HUMID

Regionalization approach to link physical characteristics with reservoir model parameters using a genetic algorithm.

Omar Cenobio-Cruz, Pere Quintana-Seguí
Observatori de l'Ebre, (URL - CSIC).

ACCWA AI Workshop, 03/12/2021

1. SASER model

Objective: Improving low flows in modeling chain SASER

SASER = **SA**FRAN + **S**URFEX + **E**audyssée + **R**APID.

Meteorological forcing (**SA**FRAN)

- SAFRAN analyses daily observed precipitation, which is interpolated to the hourly scale.
- SAFRAN Iberia, 5 km
- SAFRAN Pyrenees, 2.5 km (Adour, Garonne, Ebro, etc).

Land Surface Model (**S**URFEX) and Routing Scheme (**E**audysse +**R**APID).

Currently, SURFEX uses ISBA-DIF.

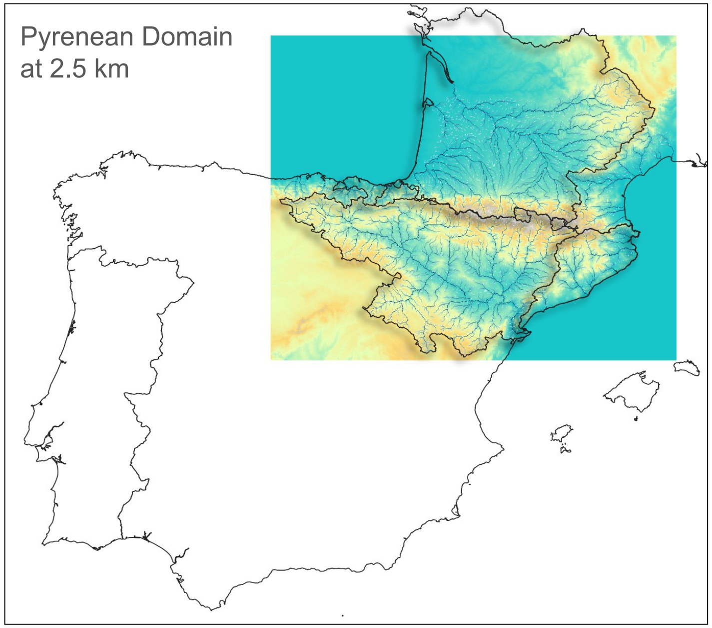
- It describes the vertical processes in the soil column and the vegetation.
- It generates the outflows that will allow us to simulate the river-flow.
- SURFEX does not simulate river flow.
-

We need to transport SURFEX's runoff and drainage to the river and then compute the river flow.

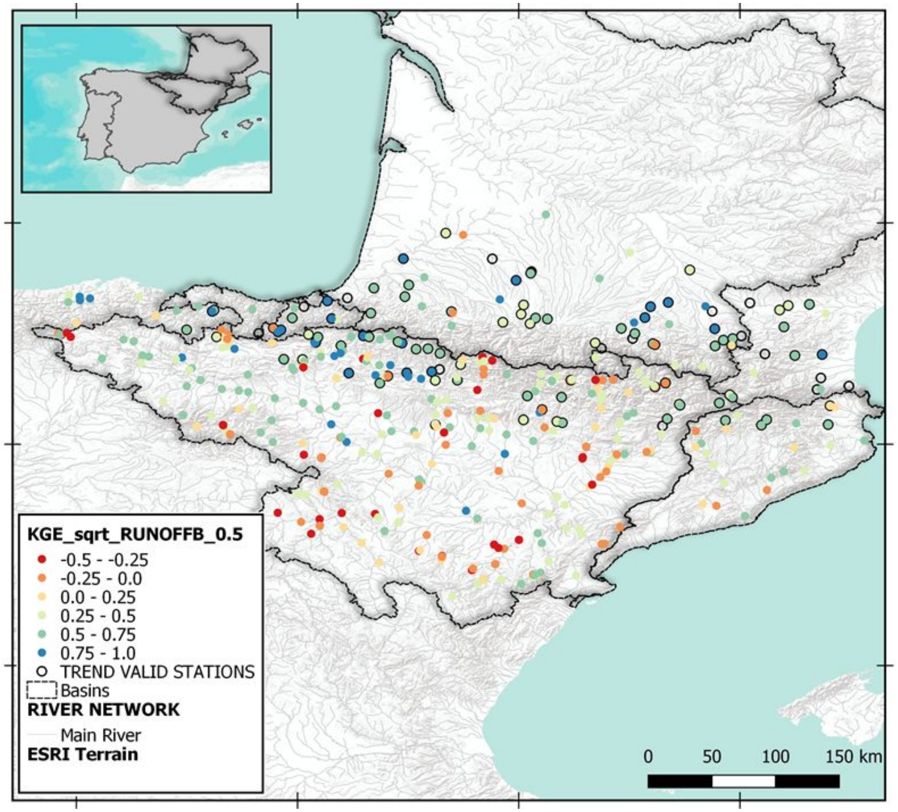
- Eau-dyssée performs this task.

