

Comparing Data Assimilation algorithms for parameter estimation on TELEMAC-2D

J. Travert^{1,2}

C. Goeury^{1,2} V. Bacchi¹ F. Zaoui¹ S. Boyaval^{2,3}

¹EDF R&D, LNHE, ²LHSV (Saint-Venant Hydraulics Laboratory)

³MATHERIALS, Inria

XXIXth TELEMAC-MASCARET User Conference
Karlsruhe • 11th - 13th October 2023



Outline

1 Context

2 Data-driven parameter estimation

3 Conclusions and perspectives

Friction coefficients in flood models

Context

- ▶ Modelling floods on floodplains

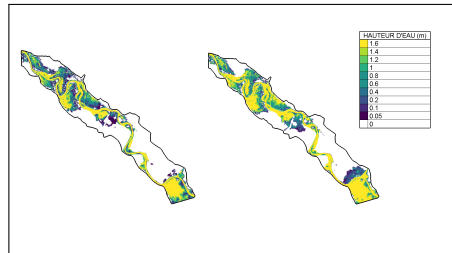


Figure 1: Example of two simulations with different sets of coefficients on the floodplains

Friction coefficients in flood models

Context

- ▶ Modelling floods on floodplains
- ▶ Shallow-Water Equations modelling:

$$\left\{ \begin{array}{l} \frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} = 0 \\ \frac{\partial(hu)}{\partial t} + \frac{\partial(hu^2)}{\partial x} + \frac{\partial(huv)}{\partial y} = -gh \frac{\partial \eta}{\partial x} - \frac{1}{\rho} \tau_{bx}(K) + \frac{h}{\rho} F_x + \nabla \cdot (h\nu_e \nabla u) \\ \frac{\partial(hv)}{\partial t} + \frac{\partial(huv)}{\partial x} + \frac{\partial(hv^2)}{\partial y} = -gh \frac{\partial \eta}{\partial y} - \frac{1}{\rho} \tau_{by}(K) + \frac{h}{\rho} F_y + \nabla \cdot (h\nu_e \nabla v) \end{array} \right.$$

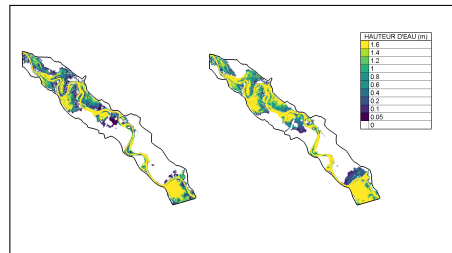


Figure 1: Example of two simulations with different sets of coefficients on the floodplains

Friction coefficients in flood models

Context

- ▶ Modelling floods on floodplains
- ▶ Shallow-Water Equations modelling:

$$\begin{cases} \frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} = 0 \\ \frac{\partial(hu)}{\partial t} + \frac{\partial(hu^2)}{\partial x} + \frac{\partial(huv)}{\partial y} = -gh \frac{\partial \eta}{\partial x} - \frac{1}{\rho} \tau_{bx}(K) + \frac{h}{\rho} F_x + \nabla \cdot (h\nu_e \nabla u) \\ \frac{\partial(hv)}{\partial t} + \frac{\partial(huv)}{\partial x} + \frac{\partial(hv^2)}{\partial y} = -gh \frac{\partial \eta}{\partial y} - \frac{1}{\rho} \tau_{by}(K) + \frac{h}{\rho} F_y + \nabla \cdot (h\nu_e \nabla v) \end{cases}$$

- ▶ Role of friction coefficients K on hydraulics variables

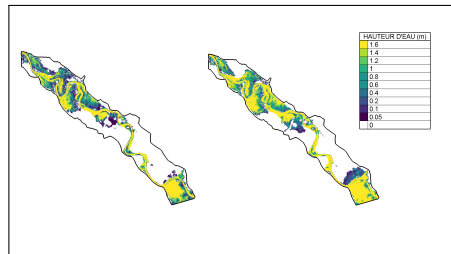


Figure 1: Example of two simulations with different sets of coefficients on the floodplains

Friction coefficients in flood models

Context

- ▶ Modelling floods on floodplains
- ▶ Shallow-Water Equations modelling:

$$\begin{cases} \frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} = 0 \\ \frac{\partial(hu)}{\partial t} + \frac{\partial(hu^2)}{\partial x} + \frac{\partial(huv)}{\partial y} = -gh \frac{\partial \eta}{\partial x} - \frac{1}{\rho} \tau_{bx}(K) + \frac{h}{\rho} F_x + \nabla \cdot (h\nu_e \nabla u) \\ \frac{\partial(hv)}{\partial t} + \frac{\partial(huv)}{\partial x} + \frac{\partial(hv^2)}{\partial y} = -gh \frac{\partial \eta}{\partial y} - \frac{1}{\rho} \tau_{by}(K) + \frac{h}{\rho} F_y + \nabla \cdot (h\nu_e \nabla v) \end{cases}$$

