IMPROVING WATER LEVELS FORECAST IN THE GIRONDE ESTUARY USING DATA ASSIMILATION ON A 2D NUMERICAL MODEL: CORRECTION OF TIME-DEPENDENT BOUNDARY CONDITIONS THROUGH A TRUNCATED KARHUNEN-LOÈVE DECOMPOSITION WITHIN AN ENSEMBLE KALMAN FILTER

V. LABORIE\(^{(1,2)}\), N. GOUTAL\(^{(1,3)}\), S. RICCI\(^{(4)}\)

(1) Laboratoire d’Hydraulique Saint-Venant
(2) Cerema Eau, Mer et Fleuves (EMF)
(3) EDF R&D/LNHE
(4) CERFACS, GLOBC/CECI CNRS
Plan

- Context and motivation
- Uncertainty quantification
- Improving water levels forecast with data assimilation
- Conclusions and perspectives
Plan

Context and motivation

Uncertainty quantification

Improving water levels forecast with data assimilation

Conclusions and perspectives
The Gironde estuary
The Gironde estuary

- 635 km²
- Confluence between 2 rivers
  - Dordogne : 380 m³/s
  - Garonne : 630 m³/s
- 75 km long (Bec d’Ambès to the mouth)
- Downstream width : 12 km
- Maritime influence
- Inflows from the Atlantic Ocean :
  15 à 25000 m³/tide cycle
Stakes

• Human: 185 cities, 1 million inhabitants; agriculture
• Industrial stakes: Blayais hydroelectric power station, Pauillac petrol terminal

Natural Hazards

• Sea river flooding
• Climate change
  => Marine submersions
  => Floods

Context and motivation

Uncertainty quantification

Data assimilation

Conclusions and perspectives

Crédit photo: Archives Sud-Ouest

Source: Schéma directeur de prévision des crues Adour-Garonne, DREAL Midi-Pyrénées, 29/12/2015
Water levels forecast in Gironde estuary using a Telemac2D numerical model

- Based on:
  - 2D shallow water equations
  - Unstructured mesh (space discretization): 7351 nodes, 12838 elements (mesh0)
  - Output on each node: (H,U,V)
Water levels forecast in Gironde estuary using a Telemac2D numerical model

Calibration parameters
• Friction coefficients: $K_{s1}$, $K_{s2}$, $K_{s3}$, $K_{s4}$
• Wind influence coefficient: $C_{dz}$

Forcings
• Meteorological: wind and pressure
• Maritime boundary conditions (CLMAR)
• River discharges: in Garonne ($Q_{GAR}$) and Dordogne ($Q_{DOR}$)

Other inputs: topography and bathymetry
• No overflowing
• Bathymetry provided by GPMB
**Water levels forecast in Gironde estuary using a Telemac2D numerical model**

**EVENTS**
- **calibration**: 4 events among which 2003
- **validation**: 6 events among which 1999

<table>
<thead>
<tr>
<th>Événement</th>
<th>Débits sur la Dordogne</th>
<th>Garonne</th>
<th>Plage de variation du coefficient de marée</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 au 14/03/2006</td>
<td>850 à 1550 m³/s</td>
<td>1700 à 4700 m³/s</td>
<td>37 à 83</td>
</tr>
<tr>
<td>27 au 31/03/2006</td>
<td>500 à 950 m³/s</td>
<td>650 à 1200 m³/s</td>
<td>78 à 115</td>
</tr>
<tr>
<td>3 au 6/05/2004</td>
<td>450 à 700 m³/s</td>
<td>1650 à 2450 m³/s</td>
<td>86 à 105</td>
</tr>
<tr>
<td>2 au 9/02/2003</td>
<td>600 à 2200 m³/s</td>
<td>1200 à 5900 m³/s</td>
<td>43 à 90</td>
</tr>
<tr>
<td>25 au 31/12/1999</td>
<td>430 à 1830 m³/s</td>
<td>400 à 3600 m³/s</td>
<td>45 à 102</td>
</tr>
<tr>
<td>24 au 30/04/1998</td>
<td>300 à 1200 m³/s</td>
<td>630 à 3100 m³/s</td>
<td>70 à 105</td>
</tr>
<tr>
<td>3 au 10/02/1996</td>
<td>300 à 850 m³/s</td>
<td>880 à 2550 m³/s</td>
<td>71 à 87</td>
</tr>
<tr>
<td>20 au 26/12/1995</td>
<td>250 à 550 m³/s</td>
<td>500 à 1200 m³/s</td>
<td>79 à 107</td>
</tr>
<tr>
<td>17 au 20/03/1988</td>
<td>500 à 1800 m³/s</td>
<td>1250 à 5500 m³/s</td>
<td>99 à 115</td>
</tr>
<tr>
<td>12 au 17/12/1981</td>
<td>1000 à 2350 m³/s</td>
<td>1700 à 7050 m³/s</td>
<td>57 à 106</td>
</tr>
</tbody>
</table>
Water levels forecast in Gironde estuary using a Telemac2D numerical model

