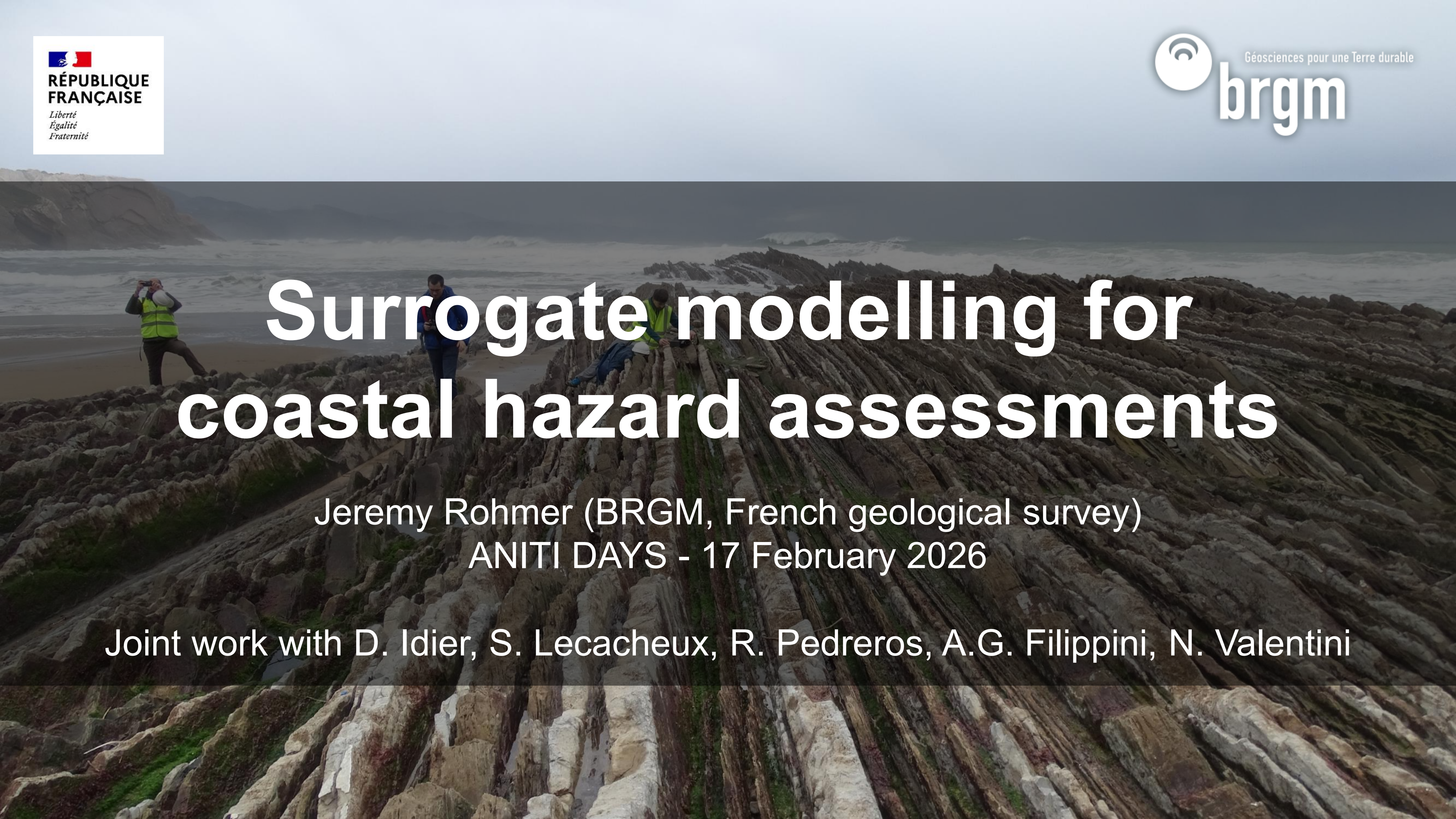
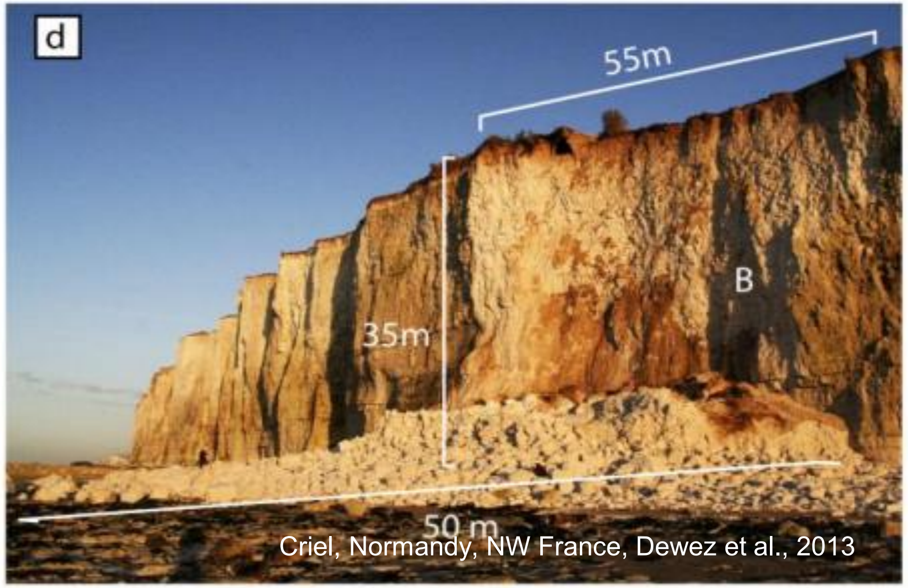


Surrogate modelling for coastal hazard assessments

Jeremy Rohmer (BRGM, French geological survey)
ANITI DAYS - 17 February 2026

Joint work with D. Idier, S. Lecacheux, R. Pedreros, A.G. Filippini, N. Valentini

