ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Thanks to Tsinghua University for hosting me this week

From hydrological complexity to hydrological predictions, through the exploration of uncertainty, randomness and stochastic theory

Alberto Montanari[1]

[1]Department of Civil, Chemical,
Environmental and Material Engineering
University of Bologna

October 22, 2025

"... although the future is not predictable in any detail, it is manageable as an aggregate phenomenon."
Herbert A. Simon, The Sciences of the Artificial

Hydrology@Unibo

Hydrology, Hydraulic Works, Coastal Engineering, Water Resources @Unibo

Armando Brath
Water resources flood risk

Elena Toth
Floods and droughts, Water supply systems, Machine Learning

Attilio Castellarin
Extreme events, Hydrological regionalisation

Günter Blöschl
Climate change, flood risk, socio-hydrology

Serena Ceola
Big data, large scale hydrology, artificial intelligence

Marco Maglionico
Urban hydrology, Urban flooding, Urban water quality

Alessio Domeneghetti
Flood models,
Civil protection

Cristiana Bragalli
Water supply,
Experimental hydraulics

Mattia Neri
Rainfall-runoff model regionalisation,
Artificial intelligence,
Climate change

Alberto Montanari
Hydrological models, Hydrological change,
Climate change, Uncertainty

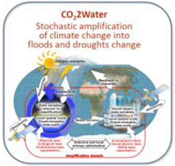
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Hydrology@Unibo

Research fields

- Flood and drought prediction;
- Climate change and climate impact amplification;
- Big-data and artificial intelligence in hydrology;
- Renewable energy, decarbonisation;
- Hydrological cycle under climate change;
- Advanced monitoring systems for water resources and floods;
- Advanced methods for the simulation of climate scenarios;
- Uncertainty assessment for environmental models;
- Stochastic theories for hydrological models and uncertainty assessment.



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Master Degree Climate Change Adaptation

Master Degree Programme in Civil Engineering – Climate Change Adaptation

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Second Cycle Degree/Two Year Master in
Civil Engineering

Home | Programme | Admission | Studying | Opportunities | Notice Board | Contacts

Design and innovate

An advanced international path between *structural engineering, hydraulics, transport and adaptation* to climate change.

Overview

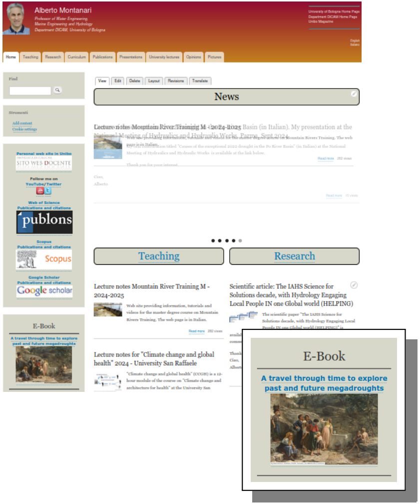
Place of teaching Bologna, Imola, Ravenna	Language English	Degree Programme Class LM-23 R - Civil engineering	Degree Programme Director Nicola Buratti
Type of access Open access	International programmes Delivered entirely in English, Double/Multiple degree	Department Civil, Chemical, Environmental, and Materials Engineering - DICAM	Proposed paths (Curricula) SUSTAINABLE MOBILITY IN URBAN AREAS • WATER ENGINEERING • CLIMATE CHANGE ADAPTATION • STRUCTURAL ENGINEERING

https://corsi.unibo.it/2cycle/CivilEngineering

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

My web site: open education, teaching, research, software, data

- This presentation is available in my website at www.albertomontanari.it.
- Everything is open, can be copied and reproduced (citation is appreciated).
- Everything I do is there. Lecture notes (in English), videos of lectures, scientific papers, presentations, software.
- Highlights: A travel through time to explore past and future megadroughts (interactive ebook); Floods, ten concepts to get protected; Climate change and water: lecture to students.
- Openness stands for dissemination and inclusivity.