- ▶ Role of friction coefficients K on hydraulics variables

Objective

- ▶ Data-driven quantification of roughness coefficients on floodplains

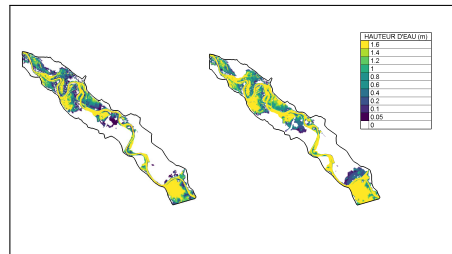


Figure 1: Example of two simulations with different sets of coefficients on the floodplains

Observations

Global context

- ▶ Overall drop in hydrometric data from in-situ stations

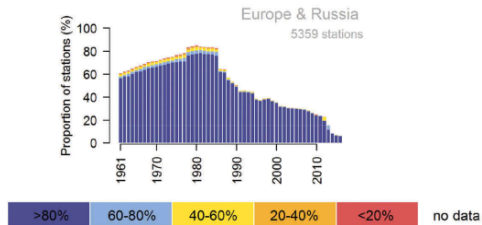


Figure 2: Proportion of available data between 1961 and 2019 by Crochemore et al. [2020]

- ▶ Frequent and accurate but limited data on floodplains

Observations

Global context

- ▶ Overall drop in hydrometric data from in-situ stations

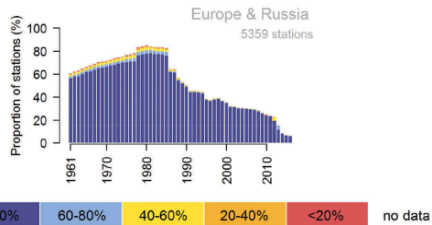


Figure 2: Proportion of available data between 1961 and 2019 by Crochemore et al. [2020]

- ▶ Frequent and accurate but limited data on floodplains

- ▶ Remote sensing: A "flood" of information



Observations

Global context

- ▶ Overall drop in hydrometric data from in-situ stations

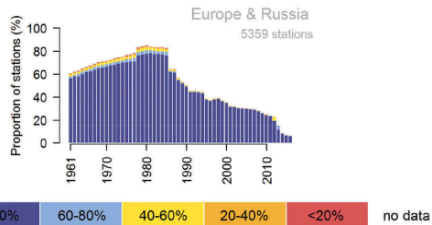


Figure 2: Proportion of available data between 1961 and 2019 by Crochemore et al. [2020]

- ▶ Frequent and accurate but limited data on floodplains

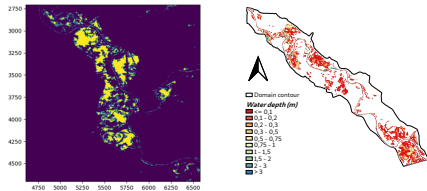
- ▶ Remote sensing: A "flood" of information



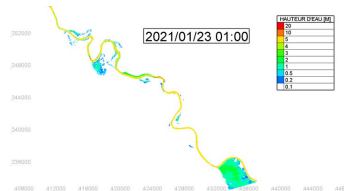
- ▶ More than 150 Earth Observation satellites
- ▶ Possible extraction of water depths or flood extent on the floodplains

Data Assimilation

Uncertain observations



Uncertain numerical modelling



Data Assimilation
Uncertain compromise

Outline

1 Context

2 Data-driven parameter estimation

3 Conclusions and perspectives

Numerical model

Objective:

- Evaluate K based on observations of h and the model (TELEMAC-2D):

$$\left\{ \begin{array}{l} \frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} = 0 \\ \frac{\partial(hu)}{\partial t} + \frac{\partial(hu^2)}{\partial x} + \frac{\partial(huv)}{\partial y} = -gh \frac{\partial \eta}{\partial x} - \frac{1}{\rho} \tau_{bx}(K) + \frac{h}{\rho} F_x + \nabla \cdot (h\nu_e \nabla u) \\ \frac{\partial(hv)}{\partial t} + \frac{\partial(huv)}{\partial x} + \frac{\partial(hv^2)}{\partial y} = -gh \frac{\partial \eta}{\partial y} - \frac{1}{\rho} \tau_{by}(K) + \frac{h}{\rho} F_y + \nabla \cdot (h\nu_e \nabla v) \end{array} \right.$$

System's unknowns: (quantities to correct)

- $h(x, y, t)$: water depths
- $u(x, y, t)$ et $v(x, y, t)$: horizontal velocity components integrated over vertical, along x and y
- $\eta = z(x, y, t) + h(x, y, t)$: elevation of free surface

Terms to be defined: (closing / parameters)

