European Geosciences Union General Assembly, Online, 4-8 May 2020

HS3.3/ERE6: Stochastic modelling and real-time control of complex environmental systems

Stochastic modelling of hydropower generation from small hydropower plants under limited data availability: from post-assessment to forecasting

Georgia-Konstantina Sakki, Vassiliki-Maria Papalamprou, Ioannis Tsoukalas, Nikos Mamassis, and Andreas Efstratiadis

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

Presentation available online: www.itia.ntua.gr/2022/

Problem setting

- Due to their negligible storage capacity, **small hydroelectric plants (SHPs) cannot offer regulation of flows**, thus making the prediction of energy production a very difficult task, even for small time horizons.
- Further uncertainties arise due to the limited hydrological information, in terms of upstream inflow data, since usually **the sole available measurements refer to the power production**, which is a **nonlinear transformation of the river discharge**.
- This transformation comprises several uncertain elements, including the estimation of energy losses by using empirically-derived **efficiency curves**.
- The retrieval of flows from energy data may be referred to as the **inverse problem of hydropower**, which is the focus of this research.
- The inverse modelling problem involves three flow ranges:
  - Low flows, below the minimum operational discharge of turbines;
  - Intermediate flows, which are directly estimated on the basis of observed hydropower data.
  - High flows, exceeding the nominal discharge of turbines;
- In all cases, the **model error** is expressed in **stochastic terms**, which allows for embedding uncertainties within calculations (Efstratiadis *et al.*, 2015).
- These uncertainties are next transferred to **energy predictions that are based on imperfect past flow data**.

The forward problem: from discharge to power

- Given data for small hydroelectric plants (SHPs) :
  - Streamflow upstream of the intake, q;
  - Gross head, h (practically constant);
  - Power plant efficiency, η, expressed as function of discharge;
  - Maximum discharge that can pass from the turbines (nominal flow), q_{max}
  - Minimum discharge for energy production, q_{min} (typically, 10-30% of q_{max})
- Flow passing through the turbines:

$$q_T = min(q, q_{max})$$

- Power produced for $q_T > q_{min}$:

$$P = \gamma\, \eta\, q_T\, h_n$$

where γ is the specific weight of water (9.81 KN/m^3) and h_n is the net head, i.e. the gross head, h, after subtracting hydraulic losses, h_L.

- Hydraulic losses include friction and local ones, which are function of discharge and the penstock properties (roughness, length, diameter, geometrical transitions).
- Large hydroelectric reservoirs allow for controlling outflows, thus their turbines are normally working with the nominal flow (which maximizes η). In contrast, SHPs are operating with any flow conditions, thus η is strongly varying across the feasible flow range (q_{min}, q_{max}).

The inverse problem: from power to discharge

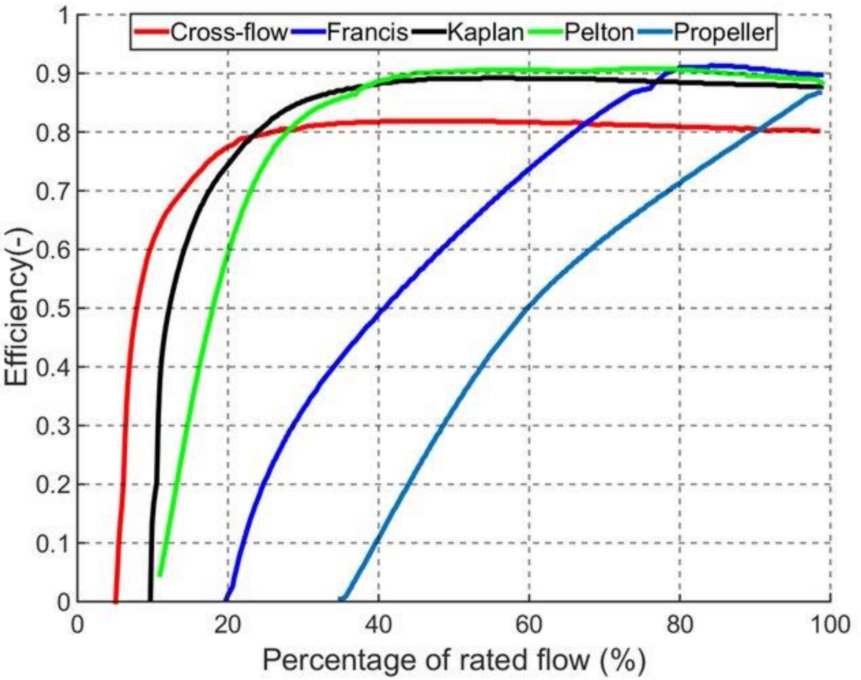
- Inverse formula, for a given power production *P*:

$$q_T = \frac{P}{\gamma\, \eta(q_T)\, h_n(q_T)}$$

- The unknown flow, q_T, that passes through the turbines can be estimated through an **iterative numerical scheme**, accounting for nonlinearities induced by efficiency and net head formulas, $\eta(q_T)$ and $h_n(q_T)$, respectively.
- Since $q_{min} \leq q_T \leq q_{max}$, this approach only allows for estimating the intermediate part of an inflow time series, thus:
  - If the power production is zero, then $q \leq q_{min}$;
  - If the system produces its power capacity (thus operating with its nominal discharge, which also ensures maximization of efficiency), then $q \geq q_{max}$;
- Measurement errors and uncertainties within any element of the governing formula $q_T = f(P)$ are transferred to discharge estimations, while low and high flows remain by definition unknown.
- Potential **sources of uncertainty**:
  - Power data per se (observational errors);
  - Hydraulic calculations (become less important, as the gross head increases);
  - Flow-efficiency relationship;

Some remarks on turbine efficiency

- The efficiency curve for specific turbine dimensions (e.g., diameter runner) is usually expressed by means of **nomographs**, as percentage of rated flow, q_T/q_{max} (Anagnostopoulos & Papantonis, 2007).
- Nomographs are provided by the turbine manufacturer and they are obtained by data extrapolation from a **reduced scale model**. Since it is not possible to exactly preserve dynamical, geometrical, and kinematical similarity between the model and the prototype, it is also not possible to precisely estimate the efficiency.
- Although empirical corrections are employed to better reflect the **prototype performance**, actual efficiency is unknown, since it also depends on constructive and operational characteristics of the power plant, as well as changes due to deterioration, damage and aging of the equipment over time (Paish, 2002).
- In general, **efficiency increases with scale**, i.e. discharge and turbine size.
- Pelton, Crossflow and Kaplan machines retain high efficiency even when running below their design flow; in contrast the efficiency of Francis turbines falls away sharply if run at below half its normal flow.



Analytical formula for turbine efficiency

- The **efficiency-discharge** relationship can be well-approximated by the following analytical formula, inspired by the Kumaraswamy distribution model:

$$\eta = \eta_{min} + \left(1 - \left(1 - \left(\frac{q - q_{min}}{q_{max} - q_{min}}\right)^a\right)^b\right)(\eta_{max} - \eta_{min})$$

- The efficiency formula uses a dimensionless expression of discharge, based on q_{min} and q_{max}, two efficiency limits, η_{min} and η_{max}, and two shape parameters, a and b.
- We remark that the efficiency curve has in fact four free parameters, since for a given power capacity P we get:

$$q_{max} = \frac{P}{\gamma\, \eta_{max}\, h_n(q_{max})}$$

$$q_{min} = \frac{P}{\gamma\, \eta_{min}\, h_n(q_{min})}$$

- By tuning these parameters we can fit the model to any empirically-derived curve, and we can also establish a **calibration framework**, to extract efficiency curves from given power and turbine flow data (cf. Hidalgo *et al.*, 2014).
- Another major advantage is the opportunity for expressing **efficiency under uncertainty**, by considering the four model parameters as random variables that follow a known distribution function.

