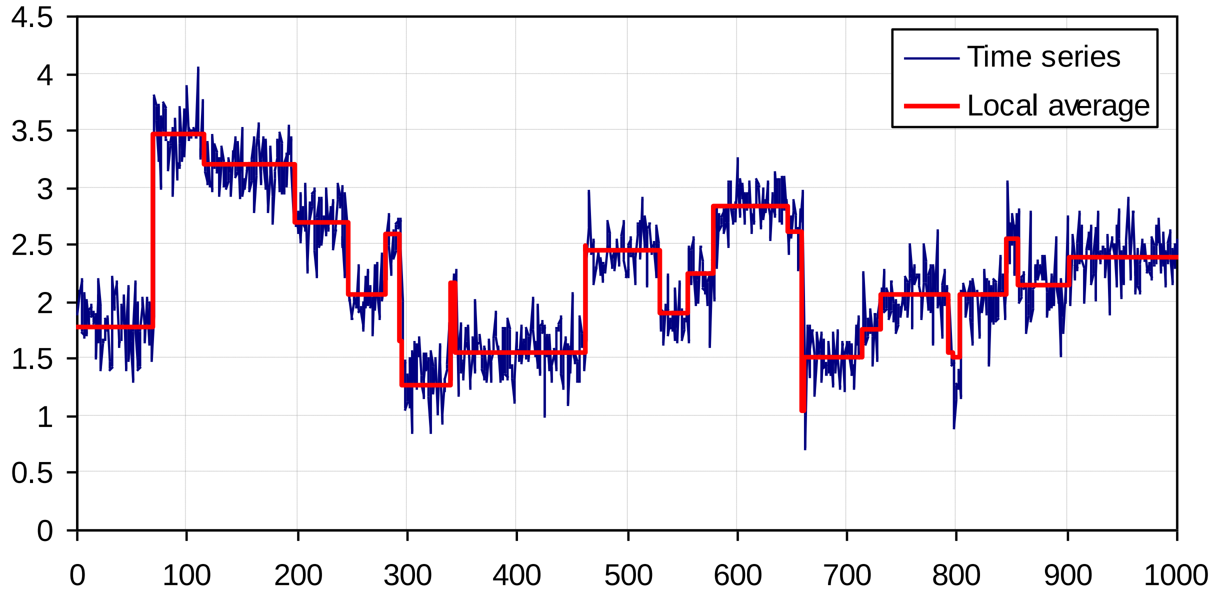


A mental copy generated by a nonstationary model (assuming the red line is a deterministic function)

