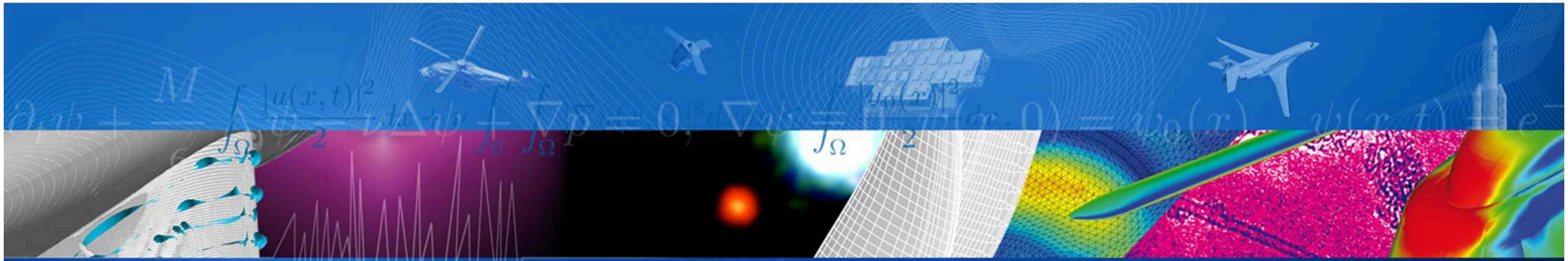


Some recent issues in numerical exploration of complex models: illustration in the case of objects infrared or radar signature estimation

Seminar on Complex Model Exploration – GDR Mascot Num

13/11/2015

S. Lefebvre - sidonie.lefebvre@onera.fr

ONERA

THE FRENCH AEROSPACE LAB

retour sur innovation

Outline

- Context
- Dimensionality reduction
- Metamodels for multi-fidelity analysis
- Concluding remarks

Context: Aircraft IRS

- Optimization of optronics sensor
  sensor must detect aircraft far ahead => low resolution images
- Computer program CRIRA => aircraft IRS according to aircraft properties
  weather conditions
  attack profiles

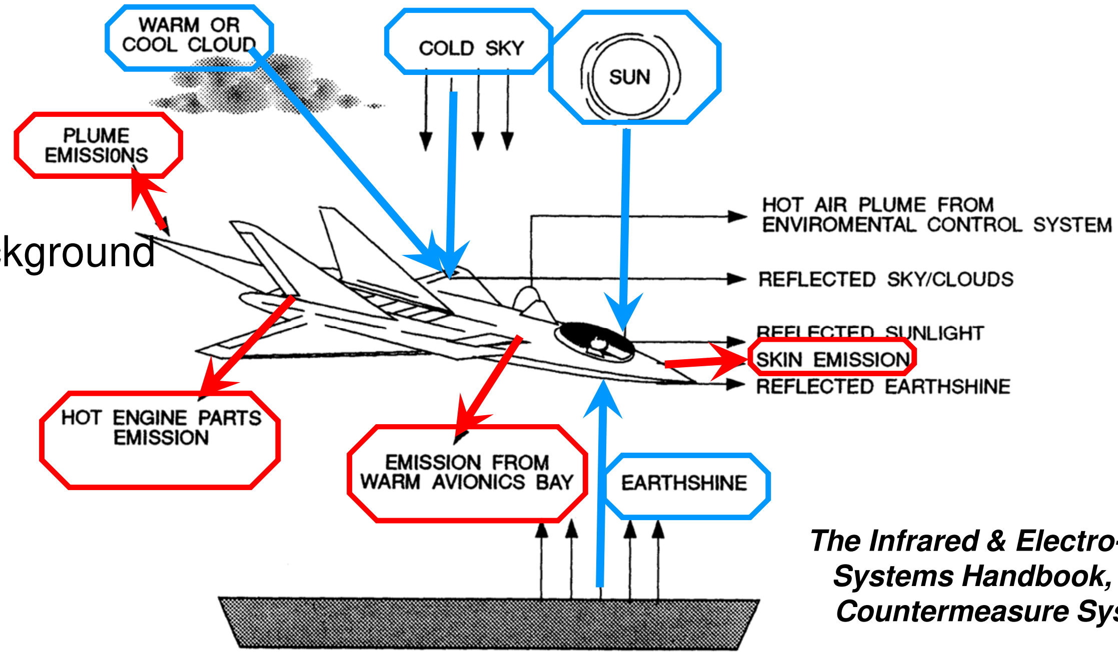
Reasonable cost for 1 basic simulation ~ **5000 simu**

Several contributions to IRS:

heat source emission

airframe reflected light from background

Uncertainties on input data



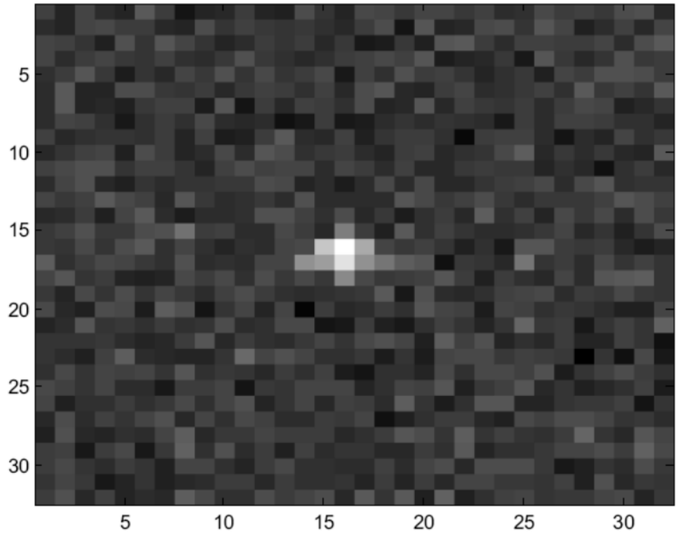
The Infrared & Electro-Optical Systems Handbook, Vol 7, Countermeasure Systems

Take IRS dispersion into account to estimate optronics sensor properties:

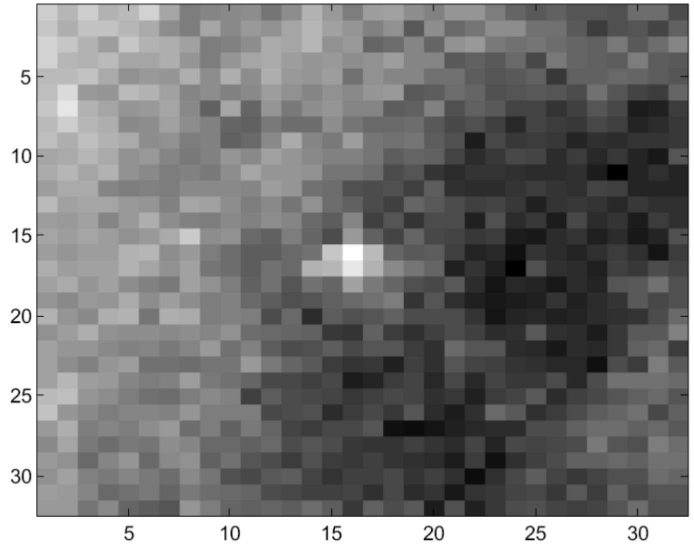
=> detection and correct classification probabilities for different aircraft