- τ_{bx} and τ_{by} : bottom shear stress, along x and y
- F_x and F_y : external forces components (Coriolis, wind, etc.)
- ν_e : effective viscosity (molecular, turbulent, dispersive)
- **Boundary conditions (BC)**: denoted u_{CL}, v_{CL} and h_{CL}

Numerical model

Objective:

- Evaluate K based on observations of h and the model (TELEMAC-2D):

$$\left\{ \begin{array}{l} \frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} = 0 \\ \frac{\partial(hu)}{\partial t} + \frac{\partial(hu^2)}{\partial x} + \frac{\partial(huv)}{\partial y} = -g h \frac{\partial \eta}{\partial x} - \frac{1}{\rho} \tau_{bx}(K) + \frac{h}{\rho} F_x + \nabla \cdot (h \nu_e \nabla u) \\ \frac{\partial(hv)}{\partial t} + \frac{\partial(huv)}{\partial x} + \frac{\partial(hv^2)}{\partial y} = -g h \frac{\partial \eta}{\partial y} - \frac{1}{\rho} \tau_{by}(K) + \frac{h}{\rho} F_y + \nabla \cdot (h \nu_e \nabla v) \end{array} \right.$$

System's unknowns: (quantities to correct)

- $h(x, y, t)$: water depths
- $u(x, y, t)$ et $v(x, y, t)$: horizontal velocity components integrated over vertical, along x and y
- $\eta = z(x, y, t) + h(x, y, t)$: elevation of free surface

Terms to be defined: (closing / parameters)

- τ_{bx} and τ_{by} : bottom shear stress, along x and y
- F_x and F_y : external forces components (Coriolis, wind, etc.)
- ν_e : effective viscosity (molecular, turbulent, dispersive)
- **Boundary conditions (BC)**: denoted u_{CL}, v_{CL} and h_{CL}

Example of closing equation: $\vec{\tau}_b = \frac{\rho g}{K(x,y)^2} \left(\frac{1}{h}\right)^{1/3} |\vec{u}| \vec{u}$ with K an uncertain parameter

Setting of the parameter estimation problem

Dynamic and observation models

Let $\mathcal{M}$ be a numerical model and $\mathbf{x}_k \in \mathbb{R}^{N_x}$ the hydrodynamic variables at time k :

$$\mathbf{x}_k = \mathcal{M}_{k:k-1}(\mathbf{x}_{k-1}, \boldsymbol{\theta}) + \boldsymbol{\eta}_k$$

where $\boldsymbol{\theta} \in \mathbb{R}^{N_p}$ are model parameters (e.g. friction) and $\boldsymbol{\eta}_k \in \mathbb{R}^{N_x}$ are model errors

Noised observations $\mathbf{y}_k \in \mathbb{R}^{N_y}$ available at discrete times k :

$$\mathbf{y}_k = \mathcal{H}(\mathbf{x}_k) + \boldsymbol{\epsilon}_k$$

where $\boldsymbol{\epsilon}_k \in \mathbb{R}^{N_y}$ are observation errors and $\mathcal{H} : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_y}$ an operator

Parameter estimation (Bayes inversion)

$$\mathbb{P}(\boldsymbol{\theta}|\mathbf{y}_k) = \frac{\mathbb{P}(\mathbf{y}_k|\boldsymbol{\theta})\mathbb{P}(\boldsymbol{\theta})}{\mathbb{P}(\mathbf{y}_k)}$$

Application

Description of the model

- ▶ $\mathcal{M}$: Shallow-Water Equations (TELEMAC-2D)
- ▶ Realistic simplified twin experiment
 - ▶ Minor bed of 100 m wide on 10 km long with two floodplains
 - ▶ Unstructured triangular mesh with $\mathcal{N} = 239\,486$ nodes
 - ▶ Three roughness sub-domains (bottom friction law: Manning-Strickler)
 - ▶ Boundary conditions: rating curve at the outlet and varying discharge at the inlet (from $30\text{ m}^3 \cdot \text{s}^{-1}$ to $620\text{ m}^3 \cdot \text{s}^{-1}$ for the first five minutes then constant)

openTELEMAC

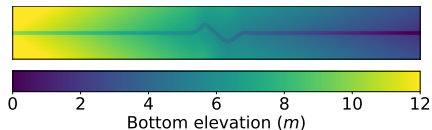


Figure 3: Bottom elevation

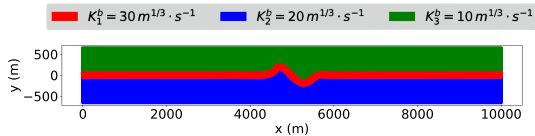


Figure 4: Roughness zones and values

State and observation vectors

- ▶ Simulations are controlled by the uncertain parameters $\theta = [K_1, K_2, K_3]$
- ▶ Noised observations: 1 hydrodynamic variable (water depth) at N_y points (9 and 11) every T_{obs} minutes (15 min and 90 min)
- ▶ The observation and state vectors at time k :