SASER doesn't take account:

- Underground water and underground interactions.



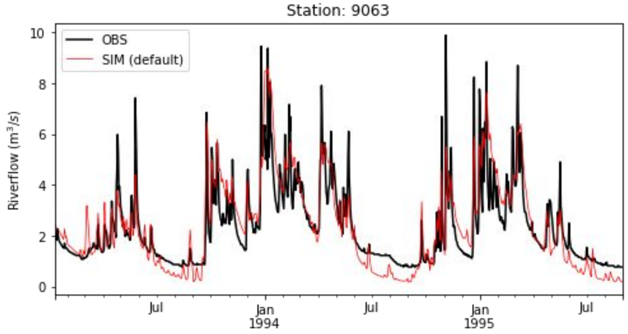
2. Model Performance

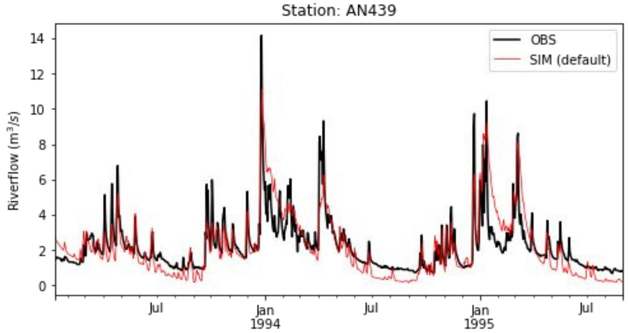


To evaluate the model performance we use KGE(sqrt(Q)).

Flow transformation	Increased low flow weight	No issue with zero flows	Dimensionless	No issue when flows average around 1
Square root	+	+	+	+
Inverse	+++	−	+	+
Reciprocal of root	++(+)	−	+	+
Logarithm	++	−	−	−
Box–Cox	+(+)	+	+ (if using Eq. 10)	+ (if using Eq. 10)

Pros and cons of different KGE functions (*L. Santos, et al., 2018*)





The model works remarkably well.

…but due to the lack of an underground water model, SASER tends to **underestimate low flows**

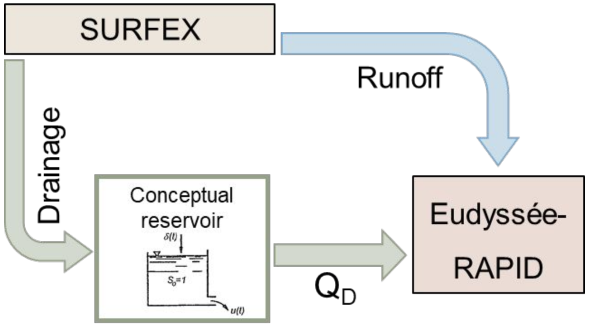
3. Improving low-flows

Reservoir implementation.

1. Common in hydrological models.

Already tried with SURFEX/ISBA:

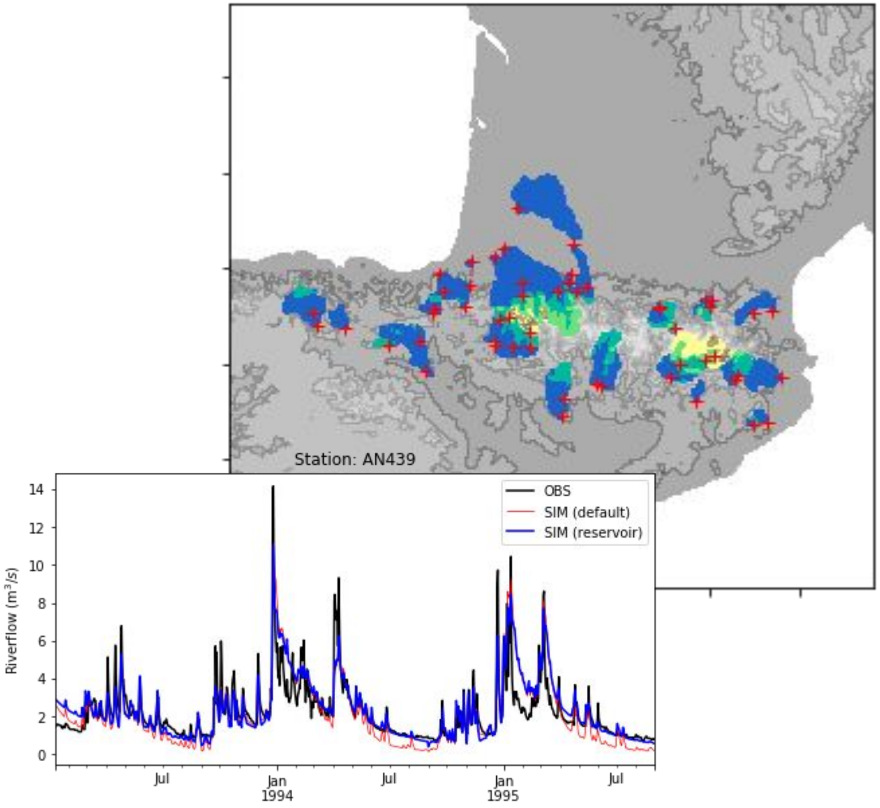
- (Artinyan et al., 2008) uses this approach in Bulgaria.
- (Getirana et al., 2014, 2017) also tried this in Africa.