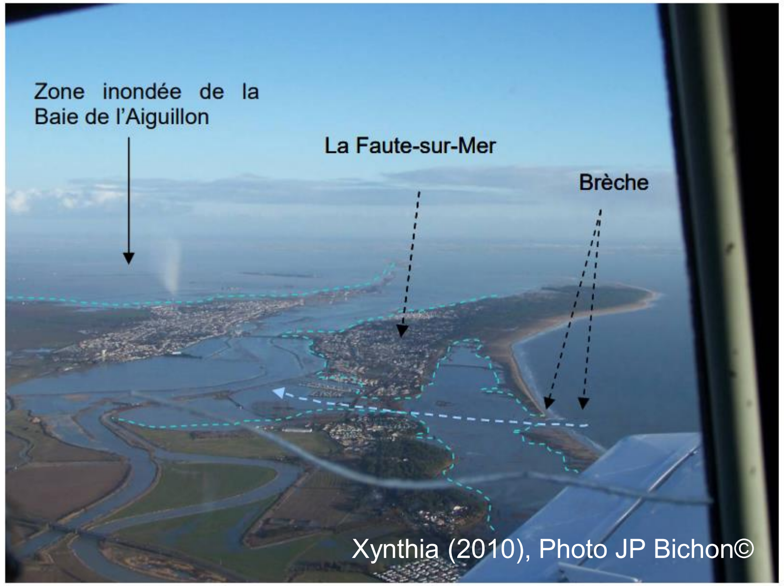
Coastal hazards



Criel, Normandy, NW France, Dewez et al., 2013


Xynthia (2010), La Tranche s. Mer Photo M Garcin©


La Reunion Island, Lecacheux et al., 2016, SPICY http://spicy.brgm.fr/



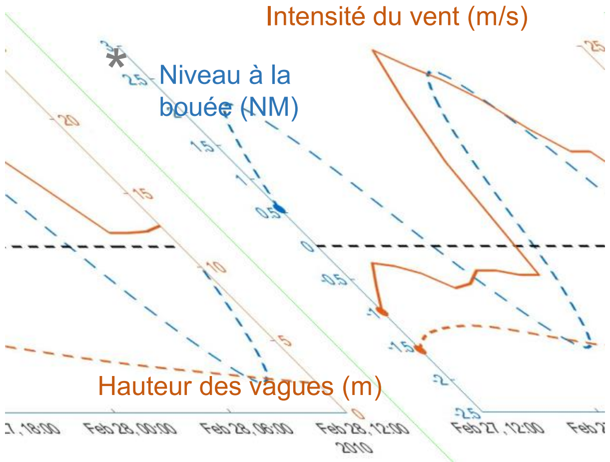
Xynthia (2010), Photo JP Bichon©

The key role of computer codes

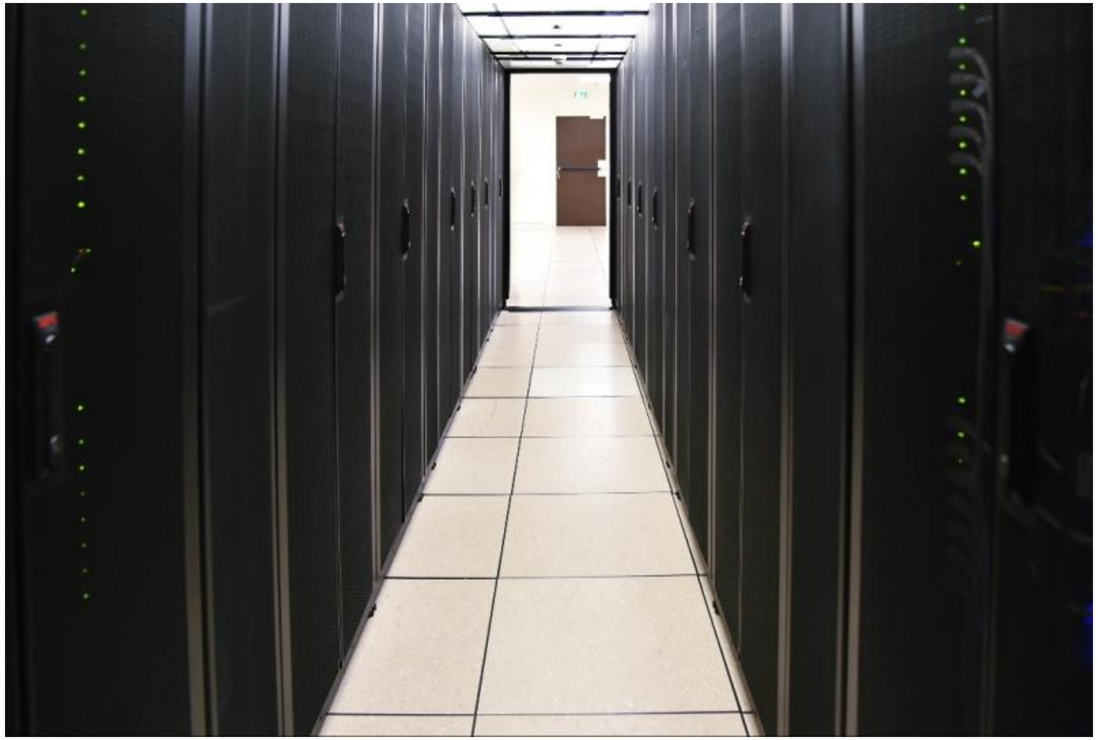
Consequences of Xynthia storm at Arcachon

Computation time cost ~ 1 day even using a HPC architecture

A. Filippini, R. Pedreros, S. Lecacheux



A context of ‘small data’… due to high computation time cost

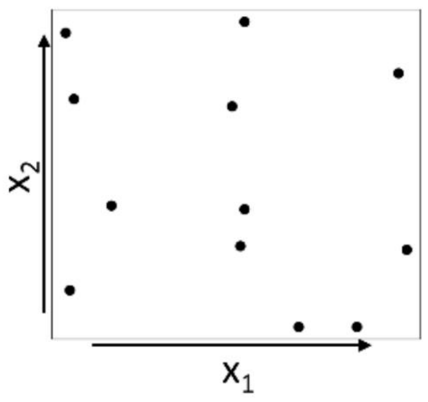


SAPHIR HPC archotecture at BRGM

- >2800 CPUs
- **If all of them were used, computation would last** <1h30
- What about other sites, other storms?

Idea: use results of high fidelity, high resolution numerical simulations <u>already done</u> to train a machine learning model

Design of Experiments (DoE)
Points to perform simulations

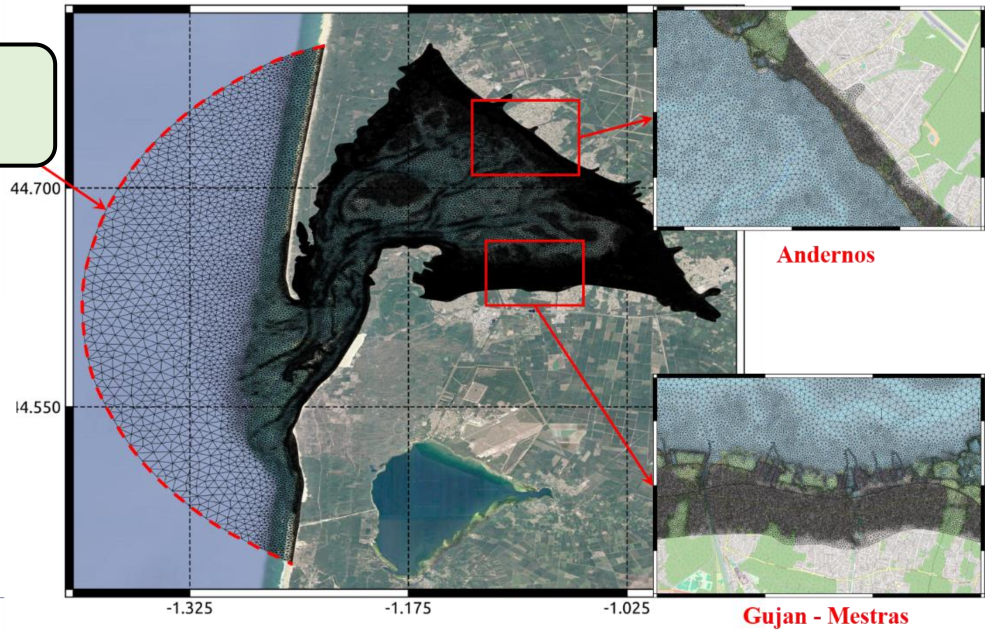


X1, X2 :
storm conditions

Surge (S)

Waves :
- **Height (Hs)**
- **Period (Tp)**
- **Direction (Dp)**

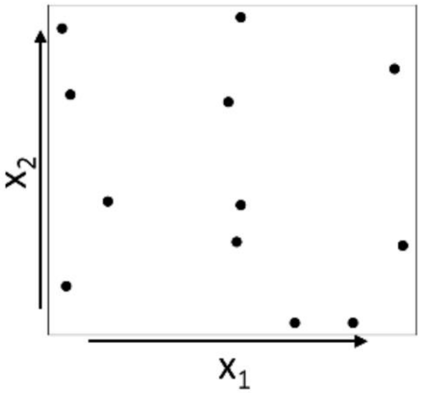
Wind :
- **Speed (U)**
- **Direction (Du)**



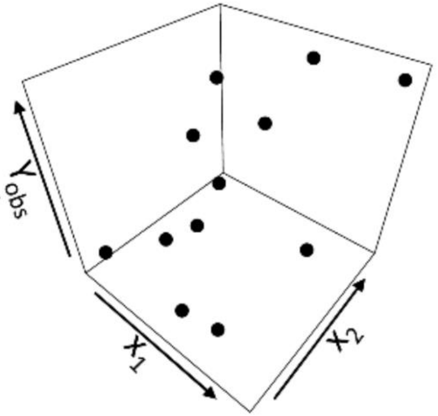
Computation time cost !!

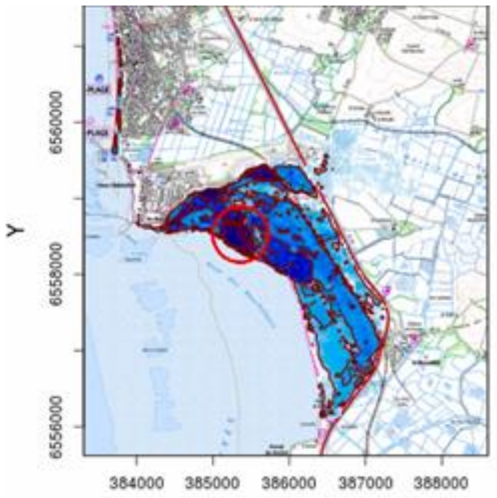
Design of Experiments (DoE)
Points to perform simulations

Simulation
Performing simulations for the points of the DoE



Computer code



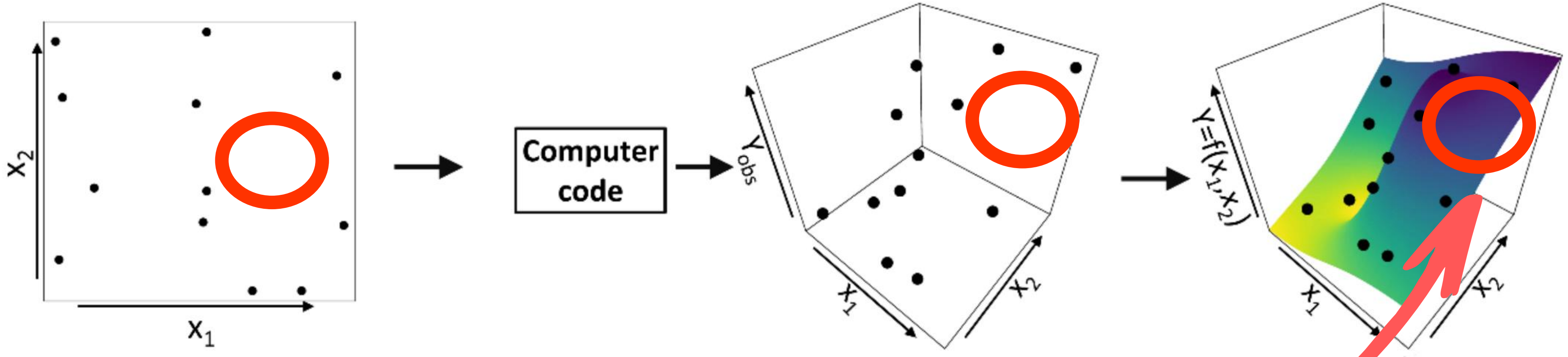


Y =
Local risk indicator e.g. flooded area, maps of water depths, spatiotemporal field of water velocity, etc.

