EUROPEAN CENTRE FOR RESEARCH AND ADVANCED TRAINING IN SCIENTIFIC COMPUTING

Reduced model for sensitivity analysis and ensemble based data assimilation with 1D and 2D hydraulic modeling

Isabelle Mirouze (CECI), Sophie Ricci (CECI),
Matthias De Lozzo (CERFACS), Nicole Goutal (LNHE, LHSV),
Siham El Garroussi (CERFACS), Simon Saysset (CERFACS),
François Moussu (CERFACS)

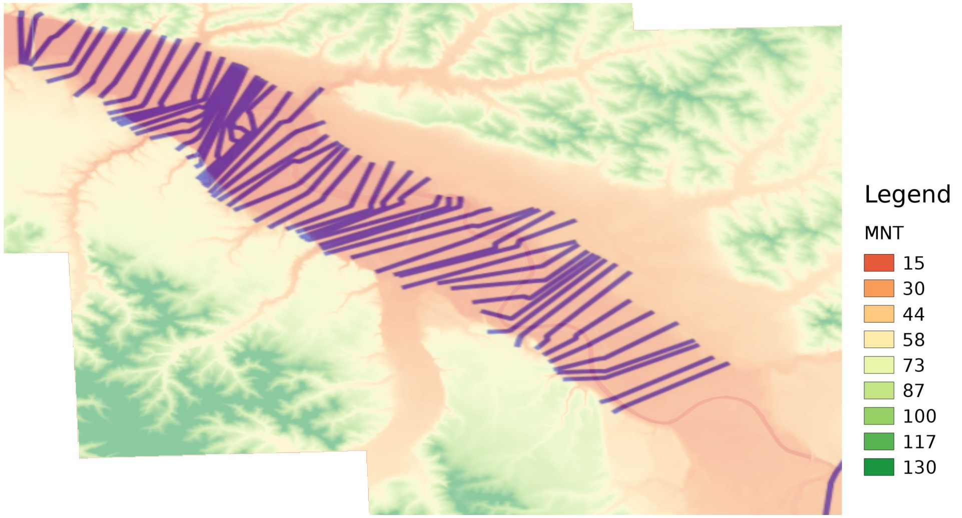
Outline

Objective: to improve the estimate of the discharge in 1D and 2D models by assimilating SWOT data

Part I – To generate inputs for the simulator	Part II – To produce SWOT data	Part III – To use SWOT data
Step 1: *Hydraulic simulations with SIC², Mascaret, Rubar, Telemac 2D* Step 2: To spatialise / interpolate water levels from the hydraulic 1D / 2D model into lat-lon rasters	Step 1: *To select tracks* *To produce the pass plan* Step 2: *To simulate the interferometric measurement* Step 3: *To calculate water levels*	Step 1 : *To average water levels (pcg)* Step 2: *Comparison SWOT_HR simulator vs Model data* Step 3: *Data Assimilation*

Part I: To generate inputs for the simulator

Requires:
- DEM of the area of interest
- Profiles of the bathymetry
- Outputs from the hydraulic model



<u>1D Model</u>

To interpolate 1D outputs on a 2D domain
- Requires georeferenced lines to project the 1D outputs on
  - ✓ *Method developed at CERFACS using Grass/Qgis scripts*

<u>2D Model</u>

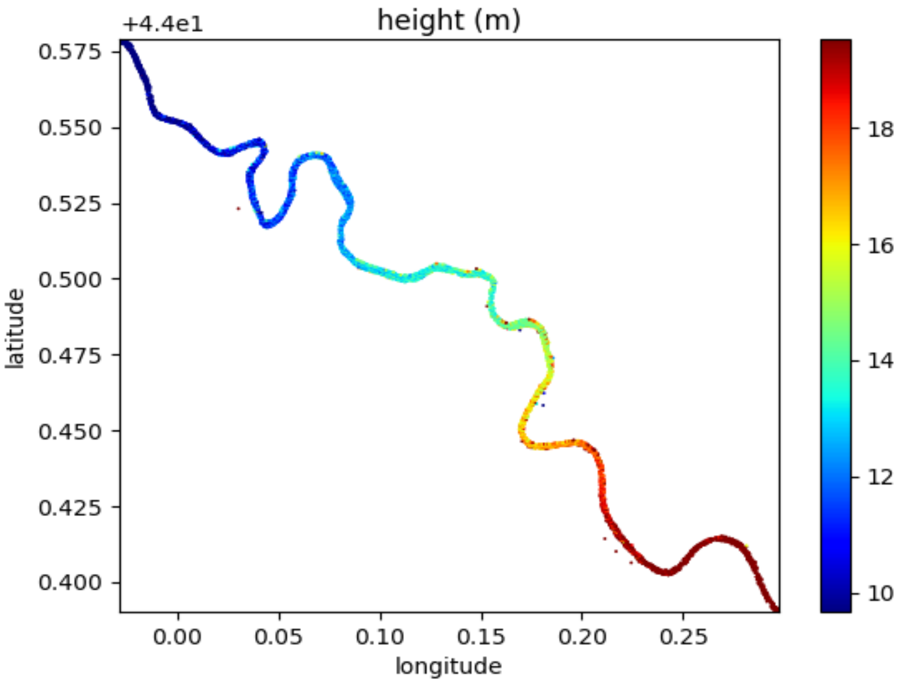
To interpolate on SWOT-HR input grid
- Transform an unstructured grid into a structured grid (Telemac)