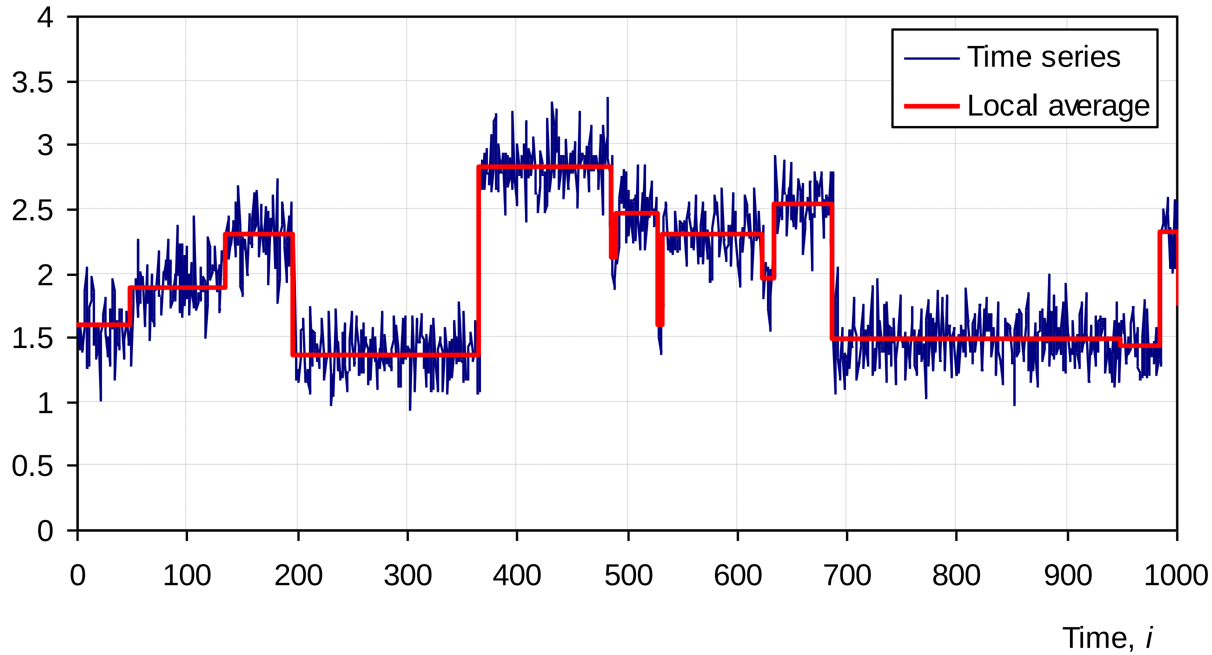


Unexplained variance (differences between blue and red line): $0.2^2 = 0.04$.

The big difference of nonstationarity and **stationarity** (2)



The initial time series



A mental copy generated by a stationary model (assuming the red line is a stationary stochastic process)

Unexplained variance (the "undecomposed" time series): 0.38 (~10 times greater).

Justified use of nonstationary descriptions: Models for the past

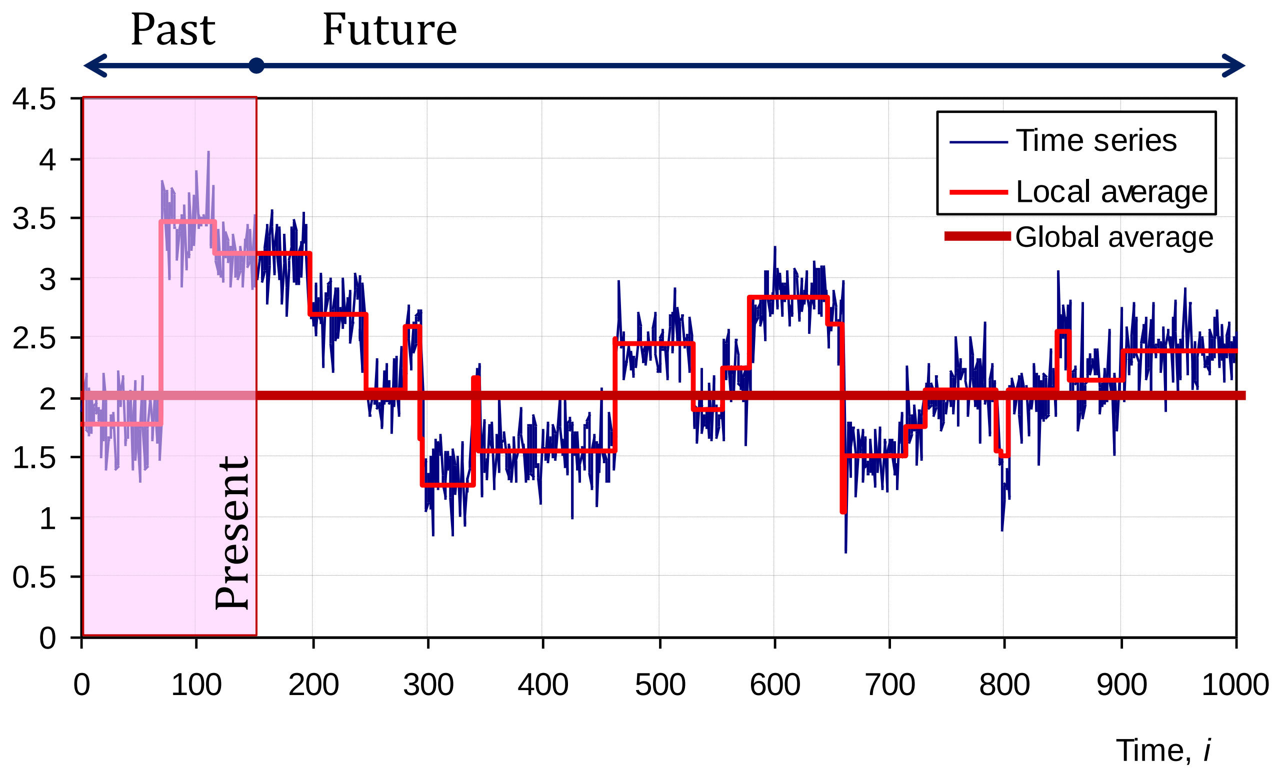
- Changes in catchments happen all the time, including in quantifiable characteristics of catchments and conceptual parameters of models.
- If we know the evolution of these characteristics and parameters (e.g. we have information about how the percent of urban area changed in time), then we build a nonstationary model:
  - Information → Reduced uncertainty → Nonstationarity.
- If we do not have this quantitative information, then:
  - We treat catchment characteristics and parameters as random variables.
  - We build stationary models entailing larger uncertainty.