Context: Aircraft IRS

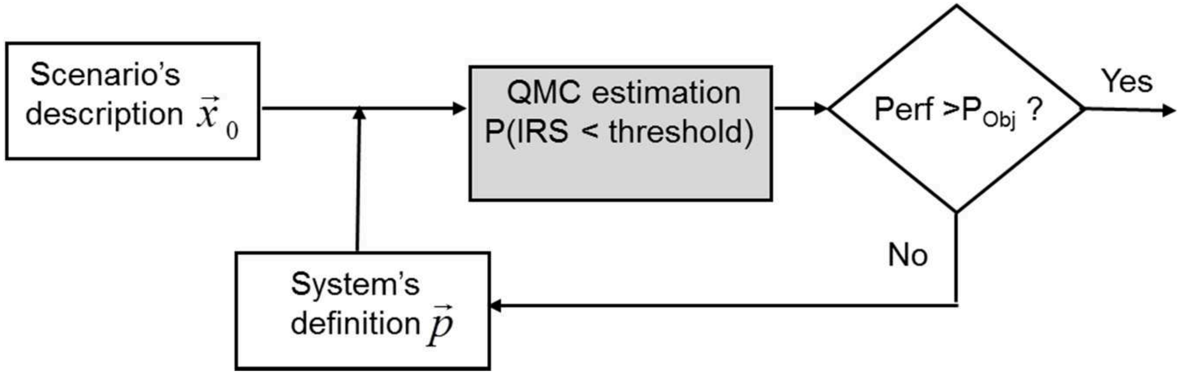
IRS examples



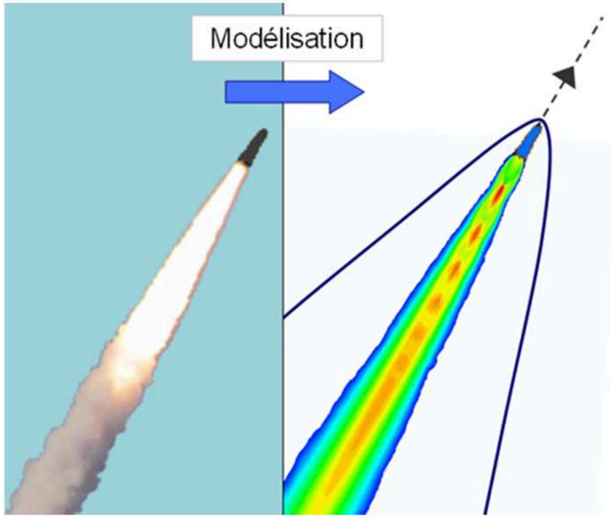
White noise background



Fractional Brownian background



Context: Sounding Rockets IRS and Radar Signature

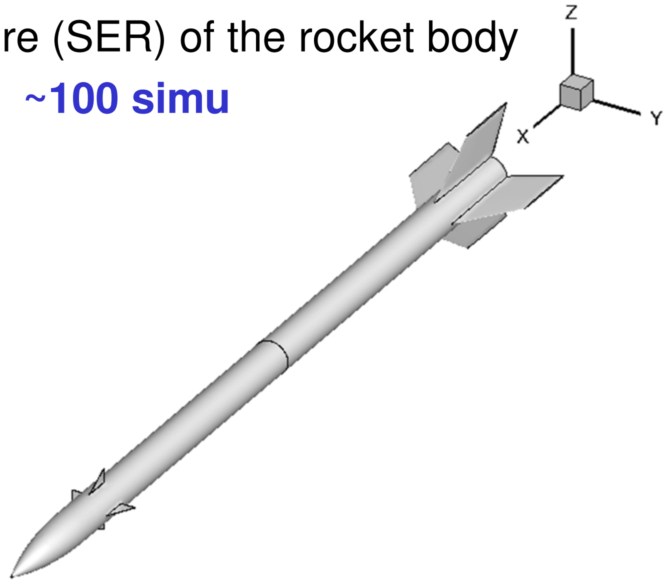


Simulation chain PRECISE (Plateforme de Recherche pour l'Évaluation de la Cinématique et des Signatures d'Engins) dev. at ONERA

=> Sizing systems for missile warning, surveillance or multisensor detection

Focus on:

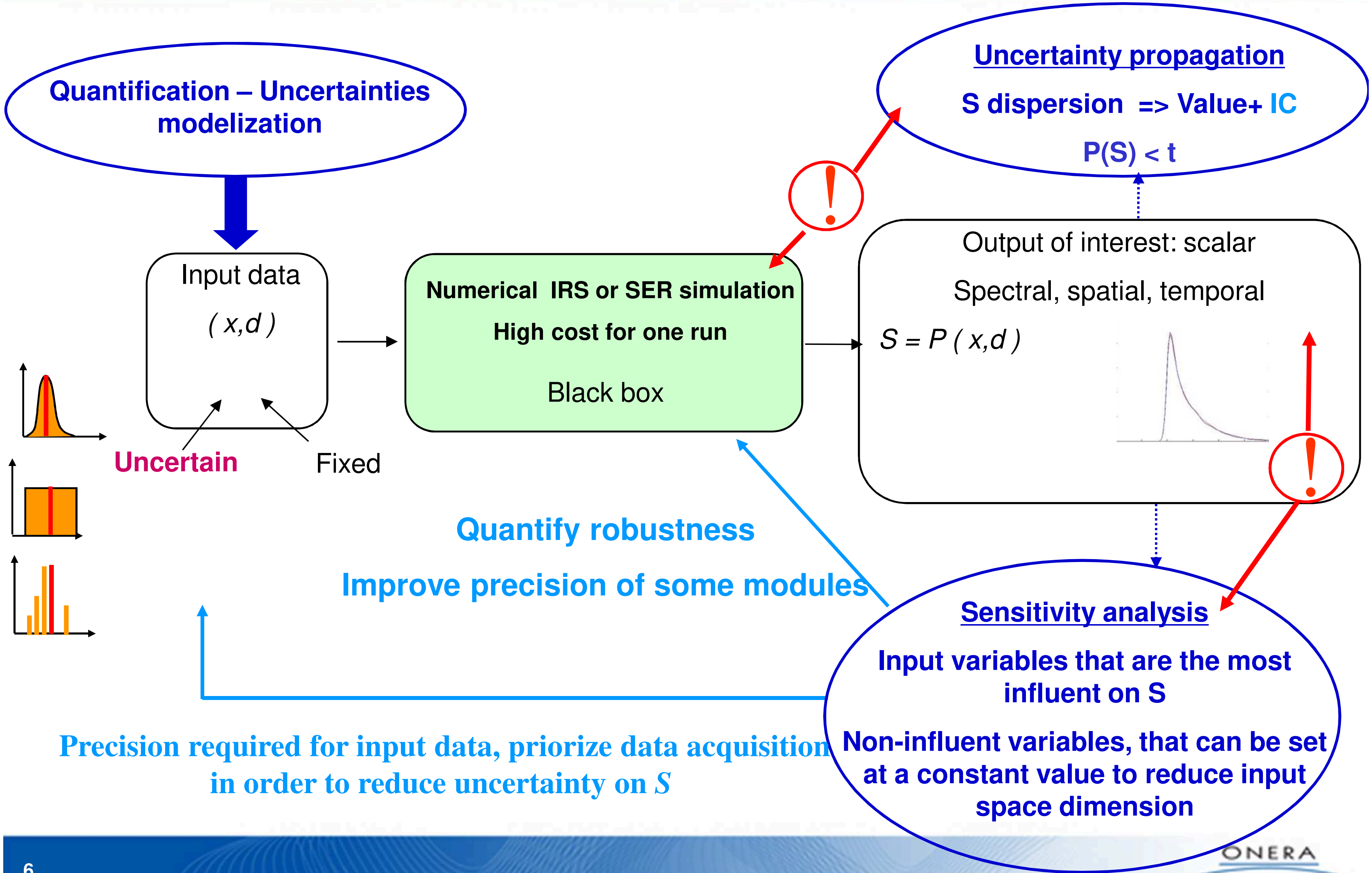
- rocket jet plume IRS – high cost – **simu <10**
- radar signature (SER) of the rocket body moderate cost **~100 simu**



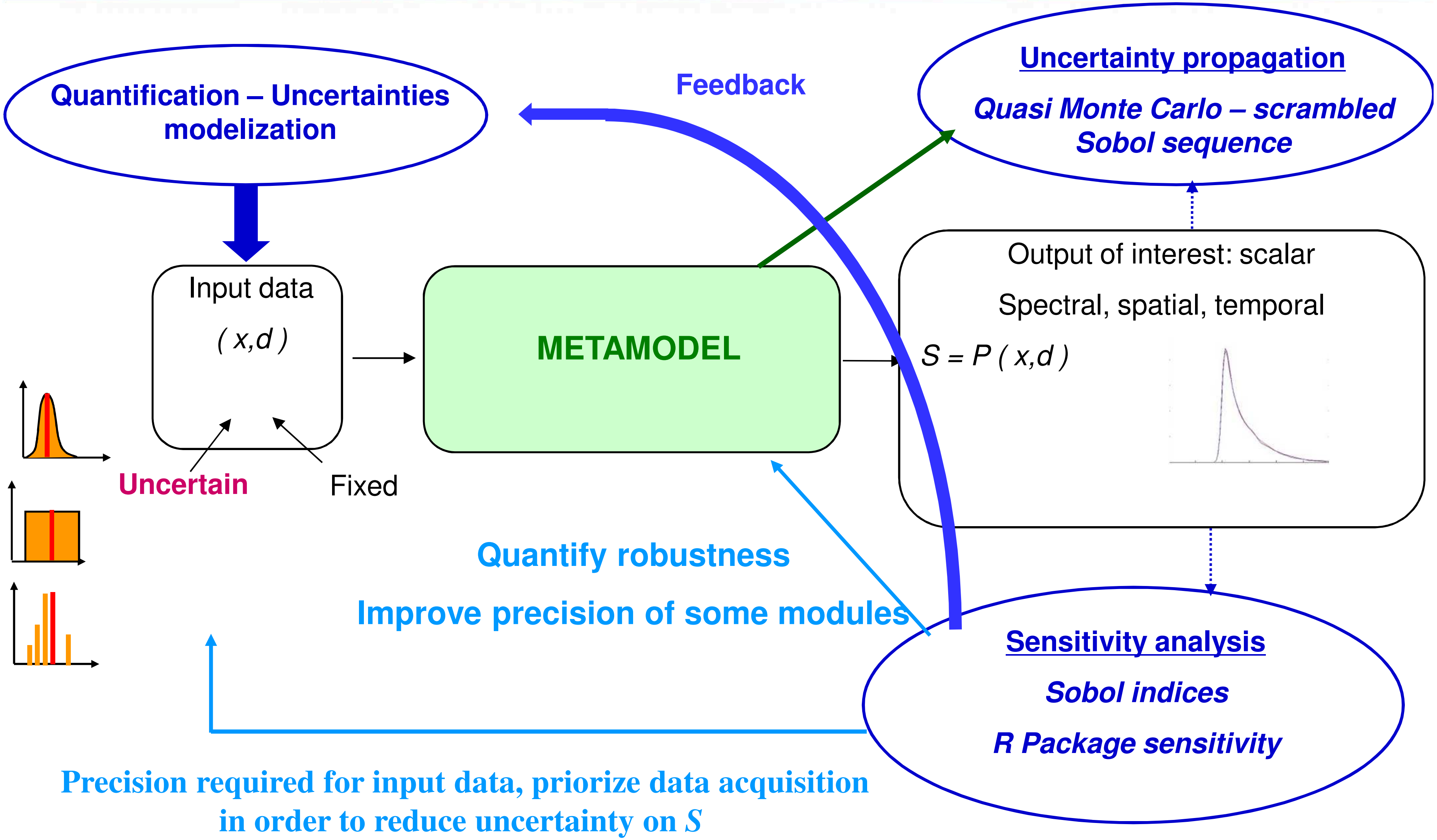
Uncertainties on input data:

CAO of rocket, aspect angles… poorly known

Issues of IRS and SER models exploration



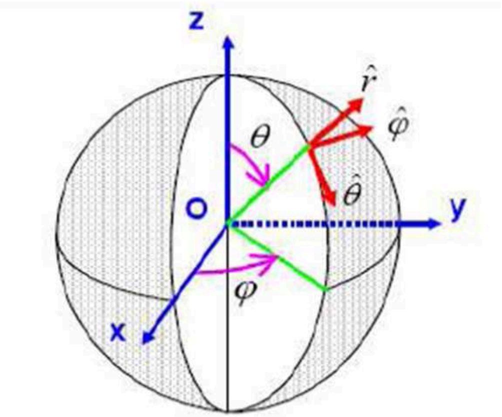
Issues of IRS and SER models exploration

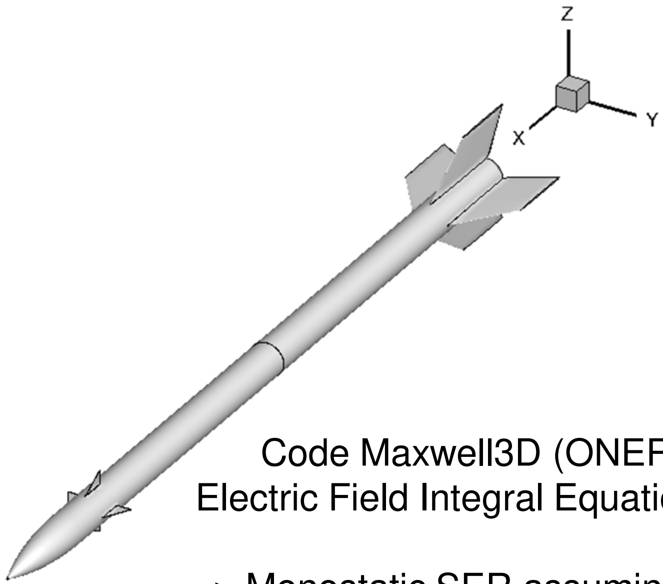


- Context
- Dimensionality reduction
- Metamodels for multi-fidelity analysis
- Concluding remarks

Study case: Radar Signature of rocket body

θ site	10 values from 45°	up to 90°	by 5 °
φ gisement	10 values from 0°	up to 45°	by 5 °





Code Maxwell3D (ONERA DEMR)
Electric Field Integral Equation formulation

=> Monostatic SER assuming plane waves

We analyze $10\log_{10}$ |polarization φφ| : **100** outputs at 400 MHz

Strong variability / aspect angles: from – 12 up to 12 dB

⇒ Dimensionality reduction

1st test PCA: 100 -> 40, not enough

⇒ **Non linear dimensionality reduction**

Nonlinear dimensionality reduction

We want to be able to:

- Calculate the coordinates of some new point in the low dimension space
- Go back from new components -> initial outputs without approx.

$\Rightarrow$ **Few nonlinear techniques can be used**

Two tests:

- Kernel PCA (drtoolbox Matlab): not very satisfactory results
- **Autoencoders** (NLPCA Matlab): MLP with 3 hidden layers and 1 output layer
  Kramer 1991, DeMers et Cottrell 1993, Hinton et Salakhutdinov 2006

Metamodel : Multilayer Perceptron

MLP = directed graph of neurons, i.e. non linear, parametric, bounded function, combined in successive layers that exchange informations through connections. All neurons in one layer work in //.

- Can be used for classification or regression
- One or several hidden layers
- Can account for multidimensional outputs

For each output

Linear activation function

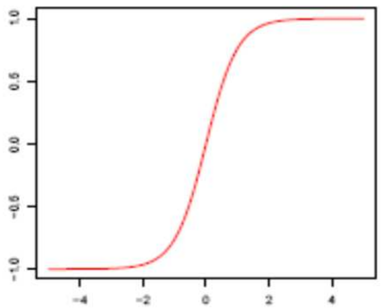
$$\varphi(x,w) = \sum_{i=1}^{n_c} w_{n_c+1,i}\, th\left(\sum_{j=1}^{n} w_{ij}x_j + w_{i0}\right) + w_{n_c+1,0}$$

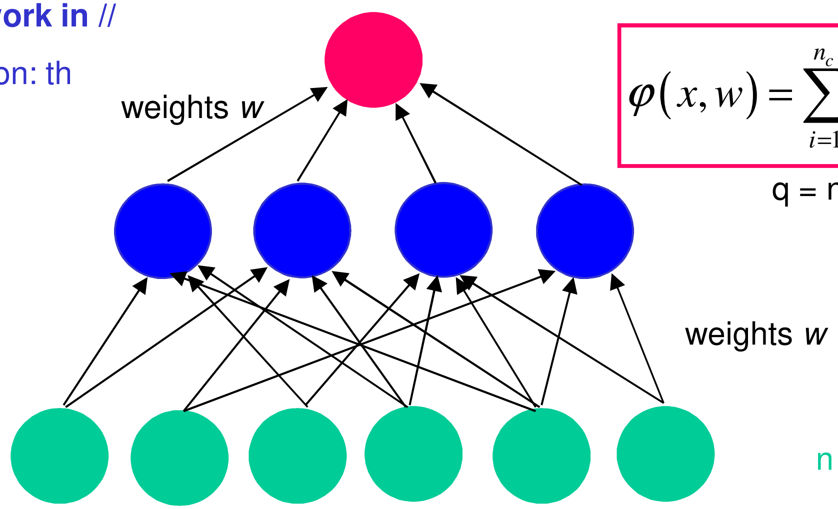
$q = n*n_c + 2*n_c + 1$ parameters

Bias units: capture the intercepts

Hidden layer n_c neurons – **work in //**

Sigmoid activation function: th



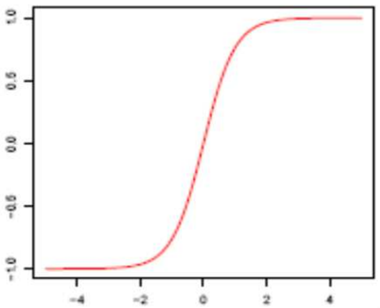


n input data

Autoencoders

Hidden layer n_c neurons

Sigmoid activation function: th



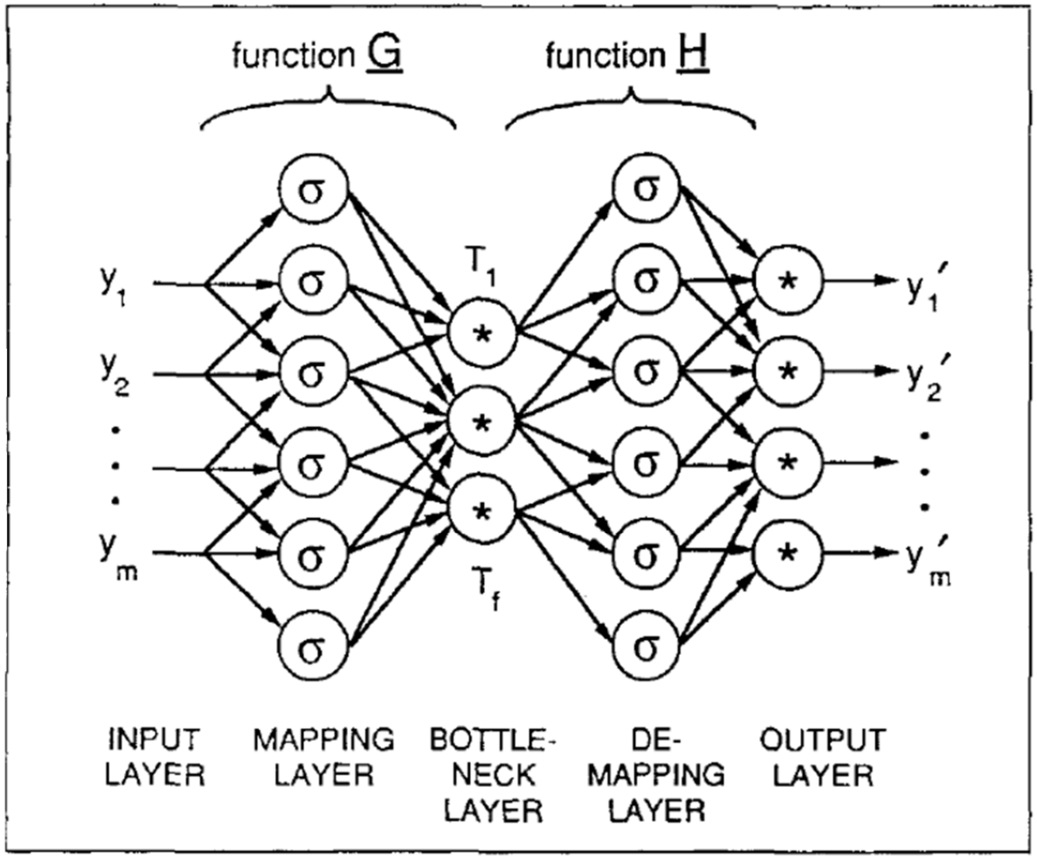


Figure 2. Network architecture for simultaneous determination of *f* nonlinear factors using an autoassociative network.