✓ *Method developed at CERFACS using Open-Palm/CWIPI coupler*

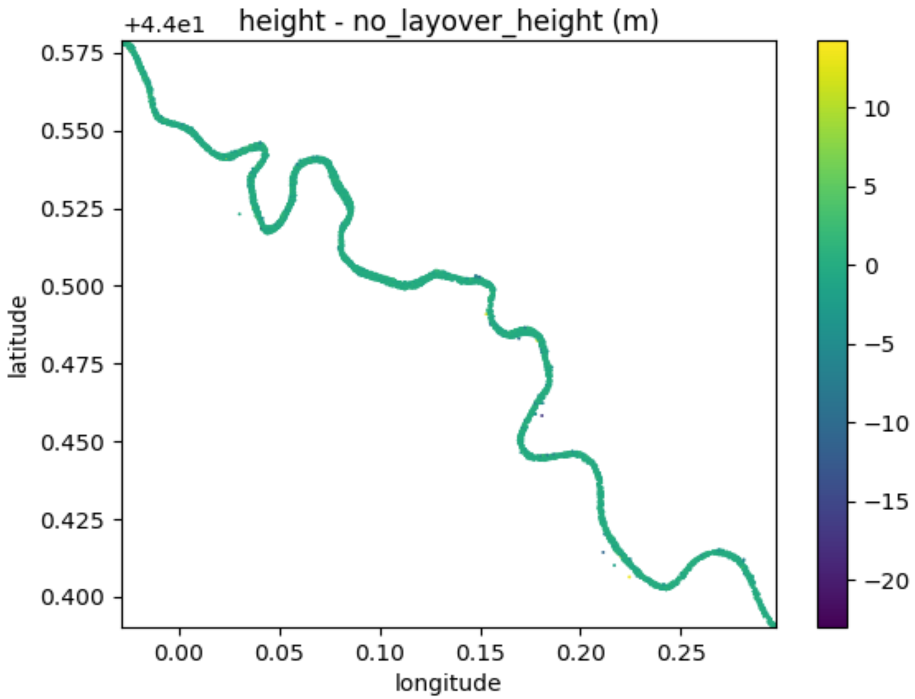
Generate a water mask

Part III – To use SWOT data

Free surface elevation map



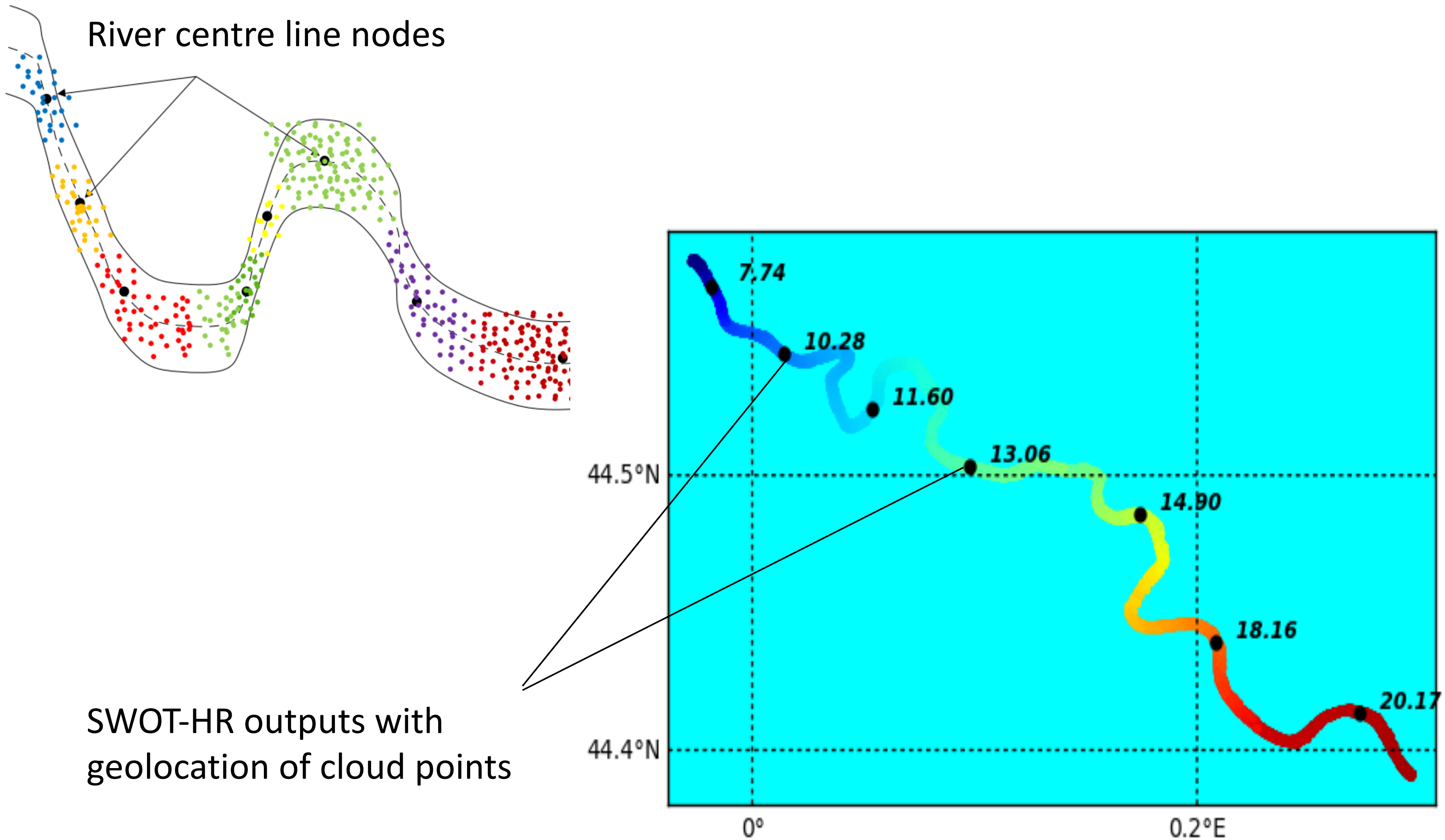
Associated errors



- Comparison between SWOT-HR and hydraulic model outputs : requires consistency between spatial scales
- Reduction of the pixel errors from metres to centimetres

