Stochastic methods in water resources

Lecture 1: Course introduction, definition of stochastic hydrology, other definitions

Luis Alejandro Morales, Ph.D.
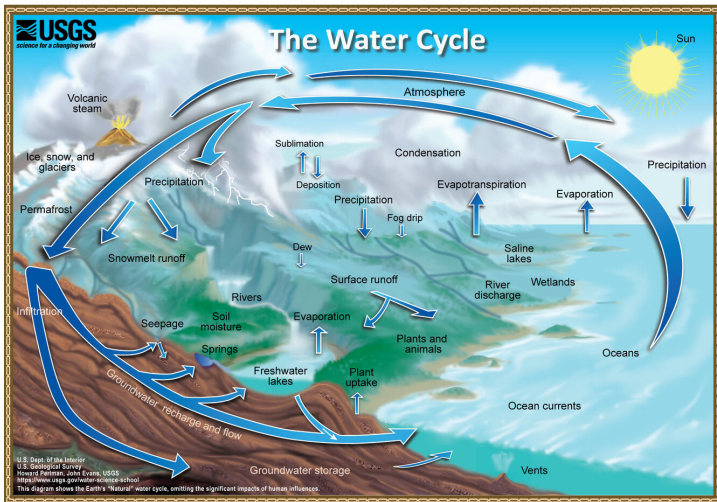
Universidad Nacional de Colombia

March 21, 2025

Some definitions

Hydrological sciences

Hydrology is an earth science concerned with the dynamics in space and time of water across different compartments (**water cycle**). Hydrology studies the source and fate of water in liquid, gaseous, and solid states on earth. It also studies water quality.

Some definitions

Hydrological environments

Hydrological environments is where hydrological processes occur and are part of the
water cycle. Here are the relationships of hydrological environments and processes:

Environment		Process
Oceans and seas	→	Evaporation and storage
Atmosphere	→	Evapotranspiration storage
Continental solid (e.g. glaciers) and liquid (e.g. river and lakes) waters	→	Storage
Land surface soils and plants	→	Infiltration, canopy interception, evapotranspiration, water retention, snow and ice accumulation, melting, interflow, overland flow and base flow
Earth's crust surface layer	→	Storage

Water quality is also affected within the hydrological environments. Water and energy
fluxes are interchanged among environments:

- ▶ Oceans and seas ⇌ atmosphere
- ▶ Land surface soils and plants ⇌ Atmosphere
- ▶ Land surface soils and plants ⇌ Continental waters

Any hydrological environment **receives**, **releases** and/or **storage** flow. This is why the
sequence of **input → storage → output** constitute the basic components of a
hydrological system that represents a **hydrological process** in space and time.
Hydrological processes in nature are **stochastic** or **deterministic** or a mix of both. A
pure deterministic process es one that occurs under controlled artificial conditions.

Some definitions

Model definition according to system's theory

A model is a simplified representation of a prototype (reality), and they can be physical or mathematical models. Mathematical models are composed by:

- **Boundary conditions**: Physical boundaries of the prototype that define how the system interact with the exterior.
- **State variables**: Describe and define the model at a certain time and space. These variables change in space and time.
- **Input variables**: Variables originated outside model's boundaries and are observed data. This includes initial condition (when $t = t_0$).
- **Parameters**: Parameters are usually model constant that change spatially only. Serve to represent model processes or properties.
- **Output variables**: Or response variables are originated from the input and the state variables throught the parameters. They change in space and time.
- **Constant**: Constants are properties that do not change in space and time. E.g. fluid properties and gravity.