Justified uses of nonstationary descriptions: Models for the future

- It is important to distinguish explanation of observed phenomena in the past from modelling that is made for the future.
- Except for trivial cases, the future is not easy to predict in deterministic terms.
- If changes in the recent past are foreseen to endure in the future (e.g. urbanization, hydraulic infrastructures), then the model of the future should be adapted to the most recent past.
  - This may imply a stationary model of the future that is different from that of the distant past (prior to the change).
  - It may also require “stationarizing” of the past observations, i.e. adapting them to represent the future conditions.
- In the case of planned and controllable future changes (e.g. catchment modification by hydraulic infrastructures, water abstractions), which indeed allow prediction in deterministic terms, nonstationary models are justified.

Conditional nonstationarity arising from stationarity models

- If the prediction horizon is long, then in modelling we will use the global average and the global variance.



- If the prediction horizon is short, then we will use the local average at the present time and a reduced variance.
- This is not called nonstationarity; it is dependence in time.
- When there is dependence (i.e., always) observing the present state and conditioning on it looks like local nonstationarity.

In nonstationary models stationarity is again important

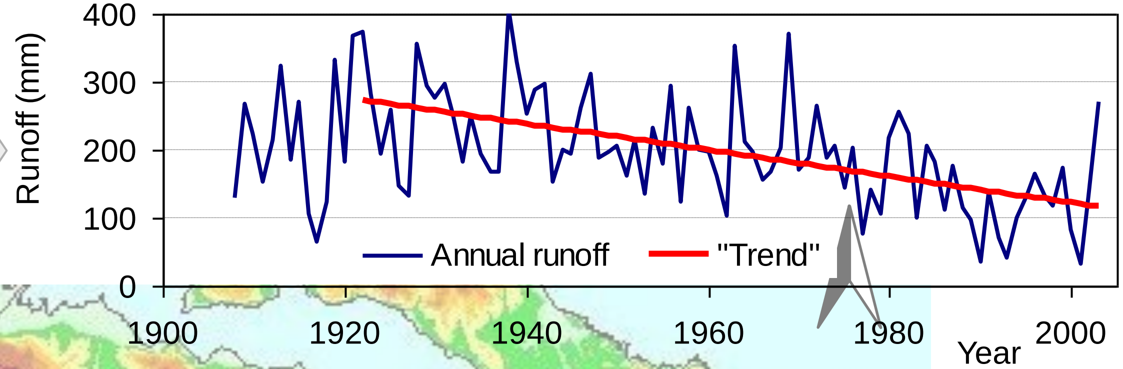
- Even if we have a good deterministic model applicable for future times, we can never hope that it will describe the future in full detail and precision.
- Uncertainty will ever be present.
- That uncertainty (unexplained variability) should be represented as a random component superimposed to the deterministic change given by the deterministic model; that random component is necessarily stationary.
- Thus, even if a process is nonstationary, it will necessarily include a stationary component, and therefore any future prediction needs to ultimately rely on the assumption of stationarity of that random part.

