



Atelier

“Réflexions sur la problématique des données fonctionnelles dans la planification et l’analyse d’expériences numériques”

B. looss et A. Marrel (CEA Cadarache)

Modérateur : J-C. Fort (Université Toulouse III)

**Rencontres GdR MASCOT NUM
IFP-Lyon**

Données fonctionnelles ?

- Le modèle usuel s'écrit $Y = f(X)$, où Y est une variable de sortie scalaire et X est un vecteur de variables d'entrée scalaires.



X est traité comme un vecteur de variables aléatoires.
Par conséquent, Y est une variable aléatoire.

Données fonctionnelles ?

- Le modèle usuel s'écrit $Y = f(\mathbf{X})$, où Y est une variable de sortie scalaire et \mathbf{X} est un vecteur de variables d'entrée scalaires.



\mathbf{X} est traité comme un vecteur de variables aléatoires.

Par conséquent, Y est une variable aléatoire.

- Le modèle avec variables fonctionnelles s'écrit $Y(v) = f(X_1(u_1), \dots, X_p(u_p))$, où
 - v et u_i sont des paramètres éventuellement multidimensionnels,
 - $Y(v)$ est une fonction de sortie,
 - $X_i(u_i)$ est une fonction d'entrée (éventuellement constante).

Ex. pour u et v : temps t , coordonnées spatiales (x, y, z) , température T , ...

Ex. pour Y et X : $Y(v)$ = concentration d'un polluant en fct du temps v

$X(u)$ = porosité du sol en fct de la coordonnée d'espace u

Les $X_i(u_i)$ sont traités comme des fonctions aléatoires
(processus stochastiques, champs aléatoires, ...).

Par conséquent, $Y(v)$ est une fonction aléatoire.

Atelier « données fonctionnelles »

Données qui apparaissent dans de nombreuses applications.

Solutions génériques à trouver pour étendre l'utilisation des analyses d'incertitudes à ces applications.



Objectifs :

- Initier des discussions et du partage d'expériences,
- Point sur l'état de l'art,
- Dégager des axes de recherche ouverts.

A partir de nos exemples, on illustre

- différentes typologies de données fonctionnelles,
- de nouvelles problématiques dues au caractère fonctionnel des données, et liées à un certain type de problèmes (propagation d'incertitudes, sensibilité, métamodèles, planification, ...).

Plan de cette présentation :²

1. Entrées fonctionnelles
2. Sorties fonctionnelles

Entrées fonctionnelles

On considère le modèle $Y = f(X_1(u_1), \dots, X_p(u_p))$



Une fonction $X_i(u_i)$ peut être :

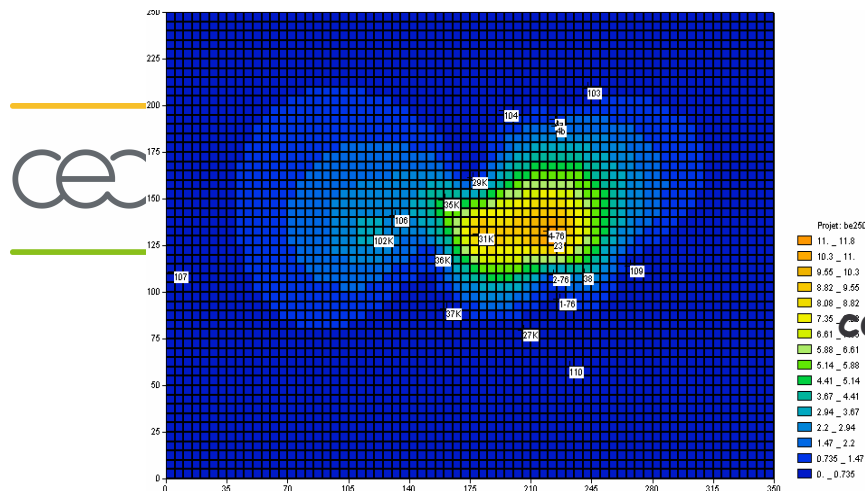
- **Un processus stochastique ou un champ aléatoire issu d'une simulation (gouvernée par un germe).**

- Une fonction résultant d'un scénario donné

- ...

