CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

CECI, CERFACS/CNRS
June 21th, 2016

OpenTurns Users Meeting
Uncertainty quantification with OpenTurns
Application sin CFD and hydrodynamics @CERFACS

Sophie Ricci
Global Change and Climate modeling team

Acknowledgments to: N. El Mocayd, P. Roy, N. Goutal, C. Goeury, B. Iooss, J.-C. Jouhaud, O. Thual, M. Rochoux, M. Baudin, A.-L. Popelin, G. Blatman, R. Ata, MRI team@EDF, OpenPALM team@CERFACS

Uncertainty quantification with OpenTurns
Applications in CFD and hydrodynamics @ CERFACS

UQ main ideas

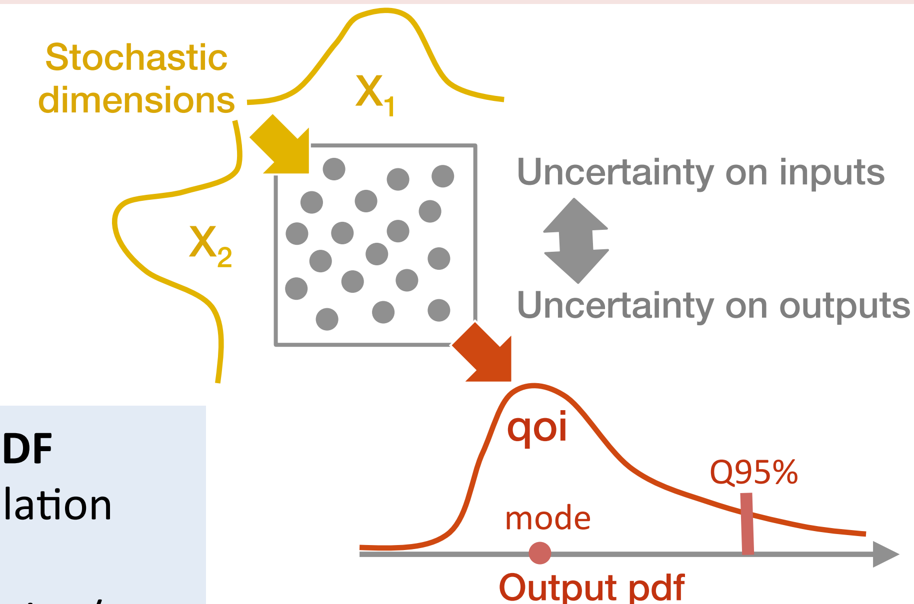
- Global uncertainty analysis: Describe the pdf of the outputs given the pdf of the inputs
- Sensitivity analysis (local/global): which input variable has the most impact on the output ?
- Reliability analysis: what is the probability that the qoi exceeds a given threshold ?
- Ensemble-based approach rather than deterministic approach

- **Non-intrusive Monte Carlo approach**

easy to implement
- **Direct Monte-Carlo simulations**

expensive especially for extreme events
- **Surrogate model**

low cost solution for pdf and statistics estimation



Applications at CERFACS, in collaboration with EDF

- in CFD with AVBP (LES) for turbine/blade simulation

(CFD team @CERFACS, B. Iooss@MRI)
- in hydraulics with MASCARET for flood forecasting/ water resources management

(Globc team @CERFACS, N. Goutal+C. Goeury @EDF/LNHE

A POD Surrogate model of the *LS89* cases for Statistical Analysis

Experimental design

Experiment LS89 with experimental heat fluxes measurements at VKI:

- 5 pales in cascade to mimic periodicity
- Heat fluxes are mesaured for different turbulence intensity

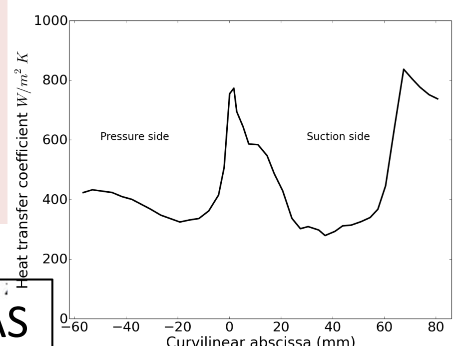
Post-combustion air Inflow T0 ≈ 400 K

Input random variables for LES:

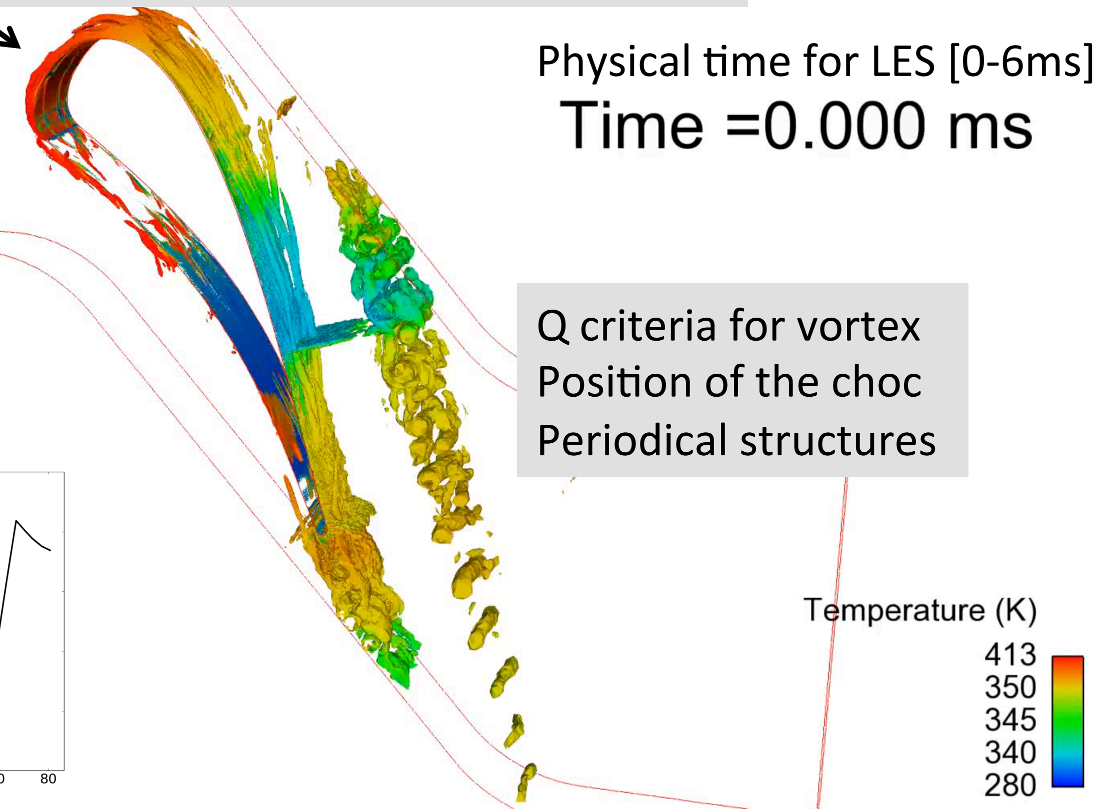
- turbulence intensity Tu
- inflow angle α

Output quantity of interest:
heat transfert coefficient along the curvilinear abscissa

$$h = \frac{q_{wall}}{T_0 - T_{wall}}$$



Physical time for LES [0-6ms]



Q criteria for vortex
Position of the choc
Periodical structures

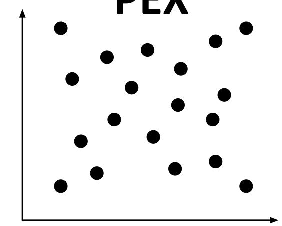
P. Roy, M2@CERFCAS

A POD Surrogate model of the *LS89* cases for Statistical Analysis
General workflow

Using < 100 *LES* simulations resolved on a minimal mesh, a representative surrogate model is created combining a POD and a Gaussian Process estimator.

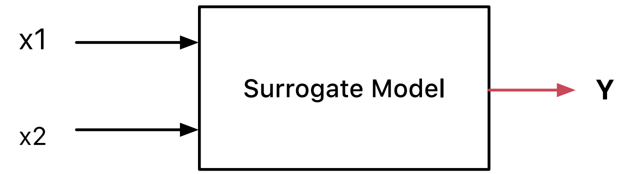
- No hypothesis on the input pdf
- Sobol or Halton law with Openturns
- otlhs will be tested with discrepancycriteria
- Investigate re-sampling based on minimal error

PEX



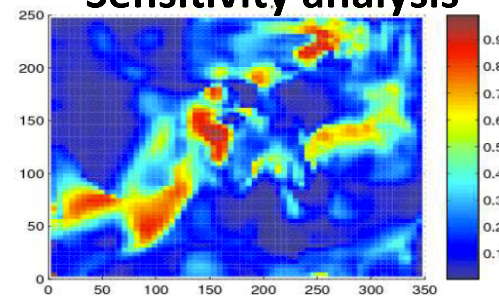
- The qoi degree of freedom is reduced with a POD using SVD decomposition (JPOD).
- h(x) (1000 grid points) is reduced to a limited number of POD coefficients (100 coeficients)
- Kriging on the POD coefficients (Scikit learn with Gaussian process, Openturns will be used) to provide full description of the surrogate model over space