σ indicates sigmoidal nodes, ∗ indicates sigmoidal or linear nodes.

Kramer 1991

K linear neurons 2nd layer – K = new dimension

100 linear neurons => output

ONERA

Study case: Definition

Fixed parameters:

- Black Brant IX sounding rocket
- Fixed launch parameters: launch date and moment given, launch site is known
- Monostatic radar at 400 MHz
- Payload is known + some dimensions

Uncertain parameters: **11 variables CAO** – uniform law in millimeters except from flèches (°)

Variable	Min	Max
Longueur ogive	1103	1348
Rayon de nez	0.5	10
Largeur fente de jonction entre charge utile et deuxième étage propulsif	2.7	3.3
Profondeur fente de jonction entre charge utile et deuxième étage propulsif	0.9	1.1
Diamètre de l'engin au niveau de l'étage Black Brant et de l'ogive	429	447
Longueur de l'étage Black Brant	4880	5728
Emplacement des empennages centraux sur le deuxième étage propulsif (donné par rapport à l'arrière de l'étage)	0	100
Flèche au bord d'attaque des empennages centraux	54	66
Flèche au bord de fuite des empennages centraux	27	33
Corde à l'emplanture des empennages centraux	884	1325
Hauteur des empennages centraux	430	646

Limited nb of CAO (manual consistency check)

= > space filling design (LHS optimized / centered L2 discrepancy) with 110 runs + 1 reference simulation

Study case: Results of dimensionality reduction

	Median of errors obtained when rebuilding 100 outputs associated to $10\log_{10}\|\varphi\varphi\|$ (dB)
PCA 20 components	0.4
kPCA 15 components	0.27
Autoencoder 7 components	0.14
Autoencoder 15 components	0.05

Leave 11 out estimation

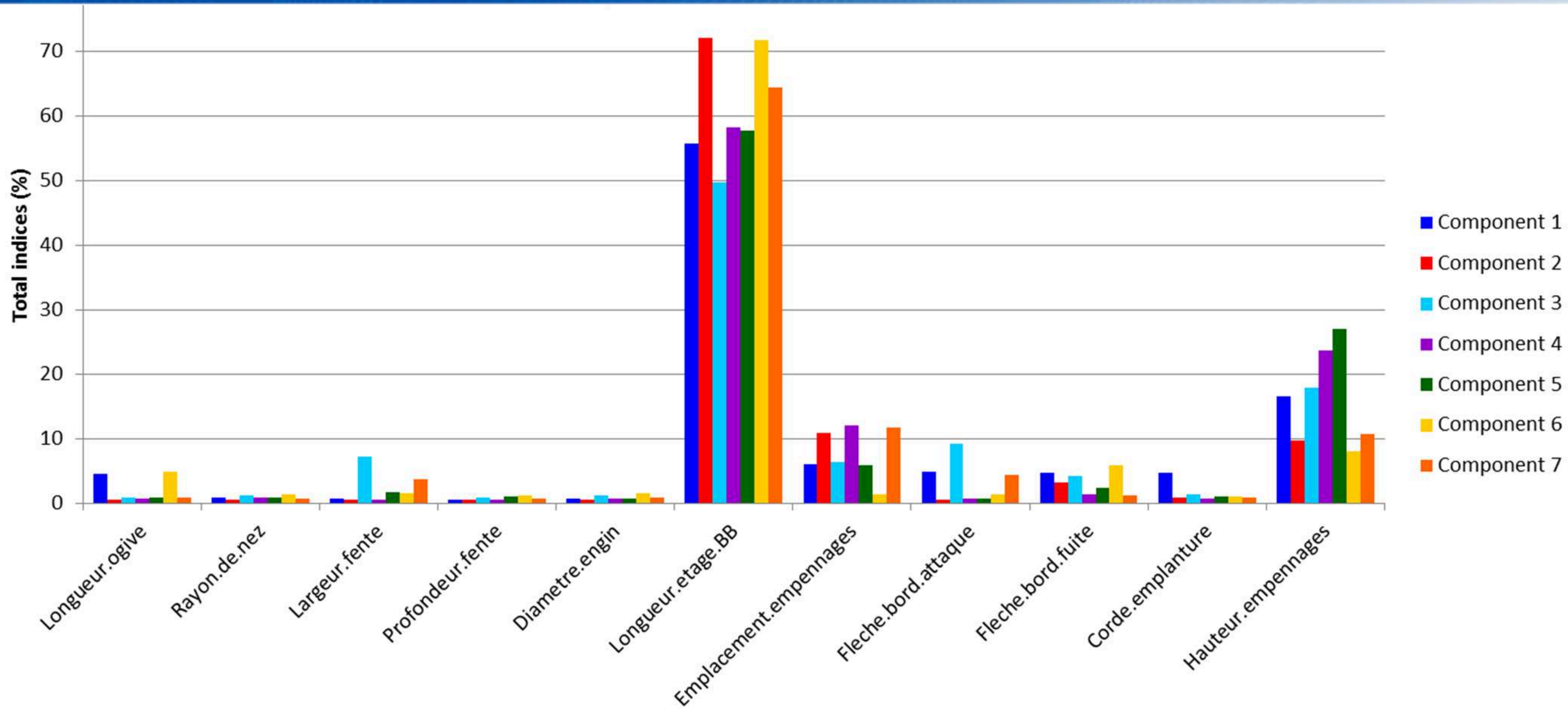
Then we learn **1 metamodel for each component**
=> Sensitivity Analysis (after nonlinear transfo…), SER dispersion

- Poor prediction performances with MLP metamodels

- Kriging metamodels (R package DiceKriging – Matern 5/2)

50 initial pts + 50 pts chosen with adaptive construction
L11out validation: some poor initial choices, but R2 median > 0.7 (resp. 0.5 for components 1 and 3)
max~0.9

Study case: Results of Sensitivity Analysis



|φφ| (dB)

Most predominant variable: **longueur étage BB**

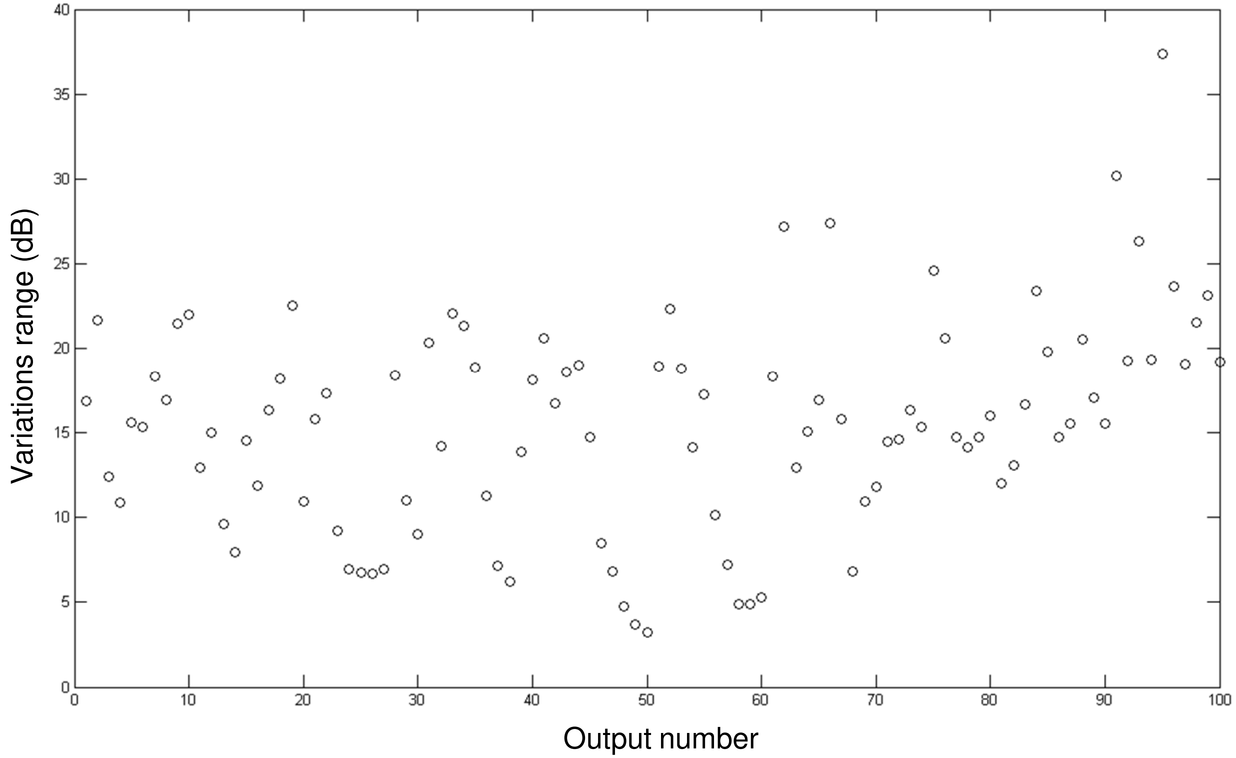
Also significant: **hauteur empennages – emplacement empennages - flèche bord d'attaque –** corde emplanture - longueur de l'ogive - flèche bord de fuite – largeur de fente