EVENTS
- **calibration**: 4 events among which 2003
- **validation**: 6 events among which 1999

Stations
- 13 stations

Criteria
- Root mean square error (RMSE)
- **High tide nash (PM)**

=> Data assimilation techniques

**Source**: Hissel (2010), Projet Gironde : rapport final d’évaluation du modèle Gironde
Data assimilation

Observations

• the "true" state of the system is unknown and must be estimated
• measurements and models are imperfect

What do we want?

• Identify the most influential variables in time and space
• find an optimal combination of measurements and simulations

(from Rochoux & al, 2015)
Improve **the prediction of water levels at the most sensitive stations of the estuary using ensemble data assimilation techniques.**

**Partie I : uncertainty quantification (UQ-GSA)**
- Objective: identify the **most influential variables and establish a space-time hierarchy**
- Scientific latch: space-time variables / forcings (maritime influence)
- Technological lock: 2D code (resources / environment)

**Partie II : data assimilation using ensemble method (EnKF-γ-KLBC)**
- Objective: **correct relevant variables** by optimizing the observation network
- Scientific locks:
  - ✔ Dispersion/characterization of the ensemble
  - ✔ Interactions between variables and equifinality
- Technology lock: 2D code (HPC / sequential task farming)
State of the art for the control vector

Boundary conditions

Parameters

Estuaries

Rivers

Habert & al (2016)

Ricci & al (2011)

Vrugt & al (2006)

Tamura & al (2014)

Raboudi & al (2019)

Estuaries

Meteo

Defforge & al (2019)

Siripatana (2018)

Almeida & al (2015)

Canas & al (2019)

Uncertainties put on

Boundary conditions

Rivers

Canas & al (2017)

Bertino & al (2002)

Frolov & al (2009)

Context and motivation

Uncertainty quantification

Data assimilation

Conclusions and perspectives

Data assimilation

Uncertainty quantification

Conclusions and perspectives

Context and motivation
Plan

Context and motivation

Uncertainty quantification: most influential inputs in time and space

Improving water levels forecast with data assimilation

Conclusions and perspectives
Global sensitivity analysis (GSA) using variance decomposition (ANOVA): Methodology for perturbing inputs

- Sobol' sequence
- 8 uncertain variables:
  - scalar: $K_s$, $CD_z$
  - Time-dependent: $CLMAR$, $QDOR$, $QGAR$

- Mesh and number convergence study
Global sensitivity analysis (GSA) using variance decomposition (ANOVA)

Why « global »?

- Statistical approach based on the resampling of the input space
- Parameters varying simultaneously in the complete range of values
- Very large number of simulations

\[ \text{Ne} \times (d+2) \]

\[ ST_i = E_{x_i} (V_{x_i}(Y | X_i)) \]

D'après Saltelli & al, 2000
Global sensitivity analysis (GSA) using variance decomposition (ANOVA): Methodology for perturbing time-dependent inputs

Objective
Preserve the temporal error correlation

Assumption
Time chronicles represented by Gaussian processes

Method
Reduction of the input space with a Karhunen Loève decomposition

Temporal vector: perturbed member $p$ over $N$

$$q_p = (q_p(t_1), \ldots, q_p(t_N))^T = \sum_{i=1}^{n_{\text{modes}}} \sqrt{\lambda_i} \Phi_i \alpha_{i,p}.$$
Global sensitivity analysis (GSA) using variance decomposition (ANOVA):

Methodology for perturbing time-dependent inputs

**Objective**
Preserve the temporal error correlation

**Assumption**
Time chronicles represented by Gaussian processes

**Method**
Reduction of the input space with a Karhunen Loève decomposition

\[
q_p = (q_p(t_1), \ldots, q_p(t_N))^T = \sum_{i=1}^{n_{\text{modes}}} \sqrt{\lambda_i} \Phi_i \alpha_i.
\]
Global sensitivity analysis (GSA) using variance decomposition (ANOVA): Methodology for perturbing time-dependent inputs

