Effective combination of stochastic and deterministic hydrological models in a changing environment

EGU General Assembly 2018, Vienna, Austria, 8-13 April 2018; Session HS2.1.7: *What is a «good» hydrological model for impact study?*

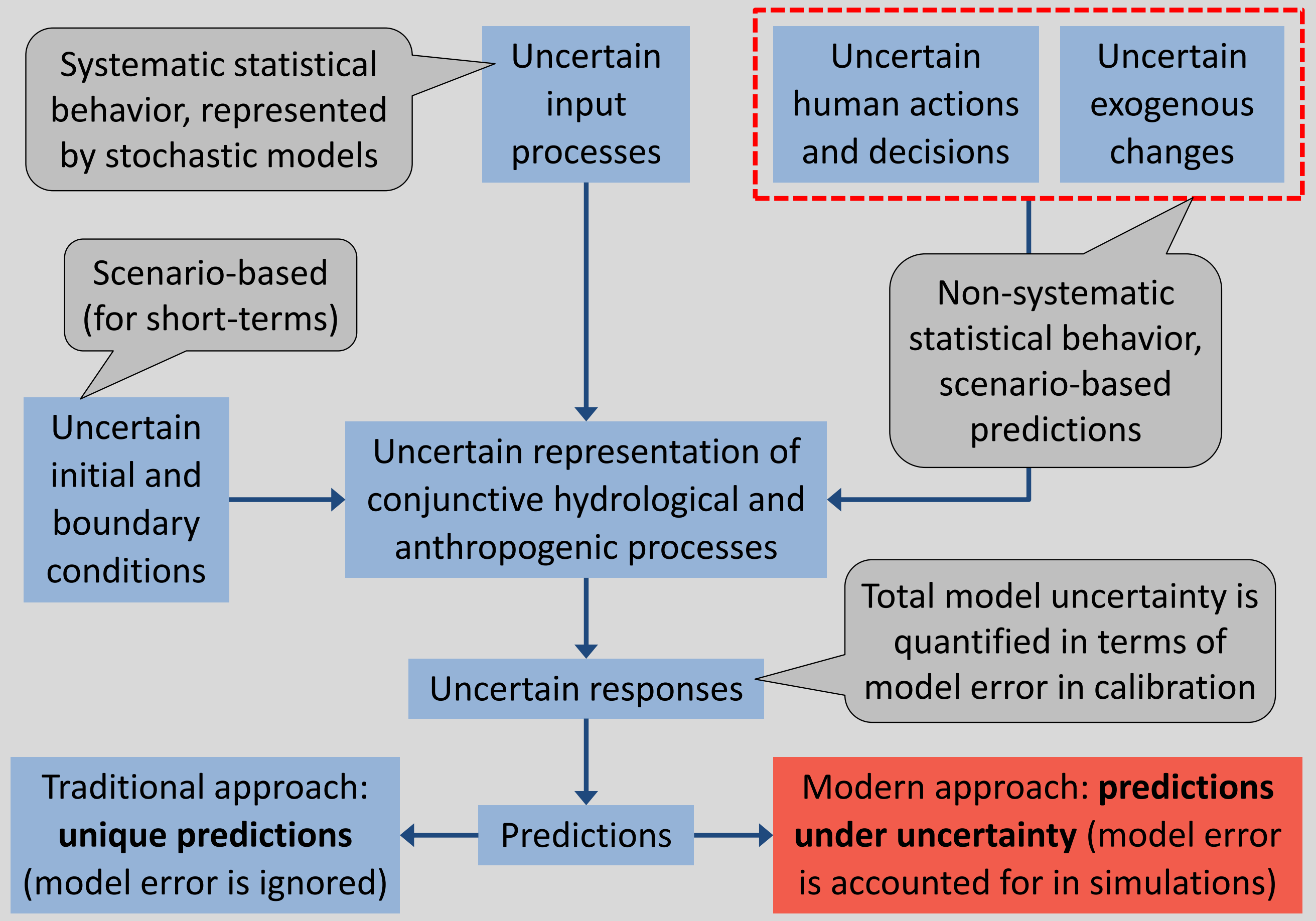
Andreas Efstratiadis[1], Ioannis Nalbantis[2], and Demetris Koutsoyiannis[1]

(1) Department of Water Resources & Environmental Engineering, National Technical University of Athens, Greece; (2) Department of Infrastructure & Rural Development, National Technical University of Athens, Greece

1. Abstract

Water resource systems are subject to **continuous changes**, at all temporal scales. Changes are induced due to the **inherently varying meteorological processes**, **anthropogenic interventions** of all kinds, as well as other **exogenous factors** modifying the system characteristics. Traditionally, **stochastic models**, for generating synthetic input data, and **deterministic hydrological models**, for representing anticipated or hypothesized environmental changes, have been regarded as alternative approaches to provide future projections of the system responses. Given that both approaches are **driven by historical data**, they are restricted by the limited, and sometimes misinterpreted, information of past observations. Using examples from real-world hydrosystems, we propose a **nonlinear stochastic framework**, by coupling stochastic and deterministic models, which aims to take full advantage of the existing **data** and **understanding**. A central assumption is that all **key uncertain aspects of the overall simulation procedure are expressed in stochastic terms** (including model parameters and water demands, among others), while major uncertainties with respect to changing processes that cannot be captured by past data are consistently represented through the **Hurst-Kolmogorov paradigm**.

2. Hydrological simulations under uncertainty



Systematic statistical behavior, represented by stochastic models

Uncertain input processes

Uncertain human actions and decisions

Uncertain exogenous changes

Scenario-based (for short-terms)

Non-systematic statistical behavior, scenario-based predictions

Uncertain initial and boundary conditions

Uncertain representation of conjunctive hydrological and anthropogenic processes

Total model uncertainty is quantified in terms of model error in calibration

Uncertain responses

Traditional approach: **unique predictions** (model error is ignored)

