EGU General Assembly 2020
Online | 4–8 May 2020

**Session HS3.3: Stochastic modelling and real-time control of complex environmental systems**

# A brisk local uncertainty estimator for hydrologic simulations and predictions (Blue Cat)

**BLUE CAT: Brisk Local Uncertainty Estimation by Conditioning And Tallying**

Demetris Koutsoyiannis[1,2] & Alberto Montanari[1]

[1]Dipartimento Ingegneria Civile, Chimica, Ambientale e dei Materiali (DICAM), Università di Bologna

[2] School of Civil Engineering, National Technical University of Athens, Greece (dk@itia.ntua.gr, http://www.itia.ntua.gr/dk/)

# Introduction

Hydrological models are transformations of inputs $x_\tau$ (e.g. rainfall) at discrete time $\tau$ to outputs $Q_\tau$ (e.g. river discharge) by means of a model:

$$Q_\tau = G(\boldsymbol{x}_\tau) \quad (1)$$

where $\boldsymbol{x}_\tau$ is a vector containing a number of consecutive input variables, or even a matrix consisting of several input variables (such as rainfall, evapotranspiration, perhaps river discharge in an upstream basin, etc.). The transformation function is generally a complicated one, also involving additional state variables (e.g. soil moisture).

A model is never identical to reality and thus the true value of the output $q_\tau$ will be different from the model prediction $Q_\tau$.

In a blueprint, Montanari and Koutsoyiannis (2012) provided a framework to upgrade a deterministic model into stochastic, which provides the probability distribution of the output given the inputs and the deterministic model output, considering the uncertainty in model parameters and in input variables. This work has been discussed (Nearing, 2014; Koutsoyiannis and Montanari, 2014a) and advanced in other studies (Sikorska et al., 2015; Papacharalampous et al. 2019a,b)

Here we have the same aim but we study a different setting, whose main characteristic is the upgrade of the deterministic model into stochastic based on the data only.

# Background

The hypothesis set is that the information contained in the true outputs $q_\tau$ and the concurrent predictions by the deterministic model $Q_\tau$ is sufficient to support this upgrade.

Simplicity is a principal objective and therefore we do not involve multiple simulations, Bayesian methods (e.g. MCMC) etc.; rather we aim at a computational framework that can run even in a worksheet software.

We do not consider parameter uncertainty in the deterministic model on the basis that another parameter set is in fact another model and the approach here is intentionally single-model rather than multi-model. Given that the final product is a stochastic model, there is no need to consider separately the parameter uncertainty. Rather, the stochastic model per se provides a basis to compare different deterministic models in terms of their overall efficiencies.

We do not subdivide uncertainty in different components. The framework should automatically incorporate all types, including the uncertainty in input data, for which no particular provision is necessary.

The framework assumes stationarity (cf. Montanari and Koutsoyiannis, 2014b; Koutsoyiannis and Montanari, 2015). If different subperiods are characterized by different model parameters or different input uncertainty, then we can split the entire simulated period in subperiods in which stationarity can be safely assumed.

# Premises

For advancing a deterministic model into stochastic, we regard all related quantities as stochastic (random) variables and their sequences as stochastic processes. For notational clarity we underline stochastic variables, stochastic processes and stochastic functions. We use non-underlined symbols for regular variables and deterministic functions, as well as for realizations of stochastic variables and of stochastic processes, where the latter realizations are also known as time series.

We assume that the inputs $\underline{x}_\tau$, at discrete times $\tau$, have a stationary probability density function $f_x(x)$ and distribution function $F_x(x)$. The output $\underline{q}_\tau$ depends on the inputs $\underline{x}_\tau$ and is given though some stochastic function (S-model) as:

$$\underline{q}_\tau = \underline{g}(\underline{\boldsymbol{x}}_\tau) \tag{2}$$

The stochastic process $\underline{q}_\tau$ is assumed to correspond to the real process, while the outcome of the deterministic model (D-model) of equation (1) is an estimate thereof. By writing the latter equation in stochastic terms, retaining however the function $G$ $(\neq \underline{g})$ as a deterministic function, we obtain the estimator $\underline{Q}_\tau$ of the output $\underline{q}_\tau$ as:

$$\underline{Q}_\tau \coloneqq G(\underline{\boldsymbol{x}}_\tau) \tag{3}$$

To advance from the D-model, in its form (3), to the S-model in (2) we just need to specify the conditional distribution:

$$F_{q|Q}(q|Q) = P\left\{\underline{q} \le q | \underline{Q} = Q\right\} \tag{4}$$

with $q$ and $Q$ assumed concurrent and referring to discrete time $\tau$. In other words, here conditioning is meant in scalar setting. An extension where $Q$ is a vector containing the current and earlier predictions by the D-model is possible but more laborious, thus not complying with our simplicity target.

# Basic requirements

It is relatively easy to infer from data the marginal distribution and density functions of the S-variable $\underline{q}$ and D-predicted variable $\underline{Q}$. Therefore we may assume that $f_q(q)$ and $f_Q(Q)$ are known.

Then the conditional density sought should obey:

$$\int_{-\infty}^{\infty} f_{q|Q}(q|Q)\mathrm{d}q = 1, \qquad \int_{-\infty}^{\infty} f_{q|Q}(q|Q) f_Q(Q)\mathrm{d}Q = f_q(q) \tag{5}$$

The former equation is trivial. The latter one, if we set $z = F_Q(Q)$, with $Q = F_Q^{-1}(z)$, so that $f_Q(Q)\mathrm{d}Q = \mathrm{d}z$, can be written as:

$$\int_0^1 f_{q|Q}\left(q\middle|F_Q^{-1}(z)\right)\mathrm{d}z = f_q(q) \tag{6}$$

Integrating we find:

$$\int_0^q \int_0^1 f_{q|Q}\left(a\middle|F_Q^{-1}(z)\right)\mathrm{d}z\mathrm{d}a = F_q(q) \tag{7}$$

and changing the order of the integrals we finally find:

$$\int_0^1 F_{q|Q}\left(q\middle|F_Q^{-1}(z)\right)\mathrm{d}z = F_q(q) \tag{8}$$

# Approximation using data

If we have time series of concurrent $Q$ and $q$, each of size $n$, and if $Q_{(i:n)}$ is the $i$th smallest value in the time series of $Q$ and $q_{(j:n)}$ is the $j$th smallest value in the time series of $q$, then we can use the approximations $F_Q(Q_i) \approx i/n$ and $F_q(q_j) \approx j/n$, and thus approximate $F_q(q)$ in equation (8) as:

$$\frac{1}{n}\sum_{i=1}^{n} F_{q|Q}\left(q \middle| Q_{(i:n)}\right) \approx F_q(q) \tag{9}$$

and for $q = q_j$

$$\frac{1}{n}\sum_{i=1}^{n} F_{q|Q}\left(q_{(j:n)} \middle| Q_{(i:n)}\right) \approx \frac{j}{n} \tag{10}$$

Hence:

$$B_j := \sum_{i=1}^{n} F_{q|Q}\left(q_{(j:n)} \middle| Q_{(i:n)}\right) = j \tag{11}$$

We can thus attempt to determine $F_{q|Q}$ by minimizing:

$$A := \sum_{j=1}^{n}\left(B_j - j\right)^2 = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} F_{q|Q}\left(q_{(j:n)} \middle| Q_{(i:n)}\right) - j\right)^2 \tag{12}$$

This may admit infinite solutions in case that the minimum value can be zero. Strategies to approach possible solutions will be discussed in section "From ideal to real cases".

# Specific cases of D-model: good models

The **perfect D-model** is when $\underline{q} = \underline{Q}$. In this case:

$$F_{q|Q}(q|Q) = P\left\{\underline{q} \le q \middle| \underline{Q} = Q\right\} = P\left\{\underline{Q} \le q \middle| \underline{Q} = Q\right\} = \begin{cases} 0 & Q > q \\ 1 & Q \le q \end{cases} \tag{13}$$

The inequality $Q \le q$ can be written in terms of $z = F_Q(Q)$ as $z \le F_Q(q) = F_q(q)$ and thus equation (8) holds true:

$$\int_0^1 F_{q|Q}\left(q \middle| F_Q^{-1}(z)\right) \mathrm{d}z = \int_0^{F_q(q)} 1 \mathrm{d}z = F_Q(q) = F_q(q) \tag{14}$$

The data-based approximation in (11) also holds true:

$$F_{q|Q}\left(q_{(j:n)} \middle| Q_{(i:n)}\right) = \begin{cases} 0 & j \le i \\ 1 & j > i \end{cases}, \qquad B_j = j, \qquad A = 0 \tag{15}$$

A **certain D-model** is one with bias but no uncertainty. In this case there holds a monotonically increasing deterministic relationship between $\underline{q}$ and $\underline{Q}$ without uncertainty, i.e., $\underline{q} = h\left(\underline{Q}\right)$ or equivalently $\underline{Q} = h^{-1}(\underline{q})$. Hence:

$$F_{q|Q}(q|Q) = P\left\{\underline{q} \le q \middle| \underline{Q} = Q\right\} = \begin{cases} 0 & Q > h^{-1}(q) \\ 1 & Q \le h^{-1}(q) \end{cases} \tag{16}$$

The inequality $Q \le h^{-1}(q)$ can be written in terms of $z = F_Q(Q)$ as $z \le F_Q\left(h^{-1}(q)\right) = F_q(q)$ and equation (8) holds true:

$$\int_0^1 F_{q|Q}\left(q \middle| F_Q^{-1}(z)\right) \mathrm{d}z = \int_0^{F_q(q)} 1 \mathrm{d}z = F_q(q) \tag{17}$$

The data-based approximation in (11) holds true again:

$$F_{q|Q}\left(q_{(j:n)} \middle| Q_{(i:n)}\right) = \begin{cases} 0 & j \le i \\ 1 & j > i \end{cases}, \qquad B_j = j, \qquad A = 0 \tag{18}$$

# Specific cases D-model: bad models

A **terrible D-model** resembles the certain D-model in the sense that it is not affected by uncertainty but in this case the monotonic deterministic relationship between $\underline{q}$ and $\underline{Q}$, $\underline{q} = h\left(\underline{Q}\right)$ or equivalently $\underline{Q} = h^{-1}(\underline{q})$, is decreasing. Obviously this is not a good model and this case is not usually met in hydrological practice, except in cases of using climate model outputs, which sometimes have negative correlation with reality (Tyralis and Koutsoyiannis, 2017). In this case:

$$F_{q|Q}(q|Q) = P\left\{\underline{q} \le q \middle| \underline{Q} = Q\right\} = \begin{cases} 0 & Q < h^{-1}(q) \\ 1 & Q \ge h^{-1}(q) \end{cases} \tag{19}$$