Surrogate model



- Use with OpenTurns
- Sobol indices (FAST method) along curvilinear abscissa

Sensitivity analysis



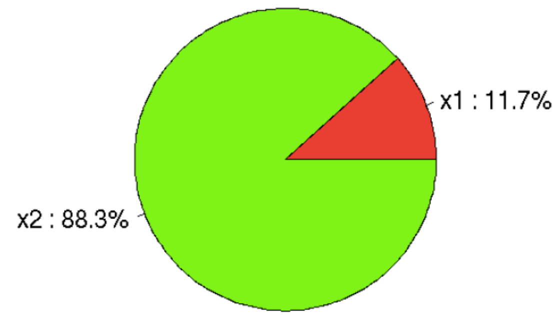
A POD Surrogate model of the *LS89* cases for Statistical Analysis

Backwater curves in a channel Flow case

This test case is representative of a real and non-linear case which allows to demonstrate the ability of JPOD to perform the analysis necessary of LS89

- Channel test case with constante slope and width
- POD Surrogate model w.r.t Ks and Q

Q and Ks Sobol's indices



$$\frac{dh}{ds} = I\frac{1-\left(\frac{h}{h_n}\right)^{-10/3}}{1-\left(\frac{h}{h_c}\right)^{-3}}$$

$$h_c = \left(\frac{q^2}{g}\right)^{1/3} \qquad h_n = \left(\frac{q^2}{IK_s^2}\right)^{3/10}$$

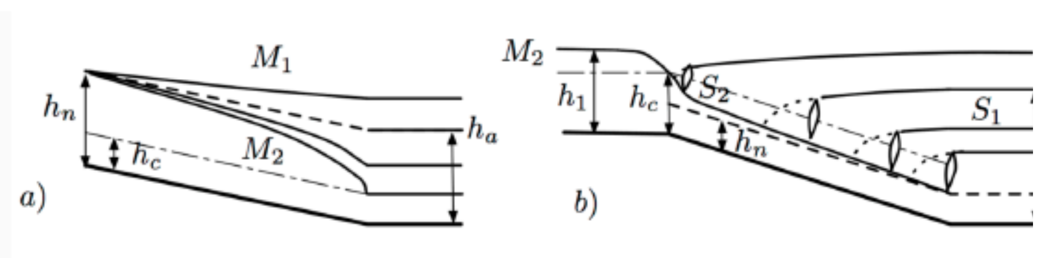
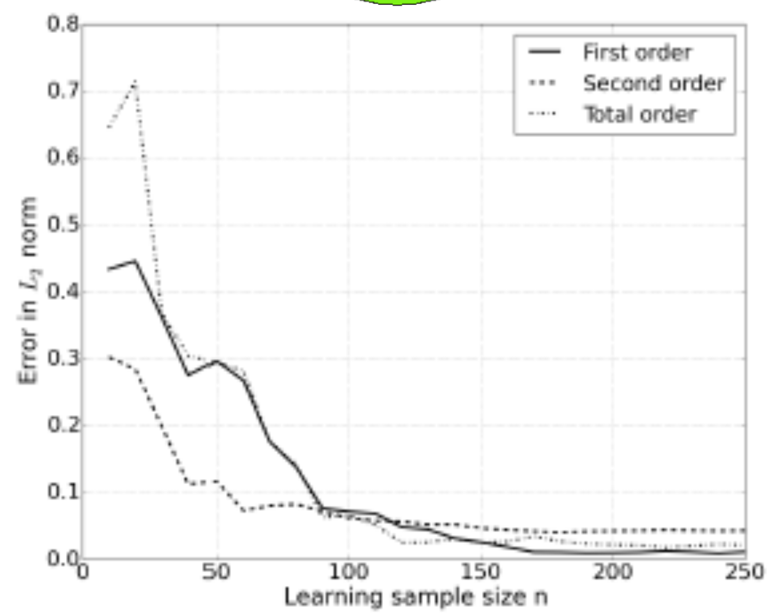


Figure: a) Low slope. b) High slope.



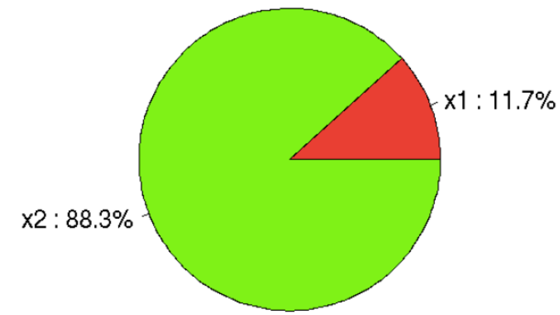
Error function of the learning sample size

A POD Surrogate model of the *LS89* cases for Statistical Analysis
Backwater curves in a channel Flow case

This test case is representative of a real and non-linear case which allows to demonstrate the ability of JPOD to perform the analysis necessary of LS89

- Channel test case with constante slope and width
- POD Surrogate model w.r.t Ks and Q