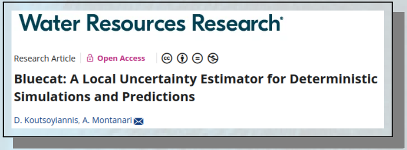


Uncertainty assessment for environmental models: looking for simplicity and operational efficiency

www.albertomontanari.it/bluecat

Koutsoyiannis and Montanari (2022) & Montanari and Koutsoyiannis (2025). **BLUECAT transforms a deterministic prediction model into a stochastic prediction model**.

- From a point prediction we obtain the probability distribution of the predictand. From the above probability distribution we estimate the average (or median) prediction along with its confidence band for an assigned confidence level.
- BLUECAT is more than an uncertainty assessment method: it is rather a new prediction model with a stochastic structure. BLUECAT can be applied in conjunction with any deterministic prediction model.



Water Resources Research

Research Article | Open Access

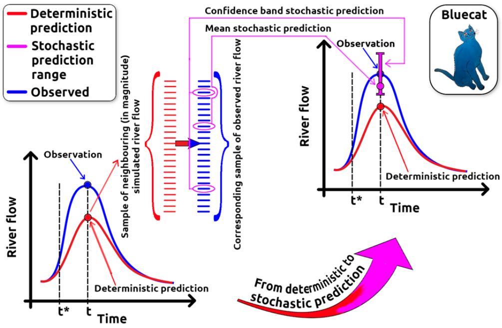
Bluecat: A Local Uncertainty Estimator for Deterministic Simulations and Predictions

D. Koutsoyiannis, A. Montanari

BLUECAT: workflow

How does Bluecat work?

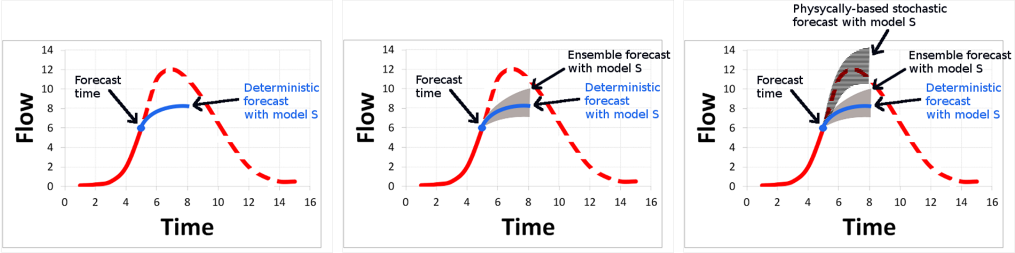
- A point prediction is obtained with a deterministic model.
- A sample of neighbouring (in magnitude to the point prediction) simulated river flows is collected from the calibration period.
- A corresponding sample of observed river flows is collected. This is used to estimate (from data) the probability distribution of the predictand conditioned on the point prediction.
- The stochastic prediction is obtained along with uncertainty assessment.
- **Targeted selection of neighbours leads to identifying what is causing uncertainty.**



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

BLUECAT: features

Note: the previous point prediction is not necessarily included in the confidence band. The latter is displaced around the stochastic prediction (not around the point prediction).



- BLUECAT updates the model.
- Uncertainty is not necessarily displaced around the deterministic prediction. BLUECAT corrects the bias, which may vary with the river flow.

BLUECAT: assumptions

Introduction
Uncertainty and Bluecat
Water Cycle and Climate
Reasons for Complexity in Hydrology
Complexity implies Limited Predictability
Statistical Predictability
Leverage Points
Conclusions

Note: in what follows, deterministic and stochastic model are D-model and S-model, respectively. BLUECAT assumptions:

- The stochastic processes describing the modelled variables are stationary during the calibration and application period. Non-stationarity can be accounted for by using non-stationary D-models.
- The calibration data set is extended enough to ensure that sufficient information is available to upgrade the D-model into the S-model.
- Uncertainty is not subdivided in different components as BLUECAT is assumed to automatically incorporate all types, including the uncertainty in input data and parameters, for which no particular provision is necessary.

Why BLUECAT?
"Brisk Local Uncertainty Estimator for generiC simulations And predicTions"

BLUECAT refers to the pop-art by Andy Warhol (1928–1987), a success creation stimulated by a simple idea that gives a feeling of positive thinking and optimism.