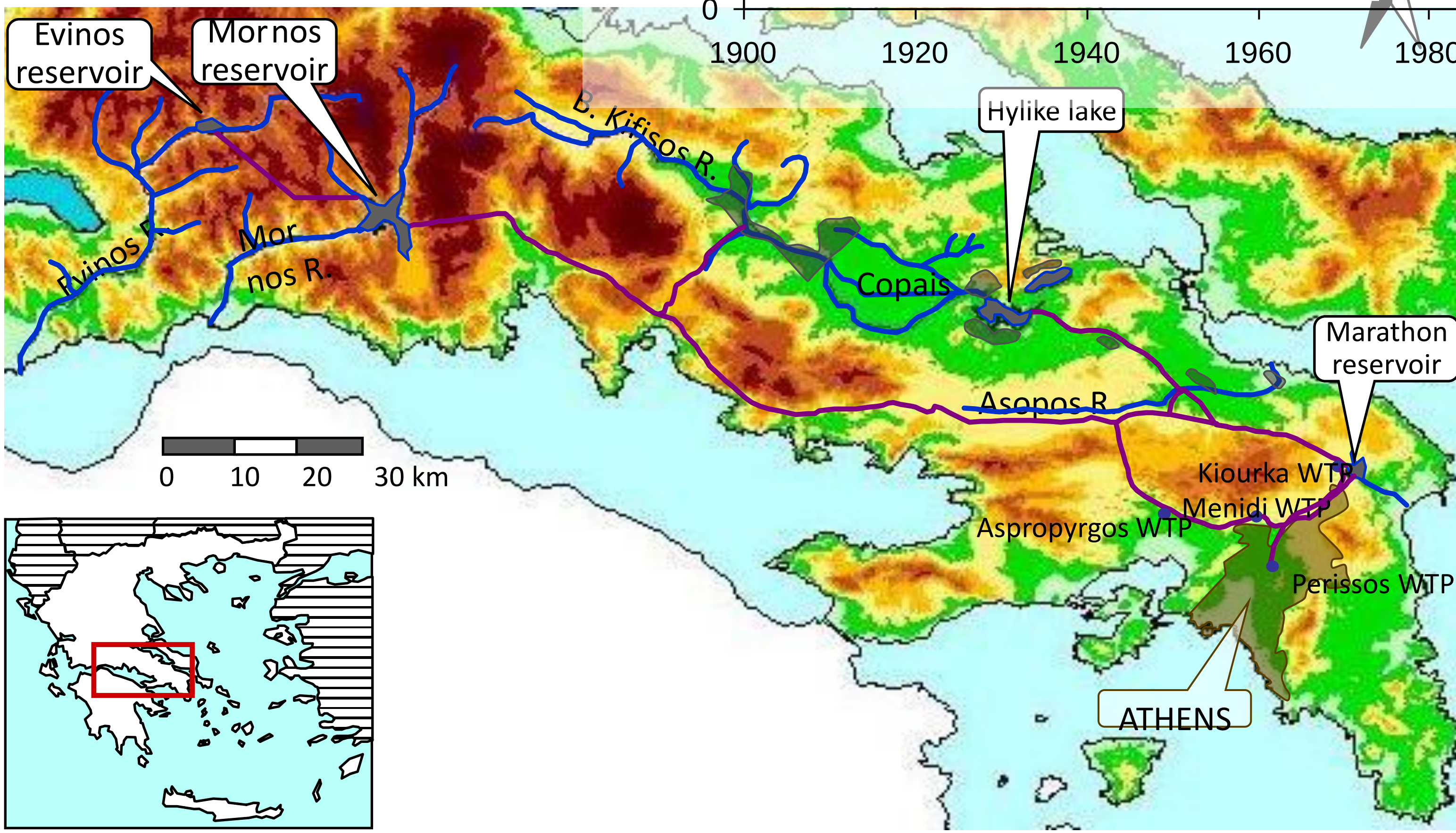
The discussion about stationarity is beyond semantics

- For mitigation of natural hazards, solving practical problems implies the design of management policies and engineering structures that need to be based on the estimation of design variables and their uncertainty, which is also related to economical feasibility of solutions.
- The stationarity concept is useful because it highlights the fact that, whatever deterministic controls and mechanisms are identified and whatever progress is made in deterministic modelling, there will always be unexplainable variability in any system for which a probabilistic description assuming stationarity is needed.
- Both **exact predictability** (particularly for distant times) and **inference without data** are **impossible**.
- Only (physically-based) stochastic modelling using real-world data offers a pragmatic solution.
- Thus, it is not paradoxical to conclude that **stationarity is immortal**, as immortal is the need for statistical descriptions and the need to seek robust solutions to practical problems.