Q and Ks Sobol's indices



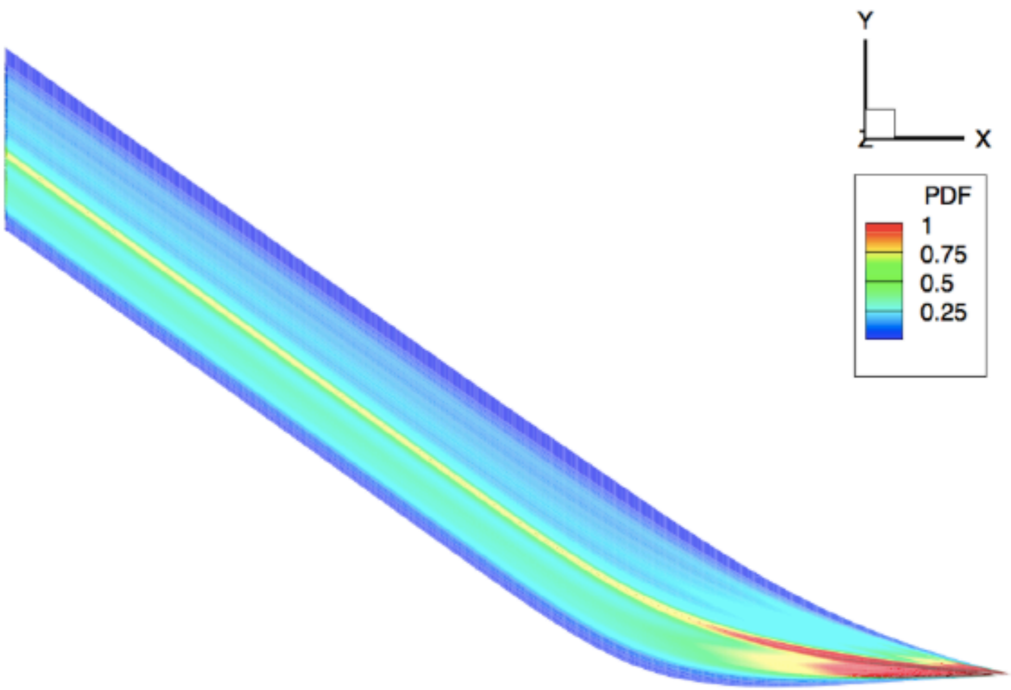
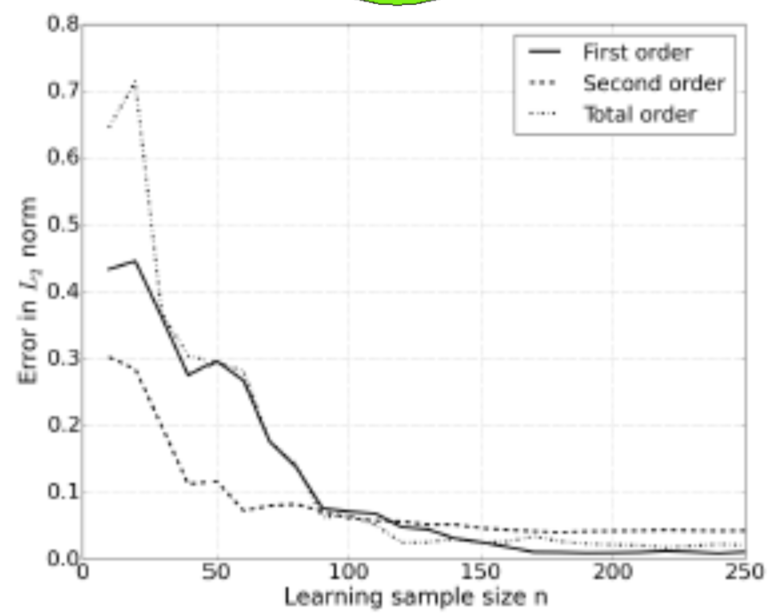


Figure: PDF of the outputs



Error function of the learning sample size

Uncertainty quantification in hydraulics

Why using a Surrogate model ?

Input data /uncertainty sources

- Up and downstream boundary conditions
- River and flood plain geometry
- Hydraulic parameter (friction)
- Initial condition

TELEMAC/MASCARET

www.opentelemac.org

Output QOIs

- Water level
- Discharge

The PC-surrogate strategy is implemented on the test case channel and on the Garonne river 50km reach between Tonneins and La Réole

Sensitivity analysis: reduced-cost computation of statistics, pdfs, quantiles

Data assimilation: reduced cost-computation of background error covariance matrices (stochastic estimate in EnKF *Rochoux et al. 2014*)

In stationnary and non-stationnary conditions with time (PC-surrogate at each point of the domain with common PEX) and space-varying coefficients (spectral representation of time-varying input flow).

