Introspective metamodeling and thrifty optimization for simulated physical phenomena

IRSN
INSTITUT
DE RADIOPROTECTION
ET DE SÛRETÉ NUCLÉAIRE

Faire avancer la sûreté nucléaire

MINES
Saint-Étienne

PhD student: Nicolas Garland
EMSE / IRSN

EMSE supervisors:
Rodolphe Le Riche
Nicolas Durrande
Xavier Bay

IRSN supervisor:
Yann Richet

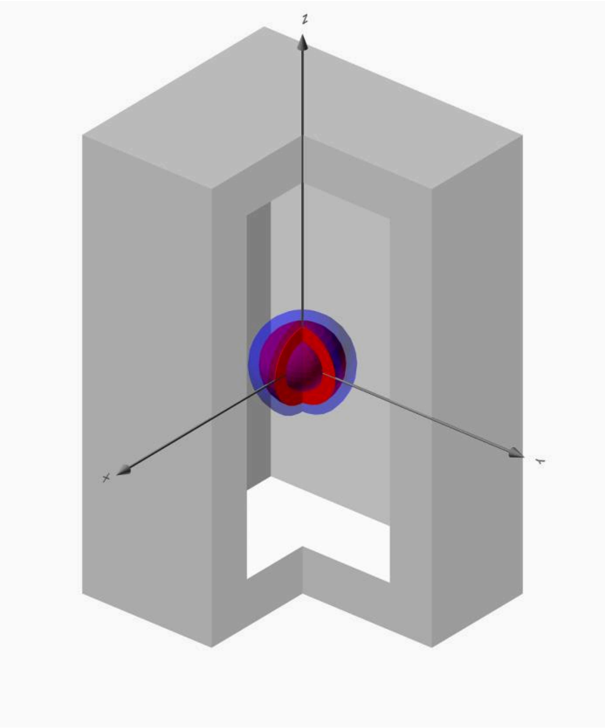
MascotNum Annual Conference, March 22-24 2017,
Paris-Massy, France

Contents

Engineering Problem

↗ Illustrative example

- A company wants to stock nuclear material
- IRSN is asked if the planned project is safe
- IRSN's experts in criticality have to find the worst case scenario



↗ Optimization problem

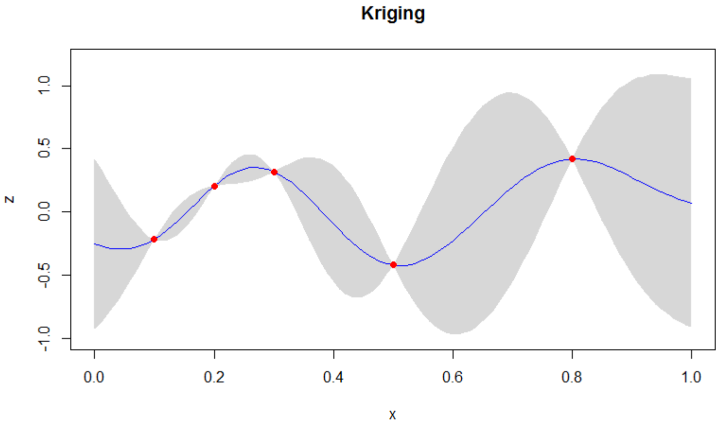
- But simulations codes take too much times for classical optimization

↗ Thrifty optimization problem

Efficient Global Optimization

↗ Iterative algorithm

1. *N* known points.
2. Kriging model => stochastic predictions.