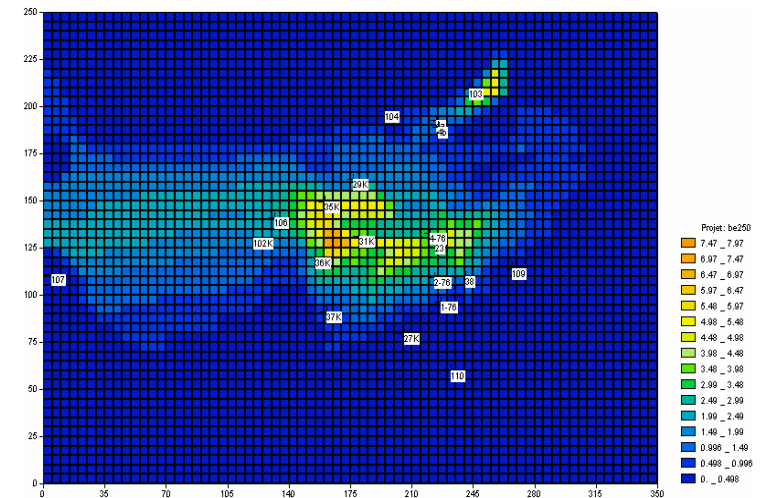
Exemple – Simulations d'écoulement de polluants dans les sols

August 2002



Cartes de concentrations

December 2010



Étude initiale :

- 20 paramètres d'entrée scalaires,
- on s'intéresse à 20 sorties scalaires (concentrations aux piézomètres),
- 300 calculs Monte-Carlo sont réalisés,
- construction de 20 métamodèles (polynôme, boosting trees, krigeage, ...)
- analyse de sensibilité avec les métamodèles.

Une des conclusions :

la perméabilité de la 2ème couche semble influente

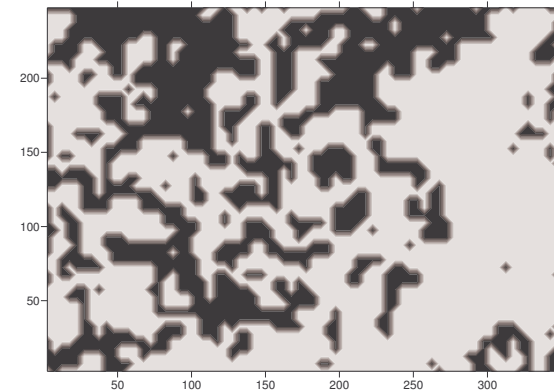
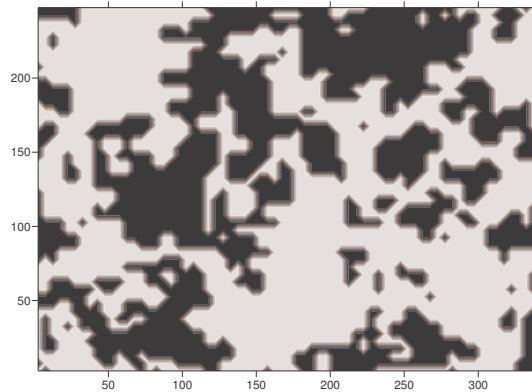
Exemple – Simulations d'écoulement de polluants dans les sols

Nouvelle étude :



- On veut modéliser la perméabilité de manière plus réaliste en prenant en compte son hétérogénéité spatiale,
- Celle-ci est donc un champ aléatoire, dont les réalisations sont obtenues par des méthodes de simulations géostatistiques.

2 réalisations possibles de la perméabilité



- 300 calculs Monte-Carlo sont réalisés,
- Le champ aléatoire n'est pas un paramètre scalaire, les méthodes classiques pour construire un métamodèle ne s'appliquent plus.

Entrées fonctionnelles

On considère le modèle $Y = f(X(u))$



Une fonction $X(u)$ peut être :