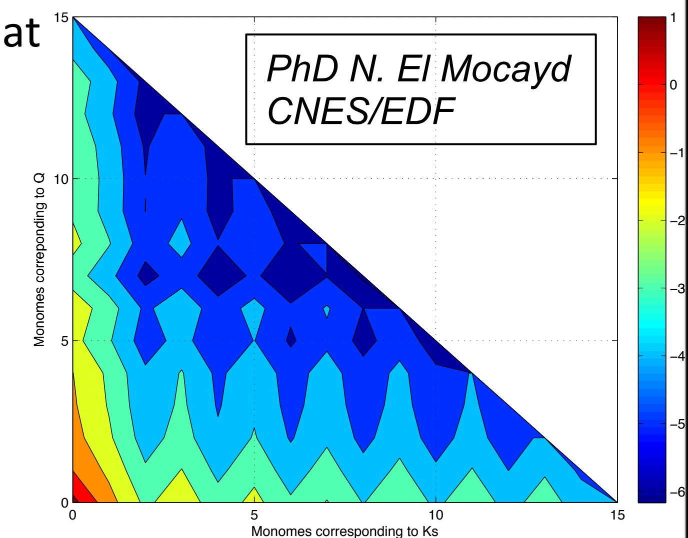
Polynomial Chaos Expansion (PC)

- 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):

$$h(x_k, \epsilon_Q, \epsilon_{Ks}) = \sum a_{i,j}(x_k)\psi_i(\epsilon_{Ks}).\phi_j(\epsilon_Q)$$

- Number of coefficients in the PC expansion for h(Ks,Q) as a function of the polynomial order

P	4	5	6	7	8	9	10	11	12	13	14	15
N_{pc}	15	21	28	36	45	55	66	78	91	105	120	136



Coeff spectrum for h(Ks,Q) PC expansion

PC coefficients estimation: two different methods

- Least square (PC-LS)

$$A = (\Psi\Psi^T)^{-1}(VF\Psi)$$

- Quadrature (PC-Quad)

$$a_j = \frac{<F(X),\psi_j(X)>}{\|\psi_j(X)\|^2}$$

$$< F(x), \psi_j > = \int_{x \in \mathcal{P}} F(x).\psi_j(x) d\mathcal{P}$$

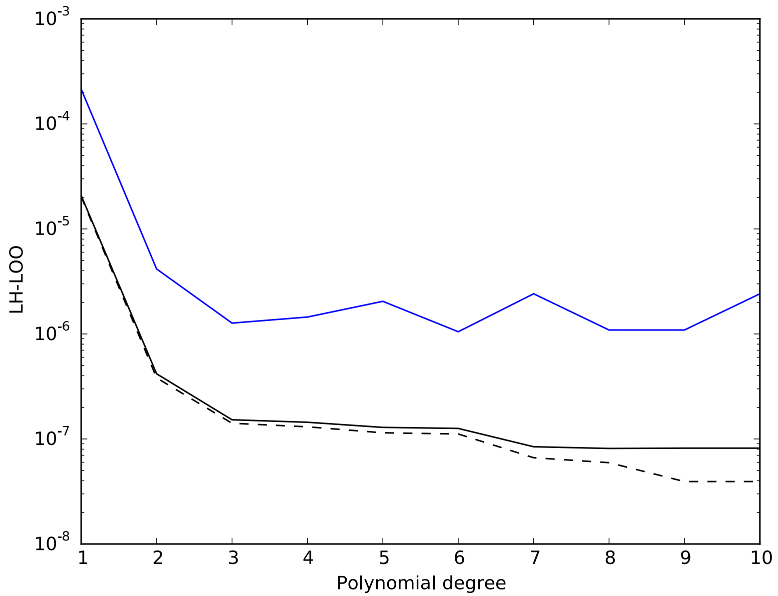
Use OpenTurns for :
- basis definition (OrthogonalProductPolynomialFactory)
- coefficients estimation (Least square strategy, Integration Strategy)
- meta model formulation (functionalChaosAlgorithm)
- post processing (Sobol, pdf, statistical moments)

- Implementation with SALOME-HYDRO using MASCARET as a Python function and with OpenPALM using MASCARET as a external code

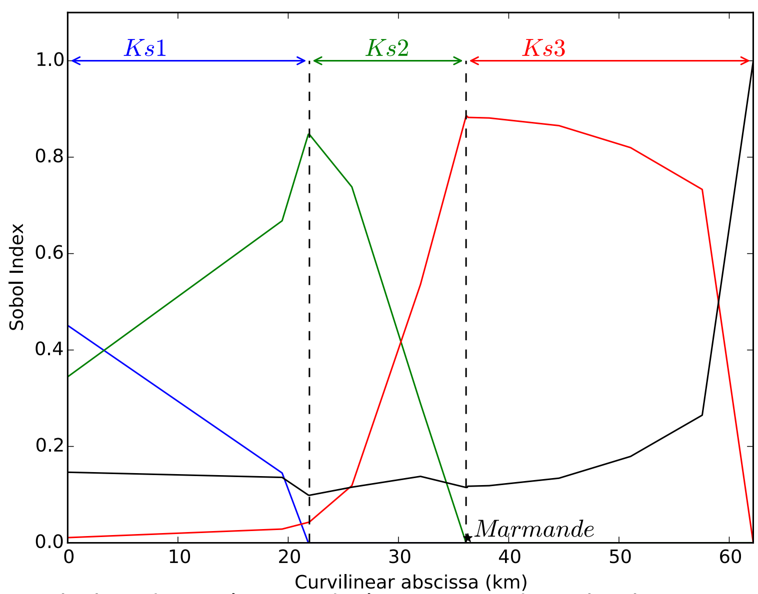
Need to adjust the call of OpenTurns methods.

