Impact of sample uncertainty of inflows to stochastic simulation of reservoirs

EGU General Assembly 2019, Vienna, Austria, 7-12 April 2019; Session HS5.3.1/ERE2.8: *Advances in modeling and control of environmental systems: from drainage and irrigation to hybrid energy generation*

E. Zacharopoulou, I. Tsoukalas, A. Efstratiadis, and D. Koutsoyiannis

Department of Water Resources & Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Greece

1. Abstract

Design and management of water resource systems are arguably challenging tasks, as they are mainly driven by hydrological processes that are dominated by **"structured" randomness**. In this vein, the **stochastic simulation** of the input processes is regarded an essential component for such studies.

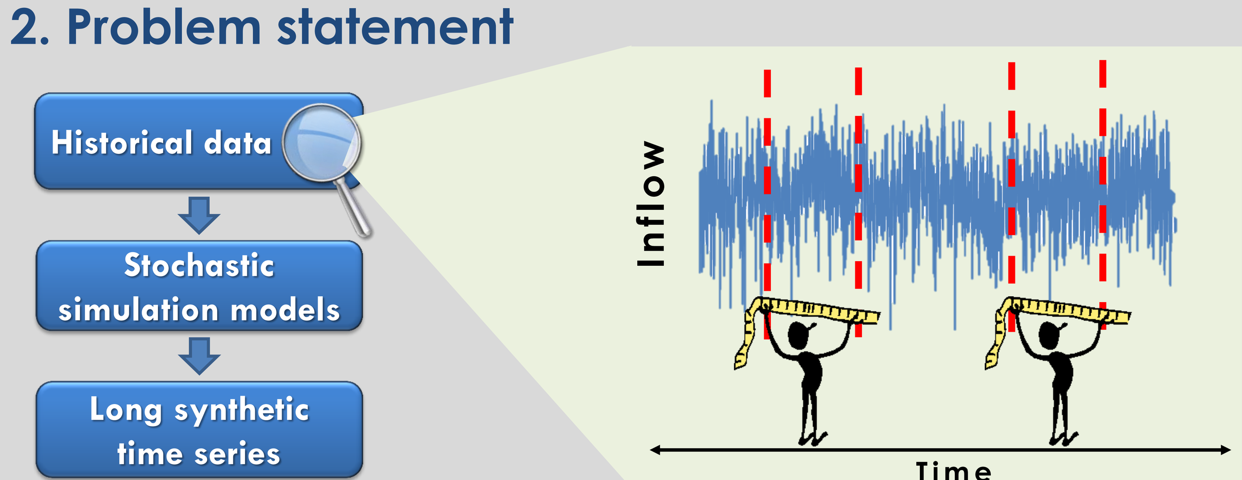
Typically, the objective of stochastic models is the generation of **long synthetic time series** that reproduce the **statistical and dependence properties** of the historical data, ideally at multiple time scales (including long-term changes, such as those induced by the Hurst-Kolmogorov behavior). However, the **sample statistical characteristics** that are forced to be reproduced entail an **inherent uncertainty**, due to the generally **short length of historical data**. This key shortcoming is not typically accounted for within the current practices.

This work is an attempt to investigate and quantify the **input uncertainty within stochastic models**, and eventually assess its impact on reservoir systems. Towards this, we establish a methodology for the quantification of the sample uncertainty, involving the essential statistical characteristics of historical inflows in a multiscale context, by using as **background stochastic simulator** the **CastaliaR model**.

Initially, this model is employed for the generation of a large set of synthetic time series with the same length with the historical sample, and thus provide multiple "pseudo-historic" realizations. Subsequently, the statistical properties of the ensemble of pseudo-historic data are extracted and employed to generate long synthetic time series, which are finally used as **inputs to a reservoir simulation model**.

The above procedure is demonstrated for the derivation of **ensembles of storage-yield-reliability relationships**. Furthermore, multiple analyses for different sample sizes and Hurst coefficients are performed, aiming to investigate the uncertainty imposed by the **sample size** and the **long-term persistence** of the inflow processes.