BLUECAT: a bit of theory

To advance from the D-model to the S-model in we need to specify the conditional distribution:

$$\boxed{F_{q|Q}(q|Q) = P\left\{\underline{q} \leq q|\underline{Q} = Q\right\},}$$

where q and Q are concurrent observed and simulated flow, respectively, and stochastic variables are underlined.
If Q and q are concurrent time series, each of size n, and if $Q_{(i:n)}$ is the *ith* smallest value in Q and $q_{(j:n)}$ is the *jth* smallest value in q, then the approximations $F_Q(Q_i) \approx i/n$ and $F_q(q_j) \approx j/n$ can be used. After mathematical development we we prove that $F_{q|Q}$ can be obtained by minimizing the quantity

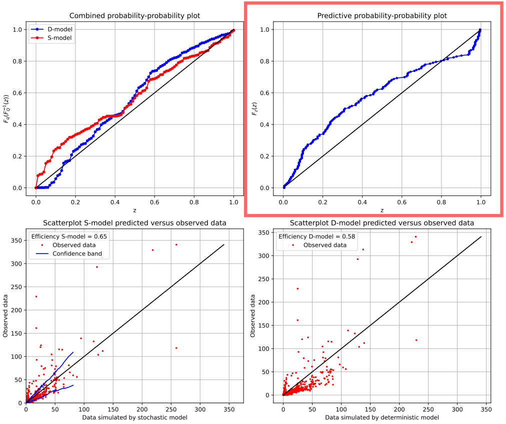
$$\boxed{A := \sum_{j=1}^{n} (B_j - j)^2 = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} F_{q|Q}\left(q_{(j:n)}|Q_{(i:n)}\right) - j \right)^2,}$$

therefore getting the desired conditional distribution which leads to the formulation of the S-model corresponding to the D-model.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

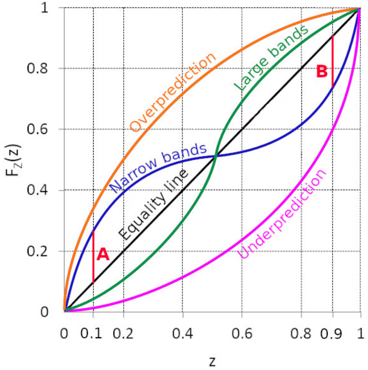
Hypothesis testing

- Testing stochastic physically-based assumptions through validation of the confidence band of the prediction.
- Verification of the reliability of uncertainty assessment.
- The only solution to quantitative testing of uncertainty assessment is comparison with observed data.
- Combined probability-probability plot (CPP), Predictive probability-probability plot (PPP).
- Embedded in the BLUECAT software.



Predictive probability-probability plot (PPP)

- Rigorous approach, surprisingly not much popular (perhaps because it is rigorous).
- First proposed by Laio and Tamea (2007). And then discussed by several authors, including Eslamian (2014), K&M (2022), M&K (2024).
- It tests whether probability of the observed data evaluated by BLUECAT are uniform. A straightforward and quick test.
- PPP plots the above probabilities against the sample frequency of the observed values. These are expected to fall within the BLUECAT confidence band with probability $1-\alpha$ where α is the significance level.

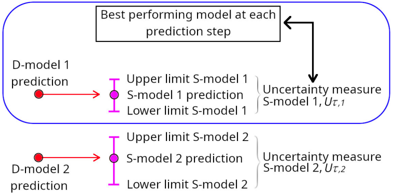


ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Ensemble simulation with BLUECAT

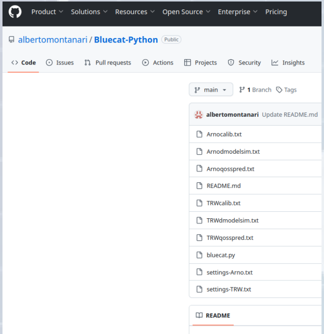
- Multimodel simulation is attractive (for example, a rainfall-runoff model with different parameters for different regimes).
- Two challenges: (a) how to combine different predictions (Bayesian averaging is an option) (b) how to estimate uncertainty for the obtained combination (remember: the pure ensemble spread does not suffice to provide a comprehensive estimate of uncertainty).
- BLUECAT: uncertainty of each model estimated at each prediction step as a criteria to select the optimal ensemble member (M&K, 2024).