- When the degree of freedom increases (spatialized Ks), a Least Angle Regression method is implemented with LOO error metrics to avoid resampling of the forward model.

Sensitivty analysis to identify most significant input parameters.



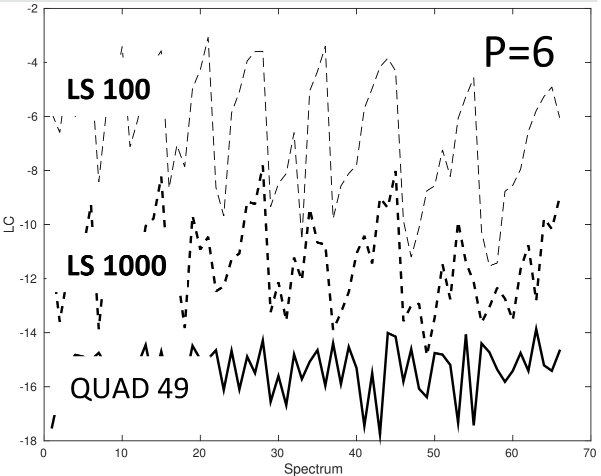
LOO (dashed line) and normalized LH (solid line) at Marmande computed with a LAR surrogate model - thin dashed line with 100 eval - and - thick dashed line with 1000 eval - for a maximum polynomial order P = 1, ..., 10.



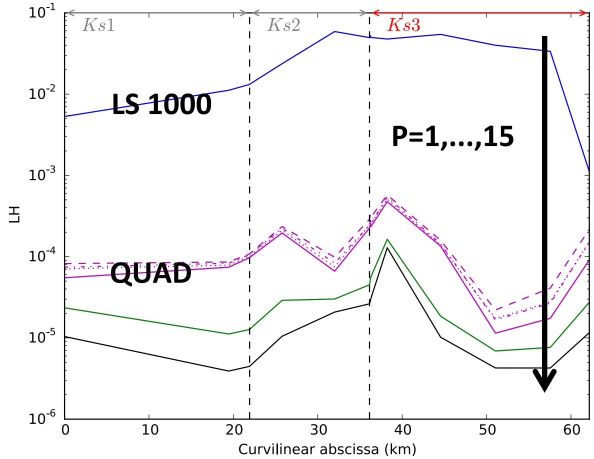
Sobol indices (1st order) computed with the LAR surrogate model w.r.t. Ks1 (blue line), Ks2 (green line), Ks3 (red line) and Q (black line). The vertical dashed lines indicate the limits between the different sections of Ks.

Uncertainty quantification in hydraulics

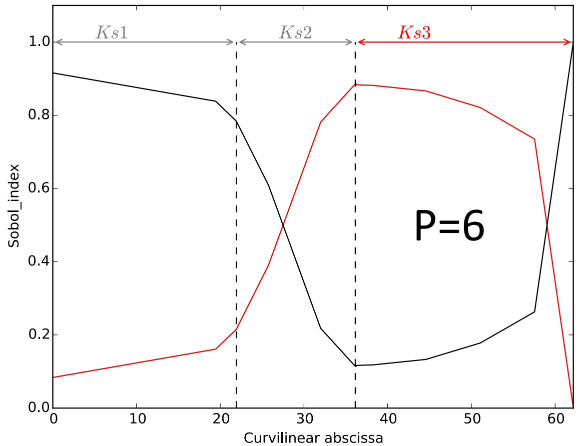
Towards Ensemble-based Data Assimiltion



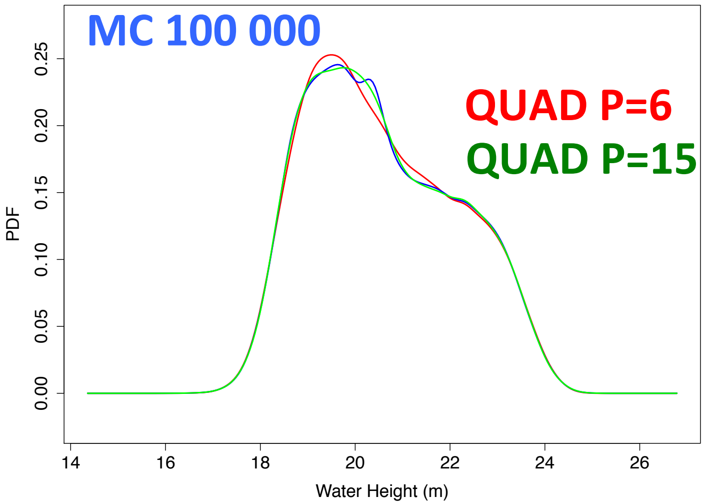
LC for surrogate with LS 100 (thin dashed), LS 1000 (thick dashed), QUAD 49 (thick solid)