The inequality $Q \ge h^{-1}(q)$ can be written in terms of $z = F_Q(Q)$ as $z \le F_Q\left(h^{-1}(q)\right) = F_q(q)$ and equation (8) holds true:

$$\int_0^1 F_{q|Q}\left(q \middle| F_Q^{-1}(z)\right)\mathrm{d}z = \int_0^{F_q(q)} 1\mathrm{d}z = F_q(q) \tag{20}$$

The data-based approximation in (11) holds true again:

$$F_{q|Q}\left(q_{(j:n)} \middle| Q_{(i:n)}\right) = \begin{cases} 0 & j \le i \\ 1 & j > i \end{cases}, \qquad B_j = j, \qquad A = 0 \tag{21}$$

The **irrelevant D-model** is one in which $\underline{q}$ is independent from $\underline{Q}$ and thus

$$F_{q|Q}(q|Q) = F_q(q) \tag{22}$$

Equation (8) holds true again:

$$\int_0^1 F_{q|Q}\left(q \middle| F_Q^{-1}(z)\right)\mathrm{d}z = F_q(q)\int_0^1 1\mathrm{d}z = F_q(q) \tag{23}$$

The data-based approximation in (11) holds true again:

$$B = \sum_{i=1}^{n} F_{q|Q}\left(q_{(j:n)} \middle| Q_{(i:n)}\right) = \sum_{i=1}^{n} F_q\left(q_{(j:n)}\right) = nF_q\left(q_{(j:n)}\right) = n\frac{j}{n} = j, \qquad A = 0 \tag{24}$$

# From ideal to real cases

In the ideal case of a perfect model there is nothing to do—the prediction equals the true value.

The certain and terrible models shift the predictions $Q$ to the actual values $q$ and the prediction interval has zero width—again the prediction is deterministic. Our only desideratum in this case is to determine the deterministic function $q = h(Q)$.

In the irrelevant model the predictions $Q$ are essentially discarded and the prediction interval becomes constant, fully determined by the marginal distribution $F_q(q)$.

In real world cases we may assume that the D-model is neither irrelevant, nor certain, let alone perfect, and hopefully nor terrible. In these cases, as already discussed, our aim is to derive the conditional distribution $F_{q|Q}(q|Q)$, which incorporates both a shift of the prediction $Q$ toward the real value $q$ (bias correction) and the probabilistic assessment of the stochastic error (uncertainty assessment). Note that in the certain model, the conditional distribution sought becomes $F_{q|Q}(q|Q) = U\big(h(Q)\big)$, where $U(x)$ is the unit step function ($U(x) = 0$ for $x < 0$ and $U(x) = 1$ for $x \geq 0$).

One strategy to tackle the problem is to use a simple parametric relationship for the function $F_{q|Q}(q|Q)$ and determine its parameters by minimizing the quantity $A$ in equation (12). An example would be to assume $F_{q|Q}(q|Q)$ to be a Pareto-Burr-Feller (PBF) distribution (see Appendix 2) with constant tail indices $\xi$ and $\zeta$ and scale parameter varying with $Q$; the manner to handle this is provided in section "Simple simulation".

A similar approach would be to assume a copula $C\left(F_q(q), F_Q(Q)\right)$ and determine $F_{q|Q}(q|Q)$ as:

$$F_{q|Q}(q|Q) = \frac{F_{qQ}(q,Q)}{f_Q(Q)}, \qquad F_{qQ}(q,Q) = C\left(F_q(q), F_Q(Q)\right) \tag{25}$$

If $C$ is a copula then equation (12) should hold automatically.

# A simple fully data-based alternative

While a parametric approach like the above is attractive from many aspects, here we try an even simpler approach, i.e. we try to determine $F_{q|Q}(q|Q)$ from the data alone without assuming a specific expression for the distribution. As the variables of interest in hydrology are of continuous type, we may expect that each value $Q_\tau$ in the available time series appears only once. Thus we cannot form a sample for a particular value of $Q$. However, as a simple approximation of $F_{q|Q}(q|Q)$, we can form a sample of $Q$-neighbours based on:

$$F_{q|Q}(q|Q) = P\left\{\underline{q} \le q \middle| \underline{Q} = Q\right\} \approx P\left\{\underline{q} \le q \middle| Q - \Delta Q_1 \le \underline{Q} \le Q + \Delta Q_2\right\} =$$
$$\approx P\left\{\underline{q} \le q \middle| F_Q(Q) - \Delta F_1 \le F_Q\left(\underline{Q}\right) \le F_Q(Q) + \Delta F_2\right\} =: F_{q|[Q]}(q|Q, \Delta F_1, \Delta F_2) \quad (26)$$

where the increments $\Delta Q_i$ and $\Delta F_i$ can be chosen based on the requirement that the intervals below and above the values $Q$ or and $F_Q(Q)$ contain appropriate numbers of data values, $m_1 := \Delta F_1 n$ and $m_2 := \Delta F_2 n$, respectively. The numbers $m_1$ and $m_2$ should not be too large, so that $F_Q(Q) \pm \Delta F_{1,2}$ be close to $F_Q(Q)$, nor too small, so that the probability $P\left\{\underline{q} \le q \middle| F_Q(Q) - m_1/n \le F_Q\left(\underline{Q}\right) \le F_Q(Q) + m_2/n\right\}$ can be estimated empirically, from a sample of size $m_1 + m_2 + 1$, as reliably as possible.

In general, we may choose $\Delta F_1 = \Delta F_2 = \Delta F$ and $m_1 = m_2 = m$. For example, setting $m_1 = m_2 = m = 20$, i.e. $m_1 + m_2 + 1 = 41$, the lowest empirical probability we can estimate would be $1/41 \approx 2.5\%$ and the highest one $1 - 1/40 \approx 97.5\%$. Conversely, for probabilities 2.5% and 97.5% we can empirically estimate the corresponding quantiles of $q$ as the minimum and the maximum observed value, respectively, in the sample of $m_1 + m_2 + 1$ values. Details on the choice of numbers $m_1$ and $m_2$ are given in Appendix 4.

For more reliable estimates that do not depend on one data point only as above, we could use a larger $m$ along with order statistics different from the lowest and the highest ones. The newly introduced concept of knowable moments (K-moments, Koutsoyiannis, 2018, 2020) offers an alternative for empirical quantile estimates, more general and reliable than order statistics as it combines many data points in each estimate.

# A summary of the K-moments approach

The *noncentral knowable moment* (or *noncentral K-moment*) *of order* $(p, q)$ is defined as (Koutsoyiannis, 2018):

$$K'_{pq} \coloneqq (p - q + 1)\mathrm{E}\left[\left(F(\underline{x})\right)^{p-q} \underline{x}^q\right], \qquad p \geq q \tag{27}$$

A most interesting special case is for $q = 1$. The *noncentral knowable moment of order* $(p, 1)$ is:

$$K'_p \coloneqq p\mathrm{E}\left[\left(F(\underline{x})\right)^{p-1} \underline{x}\right], \qquad p \geq 1 \tag{28}$$

A basic property that connects the K-moments with expectations of maxima is:

$$K'_p = \mathrm{E}[\underline{x}_{(p)}] = \mathrm{E}\left[\max(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p)\right] \tag{29}$$

For expectations of minima another type of K-moments is defined, as described in Appendix 1.

For a sample of size $n$, $K'_p$ has an unbiased estimator (Koutsoyiannis, 2020):

$$\widehat{K}'_p = \sum_{i=1}^{n} b_{inp}\, \underline{x}_{(i:n)}, \qquad b_{inp} = \begin{cases} 0, & i < p \\ \dfrac{p}{n} \dfrac{\Gamma(n-p+1)}{\Gamma(n)} \dfrac{\Gamma(i)}{\Gamma(i-p+1)}, & i \geq p \geq 0 \end{cases} \tag{30}$$

where $\Gamma(\,)$ is the gamma function and $\underline{x}_{(i:n)}$ is the *i*th *order statistic,* defined to be the *i*th smallest of *n* iid stochastic variables arranged in increasing order of magnitude, i.e.: $\underline{x}_{(1:n)} \leq \underline{x}_{(2:n)} \leq \cdots \leq \underline{x}_{(n:n)}$. The minimum and maximum are, respectively,

$$\underline{x}_{(1:n)} = \min(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n), \qquad \underline{x}_{(n)} \coloneqq \underline{x}_{(n:n)} = \max(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n) \tag{31}$$

and represent special cases of the order statistics, the lowest and the highest.

# Assigning probabilities to K-moments

By definition, $K'_p$ represents the expected value of the maximum of $p$ copies of $\underline{x}$. Koutsoyiannis (2020) has introduced the Λ-coefficient of order $p$ as:

$$\Lambda_p := \frac{1}{p\left(1 - F\left(K'_p\right)\right)} \tag{32}$$

$\Lambda_p$ happens to vary only slightly with $p$. Any symmetric distribution will give exactly $\Lambda_1 = 2$ because $K'_1$ is the mean, which will equal the median and thus yield $F\left(K'_p\right) = 1/2$. Thus, a rough approximation is the rule of thumb:

$$\Lambda_p \approx 2 \tag{33}$$

Generally, the exact value $\Lambda_1$ is easy to determine, as it is directly related to the mean:

$$\Lambda_1 = \frac{1}{1 - F(\mu)} \tag{34}$$

The exact value of $\Lambda_\infty$ depends only on the tail index $\xi$ of the distribution:

$$\Lambda_\infty = \begin{cases} \Gamma(1-\xi)^{\frac{1}{\xi}}, & \xi \neq 0 \\ \mathrm{e}^{\gamma}, & \xi = 0 \end{cases} \tag{35}$$

where $\gamma = 0.577$ is the Euler's constant.