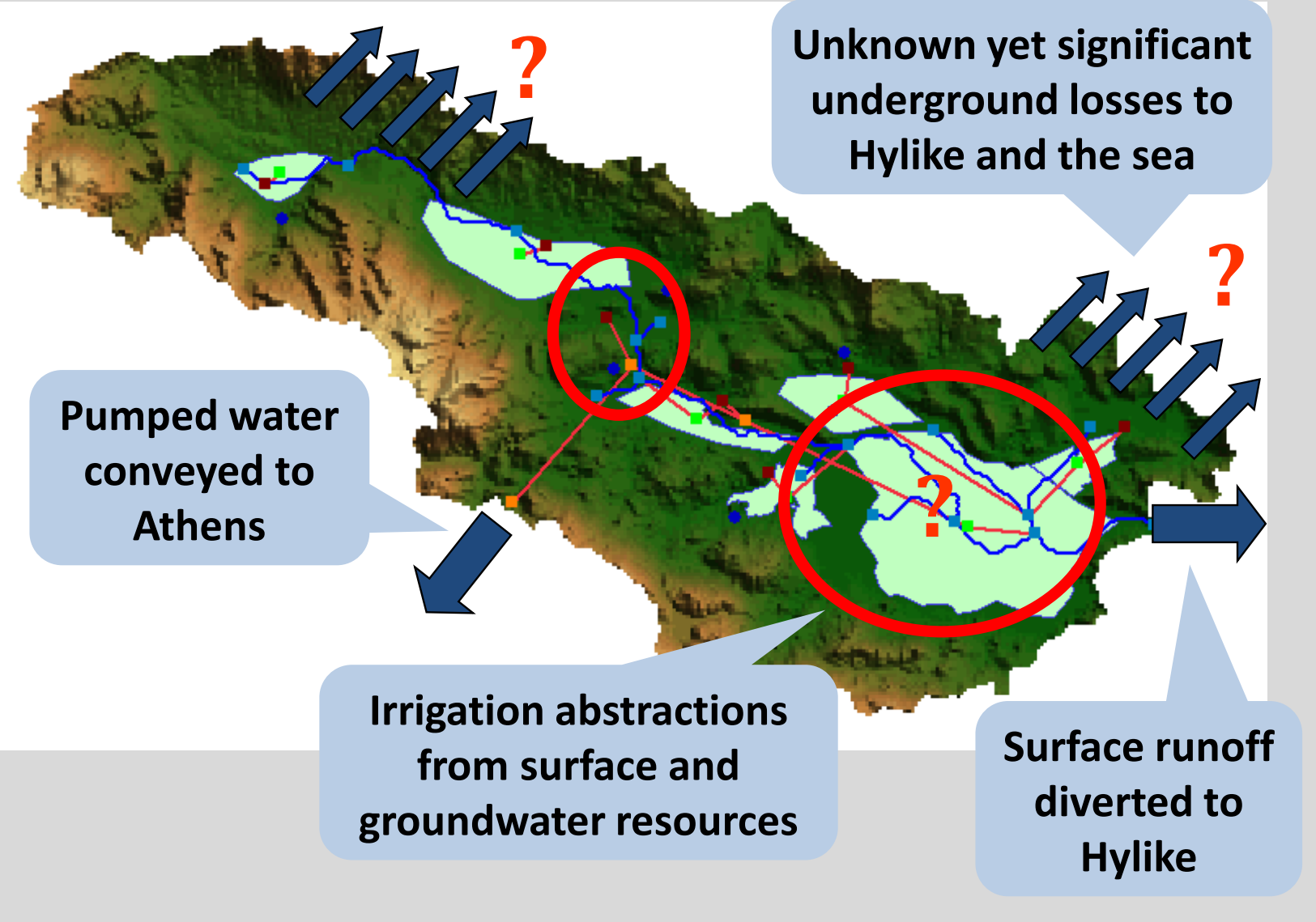
Predictions

Modern approach: **predictions under uncertainty** (model error is accounted for in simulations)

See also discussion in Efstratiadis *et al.* (2015)

3. Case study: The human-modified Boeoticos Kephisos basin

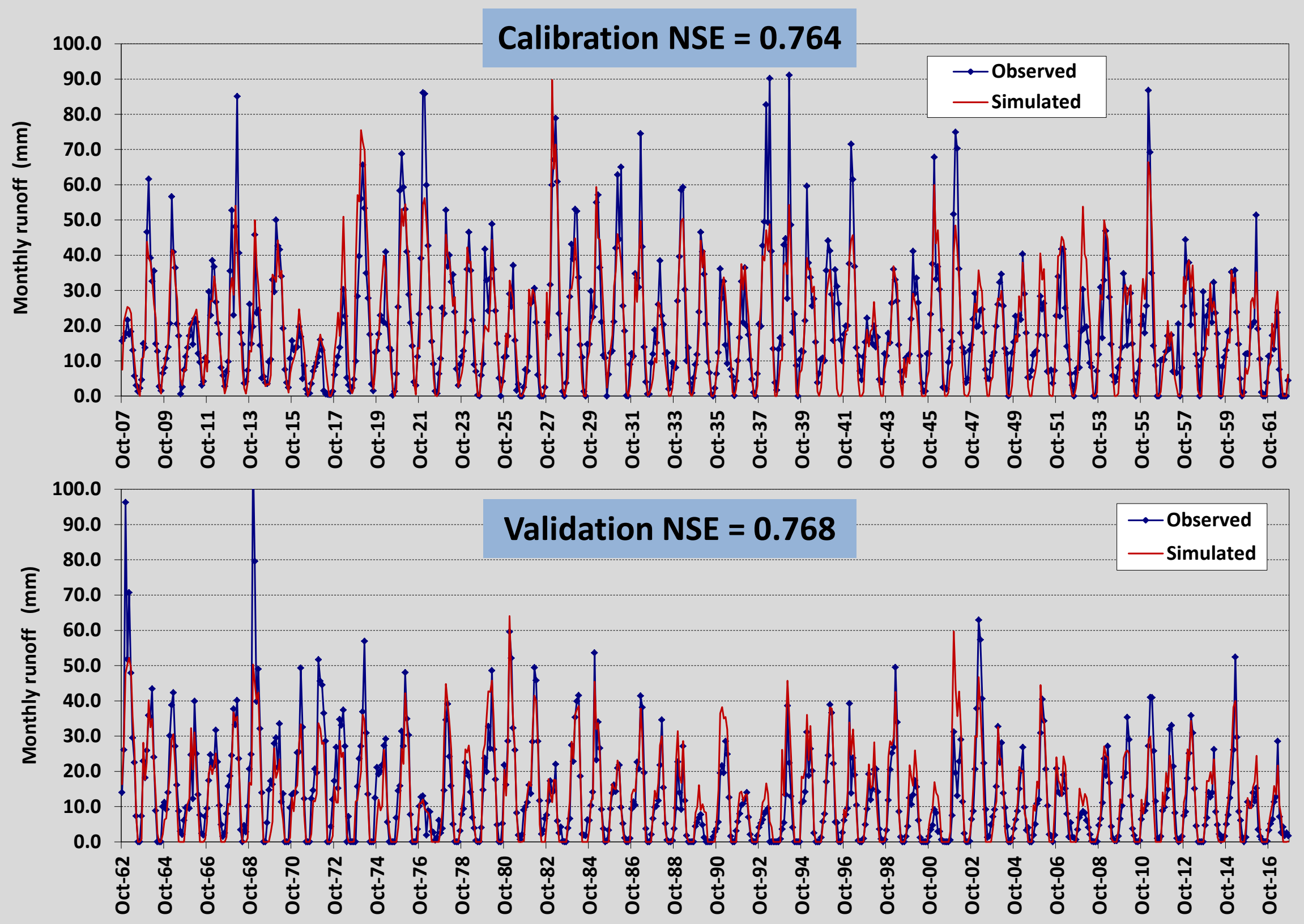
- **Basin history**: Formerly being a closed system, drained into Copais lake; from the late 19th century, all surface runoff is diverted to the neighboring lake Hylike (2nd largest reservoir of the water supply system of Athens).
- **Key characteristics**: Total area 1950 km^2; extended areas of high permeability, due to domination of karst; significant portion of runoff generated by large karst springs; modified status, due to both surface and groundwater abstractions; extended yet unknown groundwater losses to lake Hyilike and the sea.
- **Problem statement**: Predict monthly inflows to Hyilike, under different water demand scenarios.
- **Historical hydrological data**: Monthly precipitation, potential evapotranspiration, and runoff at the outlet – **the longest hydrological records in Greece** (110 years; Oct. 1907 to Sep. 2017).
- **Water management data**: From 1970, estimations of irrigation demand, based on theoretical crop needs and associated irrigated areas; assumption of 2% annual backward decrease up to the beginning of simulation (1907); assumption that 40% of irrigation demand is fulfilled via pumping and 60% from surface water abstractions.

