CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

CENTRE NATIONAL D'ÉTUDES SPATIALES

SWOT Science Team meeting

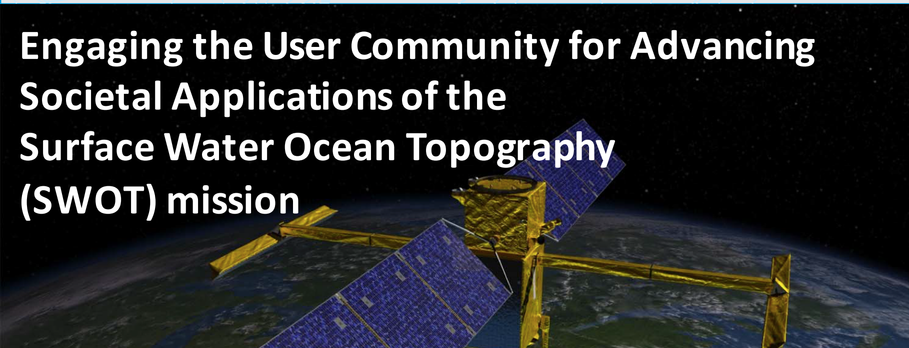
June, 28th, 2017

Ensemble-based Data assimilation with surrogate models – Garonne test case study

Sophie Ricci – CECI, CERFACS/CNRS
N. Goutal (EDF/LHSV), N. El Mocayd (CECI), F. Moussu (CECI)

Acknowledgement for funding: CNES (TOSCA) and SWOT project

Context



Potential applications: (S. Cherchali, April 2017)

- Transboundary rivers management (international & inter-regional)
- A better modelling of flood
- Clear water management for urban, industrial and agricultural needs
- Hydroelectricity production management
- Prevention of the propagation of epidemics
- Fluvial navigation support
- Integrated management for estuaries

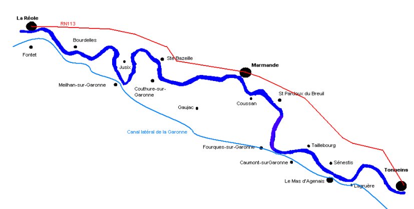
Current data latency of 45j is a big showstopper for applications, need for short term delivery, temporal series, delivery of processed data with errors and accuracy information at 2-3 days latency

- Models and DA algorithms compatible with applications' constraints

PhD N. El Mocayd (CNES-EDF), « Polynomial surrogates for open-channel flows in random steady state » (Review ENMO)

- Demonstrative test cases study

Garonne test case (Tonneins-La Réole), 1D MASCARET, 2D TELEMAC



Surrogate model in hydraulics – Polynomial Chaos

A - Context:
- Water resources management at EDF
- Flood forecasting at SCHAPI

Sources of uncertainty
- epistemic errors: friction Ks
- random errors: upstream forcing Q

Quantity of interest
- water level
- discharge

B - Motivation
- Low cost estimation of statistical moments and pdfs
- Description of water level cov. matrix for EnKF
- Reduced-cost EnKF

Li, Xiu, J. of Comp. Phys., 2009.

Non-intrusive PC surrogate model

$$h \cong \sum_i \widehat{h}_i \Psi_i$$

- Water level is expressed as a truncated sum of polynoms that form an orthogonal basis w.r.t. the uncertain input random variables (Ks,Q):

$$h_{pc}(a) \cong \mathcal{M}_{pc}(\mathbf{x}(\zeta)) = \sum_{i=0}^{N_{pc}-1} \widehat{h}_i(a)\,\Psi_i(\zeta)$$