Discharge retrieval from hydropower data

1. Computation of turbine flows for time steps $t = 1, \ldots, n$, by using the (deterministic) inverse formula:

$$q_t = f(P_t)$$

2. Estimation of model residuals, by comparing with real discharge data:

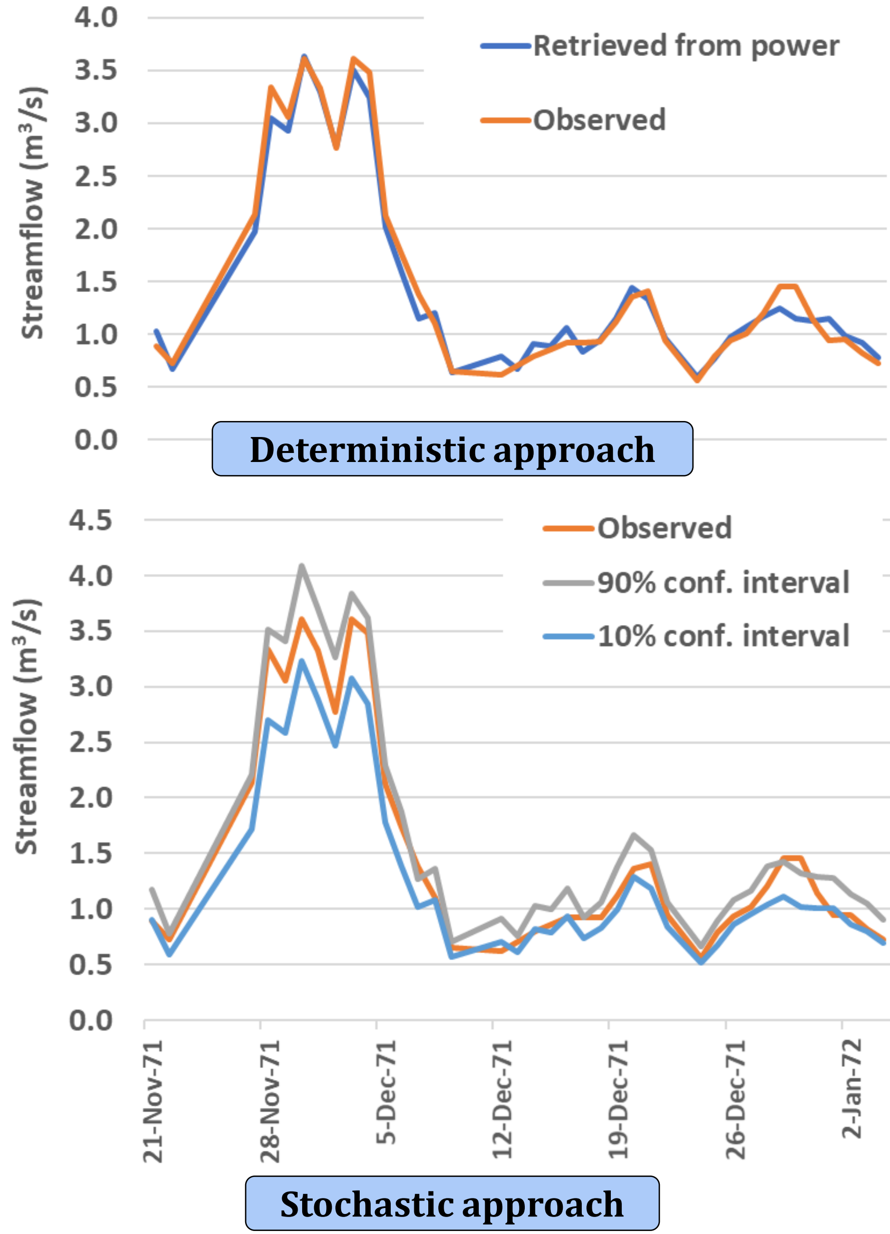
$$w_t = q_{T,t} - q_{obs,t}$$

3. Formulation of stochastic model for residuals, accounting for their marginal and dependence properties.

4. Generation of m synthetic error realizations ("ensembles") and associated discharge scenarios for each ensemble $j = 1, \ldots, m$:

$$q_{t,j} = f(P_t) + w_{t,j}$$

5. Empirical estimation of confidence intervals for each time step t, using the sample of synthetic flow data, $q_{t,j}$.



Stochastic modelling of errors

- The representation and synthesis of model residuals w_t is employed through a first order autoregressive model, AR(1), i.e.:

$$w_t = \varphi\, w_{t-1} + z_t$$

where w_t is the error process, with mean μ, standard deviation σ, skewness γ, and lag-1 autocorrelation coefficient ρ; $\varphi = \rho$ is the first order autoregression coefficient; and z_t is i.i.d. white noise with mean μ_z, standard deviation σ_z and skewness coefficient γ_z.

- The statistical characteristics of the white noise z_t are related with those of w_t by:

$$\mu_z = \mu_w\,(1-\varphi) \qquad \sigma_z = \sigma_w\sqrt{1-\varphi^2} \qquad \gamma_z = \gamma_w \frac{1-\varphi^3}{(1-\varphi^2)^{3/2}}$$

- We assume that z_t follows a three-parameter gamma distribution:

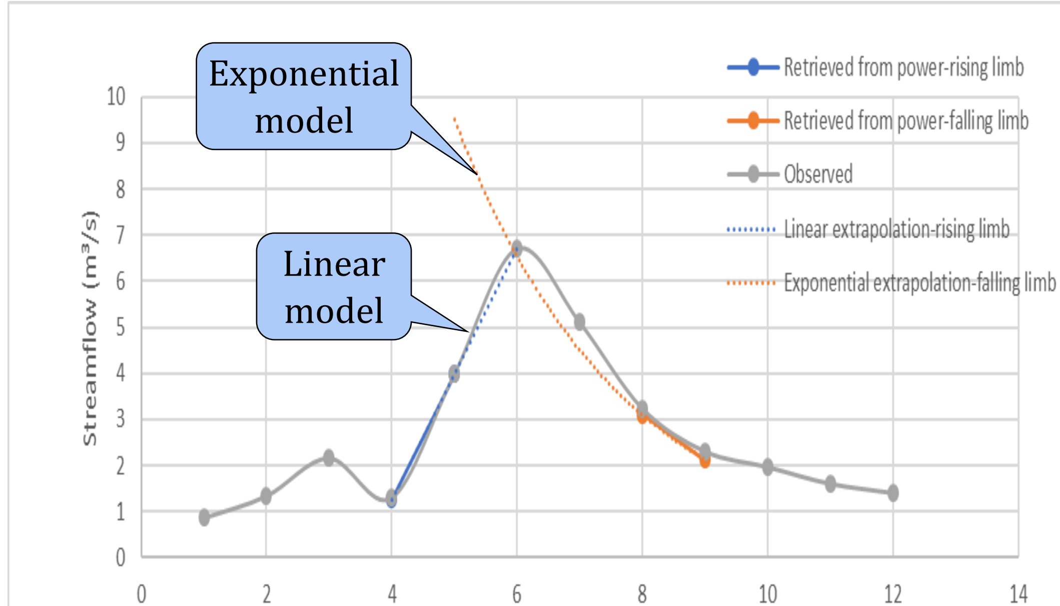
$$f_x(x) = \frac{\lambda^\kappa}{\Gamma(\kappa)}(x-c)^{\kappa-1}e^{-\lambda(x-c)}$$