$$\mathbf{y}_k = [h_1^{(o)}, \dots, h_{N_y}^{(o)}]$$

$$\mathbf{x}_k = [h_1, \dots, h_N, \mathbf{u}_1, \dots, \mathbf{u}_N]$$

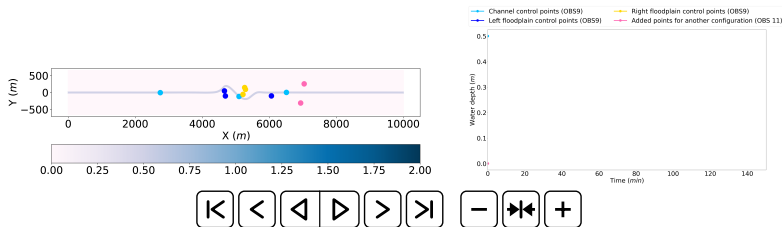


Figure 5: Example for the reference simulation with $N_y = 11$ and $T_{obs} = 15 \text{ min}$

Algorithms

Tested algorithms

- ▶ Variational formulation: 3D-VAR
- ▶ Ensemble methods: Ensemble Kalman Filter (EnKF)
- ▶ Sampling of $\mathbb{P}(\mathbf{x}|\mathbf{y})$: Particle Filter (PF) with resampling

Algorithms

Tested algorithms

- ▶ Variational formulation: 3D-VAR
- ▶ Ensemble methods: Ensemble Kalman Filter (EnKF)
- ▶ Sampling of $\mathbb{P}(\mathbf{x}|\mathbf{y})$: Particle Filter (PF) with resampling

Errors statistics

- ▶ $\epsilon_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ with $\mathbf{R} = (\sigma^I)^2 \cdot \mathbf{I}_{N_y}$ an error covariance matrix
- ▶ $\eta_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ with $\mathbf{Q}_k = \mathbf{0}_{N_x, N_x}$
- ▶ Background errors on the parameters $\sim \mathcal{N}(\mathbf{0}, \mathbf{B})$
- ▶ Following illustrations with $\theta^b = [35, 12, 18]$ and $\mathbf{B} = 100 \cdot \mathbf{I}_3$

Advantadges and disadvantages of the algorithms

	3D-VAR	EnKF	PF
Handling model non-linearity	✓	✓	✓
Handling non-Gaussian errors	✓	✗	✓
Computation resources	Low	Moderated	High
Parallel	✗	✓	✓

- ▶ 3D-VAR: More sensitive to the initialisation and local minimum
- ▶ PF: Need of resampling with hyperparameters
- ▶ EnKF: Subject to ensemble collapse

Variational formulation: 3D-VAR

3D-VAR: equations

- ▶ Minimise a cost function $J(\boldsymbol{\theta}) = J_b(\boldsymbol{\theta}) + J_o(\boldsymbol{\theta})$

$$J_b(\boldsymbol{\theta}) = \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}^b)^T \mathbf{B}^{-1}(\boldsymbol{\theta} - \boldsymbol{\theta}^b)$$

$$J_o(\boldsymbol{\theta}) = \frac{1}{2}(\mathbf{y} - \mathcal{H}(\mathcal{M}(\boldsymbol{\theta})))^T \mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathcal{M}(\boldsymbol{\theta})))$$

- ▶ Gradient descent with Broyden-Fletcher-Goldfarb-Shanno
- ▶ Implemented with ADAO (data assimilation and optimisation)