Design of Experiments (DoE)
Points to perform simulations

Simulation
Performing simulations for the points of the DoE

Metamodeling
Approximation of the computer code over the Xi space



Fill the gap with a machine learning (=surrogate model)

A flexible surrogate model: Gaussian process (Gp) regression

[Williams and Rasmussen, 2006]

- **Objective:** predict Y_* for new storm conditions (inputs) x_* **without** running the computer code

- **Assumptions**
  - distribution of Y_* depends on n simulation results $\boldsymbol{Y}_{\mathrm{n}} = \boldsymbol{y}$
  - distribution of $Y_* | \{\boldsymbol{Y}_{\mathrm{n}} = \boldsymbol{y}\}$ is Gaussian

Uncertainty quantification

With parameters (simple kriging equations)

- Mean $m = {k^*}^{\mathrm{t}} . \mathbf{K}^{-1} \boldsymbol{y}$
- Variance $\sigma^2 = k(x_*, x_*) - {k^*}^{\mathrm{t}} . \mathbf{K}^{-1} . k^*$

With

Covariance matrix $\boldsymbol{K} = \left(k(x_{\mathrm{i}}, x_{\mathrm{j}})\right)_{1 \leq \mathrm{i,j} \leq \mathrm{n}}$

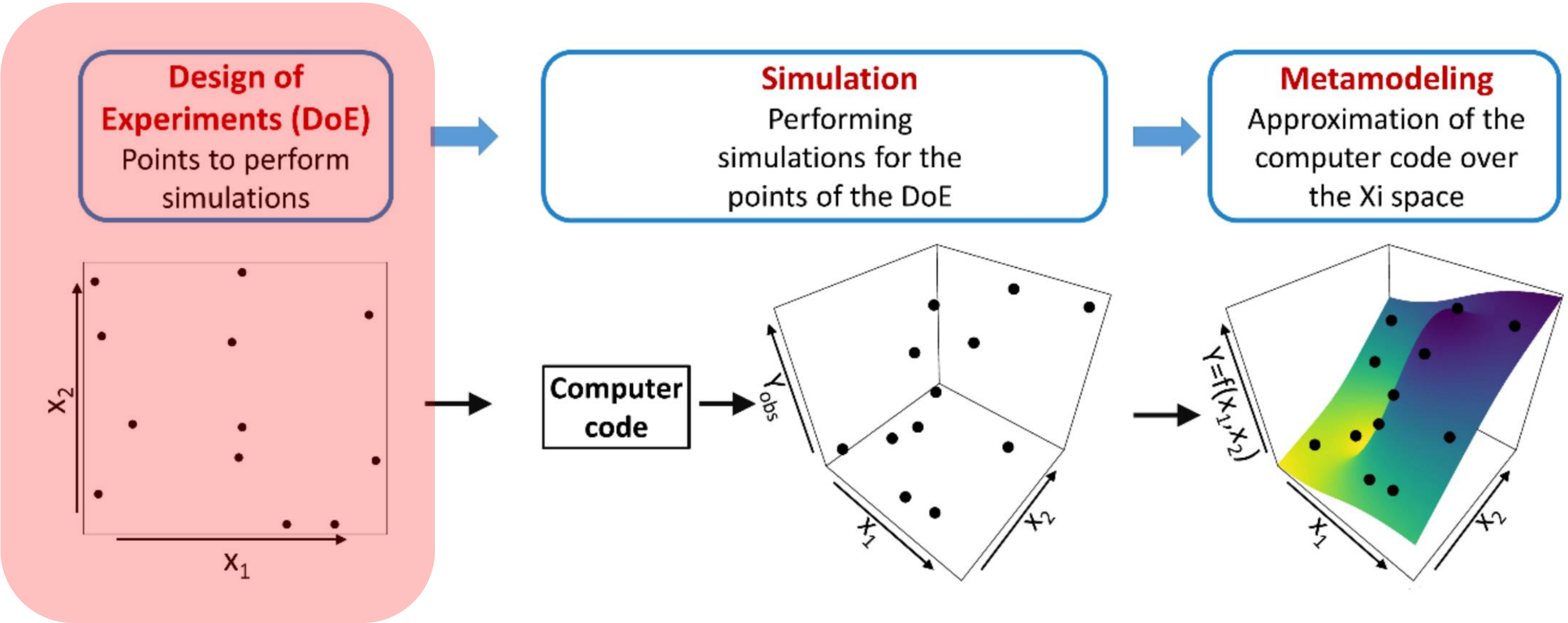
Flexibility via kernels

Covariance vector $k^* = (k(x_*, x_1), k(x_*, x_2), \dots, k(x_*, x_{\mathrm{n}}))$

Géosciences pour une Terre durable

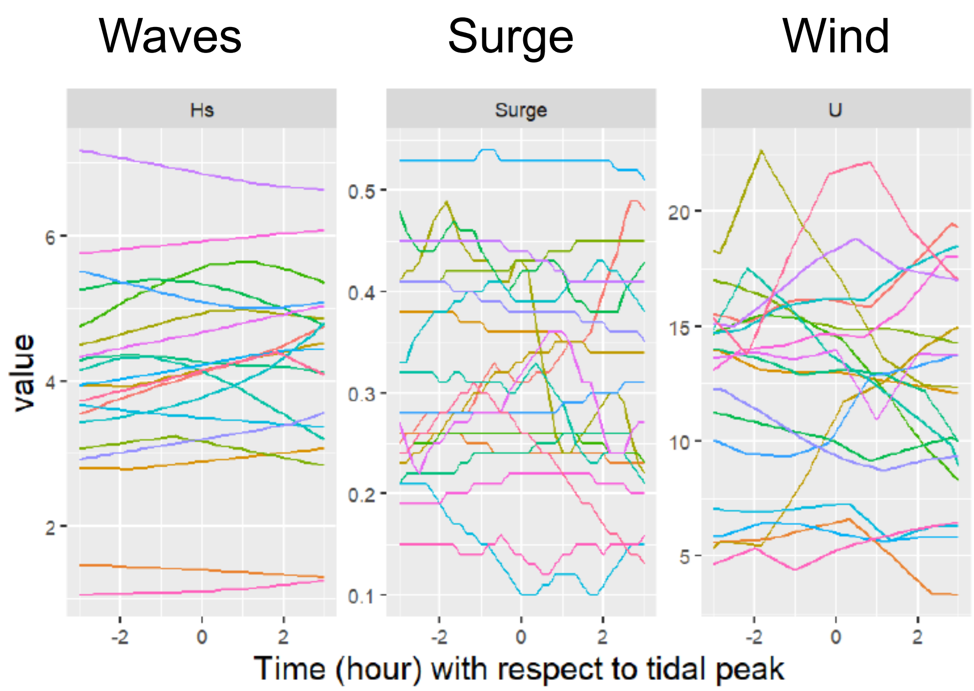
brgm

Key challenges for coastal hazards



The inputs are
extreme time series

The inputs are extreme time series

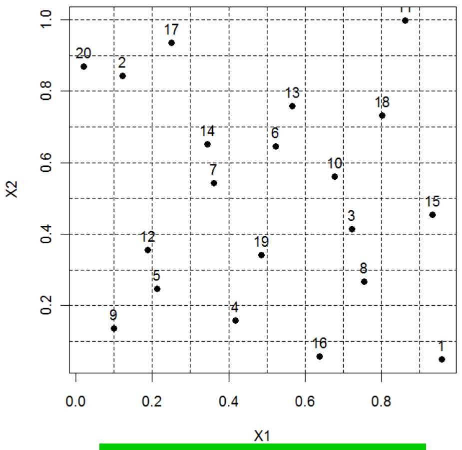


Storm Goretti, 9/1/26 - T. CREUX / OUEST-FRANCE

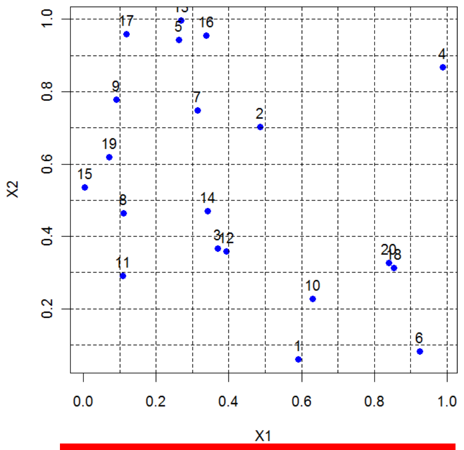
Storm Nils, close to Ajaccio (Corse-du-Sud), 12/2/26 - PASCAL POCHARD-CASABIANCA/AFP

Design Of Experiments (DOE)