# Assigning probabilities to K-moments (2)

These enable the simple approximation of $\Lambda_p$ and hence of the non-exceedance probability:

$$\Lambda_p \approx \Lambda_\infty + \frac{\Lambda_1 - \Lambda_\infty}{p}, \qquad F\left(K'_p\right) \approx 1 - \frac{1}{\Lambda_\infty p + (\Lambda_1 - \Lambda_\infty)} \tag{36}$$

Conversely, for a given non-exceedance probability $F$, we can calculate the quantile $x$ as the $K'_p$ that corresponds to:

$$p \approx \frac{1}{\Lambda_\infty (1 - F)} + 1 - \frac{\Lambda_1}{\Lambda_\infty} \tag{37}$$

The estimate of $K'_p$ based on the typical estimator $\underline{\hat{K}}'_p = \sum_{i=1}^{n} b_{inp}\, \underline{x}_{(i:n)}$ is more reliable than that based on a single $x_{(i:n)}$ because it is derived from many data points (except when $i = n$, where the two approaches are precisely identical).

# Stochastic assessment of a prediction model (1): marginal distributions

Laio and Tamea (2007) have introduced a diagnostic plot combining probability distributions of predictions and true values, which has become later popular in similar studies, having been termed *predictive quantile-quantile* (PQQ) *plot*—even though in the original paper it has been called simply probability plot. Here we refer to it as *combined probability-probability* (CPP) *plot* because the popular term "quantile" is wrong (the plot represents probabilities rather that quantiles). In stochastic language, CPP is a plot of the empirical distribution function $F_z(z)$ of a stochastic variable $\underline{z}$ against its value $z$. The variable is defined as the non-exceedence probability:

$$\underline{z} \coloneqq F_Q(\underline{q}) \tag{38}$$

Its distribution function is $F_z(z) = P\{\underline{z} \le z\} = P\left\{F_Q\left(\underline{q}\right) \le z\right\} = P\left\{\underline{q} \le F_Q^{-1}(z)\right\}$ and hence:

$$F_z(z) = F_q\left(F_Q^{-1}(z)\right) \tag{39}$$

In other words, $F_z(z)$ combines the distribution functions of predictions $Q$ and real quantities $q$. The predictions are regarded as good if the plot $F_z(z)$ vs. $z$ is the equality line, i.e., if:

$$F_z(z) = z \tag{40}$$

which means that the distribution of $\underline{z}$ is uniform. In this case $z = F_q\left(F_Q^{-1}(z)\right)$ or $F_q^{-1}(z) = F_Q^{-1}(z)$. This is possible only if $F_Q(x)$ is identical to $F_q(x)$. Therefore the plot in essence tests whether the two distributions, estimated from the data, are identical. The procedure is quite similar to plotting $F_Q(x)$ vs. $F_q(x)$.

# Stochastic assessment of a prediction model (2): Rank correlations

It is noted that the CPP plot, except for assessing the proximity of the two marginal distributions, does not give any other indication if the predictions are good. For example if $\underline{Q}$ is completely independent from $\underline{q}$ (an obviously irrelevant model) but the two distributions are identical, again the distribution of $\underline{z}$ will be uniform.

Therefore another metric should also be used in addition to CPP plot. A simple one used here is the Spearman's rank correlation coefficient, i.e. the correlation coefficient between the ranks of $Q$ and $q$. If its value is close to 1, the model would be close to certain, regardless of the value of the Pearson correlation coefficient of $Q$ and $q$.

# Simple simulation: Method A

If we want to generate a simulation series from the S-model for a period for which there exist D-model predictions, we can apply the following method that is a simplified form of that in Sikorska et al. (2015); the latter study utilized nearest neighbours in the D-model outputs to make a stochastic prediction.

The simple Method A used here contains the following steps.

1. Given a model prediction $Q$, we locate in the ordered sample $Q_{(i:n)}, i = 1, \dots, n$, the rank $i$ for which $Q$ is closest to $Q_{(i:n)}$.
2. We retrieve the sample sizes $m_1$ and $m_2$ that have been used in the calibration for the particular $Q_{(i:n)}$.
3. We generate a random number $j$ in the interval $[i - m_1, i + m_2]$ from the uniform distribution.
4. We locate the time $k$ ($k = 1, \dots, n$) which corresponds to the value $Q_{(j:n)}$ and take as simulated value of the true discharge the value $q_k$.

The problem of Method A is that it does not extrapolate the output beyond the maximum and minimum values contained in the observations $q_\tau$ of the calibration period.

# Simple simulation: Method B

The problem of method A can be tackled with an alternative parametric Method B, whose steps are the following.

1. We assume that the conditional distribution function $F_{q|Q}(q, Q)$ is PBF (see Appendix 2) with scale parameter depending on $Q$ but with tail indices constant. Then the tail indices could be estimated by bracketing the conditional distribution function as described in Appendix 3.
2. Given a model prediction $Q$, we locate in the ordered sample $Q_{(i:n)}, i = 1, \dots, n$, the rank $i$ for which $Q$ is closest to $Q_{(i:n)}$ and we retrieve from the calibration phase the expectation $\mathrm{E}[q|Q = Q_{(i:n)}]$.
3. We determine the scale parameter $\lambda$ by equating the theoretical mean with the estimated $\mathrm{E}[q|Q = Q_{(i:n)}]$. Specifically, we use equation (66) with $p = 1$ and $\mu_1' = \mathrm{E}[q|Q = Q_{(i:n)}]$ and solve it for $\lambda$.
4. We generate $q$ from the PBF distribution.

Both Methods A and B ensure preservation of the true distribution function of $q$ (i.e., a perfect CPP plot) regardless of how good or bad the D-model is.

# A toy model for checking the framework

We use a toy model to check the framework, which is intentionally constructed as simple as possible. The D-model is intentionally diverging from reality, with large bias, which alternates between positive and negative values (for different ranges of D-predictions) and substantial heteroscedasticity. These negative features help to check whether the framework is able to recover reality from a D-model that is biased, with inconsistent tail behaviour and heteroscedasticity.

The input process $\underline{x}_\tau$ at discrete time $\tau$ is assumed independent in time, stationary, and nonnegative ($\underline{x}_\tau \geq 0$) with exponential distribution:

$$f_x(x) = \mathrm{e}^{-x} \tag{41}$$

The output $\underline{q}_\tau$ is assumed to be a deterministic function of two consecutive $\underline{x}_\tau$:

$$\underline{q}_\tau = g\left(\underline{x}_\tau, \underline{x}_{\tau-1}\right) := c(\mathrm{e}^{a\,\underline{x}_\tau + b\underline{x}_{\tau-1}} - 1) \tag{42}$$

where $a$ and $b$ are independent parameters satisfying $0 \leq a, b \leq 1$ and $c$ is an additional parameter determined from $a$ and $b$ on the basis that $\mathrm{E}\left[\underline{q}_\tau\right] = \mathrm{E}\left[\underline{x}_\tau\right] = 1$ (see below). It may be noticed that $\underline{x}_\tau \geq 0, \underline{q}_\tau \geq 0, h(0,0) = 0$ and that the output has a heavy tailed distribution while the input has not.

To make the output a stochastic function $\underline{g}$ of input, we omit the input $\underline{x}_{\tau-1}$ and write

$$\underline{q}_\tau = \underline{g}(\underline{x}_\tau) \tag{43}$$

We estimate $\underline{q}_\tau$ by the simplest possible D-model:

$$\underline{Q}_\tau = G\left(\underline{x}_\tau\right) = \underline{x}_\tau \tag{44}$$

# Specification of the toy S-model

To specify the S-model for $q_\tau$ we need to determine the conditional distribution $F_{q_\tau|Q_\tau}(q|Q)$, which we assume stationary, i.e. $F_{q_\tau|Q_\tau}(q|Q) \equiv F_{q|Q}(q|Q)$. We have:

$$F_{q|Q}(q|Q) = F_{q_\tau|Q_\tau}(q|Q) = P\left\{\underline{q}_\tau \le q|\underline{Q}_\tau = Q\right\} = P\left\{\underline{q}_\tau \le q|\underline{x}_\tau = Q\right\} = P\{c(\mathrm{e}^{aQ+b\underline{x}_{\tau-1}} - 1) \le q\}$$

$$= P\left\{\underline{x}_{\tau-1} \le \ln\left(\frac{q}{c}+1\right)^{\frac{1}{b}} - \frac{aQ}{b}\right\} = F_x\left(\ln\left(\frac{q}{c}+1\right)^{\frac{1}{b}} - \frac{aQ}{b}\right) = 1 - e^{\frac{aQ}{b}}\left(\frac{q}{c}+1\right)^{-\frac{1}{b}} \quad (45)$$

where we notice that the rightmost side is positive when

$$q \ge c(\mathrm{e}^{aQ} - 1) \ge 0, \qquad 0 \le Q \le \frac{1}{a}\ln\left(\frac{q}{c}+1\right) \quad (46)$$

while otherwise $F_{q_\tau|Q_\tau}(q|Q) = 0$.

Based on the above two equations, we can find every property of the model, conditional and marginal, using the standard algebra of stochastics. The results are listed in the table of next page.