LH for LS 1000 and QUAD, P = 1,...,15

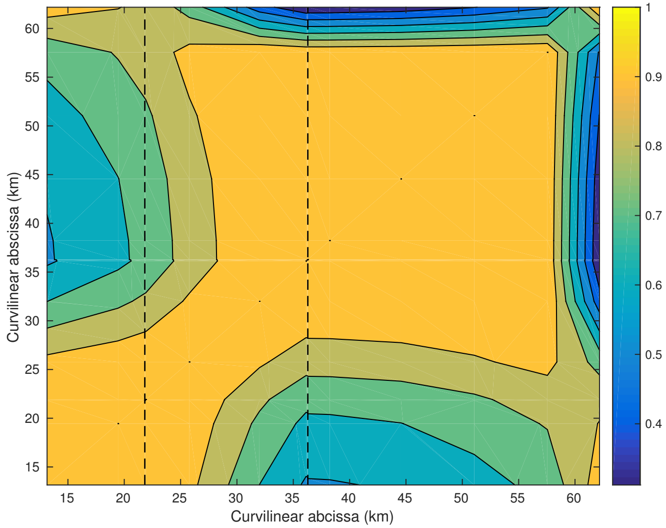


Sobol for QUAD 49 w.r.t. Q, Ks3



Pdf H at Marmande with QUAD 49, QUAD 256

El Mocayd et al., in preparation



<H,H^T> correlation matrix along space with QUAD 49

Uncertainty quantification in hydraulics

Time-varying PC-expansion

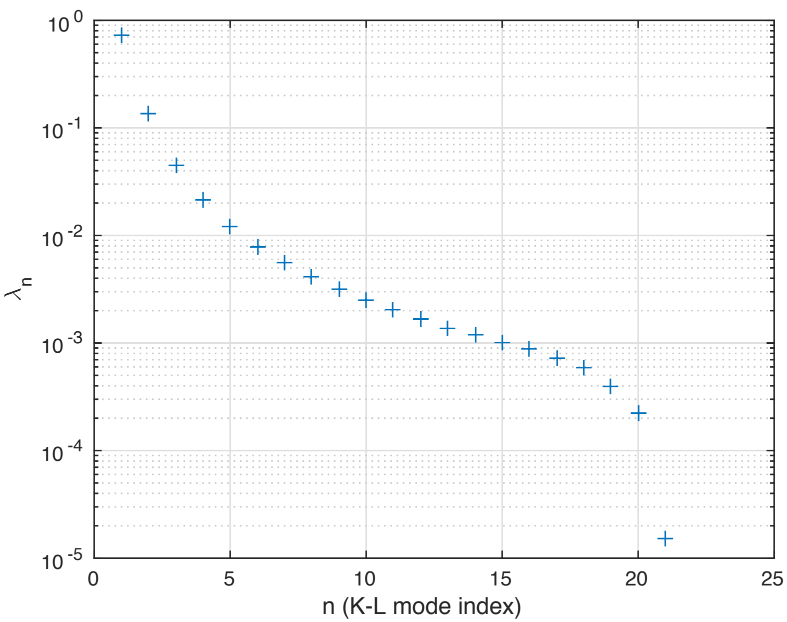
Towards large dimension problems:
Assume the discharge error is decomposed into a non-aleatory variable that is time-dependent (exponential) and a non time-dependent aleatory variable

$$Q(t,\theta) \approx \bar{Q}(t) + \sum_{i=1}^{d} \sqrt{\lambda_i} f_i(t) \zeta_i(\theta)$$

$$\int_{t_2 \in T} \mathcal{C}(t_1,t_2) f_i(t_2) dt_2 = \lambda_i f_i(t_1), t_1 \in T$$

λ_i,$f_i(t)$ are eigen values and functions of time autocorrelation matrix $\mathcal{C}$ prescribed for Q(t)

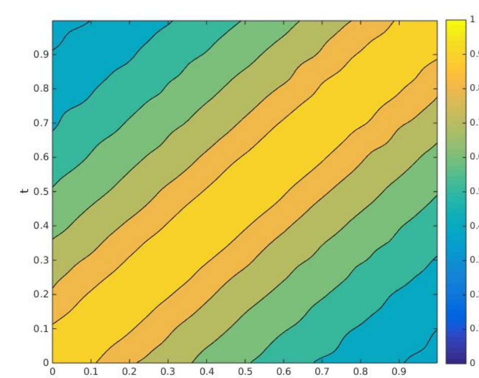
λ_i,$f_i(t)$ are identified from a Lanczos algo in the SVD decomposition coded in Matlab
Persp : Use OpenTurns KarhunenLoeveFactory