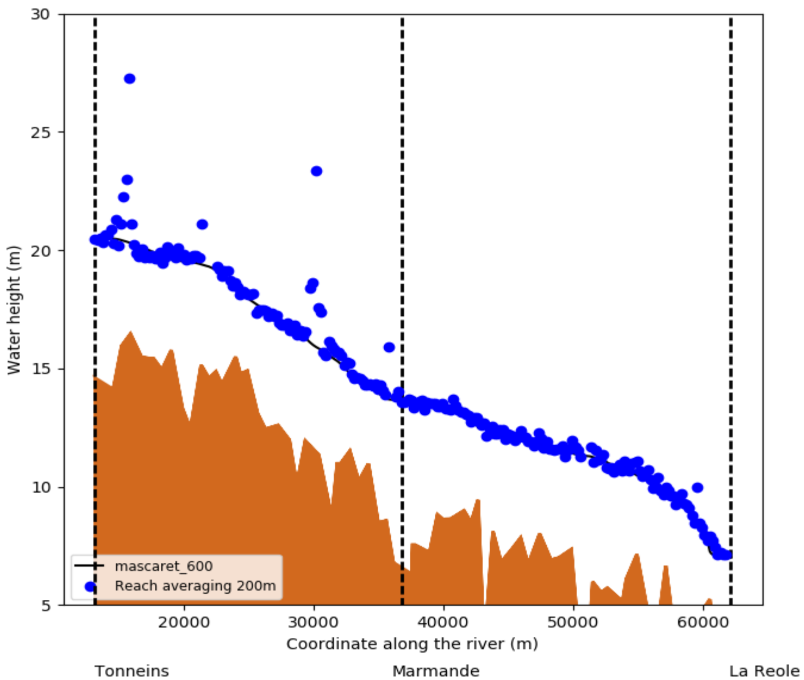
Averaging the elevations by sections

Consistency between the spatial scale



Comparison with the 1.5D model Mascaret

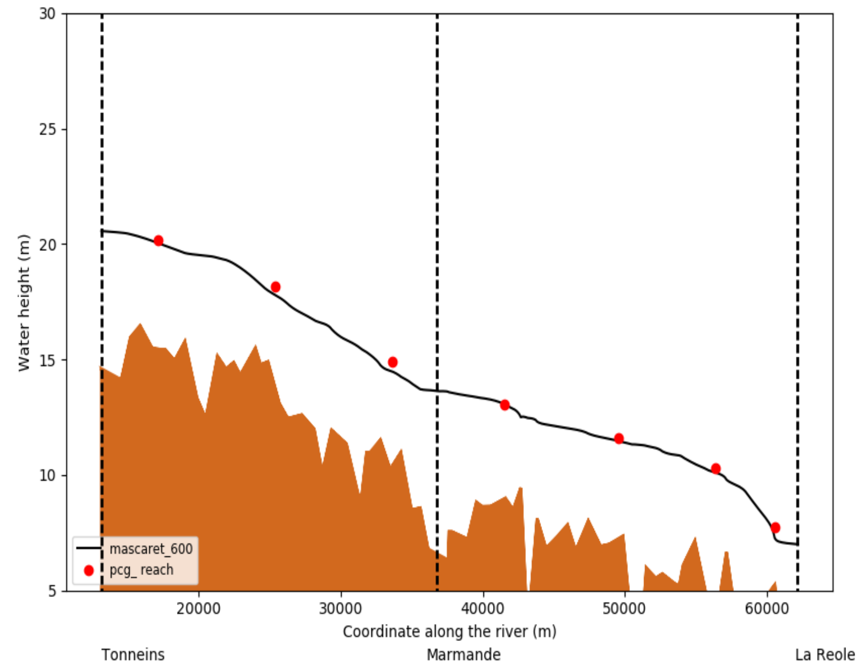
Water level average every 200 m



Mascaret

Averaged data

Water level average by 5 km sections



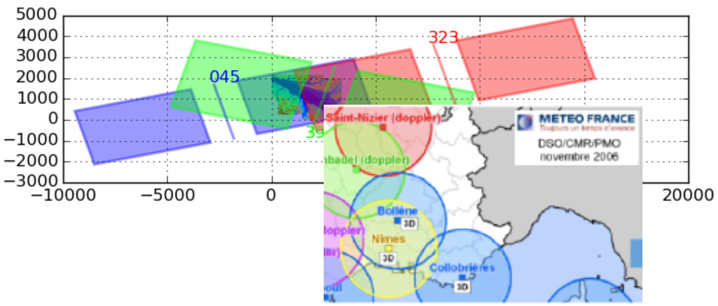
Mascaret

Averaged data

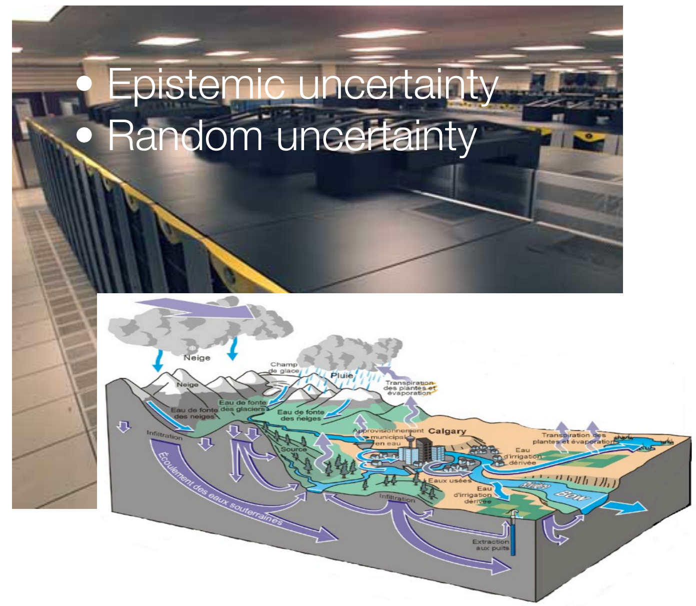
To identify and to quantify the major sources of uncertainty in hydrology solver

Observations

- Sparse and imperfect
- Relation between observations and model outputs



Model

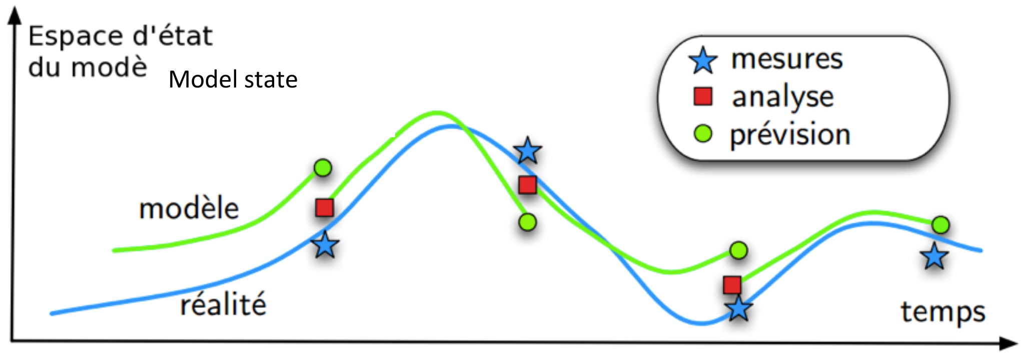


- Epistemic uncertainty
- Random uncertainty

Data assimilation: optimal combination of observations and model a priori to determine the best estimate of a dynamical system

Uncertainty reduction with Data Assimilation

To reduce the major sources of uncertainty in hydrology solver



- Variational methods

→ **Minimisation of a cost function**

$$J(\blacksquare) = \frac{1}{2}\| \blacksquare - \bullet \|_B^2 + \frac{1}{2}\| \bigstar - \mathcal{G}(\blacksquare)\|_R^2$$