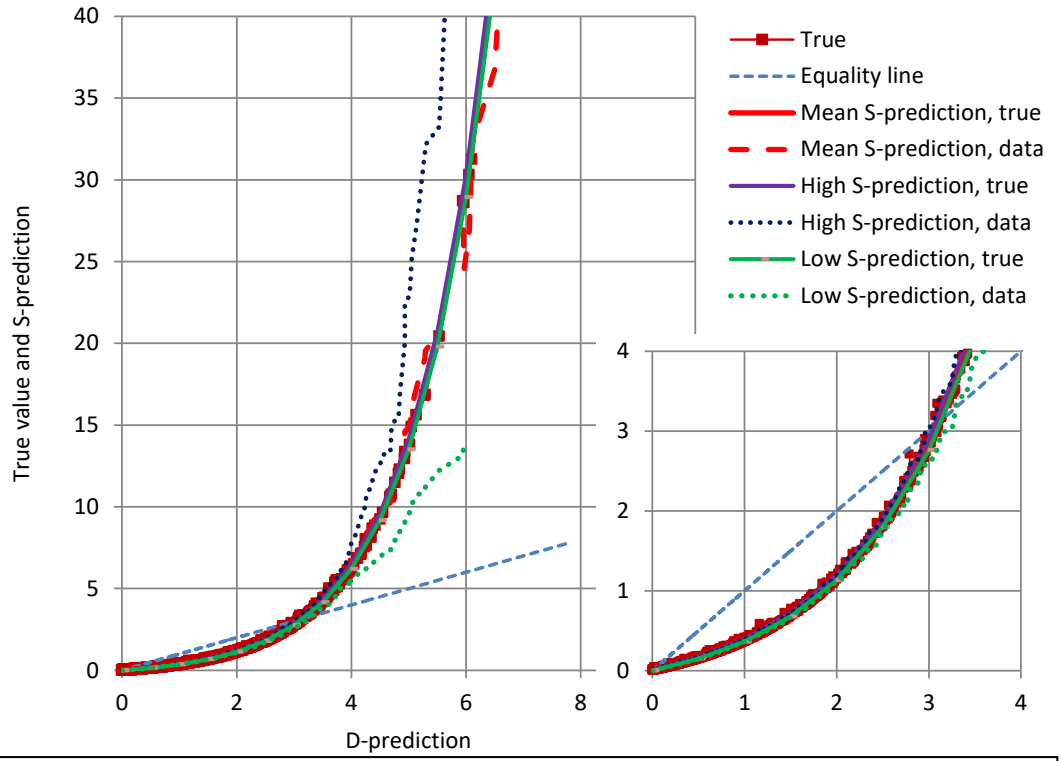
# Toy S-model equations

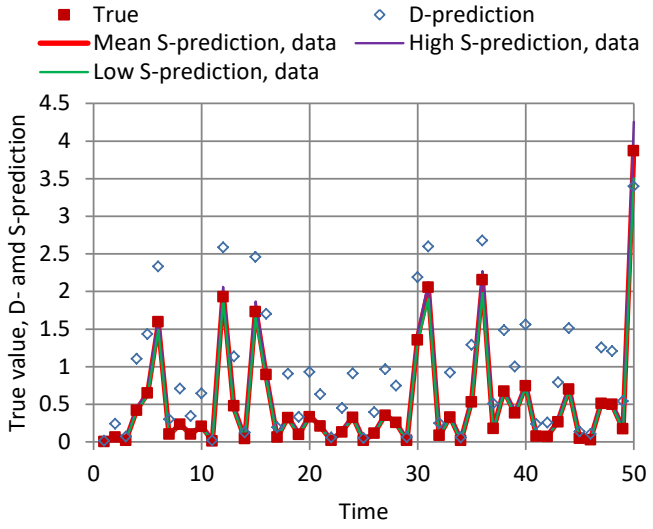
| Quantity | Related Equations* | Ref. |
|---|---|---|
| Parameters† | $0 < a < 1, \quad 0 < b < 1, \quad c = \frac{(1-a)(1-b)}{a+b-a\,b}$ | (47) |
| Joint density | $f_{q,Q}(q,Q) = \frac{1}{bc}\mathrm{e}^{\left(\frac{a}{b}-1\right)Q}\left(\frac{q}{c}+1\right)^{-\frac{1}{b}-1}$ | (48) |
| Conditional density | $f_{q\|Q}(q\|Q) = \frac{1}{bc}\mathrm{e}^{\frac{aQ}{b}}\left(\frac{q}{c}+1\right)^{-\frac{1}{b}-1}$ | (49) |
| Conditional distribution | $F_{q\|Q}(q\|Q) = 1-\mathrm{e}^{\frac{aQ}{b}}\left(\frac{q}{c}+1\right)^{-\frac{1}{b}}$ | (50) |
| Conditional quantile | $q = c\left(\mathrm{e}^{\frac{aQ}{b}}\left(1-F_{q\|Q}(q\|Q)\right)^{-b}-1\right)$ | (51) |
| Marginal densities | $f_Q(Q) = e^{-Q}, \quad f_q(q) = \frac{1}{c(a-b)}\left(\left(\frac{q}{c}+1\right)^{-\frac{1}{a}-1}-\left(\frac{q}{c}+1\right)^{-\frac{1}{b}-1}\right)$ | (52) |
| Marginal distributions | $F_Q(Q) = 1-e^{-Q}, \quad F_q(q) = \frac{1}{a-b}\left(a\left(1-\left(\frac{q}{c}+1\right)^{-\frac{1}{a}}\right)-b\left(1-\left(\frac{q}{c}+1\right)^{-\frac{1}{b}}\right)\right)$ | (53) |
| Marginal means | $\mathrm{E}\left[\underline{Q}\right] = 1, \mathrm{E}\left[\underline{q}\right] = \frac{c(a+b-a\,b)}{(1-a)(1-b)} = 1$ | (54) |
| Conditional mean | $\mathrm{E}\left[\underline{q}\|Q\right] = \frac{c(\mathrm{e}^{aQ}-1+b)}{1-b}$ | (55) |
| Tail indices | $\zeta = 1; \xi = \max(a,b)$ for marginal distribution; $\xi = b$ for conditional distribution | (56) |

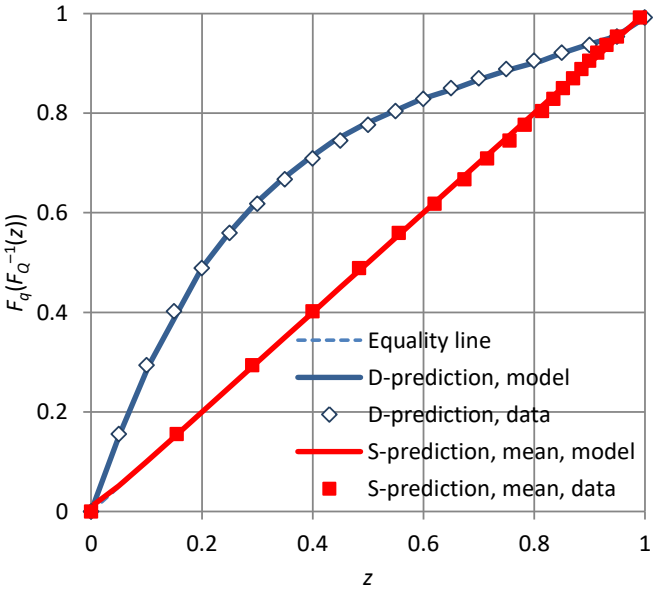
* The equations are valid for $a \neq b$. The limit for $a = b$, as well as those for $a, b = 0, 1$ can be easily found but are omitted for brevity. The equations are valid for $q$ and $Q$ satisfying (46); out of the limits of (46), the probability densities are zero.

† The upper limits for $a$ and $b$ assure a finite mean; for finite variance they should be replaced by 1/2. The value of $c$ (> 0) ensures mass conservation, i.e., $\mathrm{E}\left[\underline{Q}\right] = \mathrm{E}\left[\underline{q}\right]$ (see equation (54)).

# Toy model results: almost certain D-model




Independent parameters: $a = 0.75, b = 0.02$

Choices: $M = 10, F_{\mathrm{L}} = 1 - F_{\mathrm{H}} = 0.1$

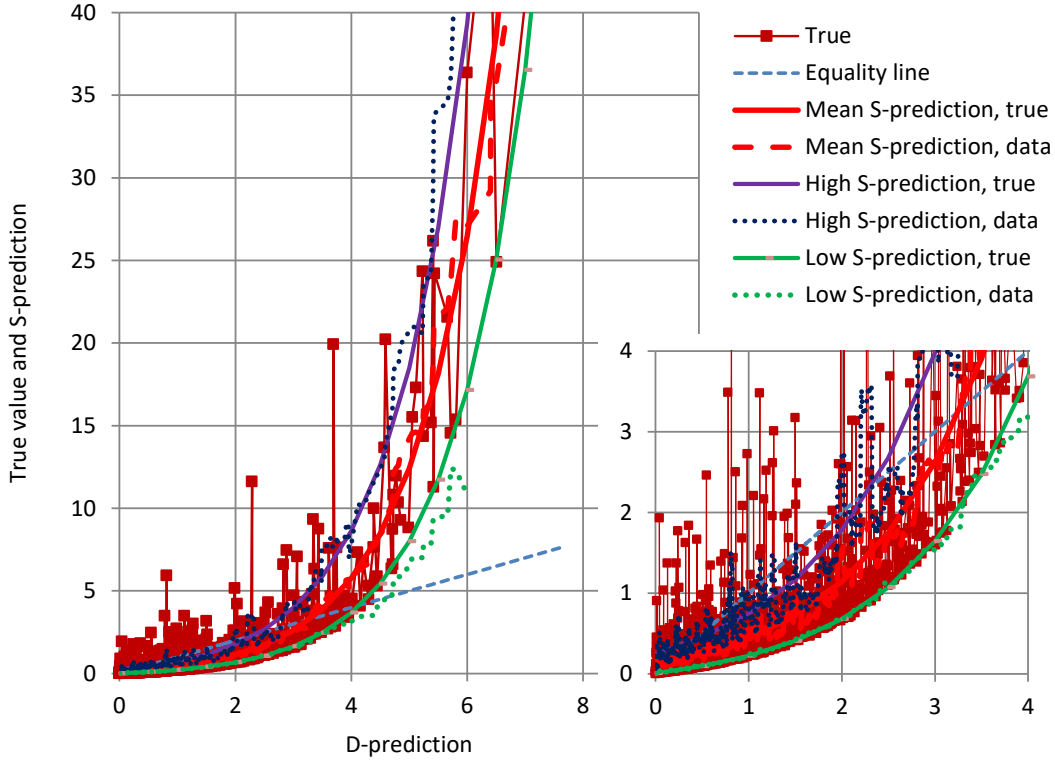
Resulting parameters: $c = 0.325, \xi = 0.02, \zeta = 1$

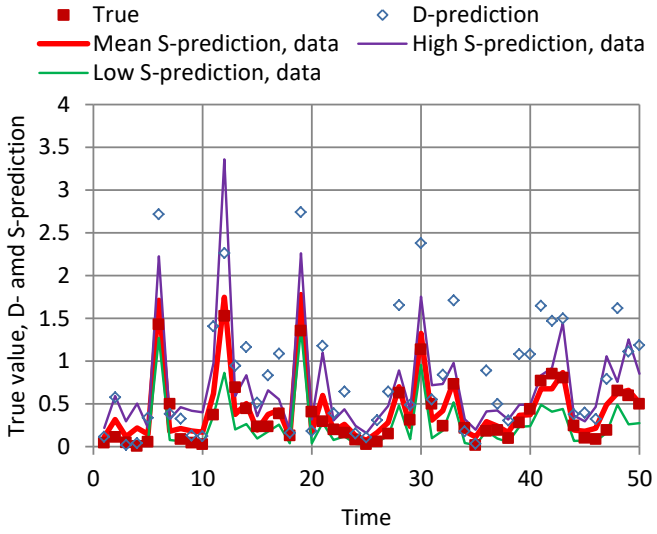
$$\Lambda_1 = 2.75, \Lambda_\infty = 1.81, p_{\mathrm{H}} = 5.01$$