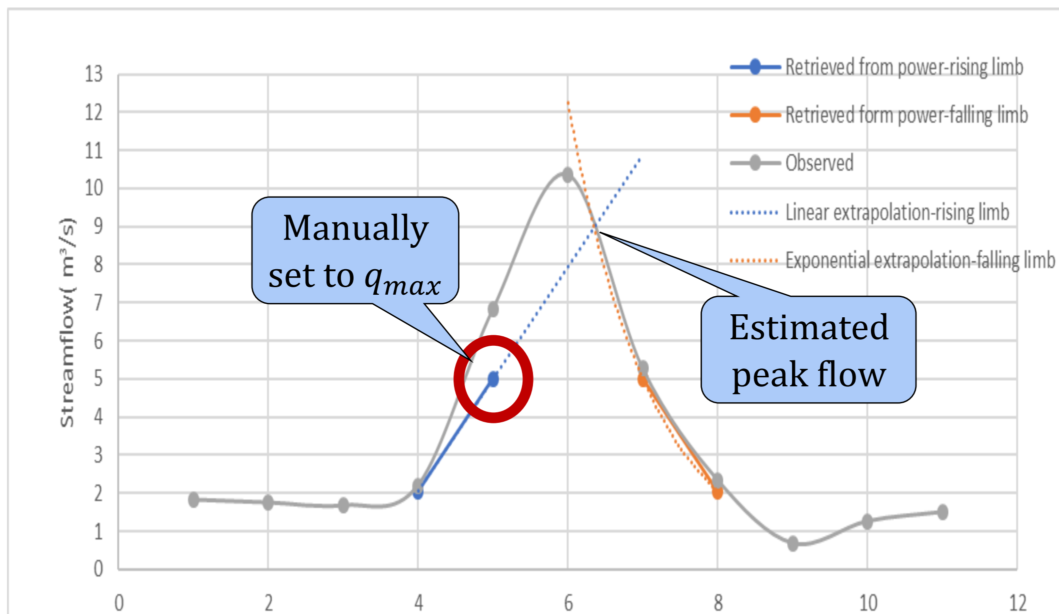
where κ, λ and c are shape, scale and location parameters, respectively, which in this case are estimated by the method of moments as follows:

$$\kappa = \frac{4}{{\gamma_z}^2} \qquad \lambda = \frac{\sqrt{\kappa}}{\sigma_z} \qquad c = \mu_z - \kappa/\lambda$$

Extrapolation for high flows

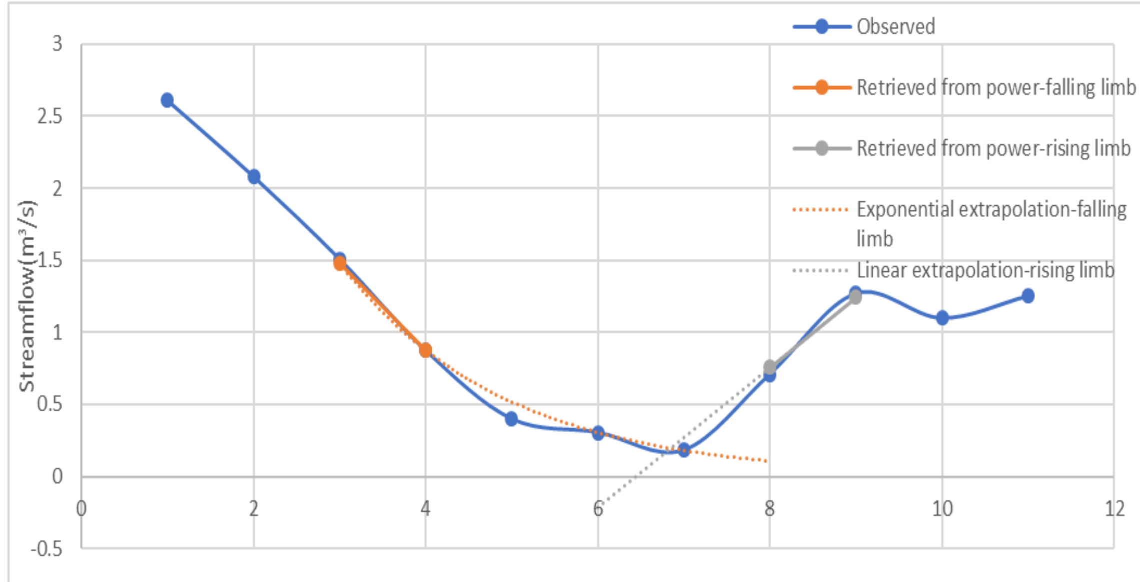
- Hydrograph extrapolation for $q > q_{max}$, indicating periods that the **turbines operate in their maximum capacity**, thus the flow passing is q_{max}.
- **Linear extrapolation** for the rising limb, by linking forward the last two known discharge values; slope is adjusted to ensure that all estimated discharge values exceed q_{max}.
- **Exponential extrapolation** for the falling limb, by linking backward the first two known discharge values, which ensures a recession rate that is representative of the flood propagation over the basin.
- Peak flow appears in their intersection.