- Sequential methods

→ **Kalman Filter analysis**

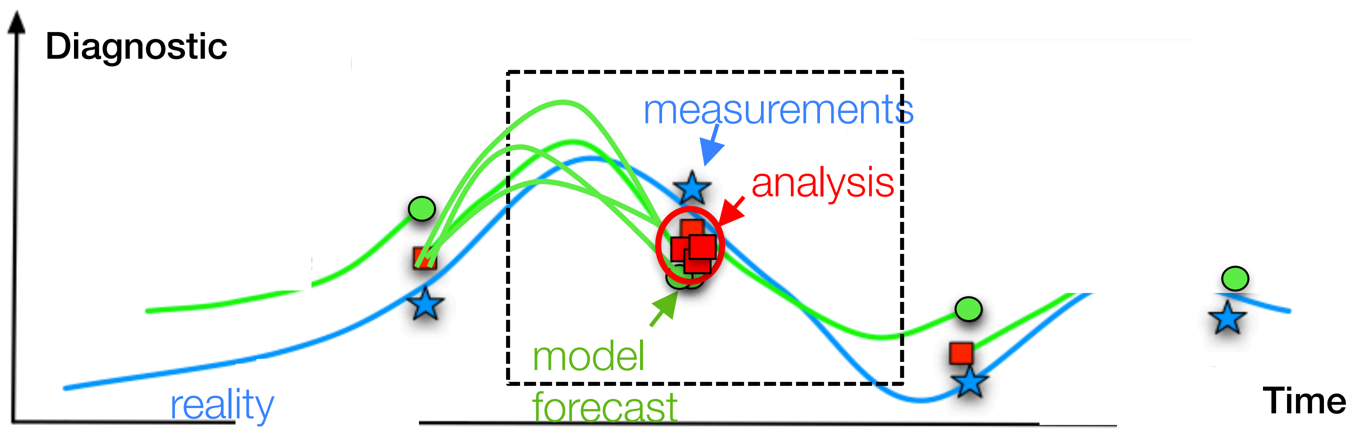
$$\mathbf{K} = \mathbf{C}_{xy}(\mathbf{C}_{yy} + \mathbf{R})^{-1}$$

$$\blacksquare = \bullet + \mathbf{K}\left[\bigstar - \mathcal{G}(\bullet)\right]$$

Weight (K) — Analysis (■) — Forecast (●) — Distance to observations ([★ - $\mathcal{G}$(●)])

Uncertainty reduction with Data Assimilation

Stochastic estimate of the covariance matrix: Ensemble Kalman Filter (EnKF)

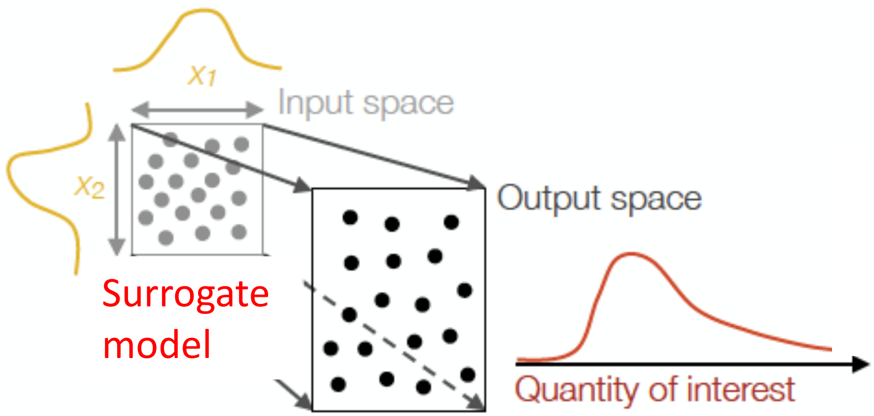


$$\mathbf{C}_{xy} = \mathbf{P}_t^{\mathrm{f}} \mathbf{G}_t^{\mathrm{T}} = \sum_{k=1}^{N_{\mathrm{e}}} \frac{\left(\boldsymbol{x}_t^{\mathrm{f},(k)} - \overline{\boldsymbol{x}_t^{\mathrm{f}}}\right)\left(\mathcal{G}_t(\boldsymbol{x}_t^{\mathrm{f},(k)}) - \overline{\mathcal{G}_t(\boldsymbol{x}_t^{\mathrm{f}})}\right)^{\mathrm{T}}}{N_{\mathrm{e}} - 1}$$

Classical way:
Monte Carlo sampling
with N_e members

Implementation: EnKF requires numerous model integrations

Stochastic estimate of the covariance matrix with a surrogate model



Polynomial Chaos

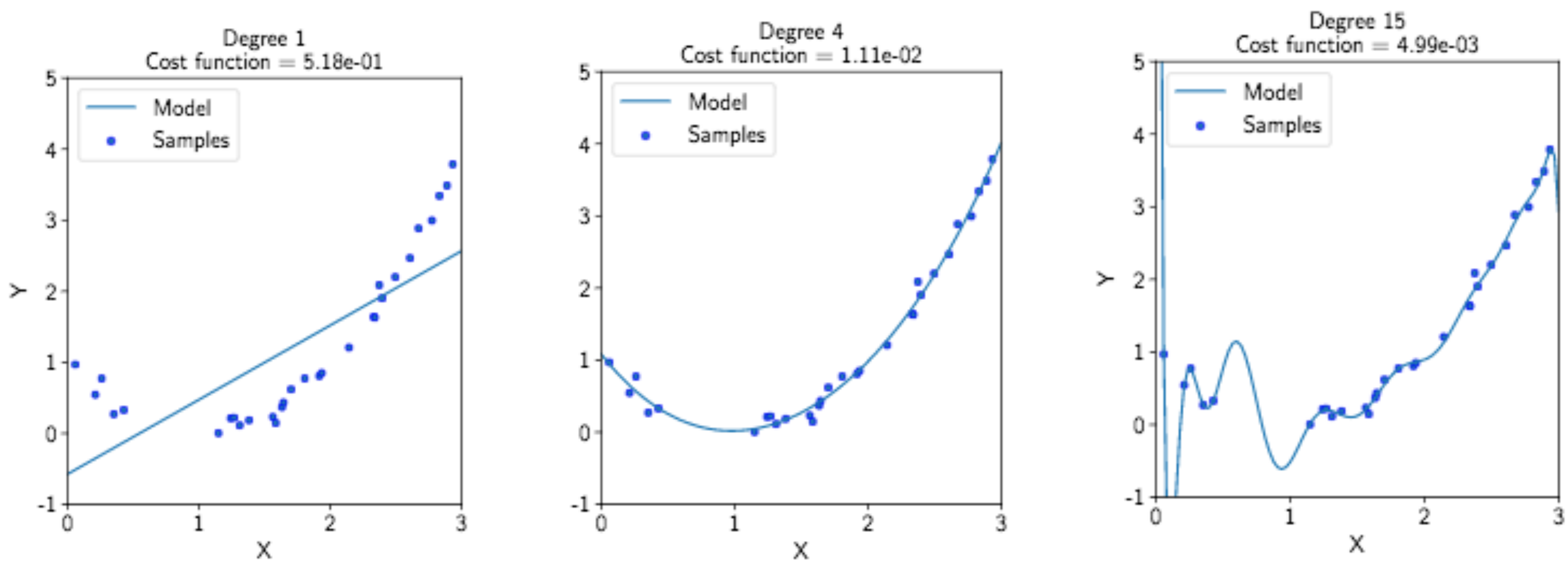
$$\tilde{f}(X) = \sum_{k=0}^{m} w_k.X^k = w_0 + w_1.X + \ldots + w_m.X^m$$