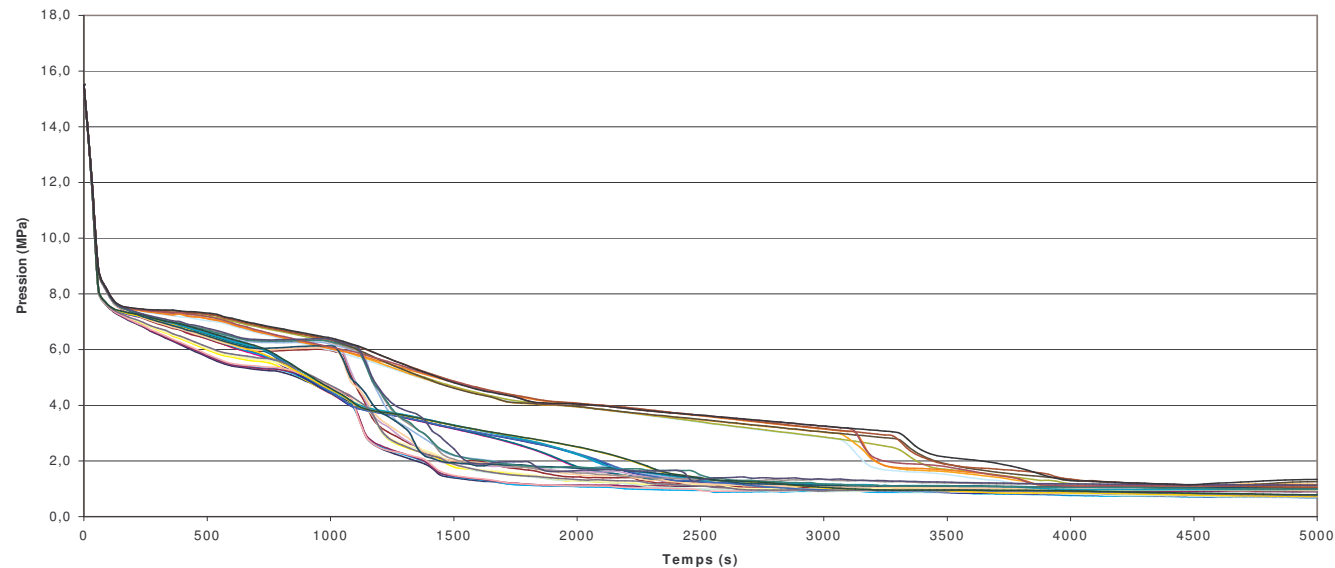
- Un processus stochastique ou un champ aléatoire issus d'une simulation (gouvernée par un germe).

- **Une fonction résultant d'un scénario donné**

- ...

Exemple : choc thermique pressurisée sur la cuve de réacteur

32 profils temporels de la pression $P(t)$



Code de calcul mécanique → facteur de marge à la rupture de la cuve.

7 v.a. d'entrée scalaires + 1 profil

Une propagation d'incertitudes par Monte Carlo peut être réalisée en tirant aléatoirement un profil à chaque simulation.

Quid de la construction d'un métamodèle ?

Quelques problématiques sur les entrées fonctionnelles

Plans d'expériences pour des simulations de processus stochastiques

Ex : LHS de champs aléatoires gaussiens (*Pebesma et al., JASA, 1999*)



Analyse de sensibilité (sans métamodèle) :

- Remplacer la fonction par quelques paramètres d'intérêt (valeur finale, max, moyenne, intégrale, ...)
- Utiliser la discrétisation complète de la fonction ?!?!?
- Décomposer la fonction dans une base appropriée (peut s'avérer difficile dans certains cas, par ex. bruit blanc).
- **Considérer la fonction comme un seul paramètre** (possible pour les méthodes qui utilisent des tirages indépendants des entrées,
Ex : indices de Sobol estimés par Monte-Carlo simple, Jacques, RESS, 2006).
- ...

Problématiques liées aux métamodèles

- Remplacer la fonction par un paramètre scalaire qui correspond à un numéro de scénario (*Ruffo et al., RESS, 2006*)

→ *variable qualitative en entrée du métamodèle*



- **Décompositions dans des bases de fonctions** - Régression fonctionnelle.
Cela peut s'avérer difficile dans certains cas car les processus en entrée peuvent être particulièrement complexes.

- Si on a des entrées scalaires \mathbf{X} et des entrées fonctionnelles $\varepsilon(u)$:
modélisation de la moyenne $E_{\varepsilon}(Y|\mathbf{X})$ et de la variance $Var_{\varepsilon}(Y|\mathbf{X})$.

Modélisation duale, modélisation jointe (GLM, GAM), Zabalza (2000).

ε est considéré comme un paramètre incontrôlable,

La propagation d'incertitudes et l'analyse de sensibilité peuvent être réalisés à l'aide de ce double métamodèle.

- ...