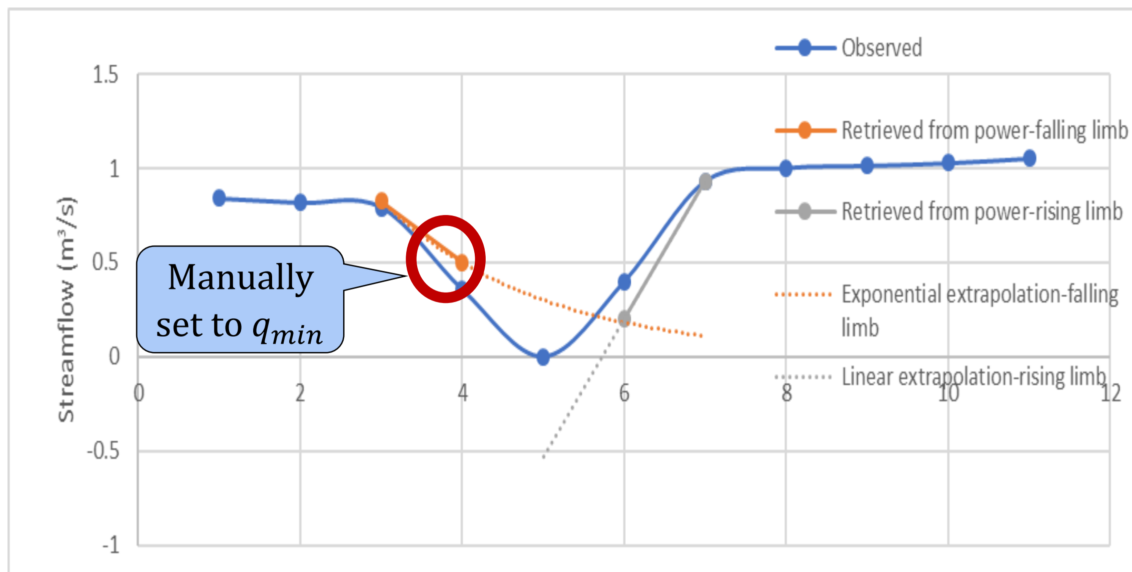




Extrapolation for low flows

- Hydrograph extrapolation for $q < q_{min}$, indicating periods that the turbines do not operate, thus the **power production is zero**.
- Exponential extrapolation for the falling limb, based on the last two known discharge values.
- Linear extrapolation for the rising limb, by linking backwards the first two known discharge values.
- Adjustment to ensure that all estimated discharge values do not exceed q_{min}.
- <u>**Important hint**</u>: Different error models are established for low, high and intermediate flows.





Theoretical example

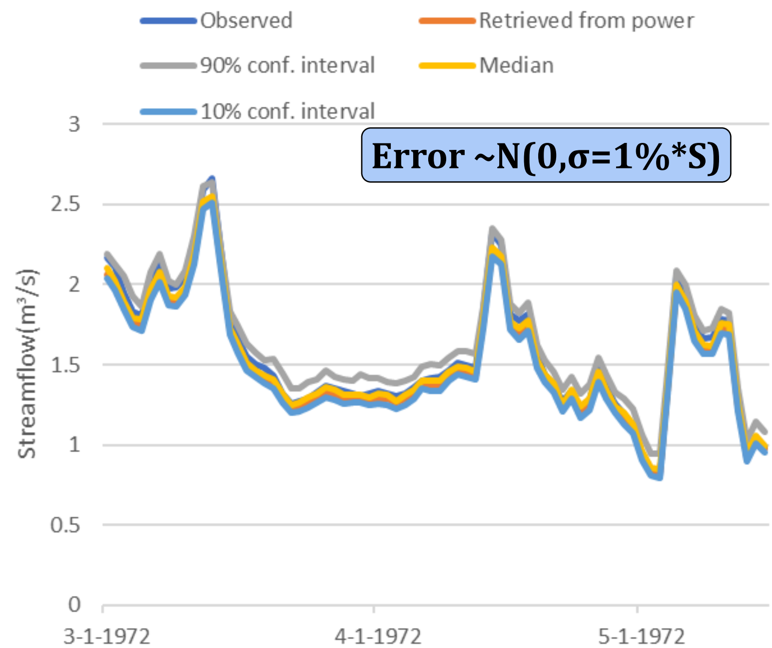
- Hypothetical small hydroelectric plant, with known daily inflows (10 year data), comprising a single turbine of 10.8 MW power capacity.
- Net head is considered constant, i.e. $h_n = 260$ m.
- Two alternative turbines are considered, i.e., Pelton or Francis, operating at low flow limits 10 and 20%, respectively, and having different efficiency curves that are expressed through the four-parametric analytical function.
- Forward problem: estimation of daily energy data generated by each turbine type
- Inverse problem: retrieval of daily flows by assigning two artificial error expressions:
  - random perturbation of energy generation data, by assigning an additive error term to simulated energy that follows either a normal or a skewed (Gamma) distribution, thus accounting for **observation errors**;
  - extraction of discharge data by using a set of 100 randomly generated efficiency curves around the actual ones, to represent the inherent uncertainties of the modelling procedure (**parameter errors**).
- In the first setting, the uncertain discharge data are represented in stochastic terms, i.e. by employing the AR(1) model to residuals, while in the second setting the ensembles are directly obtained by solving the inverse problem for each uncertain efficiency curve.

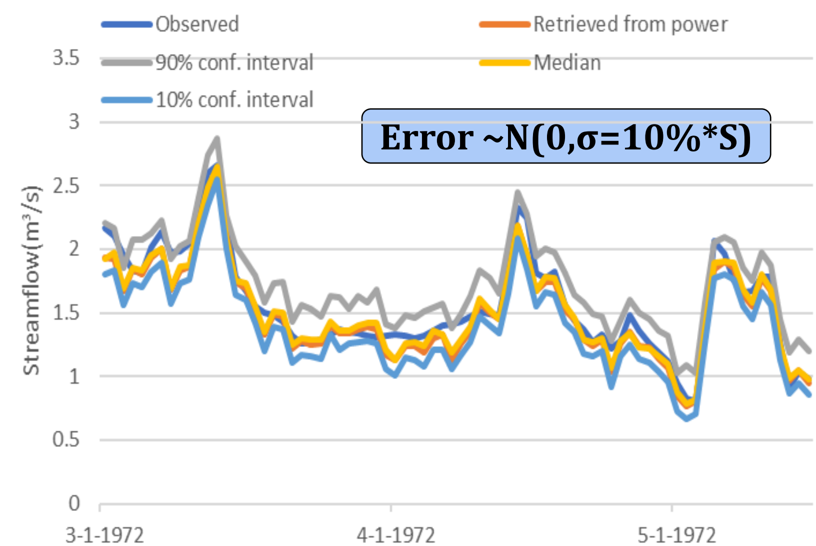
Artificial error added to simulated energy

- Uncertain energy production is expressed by adding to the actual data $e_t(q_t)$, which is obtained from known inflows q_t, the error term Δe_t, as follows:

$$e_t^* = e_t(q_t) + \Delta e_t$$

- Δe_t is expressed by means of unbiased noise, either normal $N(0, \sigma_e)$ or gamma-type, with skewness γ_e.
- σ_e is expressed as percentage of the standard deviation of simulated energy production, i.e. 1%, 5% and 10%.
- The uncertainty of the inflows that are retrieved by the inverse procedure is quantified in terms of key statistical characteristics of residuals:
  - mean, variance, skewness
  - lag-one autocorrelations
  - cross-correlations with actual flow data (heteroscedasticity?)