Efficient Global Optimization

Kriging: stochastic metamodel

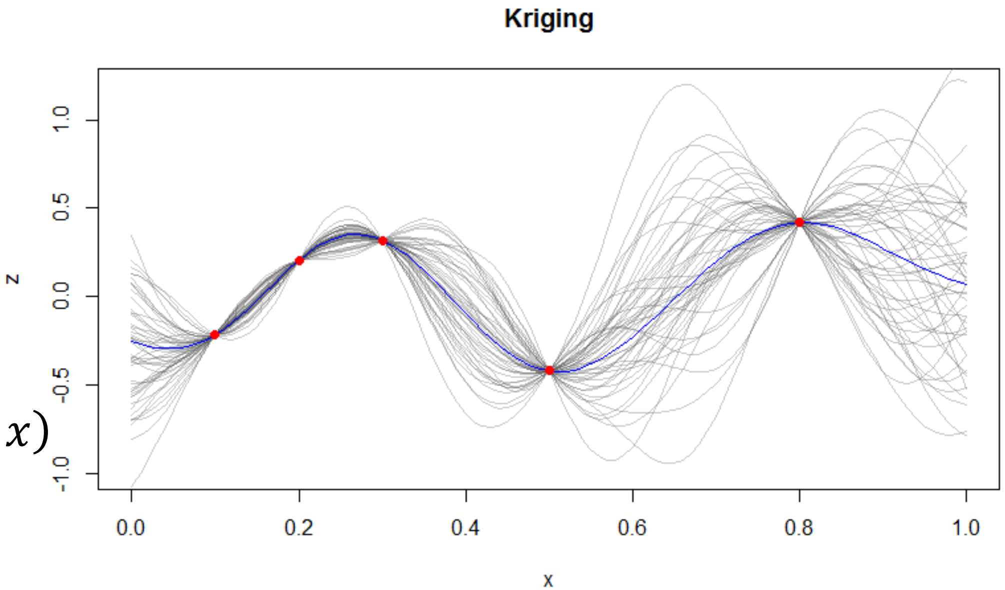
- Hypothesis about covariance: $cov(Z_{x_1}, Z_{x_2}) = k(x_1, x_2)$
- Conditionning by known points (X_N, Z_N)

➢ Prediction = distribution (gaussian)

$$Z(x) \sim N\big(\mu(x), \sigma^2(x)\big)$$

$$\mu(x) = k(x, X_N) K^{-1} Z_N$$

$$\sigma^2(x) = k(x,x) - k(x, X_N) K^{-1} k(X_N, x)$$



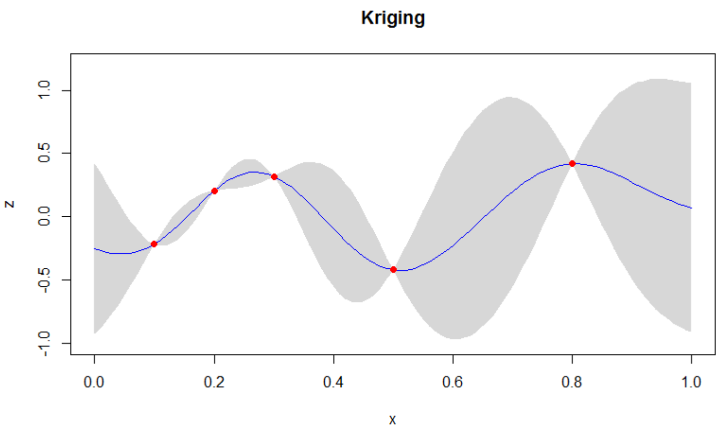
Efficient Global Optimization

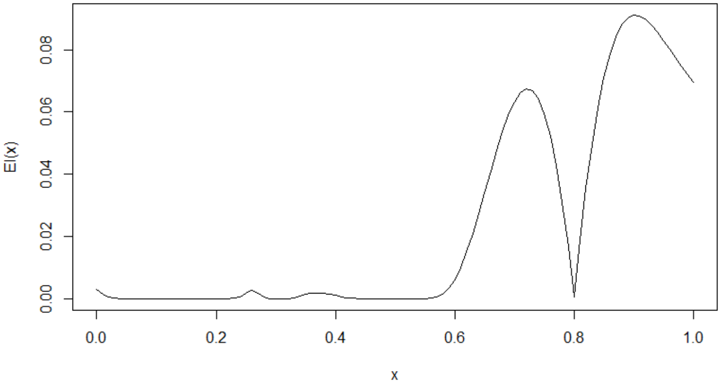
Iterative algorithm

1. *N* known points.
2. Kriging model => stochastic predictions.
3. Criterion computed from the model on each *x*: EI, Expected Improvment of the current optimum for chosing *x* as the new calculated point.