2. Problem statement



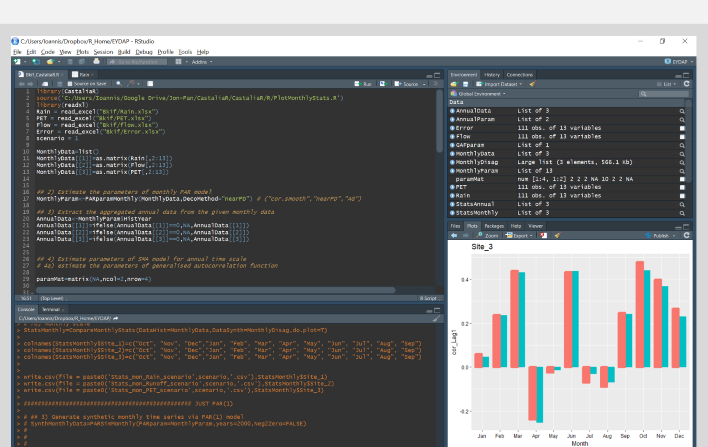
QUIZ: What would be **the response of the system**, if we use another **random sequence** with the same **length** as the **historical sample in use**?

Reservoir simulation model

Derive the **storage-yield-reliability** relationship for:

- Design purposes (estimation of reservoir capacity)
- Water management purposes (estimation of reliability for a given demand and vice versa)

3. CastaliaR package



- R-based, open-source implementation of a state-of-the-art framework for multivariate stochastic simulation of hydro-meteorological processes; its background builds upon the works of Koutsoyiannis and Manetas (1996), Koutsoyiannis (2000) and Efstratiadis et al. (2014).
- Reproduces the essential statistical characteristics (marginal and joint-moments) of the historical data at three temporal scales (annual, monthly and daily) as well as the long-range dependence (Hurst-Kolmogorov behavior) at over-annual scales.
- The generation procedure lies upon a symmetric moving average process for the annual scale and a periodic autoregressive process for the finer scales, while a Monte Carlo disaggregation approach re-establishes consistency across the three temporal scales.
- CastaliaR is fully automated, providing a user-friendly time series generation package for engineers and researchers (Tsoukalas et al., 2018a).

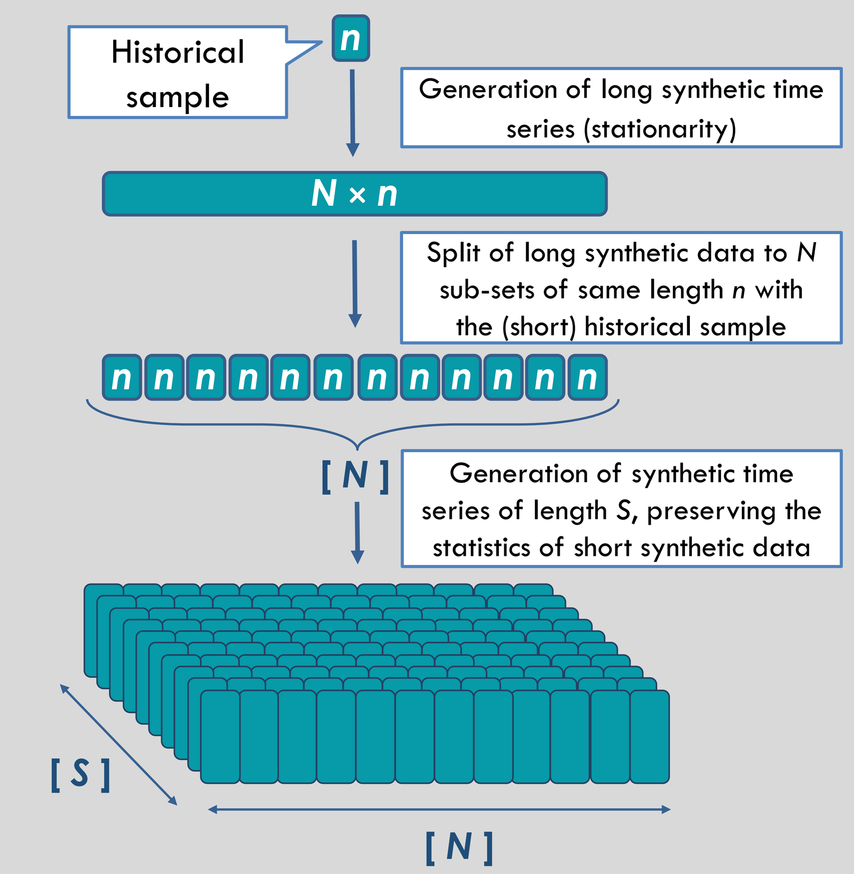
4. Evaluation of sample uncertainty within stochastic simulation