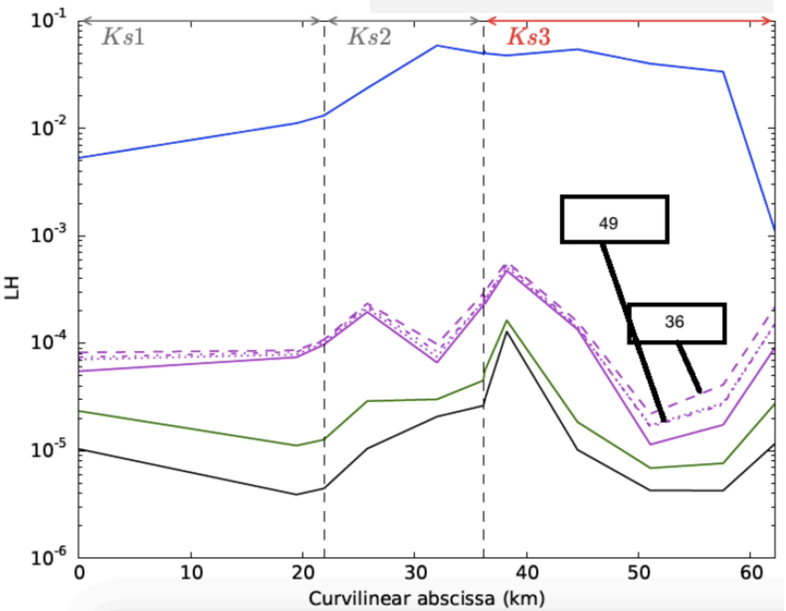
- Coefficients are computed with a quadrature method

$$N_{pc} = \frac{(n+P)!}{(n!\ P!)}$$

$$\widehat{h}_i = < h, \Psi_i(\zeta) > \cong \sum_{k=1}^{N_s} \mathcal{M}(\mathbf{x}^k)\,\Psi_i(\zeta^k)\,w^k$$

Validation of the PC surrogate over the Garonne River
L2-error through the channel
for P = 1, 6, 10, 15 w.r.t. MC (100000 samples)

PhD N. El Mocayd, 2017



Surrogate model in hydraulics – POD/Gaussian Process

A - Context:
- ○ Water resources management at EDF
- ○ Flood forecasting at SCHAPI

Sources of uncertainty
- ○ epistemic errors: friction Ks
- ○ random errors: upstream forcing Q

Quantity of interest
- ○ water level
- ○ discharge

B - Motivation
- ○ Low cost estimation of statistical moments and pdfs
- ○ Description of water level cov. matrix for EnKF
- ○ Reduced-cost EnKF Li, Xiu, J. of Comp. Phys., 2009.

Non-intrusive PC surrogate model
- The POD of the snapshot matrix provides the modes and coefficients of the decomposition of the water level.

$$h_{gp}(a) = \mathcal{M}_{\mathrm{gp}}(\mathbf{x}) = \sum_{i=1}^{p} \widehat{h}_i(a)\,\Psi^i(a)$$

$$\mathbf{S} = \sum_{k=1} \sigma_k\,\mathbf{u}_k\,\mathbf{v}_k^{\mathrm{T}} = \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^{\mathrm{T}}$$

- The POD coefficients are interpolated with Gaussian process regression to formulate the surrogate model.

$$f(\mathbf{x}) \sim GP(m(\mathbf{x}), k(\mathbf{x},\mathbf{x}')),\ \text{with}$$
$$m(\mathbf{x}) = \mathbb{E}\left[f(\mathbf{x})\right],$$
$$k(\mathbf{x},\mathbf{x}') = \mathbb{E}\left[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))\right]$$