A real-world case: The Athens water supply system

Historical time series of Boeoticos Kephisos runoff

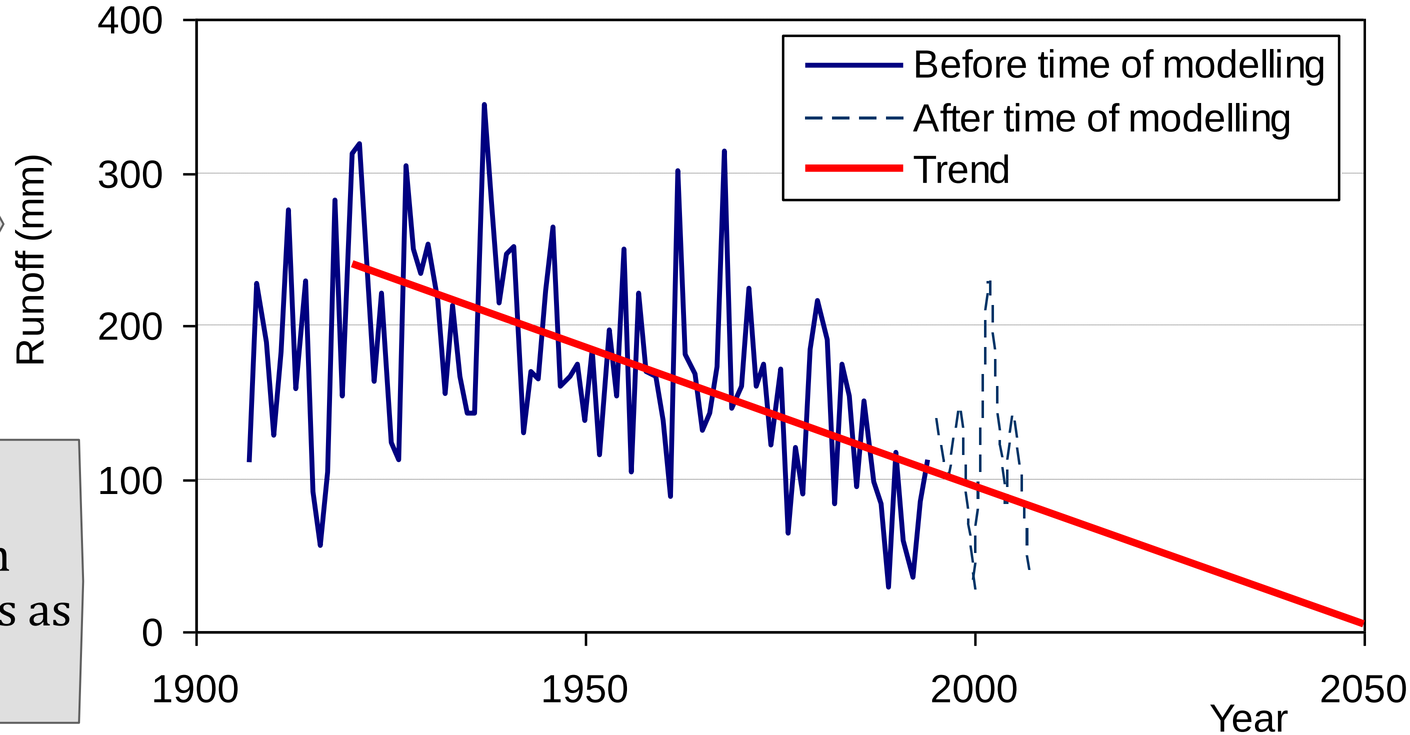




Nonstationary approach 1: trend based

- The flows would disappear at about 2050…
- The trend reduces uncertainty (because it “explains” part of variability): The initial standard deviation of 70 mm decreases to 55 mm.
- In contrast, in a stationary approach assuming Hurst-Kolmogorov dynamics (consistent with the data) the standard deviation increases to 75 mm.