$$EI(x) = (T - \mu(x))\Phi\left(\frac{T - \mu(x)}{\sigma(x)}\right) + \sigma(x)\,\varphi\left(\frac{T - \mu(x)}{\sigma(x)}\right)$$

4. Best EI => new point to be evalueted.
5. Simulation code => value on the new point.
6. *N+1* known points, loop until end-of-loop criterion.





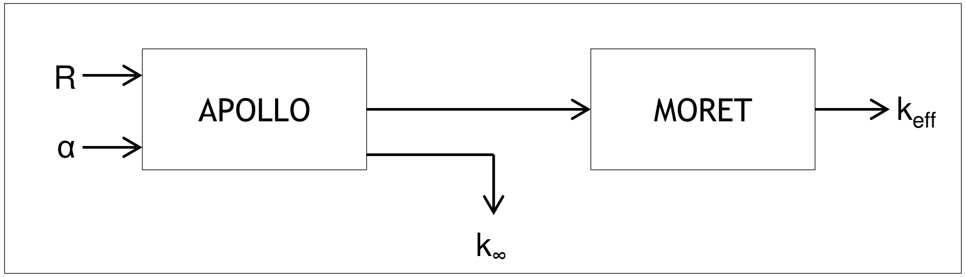
Efficient Global Optimization

- Known since 1998 (Jones, Schonlau and Welch)
- Many works about it: on handling noise, on parameters estimation, R packages, …
- But:
- The complexity of the simulated physic is often split between different codes.
- The black-box approch is not fully satisfying.

Chained Codes

Example: APOLLO-MORET

- Simulations are made by 2 successive codes



- APOLLO: compute a simplified equivalence for materials
  => reactivity for infinite materials (k_∞) ; short computations

- MORET: compute neutrons moves
  => reactivity for real materials (k_{eff}) ; long computations

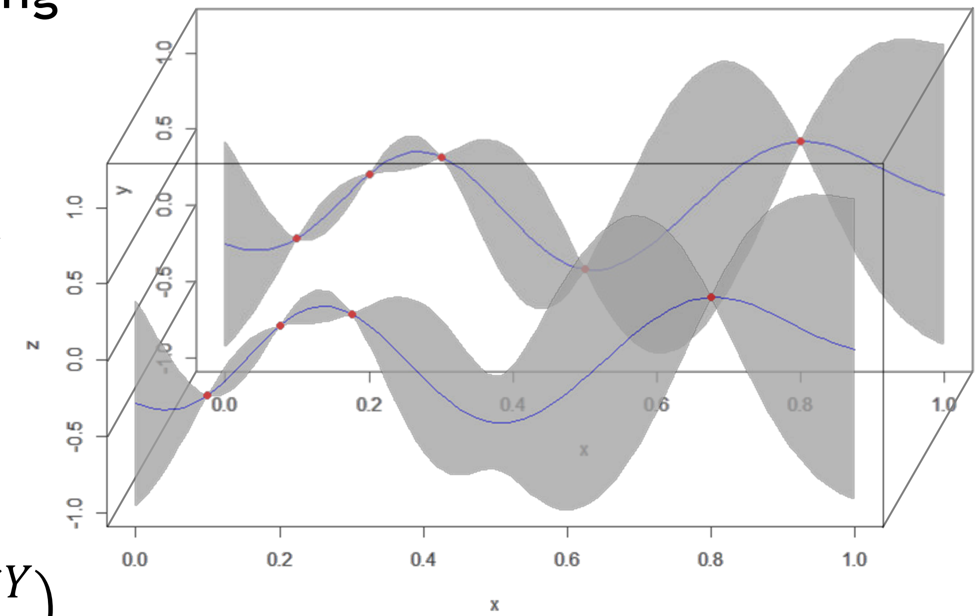
Information from k_∞ used for better metamodel