Objectives of 'classical' DOE
- Select a limited number of inputs' configurations
- Spanning a large range of values
- With space-infilling property



optimised



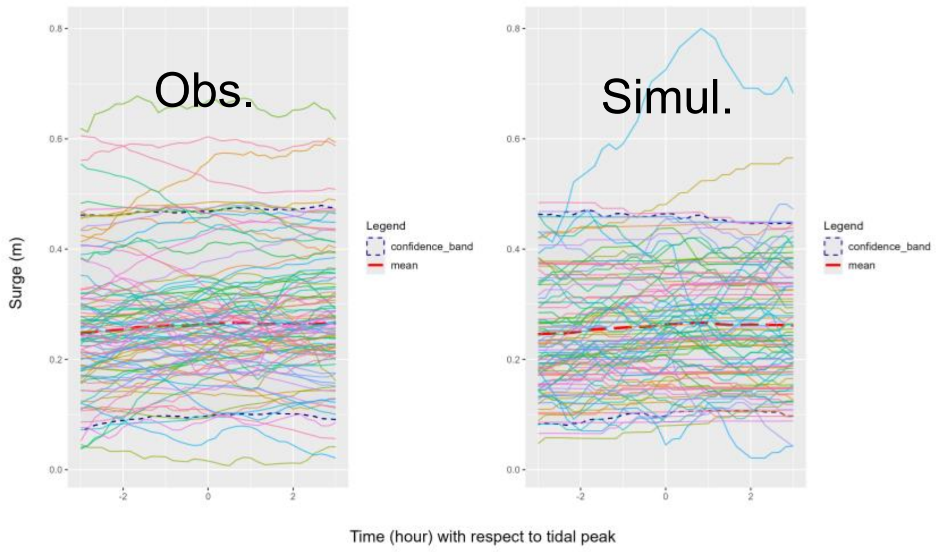
crude approach

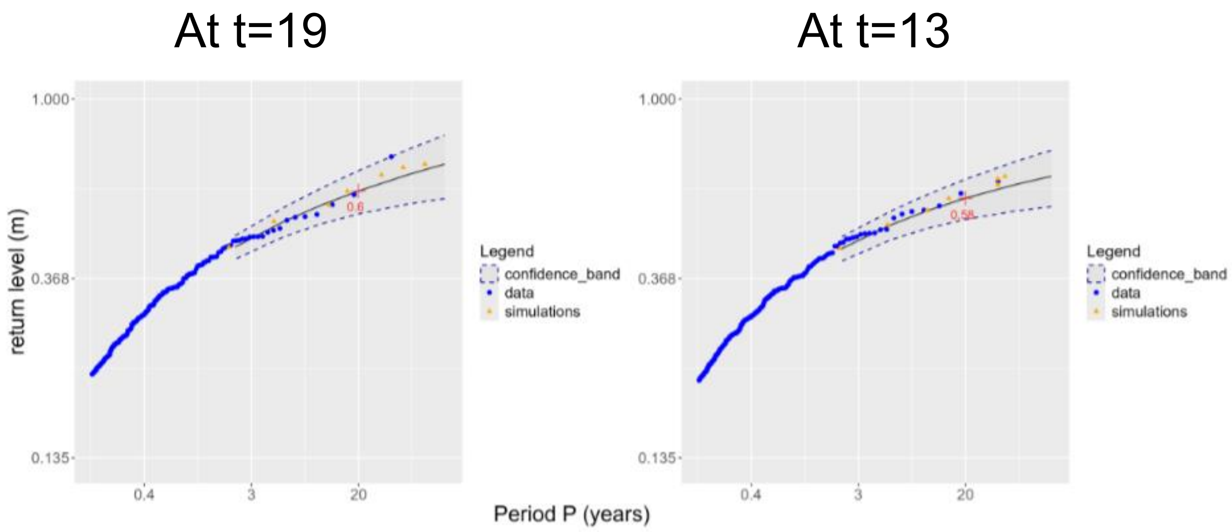
DOE and extreme time series

PhD thesis of Nathan Gorse (2024+)

Developments of an approach combining
- Stochastic simulation
- Multivariate extreme value theory
- Angular-radial decomposition
- ML-based validity check

See Nathan's poster this afternoon!





Stochastic simulation for extreme surge time series (Gorse et al., in review)

ANITI

PiLearnWater

Physics-Informed Learning Methodsfor Continental Waters and Marine Risks

Olivier Roustant – Jérôme Monier

Surrogate modelling and time series

Risk-based System For Coastal Flooding Early Warning

Collaboration IMT-BRGM

PhD thesis of Jose Betancourt (2017-2020)

Adapt standard kernels (e.g., Gaussian, Matérn) using a **dissimilarity measure adapted to** *df* functions $\boldsymbol{f}, \tilde{\boldsymbol{f}}$

(**Opt. 1**) L2 norm for functions

$$\|\boldsymbol{f} - \tilde{\boldsymbol{f}}\|_{F,\boldsymbol{\theta_f}} := \sqrt{\sum_{k=1}^{df} \frac{\int_{T_k} \left(f^{(k)}(t) - \tilde{f}^{(k)}(t)\right)^2 dt}{\left(\theta_f^{(k)}\right)^2}},$$

Length scales

(**Opt. 2**) Projection onto a new basis *B* of lower dimension p_k

$$\left\|\Pi\left(\boldsymbol{f}\right) - \Pi\left(\tilde{\boldsymbol{f}}\right)\right\|_{D,\boldsymbol{\theta_f}} := \sqrt{\sum_{k=1}^{df} \frac{\int_{T_k} \left(\sum_{r=1}^{p_k} \left(\alpha_r^{(k)} - \tilde{\alpha}_r^{(k)}\right) B_r^{(k)}(t)\right)^2 dt}{\left(\theta_f^{(k)}\right)^2}}.$$

Length scales

Géosciences pour une Terre durable

brgm

Surrogate modelling and time series

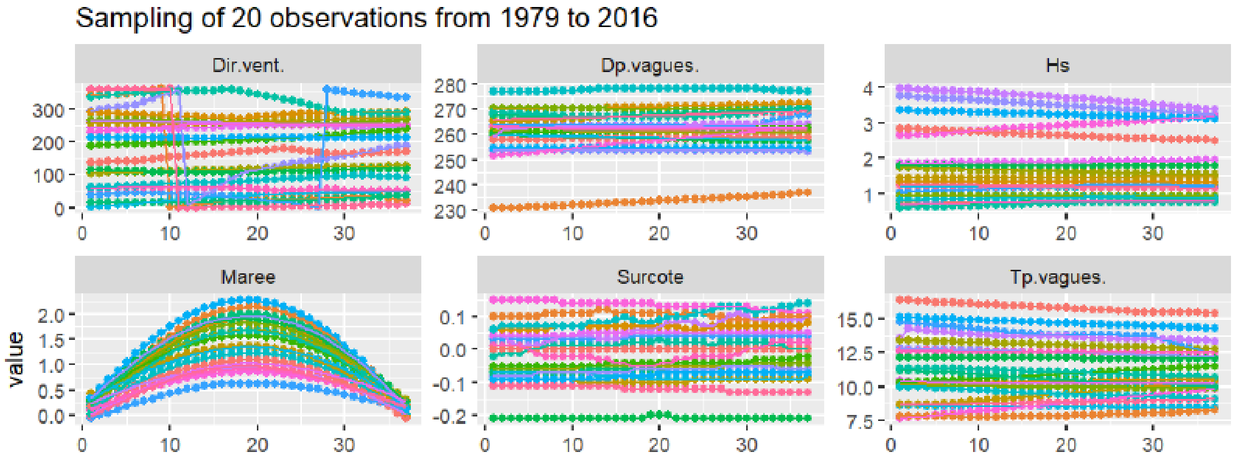
RISCOPE Risk-based System For Coastal Flooding Early Warning

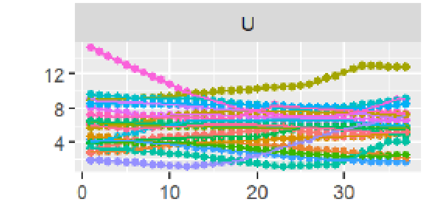
Collaboration IMT-BRGM

PhD thesis of Jose Betancourt (2017-2020)

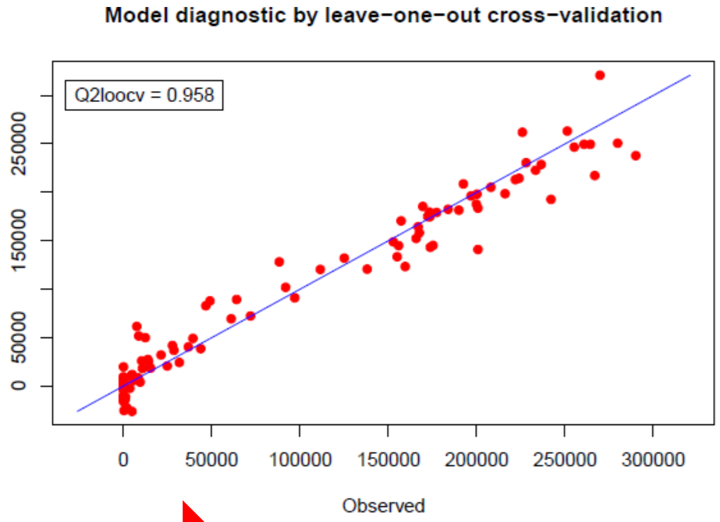
Application of Gp with adapted kernel at Gavres (French Brittany)

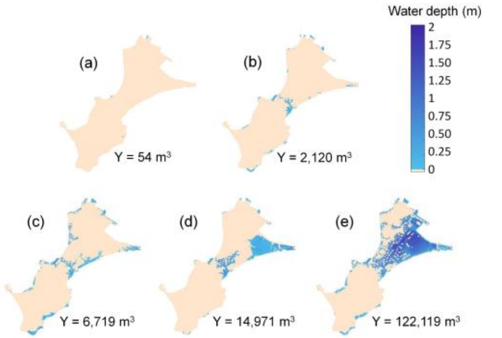
Inputs





Output: Flooded area (m²)