3 insignificant variables: profondeur de fente – diamètre engin – rayon de nez

Study case: SER dispersion

Dispersion : 50000 QMC samples

We use Kriging metamodels associated to 7 components and we rebuild the 100 initial outputs



$$10\log_{10}|\varphi\varphi|$$

Important variations range, from 3 up to 38 dB

=> CAO variables are influent and lead to high SER dispersion at 400 MHz

- Context
- Dimensionality reduction
- Metamodels for multi-fidelity analysis
- Concluding remarks

Multi-fidelity metamodels

We need about **10*d*** data in order to build a metamodel - *d* uncertain input variables
Pb: not possible for rocket jet plume IRS ~ 5-10 simu
⇒ Multi-fidelity metamodels

Basic idea (Kennedy et O'Hagan – 2000): **to combine data** coming from
- high fidelity simulations (HF): high cost, in small number,
- low fidelity simulations (LF): reasonable cost, more samples.

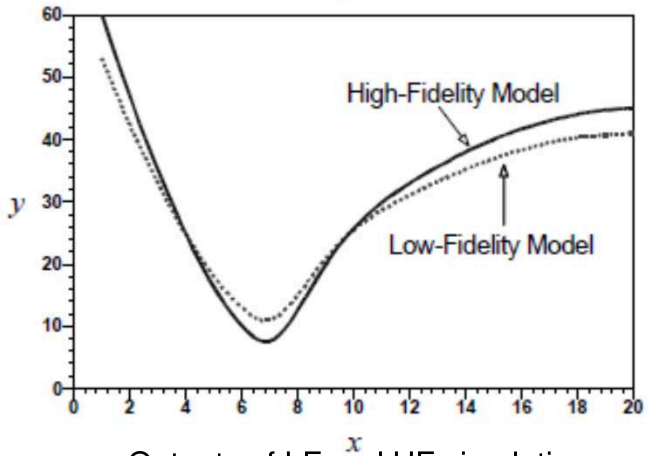
LF Simulations = physical reduced models
coarse mesh
non convergence
experimental measurements…

M. C. Kennedy, A. O'Hagan, Predicting the Output From a Complex Computer Code When Fast Approximations Are Available, Biometrika 87 (2000) 1-13.

Multi-fidelity metamodels

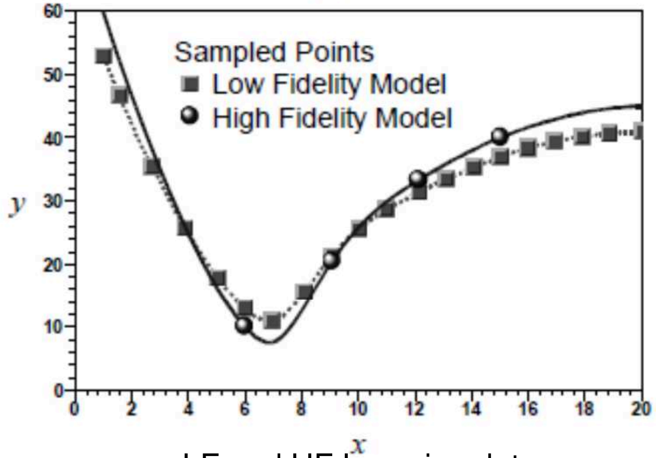
From Jones et al 1998

(a)



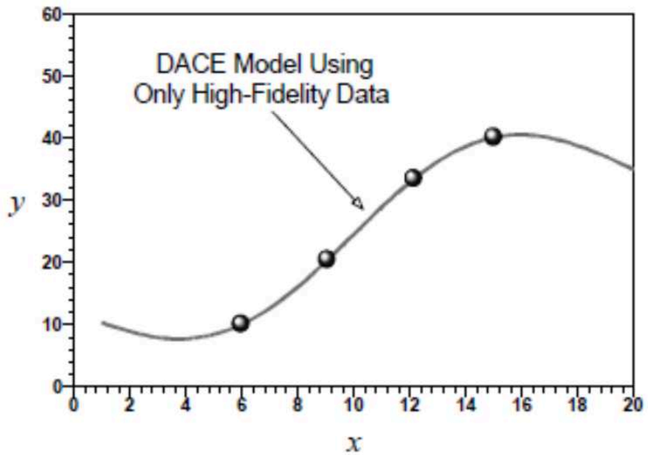
Outputs of LF and HF simulations

(b)



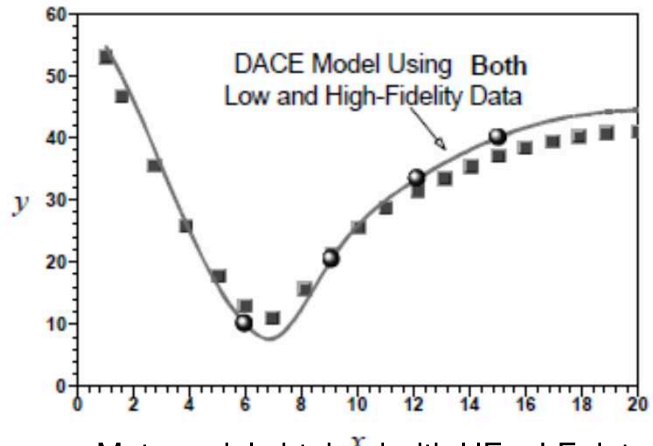
LF and HF Learning data

(c)



Metamodel obtained with HF data only

(d)



Metamodel obtained with HF + LF data

Better prediction performance

D.R. Jones, M. Schonlau, W.J. Welch
Efficient Global Optimization of Expensive Black-Box Functions
J. of Global Optimization, 13 (1998) 455-492.

Multi-fidelity metamodels

A simple multi-fidelity metamodel:

1. Build a metamodel based on LF data Y_l : linear regression, Kriging…
2. Adjust Y_l thanks to HF data, with 2 bridge functions a_0 and a_1

$$Y_h = a_0(X) + a_1(X)Y_l$$

Cokriging $$Y_h(x) = \rho_h Y_l(x) + \delta(x)$$

$$Y_l(x) \sim PG\left(g_l{}^T(x)\beta_l, \sigma_l^2 R_l\left(x, \tilde{x}, \theta_{l,0}\right)\right) \qquad \delta(x) \sim PG\left(g_h{}^T(x)\beta_h, \sigma_h^2 R_h\left(x, \tilde{x}, \theta_{h,0}\right)\right)$$

Hyp :

- $\delta \perp Y_l$: $cov(Y_h(x), Y_l(\tilde{x}) | Y_l(x)) = 0 \;\; \forall\, x \neq \tilde{x}$

  i.e. given the value of LF simu at x, we can learn no more about the value of HF simu at x from any other run LF at $\tilde{x} \neq x$.
- (not necessary) HF design $\subset$ LF design

There exists R Package: MuFiCokriging

Pb: - some variables are fixed in our HF simulations
- HF design $\not\subset$ LF design

=> **We make use of Han, Görtz and Zimmermann Hierarchical Kriging Model**

Stefan Görtz Zhong-Hua Han. Hierarchical kriging model for variable-fidelity surrogate modeling. AIAA Journal, 50(9) :1885 – 1896, September 2012.

Multi-fidelity metamodels

Hierarchical Kriging $Y_h(x) = \rho Y_l(x) + \delta(x)$

Y_l Kriging metamodel based on LF data

~ Kriging with LF metamodel as a model trend

Does not require modeling the cross correlation between LF and HF functions

$$\hat{\rho} = \left(F^T R_h^{-1} F\right)^{-1} F^T R_h^{-1} y_h$$

$$\widehat{Y_h}(x_0) = \hat{\rho} Y_l(x_0) + r_{0,h}^T R_h^{-1}(y_h - \hat{\rho} F)$$

$$\widehat{\sigma^2}(x_0) = \sigma^2 \left\{1 - r_{0,h}^T R_h^{-1} r_{0,h} + \left[r_{0,h}^T R_h^{-1} F - Y_l(x_0)\right](F^T R_h^{-1} F)^{-1}\left[r_{0,h}^T R_h^{-1} F - Y_l(x_0)\right]^T\right\}$$

where F = predictions LF metamodel on HF design,

R_h = δ covariance matrix on HF design,

$r_{0,h}$ covariances vector associated to new point HF x_0 and HF design

y_h = outputs HF simu.