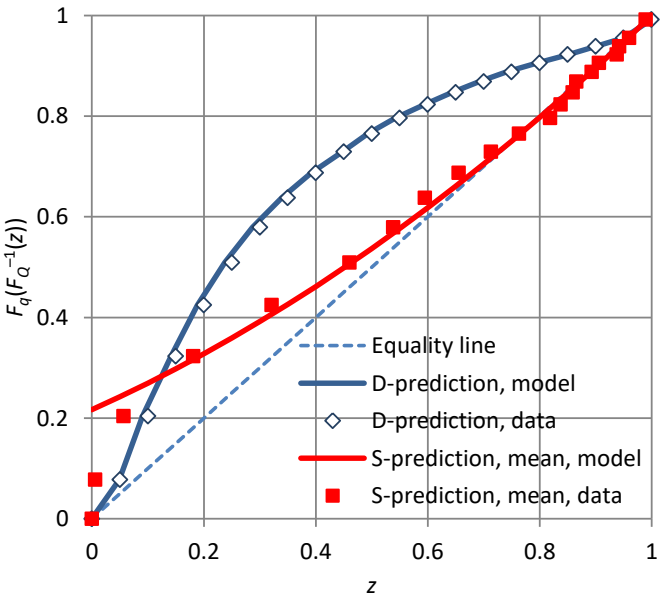
$$\overline{\Lambda}_1 = 1.57, \overline{\Lambda}_\infty = 1, p_{\mathrm{L}} = 9.43$$

Rank correlations: $\hat{r}_{qQ} = 0.998, \hat{r}_{q\mathrm{E}[q|Q]} = 0.9999$

# Toy model results: good D-model




Independent parameters: $a = 0.75, b = a/2 = 0.375$

Choices: $M = 10, P_{\mathrm{L}} = 1 - P_{\mathrm{H}} = 0.1$

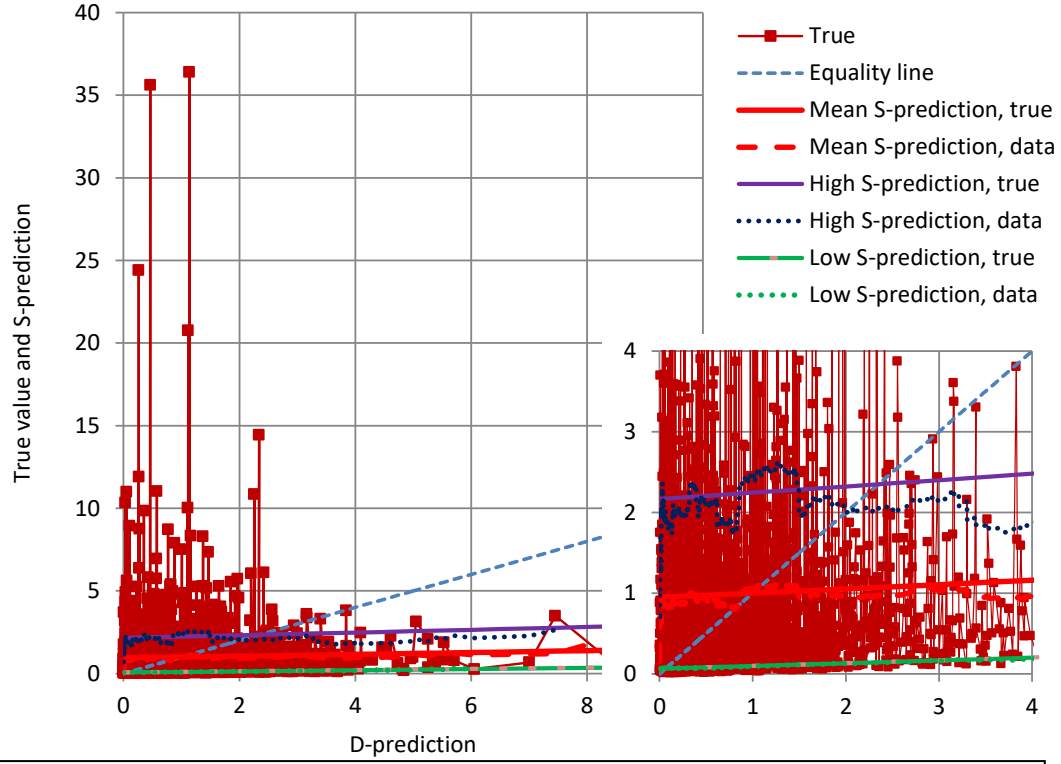
Resulting parameters: $c = 0.185, \xi = 0.375, \zeta = 1$

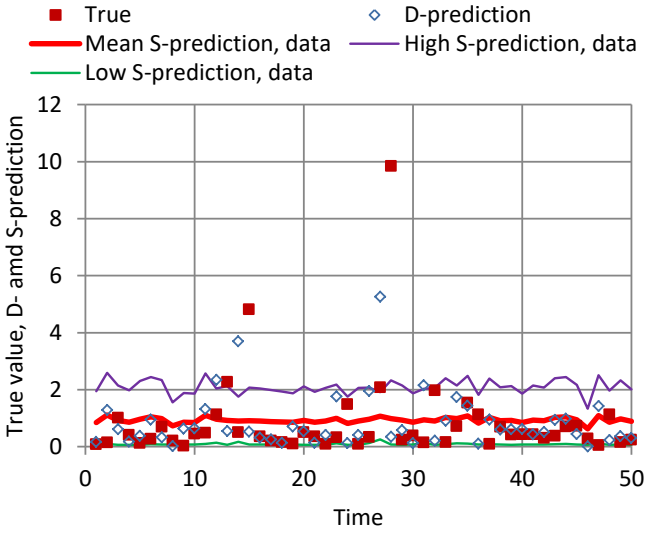
$$\Lambda_1 = 3.50, \Lambda_\infty = 2.62, p_{\mathrm{H}} = 3.48$$

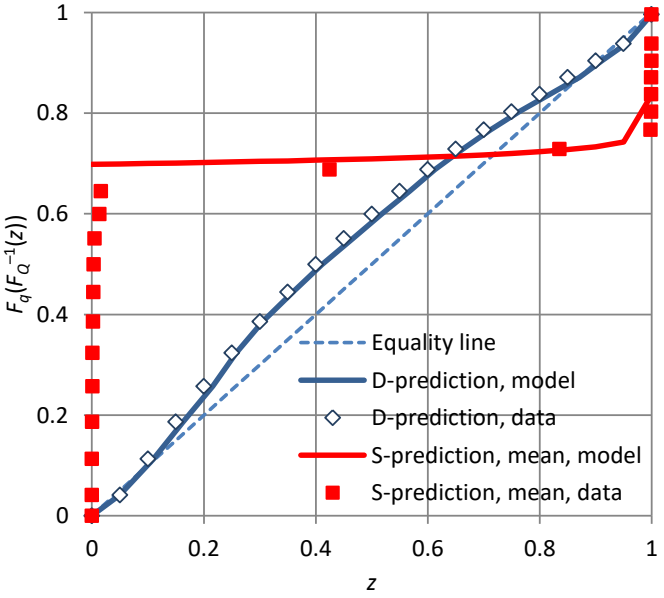
$$\overline{\Lambda}_1 = 1.40, \overline{\Lambda}_\infty = 1, p_{\mathrm{L}} = 9.60$$

Rank correlations: $\hat{r}_{qQ} = 0.84, \hat{r}_{q\mathrm{E}[q|Q]} = 0.97$

# Toy model results: almost irrelevant D-model




Independent parameters: $a = 0.02, b = 0.375$

Choices: $M = 200, F_{\mathrm{L}} = 1 - F_{\mathrm{H}} = 0.1$

Resulting parameters: $c = 1.58, \xi = 0.375, \zeta = 1$

$$\Lambda_1 = 3.50, \Lambda_\infty = 2.62, p_H = 3.48$$

$$\overline{\Lambda}_1 = 1.40, \overline{\Lambda}_\infty = 1, p_{\mathrm{L}} = 9.60$$