Computation times used for better algorithm

Introspective Metamodels

↗ Cokriging

- Same principle than kriging
- Covariance function depends of the nature of the points (Y/Y or Z/Z or Y/Z or Z/Y)



$$Z(x) \sim \mathcal{N}\left(\mu_Z(x)\,, {\sigma_Z}^2(x)\right)$$

$$\mu_Z(x) = {}^t\begin{pmatrix} k_{YZ}(X_Y, x) \\ k_{ZZ}(X_Z, x) \end{pmatrix} \begin{pmatrix} K_{YY} & K_{YZ} \\ K_{ZY} & K_{YY} \end{pmatrix}^{-1} \begin{pmatrix} Y \\ Z \end{pmatrix}$$

$${\sigma_Z}^2(x) = k_{ZZ}(x,x) \; - {}^t\begin{pmatrix} k_{YZ}(X_Y, x) \\ k_{ZZ}(X_Z, x) \end{pmatrix} \begin{pmatrix} K_{YY} & K_{YZ} \\ K_{ZY} & K_{YY} \end{pmatrix}^{-1} \begin{pmatrix} k_{YZ}(X_Y, x) \\ k_{ZZ}(X_Z, x) \end{pmatrix}$$

Introspective Metamodels

Cokrigings

- Covariance still has to be positive-definite

- Various ways to choose covariance function:

  - Same range parameters (CoThId): $k_{YY} \propto k_{ZZ}$; $k_{YZ} = k_{ZY} = \rho \sqrt{k_{ZZ}\, k_{YY}}$
  - Transformed space (CoAffi): $x \rightarrow \frac{x - m}{\theta_W}$ $with\ W = Y\ or\ Z$; then similar to CoThId
  - Convolution method (CoConv): $k_{WW'} \propto \int k_W(x_1, a)\, k_{W'}(a, x_2)\, da \quad with\ \ W\ and\ W' = Y\ or\ Z$
  - Linear model of coregionalization (LMC), balanced (CoDoub) :

    $Y(x) = \sigma_Y{}^2 U(x) + \rho\sigma_Y\sigma_Z V(x) \quad et \quad Z(x) = \rho\sigma_Y\sigma_Z U(x) + \sigma_Z{}^2 V(x) \qquad with\ \ U \perp V$
  - LMC, multifidelity (CoMark): $Z(x) = \rho\, Y(x) + \delta(x)\ \ with\ \ \delta \perp Y$

(Ref: Fricker, 2013, Multivariate gaussian process emulators with nonseparable covariance structures.)

Introspective Metamodels

Cokriging: parameters estimation

- All those cokrigings have similar structure

$$cov\big(Y(x_1),Y(x_2)\big) = \sigma_Y^2\, k_{YY}(x_1,x_2,(\rho),\Theta)$$
$$cov\big(Z(x_1),Z(x_2)\big) = \sigma_Z^2\, k_{ZZ}(x_1,x_2,(\rho),\Theta)$$
$$cov\big(Y(x_1),Z(x_2)\big) = \rho\, \sigma_Y\sigma_Z\, k_{YZ}(x_1,x_2,\Theta)$$

- Same parameters interpretation

- Explicit log-likelihood maximizer available for σ_*:

$$\sigma_Y^2 = \frac{-(n_Y - n_Z)w_{YZ}^2 + 2n_Y w_{YY} w_{ZZ} + w_{YZ}\sqrt{\Delta^*}}{2n_Y^2 w_{ZZ}}$$

- (publication in progress)

Introspective Metamodels

Hyperkriging

1) Kriging of Y => $Y(x)$

$$Y(x) \sim \mathcal{N}\left(\mu_Y(x), {\sigma_Y}^2(x)\right)$$
$$\mu_Y(x) = k_Y(x, X){K_Y}^{-1}\, Y$$
$${\sigma_Y}^2(x) = k_Y(x, x) - k_Y(x, X)\, {K_Y}^{-1} k_Y(X, x)$$