2. Very simple to implement (few lines in python).
3. Not physical, and needs a calibration/validation process to compensate for other error sources.

We introduce a conceptual reservoir at grid-scale to modify only the drainage from SURFEX

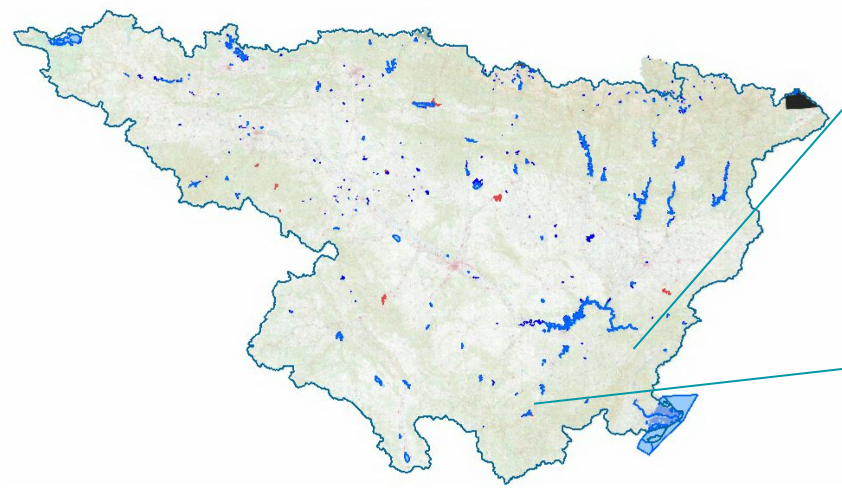
We select near-natural sub-basins (headwaters) and implemented the linear reservoir



3.1 Limitations to implement linear reservoirs

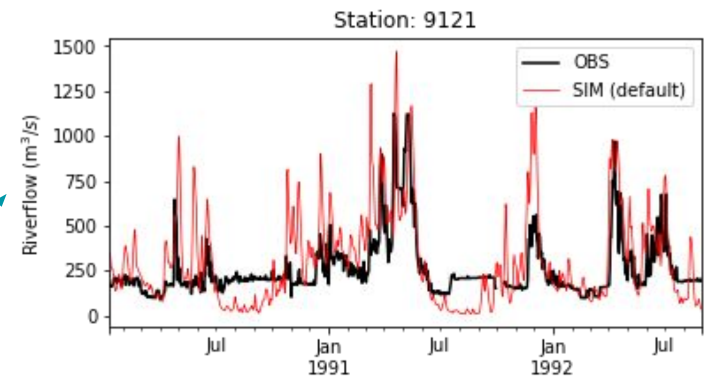
Ebro basin is data-rich, but also there are large hydraulic infrastructures and irrigation areas.

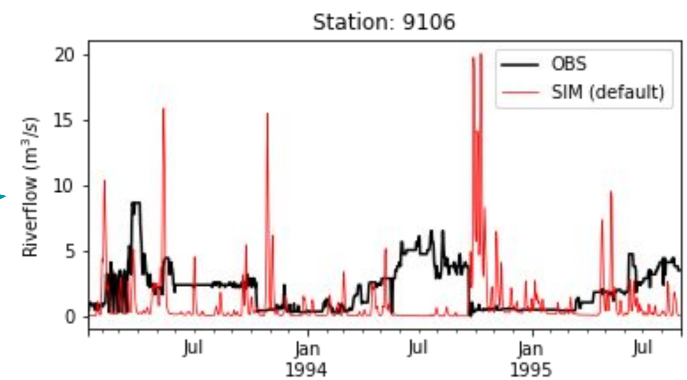
- Dams: water storage during the wet season, to be used during the dry season.
- Canals: Transport of water from the runoff generating areas to the irrigation areas.