$$\mathbf{C}_{xy} = \mathbf{P}_t^{\mathrm{f}}\mathbf{G}_t^{\mathrm{T}} = \sum_{k=1}^{N_{\mathrm{e}}} \frac{\left(\boldsymbol{x}_t^{\mathrm{f},(k)} - \overline{\boldsymbol{x}_t^{\mathrm{f}}}\right)\left(\mathcal{G}_t(\boldsymbol{x}_t^{\mathrm{f},(k)}) - \overline{\mathcal{G}_t(\boldsymbol{x}_t^{\mathrm{f}})}\right)^{\mathrm{T}}}{N_{\mathrm{e}} - 1}$$

New way:
Stochastic estimation with the surrogate model

Implementation: using a surrogate model reduces the cost

Constructing a surrogate model with polynomial chaos

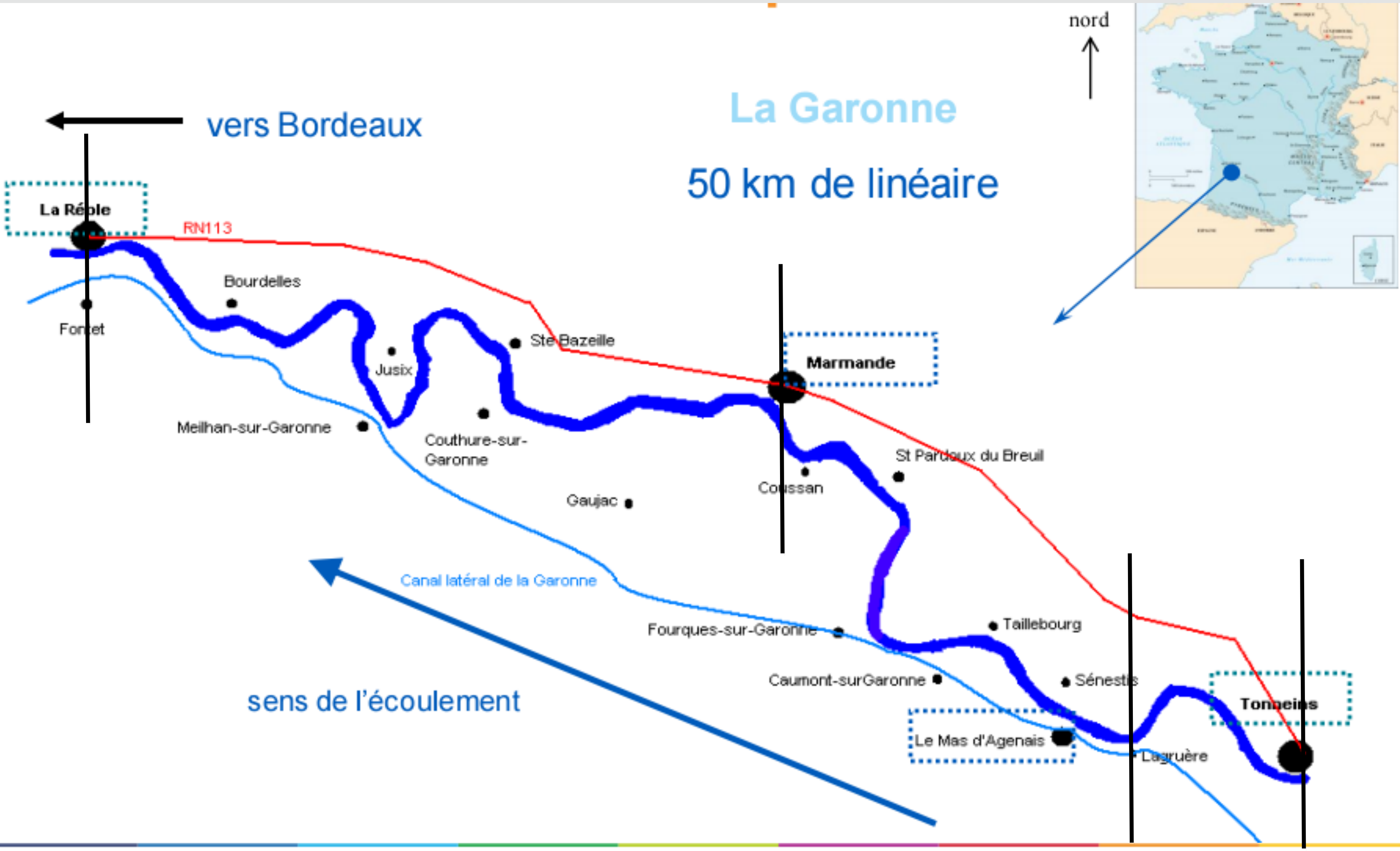