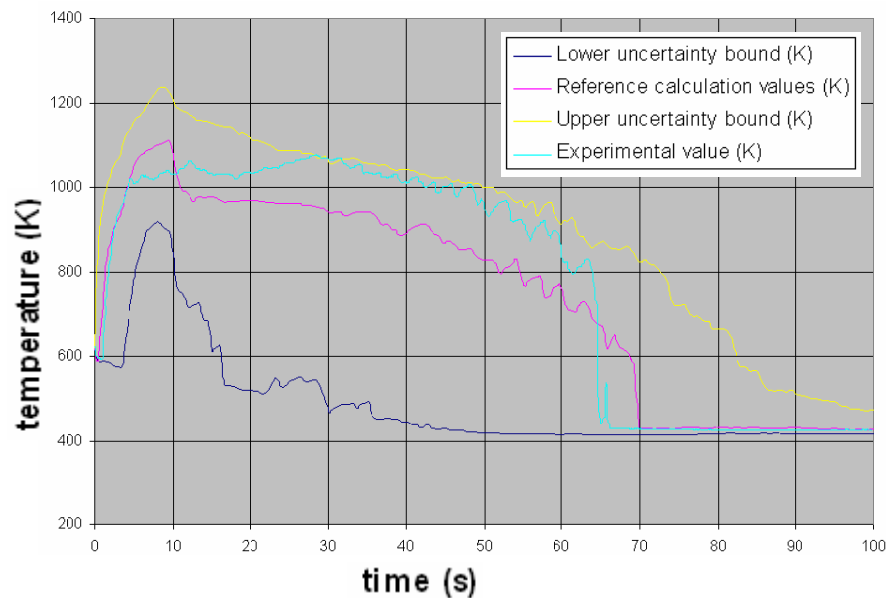
Sorties fonctionnelles

On considère le modèle $Y(v) = f(X)$



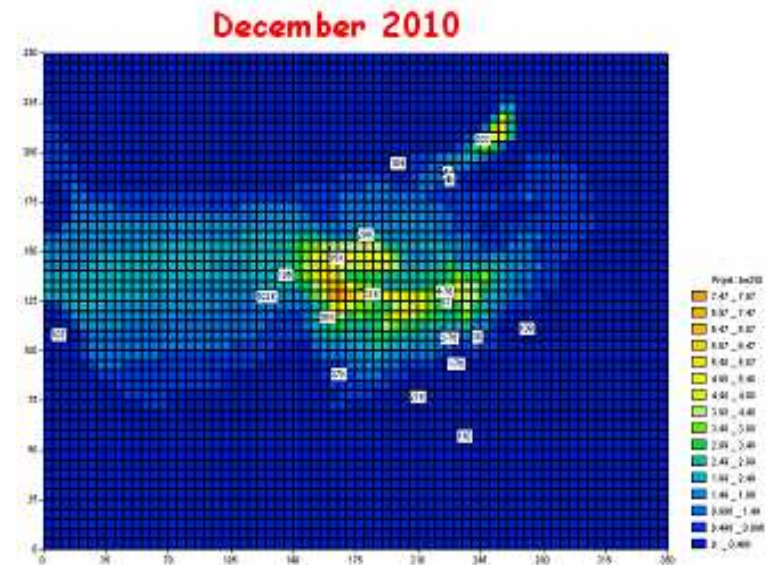
Sortie temporelle

Ex : évolution de la température d'une gaine de combustible, ...



Sortie spatiale

Ex : Concentration en polluant sur un site, ...



Un exemple : Simulation de régimes accidentels

Analyses d'incertitudes et de sensibilité sur les calculs thermohydrauliques de sûreté des réacteurs nucléaires.

CEA Grenoble/LDAS (A. de Crecy, P. Bazin)

Scénario :

Accident de Perte de Réfrigérant Primaire due à une Grosse Brèche

Variable d'intérêt :