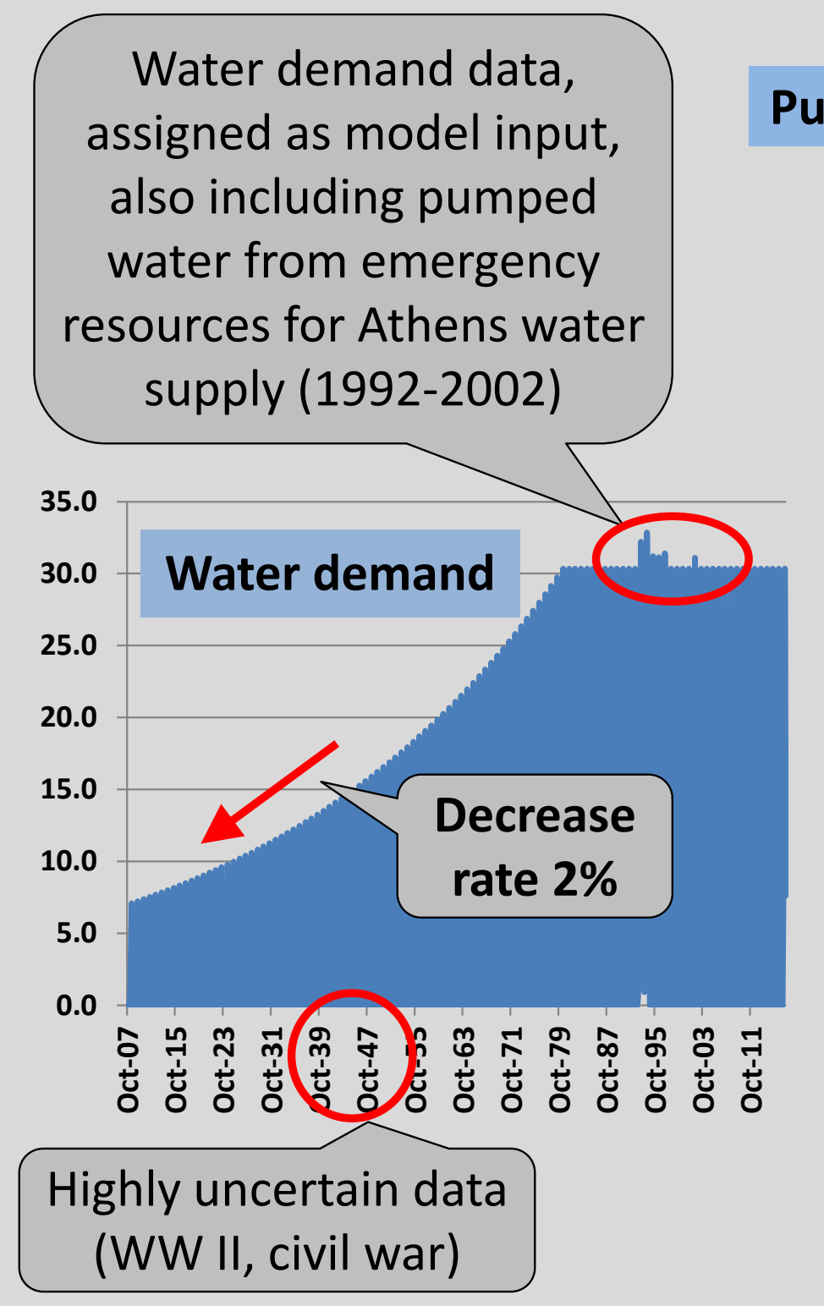


Boeoticos Kephisos hydrosystem

4. Deterministic hydrological model in (classical) calibration setting

- Conceptual model, comprising three interconnected tanks that account for overland, soil and groundwater processes, using nine parameters .
- Simplified representation of human interventions, by means of water removal from simulated groundwater storage and runoff at the basin outlet, to fulfill corresponding demands.
- Typical split-sample scheme for calibration (1907-1962) and validation (1962-2017), using Nash-Sutcliffe efficiency (NSE) as goodness-of-fitting criterion.

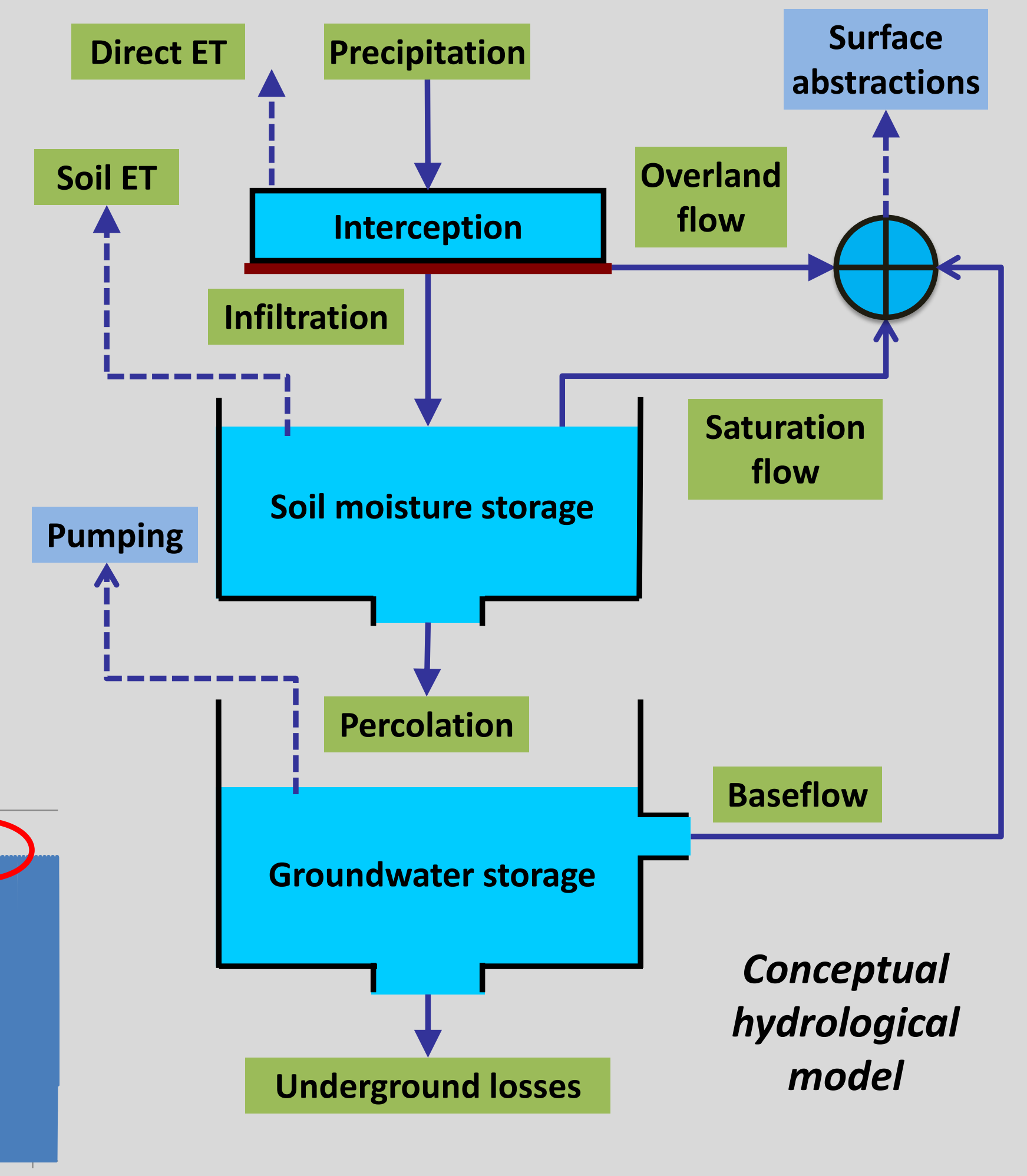




Water demand data, assigned as model input, also including pumped water from emergency resources for Athens water supply (1992-2002)