Polynomial Chaos

$$\tilde{f}(X) = \sum_{k=0}^{m} w_k.X^k = w_0 + w_1.X + \ldots + w_m.X^m$$

Implementation: the surrogate model should be accurate but cheap to construct

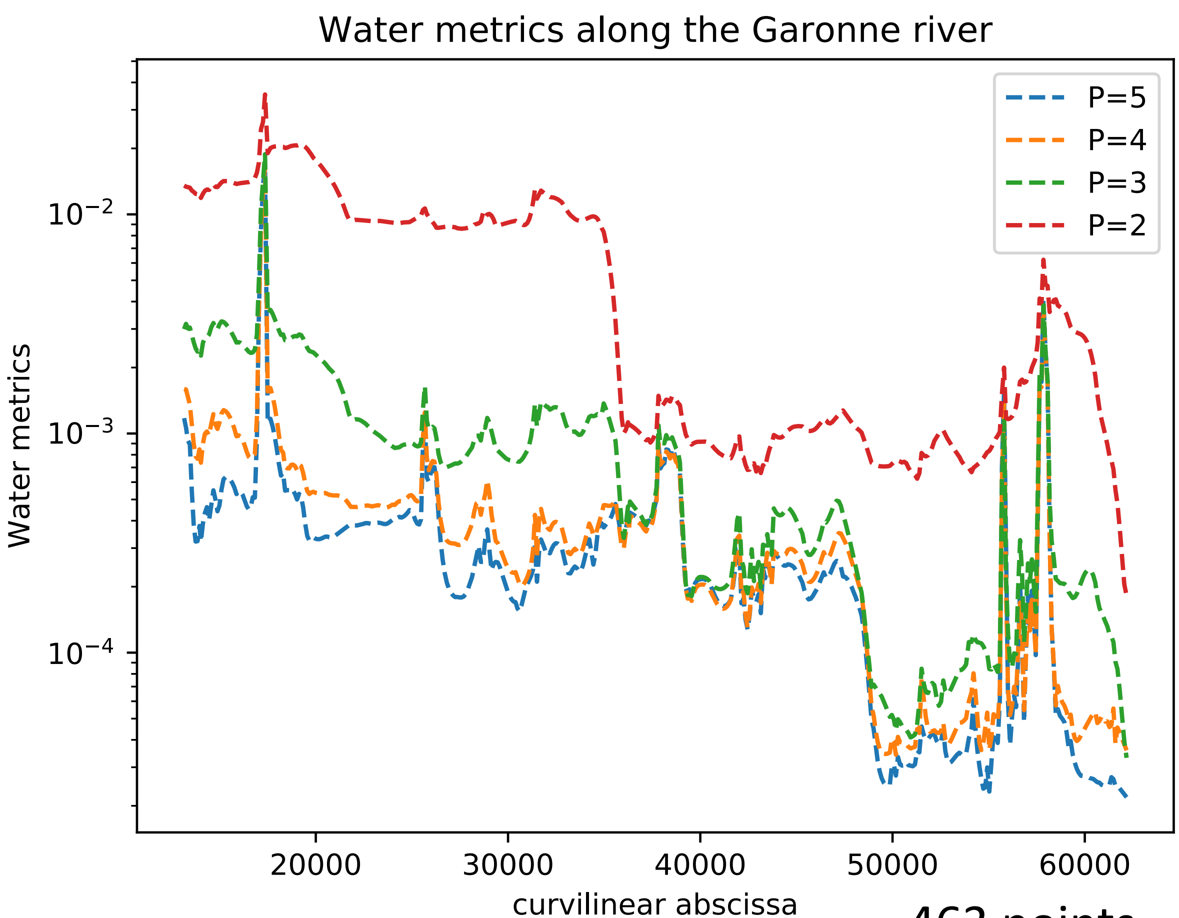
Surrogate model on the Garonne



Surrogate model - Garonne (1.5D)

Choosing the appropriate order of the polynomial

Inputs: Ks1, Ks2, Ks3, Q



463 points

$$e = \frac{1}{M}\sum_{s=1}^{M}(\tilde{h}_s - h_s)^2$$

$\tilde{h}$:

Various P

N = 1000

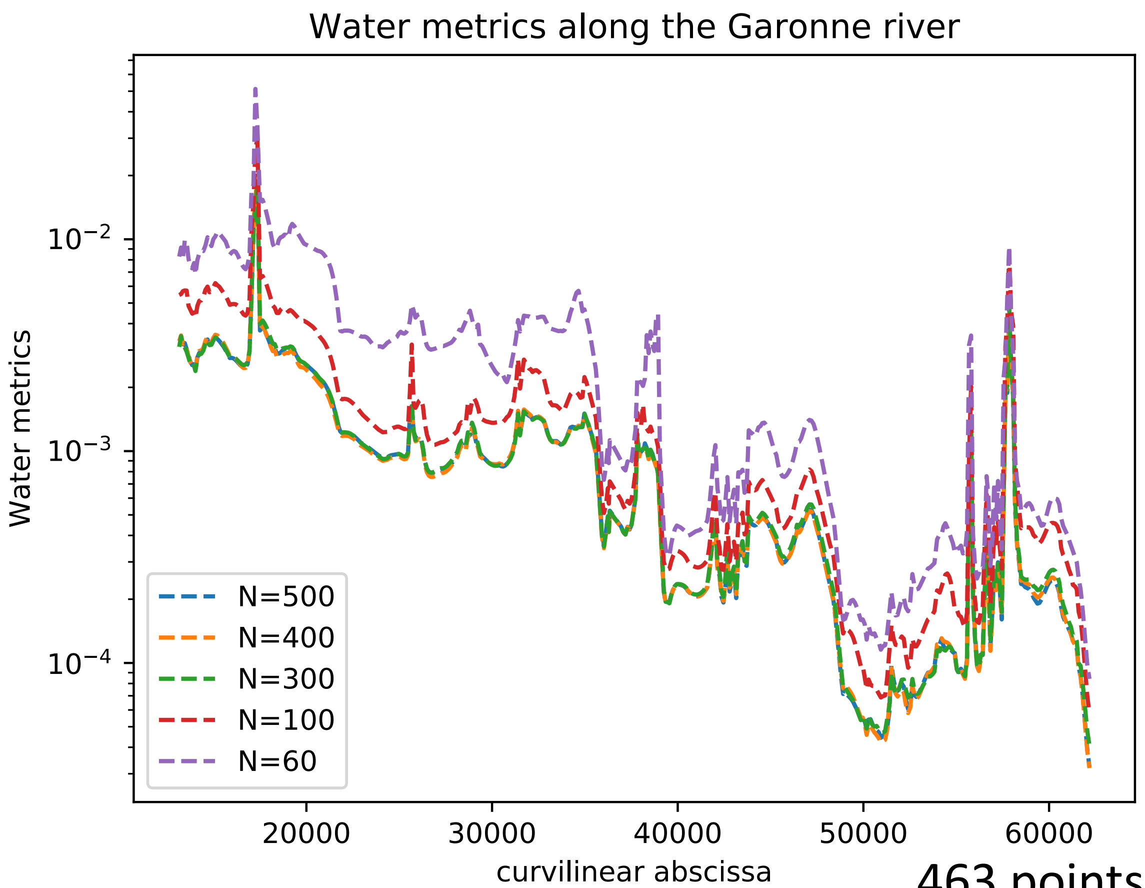
h:

Direct model

Surrogate model - Garonne (1.5D)