Rank correlations: $\hat{r}_{qQ} = 0.09, \hat{r}_{q\mathrm{E}[q|Q]} = 0.66$

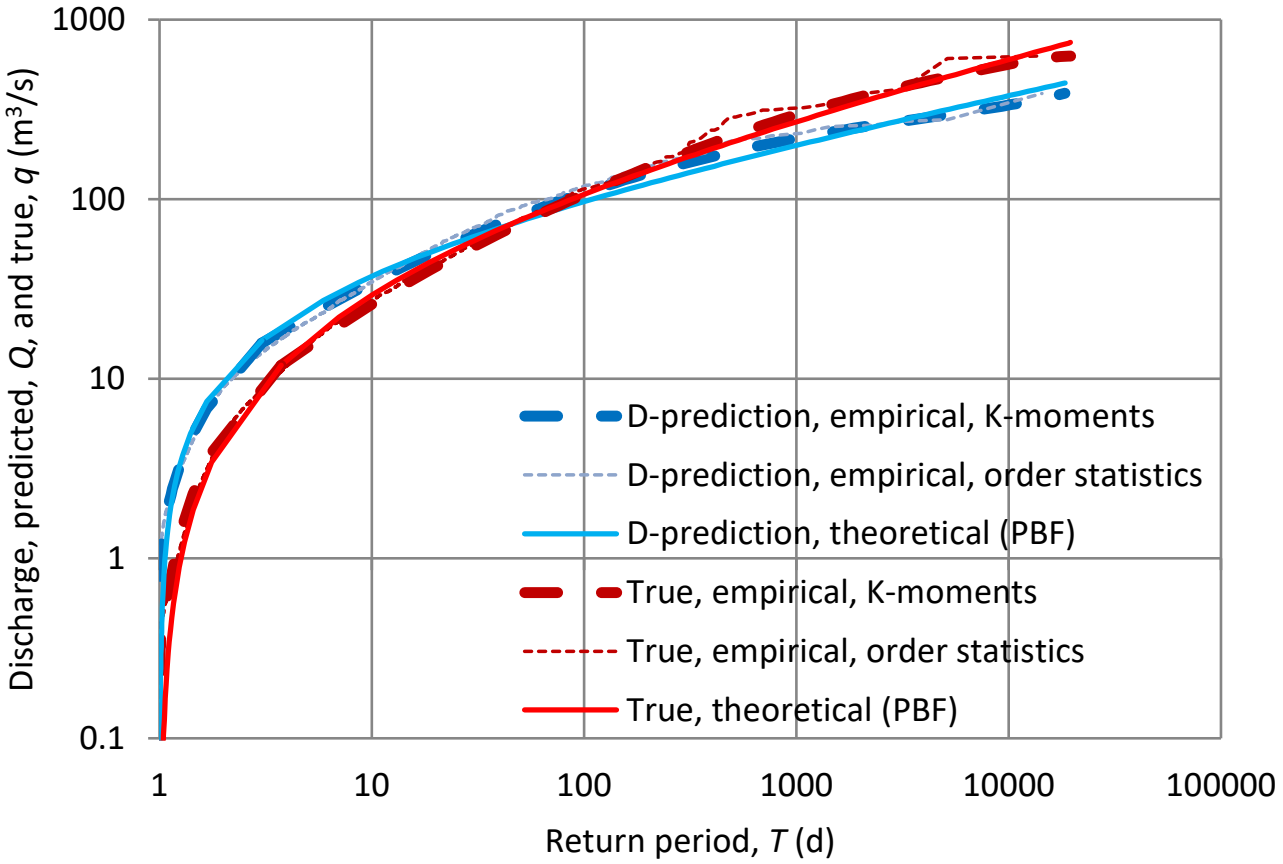
# Real-world case study: Arno River

As a real world case, the observations of daily time series of the Arno River at Subbiano were used. The catchment area is 752 km$^2$. The observations span the 22-year period 1992-2013. The first 20 years are used for model calibration and the last two for model validation. The D-model is the Hymod model (Boyle, 2000; Montanari, 2005) with 5 parameters. For the calibration period the correlation coefficient between the D-model outputs $Q$ and the true values $q$ is 0.87, which means that the model is able to explain $0.87^2 = 75\%$ of the total variance. The Spearman's rank correlation coefficient is also 0.87. These characteristics justify the characterization of the D-model as a good one.

In contrast to the toy-model case, here we do not know the true marginal and conditional distribution functions. It is rather easy to infer the marginal distributions (see next page) and we will use the method proposed to infer the conditional distribution. In absence of knowledge of the true distribution, we assess the appropriateness of the method by testing it for the validation period, where in addition to expected values and confidence limits we perform also a simulation using Method A, whose results are assessed by the CPP plot.

# Real-world case study: Marginal distributions

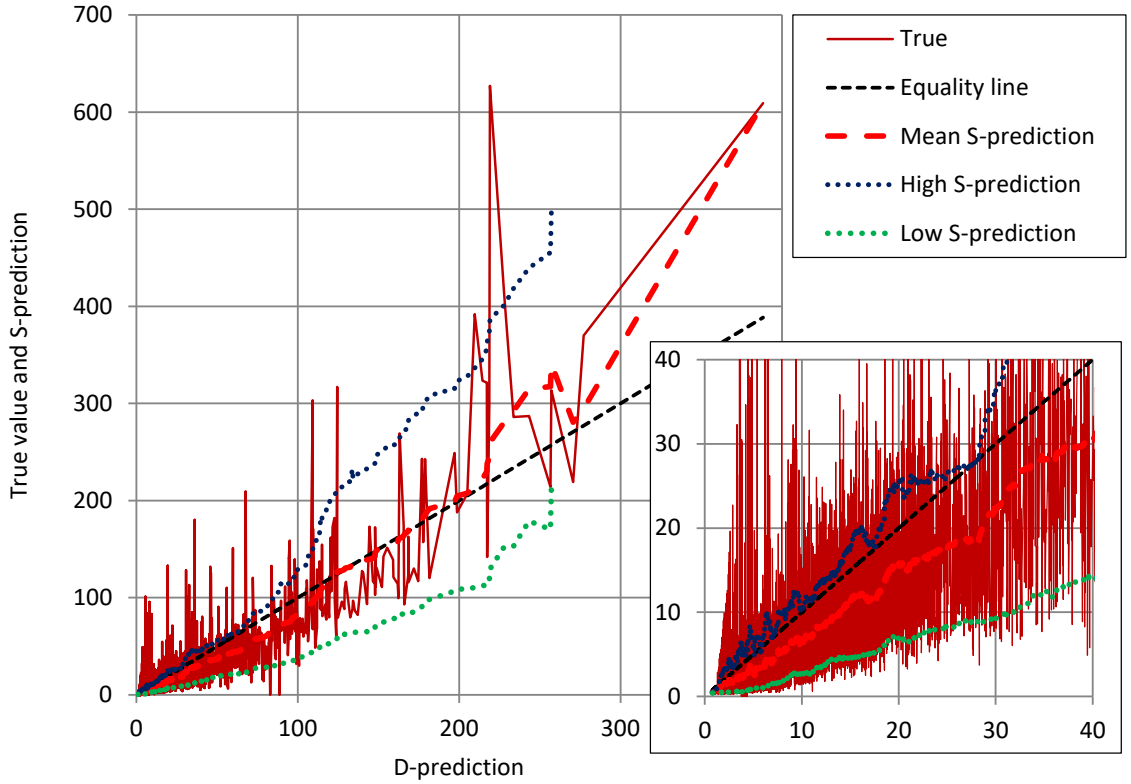
The plot on the right suggests that the PBF distribution describes very well both the true and D-predicted discharge. The fitting was made by a least square method of theoretical and empirical K-moment-based quantiles, using both noncentral K-moments for quantiles larger than the mean and tail-based K-moments for quantiles smaller than the mean. The fitted parameters are shown in the table. The graph also contains quantile estimates based on order statistics (i.e. the well-known plotting positions).

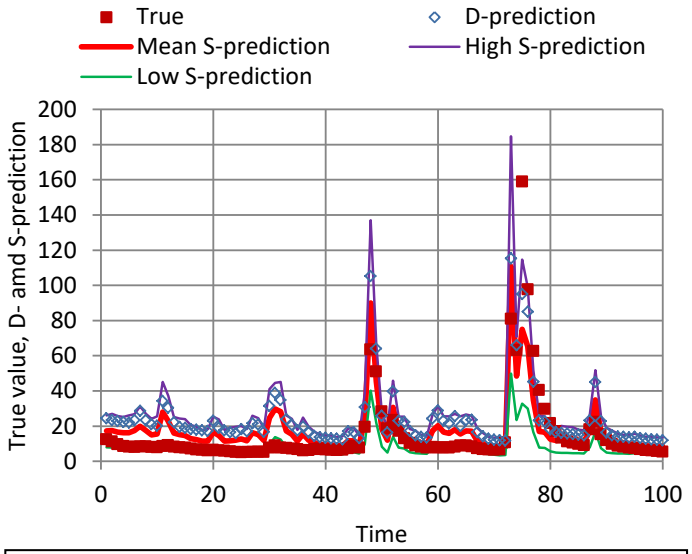



**Parameter estimates**

| ↓Variable - Parameter→ | $\xi$ | $\zeta$ | $\lambda$ (m³/s) |
|---|---|---|---|
| $\underline{q}$ | 0.27 | 0.72 | 6.59 |
| $\underline{Q}$ | 0.24 | 1.11 | 13.64 |

# Real-world case study: Calibration




Choices: $M = 100, F_{\mathrm{L}} = 1 - F_{\mathrm{H}} = 0.1$

Calibrated parameters: $\xi = 0.34, \zeta = 2.49$

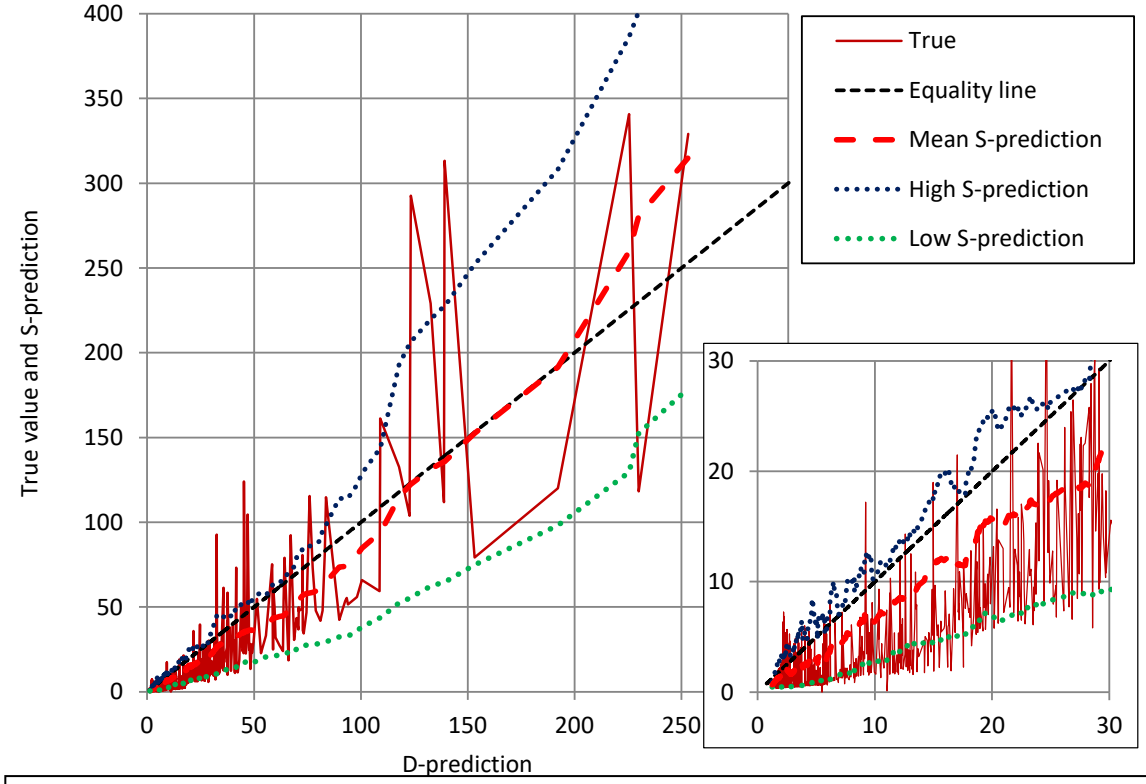
$$\Lambda_1 = 2.81, \Lambda_\infty = 2.52, p_H = 3.8$$