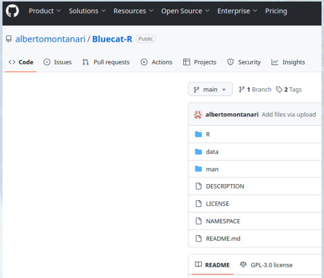
A single model prediction corresponding to the least uncertain ensemble member, that is identified through a proper measure, is picked up at each prediction step.

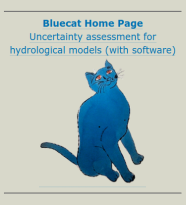


BLUECAT software (www.albertomontanari.it/bluecat)

- Python: https://github.com/albertomontanari/Bluecat-Python
- R: https://github.com/albertomontanari/Bluecat-R







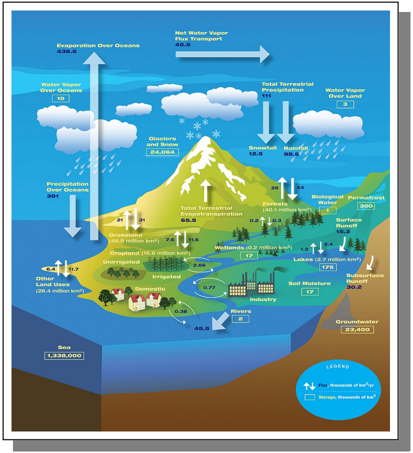
Software are accompanied by embedded help and examples of applications for full reproducibility.

Going beyond the data driven world

- BLUECAT is data driven. What if data are not available or limited?
- How to estimate uncertainty for PUB?
- A broader understanding of uncertainty is needed.
- What is causing uncertainty?
- Uncertainty can be eliminated?
- Uncertainty means limited predictability which in turn is due to complexity. We need to inspect the link between complexity and predictability.

Does uncertainty have a physical meaning?

The Water Cycle: a system with complex dynamics

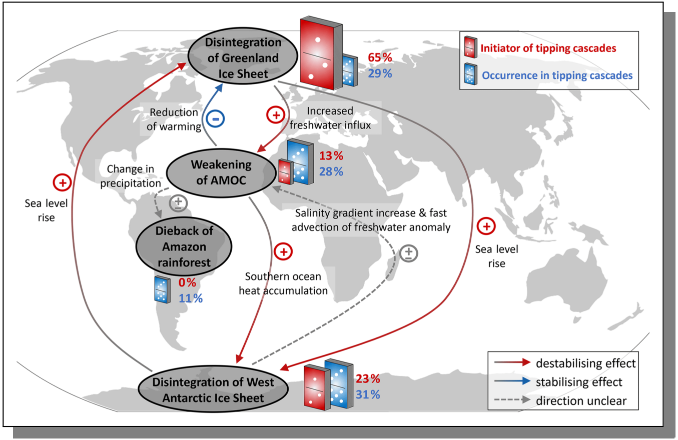


By Atmospheric Infrared Sounder, Public domain, via Wikimedia Commons, NASA/JPL

The water cycle: a complex system

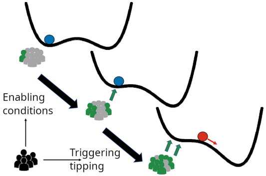
- River flow is determined by several connected systems forming an integrated entity;
- Traditionally these systems are studied and modelled separately. Today we know that to understand patterns we need a holistic approach;
- Processes in a river basin are mostly governed by physical laws that can be expressed in **deterministic form;**
- Models of the water cycle are essentially deterministic.

Complexity in Climate



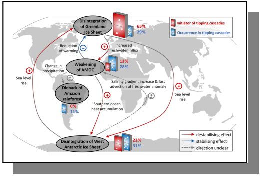
By N. Wunderling et al., CC BY 4.0, via Wikimedia Commons

Complexity in the climate systems turns into **tipping points**: critical thresholds that, when crossed, lead to large, accelerating and may be irreversible changes. Tipping points may occur in the hydrosphere and the anthroposphere as well.