Choosing the appropriate learning sample

Inputs: Ks1, Ks2, Ks3, Q



463 points

$$e = \frac{1}{M}\sum_{s=1}^{M}(\tilde{h}_s - h_s)^2$$

$\tilde{h}$:

P = 3

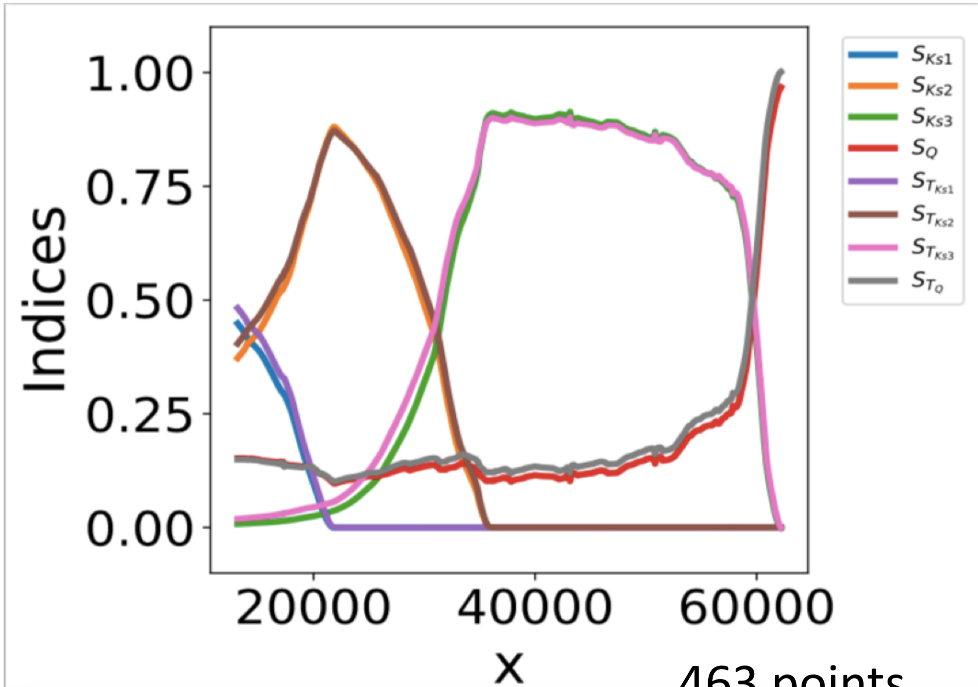
Various N

h:

P = 3

N = 1000

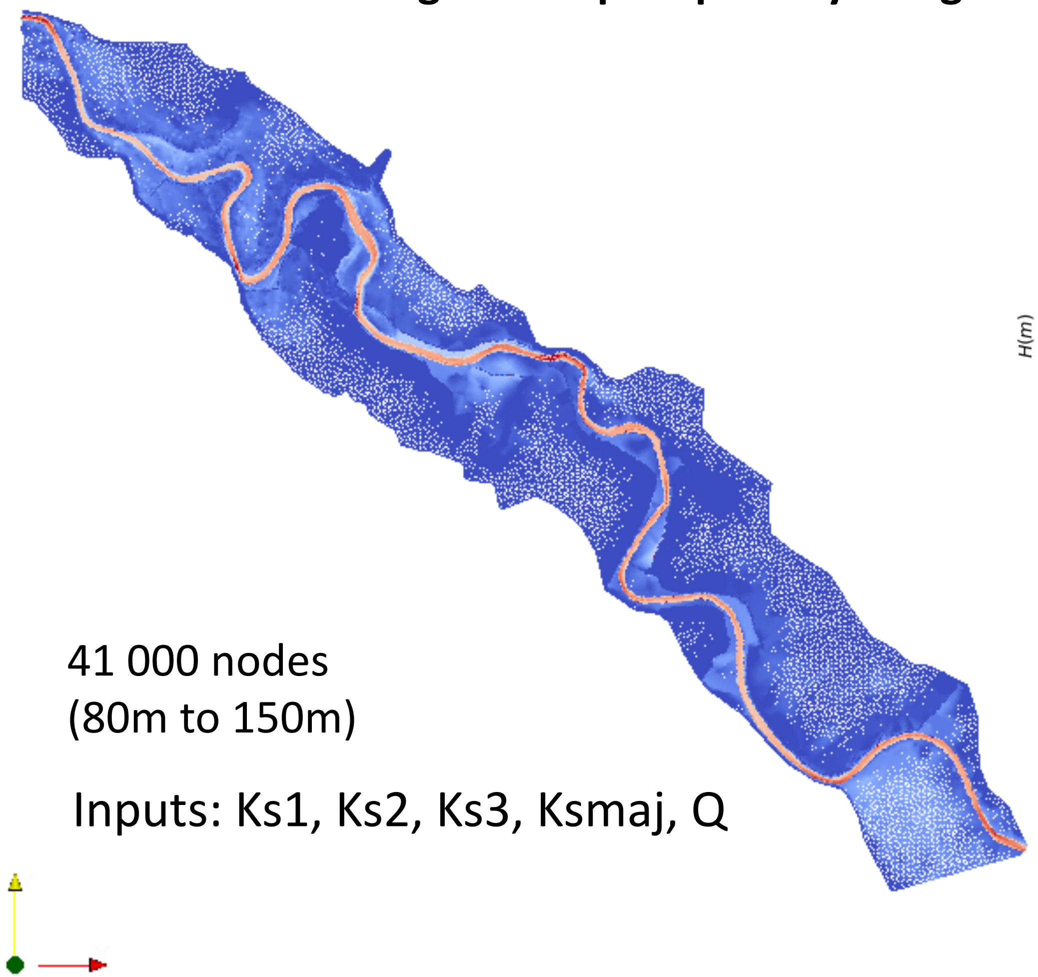
Sensitivity index for medium flow (1000 simulations)



463 points

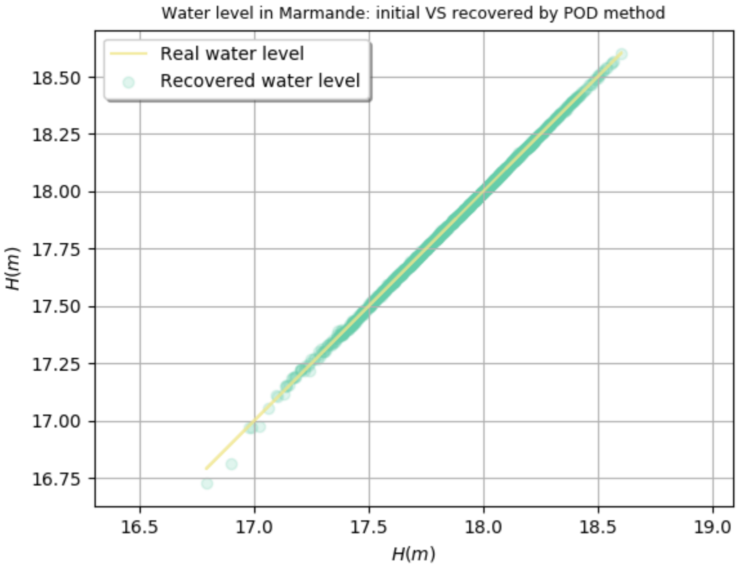
Discharge consistent with boundary conditions, each Ks leads its own section, but transitions between sections are visible