2) Kriging of Z as function of (x, y) => $Z(x, y)$

$$y \equiv supplementary\ dimension \quad ; \quad x' = (x, y)$$
$$Z(x, y) \sim \mathcal{N}\left(\mu_Z(x, y)\, , {\sigma_Z}^2(x, y)\right)$$
$$\mu_Z(x, y) = k_Z(x',\ X'){K_Z}^{-1}\, Z$$
$${\sigma_Z}^2(x, y) = k_Z(x', x') - k_Z(x', X')\, {K_Z}^{-1} k_Z(X', x')$$

3) Composition: distribution of $Z(x)$ is not gaussian

$$f_{Z(x)}(z) = \int f_{Z(x,y)}(z) f_{Y(x)}(y) dy$$

Introspective Metamodels

Hyperkriging: Annexe (non-gaussian dealing)

- Approximation by Taylor series:

$$\int_{-\infty}^{+\infty} g(x,y)\frac{1}{\sigma_x}\varphi\left(\frac{y-\mu_x}{\sigma_x}\right)dy = g(x,\mu_x) + \sum_{j=1}^{+\infty}\frac{1}{j!\ 2^j}\sigma_x^{2j}\frac{\partial^{2j} g}{\partial y^{2j}}(x,\mu_x)$$

Introspective Metamodels

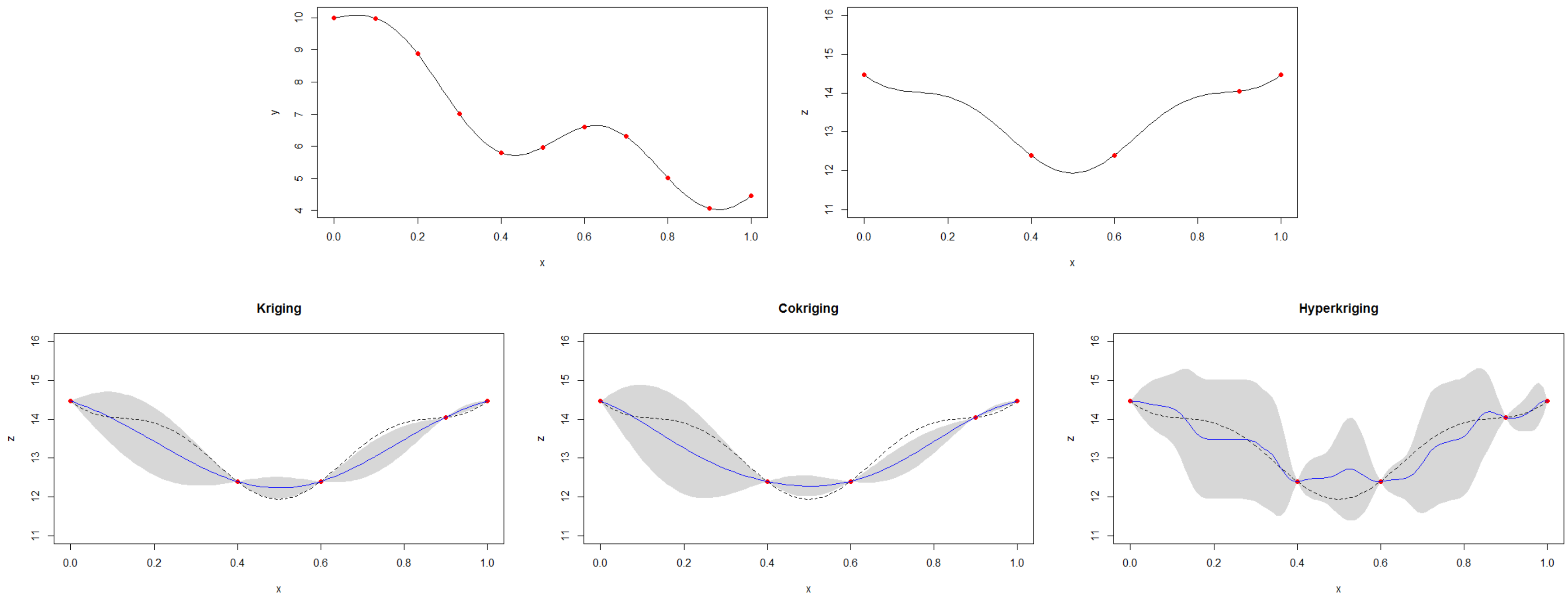
↗ Hyperkriging: Annexe (non-gaussian dealing)