Dams in the Ebro basin (source: CHE)

Water management and anthropogenic processes modify the hydrology of river basins.





In this way, we can **only calibrate in natural/near-natural** sub-basins (limited areas over most of the basins)

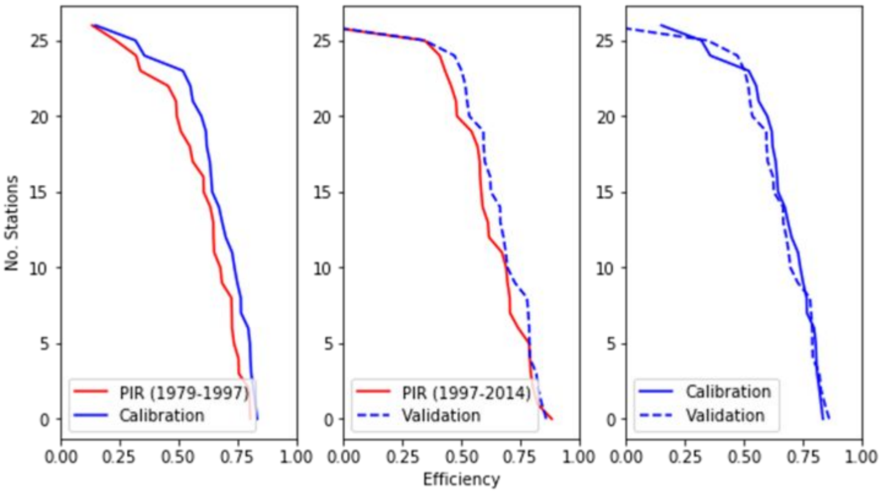
3.2 Reservoir implementation

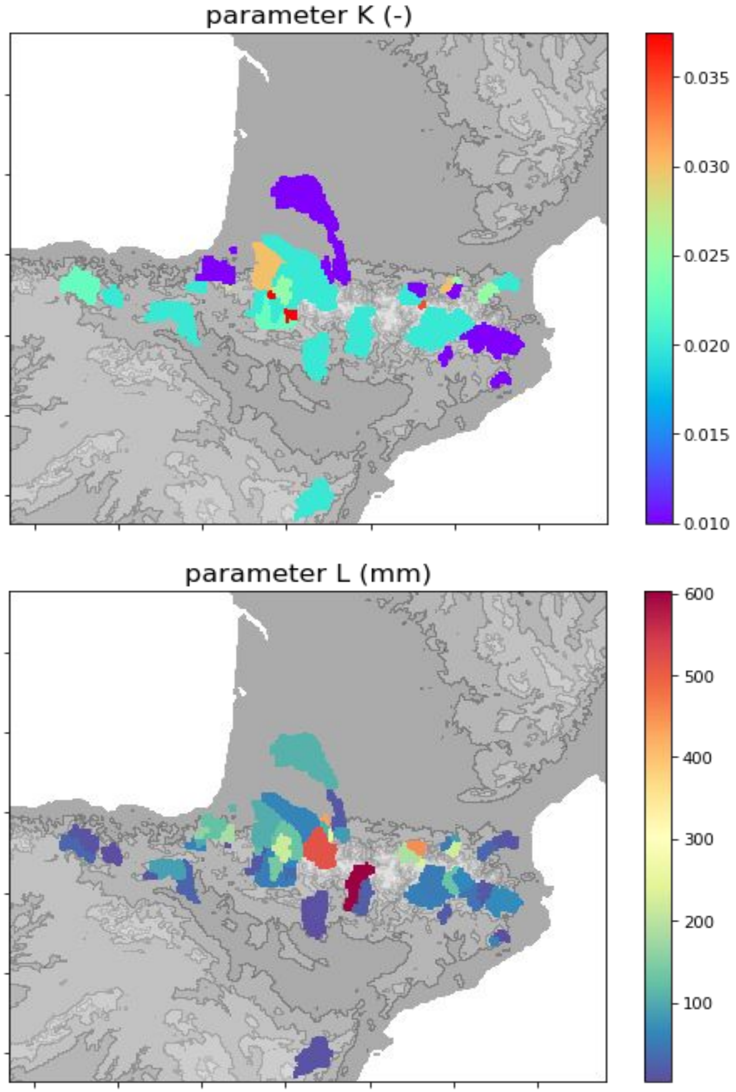
We calibrate reservoir parameters catchment by catchment (nested catchments).