Transforming *a priori* errors (assigned to energy) to *a posteriori* errors of simulated discharge

Statistical characteristics of simulated discharge errors, after adding a normal error term to actual energy data (zero bias, standard deviation 1, 5 and 10% of energy standard deviation)

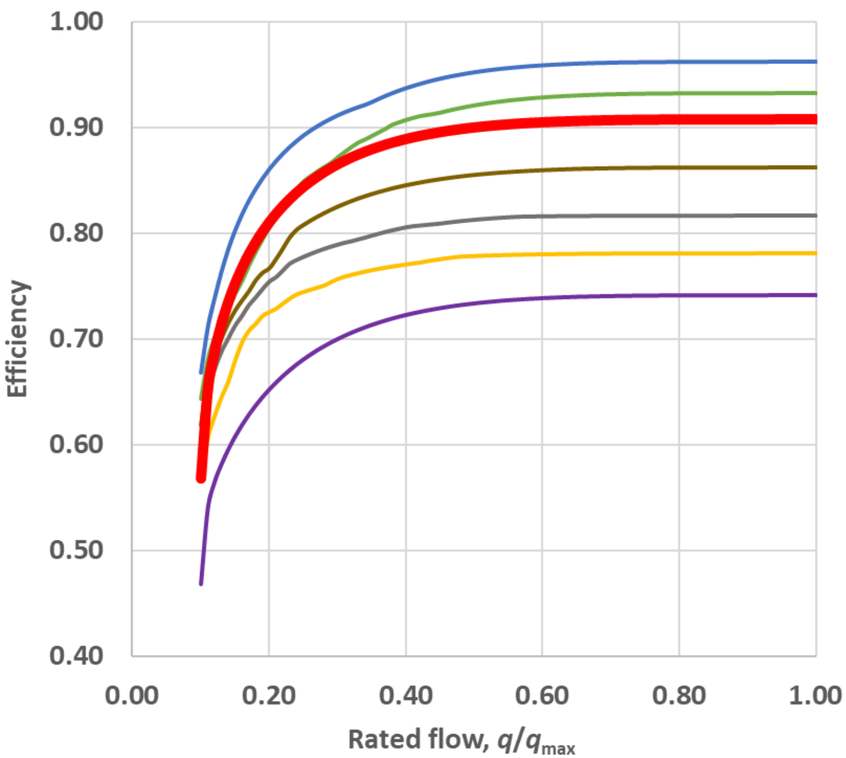
	1%		5%		10%	
	Pelton	Francis	Pelton	Francis	Pelton	Francis
mean	0.037	-0.112	0.044	-0.115	0.049	-0.088
stdev	0.065	0.193	0.100	0.196	0.139	0.118
skewness	1.411	1.213	1.968	1.225	1.154	-0.441
autocorrelation	0.619	0.769	0.243	0.736	0.125	0.703
cross-correlation	0.777	0.965	0.398	0.947	0.310	0.826

Statistical characteristics of simulated discharge errors, after adding a gamma-distributed error to actual energy data (zero bias, standard deviation 1% of energy, skewness coefficients 0.3, 1, 5)

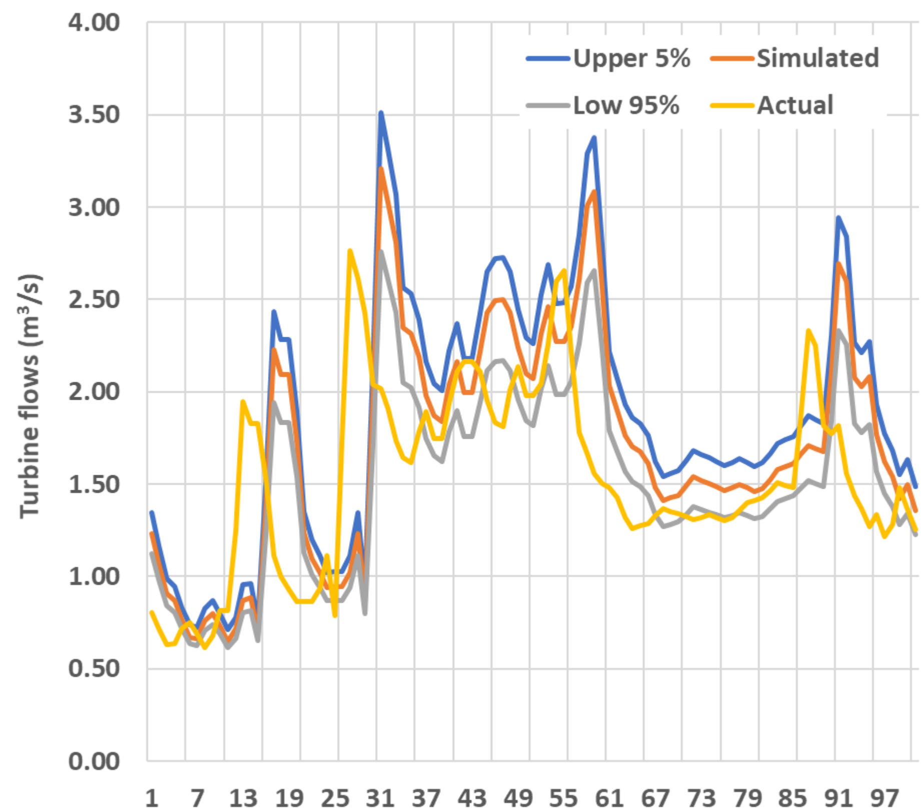
	0.3		1		5	
	Pelton	Francis	Pelton	Francis	Pelton	Francis
mean	0.037	-0.117	0.037	-0.116	0.036	-0.116
stdev	0.064	0.179	0.064	0.180	0.060	0.179
skewness	1.442	0.674	1.174	0.683	0.573	0.680
autocorrelation	0.600	0.794	0.633	0.795	0.723	0.796
cross-correlation	0.773	0.968	0.780	0.968	0.862	0.968

Inverse problem under uncertain efficiency

- **Deterministic approach**: extraction of flow data from energy, considering a Francis turbine with known efficiency curve (given in analytical form);
- **Stochastic approach**: Generation of 100 synthetic efficiency curves around the known one (red line; left figure), by generating random parameter values, and inverse modeling approach, to extract ensembles of stochastic flow series.



Synthetic efficiency curves (six out of 100) around the "true" one (red line)

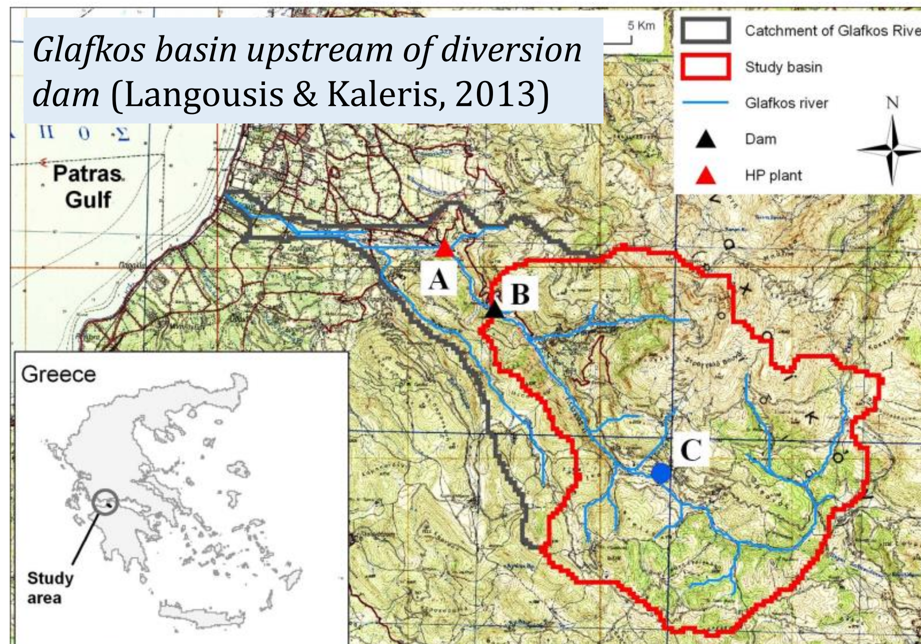


Flow data for a 100-day period

Real-world study: Glafkos power plant

- Glafkos (Greek: *Γλαύκος*, Latin: *Glaucus*) is a small river in thd city of Patras, Greece, flowing into the Gulf of Patras (Ionian Sea), south of the city centre.
- The hydroelectric power plant was built in 1927 and fully renovated in 1997.
- It is a typical run-of-river scheme, comprising:
  - a small diversion dam, receiving a mean annual inflow of ~39 hm^3;
  - a diversion tunnel, conveying ~31 hm^3 to the forebay tank;
  - a penstock of 1695 m length, taking advantage of a head of 150 m;
  - two turbines, Francis (2.3 MW) and Pelton (1.4 MW).
- The mean annual energy production is 10.4 GWh (capacity factor 31%).