Boeoticos Kephisos runoff and projected trend



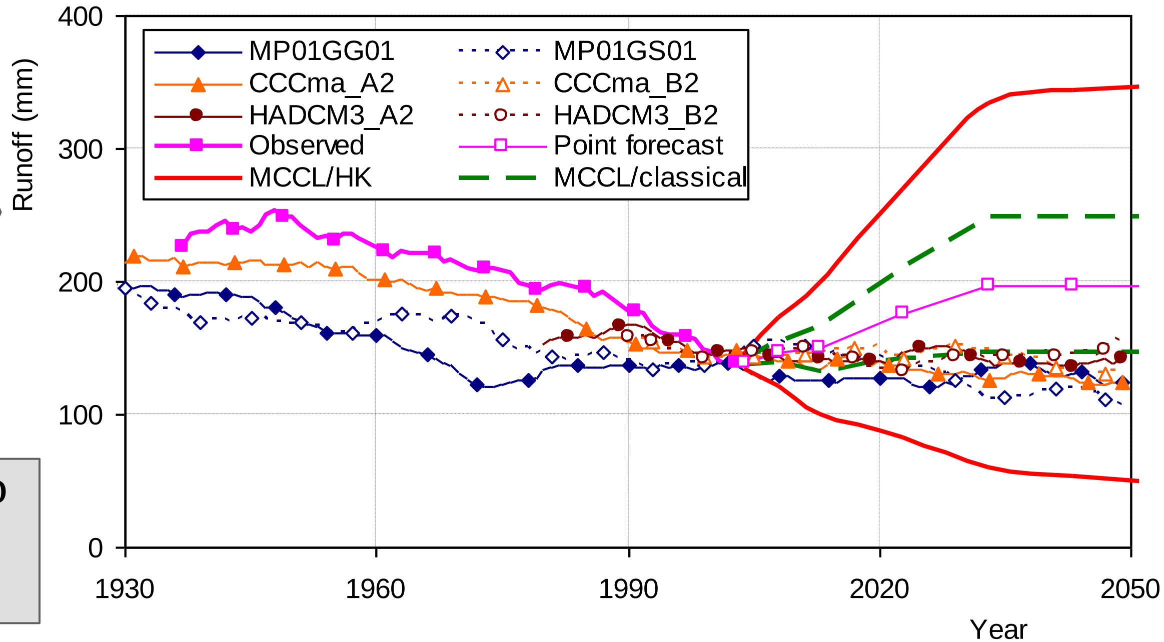
Conclusion: It is absurd to use such simplistic methods as trend projection

Nonstationary approach 2: GCM based

- Outputs from three GCMs for two scenarios were used.
- The original GCM outputs (not shown) had no relation to reality (highly negative efficiencies at the annual time scale and above).
- After corrections (also known as “downscaling”) the GCM outputs improved with respect to reality (to about zero efficiencies at the annual time scale).
- For **the past**, despite adaptations, the proximity of models with reality is not satisfactory.
- For **the future** the runoff obtained by adapted GCM outputs is too stable.