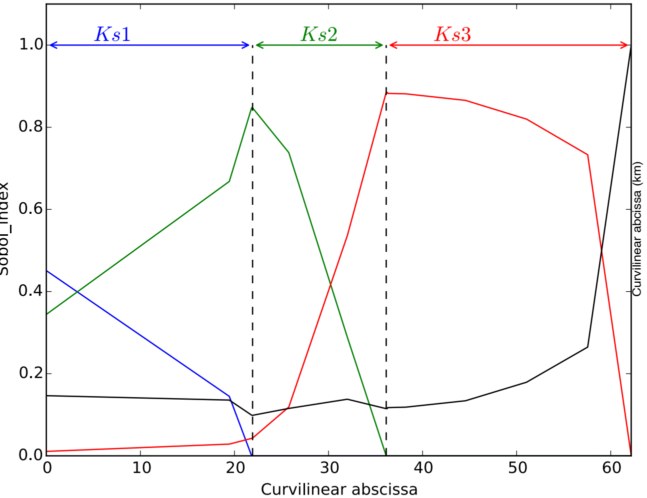
Validation of the POD-GP and PC surrogates w.r.t. MC

Model	Snapshots	Statistics D	*p-value*
GP	49	$7{,}95 \cdot 10^{-3}$	0.004
	121	$3{,}97 \cdot 10^{-3}$	0.409
	256	$3{,}02 \cdot 10^{-3}$	0.751
PC	49	$7{,}15 \cdot 10^{-3}$	0.012
	121	$4{,}95 \cdot 10^{-3}$	0.172
	256	$4{,}93 \cdot 10^{-3}$	0.175

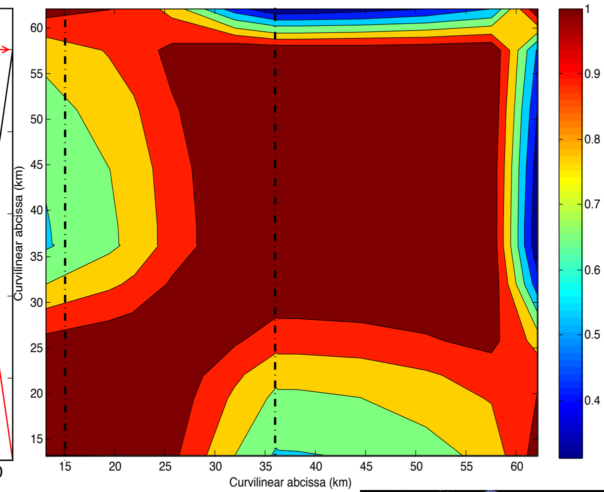
	RMSE	
Model	S_{K_S}	S_Q
GP	$2{,}17 \cdot 10^{-2}$	$6{,}71 \cdot 10^{-2}$
PC	$1{,}15 \cdot 10^{-2}$	$4{,}83 \cdot 10^{-2}$

Sensitivity Analysis and Data Assimilation with surrogate models – The Garonne river

Sobol' indices w.r.t Q, Ks1, Ks2, Ks3 over the Garonne River



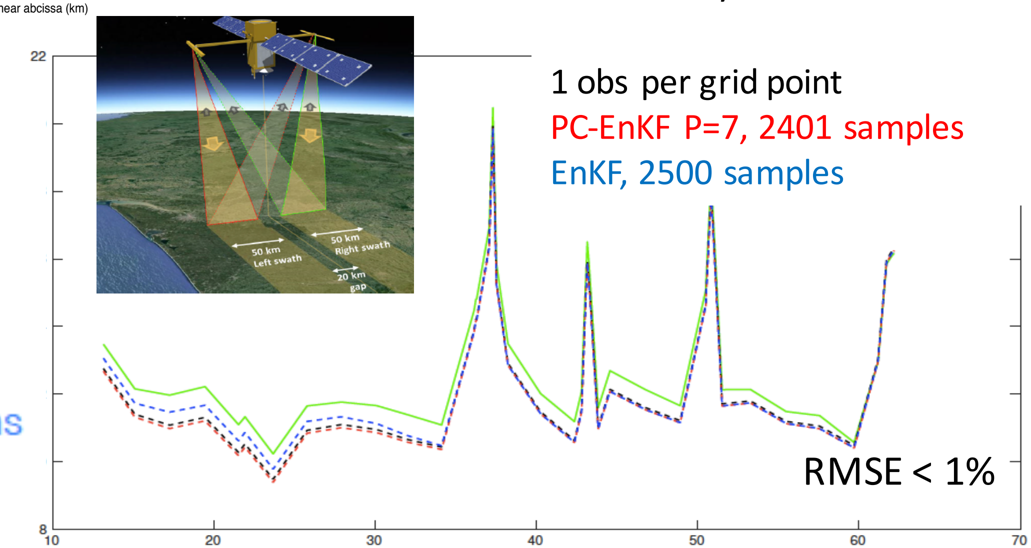
Water level corr. matrix over the Garonne River