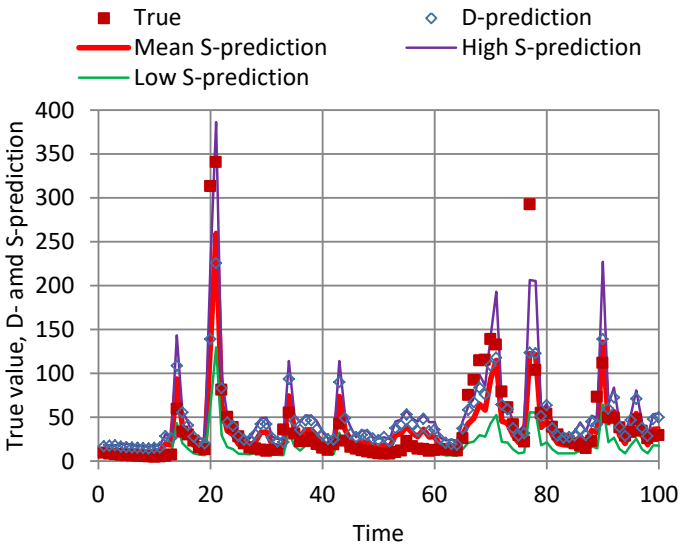
$$\overline{\Lambda}_1 = 1.55, \overline{\Lambda}_\infty = 1.34, p_{\mathrm{L}} = 7.0$$

Rank correlations: $\hat{r}_{qQ} = 0.87, \hat{r}_{q\mathrm{E}[q|Q]} = 0.998$

**Note**: The graph above depicts 100 days of the calibration period, where the first day is 2011-01-01. As shown in all graphs and particularly in the inset, the D-model overpredicts low discharges and underpredicts high ones (this behaviour is also seen in the graph of marginal distributions).

# Real-world case study: Validation



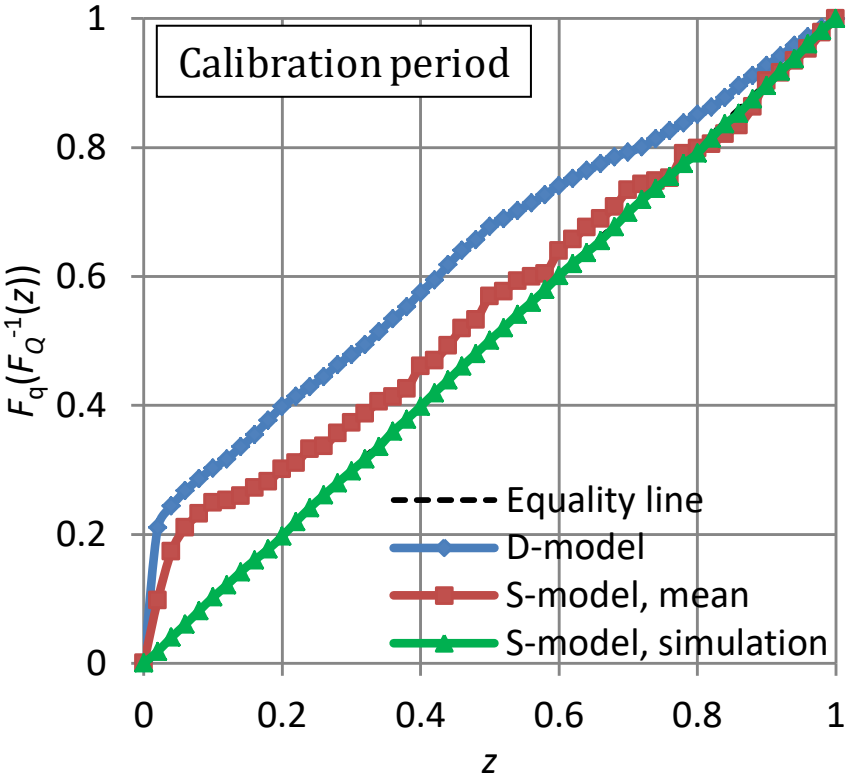



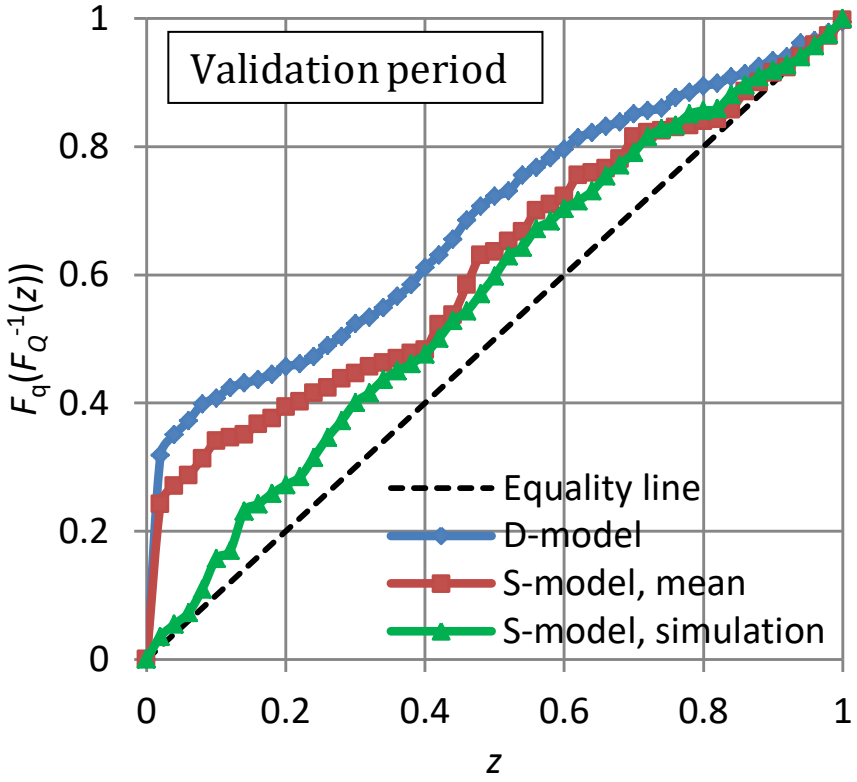

**Note**: The graph on the right depicts 100 days of the calibration period, where the first day is 2013-01-01. All required parameters have been determined in the calibration phase.

Rank correlations: $\hat{r}_{qQ} = 0.87, \hat{r}_{q\mathrm{E}[q|Q]} = 0.87$

Pearson correlations: $\hat{r}_{qQ} = 0.85, \hat{r}_{q\mathrm{E}[q|Q]} = 0.86$

# Real-world case study: CPP plots




In addition to D-predictions and S-expectations ($E[\underline{q}|Q]$), the graphs also show CPP plots for simulated time series, produced by Method A. In the calibration period the CPP of the simulated series aligns perfectly over the equality line.

In the validation period there is a slight discrepancy of the CPP of the simulated series. The reason is the fact that the CPP of the D-model is worse than that of the calibration period and this has some effect on that of the simulated series.

# Conclusions

- Using only observational data along with predictions of a deterministic model (D-model), we can advance the latter into a stochastic model (S-model), with a simple computational framework.
- The stochastic counterpart of the deterministic model accomplishes two important targets:
  a. It corrects systematic discrepancies (biases) of the D-model, whether these are constant or vary with the value of the predictand.
  b. It quantifies the uncertainty of each prediction.
- The framework provided fully adjusts the marginal distribution function of the predictions to that of the true values, thus making a perfect CPP plot. Also it generally improves the Spearman and Pearson correlation coefficients when the conditional expectations of the S-model are used in place of the D-model outputs.
- In the hydrological case study performed, it appears that the Pareto-Burr-Feller distribution can serve as a good model for the marginal distributions of predictands and predictions, as well as for the conditional distribution of predictand given the prediction.
- The newly introduced concept of knowable moments has proved very helpful in the foundation and the application of the framework.

# Appendix 1: Tail-based K-moments

By analogy to the *noncentral knowable moment* (section "A summary of the K-moments approach", p. 10), the *tail-based noncentral knowable moment of order* $(p, q)$ is defined as (Koutsoyiannis, 2020):

$$\overline{K}'_{pq} \coloneqq (p-q+1)\mathrm{E}\left[\left(1-F\left(\underline{x}\right)\right)^{p-q}\underline{x}^{q}\right] = (p-q+1)\mathrm{E}\left[\left(\overline{F}\left(\underline{x}\right)\right)^{p-q}\underline{x}^{q}\right], \qquad p \geq q \tag{57}$$

The most interesting special case is again for $q = 1$, i.e.:

$$\overline{K}'_{p} \coloneqq p\mathrm{E}\left[\left(1-F\left(\underline{x}\right)\right)^{p-1}\underline{x}\right] = p\mathrm{E}\left[\left(\overline{F}\left(\underline{x}\right)\right)^{p-1}\underline{x}\right], \qquad p \geq 1 \tag{58}$$

The tail-based K-moments are connected to expectations of minima by:

$$\overline{K}'_{p} = \mathrm{E}[\underline{x}_{(p)}] = \mathrm{E}\left[\min\left(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_p\right)\right] \tag{59}$$

The same equation (30) will give an unbiased estimator $\overline{K}'_{p}$ if we revere the order of the sample, i.e. if we replace $\underline{x}_{(i:n)}$ with $\underline{x}_{(n-i+1:n)}$. Likewise, we can introduce the tail-based $\Lambda$-coefficient of order $p$ as:

$$\overline{\Lambda}_p \coloneqq \frac{1}{p\,F\left(K'_p\right)} \tag{60}$$

$\overline{\Lambda}_p$ has similar properties with $\Lambda_p$ and in particular varies only slightly with $p$. For $p = 1$ it is readily seen that

$$\overline{\Lambda}_1 = 1/F(\mu) = \Lambda_1/(\Lambda_1 - 1) \tag{61}$$

The limiting value $\overline{\Lambda}_\infty$ depends only on the lower tail index $\zeta$ of the distribution:

$$\overline{\Lambda}_\infty = \Gamma(1+1/\zeta)^{-\zeta} \tag{62}$$

# Appendix 1: Tail-based K-moments (2)

A simple approximation of $\overline{\Lambda}_p$ and hence of the non-exceedance probability is:

$$\overline{\Lambda}_p \approx \overline{\Lambda}_\infty + \frac{\overline{\Lambda}_1 - \overline{\Lambda}_\infty}{p}, \qquad F\left(K'_p\right) \approx \frac{1}{\overline{\Lambda}_\infty p + \left(\overline{\Lambda}_1 - \overline{\Lambda}_\infty\right)} \tag{63}$$