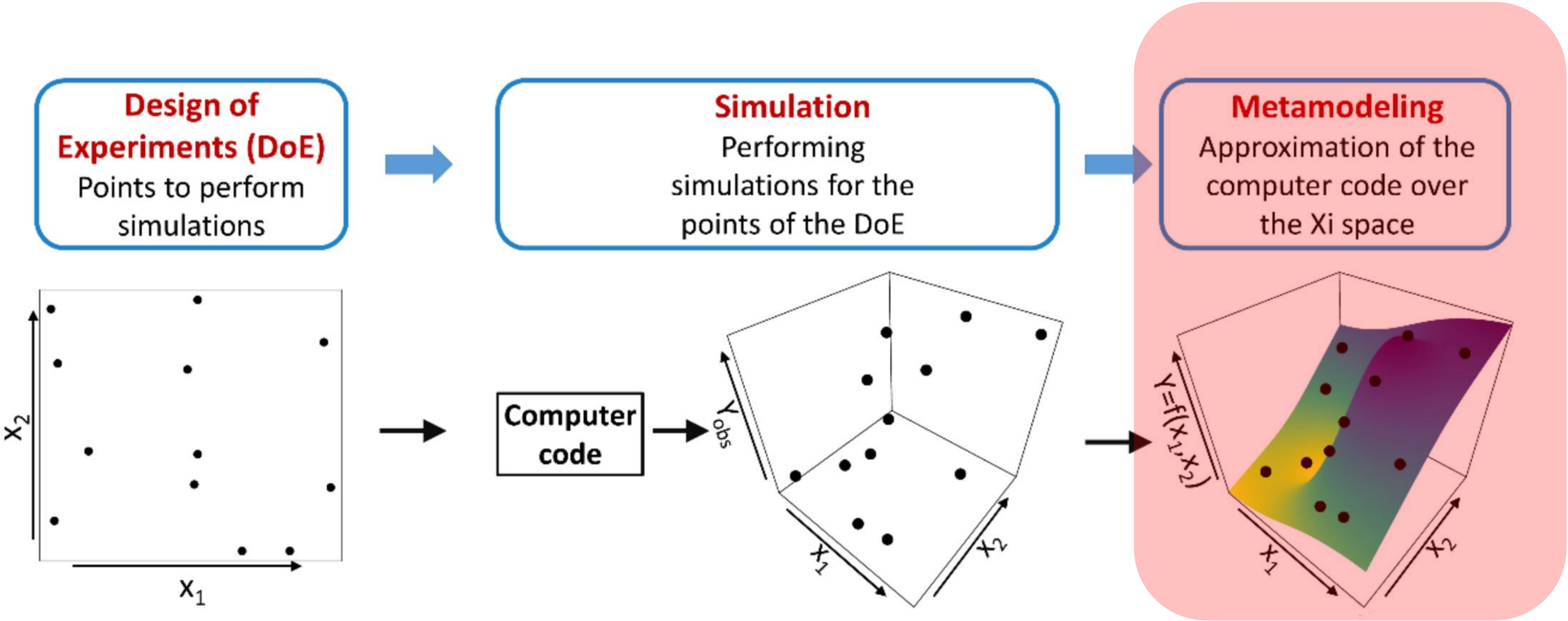




Most optimal model includes temporal effects (F)

	X1	X2	X3	X4	X5	X6	X7	X8	F1	F2	F3	F4	F5	F6	F7	Kern	Q2
1	x	x	x		x	x		x	x	x	x			x	x	gauss	0.970
2	x	x	x		x	x		x	x	x	x			x	x	gauss	0.968
3	x	x	x		x	x		x			x			x	x	gauss	0.968
4	x	x	x		x	x		x	x	x	x			x	x	gauss	0.968
5	x	x	x		x	x		x		x	x			x	x	gauss	0.966

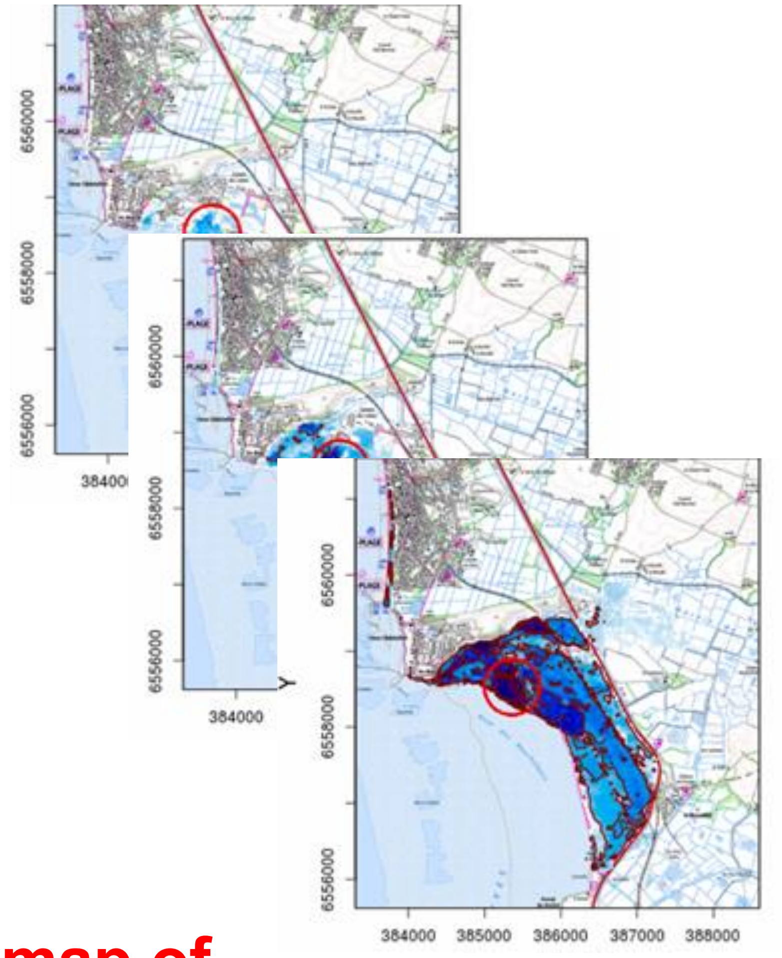
Key challenges for coastal hazards



The outputs are high dimensional

Forcing conditions

Flood map (max water depth)



$$\mathbf{x}^{(1)} = (S^{(1)}, T^{(1)}, t_0^{(1)}, t_+^{(1)}, t_-^{(1)})$$ → Computer code →

$$\mathbf{x}^{(2)} = (S^{(2)}, T^{(2)}, t_0^{(2)}, t_+^{(2)}, t_-^{(2)})$$ → Computer code →

$$\mathbf{x}^{(3)} = (S^{(3)}, T^{(3)}, t_0^{(3)}, t_+^{(3)}, t_-^{(3)})$$ → Computer code →

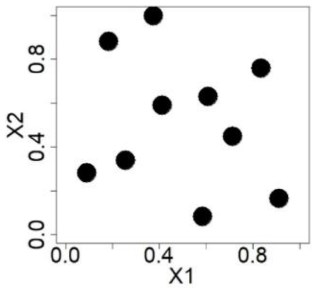
Etc…

Flood map of >10,000 of pixels

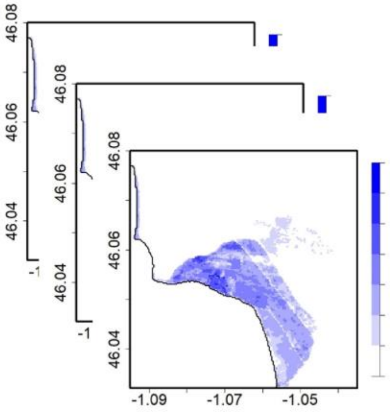
A combined Dimension Reduction – Surrogate Modelling approach

PhD thesis of Elodie Perrin – collaboration Mines St Etienne – CCR - BRGM

Step 1. Design of experiments



Step 2. Numerical simulation



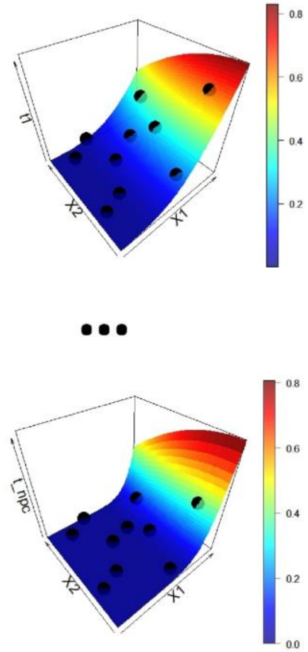
Set of **High dimensional** flood maps of size N

Step 3a. Dimension reduction

$$\begin{pmatrix} z_1 \\ z_2 \\ \dots \\ z_{n_{pc}} \end{pmatrix}$$

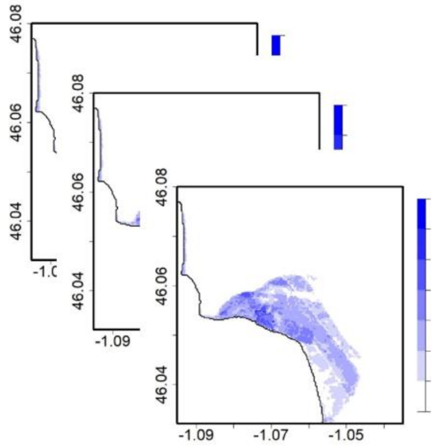
Set of latent variables of size $n_{pc} \ll N$