Modified after TM. Lenton et al., CC BY-SA 4.0, via Wikimedia Commons

Complexity: definition



By N. Wunderling et al., CC BY 4.0, via Wikimedia Commons

- Physicist Neil Johnson stated that "...even among scientists, there is no unique definition of complexity – and the scientific notion has traditionally been conveyed using particular examples..."
- Complexity arises from non-linear dynamics, sensitivity to initial conditions, chaos;
- Complexity characterizes the behavior of a system or model whose components interact in multiple ways and follow local rules (from Wikipedia);
- **Complexity implies lack of deterministic predictability**.

We will get back to it later →

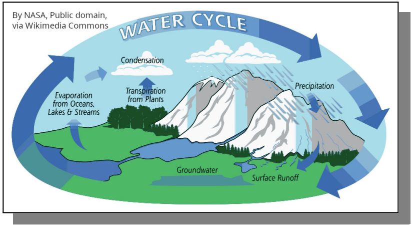
Complexity: "my" definition

- **A complex system exhibits patterns structured by deterministic laws enabling limited predictability.**
- The structure of the system cannot be fully compressed without loss, yet is not irreducible randomness.
- It is located somewhere between a fully predictable system (no complexity) and a fully random system (no complexity).
- Typically, the system is predictable up to a finite and limited time horizon.

Pictofigo, CC BY-SA 3.0, via Wikimedia Commons

Weather forecasting stone by N Chadwick, CC BY-SA 2.0, via Wikimedia Commons

Reasons for Complexity in Hydrology



By NASA, Public domain, via Wikimedia Commons

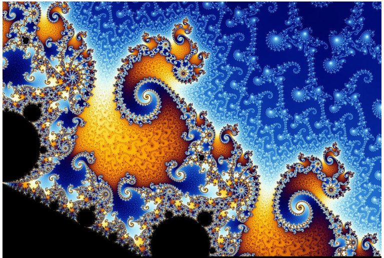
Why the water cycle is complex:

- Initial and boundary conditions and forcings cannot be precisely determined;
- Processes are sensitive to initial conditions (butterfly effect) in particular at small scales;
- Unpredictable behaviours in connected systems (social dynamics);
- Physical laws apply at the microscopic scale;
- **Scaling from microscale behaviours to macroscale needs perfect knowledge of the microscale geometry of the system.** →

By Русский: Игорь Шпиленок, CC BY-SA 3, via Wikimedia Commons

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Upscaling microscopic behaviours



Created by Wolfgang Beyer with the program Ultra Fractal 3., CC BY-SA 3.0, via Wikimedia Commons

The illusion of fractals - One of the major illusions of the new millenium

- Some 40 years ago we dreamed to discover upscaling laws for natural systems;
- A precise upscaling law would allow us to make perfect scaling of environmental systems from microscopic to macroscopic scales;
- Such deterministic description is still to be discovered;

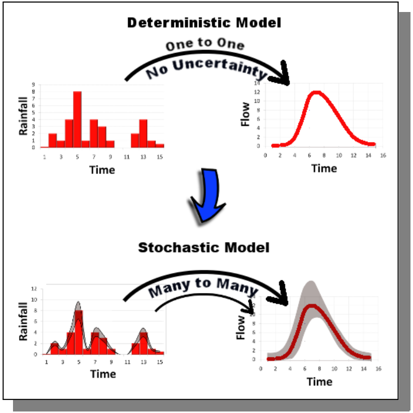
- In abscence of a deterministic relationship to upscale, we resort to a statistical (stochastic) description;
- Macroscale behaviours are computed by estimating their probability distribution. How?

A Phylosophical Consideration

New millenium illusions:

- Endless Progress: The belief that technology and globalization would automatically create peace, prosperity, and equality for all.
- Connection: we felt "more connected than ever," but loneliness, polarization, and misinformation grew.
- Security: Before events like 9/11, financial crashes, pandemics, many believed the world was stable and predictable.
- Infinite Resources: Assuming economic growth and consumption could continue without ecological limits.
- Truth in the Digital Age: With the explosion of information online, many assumed access = accuracy, but it also fueled fake news and disinformation..
- Control: Believing advanced systems (finance, climate management, AI, biotech) are fully under human control, when in reality they often spiral unpredictably.
- Stock market: The Belief that stock market had the capability to self-compensate to ensure equitable distribution of resources and liberalism.