There are many difficulties in quantifying the sample uncertainty of the **essential statistical characteristics** (marginal statistics: average, standard deviation, skewness; joint-statistics: lag-one auto- and zero-lag cross-correlations) of time series, that are typically reproduced through stochastic simulation.

Our goal is the evaluation of **sample uncertainty** induced by the use of **statistical characteristics that are empirically estimated from short samples.**

In this context, we developed a **Monte Carlo based approach** aiming to generate "pseudo-historic" realizations comprising the following steps:

- **Step 1**: Generation of **long synthetic time series with length** $N \times n$, that reproduce the statistical and dependence properties of the historical data, where n is the length of historical sample and N is the selected number of Monte Carlo analyses. In this way, **stationarity** is established.
- **Step 2**: Split of long time series to N sub-sets with same length with the historical sample (n years), using a moving time-window approach. These **short synthetic time series** represent the **multiple possible samples** ("pseudo-historic" realizations).
- **Step 3**: Generation of N **synthetic time series of length** S, where S is the years of simulation. These time series reproduce the statistical and dependence properties of another short synthetic time series each time.



5. Case study: Stochastic analysis of inflow time series at Kremasta reservoir (Western Greece)

Modelling assumptions

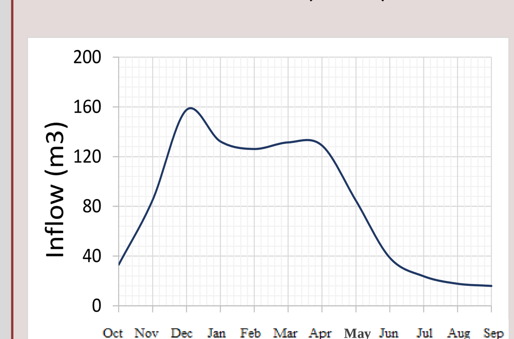
Sample size (n):
- 28 years (historical sample)
- 20 years
- 50 years