Step 3b. Metamodelling



Set of Surrogate models, one for each of the n_{pc} latent variables

Step 3c. Reconstruction



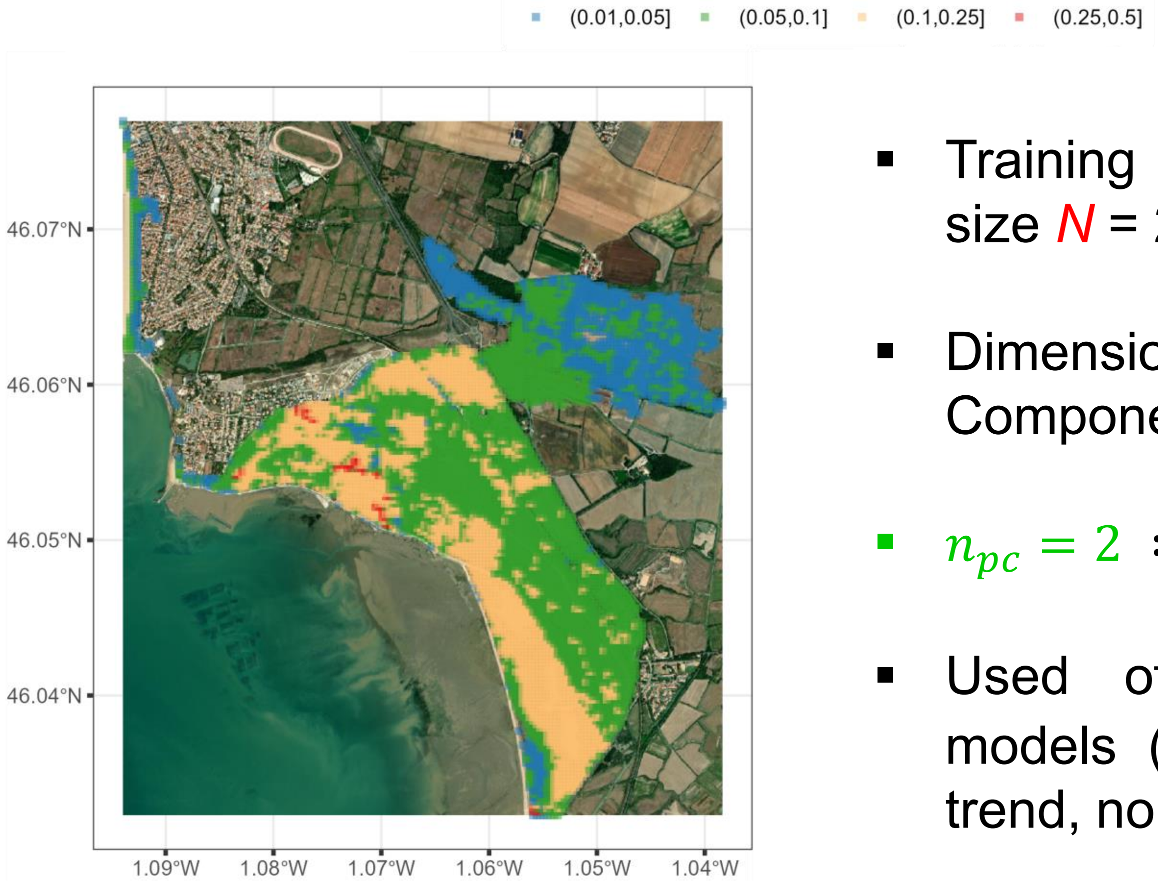
For prediction

Géosciences pour une Terre durable

brgm

Application at Boucholeurs [1]

5-fold cross validation **prediction error** (RMSE)



- Training dataset of 200 flood maps of size N = 256 x 256

- Dimension reduction using Principal Component Analysis

- $n_{pc} = 2$: 95% of explained variance

- Used of $n_{pc} = 2$ independent Gp models (Matérn 5/2 covariance, linear trend, no nugget)

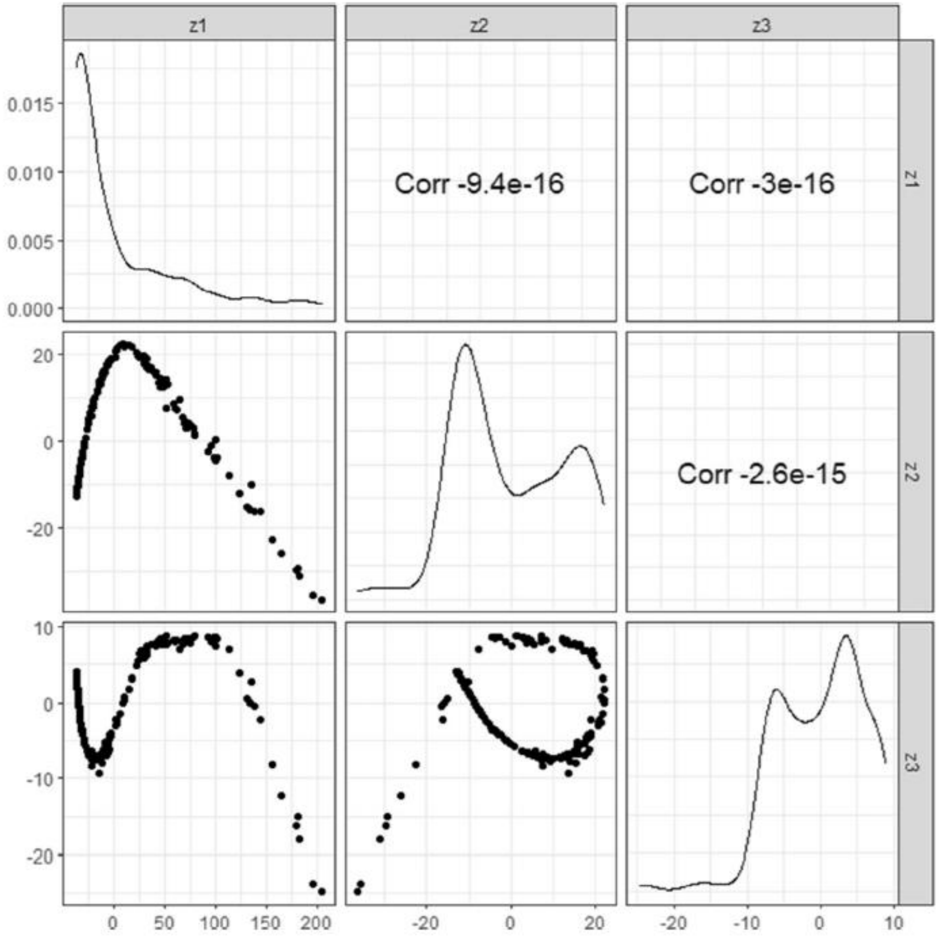
[1] Adapted from Perrin, PhD thesis (2021)

Géosciences pour une Terre durable

brgm

Improvement #1 [ROH23]

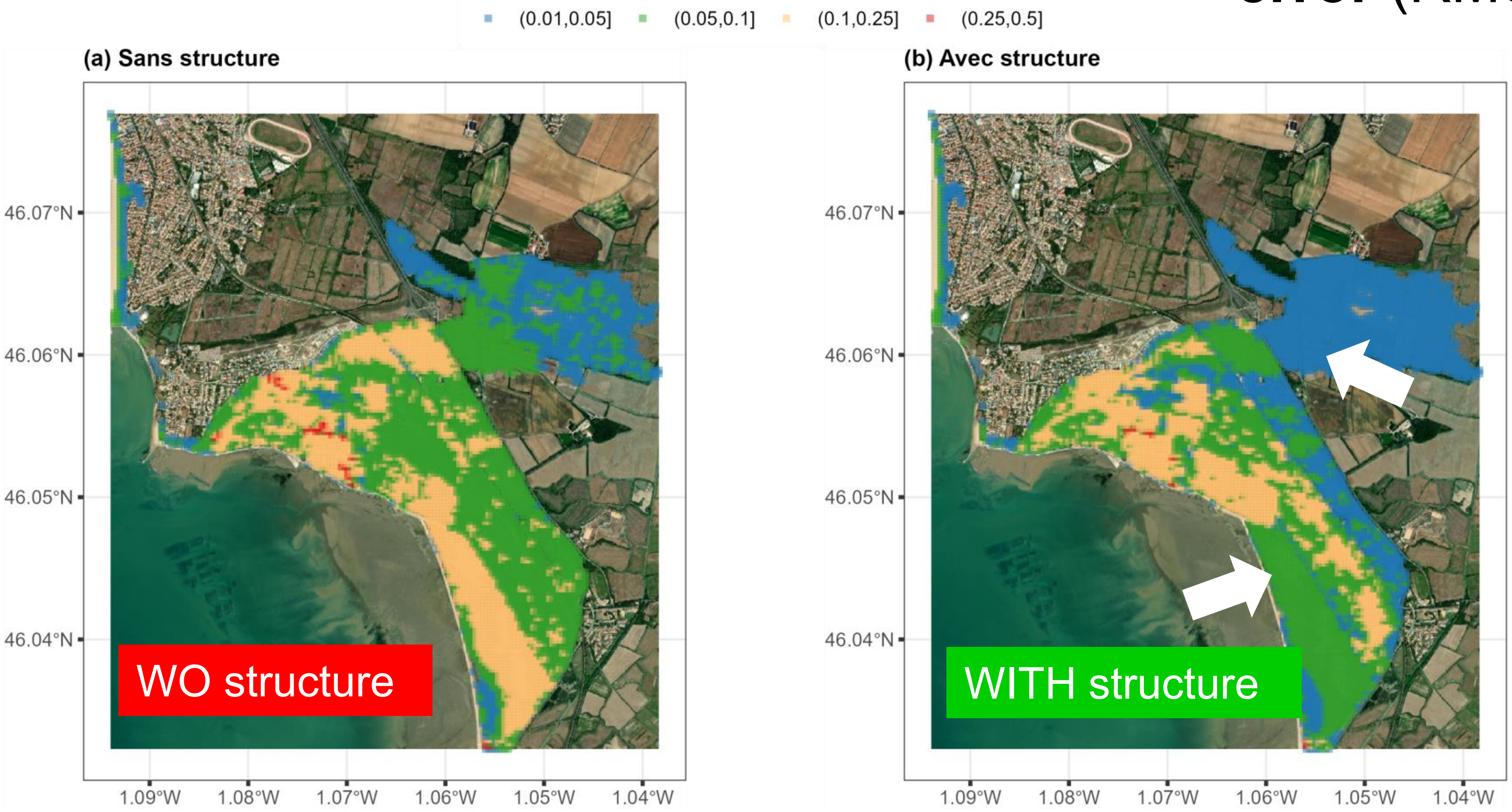
Dependence structure of the latent variables
after applying PCA to the 200 flood maps