Two parameters are calibrated:

- **L** [mm]: Threshold parameter for extra outflow from storage.
- **k** [-]: Recession Coefficient

KGE shows an improvement in both periods (calibration and validation)



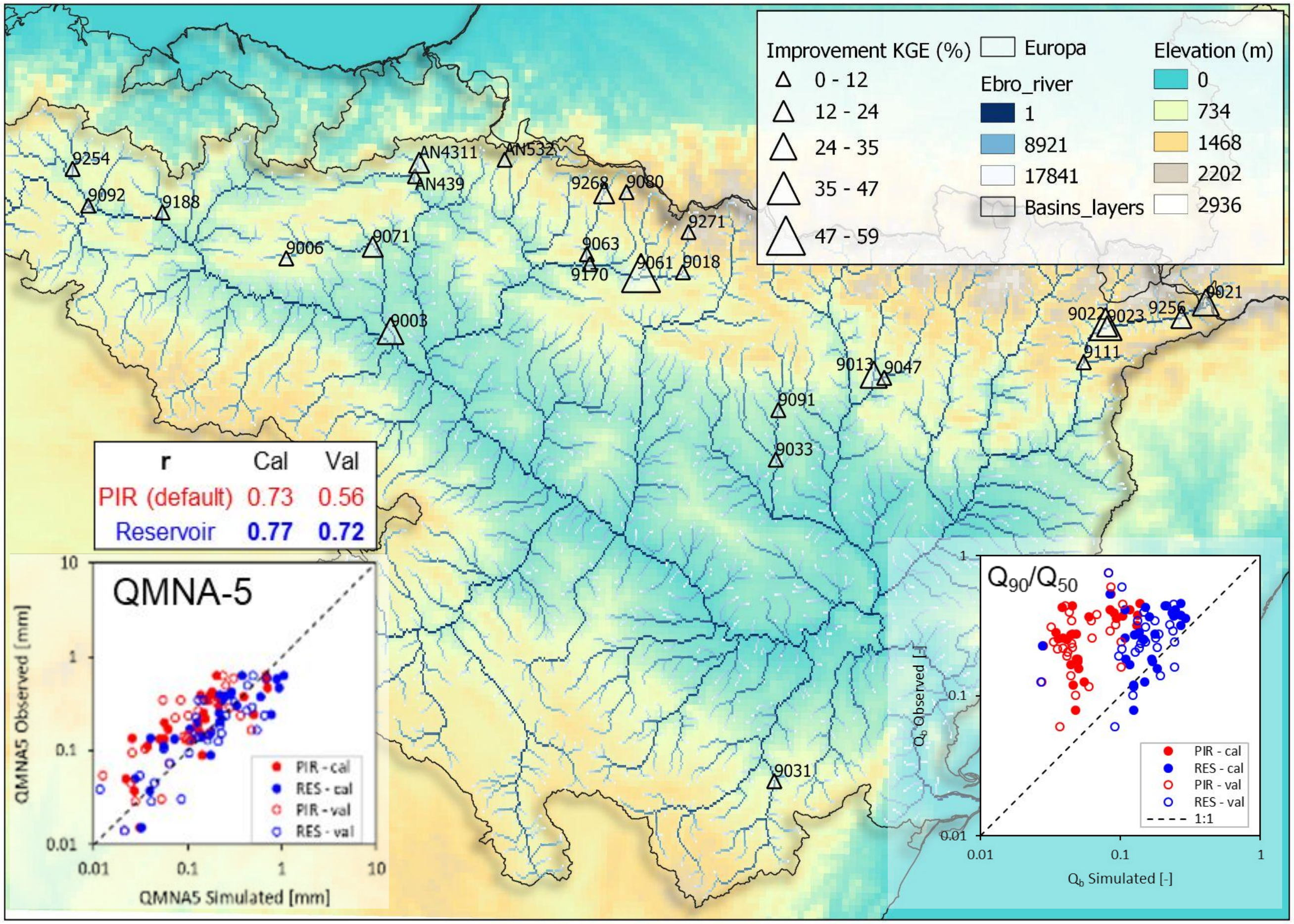


3.3 Low flow indexes

Low flow indices also indicate an improvement.

QMNA-5 provides information about low flow severity and allows statistically evaluating the lowest flow of a river flow during a given period.

Q_b (Q_{90}/Q_{50}) makes it possible to assess the dynamics of the underground flow in the basins, specifically if this dynamic plays a supporting role during periods of low flow.



4. Regionalization approach

The core of parameter regionalization is to “lend” the effective hydrological information from gauged catchments to the ungauged catchments. (Guo et al., 2020).