Monte Carlo analyses (N): 100

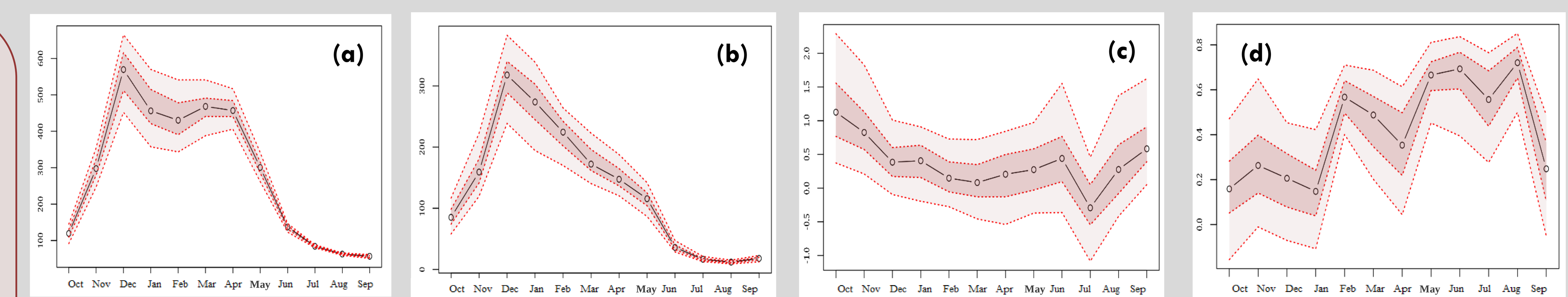
Simulation (S): 1000 years

Hurst coefficient (H):
- H=0.5
- H=0.7
- H=0.9

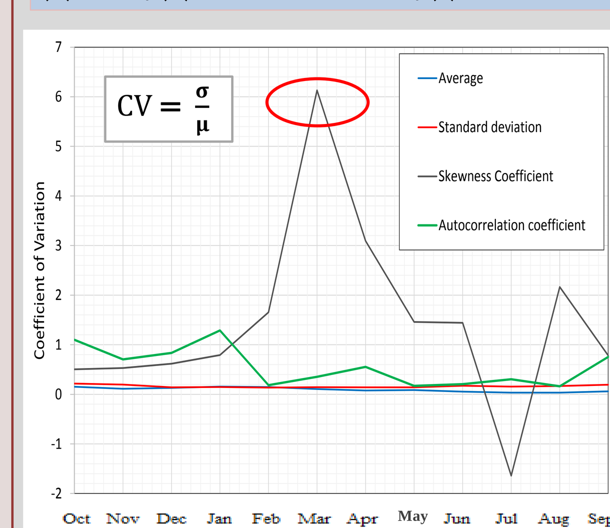
Total scenarios: 9 (3×3)



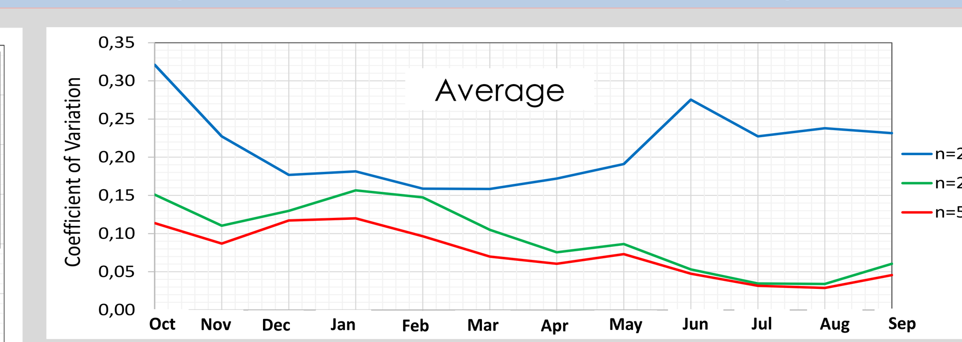
Mean monthly inflow, as estimated from the historical data (28 years)



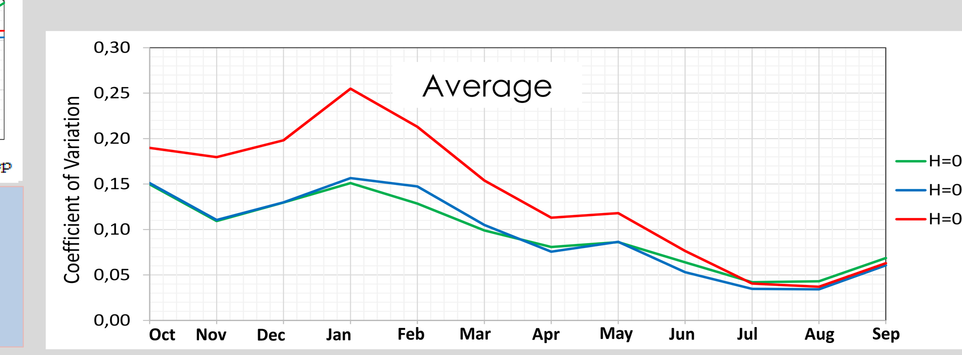
Quantiles (5%, 25%, 50%, 75%, 95%) of statistical characteristics of generated time series (pseudo-historic), based on historical sample size n = 28 years, for H = 0.70: (a) mean; (b) standard deviation; (c) skewness coefficient; (d) lag-1 autocorrelation coefficient. All statistics exhibit significant variability (even for the mean).