Decrease rate 2%

Highly uncertain data (WW II, civil war)



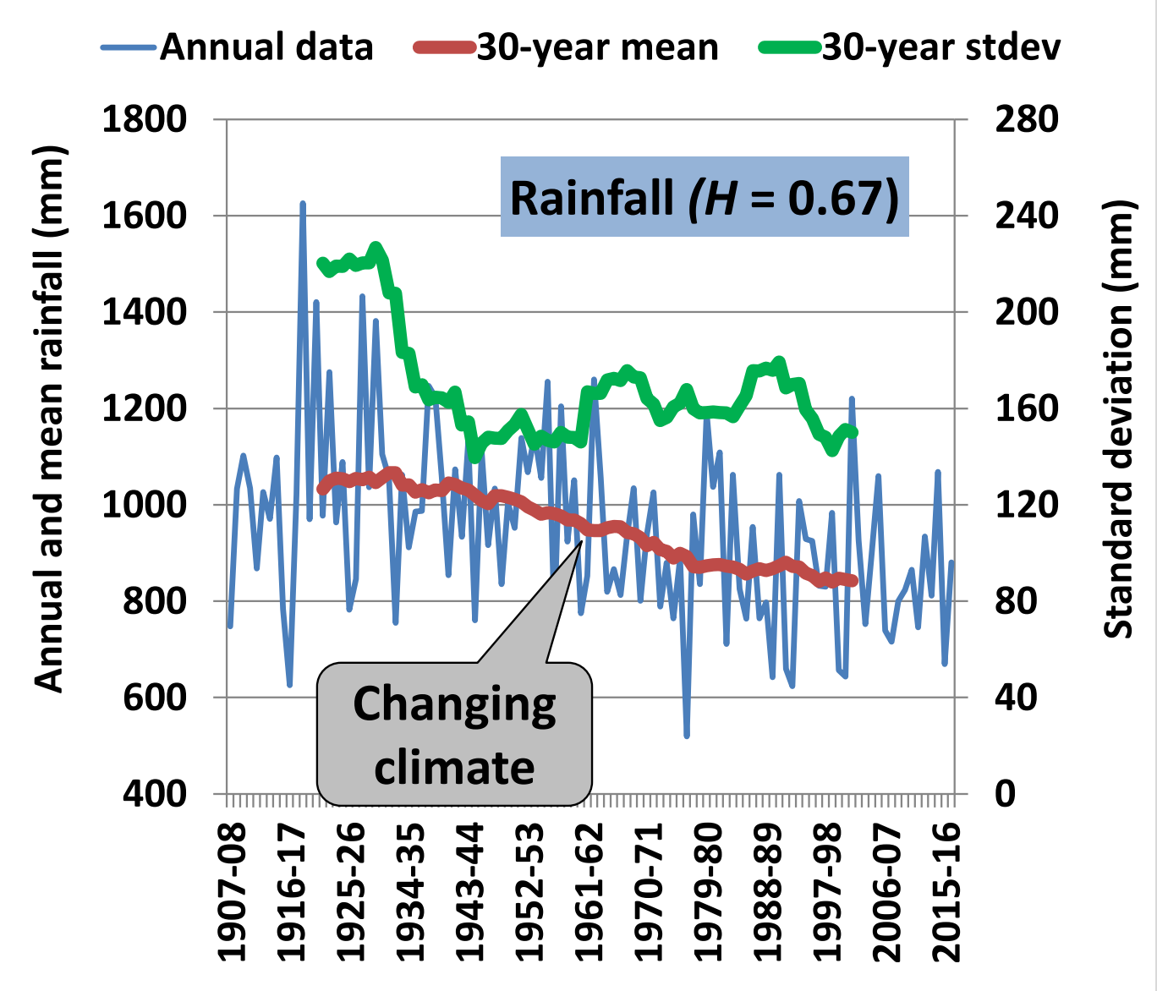
Conceptual hydrological model

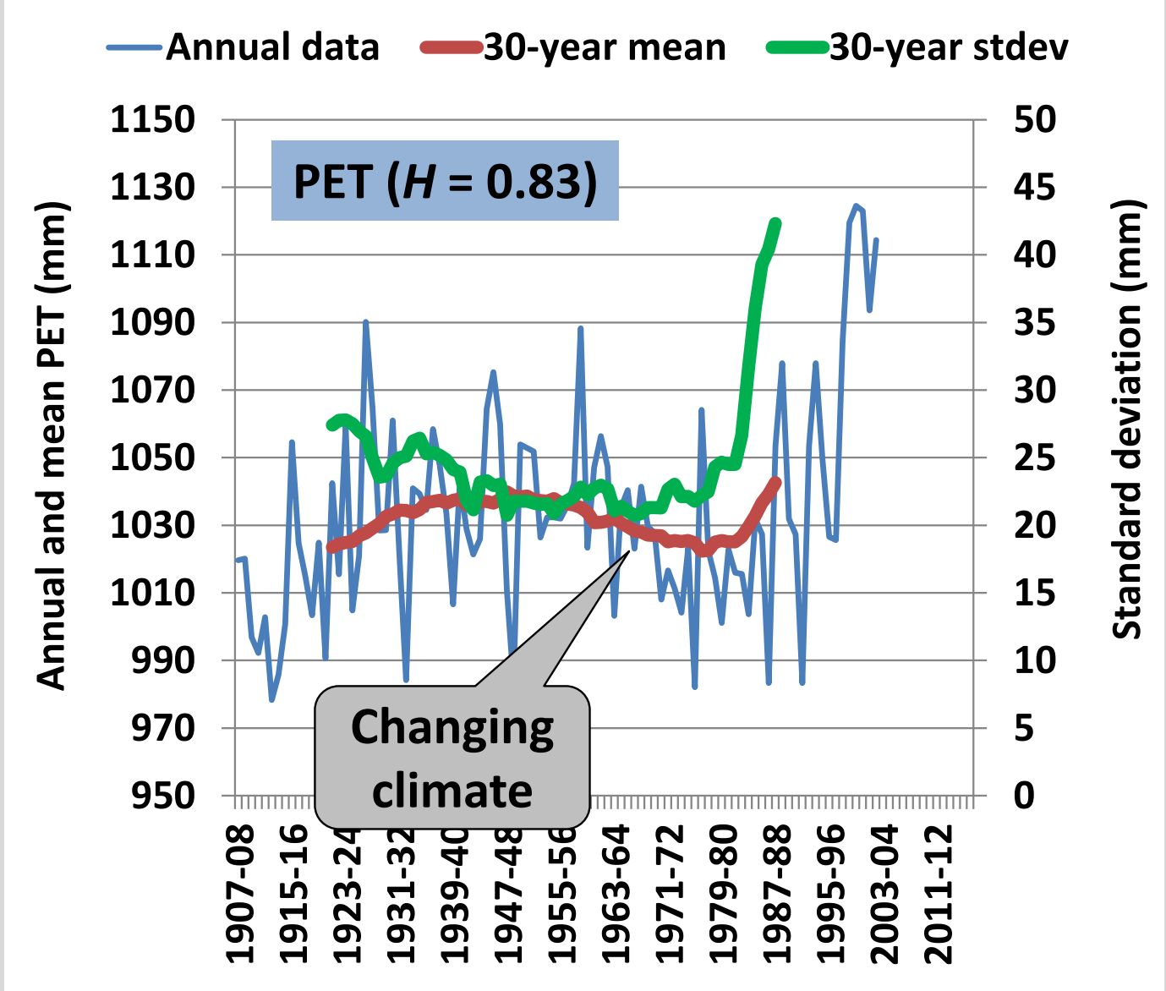
For more advanced (semi-distributed, GIS-based) modelling approaches in the study basin, please refer to Efstratiadis *et al.* (2008) and Nalbantis *et al.* (2011)