**Objective**
Preserve the temporal error correlation

**Assumption**
Time chronicles represented by Gaussian processes

**Method**
Reduction of the input space with a Karhunen Loève decomposition

Temporal vector: perturbed member \( p \) over \( N \) modes

\[
q_p = (q_p(t_1), \ldots, q_p(t_N))^T = \sum_{i=1}^{n_{\text{modes}}} \sqrt{\lambda_i} \Phi_i \alpha_i.
\]
Global sensitivity analysis (GSA) using variance decomposition (ANOVA):
Methodology for perturbing time-dependent inputs

Objective
Preserve the temporal error correlation

Assumption
Time chronicles represented by Gaussian processes

Method
Reduction of the input space with a Karhunen Loève decomposition

Temporal vector: perturbed member \( p \) over \( Ne \)

\[
q_p = (q_p(t_1), \ldots, q_p(t_N))^T = \sum_{i=1}^{\text{n_modes}} \sqrt{\lambda_i} \Phi_i \alpha_i.
\]
Global sensitivity analysis (GSA) using variance decomposition (ANNOVA): Methodology for perturbing time-dependent inputs

Objective
Preserve the temporal error correlation

Assumption
Temporel chronicals Time Chronicles represented by Gaussian Processes

Method
Reduction of the input space with a Karhunen Loève decomposition

Ensemble of $N_e$ members of perturbed temporal vectors
Global sensitivity analysis (GSA) using variance decomposition (ANOVA): results

Method
Computation of Sobol’ indices

Results
Determination of the spatio-temporal evolution of the zone of influence

Exemple
Space-time homogeneity

Plan

- Context and motivation
- Uncertainty quantification
- Improving water levels forecast with data assimilation
- Conclusions and perspectives
The Ensemble Kalman filter

- Choice of a control vector
- Construction of the ensemble

\[ K_s_i \quad \text{with} \quad i \in (1,..,4) \]

\[ \alpha_i \quad \text{with} \quad i \in (1,..,n_{\text{modes,CLMAR}}) \]

\[ C_d_z \]

\[ \alpha_{i,GAR} \quad \text{with} \quad i \in (1,..,n_{\text{modes,QGAR}}) \]

\[ \alpha_{i,DOR} \quad \text{with} \quad i \in (1,..,n_{\text{modes,QDOR}}) \]
The Ensemble Kalman filter

- Choice of a control vector
- construction of the ensemble
- Sequential in 2 steps (analysis and prediction)

\[ \text{Analysis} = \text{Prediction} + \left[ K \left( \text{observations} - y \right) \right] \]

\[ K = P_{y,y} (P_{y,y} + R)^{-1} \]

(Source: Carrassi & al (2018))
Validation with twin experiments

• Influence zone validation
• Time-varying parameters
• Simultaneous reconstruction of parameters and forcings

Evaluation on real experiments

• Evolution of parameters and forcings
• Performances
Validation of influence zones

Good estimation of constant or periodic parameters

Improvement of water level forecasting

Over-estimation of the amplitude of the parameters

Difficulties in estimating Ks3 (confluence zone)

Highlighting the equifinality on the Ks
Joint estimation of forcing by decomposition of KL and parameters

Control variable
\( K_{s1}, K_{s2}, K_{s3}, \alpha_{CLMAR} \)
\( \alpha_{GARONNE}, \alpha_{DORDOGNE} \)
Ne = 100

Observations
12 stations
Frequency
1 hour

Assimilation window
1 hour
Overlapping
1 hour

Evolution of the mean of the analysed ensemble WITH redispersion for \( K_{s1} \)

Context and motivation
Uncertainty quantification
Data assimilation
Conclusions and perspectives
Joint estimation of forcing by decomposition of KL and parameters

Control variable
Ks1, Ks2, Ks3, $\alpha_{\text{CLMAR}}$
$\alpha_{\text{GARONNE}}, \alpha_{\text{DORDOGNE}}$
Ne = 100

Observations
12 stations
Frequency
1 hour

Assimilation window
1 hour
Overlapping
1 hour

Evolution of the mean of the analysed ensemble WITH redispersion for Ks3

Context and motivation
Uncertainty quantification
Data assimilation
Conclusions and perspectives

High frequency oscillations !!!
Joint estimation of forcing by decomposition of KL and parameters

Control variable
Ks1, Ks2, Ks3, $\alpha_{\text{CLMAR}}$
$\alpha_{\text{GARONNE}}$, $\alpha_{\text{DORDOGNE}}$
Ne = 100

Observations
12 stations
Frequency
1 hour

Assimilation window
1 hour
Overlapping
1 hour

Evolution of the mean of the analysed ensemble WITH redispersion for $\alpha_{\text{CLMAR}}$