Complexity implies Limited Predictability



Whats does limited predictability means?

- The above reasoning has demonstrated that complexity implies limited **deterministic** predictability;
- However, complexity does not preclude **statistical** predictability;
- In other words, if we cannot make a perfect prediction, we can still make a statistical prediction. How to describe predictions probabilistically?
- This is an open question in hydrology since the 1960s!

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Getting to Target: Exploiting Statistical Predictabilty

Let's suppose that we deal with a generic deterministic predictive model S for a time series (can be a space time model, the only requirement is that it is deterministic):

$$\boxed{Q(t) = S\left(\boldsymbol{\Theta}, \mathbf{X}(t)\right) + e(t)}$$

where $Q(t)$ is the true value to be predicted, $\boldsymbol{\Theta}$ and $\mathbf{X}$ are model parameter vector and the vector of input variables, respectively, and e is prediction error.

In what follows index t is dropped for clarity.

To exploit statistical predictability we need to transform the above deterministic model into a stochastic model which gives the prediction in terms of probability distribution:

$$\boxed{f_Q\left(Q\right) = K_S f_{\boldsymbol{\Theta},\mathbf{X}}\left(\boldsymbol{\Theta}, \mathbf{X}\right)}$$

How to shape it?

where f indicates a (time varying) probability density function, and K_S is a stochastic operator depending on model S, which operates a shift from one (deterministic) to many (stochastic) predictions.

Stochastic physically-based framework

Getting to Target: Exploiting Statistical Predictabilty

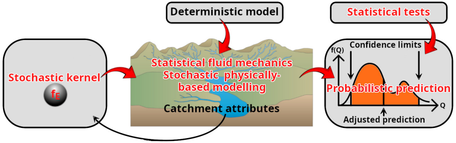
Montanari and Koutsoyiannis (2012) and Montanari (2025, unpublished) proved that the previous equation - under mild assumptions - can be written in the simple form:

$$f_Q\left(Q\right) = f_E\left(Q - S(\boldsymbol{\Theta}, \mathbf{X})\right)$$

where f_E – in simple words – is the time-varying probability density distribution of the model error.

The key question now becomes the definition of f_E - which depends on model prediction $S(\boldsymbol{\Theta}, \mathbf{X})$ - basing on our **physical knowledge of the system**, by building on the above premises.

What are the hydrological structures, features and patterns determining f_E?



Stochastic kernel f_E: Probability distribution of deterministic model prediction error

Stochastic physically-based framework

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Physics Informed Estimation of Stochastic Kernel

Distribution f_E depends on physical behaviours of the system, like for instance catchment attributes, as well as state of the system and prediction model.
Thus, we assume that quantiles of the cumulative distribution function (CDF) F_E of the stochastic kernel f_E, for probability p, are given by two additive components - in quantile function form - namely:

$$\boxed{F_E^{-1}(p) = F_I^{-1}(p) + F_S^{-1}(p)}$$

In the right hand side:

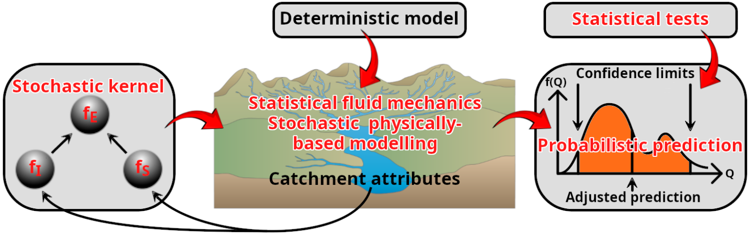
- F_I is the CDF of the prediction error of an ideally optimal model, **which is assumed to be independent of model S**
- F_S is the CDF of an additional error component which depends on model S.