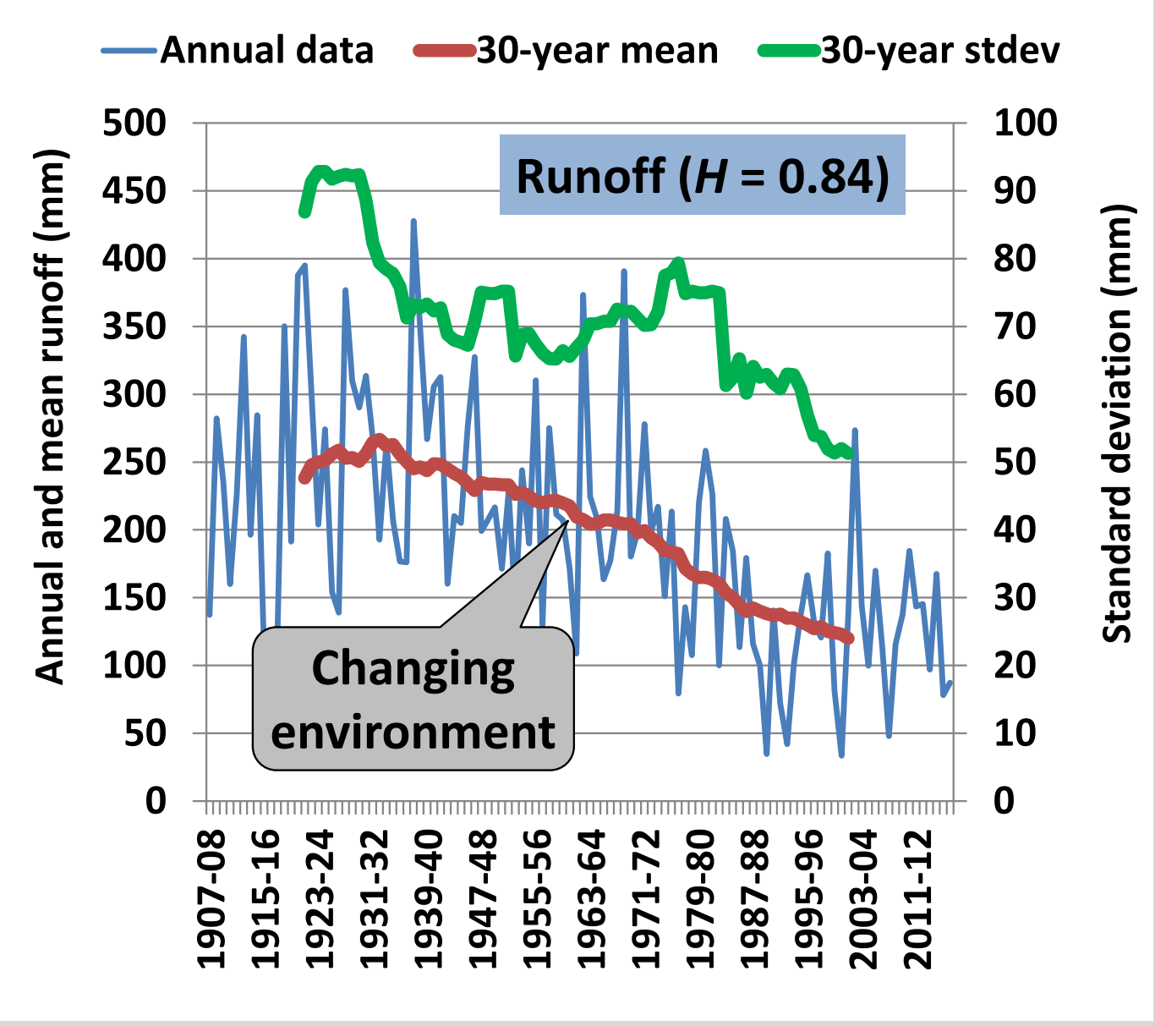
5. A careful look at the long-term data

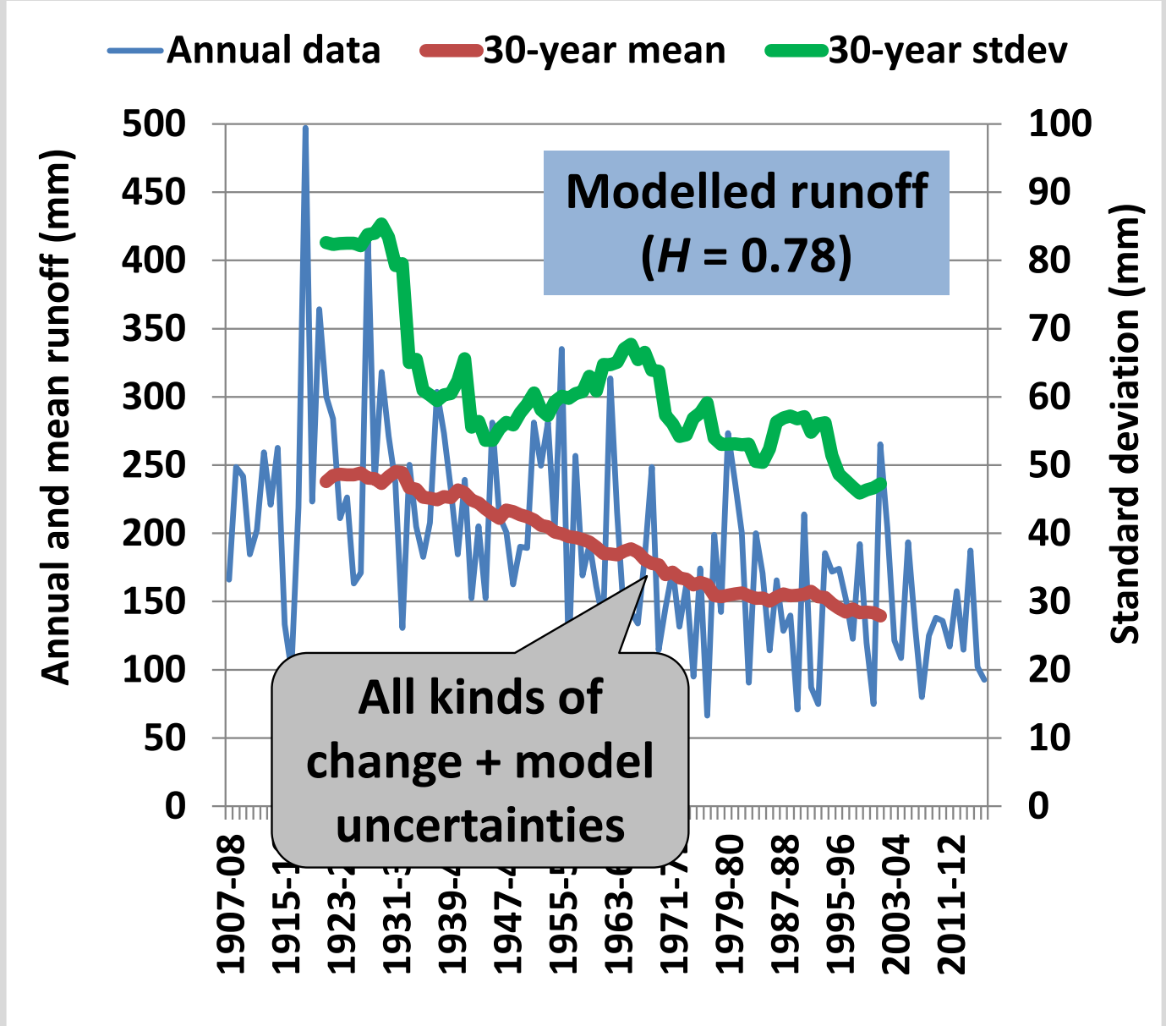
- Random yet structured changes, by means of trends, of long-term statistical characteristics of the observed precipitation, PET and runoff, at the climatic scale (i.e., 30-year moving average) → **Hurst-Kolmogorov (HK) dynamics**
- The HK behavior (quantified in terms of Hurst exponent, *H*) is apparent both in inputs (rainfall and PET) and the basin's response (runoff), but the long-term changes in runoff cannot be exclusively explained by the corresponding changes in input processes:
  - due to the complexity of the physical system;
  - due to the changes in the surface and groundwater flow regime induced by human interventions;
  - due to unknown changes of the physical characteristics of the basin (e.g. land cover changes).
- The HK behavior is less well represented in the simulated runoff, since the model cannot fully capture the observed variability (and associated peculiarities) at the climatic scale:
  - because the reservoir-based approach of conceptual models provides smoother responses than in reality;
  - due to uncertainties in the estimation of demands and the representation of associated abstractions;
  - due to calibration errors.

The one million question: Which part of the observed data (and their statistics) should be used for future simulations?









6. Mind the residuals!

- In the ideal hydrological modeling world, the residuals should be:
  - uncorrelated with the simulated runoff;
  - uncorrelated with themselves;
  - iid random variables, without periodicity or other kind of temporal variation in their statistical characteristics.
- In the real world, **the model residuals inherit the changing statistical behavior of the input data** (cf. Solomatine & Shrestha, 2009), both systematic (periodicity) and non-systematic (HK dynamics).

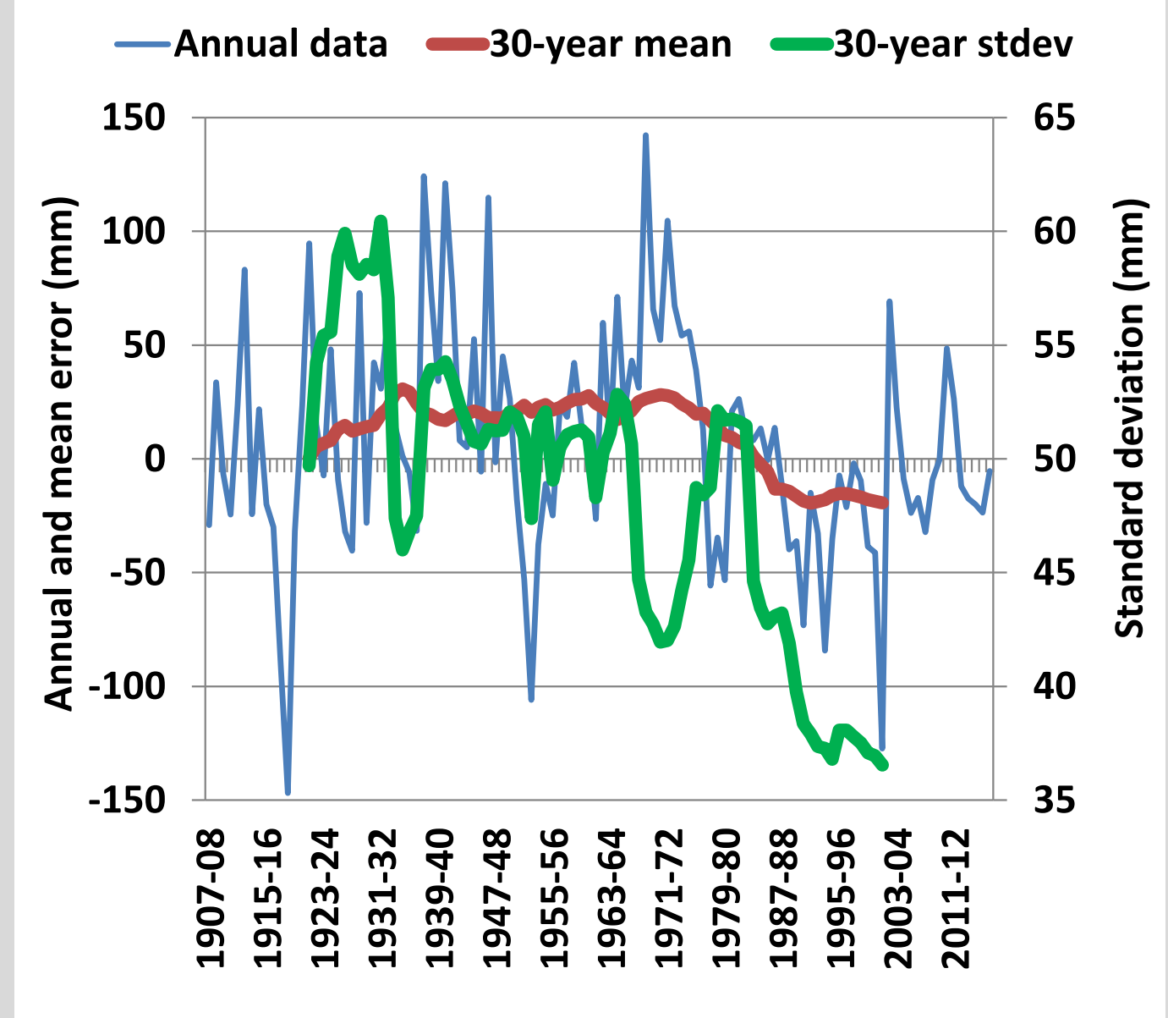
Statistical characteristics of model residuals (stationary approach)