3D-VAR Results

Sensitivity to N_y

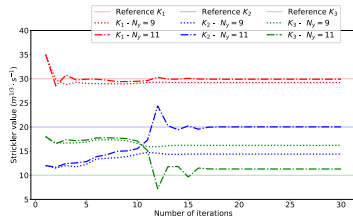


Figure 6: For $\#T_{obs} = 2$ and $\sigma^I = 0.05 m$

- Improvement in the identification with $N_y = 11$

Sensitivity to σ^I

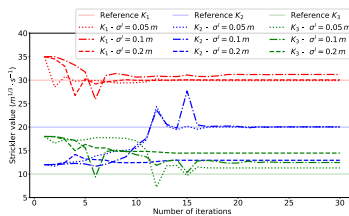


Figure 7: For $\#T_{obs} = 2$ and $N_y = 11$

- Identification of K_2 up to $\sigma^I = 0.1 m$

Sensitivity to $\#T_{obs}$

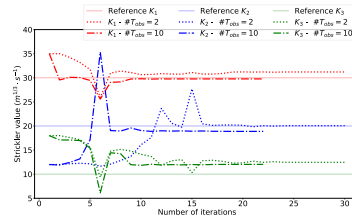


Figure 8: For $N_y = 11$ and $\sigma^I = 0.1 m$

- Little benefit from increased temporal observations

3D-VAR Results

Sensitivity to N_y

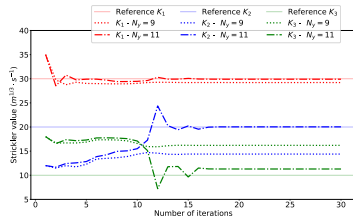


Figure 6: For $\#T_{obs} = 2$ and $\sigma^I = 0.05 m$

- Improvement in the identification with $N_y = 11$

Sensitivity to σ^I

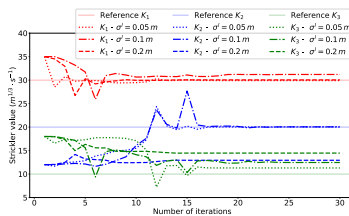


Figure 7: For $\#T_{obs} = 2$ and $N_y = 11$

- Identification of K_2 up to $\sigma^I = 0.1 m$

Sensitivity to $\#T_{obs}$

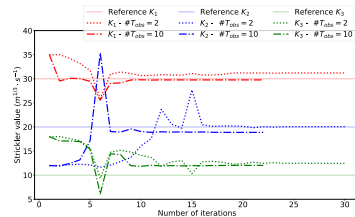


Figure 8: For $N_y = 11$ and $\sigma^I = 0.1 m$

- Little benefit from increased temporal observations

In the following, $N_y = 11$ and $\sigma^I = 0.1 m$

Ensemble Kalman Filter (EnKF)

EnKF principle

- ▶ We consider N_e members in the ensemble
- ▶ $\mathbf{H}$ and $\mathbf{M}$ are linear tangent

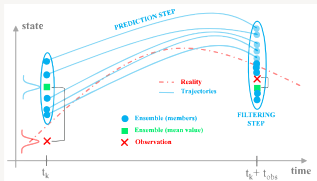


Figure 9: Schematic of the principle of EnKF by Zaoui et al. [2018]

Ensemble Kalman Filter (EnKF)

EnKF principle

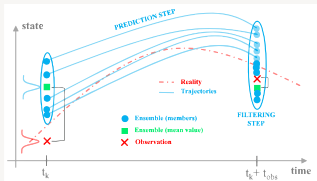


Figure 9: Schematic of the principle of EnKF by Zaoui et al. [2018]

- ▶ We consider N_e members in the ensemble
- ▶ $\mathbf{H}$ and $\mathbf{M}$ are linear tangent
- ▶ Propagation step: Each member $i \in [1, N_e]$ is propagated from t_k to t_{k+1} with the model and:

$$\mathbf{P}_{k+1}(\boldsymbol{\theta}_{k+1}^i) = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (\boldsymbol{\theta}_{k+1}^i - \bar{\boldsymbol{\theta}}_{k+1})(\boldsymbol{\theta}_{k+1}^i - \bar{\boldsymbol{\theta}}_{k+1})^T$$
$$\bar{\boldsymbol{\theta}}_{k+1} = \frac{1}{N_e} \sum_{i=1}^{N_e} \boldsymbol{\theta}_{k+1}^i$$

Ensemble Kalman Filter (EnKF)

EnKF principle

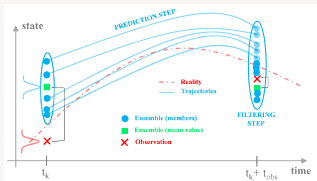


Figure 9: Schematic of the principle of EnKF by Zaoui et al. [2018]

- ▶ We consider N_e members in the ensemble
- ▶ $\mathbf{H}$ and $\mathbf{M}$ are linear tangent
- ▶ Propagation step: Each member $i \in [1, N_e]$ is propagated from t_k to t_{k+1} with the model and:

$$\mathbf{P}_{k+1}(\theta_{k+1}^i) = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (\theta_{k+1}^i - \bar{\theta}_{k+1})(\theta_{k+1}^i - \bar{\theta}_{k+1})^T$$

$$\bar{\theta}_{k+1} = \frac{1}{N_e} \sum_{i=1}^{N_e} \theta_{k+1}^i$$

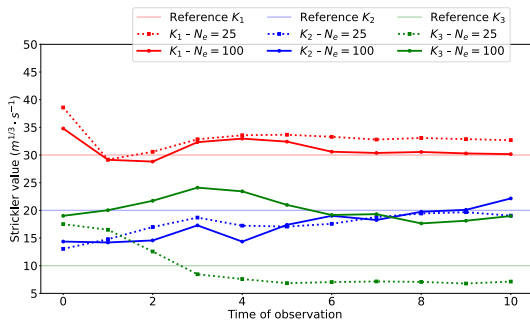
- ▶ Analysis step:

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1} \mathbf{H}^T (\mathbf{H} \mathbf{P}_{k+1} \mathbf{H}^T + \mathbf{R})^{-1}$$

$$\theta_{k+1}^{i,a} = \theta_{k+1}^i + \mathbf{K}_{k+1} (\mathbf{y}_{k+1} + \epsilon_k^i - \mathbf{H} \mathbf{x}_{k+1}^i)$$