Stefan Görtz Zhong-Hua Han. Hierarchical kriging model for variable-fidelity surrogate modeling. AIAA Journal, 50(9) :1885 – 1896, September 2012.

Study case: IRS of rocket jet plume

Black Brant sounding rocket - Calculation case reproduces the conditions of an experimental rocket launch, performed at White Sands in 1997, at altitude 7.9 km

IRS spatially and spectrally integrated over 2 wavelength bands– aspect angle 20°

5 HF data : ONERA simulations (CEDRE + SIR)

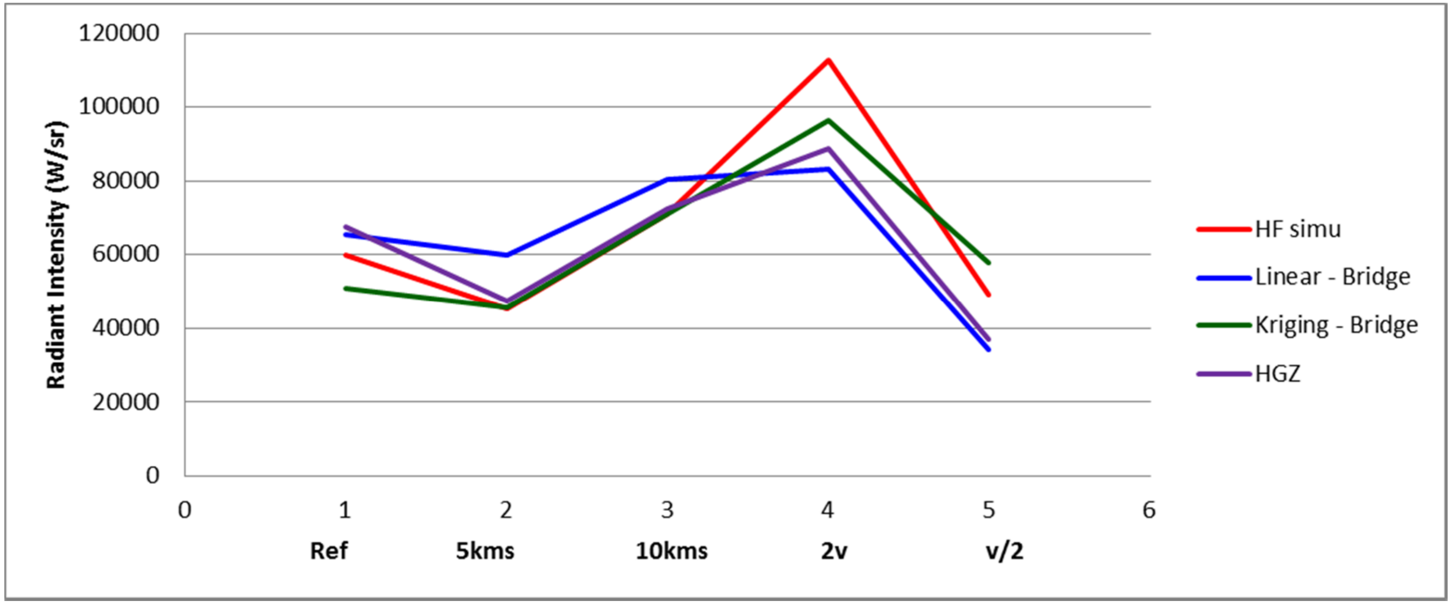
	Ref	v/2	2v	5 kms	10 kms
Speed	646.79 m/s	323.39 m/s	1293.58 m/s	646.79 m/s	646.79 m/s
Altitude	7.9 kms	7.9 kms	7.9 km	5 kms	10 kms
Atmospheric pressure	36117 Pa	36117 Pa	36117 Pa	54019 Pa	26436 Pa
Atmospheric temperature	236.8 K	236.8 K	236.8 K	255.65 K	223.15 K
Chamber pressure	45.2 atm	45.2 atm	45.2 atm	45.2 atm	45.2 atm
Aluminium Composition	19 %	19 %	19 %	19 %	19 %

44 LF data : parabolic code REP3 + SIR
4 HF data (v/2 is not compatible with the parabolic code) + 40 pts uniform distribution

Variable	Altitude	Speed	Chamber Pressure	Aluminium Composition
Min	5 kms	450 m/s (about 3v/2)	38.4 atm (- 15 % / nominal)	14 %
Max	10 kms	1293.58 m/s (2v)	52 atm (+ 15 % / nominal)	24 %

Results over 2000-2500 cm^{-1}

3 metamodels: Linear Regression + bridge functions a_0(Speed) and a_1 cst
Kriging + bridge functions a_0(Speed) and a_1 cst
Hierarchical Kriging (HGZ)