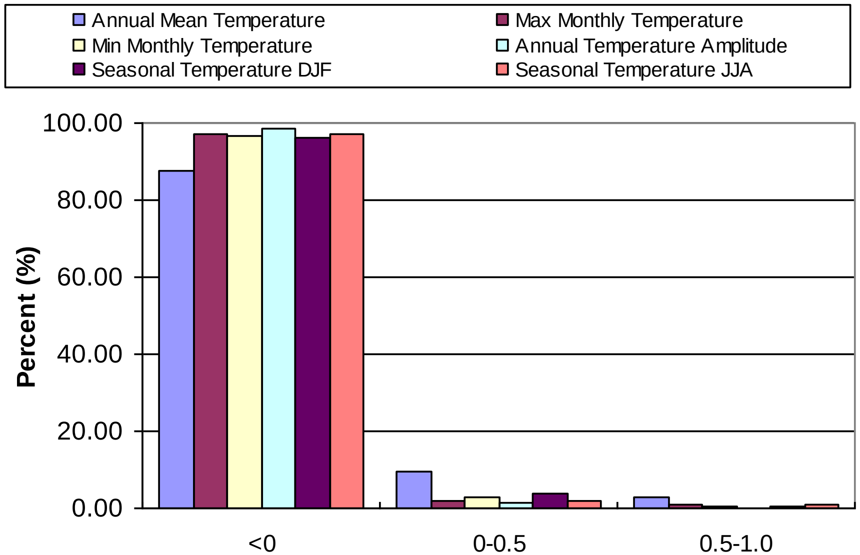
Boeoticos Kephisos runoff produced with downscaled GCM outputs, superimposed to confidence zones produced with Hurst-Kolmogorov statistics under stationarity

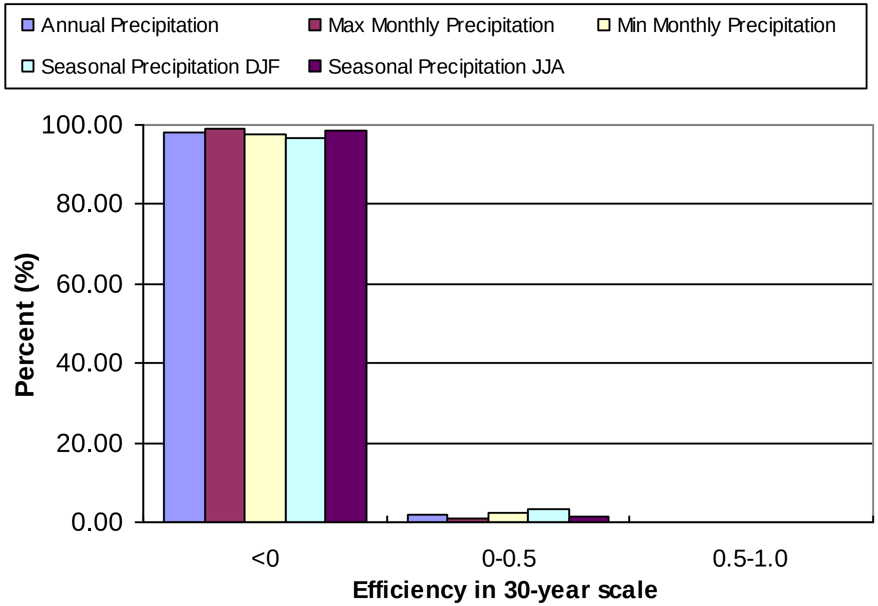
(Koutsoyiannis et al., 2007)



Conclusion: It is dangerous to use GCM future projections: they hide uncertainty.

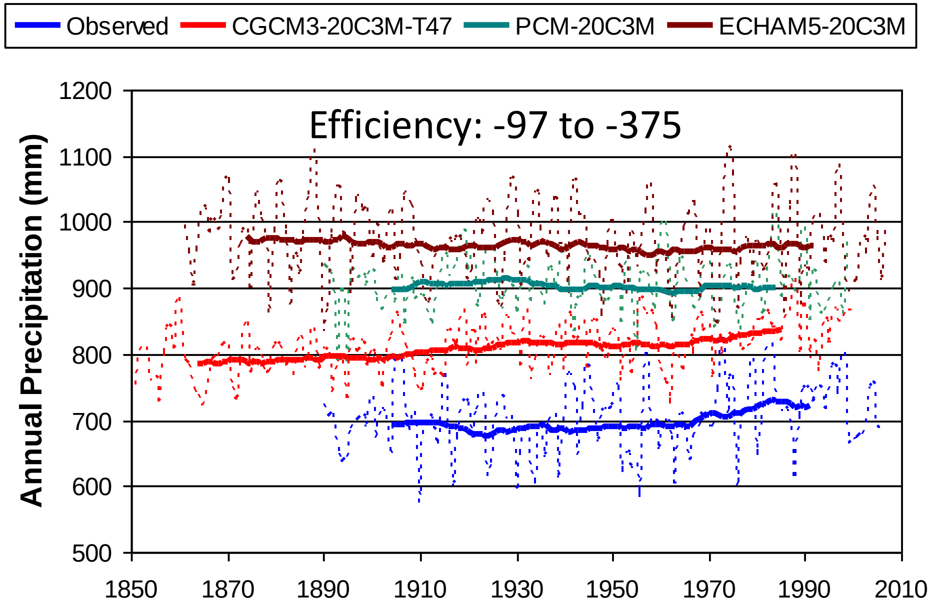
A note on the skill of GCMs in reproducing reality