No convergence
Reconstruction of the maritime boundary condition CLMAR

**Control variable**
- $K_s1$, $K_s2$, $K_s3$, $\alpha_{CLMAR}$
- $\alpha_{GARONNE}$, $\alpha_{DORDOGNE}$
- $Ne = 100$

**Observations**
- 12 stations
- **Frequency**
  - 1 hour

**Assimilation window**
- 1 hour
- **Overlapping**
  - 1 hour

**Perfect reconstruction**
Joint estimation of forcing by decomposition of KL and parameters

Control variable
Ks1, Ks2, Ks3, $\alpha_{CLMAR}$
$\alpha_{GARONNE}, \alpha_{DORDOGNE}
Ne = 100$

Observations
12 stations
Frequency
1 hour

Assimilation window
1 hour
Overlapping
1 hour

Root mean square error without and with assimilation

RMSE = 10 cm
Joint estimation of forcing by decomposition of KL and parameters

Control variable
Ks1, Ks2, Ks3, $\alpha_{\text{CLMAR}}$
$\alpha_{\text{GARONNE}}, \alpha_{\text{DORDOGNE}}$
Ne = 100

Observations
12 stations
Frequency
1 hour

Assimilation window
1 hour
Overlapping
1 hour

No convergence of modal coefficients
High-frequency oscillations for parameters
Difficulties in estimating Ks3 (confluence zone)
Equifinality on Ks and $\alpha$

Good estimation of friction parameters
Very good reconstruction of time-dependent forcings
Improvement of water levels forecast
Joint estimation of forcing by decomposition of KL and parameters

**Control variable**
- $K_s1$, $K_s2$, $K_s3$, $\alpha_{\text{CLMAR}}$
- $\alpha_{\text{GARONNE}}$, $\alpha_{\text{DORDOGNE}}$
- Ne = 100

**Observations**
- 12 stations
- Frequency: 1 hour

**Assimilation window**
- Frequency: 1 hour
- Overlapping: 1 hour

**Context and motivation**

Joint estimation of forcing by decomposition of KL and parameters

**Uncertainty quantification**

- Good estimation of friction parameters
- Very good reconstruction of time-dependent forcings
- Improvement of water levels forecast

**Data assimilation**

- No convergence of modal coefficients
- High-frequency oscillations for parameters
- Difficulties in estimating $K_s3$ (confluence zone)
- Equifinality on $K_s$ and $\alpha$

**Conclusions and perspectives**

Study of the most efficient configuration of the method

- Control variables: time-dependent (discharges, maritime boundary): modes, parameters: $K_s1$, $K_s2$, $K_s3$
- Optimal observation network: 7 tide gauges, Ne = 100
Plan

Context and motivation

Uncertainty quantification

Improving water levels forecast with data assimilation

Conclusions and perspectives
1. Identification of the spatio-temporal evolution of the zones of influence of time-dependent variables

- **Uncertainty Quantification** Study (ANOVA-GSA) for the 2003 storm
- **Ensemble kalman filter**

2. Joint Estimation of Parameters and time-dependent forcings

- **Correlations** and **specific equifinality**
- **reconstruction** of **maritime** boundary conditions
- **Sensitive area of the confluence**
- **NO convergence of modal coefficients, variability of friction coefficients**

3. Improvement of water levels estimation along the estuary under reanalysis

- **Better estimation** of high tides, storm peaks downstream of the estuary, signal amplitude
- **Most efficient** configuration in a **synthetic** setting
Validation with twin experiments
Influence zone validation
Time-varying parameters
Simultaneous reconstruction of parameters and forcings

Evaluation on real experiments
Evolution of parameters and forcings
Performances
2. Improvement of the methodology for a better representation of physical processes

- **Localisation** based on the analysis of the spatio-temporal evolution of sensitivity indices (Sobol')
- **Emulation** of the Ensemble Kalman filter (Raboudi & al, Frolov & al)
- **Iterative Kalman Filter** (Sakov & al, 2012)

3. Extension of the control vector to 2D uncertain time and space dependent fields

- **Integration of additional forcings**: meteorological forcing, bathymetry

4. Diversification of observations

- **Satellite data** (SWOT project)

5. Model reduction and operationnability

- **Metamodels**
- **Predictive mode**
Thank you for your attention.

V. LABORIE
Vanessya.Laborie@cerema.fr

Sincere thanks to: Nicole GOUTAL, Sophie RICCI, Matthias DE LOZZO, Yoann AUDOUIN, Laurent DIEVAL, Patrick OHL, Guillaume MORVAN, Gwénaële JAN, Philippe SERGENT

Collaborations: Cerema, LHSV, CERFACS, SCHAPI, Météo-France, SPC GAD, SHOM, Rouen university