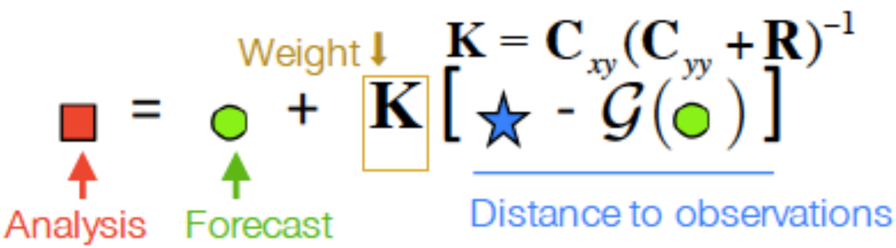
DA OSSE

- Use SWOT-HR (see Poster F. moussu)
- Use basic SWOT-like model outputs

Water level OSSE analysis with PC-EnKF

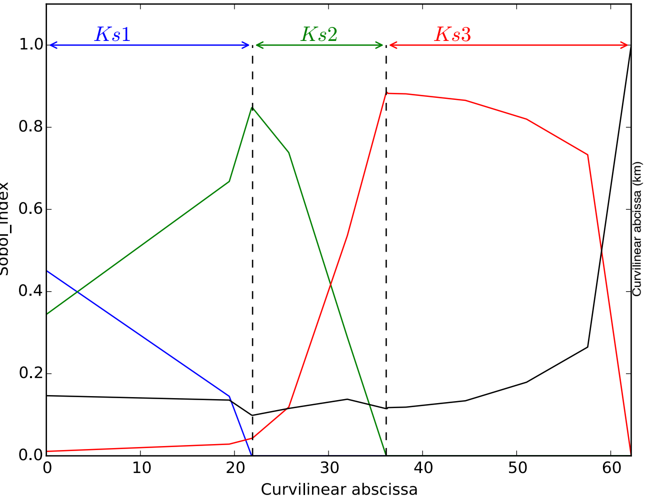


Ensemble Kalman Filter

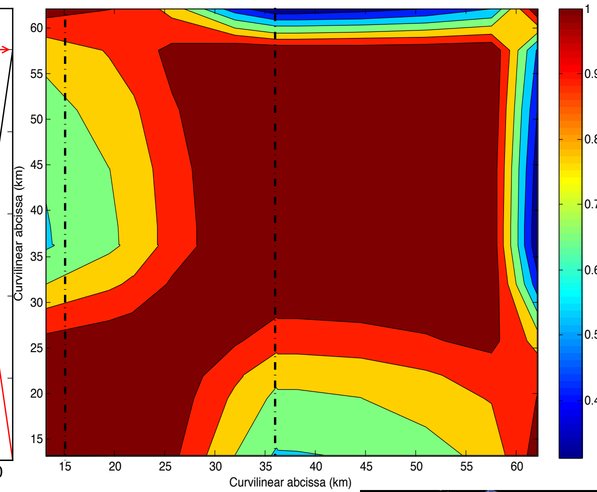


Sensitivity Analysis and Data Assimilation with surrogate models – The Garonne river