Mean	0.538
St. dev.	7.591
Skewness	1.059
Autocorrel.	0.446
Cross-correl.	-0.099

At least one good news: the residuals are not correlated with runoff

Statistical characteristics of model residuals across seasons (cyclostationary approach)

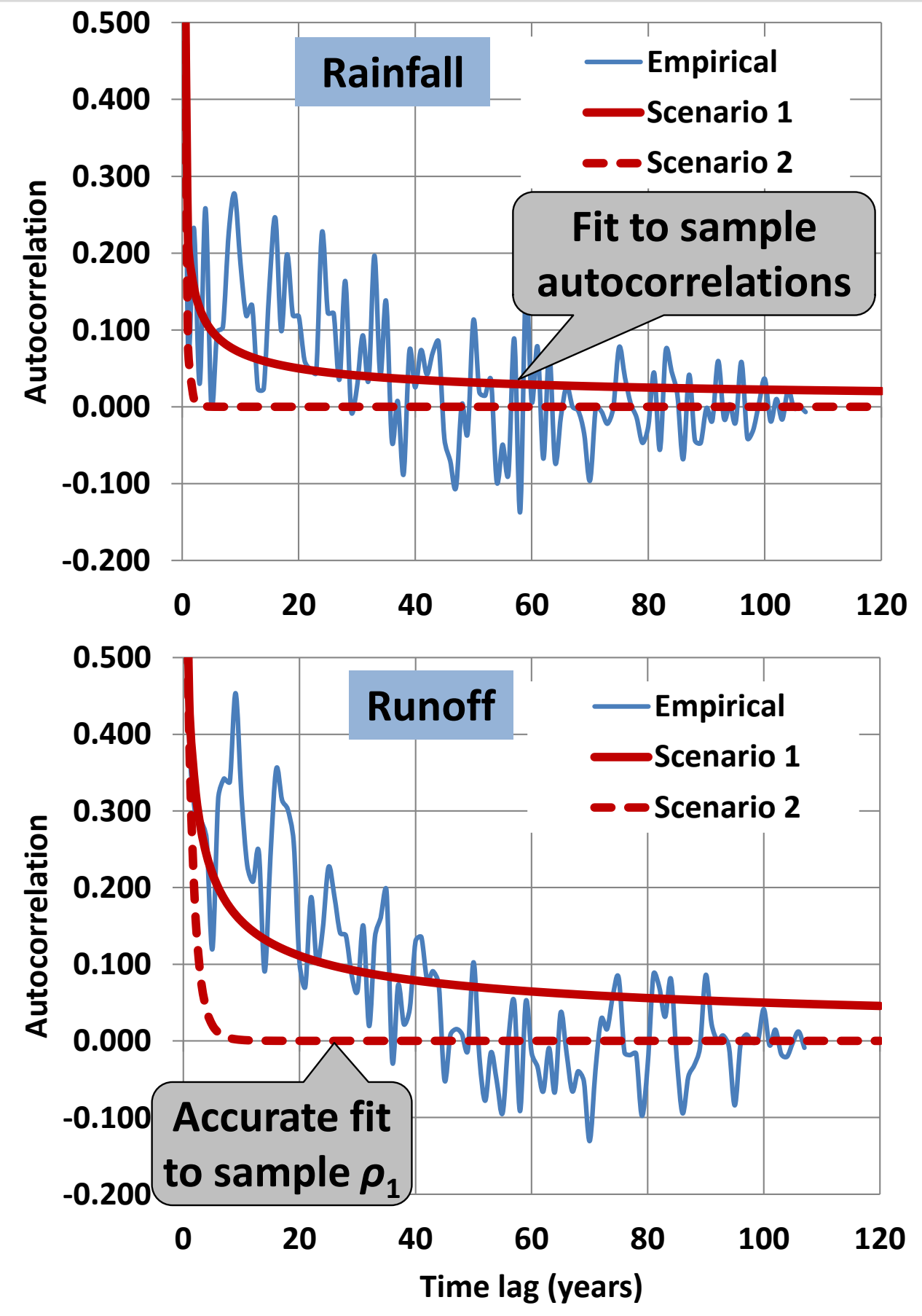
	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUNE	JULY	AUG	SEP
Mean	0.605	-0.782	-0.544	-0.511	-0.838	1.713	-0.435	-0.910	2.137	1.209	1.331	3.487
St. dev.	5.598	7.345	11.133	10.135	9.505	10.232	8.173	6.335	4.672	3.617	2.598	4.348
Skewness	-0.810	-0.460	2.341	0.519	1.007	0.808	1.648	0.643	1.591	4.585	2.636	0.339
Autocorrel.	0.397	0.378	0.497	0.573	0.394	0.328	0.573	0.440	0.531	0.566	0.498	0.500



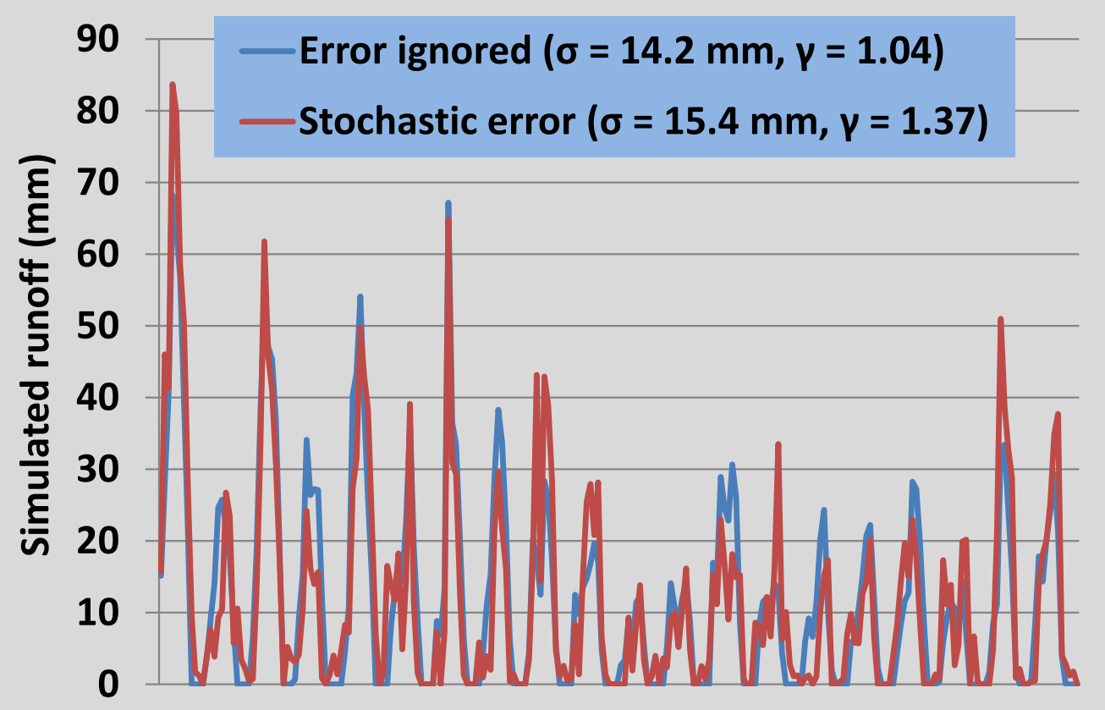
The changing statistics of the observed data, at all temporal scales, are reflected into the model residuals