[ROH23] Rohmer et al., SERRA, (2023)

Improvement #1 [ROH23]

5-fold cross validation **prediction error** (RMSE)



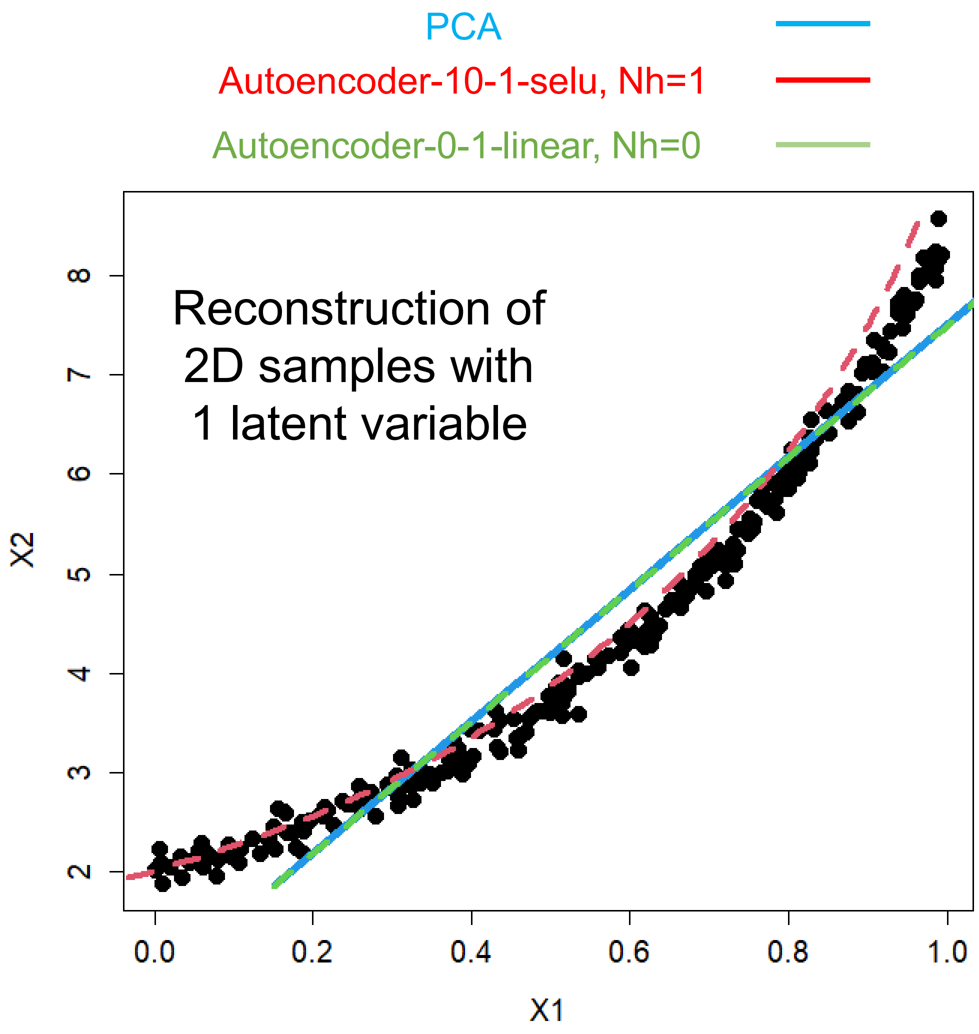
Approach via **independent** Gps

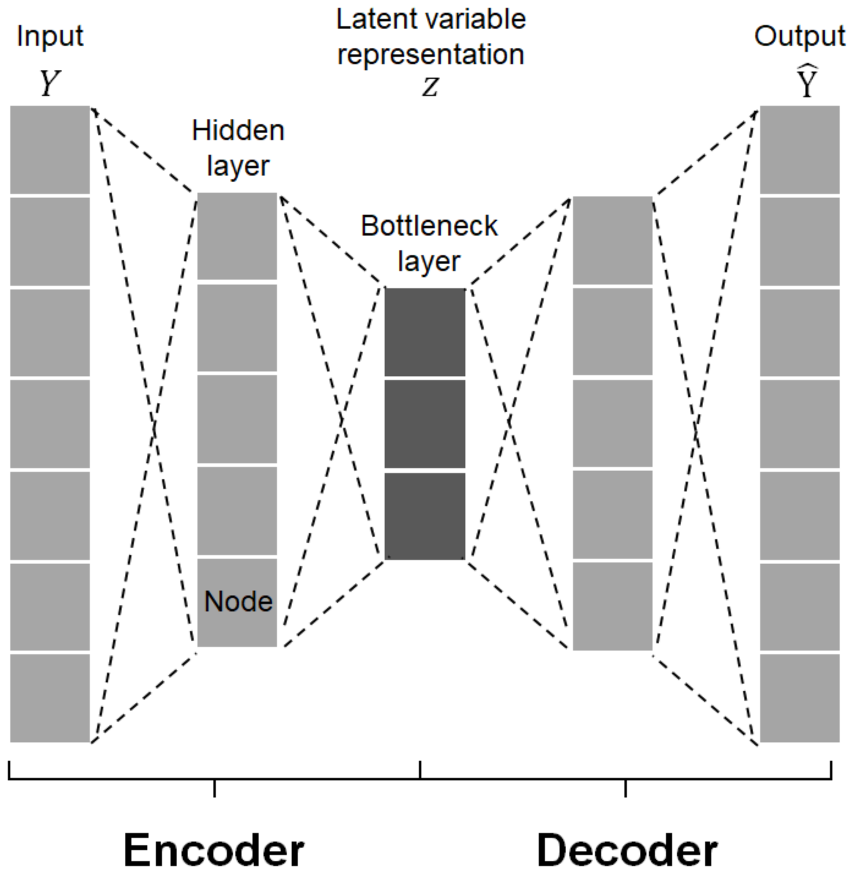
Approach via **Multi-output** MoGp

[ROH23] Rohmer et al., SERRA, (2023)

Géosciences pour une Terre durable
brgm

Improvement #2 [ROH24]





[ROH24] Rohmer et al., Symhydro, (2024)

Géosciences pour une Terre durable

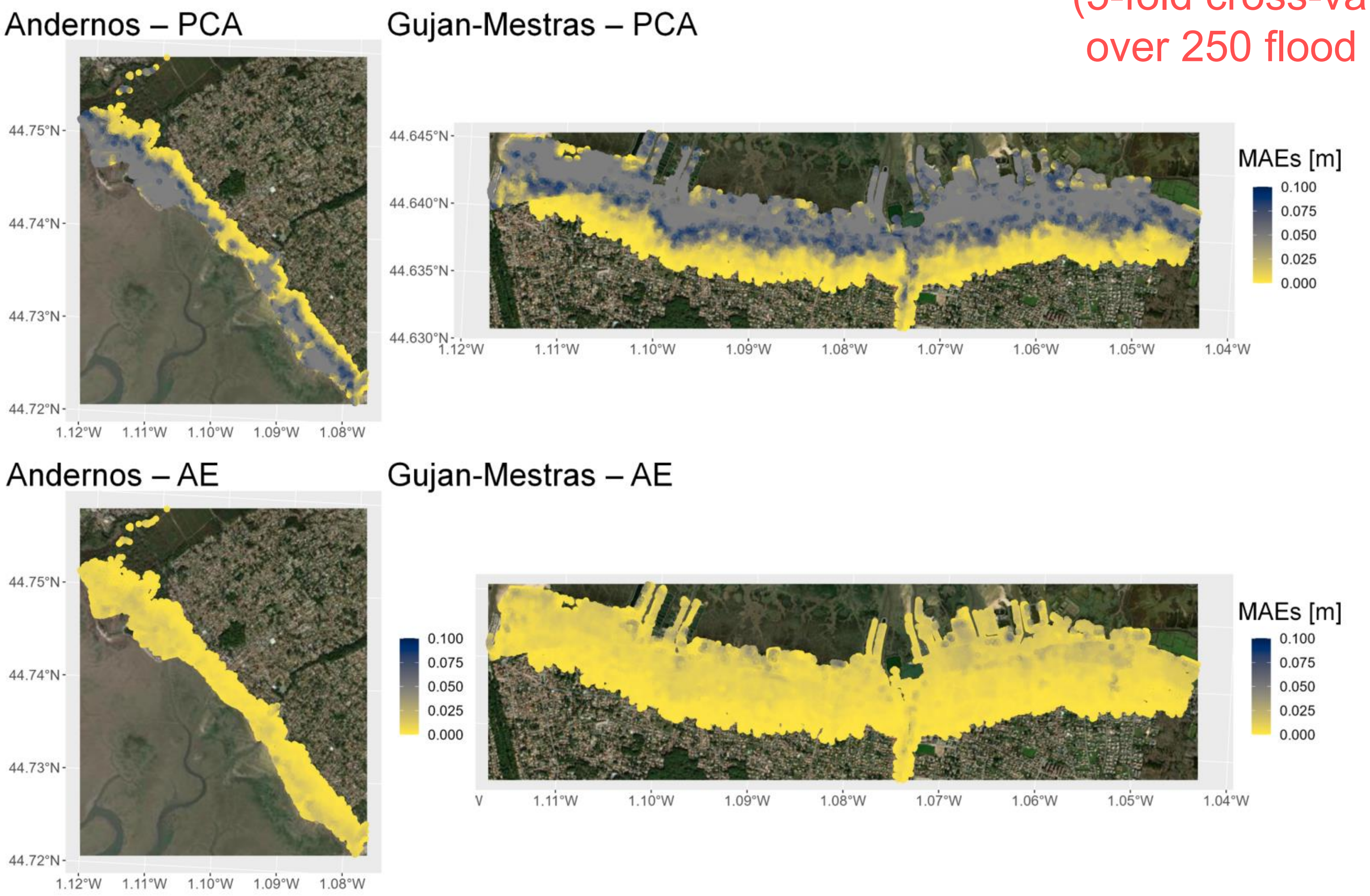
brgm

Improvement #2 [ROH24]