EnKF Results

- Parameters updating at the end of an assimilation window (2h30)



- Slightly better results for high N_e

Figure 10: Evolution of the values of the mean parameters for different ensemble sizes and $\#T_{obs} = 10$

Particle Filter (PF)

Particle Filter principle

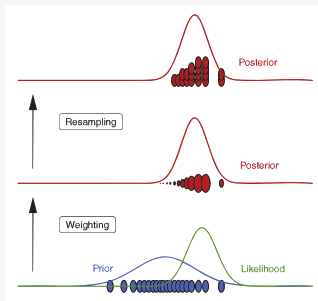


Figure 11: Schematic principle of PF by Carrassi et al. [2018]

- Initialisation: Sampling of N particles $[\theta^1, \dots, \theta^N]$

$$\mathbb{P}_k(\theta) \cong \sum_{i=1}^N w_k^i \delta(\theta - \theta_k^i)$$

with w_k^i the weight of the particle i at k and $w_0^i = \frac{1}{N}$

Particle Filter (PF)

Particle Filter principle

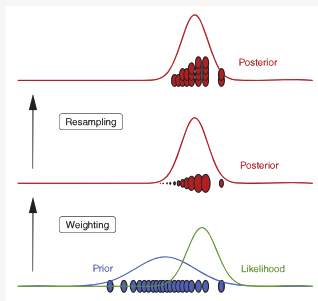


Figure 11: Schematic principle of PF by Carrassi et al. [2018]

- Initialisation: Sampling of N particles $[\theta^1, \dots, \theta^N]$

$$\mathbb{P}_k(\theta) \cong \sum_{i=1}^N w_k^i \delta(\theta - \theta_k^i)$$

with w_k^i the weight of the particle i at k and $w_0^i = \frac{1}{N}$

- Propagation of particles by the model

Particle Filter (PF)

Particle Filter principle

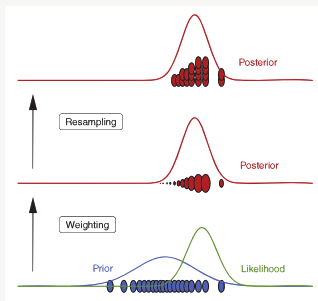


Figure 11: Schematic principle of PF by Carrassi et al. [2018]

- Initialisation: Sampling of N particles $[\theta^1, \dots, \theta^N]$

$$\mathbb{P}_k(\theta) \cong \sum_{i=1}^N w_k^i \delta(\theta - \theta_k^i)$$

with w_k^i the weight of the particle i at k and $w_0^i = \frac{1}{N}$

- Propagation of particles by the model
- Analysis step: $w_{k+1}^{a,i} \cong w_{k+1}^{f,i} \mathbb{P}(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}^i, \theta_{k+1}^i)$

$$\mathbb{P}(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}^i, \theta_{k+1}^i) = \prod_{l=1}^{N_y} \frac{1}{\sigma^l \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x_{k+1}^{i,l} - y_{k+1}^l}{\sigma^l} \right)^2}$$

Particle Filter (PF)

Particle Filter principle

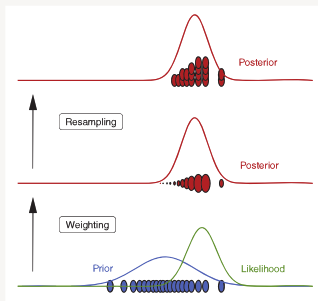


Figure 11: Schematic principle of PF by Carrassi et al. [2018]

- ▶ Initialisation: Sampling of N particles $[\theta^1, \dots, \theta^N]$

$$\mathbb{P}_k(\theta) \cong \sum_{i=1}^N w_k^i \delta(\theta - \theta_k^i)$$

with w_k^i the weight of the particle i at k and $w_0^i = \frac{1}{N}$

- ▶ Propagation of particles by the model
- ▶ Analysis step: $w_{k+1}^{a,i} \cong w_{k+1}^{f,i} \mathbb{P}(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}^i, \theta_{k+1}^i)$

$$\mathbb{P}(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}^i, \theta_{k+1}^i) = \prod_{l=1}^{N_y} \frac{1}{\sigma^l \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x_{k+1}^{i,l} - y_{k+1}^l}{\sigma^l} \right)^2}$$