Model definition according to system's theory

From the continuity equation:

$$\frac{ds}{dt} = I - O$$

where s is the sytem's storage as a function of time (t), and I and O are inputs and outputs of the system, respectively. From the figure, this equation is expressed as:

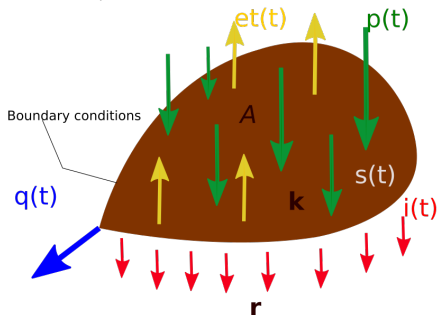$$\frac{ds}{dt} = (p(t) - et(t) - i(t))\, A - q(t)$$

where $p(t)$ is the precipitation in LT^{-1}, $et(t)$ is the evapotranspiration in LT^{-1}, $i(t)$ is the infiltration in LT^{-1}, $q(t)$ is the outflow in L^3T^{-1} and A is the catchment's area in L^2. Also in the figure, k is an storage coefficient in T^{-1} and r is the infiltration capacity in LT^{-1}. Note that $s(t)$ is the **state variable**; $p(t)$ is the **input variable**; $q(t)$, $et(t)$ and $i(t)$ are **output variables**; k and r are **parameters**; and A is a **constant**. From the equation above:
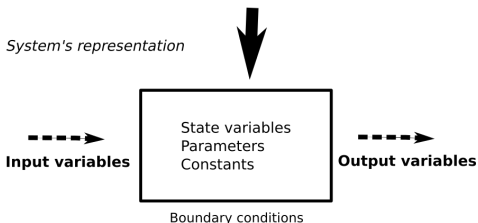
$$s^t = s^{t-1} + \left[\left(p^t - et^t - i^t \right) A - q(t) \right] \Delta t$$

where $i(t) = r$ and $q(t) = s(t)k$. So:

$$s^t = s^{t-1}(1 - k\Delta t) + \left[\left(p^t - et^t - r \right) A \right] \Delta t$$

Catchmen's representation

et(t)   p(t)

Boundary conditions

A

q(t)

k

s(t)

i(t)

r

System's representation

- - - -> **Input variables**

State variables
Parameters
Constants

- - - -> **Output variables**

Boundary conditions

Why stochastic methods in water resources?

Stochastic methods

According to [], "stochastic methods thus aim at predicting the value of some variable at non-observed times or at non-observed locations, while also stating how uncertain we are when making these predictions".
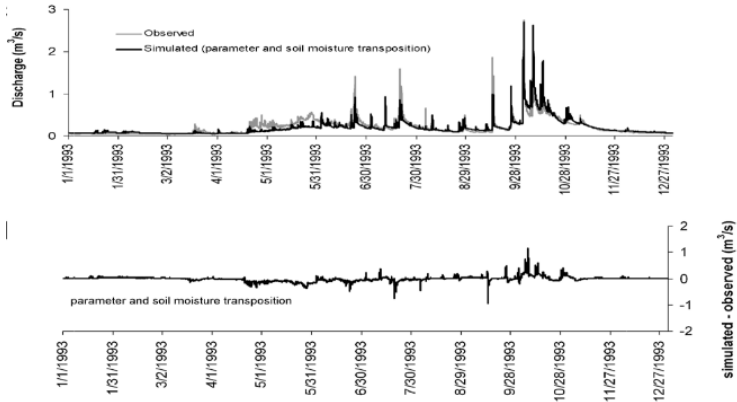

Figure: Observed vs simulated discharge (above) and residuals (below). Taken from [?]

Uncertainty is the key factor for the implementation of stochastic methods. For instance, when observed (Q) and estimated ($\hat{Q}$) discharges are compared, there are some diferences (**residuals**) between the two data sets ($e = \hat{x} - x$).

Why stochastic methods in water resources?