Mean absolute reconstructin error
(5-fold cross-validation
over 250 flood maps)

PCA
99%
explained
variance

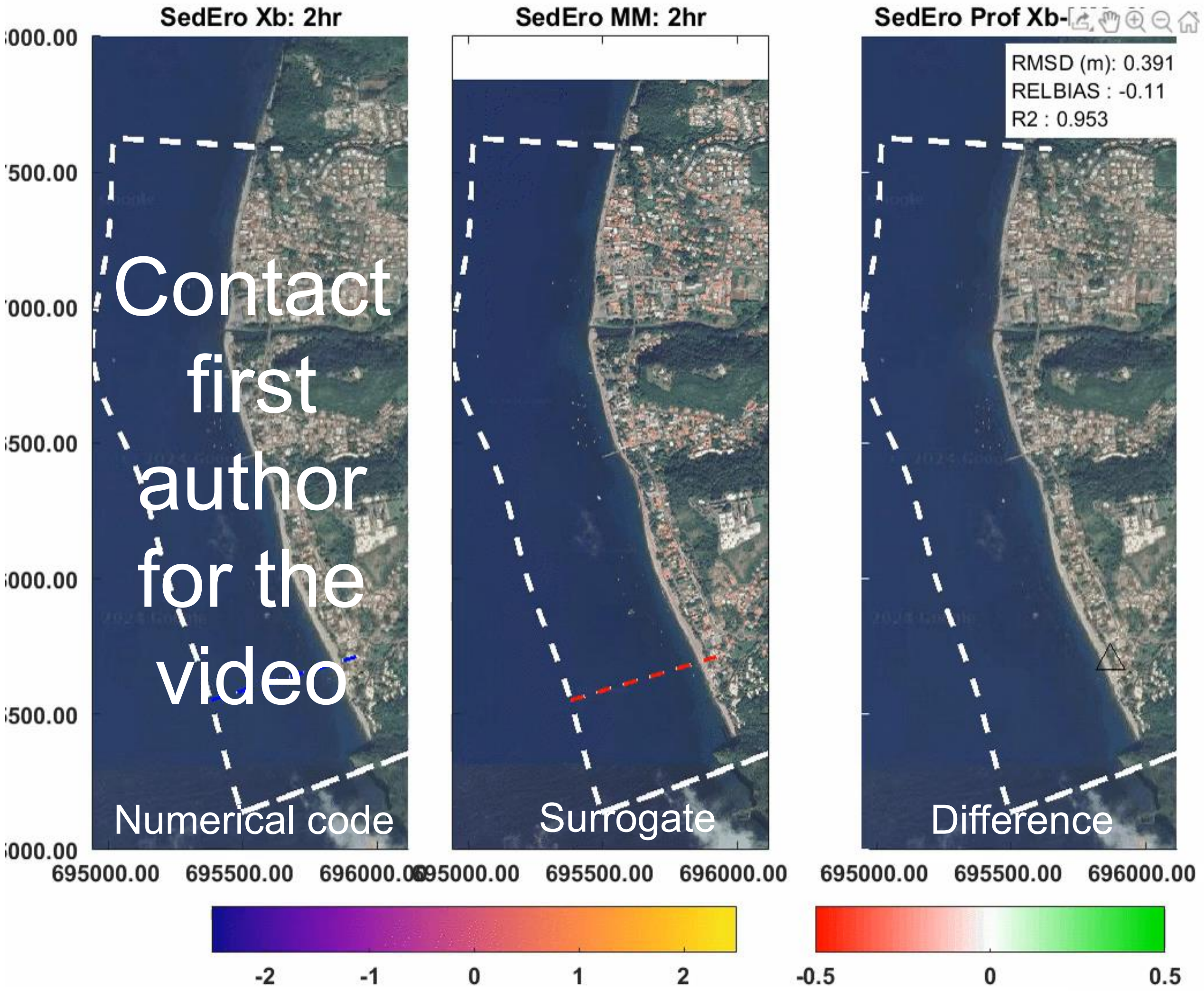
Auto-
Encoder

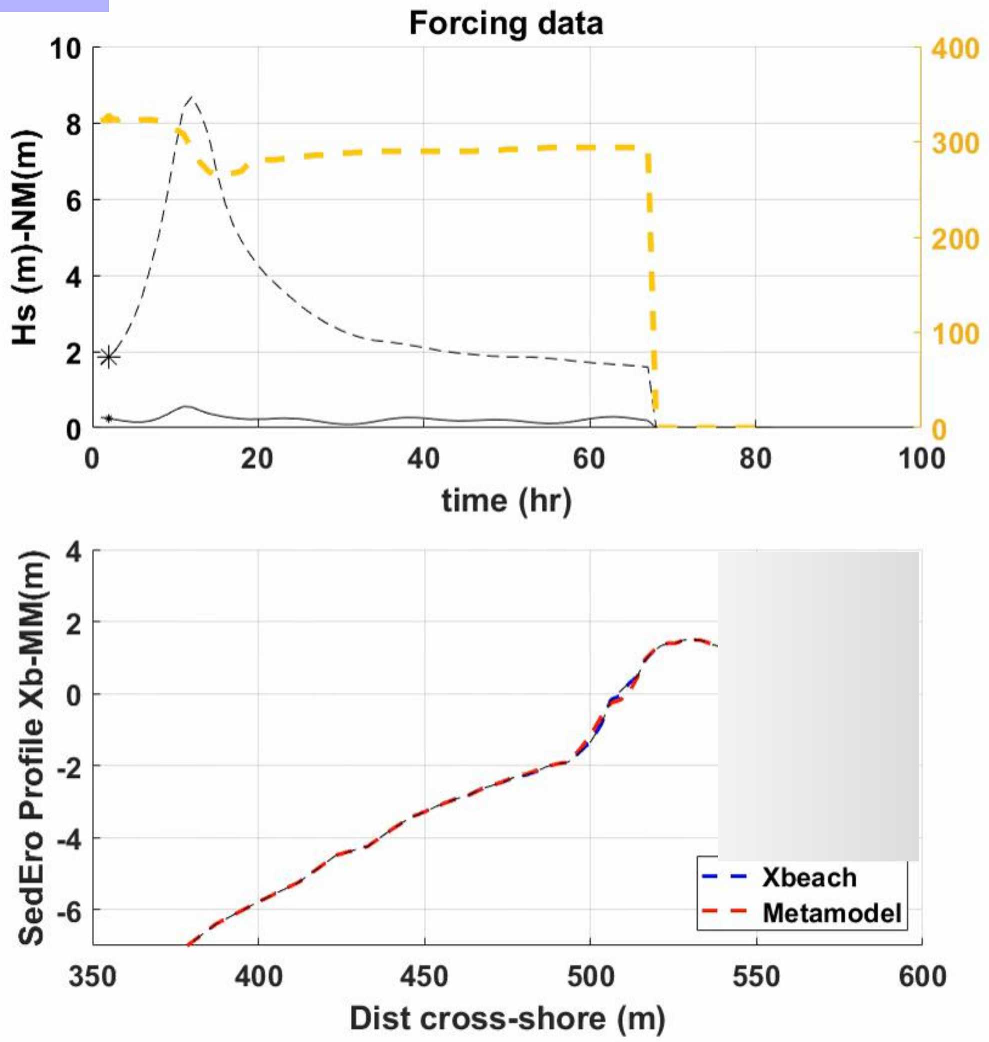


[ROH24] Rohmer et al., Symhydro, (2024)

Improvement #3 [VAL25]

Prediction of beach morphodynamic **spatio-temporal** response to tropical storms using **ConvLSTM**-based metamodels





[VAL25] Valentini et al., Coastal dynamics, (2025)

Summary

- Successful application of machine-learning based on surrogate modelling in **multiple coastal environments** (Arcachon, Gavres, Boucholeurs, Antilles)

- **Collaboration with IMT Toulouse** enabled us to overcome the '*curse of dimensionality*' (time varying inputs and spatial outputs)

- On-going developments for **DOE adapted to functional extreme inputs**
(N. Gorse's PhD within ANITI PILearnWater)

- The integration of **high dimensional spatiotemporal** variables (e.g., evolving bathymetry, space-time evolution of water depth, etc.) = **open question**

>> More coming soon for wave models with Antoine Jarry's PhD thesis (start 02/2026) - collab. ISAE-SUPAERO, INSA, BRGM

Géosciences pour une Terre durable
brgm

Merci pour votre attention!

Merci aux projets ANR ORACLES, ANR RISCOPE, ANR SHORECAST, Chaire PILearnWater

Risk-based System
For Coastal Flooding Early Warning

Géosciences pour une Terre durable