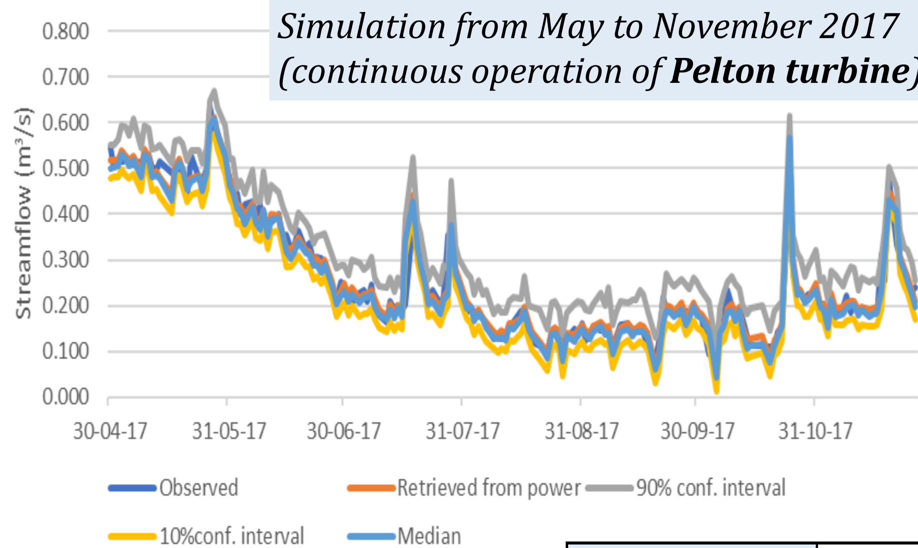

Source*:* Efstratiadis *et al.*, 2020



Glafkos basin upstream of diversion dam (Langousis & Kaleris, 2013)

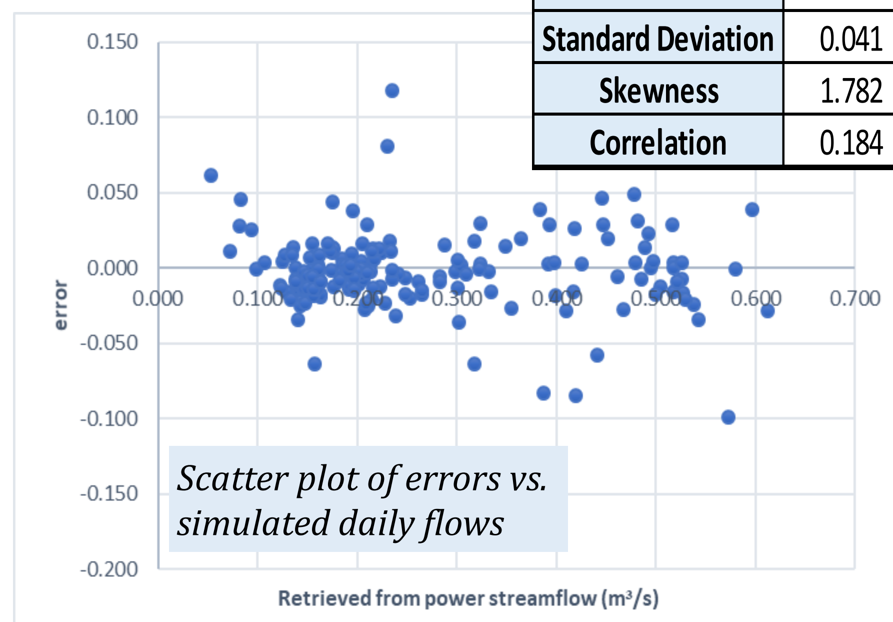
Inverse modeling procedure applied to Glafkos

- Available data from 2008 to 2018:
  - Daily water volume diverted from the dam to the power plant;
  - Hourly energy from each turbine;
- Computational procedure:
  - Retrieval of hourly flow data from hourly energy (inverse problem);
  - Extraction of error series by contrasting the aggregated daily flows to the actual ones;
  - Statistical analysis of errors and generation of long error data through an AR(1) model;
  - Synthesis of 100 ensembles of stochastic daily flow data, by adding synthetic errors to simulated data;
  - Empirical estimation of three characteristic quantiles (5, 50 and 95%), contrasted to observed flows;



*Simulation from May to November 2017 (continuous operation of **Pelton turbine**)*

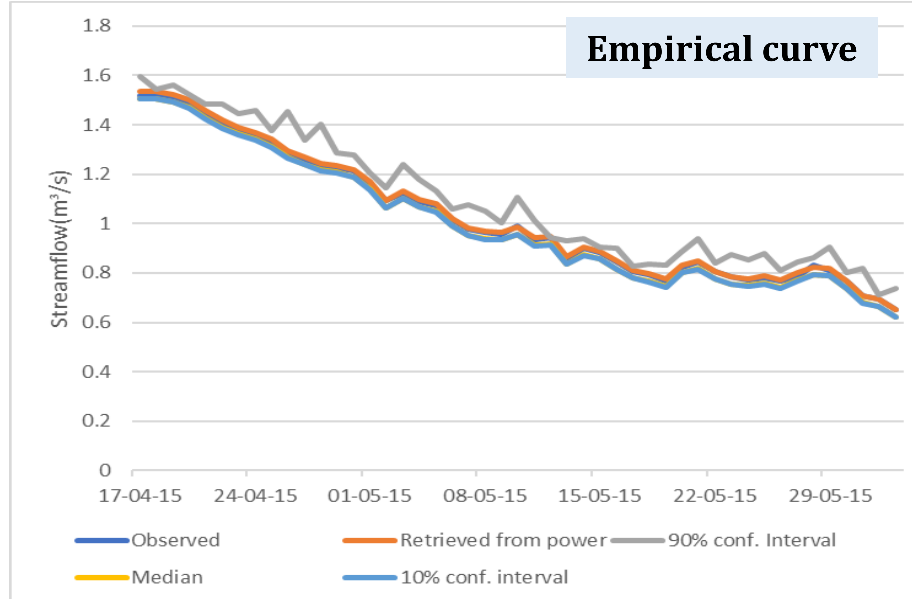
Mean	0.001
Standard Deviation	0.041
Skewness	1.782
Correlation	0.184