Eigen values for the exponential auto-correlation matrix

The DOF of the input var. is reduced to the number of most-significant spectrum values ζ_i with Gaussian statistics (about 10 independent values)

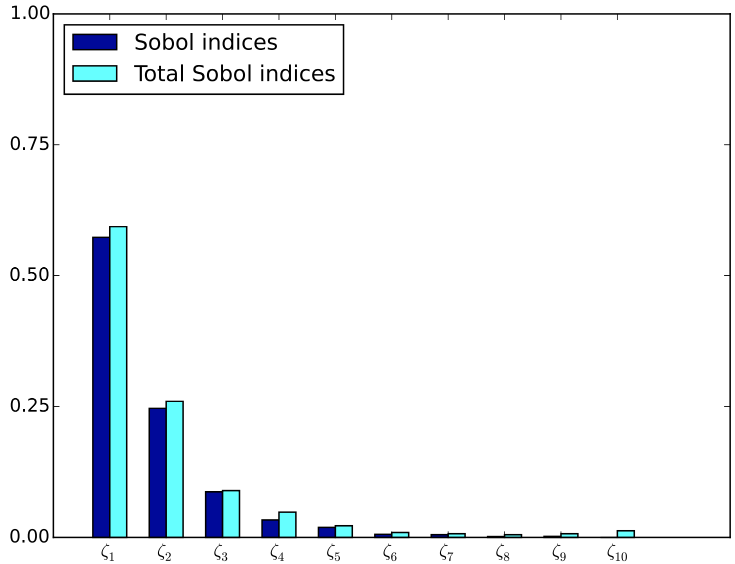
Correlation matrix retrieved from the 10 modes SVD



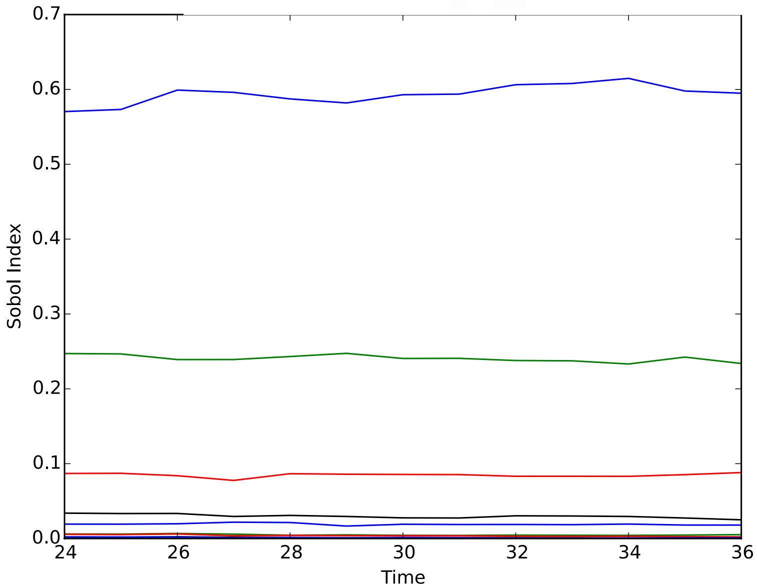
Uncertainty quantification in hydraulics
Time-varying PC-expansion

- The water level is decomposed on a Hermite polynomial basis
- A LAR strategy is implemented w.r.t. Q in non-stationnay conditions on the channel test case

$$h(t,\theta) = \sum_{i=0}^{\infty} a_i(t)\Psi_i(\zeta(\theta))$$



Sobol indices for most significant modes at 1st time step at the middle of the channel



Time varying Sobol indices at the middle of the channel

Data Assimilation strategy for ensemble-based algorithm

Define the control vector with the most-significant modes ζ_i for non-stationny flood events on the Garonne River

On going:
P. Roy M2 internship and PhD (Oct. 2016)
N. El Mocayd PhD (2013-2016)

- Different reduced-model methods are implemented with non-intrusive strategy using OpenTurns
- Small to large dimension problems with parcimonous to expensive computational codes
  - towards LES application in CFD
  - towards 1D/2D/3D model in operational context in hydrodynamics
- Real case applications in CFD and hydrodynamics
- Collaboration with EDF/MRI + EDF/LNHE
- Potential use of new developments in OpenTurns
- Potential inputs to OpenTurns are possible (LAR, SVD, re-sampling...in Python)

Key challenge:
Implement efficient strategy for UQ and DA real applications with computationally expensive codes and high dimension problems