Errors can be summarized as:

- **Observation error**: Errors are always present in observations. Also external factor such air temperature and pressure can influence on water level records.
- **Errors in boundary conditions**: In watershed hydrological modelling, the definition of boundary conditions, most of the time, does not account for flow exchange between neigbour basins, particularly, groundwater flow exchange bringing in uncertainty in water balance estimates.
- **Errors in initial conditions**: Usually, initial conditions (initial values of state variables) are unknown at the beginning of model simulations causing uncertainty in model estimates. To correct it, models are "warm up" to pre-estimate initial conditions in order to reduce uncertainty.
- **Errors in input data**: Such errors are very common in watershed modelling. Input data to rainfall-runoff models such as precipitation or evaporation are difficult to observed for large areas so large errors are possed on data causing uncertainty in model's estimates.
- **Errors due to unknown heterogeneity and parameters**: Unknown heterogeniety in processes representation is common in computational modelling as physical processes are oversimplified in models. Such simplifications, of course, do not capture process's heterogeneity inducing to errors. Vegetation properties and soil hydraulic conductivity are very heterogenous in space and time and difficul to represent correctly. Parameters that represent physical processes or properties are higly variable in space and time, however, modelling assume unique parameters for large areas and time frames leading to errors in model estimates.
- **Errors due to scale discrepancies**: Model's domains usually assume homegenety at certain scale. For instance, in a distributed hydrological models if the size of a cell in the computational grid is 1km x 1km, it is assume that processes representation does nos change is continuous and homogeneous in the cell. This ignore the changes in the cell leading ot uncertain estimates.
- **Model errors**: *all models are wrong, but some are useful* by George Box. Models are an approximate representation of reality because they do not include all the intricate mechanism and interations that occur in natural systems. For instance, when surface flow is modelled the approximation of kinewatic wave is not always according to reality since small gradients of water level dominate the flow.

Deterministic hydrology vs Stochastic hydrology

Deterministic hydrology	Stochastic hydrology
Result from physical, chemical and biological deterministic laws, particularly from fluid mechanics (e.g. Saint-Venant equations) and thermodynamics laws. Also from deterministic analysis (e.g. rating curve)	Most of physical processes in nature are governed by laws of chance (stochastic or random). Result from a physical model where variables and parameters are treated as random variables
State variables are related to each other to represent fundamental laws through mathematical expressions	The response of physical processes are governed by laws of change where every variable is a stochastic variable.
Errors are taken into account only during calibration (e.g. adjust hydrulic conductivity)	In some cases, errors are considered during simulation as the **probability distribution** of them is known
E.g. The response (runoff) of a impervious surface to a specific rainfall event with negligible evaporation and infiltration processes.	*E.g. Runoff, precipitation, evaporation, sediment transport, etc*
E.g. A rating function on the form of $Q = f(y)$ where Q represent discharge through a fix-bed river cross section and y water level stage.	*E.g. The the annual precipitation at a point or over an area or other hydrological process at annual scale where no deterministic components are presented.*

Deterministic hydrology vs Stochastic hydrology

- Note that *stochastic* ⇌ *change* ⇌ *random* ⇌ *probabilistic* are interchangeable.
- There are no pure deterministic processes in hydrology; contrary, these are a combination of conservation and change laws.
- Hydrological processes can be a combination of both: e.g. periodic behaviour (deterministic trigonometric function) + random variation → **periodic-stochastic process**
- Note that a deterministic relationship function between random variables is also a random variables. E.g. The response of a river basin (runoff) to rainfall events is better understood by analysing not only the deterministic relationships but also the stochastic properties of observed rainfall, runoff, climate and basin data.

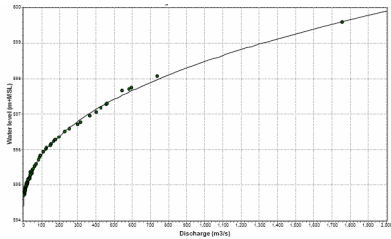# Deterministic hydrology vs Stochastic hydrology
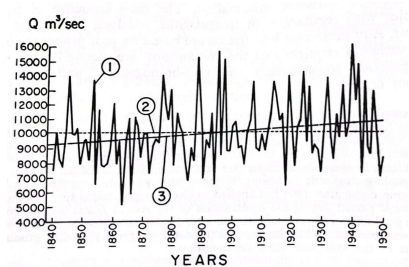

Figure: Typical rating curve.

Figure: Annual maximum discharges of the Danube River at Orshova Romania. 1) time series of maximum, 2) time series mean and 3) time series trend.