$$\mu(x) = \mu_Z(x, \mu_x) + \sum_{j=1}^{+\infty} \frac{1}{j!\ 2^j} \left(\sigma_Y^2(x)\right)^j \frac{\partial^{2j} \mu_Z}{\partial y^{2j}}(x, \mu_x)$$

$$EI(x) = EI_Z(x, \mu_x) + \sum_{j=1}^{+\infty} \frac{1}{j!\ 2^j} \left(\sigma_Y^2(x)\right)^j \frac{\partial^{2j} EI_Z}{\partial y^{2j}}(x, \mu_x)$$

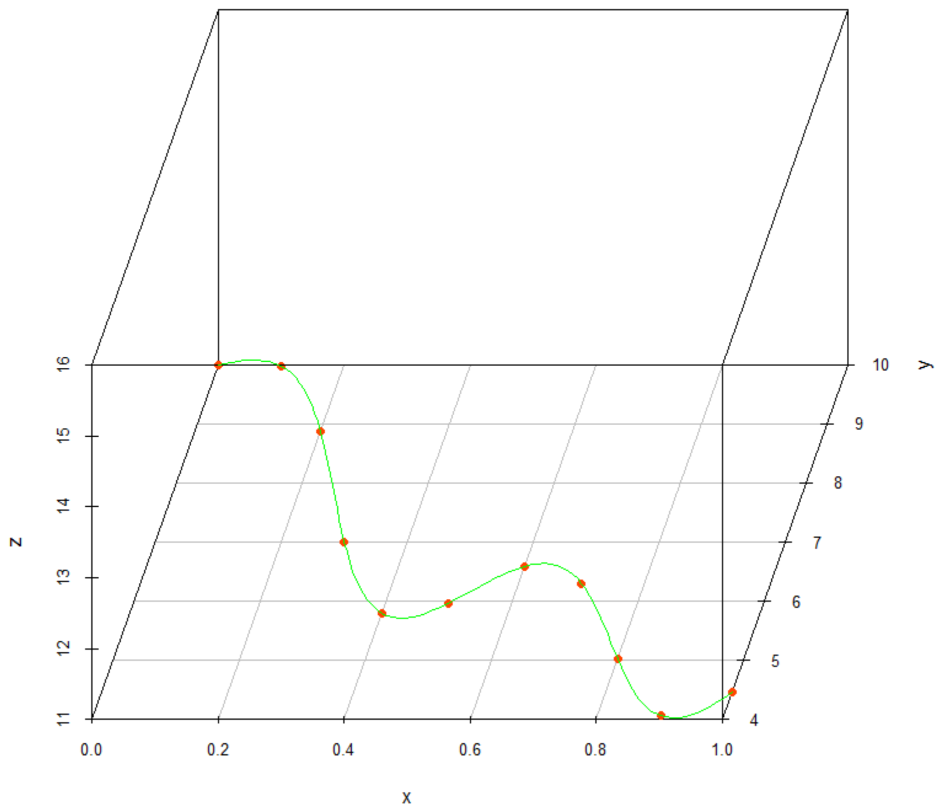
Introspective Metamodels

↗ Hyperkriging: Annexe (illustration)