Reducing the output space by using a Proper Orthogonal Decomposition (POD)



41 000 nodes
(80m to 150m)

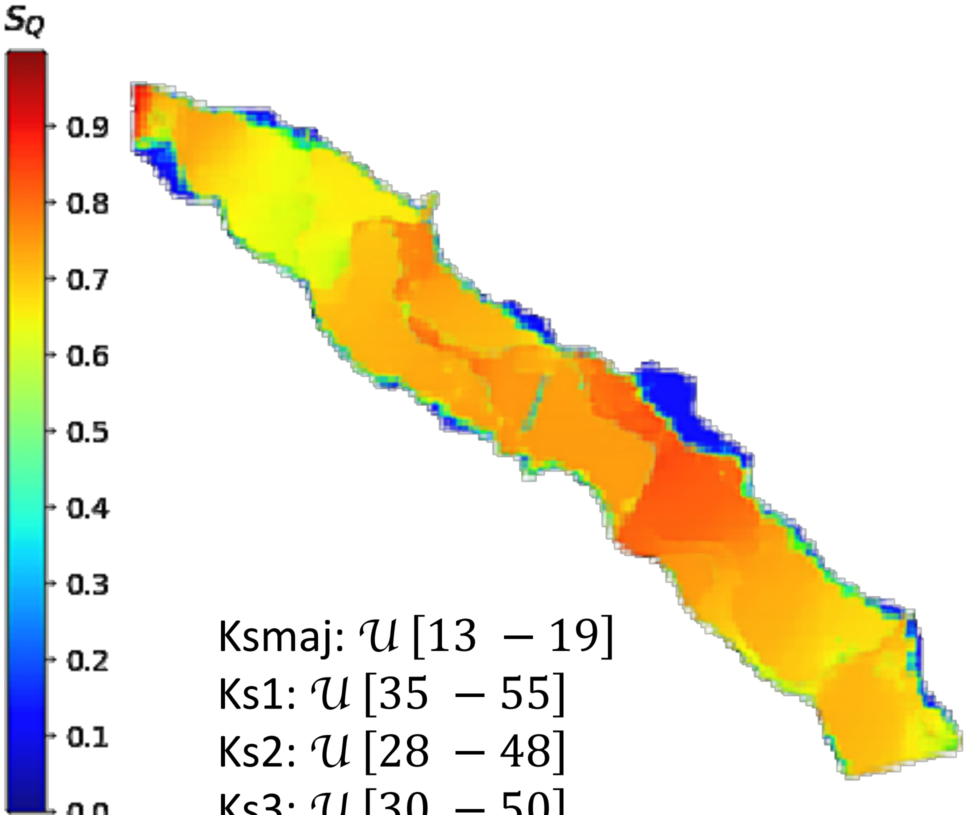
Inputs: Ks1, Ks2, Ks3, Ksmaj, Q



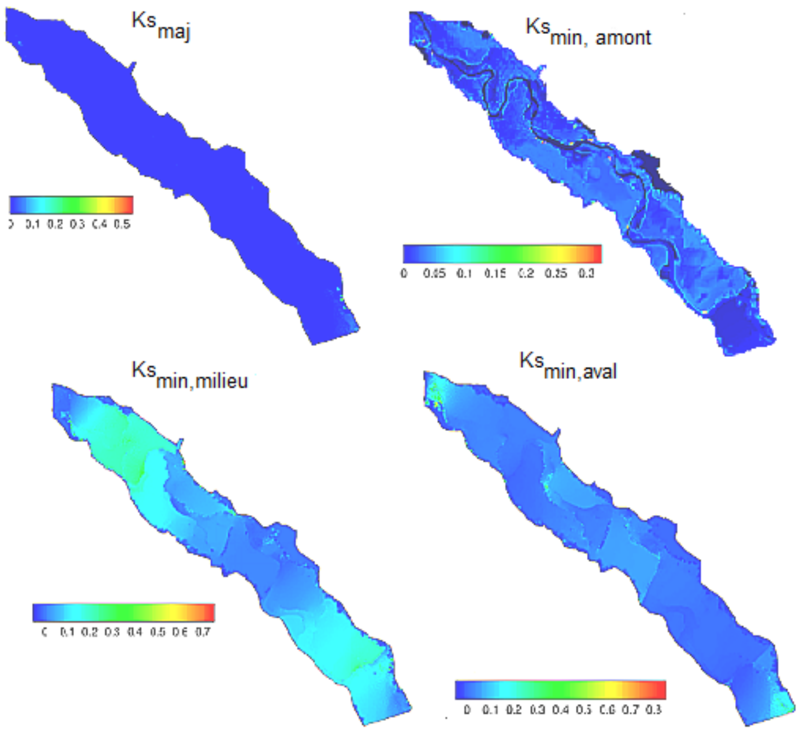
33 modes explain 99.87 % of water level variance

Surrogate construction gain : 1000 times cheaper

Sensitivity index map for high flow (1565 simulations)



Ksmaj: $\mathcal{U}\,[13\ -19]$
Ks1: $\mathcal{U}\,[35\ -55]$
Ks2: $\mathcal{U}\,[28\ -48]$
Ks3: $\mathcal{U}\,[30\ -50]$
Q: truncated $\mathcal{N}\,[6000\ -11000]$



The sensitivity depends on the range of the variations

Summary

Objective: to improve the estimate of the discharge in 1D and 2D models by assimilating SWOT data

- Part I: to generate inputs for the SWOT-HR simulator
  - Methods developed for spatializing / interpolating 1D and 2D model outputs into lat-lon rasters
- Part II: to produce SWOT data
  - Task performed by CNES
- Part III: to use SWOT data
  - Geolocation of cloud points and water elevation average
  - Design of surrogate models for 1.5D and 2D models on-going
  - Assimilation of SWOT data in an ensemble Kalman filter where the covariance matrix is estimated with surrogate models (next step)