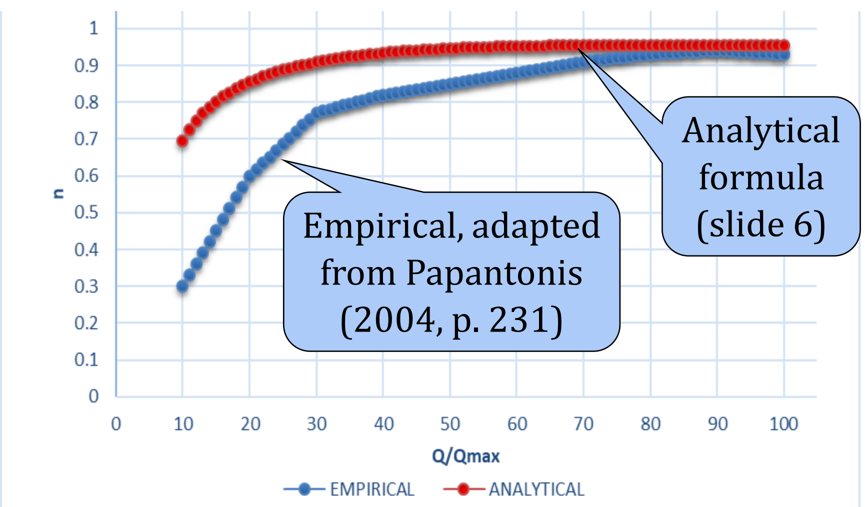


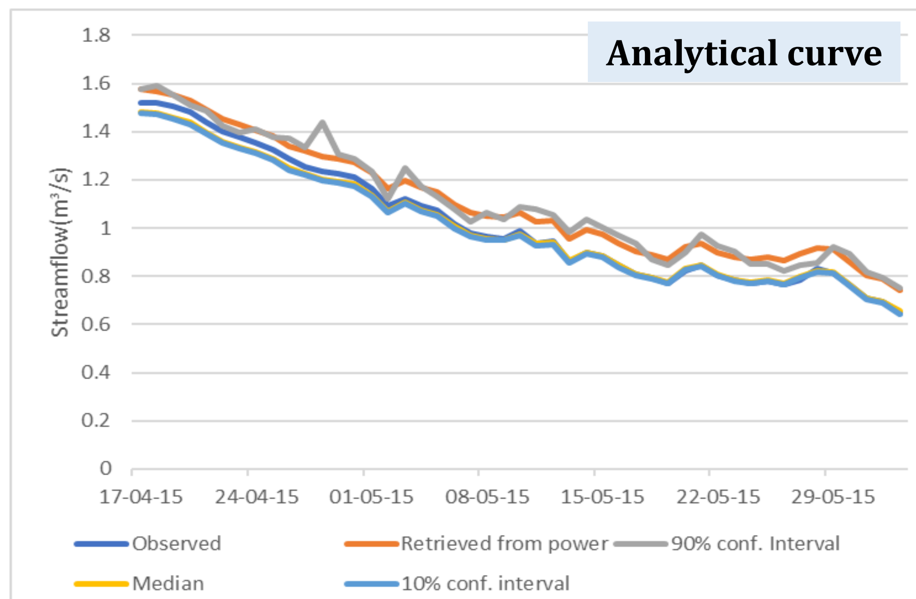
Scatter plot of errors vs. simulated daily flows

Impacts of uncertain efficiency curves

- Application of inverse modelling procedure to energy data provided by the **Francis turbine**, by applying two alternative efficiency curves:
  - Typical empirical curve for specific speed n_s = 100 rpm;
  - Analytical curve, with η_{min} = 0.70, η_{max} = 0.95, a =0.59 and b = 3.95.
- Multiplied by 0.95, to account for additional energy losses in the generator and the transformer.



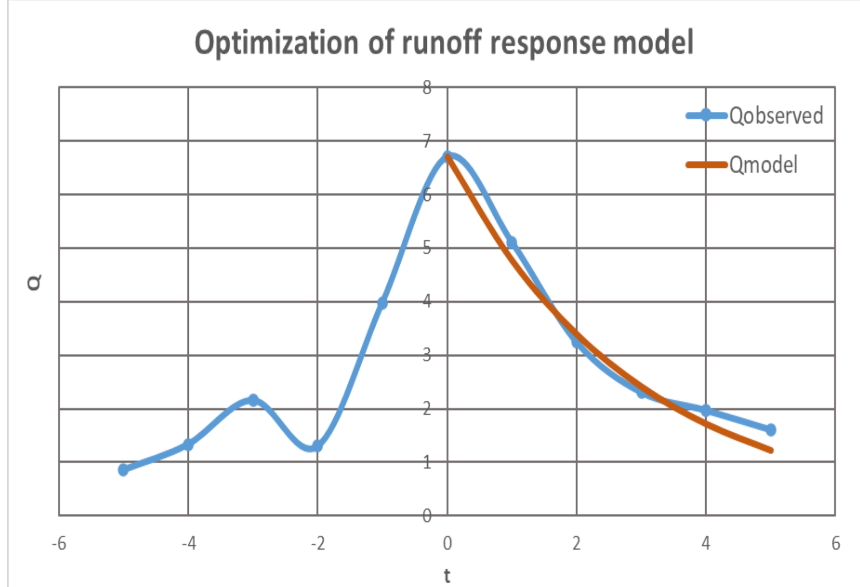


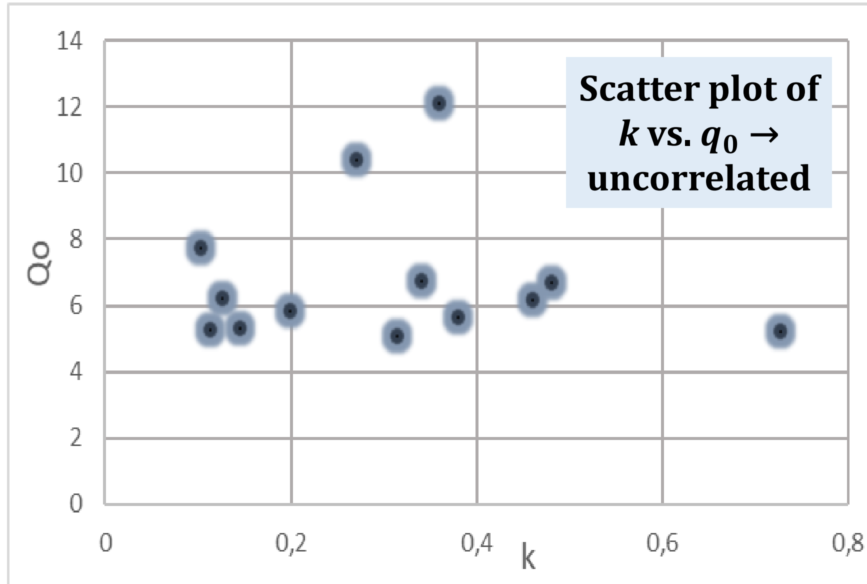


From post-analysis to forecasting

- **Forecasting** of future inflows: during the dry period (**long–term**) and the runoff response in rainfall events (**short-term**)