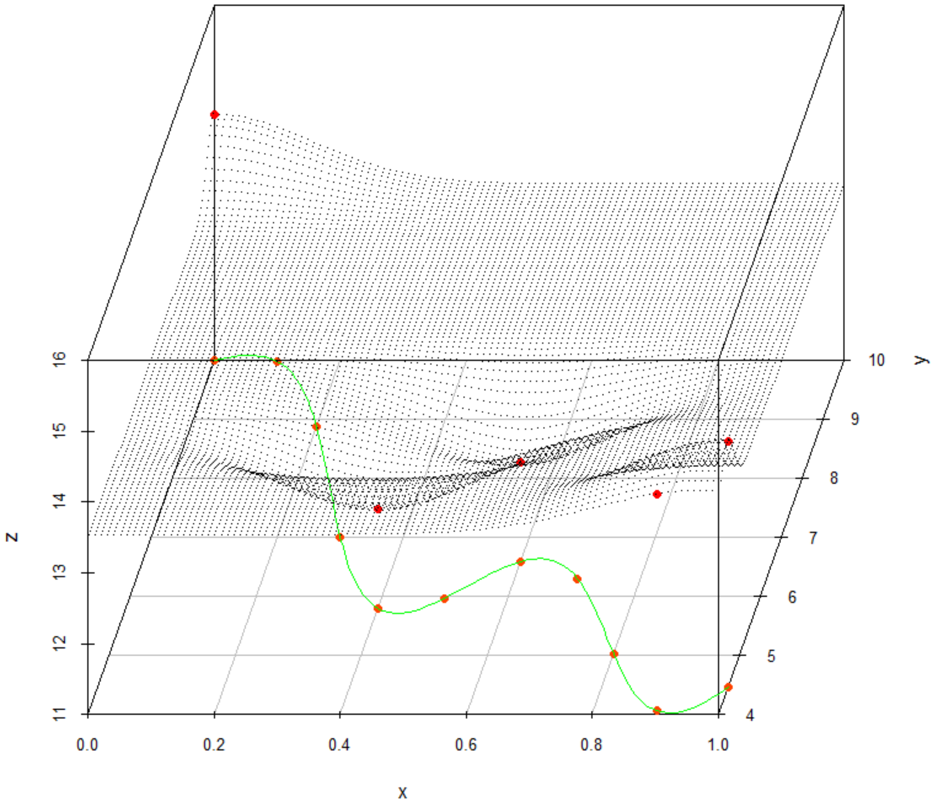
Introspective Metamodels

↗ Hyperkriging: Annexe (illustration)