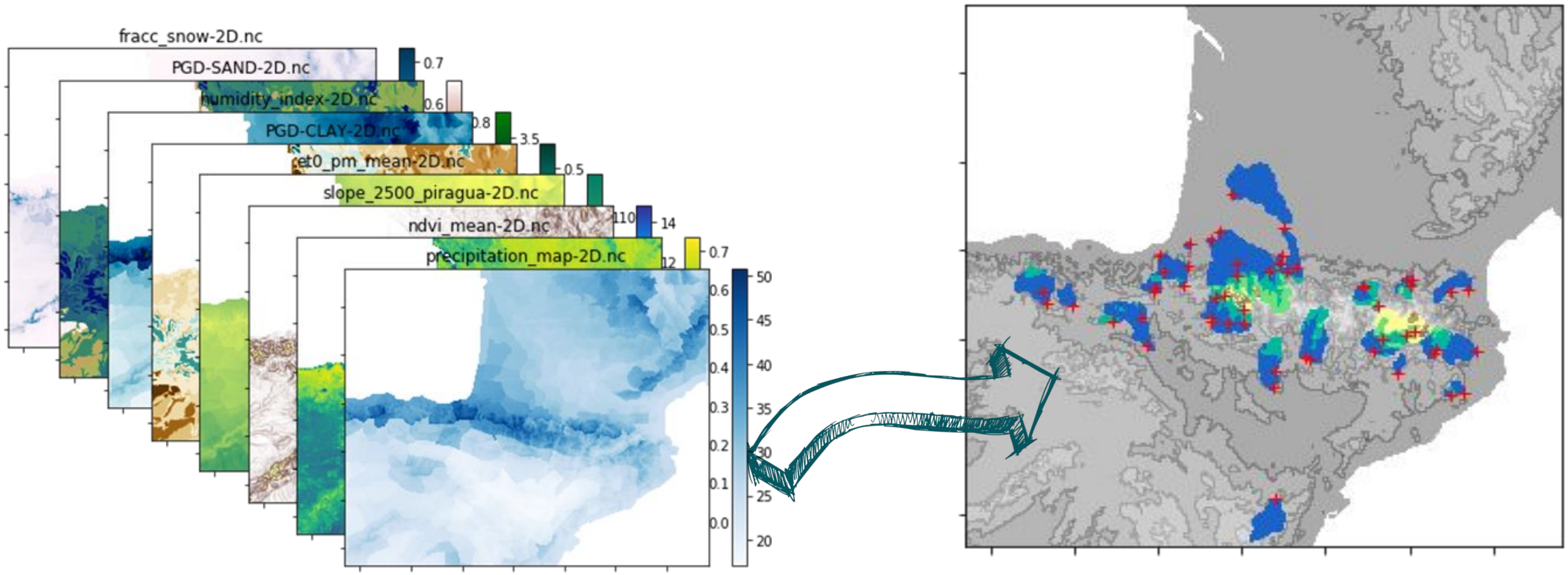
We use a regionalization approach to **link physical characteristics with reservoir parameters** through transfer equations, as it did (Beck et al., 2020)

The main goal: produce **parameter maps for our entire domain** including influenced sub-basins

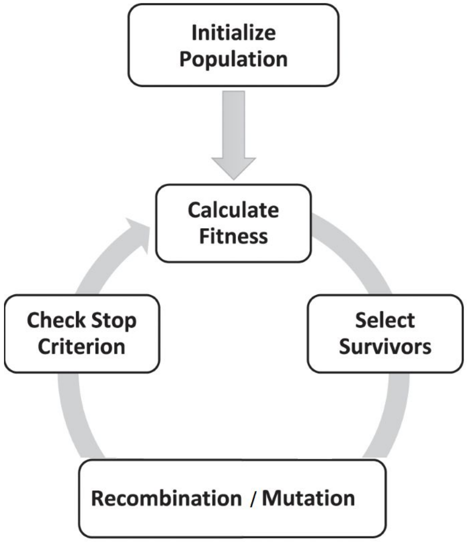
Physical information (predictors) related to climate, land cover, topography and soil.

$$parameter_i = \mathbf{w}_{i1}\, predictor_1 + \mathbf{w}_{i2}\, predictor_2 + \cdots + \mathbf{w}_{iN}\, predictor_N + \mathbf{w}_{iN+1}$$

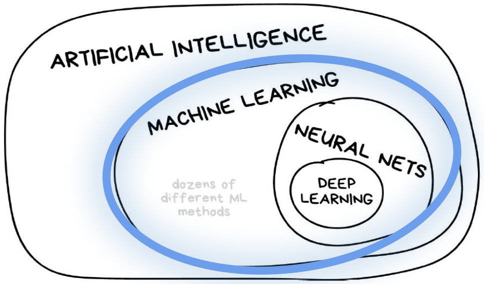
We are using a **genetic algorithm (GA)** to optimize the coefficients of the transfer equations (w)