Conversely, for a given non-exceedance probability $F$, we can calculate the quantile $x$ as the $\overline{K}'_p$ that corresponds to:

$$p \approx \frac{1}{\overline{\Lambda}_\infty F} + 1 - \frac{\overline{\Lambda}_1}{\overline{\Lambda}_\infty} \tag{64}$$

# Appendix 2: The Pareto-Burr-Feller (PBF) distribution

The *Pareto-Burr-Feller* (PBF) distribution, named thus by Dimitriadis (2017), is also known as *Pareto III and IV*, *Burr XII* and *Feller*. Its probability density and distribution functions are, respectively:

$$f(x)=\frac{\zeta}{\lambda}\left(\frac{x}{\lambda}\right)^{\zeta-1}\left(1+\xi\zeta\left(\frac{x}{\lambda}\right)^{\zeta}\right)^{-\frac{1}{\zeta\xi}-1},\qquad F(x)=1-\left(1+\xi\zeta\left(\frac{x}{\lambda}\right)^{\zeta}\right)^{-\frac{1}{\xi\zeta}},\qquad x,\xi\geq 0,\zeta,\lambda>0 \tag{65}$$

The parameter $\lambda$ is a scale parameter with units $[x]$ and the parameters $\xi,\zeta$ are dimensionless shape parameters, known as (higher) tail index and lower tail index, respectively.

Because of the analytical equations of both the density and distribution functions, it is a very convenient computationally. Because of its zero lower bound it is a realistic representation for many physical quantities. Because of its two shape parameters it is quite flexible. In fact, it contains as special cases the ditributions: Weibull ($\xi=0$), Pareto ($\zeta=1$) and exponential ($\xi=0,\zeta=1$). It admits analytical relationships for the classical noncentral moments (about the origin):

$$\mu'_p=\lambda^p\frac{p}{\zeta(\xi\zeta)^{p/\zeta}}\mathrm{B}\left(\frac{1}{\xi\zeta}-\frac{p}{\zeta},\frac{p}{\zeta}\right) \tag{66}$$

as well as for the tail-based K-moments:

$$\overline{K}'_{pq}=\lambda^q\frac{q}{\zeta}\frac{1}{(\zeta\xi)^{q/\zeta}}\mathrm{B}\left(\frac{p-q+1}{\zeta\xi}-\frac{q}{\zeta},\frac{q}{\zeta}\right) \tag{67}$$

Clearly, the classical moment of order $p$ and the K-moment of orders $(p,q)$ exist if $p<1/\xi$ or $q<1/\xi$, respectively. The characteristic $\Lambda$-coefficients are:

$$\Lambda_1=\left(1+\left(\frac{\mathrm{B}\left(\frac{1}{\zeta\xi}-\frac{1}{\zeta},\frac{1}{\zeta}\right)}{\zeta}\right)^{\zeta}\right)^{1/\zeta\xi},\qquad \overline{\Lambda}_1=\frac{\Lambda_1}{\Lambda_1-1},\qquad \Lambda_\infty=\Gamma(1-\xi)^{\frac{1}{\xi}},\qquad \overline{\Lambda}_\infty=\Gamma\left(1+\frac{1}{\zeta}\right)^{-\zeta} \tag{68}$$

# Appendix 3: Determination of tail indices from K-moment orders

We assume that we have bracketed a distribution through its quantiles for two non-exceedance probabilities $F_{\mathrm{H}} > 0.5$ and $F_{\mathrm{L}} < 0.5$. We have empirically (based on data) estimated the K-moment orders $p_{\mathrm{H}}$ and $p_{\mathrm{L}}$ that correspond to $F_{\mathrm{H}}$ and $F_{\mathrm{L}}$. We wish to estimate the tail indices $\xi$ and $\zeta$. According to (37), (64) and (61) we have:

$$p_{\mathrm{H}} \approx 1 - \frac{\Lambda_1}{\Lambda_\infty} + \frac{1}{\Lambda_\infty (1 - F_{\mathrm{H}})}, \qquad p_{\mathrm{L}} \approx \frac{1}{\overline{\Lambda}_\infty F_{\mathrm{L}}} + 1 - \frac{\overline{\Lambda}_1}{\overline{\Lambda}_\infty} = \frac{1}{\overline{\Lambda}_\infty F_{\mathrm{L}}} + 1 - \frac{\Lambda_1}{\overline{\Lambda}_\infty (\Lambda_1 - 1)} \tag{69}$$

Assuming $F_{\mathrm{L}} = 1 - F_{\mathrm{H}} = F$ we get

$$p_{\mathrm{H}} \approx 1 - \frac{\Lambda_1}{\Lambda_\infty} + \frac{1}{\Lambda_\infty F}, \qquad p_{\mathrm{L}} \approx \frac{1}{\overline{\Lambda}_\infty F} + 1 - \frac{\Lambda_1}{\overline{\Lambda}_\infty (\Lambda_1 - 1)} \tag{70}$$

If we assume that the quantity of interest follows the PBF distribution, then $\Lambda_1, \overline{\Lambda}_1, \Lambda_\infty, \overline{\Lambda}_\infty$ are given by (68). Combining (69), (70) and (68) we can find the two unknown tail indices $\xi$ and $\zeta$ from the estimated $p_{\mathrm{H}}$ and $p_{\mathrm{L}}$. A better strategy is to use several values of $F$, estimate the related $p_{\mathrm{H}}$ and $p_{\mathrm{L}}$ for each $F$ and finally calculate $\xi$ and $\zeta$ by minimizing the deviations of theoretical and estimated sets of $p_{\mathrm{H}}$ and $p_{\mathrm{L}}$.

# Appendix 4: Choice of sub-sample sizes $m_1$ and $m_2$

The method followed for the choice of sample sizes $m_1$ and $m_2$ is empirical and was tested by extended simulations. A more theoretical account is possible but would distract the focus of the approach presented which is the ultimate simplicity. Simulation results have shown that setting equal numbers $m$ gives good results for the body of the distribution but in the upper tail some differentiation is required. Furthermore, a decrease of these numbers is necessary for both tails down to the minimum possible required for a chosen confidence coefficient. Algorithmically, the procedure chosen in order to estimate $F_{q|Q}\left(q\middle|Q_{(i:n)}\right)$ from $F_{q|[Q]}\left(q\middle|Q_{(i:n)}, \Delta F_1, \Delta F_2\right)$ is described by the following steps.

1. We choose a number $M$ equal to a small multiple of the inverse of the minimum probability that we are seeking for bracketing the probabilistic predictions. For example, if we want to bracket the probabilistic predictions between 10% and 90%, the inverse of 10% is 10 and we could choose $M = 10$ to 20.
2. For $M + 1 \leq i \leq n - 2M$ we set $m_1 = m_2 = M$.
3. For $i \leq M$ we set $m_1 = m_2 = i - 1$.
4. For $i \geq n - 2M + 1$ we set $m_2 = \min\{M, n - i\}$ and $m_1 = \lfloor m_2/a_i \rfloor \geq m_2$, where $a_i = \min\{1, 1/2 + (n - i)/4M\}$; notice that $1/2 + (n - i)/4M$ is a linear function of $i$ with minimum value $a_i = 0.5$ for $i = n$ and maximum $a_i = 1$ for $i = n - 2M$.
5. If the resulting curve $F_{q|Q}\left(q\middle|Q_{(i:n)}\right)$ is too rough, we increase $M$ and repeat the procedure from step 1.

# References


Boyle, D.P., 2000. Multicriteria calibration of hydrological models. Ph.D. dissertation, Dep. of Hydrol. and Water Resour., Univ. of Ariz., Tuscon.

Koutsoyiannis, D., 2019a. Knowable moments for high-order stochastic characterization and modelling of hydrological processes. *Hydrological Sciences Journal*, 64(1), 19–33, doi:10.1080/02626667.2018.1556794.

Koutsoyiannis, D., 2020. *Stochastics of Hydroclimatic Extremes – A Cool Look at Risk* (in preparation).

Koutsoyiannis, D., and Montanari, A., 2015. Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal*, 60 (7-8), 1174–1183, doi: 10.1080/02626667.2014.959959.

Montanari, A., 2005. Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. *Water Resour. Res.*, 41, W08406, doi: 10.1029/2004WR003826

Montanari, A., and Koutsoyiannis, D., 2012. A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research*, 48, W09555, doi: 10.1029/2011WR011412.

Montanari, A., and Koutsoyiannis, D., 2014a. Reply to comment by G. Nearing on ''A blueprint for process-based modeling of uncertain hydrological systems''. *Water Resources Research*, 50 (7), 6264–6268, doi: 10.1002/2013WR014987.

Montanari, A., and Koutsoyiannis, D., 2014b. Modeling and mitigating natural hazards: Stationarity is immortal!. *Water Resources Research*, 50 (12), 9748–9756, doi: 10.1002/2014WR016092.

Nearing, G., 2014. Comment on "A blueprint for process-based modeling of uncertain hydrological systems" by Alberto Montanari and Demetris Koutsoyiannis. *Water Resources Research*, 50, doi: 10.1002/2013WR014812.

Papacharalampous, G., Tyralis, C., Koutsoyiannis, D., and Montanari, A., 2019a. Quantification of predictive uncertainty in hydrological modelling by harnessing the wisdom of the crowd: Methodology development and investigation using toy models, Advances in Water Resources, doi: 10.1016/j.advwatres.2019.103471.

Papacharalampous, G., Tyralis, C., Koutsoyiannis, D., and Montanari, A., 2019b. Quantification of predictive uncertainty in hydrological modelling by harnessing the wisdom of the crowd: A large-sample experiment at monthly timescale, Advances in Water Resources, doi: 10.1016/j.advwatres.2019.103470.

Sikorska, A., Montanari, A., and Koutsoyiannis, D., 2015. Estimating the uncertainty of hydrological predictions through data-driven resampling techniques, *Journal of Hydrologic Engineering (ASCE)*, 20 (1), doi: 10.1061/(ASCE)HE.1943-5584.0000926.