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

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



Source: Anagnostopoulos, *et al.* (2009)
See also Koutsoyiannis *et al.* (2008).

Πάντα ῥεῖ: Does change entail nonstationarity?

Reply: No
See justification in a series of papers

- Montanari, A., G. Young, H. H. G. Savenije, D. Hughes, T. Wagener, L. L. Ren, D. Koutsoyiannis, et al., “Panta Rhei – Everything Flows”, Change in Hydrology and Society – The IAHS Scientific Decade 2013-2022, *Hydrological Sciences Journal*, 58 (6), 1256–1275, 2013.
- Koutsoyiannis, D., Hydrology and Change, *Hydrological Sciences Journal*, 58 (6), 1177–1197, 2013.
- Ceola, S., A. Montanari, and D. Koutsoyiannis, Toward a theoretical framework for integrated modeling of hydrological change, *WIREs Water*, doi:10.1002/wat2.1038, 2014.
- Koutsoyiannis, D., and A. Montanari, Negligent killing of scientific concepts: the stationarity case, *Hydrological Sciences Journal*, doi:10.1080/02626667.2014.959959, 2014.
- Montanari, A., and D. Koutsoyiannis, Modeling and mitigating natural hazards: Stationarity is Immortal!, *Water Resources Research*, doi:10.1002/2014WR016092, 2014.

Change in Hydrology and Society
IAHS Scientific Decade 2013-2022

Concluding remarks

- Πάντα ῥεῖ (or: Change is Nature’s style).
- Change occurs at all time scales.
- Stationarity is a property of a process and a **process** is synonymous to **change**.
- Nonstationarity should not be confused with change, nor with dependence of a process in time.
- Stationarity and nonstationarity apply to models, not to the real world, and are defined within stochastics.
- Nonstationary descriptions are justified only if the future can be predicted in deterministic terms.
- Unjustified/inappropriate claim of nonstationarity results in underestimation of variability, uncertainty and risk!!!

Stationarity is not dead. It is immortal!!!

References

- Khintchine, A., Korrelationstheorie der stationären stochastischen Prozesse. *Mathematische Annalen*, 109 (1), 604–615, 1934.
- Kolmogorov, A. N., Uber die analytischen Methoden in der Wahrscheinlichkcitsrechnung, *Math. Ann.*, 104, 415-458, 1931. (English translation: On analytical methods in probability theory, In: Kolmogorov, A.N.,. *Selected Works of A. N. Kolmogorov - Volume 2, Probability Theory and Mathematical Statistics*, A. N. Shiryayev, ed., Kluwer, Dordrecht, The Netherlands, pp. 62-108, 1992).
- Kolmogorov, A.N., A simplified proof of the Birkhoff-Khinchin ergodic theorem, *Uspekhi Mat. Nauk*., 5, 52-56,1938. (English edition: Kolmogorov, A.N., *Selected Works of A. N. Kolmogorov - Volume 1, Mathematics and Mechanics*, Tikhomirov, V. M. ed., Kluwer, Dordrecht, The Netherlands, pp. 271-276, 1991).
- Koutsoyiannis, D., Hurst-Kolmogorov dynamics and uncertainty, *Journal of the American Water Resources Association*, 47 (3), 481–495, 2011.
- 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, 2007.