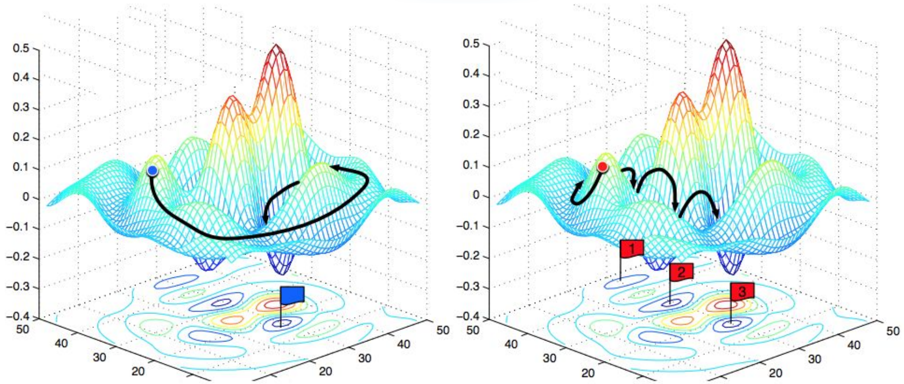


4.1 Genetic algorithm



- ➔ Genetic algorithms are based on the ideas of natural selection. *"Survival of the fittest"*.
- ➔ They are commonly used to generate high-quality solutions for optimization problems and search problems.





Advantages of genetic algorithm

The concept is easy to understand and implement.

A genetic algorithm does not need derivative information (only evaluates fitness score, obtained from objective function).

Drawback of genetic algorithm

Genetic algorithms are costly in computational terms.

These kind of algorithms can take a long time to converge because there are multiple local optima.

4.2 Genetic algorithm implementation

➔ We selected catchments up to 5,000 km^2 (to avoid routing effects in larger catchments)

We use the following predictors:

- Humidity index (P/PET)
- Mean annual precipitation (square root)
- Mean annual potential evaporation
- Fraction of snow
- NDVI
- Slope
- Sand
- Clay

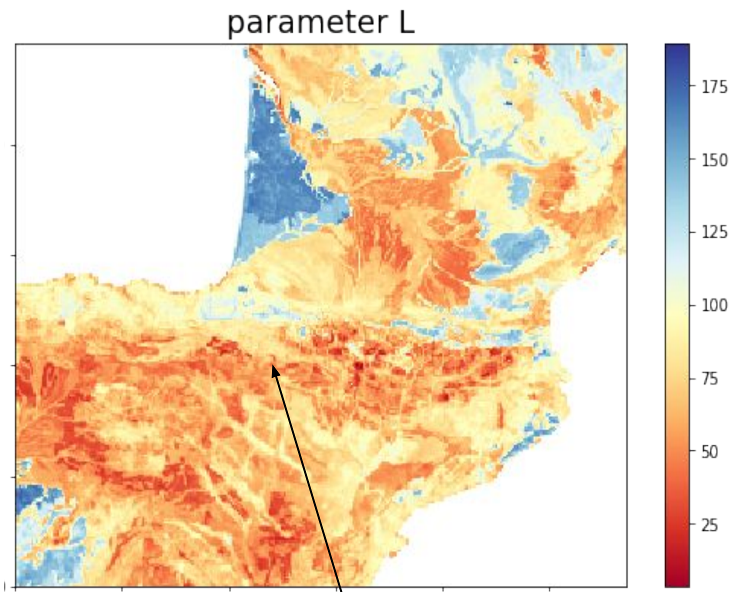
➔ Our fitness function is KGE from daily time series of observed streamflow and simulated streamflow (obtained by aggregating the runoff and drainage)

➔ Predictors were standardized by subtracting the mean and dividing by the standard deviation of the area covered by the catchments, to make all predictors comparable.

➔ **Cross Validation**. We divided the catchment set into subsets for calibration (90%) and for validation(10%)

4.3 Genetic algorithm results (1)

Our first guess was to run the algorithm to optimize only one of the reservoir parameters (L) so we set a fixed value of k parameter (0.02).



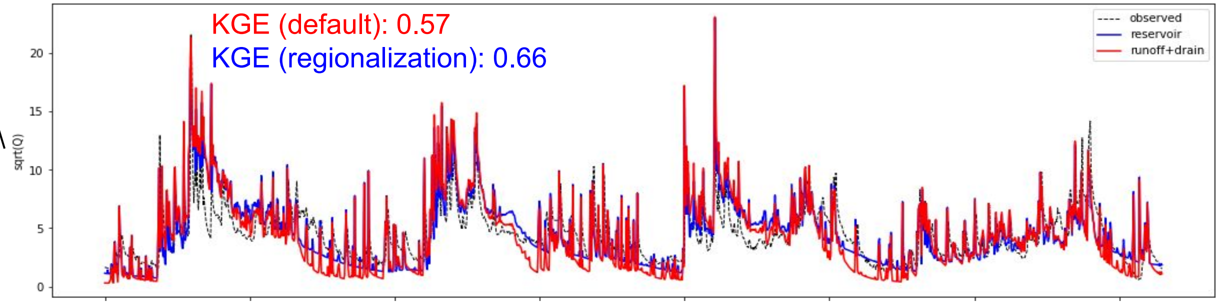
Using only Ebro catchments (median KGE):

default	regionalization	Local calibration
0.427	0.53	0.525

The map produced by the optimized parameters shows:

- higher values of L parameter in zones where sand predominate.
- by contrast, clay regions show lower values of L parameter.