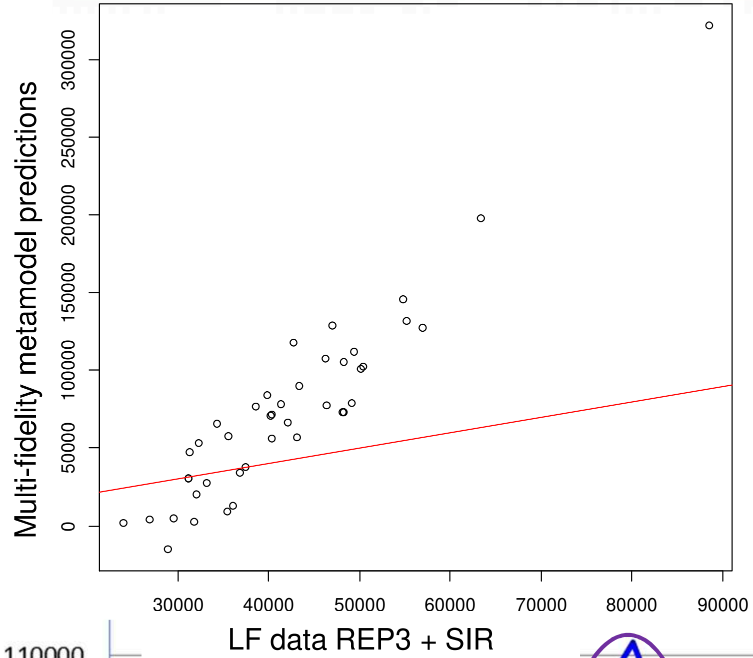


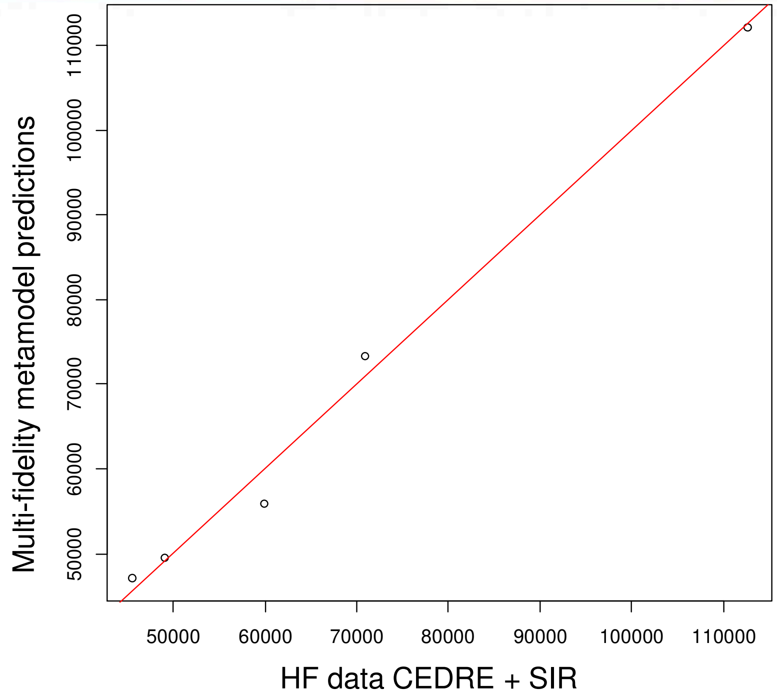
LOO predictions

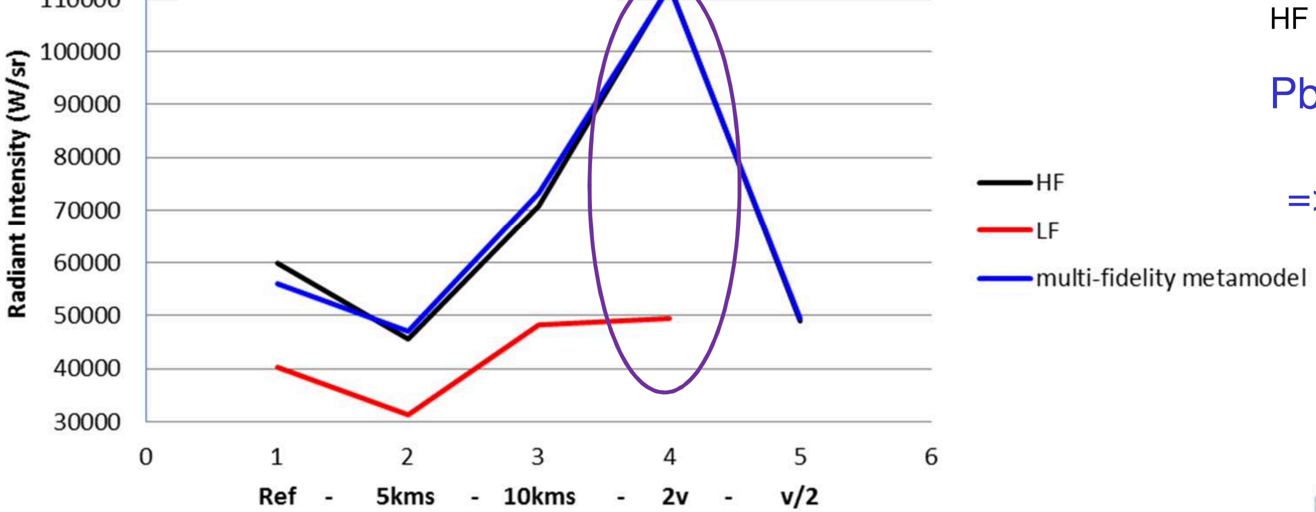
Best model: Kriging + bridge functions

Hierarchical Kriging promising

Results over 2000-2500 cm^{-1}





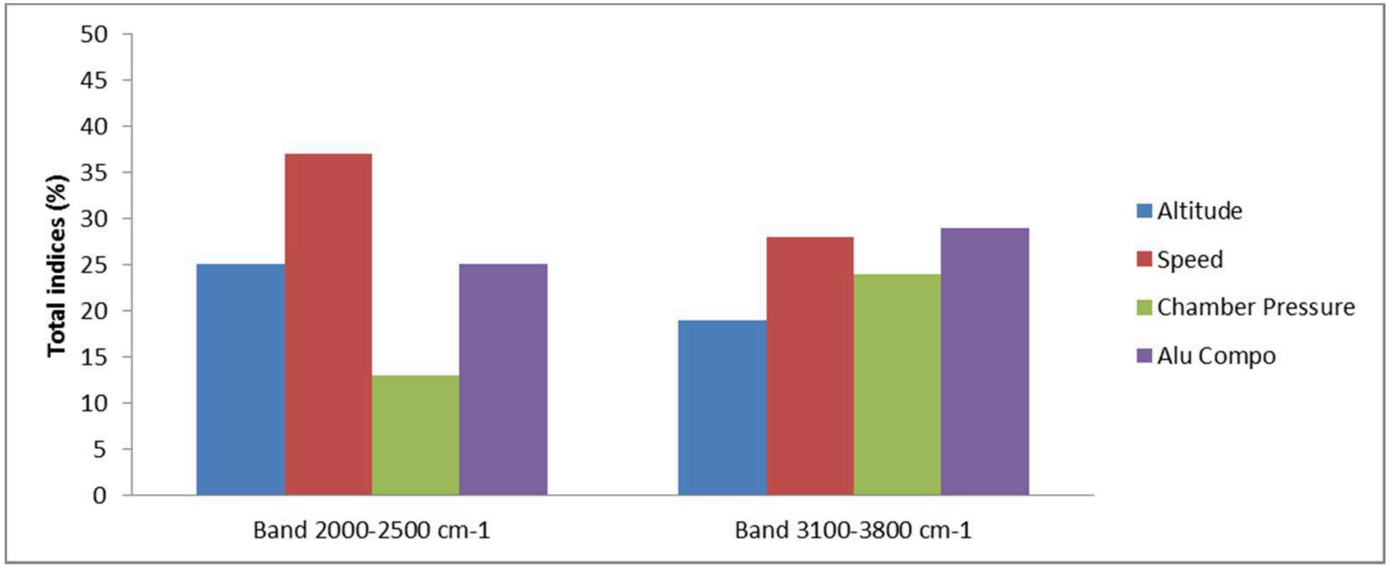


Pb: different behavior at high Speed
=> Another LF simulation ?

ONERA

Results: Sensitivity Analysis

Similar results with the 3 multi-fidelity metamodels



All variables are influent

Order differs: Chamber Pressure more influent over 3100-3800 cm^{-1}, Altitude over 2000-2500 cm^{-1}

Strong impact of speed, due to different behavior REP3/ CEDRE at 2v

- Context
- Dimensionality reduction
- Metamodels for multi-fidelity analysis
- Concluding remarks

ONERA
THE FRENCH AEROSPACE LAB

Concluding remarks

✓ First steps => non linear dimensionality reduction of high dimension outputs

✓ A first, **very simple** multi-fidelity metamodel => choose new HF simu

Future work

- Deep Learning with R package H2O
- Multi-fidelity metamodel:
  - 2 LF simu: REP3 if $v<v_{lim}$ and ?
  - Cokriging thanks to new HF simu
- Multi-fidelity metamodels for spectral or spatial IRS, categorical input variables…
- Inverse problem => threat identification

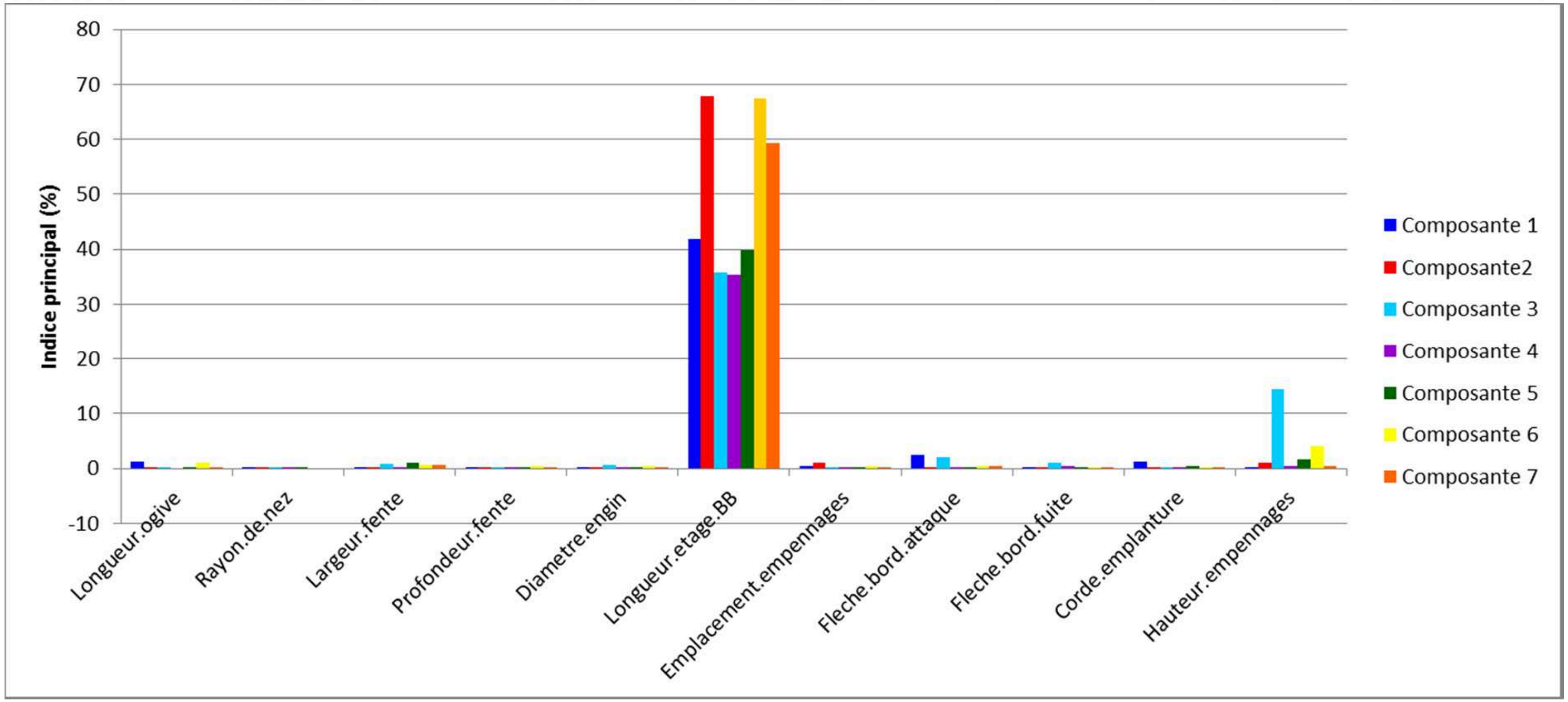
Questions ?