Coefficient of variation comparisons for the essential statistical characteristics of mean inflow (n = 28, H = 0.70). Significantly higher value for skewness coefficient.



Variation of CV of the average of mean inflow, for different sample sizes and H = 0.70. As expected, longer sample decreases the coefficient value. Similar behavior is observed for all essential characteristics.



Variation of CV of the average of mean inflow, for different Hurst coefficients (n = 28). As expected, long-term persistence increases the CV value. Similar behavior is observed for all essential characteristics.

6. Implementation on reservoir simulation model

- **Goal :** Estimation of the useful storage capacity of the reservoir, for multiple demand targets and associated reliability levels
- **Key assumptions:**
  1. Constant water demand (stationary conditions)
  2. Max annual demand ≤ Mean annual inflow
  3. Failure (deficit) of one month accounts as failure of the entire year (annual reliability)
- **Methodology:**
  1. Scenarios development
     - i) Demand scenarios
     - ii) Reliability scenarios
  2. Implementation of reservoir simulation model
  3. Derivation of storage-yield-reliability relationship

Problem setting

Demand scenarios: 12
- Expressed in terms of ratio of max demand (adjusted demand)
- Max demand = Mean annual inflow
- Starting point $D = -\frac{\mu}{\sigma}$, corresponding to 100%
- Increase by step 0.1

Reliability scenarios: 5

Reservoir model Implementation:
- R programming
- Input inflow: 100 time series of 1000 years length, generated through the CastaliaR program.
- Multiple scenarios for different sample sizes (n = 28, 20 and 50 years) and Hurst coefficient (H = 0.50, 0.70, 0.90).

Totally simulated scenarios: 540 (12×5×9)

Table 1: Demand scenarios

100.00 %	91.87 %	83.74 %	75.61 %
67.48 %	59.35 %	51.22 %	43.09%
34.96 %	26.83%	18.70 %	10.57%

Table 2: Reliability scenarios

80%	85%	90%	95%	99%

Storage-yield-reliability analysis

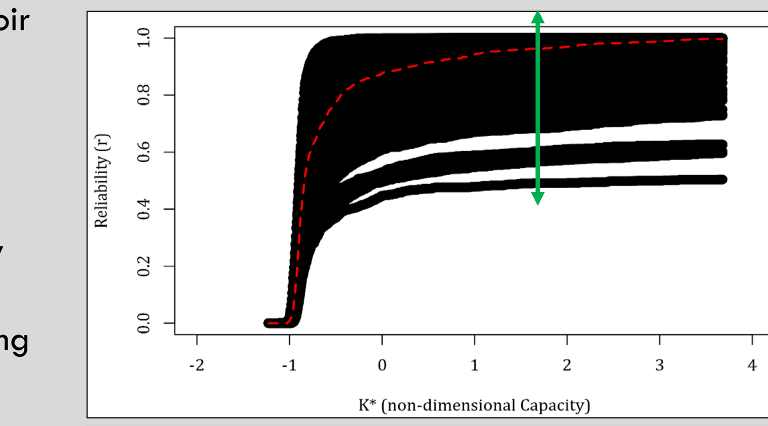
Non-dimensional storage capacity $K^* = \frac{K-\mu}{\sigma}$