Note:

- the additive property holds if and only if random variables I and S are comonotonic.
- **density f_I is assumed to be a proxy for inherent (irreducible) unpredictability** (Montanari (2025), unpublished).

Physics Informed Estimation of Stochastic Kernel

Resulting picture:



Stochastic kernel f_E: Probability distribution of deterministic model prediction error;
f_I: Probability distribution of inherent predictive uncertainty;
f_S: Probability distribution of model additional uncertainty.

Note:

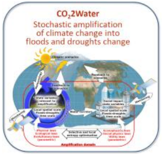
- Data driven estimation of f_I through machine learning identification of the deterministic part of the input/output transformation.
- Data Driven estimation of f_S through large sample application at the global level.
- Physics informed estimation through analysis of the results of data-driven estimation.

Looking forward to a stochastic future
Climate impact amplification

"Climate impact amplification" is defined as a step change in the impact of flood and droughts – impossible flood – triggered by a minor change in climatic drivers. It is determined by the concurrency of critical conditions in the climatic, hydrological and socio-economic systems.

FIS Advanced Grant project: $CO_2$2Water
Italian Science Fund

www.albertomontanari.it/co22water



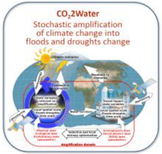
Ministero dell'Università e della Ricerca
FIS
Fondo Italiano per la Scienza

The project is in its 2 year over 5. The theory to describe climate impact amplification is being developed basing on the energy and power of floods. The key is to identify the condition under which power of floods gets increased.

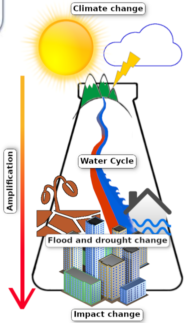
Climate impact amplification: searching for events

$CO_2$2Water is activating a search for events where climate impact got suddenly amplified and communities were taken by surprise.

- Flash and large scale floods.
- Droughts with unexpected impact.
- Debris flow events.



Ministero dell'Università e della Ricerca
FIS
Fondo Italiano per la Scienza

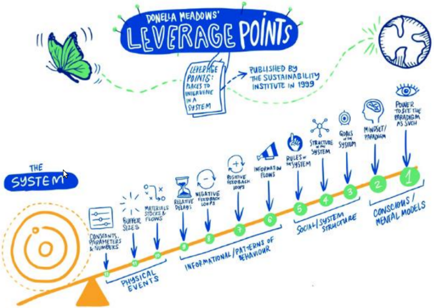


The purpose is to elaborate a large data set of surprise events at global scale to propose a checklist for identifying the climatic, hydrological and socio-economic conditions leading to amplification and surprise.
Such checklist may support the identification of areas exposed to risk.
We need data! **Interested in participating** and contribute to papers?
Stay tuned at www.albertomontanari.it/co22water!

Leverage points

The key role of leverage points to determine system's behaviour

- Pay particular attention to decipher the forcing of leverage points on system behaviours and therefore f_E.
- Leverage points are nodes in a complex system where a "small shift in one thing can produce big changes in everything" (Meadows (2008), Thinking in Systems: A Primer).



Donella Meadows' leverage points (Source: based on Meadows, 1999; credit: UNDP/Carlotta Cataldi)

- Meadows (2008) identified 12 leverage points (see figure), including, for instance, buffer size and material stock and flows.
- Leverage points may amplify a small climate signal generating an environmental disaster. They are one of the causes of "surprise".

Concluding remarks

Complexity implies lack of deterministic predictability. To manage the system a possible solution is:

- Resort to statistical predictability to maximise preparedness. Failing to take advantage frpm statistical predictability is a missing opportunity for the sake of resilience.
- A rigorous theory is to be adopted to make sure (and test) that statistical predictions are correct in a probabilistic sense. Probability is useful only if it correct and physics informed, otherwise is misleading.
- System thinking and the laws of physics should inform statistical predictability, through the stochastic physically-based estimation of the probability distribution of the prediction error.
- In complex systems, where several subsystems are interacting, and particular when the interaction is with social dynamics, leverage points are to be expected and duly interpreted to maximise (statistical) predictability.