7. Combined stochastic-deterministic modelling

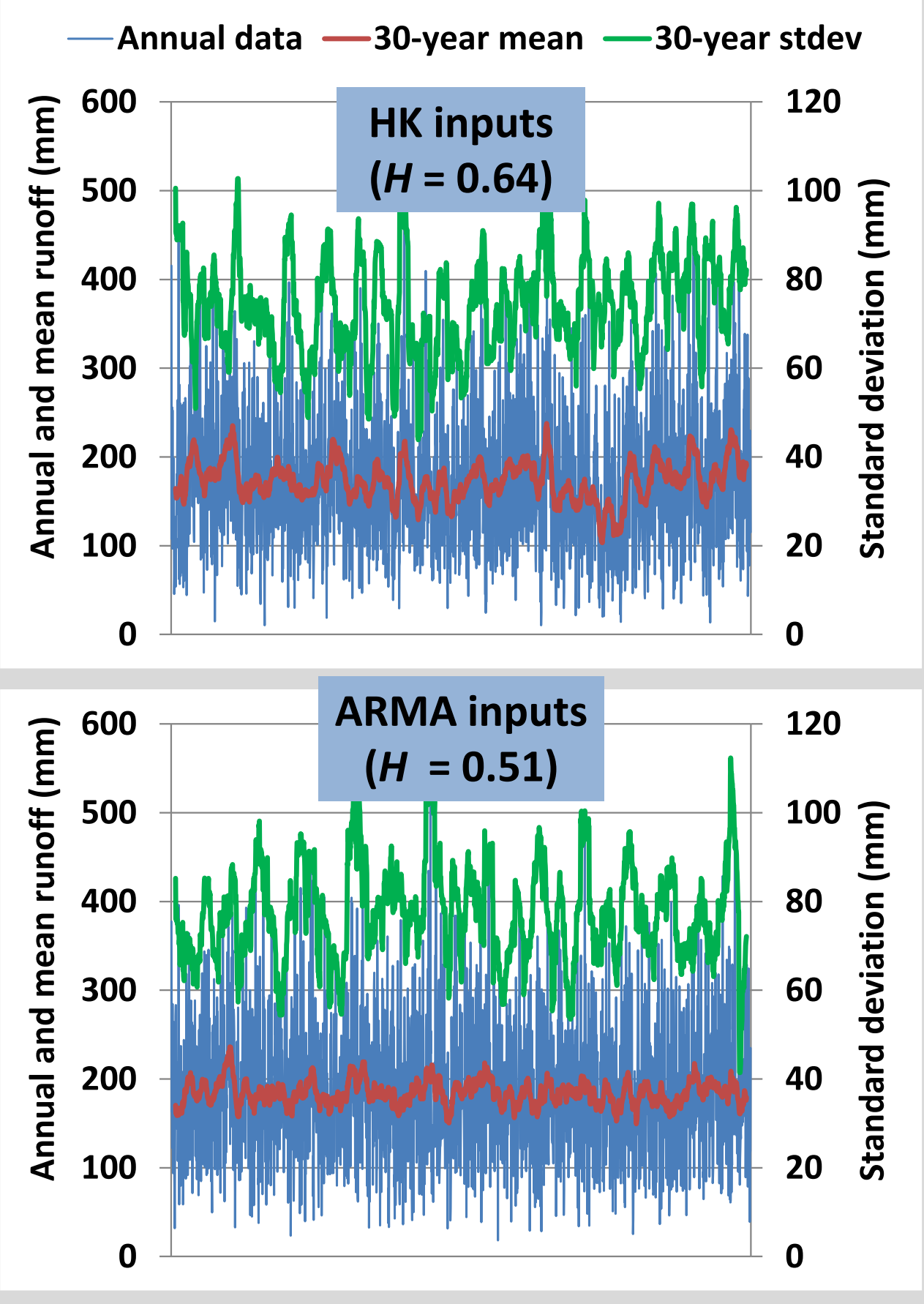
- Multivariate generation of synthetic rainfall, PET and runoff data of 2000 years length through CastaliaR model (Efstratiadis *et al.*, 2014), considering **two stochastic scenarios**, with high (HK) and low (ARMA-type) persistence.
- The representation of the desirable stochastic structure is implemented by means of a **theoretical autocorrelation function**, assigned to the annual processes (Koutsoyiannis, 2000).
- The rainfall and PET scenarios are used as inputs to the hydrological model, to provide synthetic runoff data under **stationary demand** (mean values of years 1970-2017), by adding or not a **stochastic error term**.
- The error is generated through a cyclostationary AR(1) model with gamma noise, which reproduces the monthly means and variances.
- The deterministic model without error term underestimates the seasonal variance and skewness of runoff.
- If the model is fed with ARMA-type synthetic data, it fails to capture the long-term (i.e. climatic) variability.
- Coupling a deterministic model with a stochastic error term and HK synthetic inputs ensures realistic representation of variability at all temporal scales, i.e. seasonal, annual and over-annual.



Empirical vs. theoretical autocorrelation structures for the two simulation scenarios



Two modelling approaches (without and with error term) for 20 years of synthetic data



Annual data and moving average statistics for the two simulation scenarios (HK, ARMA)

8. Conclusions

- Deterministic hydrological models can help identify the causes of long-term runoff changes, which are **combined effect of the changing climate and environment**.
- The **stochastic representation of the error term** allows for capturing the short-term (i.e., seasonal) variability of runoff, otherwise the latter may be significantly underestimated, since the common bucket-type models tend to provide smoother responses than in reality.
- The reproduction of the **long-term variability of runoff** requires the use of synthetic inputs exhibiting a **Hurst-Kolmogorov behavior**, as a result of the perpetually changing climate.

References

- Efstratiadis, A., I. Nalbantis, A. Koukouvinos, E. Rozos, & D. Koutsoyiannis, HYDROGEIOS: A semi-distributed GIS-based hydrological model for modified river basins, *Hydrol. Earth Sys. Sci.*, 12, 989–1006, 2008.
- Efstratiadis, A., I. Nalbantis, & D. Koutsoyiannis, Hydrological modelling of temporally-varying catchments: Facets of change and the value of information, *Hydrol. Sci. J.*, 60(7-8), 1438–1461, 2015.
- Efstratiadis, A., Y. Dialynas, S. Kozanis, & D. Koutsoyiannis, A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence, *Envir. Model. Soft.*, 62, 139-152, 2014.
- Koutsoyiannis, D., A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resour. Res.*, 36(6), 1519–1533, 2000.
- Nalbantis, I., A. Efstratiadis, E. Rozos, M. Kopsiafti, & D. Koutsoyiannis, Holistic versus monomeric strategies for hydrological modelling of human-modified hydrosystems, *Hydrol. Earth Sys. Sci.*, 15, 743–758, 2011.
- Solomatine, D. P., & D. L. Shrestha, A novel method to estimate model uncertainty using machine learning techniques, *Wat. Resour. Res.*, 45, W00B11, 2009.

Contact info: andreas@itia.ntua.gr; presentation available at: http://www.itia.ntua.gr/1779/