⇩

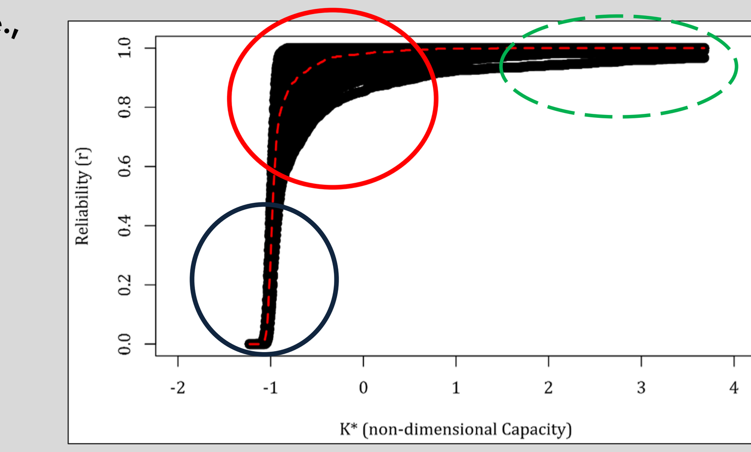
easier comparisons

7. Reservoir capacity vs. reliability relationships

- Typical requirements of reservoir sizing problems: **maximize demand, under a high reliability level.**
- Significant uncertainty in the estimations of storage capacity for scenarios around the max water demand, when considering reliability levels around 95 to 99% (red circle).
- Uncertainty is reduced for very high storage capacity values (i.e., much above the mean annual inflow) that reasonably ensure high reliability levels (green circle), as well as for very small capacities, which yet result to sharp decrease of reliability (blue circle)
- Too large reservoirs or too low reliability levels are both **non-acceptable, for socio-economic and environmental reasons.**



Storage-reliability curve for 91.87% of max demand



Storage-reliability curve for 75.61% of max demand

- "Classical" long-term stochastic simulation approach (displayed with red line), as far as the time series generation is concerned, gives quite conservative estimates, well above the average scenario.
- Similar outcomes for different sample sizes and Hurst coefficients. As the Hurst coefficient increases, the deviation from the average scenario decreases as well.

8. Discussion & future research

- We developed a generalized methodology to quantify the impacts of sample uncertainty within a typical stochastic simulation exercise, i.e. the derivation of the storage-yield-reliability relationship of a single-purpose reservoir.
- The outcomes of our analyses may be characterized **shocking**, since key design and management quantities that have been traditionally estimated via stochastic approaches (i.e. the use of synthetic data) are substantially affected by the uncertainty induced from short length of historical data.
- In the context of storage-yield-reliability estimations, the uncertainty is amplified in the area of interest of practical applications, i.e. reservoir sizing for water demands little less than the mean annual inflow, to be fulfilled with high reliability levels.
- The existence (or not) of long-range dependence at the annual scale (as quantified by the Hurst coefficient) has a significant impact in the analysis of reservoir's design.
- Future research will be focused to:
  - Use of newly emerged synthetic data generators, that beyond marginal and joint moments are capable of generating time series with any marginal distribution and correlation structure (Tsoukalas *et al.*, 2018b,c).
  - Application to multi-reservoir multi-purpose systems, involving the simulation of spatially-correlated processes.
  - Application to hydroelectric reservoirs, considering firm energy demands.

References

- 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.
- Koutsoyiannis, D., & A. Manetas, Simple disaggregation by accurate adjusting procedures, *Water Resour. Res.*, 32(7), 2105-2117, 1996.
- Tsoukalas, I., P. Kossieris, A. Efstratiadis, C. Makropoulos, & D. Koutsoyiannis, CastaliaR: An R package for multivariate stochastic simulation at multiple temporal scales, EGU General Assembly 2018, 2018a.
- Tsoukalas, I., C. Makropoulos, and D. Koutsoyiannis, Simulation of stochastic processes exhibiting any-range dependence and arbitrary marginal distributions, *Water Resour. Res.*, 54(11), 9484-9513, 2018b.
- Tsoukalas, I., A. Efstratiadis, & C. Makropoulos, Stochastic periodic autoregressive to anything (SPARTA): Modeling and simulation of cyclostationary processes with arbitrary marginal distributions, *Water Resour. Res.*, 54(1), 161-185, 2018c.
- Zacharopoulou, E., *The impacts of sample uncertainty of inflows to stochastic simulation of reservoirs*, Diploma Thesis, 128 p., Department of Water Resources & Environmental Engineering, NTUA, 2017.

Contact info: zaharel94@gmail.com; **Presentation available at:** http://www.itia.ntua.gr/1938/