Sobol' indices w.r.t Q, Ks1, Ks2, Ks3 over the Garonne River



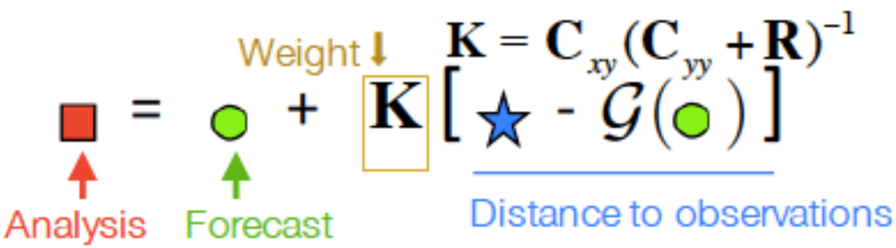
Water level corr. matrix over the Garonne River



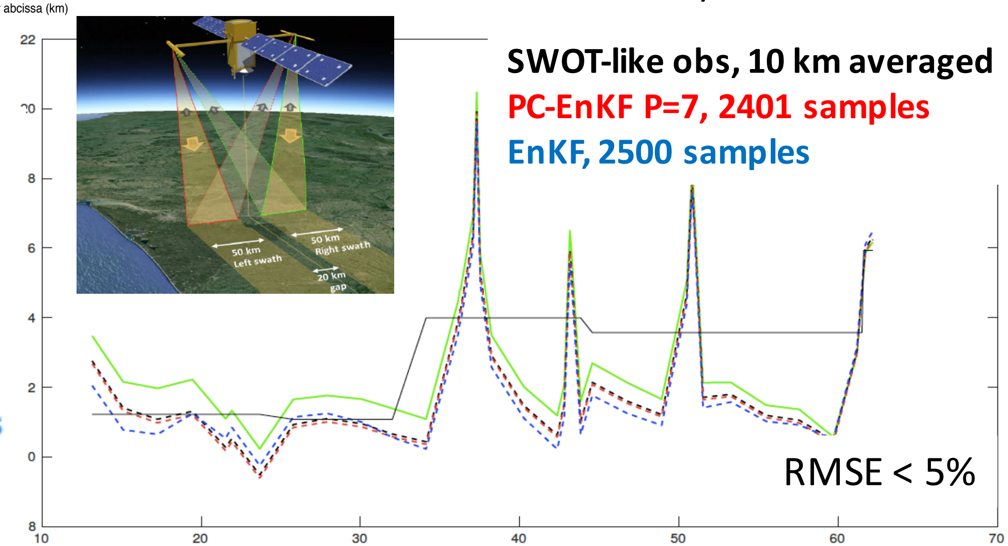
DA OSSE

- Use SWOT-HR (see Poster F. moussu)
- Use basic SWOT-like model outputs

Ensemble Kalman Filter



Water level OSSE analysis with PC-EnKF



Conclusions and Perspectives

Conclusions:

The group has focused on Ensemble-based DA algorithm and Sensitivity Analysis

The use of surrogate models is a promising lead for low-cost DA compatible with applications using SWOT data combined with other data.

The group also worked on the use of SWOT-HR simulator (see F. Moussu Poster)

Perspectives:

Improvement of the algorithm: extend the control vector to bathymetry

Pepsi & DA Pepsi Challenge

Application on the Gironde estuary with 2D model

Papers:

N. El Mocayd, S. Ricci, N. Goutal, M. C. Rochoux, S. Boyaval, C. Goeury, D. Lucor, O. Thual. Polynomial surrogate model for open-channel flows in steady state. Under review, Environmental Modeling Assessment.

P. Roy, N. El Mocayd, S. Ricci, M. De Lozzo, M. Rochoux, J.-C. Jouhaud, N. Goutal. Comparison of Polynomial Chaos and Gaussian Process surrogates for uncertainty quantification and correlation estimation of spatially-distributed open-channel steady flows. Under review, Stochastic Environmental Research and Risk Assessment.

The group thanks CNES-TOSCA funding and J.-F. Crétaux PI of the SWOT-discharge project