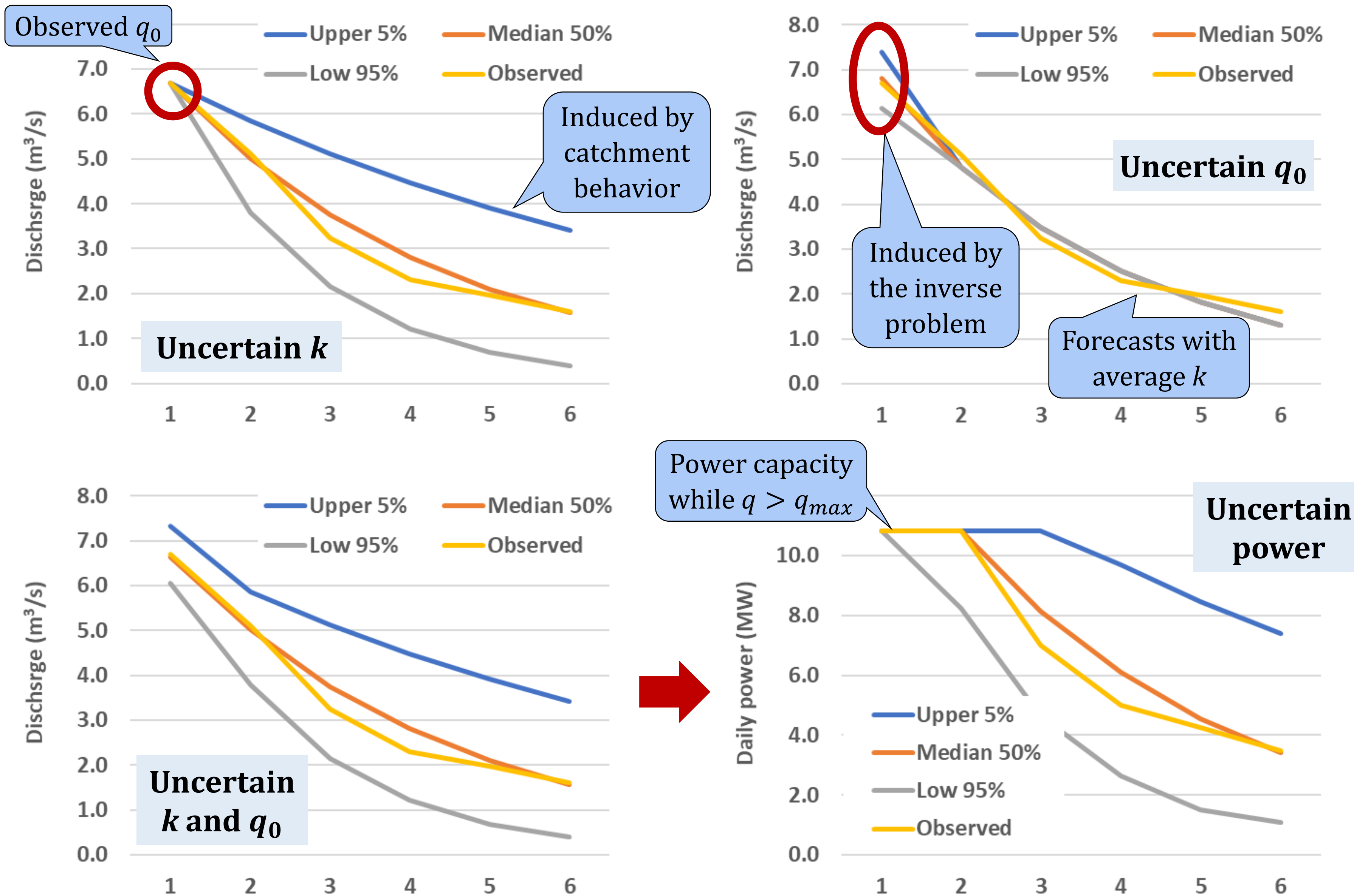
1. Modelling of the falling limb of the hydrograph via the linear reservoir recession model (Risva *et al.*, 2018):
$$q(t) = q_0 \exp(-k\ t)$$
2. Analysis of historic discharge data to estimate recession coefficients, k (different for floods and dry periods);
3. Fitting a statistical model of k, also accounting for dependencies with q_0.
4. For given q_0, generation of stochastic forecasting ensembles of discharge (random samples of k) and estimation of their confidence intervals.
5. Generation of ensemble forecasts accounting for combined uncertainty of initial flow, q_0 (derived from the inverse problem) and k.





Scatter plot of k vs. q_0 → uncorrelated

Example: 5-day forecasts during flood recession



Conclusions

- The retrieval of flows from energy data, here called **the inverse problem of hydroelectricity**, revealed many challenges, since the computational procedure exhibits multiple uncertainties.
- The **stochastic paradigm** – as the unique means for consistent quantification of uncertainty – can be easily applied to this problem, thus allowing to express the overall uncertainties in typical statistical terms (e.g. marginal statistics and confidence intervals);
- Here we focused on two key uncertain issues, i.e. the observed output (**energy production**) and the **efficiency curve of turbines**. Our analyses indicated that **efficiency is the major source of uncertainty**, particularly for the case of **Francis machines**, in which efficiency drops rapidly as discharge decreases.
- The **extrapolation of high and low flows**, outside of the range of operation of SHPs, is employed by combining empirical **hydrological rules** for representing the rising and falling limbs with stochastic approaches.
- The hydrological behavior of the catchment, as reflected in the **recession parameter** of falling limbs, plays important role in flow forecasting, both in short-term (flood recession) and in the long run (dry-period baseflow).
- Preliminary results showed that the nonlinear transformation of flow to energy seems resulting to slightly **smoothed uncertainties**, in terms of power predictions.

References

Anagnostopoulos, J. S., and D. E. Papantonis, Optimal sizing of a run-of-river small hydropower plant, *Energy Conversion & Management*, 48(10), 2663-2670, doi:10.1016/j.enconman.2007.04.016, 2007.

Efstratiadis, A., I. Nalbantis, and D. Koutsoyiannis, Hydrological modelling of temporally-varying catchments: Facets of change and the value of information, *Hydrological Sciences Journal*, 60(7-8), 1438–1461, doi:10.1080/02626667.2014.982123, 2015.

Efstratiadis, A., N. Mamassis, and D. Koutsoyiannis, *Lecture Notes on Renewable Energy and Hydroelectric Works*, Department of Water Resources and Environmental Engineering, National Technical University of Athens, 2020.

Hidalgo, I. G., D. G. Fontane, J. E. G. Lopes; J. G. P. Andrade, and A. F. de Angelis, Efficiency curves for hydroelectric generating units, *Journal of Water Resources Planning & Management*, 140(1), 86-91, doi:10.1061/(ASCE)WR.1943-5452.0000258, 2014.

Langousis, A., and V. Kaleris, Theoretical framework to estimate spatial rainfall averages conditional on river discharges and point rainfall measurements from a single location: an application to western Greece, *Hydrology & Earth System Sciences*, 17, 1241-1263, doi:10.5194/hess-17-1241-2013, 2013.

Paish, O., Small hydro power: technology and current status, *Renewable & Sustainable Energy Reviews*, 6(6), 537–556, doi:10.1016/S1364-0321(02)00006-0, 2002.

Papantonis, D. E., *Small Hydroelectric Works*, 2nd edition, Symeon, Athens, 2008.

Risva, K., D. Nikolopoulos, A. Efstratiadis, and I. Nalbantis, A framework for dry period low flow forecasting in Mediterranean streams, *Water Resources Management*, 32(15), 4911–1432, doi:10.1007/s11269-018-2060-z, 2018.