- ▶ Iterative resampling: Kernel Density Estimation

PF Results

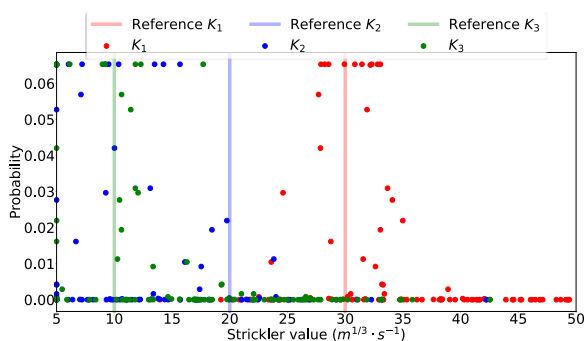


Figure 12: Without resampling with 100 particles and $\#T_{obs} = 2$

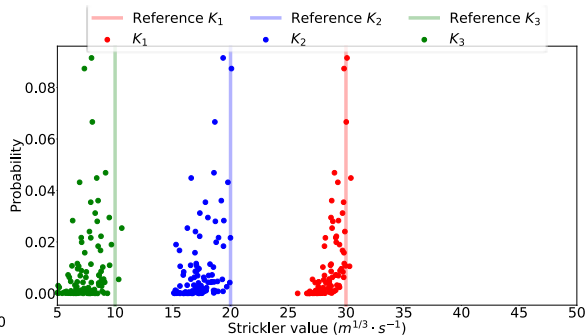


Figure 13: With resampling after 5 iterations with 100 particles and $\#T_{obs} = 2$

- ▶ Best identification of parameters with resampling
- ▶ Similar results with 25 particles and resampling

Outline

1 Context

2 Data-driven parameter estimation

3 Conclusions and perspectives

Conclusions and perspectives

3D-VAR	EnKF	PF
• Important computation time but few resources • Sensitive to initialisation	• Parallel • Need more observations	• Parallel but iterative resampling • Hyperparameters complex to determine
3D-VAR gives the best results overall, and the two others give similar results		

Conclusions and perspectives

- ▶ Possibility with TelApv and TELEMAC-MASCARET to implement Data Assimilation algorithms
- ▶ Using uncertain observations can help to improve the identification of friction parameters
- ▶ Need to quantify the performance of the algorithms on simpler cases

Thanks for your attention!
Danke schön!

- Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G. (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):e535.
- Cohen, S., Brakenridge, G. R., Kettner, A., Bates, B., Nelson, J., McDonald, R., Huang, Y.-F., Munasinghe, D., and Zhang, J. (2018). Estimating floodwater depths from flood inundation maps and topography. *JAWRA Journal of the American Water Resources Association*, 54(4):847–858.
- Crochemore, L., Isberg, K., Pimentel, R., Pineda, L., Hasan, A., and Arheimer, B. (2020). Lessons learnt from checking the quality of openly accessible river flow data worldwide. *Hydrological Sciences Journal*, 65(5):699–711.
- Turlach, B. A. et al. (1993). Bandwidth selection in kernel density estimation: a review. Technical report, Humboldt Universitaet Berlin.
- Zaoui, F., Goeury, C., and Audouin, Y. (2018). Ensemble integrations of telemac-mascaret for the optimal model calibration. In *XXVth Telemac & Mascaret User Club*.

Appendixes

Water depth extraction

How to map water depths?

- ▶ Flood extent + Topographic data
- ▶ Example: FwDET method (Floodwater Depth Estimation Tool) (Cohen et al. [2018])

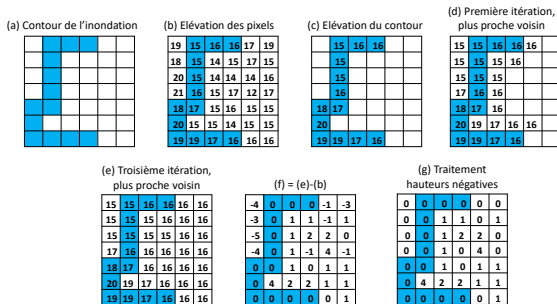


Figure 14: Description of FwDET method

Results: Water depths extraction

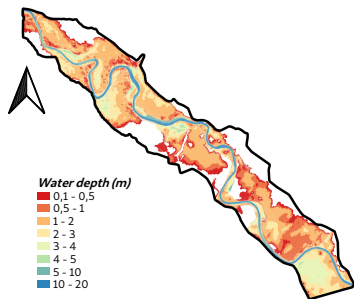


Figure 15: Simulated water depths

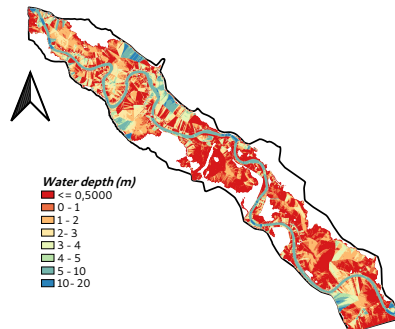


Figure 16: Application of FwDET

- Under-estimation of water depths in the floodplain except near the edges

Results: Water depths extraction on a controlled case

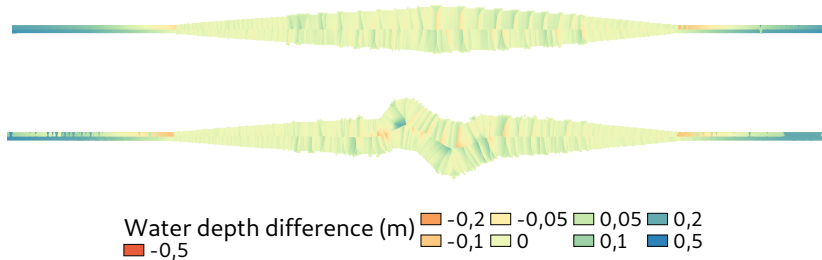


Figure 17: Differences between simulated water depths and FwDET

- ▶ Errors mainly upstream and downstream where the detected contour does not correspond to null water depth
- ▶ Rest of the domain: Low RMSE