Station: 9170



Streamflow comparison for a catchment
(only show a couple of years)

4.3 Genetic algorithm results (2)

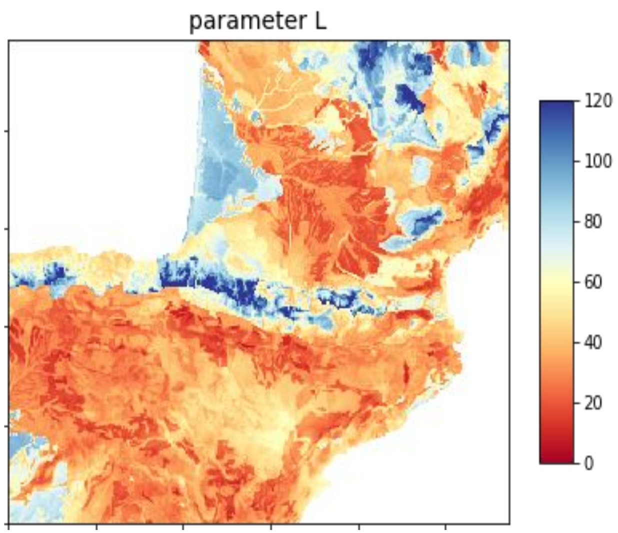
After, we run the algorithm to optimize both parameters.

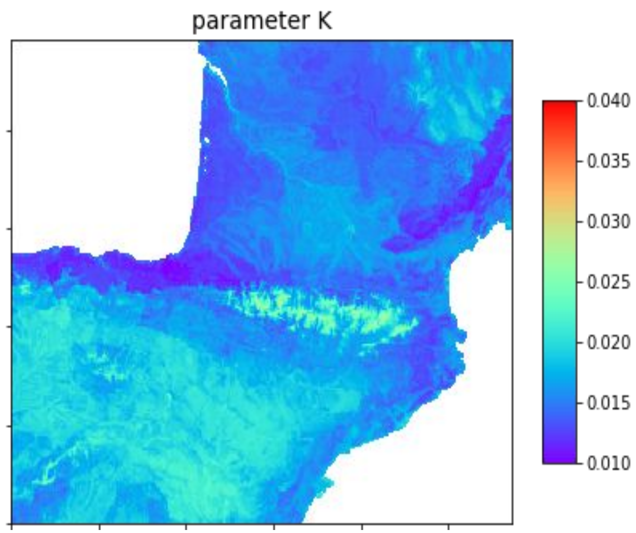
To define an upper limit to regionalization approach we use the local calibration scores.

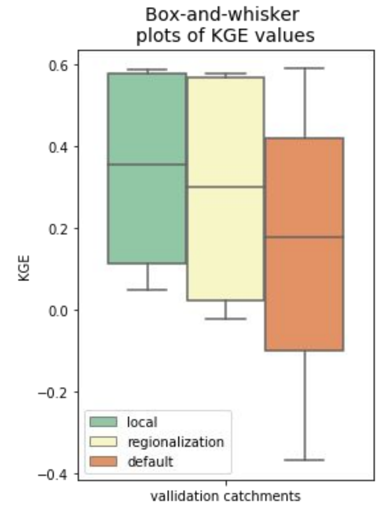
The difference in median KGE between default (0.39) and regionalized parameters for the training catchments (0.53) was thus 0.14.

The box plot of KGE values of validation catchment to local calibration (validation period), regionalization (validation period), and default simulation (without reservoir).

75% of the validation catchments show an improvement







Conclusions

- Performance of default simulation is near to the model capabilities. The model represent quite well median and peak flows
- The linear reservoir implementation noticeable improves low flows.
- We obtained maps at 2.5 km resolution (covering our entire domain) of each reservoir parameter, making it we can simulate “naturalized streamflow” in influenced catchments.
- Maps produced by optimization vary according to physical variables at each gridpoint over full domain, in contrast to local calibration (catchment by catchment)
- In parameter maps we can identify clear spatial patterns related to hydrological processes.

Work in progress

- We are running more simulations with different calibration subsets to cross-validation step.

HUMID

Thank you!

ocenobio@obsebre.es

.