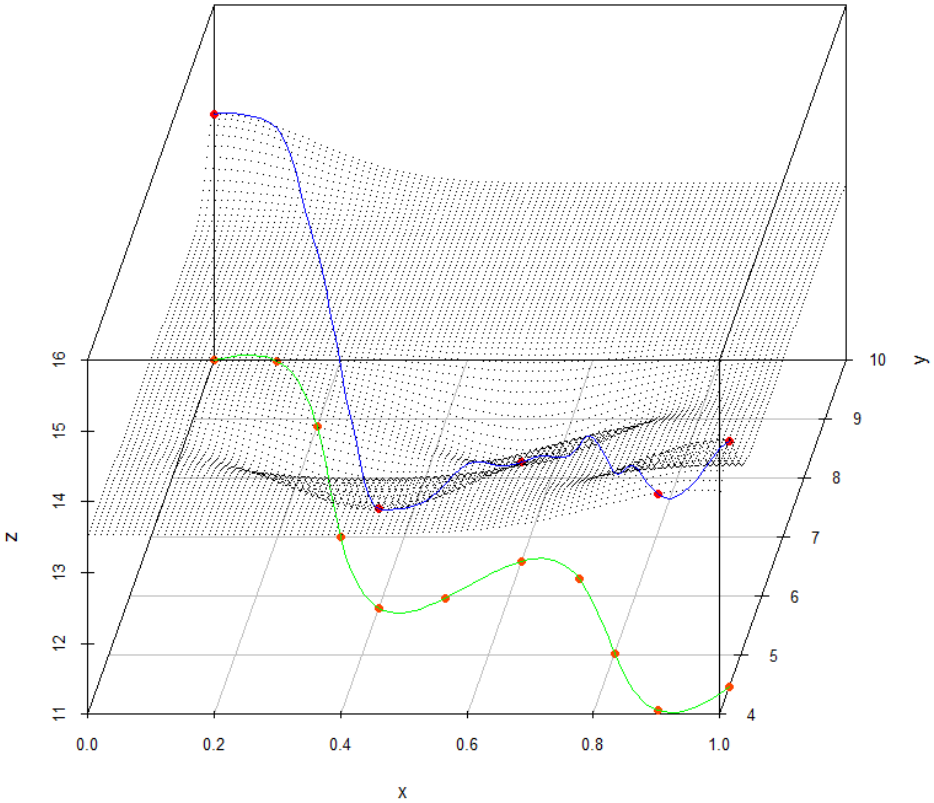
Introspective Metamodels

↗ Hyperkriging: Annexe (illustration)



Introspective Metamodels

➚ Hyperkriging: Annexe (illustration)



Introspective Metamodels

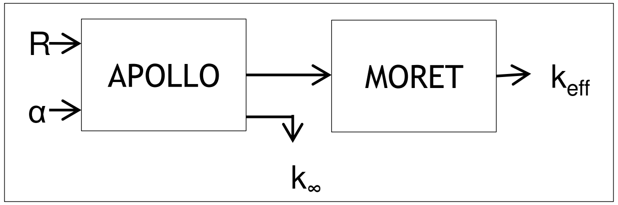
Metamodels Comparaison

- Comparaison of performance of prediction
  - Test functions and design randomly generated
  - Proportion of times when line-model better than column-model:

	Krig	HyperK	CoAffi	CoConv	CoDoub	CoMark	CoThId
Krig	X	0,9%	0,8%	0,9%	0,8%	0,9%	0,8%
HyperK	98,0%	X	93,9%	93,8%	92,9%	93,8%	93,0%
CoAffi	98,3%	5,3%	X	47,6%	41,4%	48,4%	43,8%
CoConv	98,4%	5,5%	51,5%	X	41,7%	52,3%	43,5%
CoDoub	98,3%	6,2%	57,6%	57,4%	X	58,1%	50,4%
CoMark	98,2%	5,3%	50,6%	46,9%	41,1%	X	43,5%
CoThId	98,3%	6,1%	55,2%	55,8%	48,7%	55,7%	X

Step or Stop (SoS) Algorithm

➚ Introspective variation of EGO



- Kriging switched for introspective metamodel (cokriging or hyperkriging)

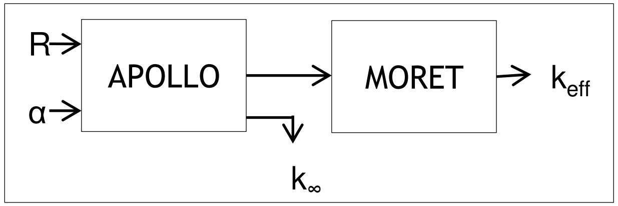
- Timely Expected Improvment (TEI) calculated by dividing EI by time ratio (last-part run time/complete run time)

- When adding a new point:
  1. Running of the first part of the chained code
  2. Re-computation of the updated Expected Improvment
  3. Choise to continue the step (finish the run) if TEI better than best EI or to stop the run.

Step or Stop (SoS) Algorithm

Introspective variation of EGO

- Benefits:
  - Closer to « physics »
  - More precise metamodel
  - Less computation time required



- Promising first results for a toy function

- More comparative work have to be done, specially in industrial cases …

Prospects

- Publication about parameters estimation in progress
- Proper implementation of hyperkriging
- More tests of the SoS algorithm, for toy functions and applicative cases
- Generalization of the SoS algorithm for more sophisticated code chains

Thank you for your attention