Data Assimilation

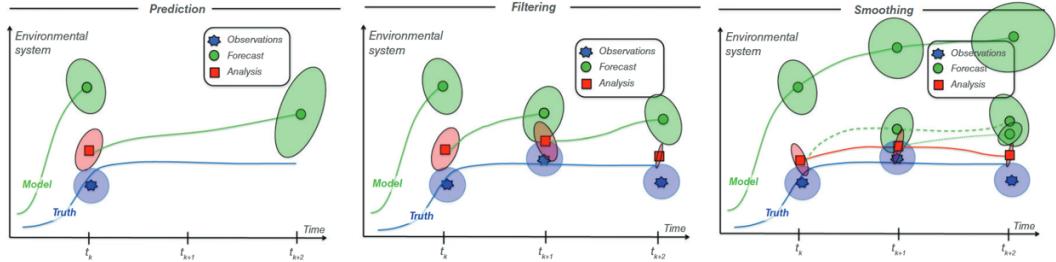


Figure 18: Data assimilation by Carrassi et al. [2018]

Bases

- ▶ Uncertain observations
- ▶ Uncertain modelling
- ▶ Analysis = uncertain compromise
- ▶ Different problems (eg. calibration)

Different methods and frameworks

- ▶ States and parameters estimation
- ▶ Ensemble approaches, variational formulations, hybrids, etc
- ▶ Limitations: Resampling, gradient derivation

Resampling procedure

- ▶ Non-parametric estimation by Kernel Density Estimation (KDE)
- ▶ The objective is to evaluate the probability density function of a sample $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ of random vector $\mathbf{X}$

Kernel reconstruction: 1D Case

Density approximation of $\mathbb{P}$ by the empirical density $\mathbb{P}_n$

$$\mathbb{P}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)$$

- ▶ K : kernel, symmetrical distribution with respect to 0. For example, Gaussian $(0, 1)$, uniform $(-1, 1)$...
- ▶ h : bandwidth

Resampling procedure - 2

- Bandwidth selection strongly influences the estimation (Turlach et al. [1993])

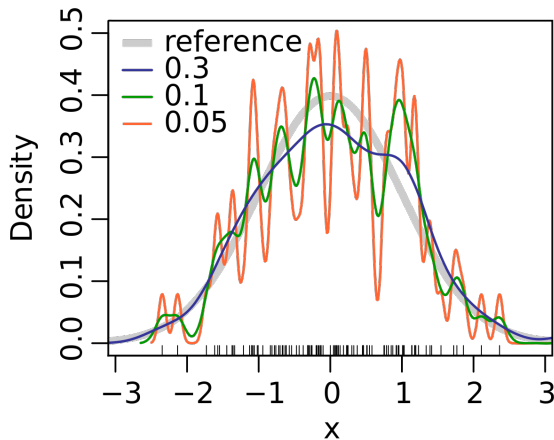


Figure 19: Kernel density estimation for different h

Particle Filter with another variable

- ▶ Observations: γ^0 for each pixel

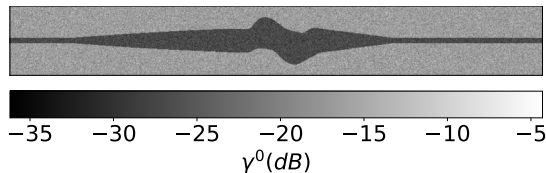


Figure 20: Transformed water depths in γ^0 based on probabilistic approaches (if a pixel is inundated, we sample in a Gaussian distribution)

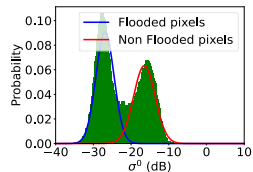


Figure 21: Distributions example for inundated and not inundated pixels

- ▶ Likelihood function: $w_{k+1}^{i,a} = w_k^i \prod_{j=1}^{N_p} \frac{1}{\sigma^j \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\mu^j - y_{k+1}^j}{\sigma^j} \right)^2}$ with N_p the number of pixels, α an hyperparameter and (μ^j, σ^j) depending if the simulated pixel is wet or not
- ▶ Minor bed identification is simpler, but more complex for floodplains