ONERA

THE FRENCH AEROSPACE LAB

retour sur innovation

Study case: Results of Sensitivity Analysis



|φφ| (dB)

Most predominant variable: **longueur étage BB**

Also significant: **hauteur empennages** - **flèche bord d'attaque** - corde emplanture

+ longueur ogive

3 insignificant variables: profondeur de fente – diamètre engin – rayon de nez

Functional ANOVA

Hyp: d input X^i indep. $E(f^2(X)) < \infty$

Unique decomposition (Hoeffding 1948) : $$f(X) = f_0 + \sum_{i=1}^{d} f_i(X_i) + \sum_{j=2}^{d} \sum_{1 \le i_1 < \ldots < i_j \le d} f_{i_1 \ldots i_j}\left(X_{i_1}, \ldots, X_{i_j}\right)$$

With $$\int_{[0,1)} f_{i_1,\ldots,i_j}\left(x^{i_1}, \ldots, x^{i_j}\right) dx^{i_k} = 0 \quad 1 \le k \le j$$

$$\int_{[0,1)^d} f_u(x) f_v(x) dx = 0 \quad u \ne v, u \subseteq \{1,\ldots,d\}, v \subseteq \{1,\ldots,d\}$$

Details:

$$f_0 = E(Y)$$

$$f_i(X_i) = E(Y / X_i) - f_0$$

$$f_{ij}(X_i, X_j) = E(Y / X_i, X_j) - f_i - f_j - f_0$$

$$f_{ijk}(X_i, X_j, X_k) = E(Y / X_i, X_j, X_k) - f_{ij} - f_{ik} - f_{jk} - f_i - f_j - f_k - f_0$$

Functional ANOVA

Orthogonality=> Variance decomposition

$$Var\left(f(X)\right)=\sum_{i=1}^{d}D_i+\sum_{j=2}^{d}\sum_{1\le i_1<\ldots<i_j\le d}D_{i_1\ldots i_j}$$

Where $D_i=\mathrm{var}\left(f_i(X_i)\right)$ and $D_{ij}=\mathrm{var}\left(f_{ij}(X_i,X_j)\right)$ …

Sobol' sensitivity indices:

$$S_I=\frac{D_I}{Var\left(f(X)\right)}$$

describe the amount of variance of Y explained by interactions involving factors whose indices are included in I

vary between 0 and 1 – sum up to 1 – equal to SRC^2 in the linear case

$1-\sum_{i=1}^{d}S_i$ measure the degree of interactions between variables

Total indices (Homma et Saltelli 1996)

$$S_{ti}=\sum_{\substack{I\subset\{1,\ldots,d\}\\ i\in I}}S_I=1-\frac{D_{-i}}{Var\left(f(X)\right)}$$

describe the total contribution of factor X^i (all interactions with other factors)

vary between 0 and 1

if S_{ti} ~0, variable not significant – if ~1 variable predominant

Sobol'indices estimation

$D_i = \mathrm{var}\left(f_i\left(X^i\right)\right)$ $S_I = \dfrac{D_I}{Var\left(f\left(X\right)\right)}$ $S_{ti} = \sum_{\substack{I \subset \{1,\ldots,d\} \\ i \in I}} S_I = 1 - \dfrac{D_{-i}}{Var\left(f\left(X\right)\right)}$

We need to estimate S_i and S_{ti} for $1 \le i \le d$ => *2d+1* integral estimations, including *I(f)* !

Recent estimation algorithms: Janon 2012 for S_i and Saltelli 2011 for S_{ti} :

Make use of 2 QMC samples (generally Sobol) with *N* pts : *X* and *Z* – we define (X^i, Z^{-i}) where the i^{th} coordinate comes from X and other coordinates come from Z (« Pick and freeze »)

$$\hat{S}_i = \frac{\dfrac{1}{N}\sum_{j=1}^{N} f(x_j) f(x_j^i, z_j^{-i}) - \left(\dfrac{1}{N}\sum_{j=1}^{N} \dfrac{f(x_j) + f(x_j^i, z_j^{-i})}{2}\right)^2}{\dfrac{1}{N}\sum_{j=1}^{N} \dfrac{f(x_j)^2 + f(x_j^i, z_j^{-i})^2}{2} - \left(\dfrac{1}{N}\sum_{j=1}^{N} \dfrac{f(x_j) + f(x_j^i, z_j^{-i})}{2}\right)^2}$$

$$\hat{S}_{ti} = \frac{\dfrac{1}{2N}\sum_{j=1}^{N} \left(f(x_j) - f(x_j^{-i}, z_j^i)\right)^2}{\dfrac{1}{N}\sum_{j=1}^{N} f(x_j)^2 - \left(\dfrac{1}{N}\sum_{j=1}^{N} f(x_j)\right)^2}$$

N (d + 1) runs

Typically N ≥ 500

Huge cost !

Results of Sensitivity Analysis : validation

	Nb of times influent / 100 outputs
Longueur ogive	25
Rayon de nez	5
Largeur fente	15
Profondeur fente	10
Diamètre engin	8
Longueur étage BB	**90**
Emplacement empennages	**35**
Flèche bord attaque	22
Flèche bord fuite	24
Corde emplanture	21
Hauteur empennages	**72**

1 Kriging metamodel and Sensitivity Analysis for each of the100 outputs

Very good agreement with our results – same predominant and insignificant variables

same order

Nonlinear dimensionality reduction: t-SNE

t-SNE (t Distributed Stochastic Neighbor Embedding) => visualize high-dimensional data by giving each datapoint a location in a 2D (or 3D) map

Non linear dimensionality reduction that seeks to preserve pairwise similarity

Two steps:

- Define joint probabilities for (x_i,x_j) in the high dimensional space

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n} \qquad p_{j|i} = \frac{exp\left(-\|x_i - x_j\|^2/2\sigma_i^2\right)}{\sum_{k\neq i} exp\left(-\|x_i - x_k\|^2/2\sigma_i^2\right)}$$

n dim of sample, σ_i Gaussian variance. Similarity of x_j to x_i corresponds to conditional probability $p_{j|i}$ that x_i would pick x_j as its neighbor if neighbors were picked prop. to their probability density under a Gaussian centered at x_i.

Joint probabilities for (y_i,y_j) in the low dimensional map are defined as:

$$q_{ij} = \frac{\left(1 + \|y_i - y_j\|^2\right)^{-1}}{\sum_{k\neq l}\left(1 + \|y_k - y_l\|^2\right)^{-1}}$$

Student t – heavy tailed distribution

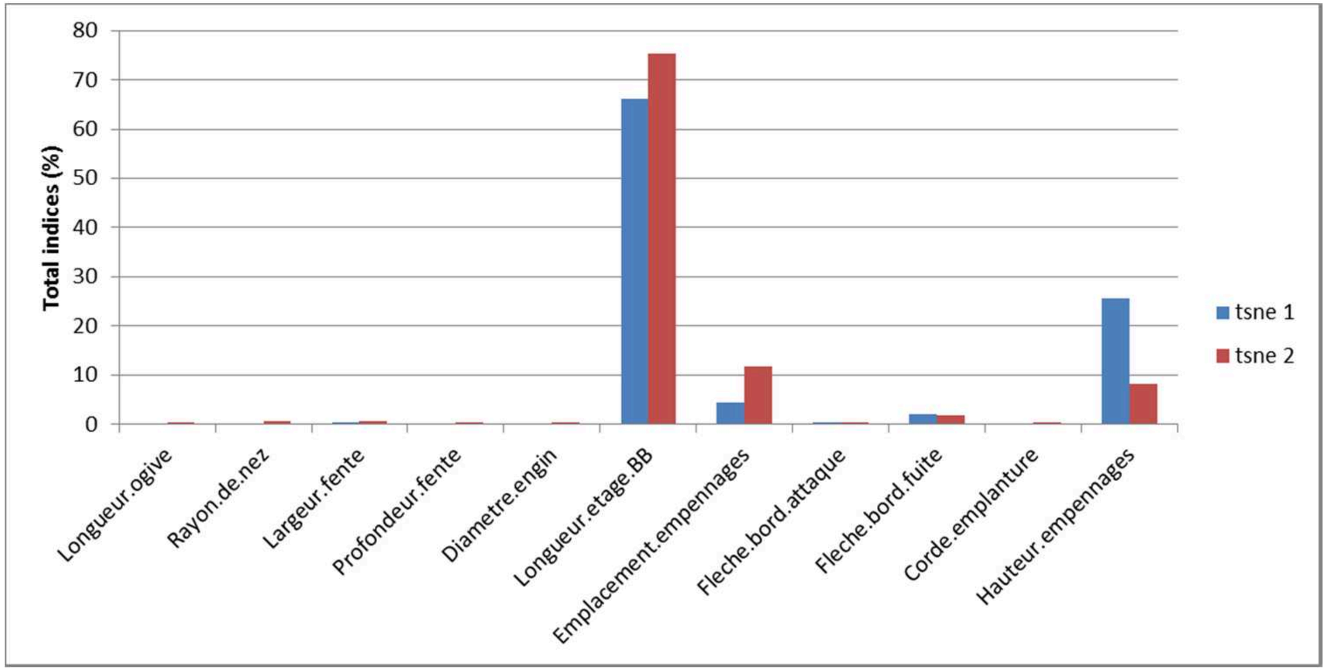
L.J.P Van der Maaten, G.E. Hinton
Visualizing High-Dimensional Data Using t-SNE
J. of Machine Learning Research 9, 2579-2605, (2008)

Nonlinear dimensionality reduction: t-SNE

- Coordinates y_i in the low dimensional map are obtained by minimizing the Kullback-Leibler divergence KL between the two probability distributions *p* et *q*, in order to preserve pairwise similarity

$$KL(p|q) = \sum_{i \neq j} p_{ij} log \frac{p_{ij}}{q_{ij}}$$

Sensitivity Analysis



3 most predominant variables: longueur étage BB - hauteur empennages - emplacement empennages – identical / autoencoders components

=> Concise analysis – **Pb: no way back to the 100 outputs**