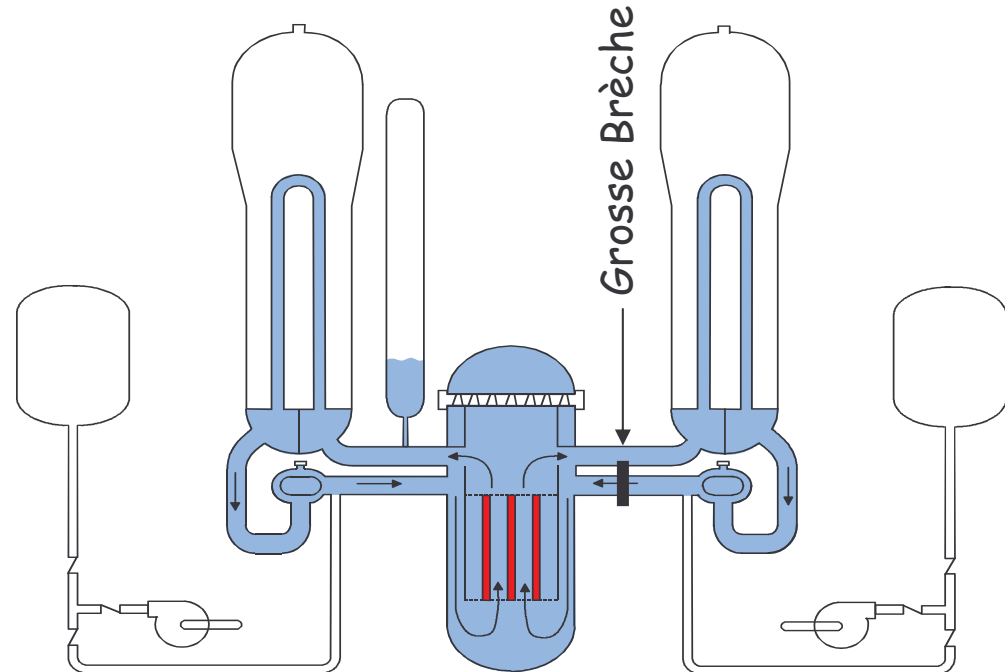
température de la gaine

Critère de sûreté :

Non dépassement d'une valeur critique

Variables explicatives:

Paramètres matériaux, ...

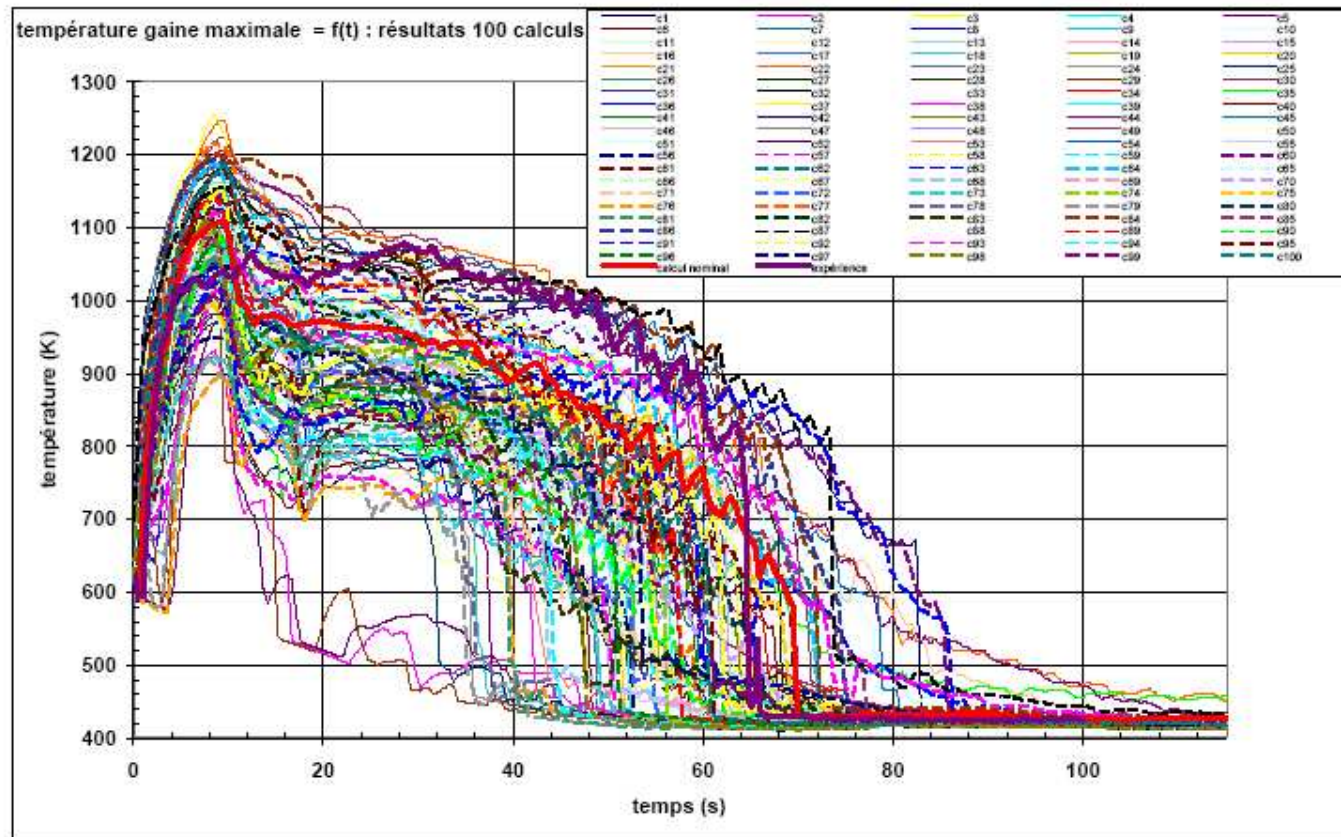


Étude CEA :

Simulation avec le logiciel CATHARE

100 calculs Monte Carlo

53 variables d'entrée (scalaires) aléatoires

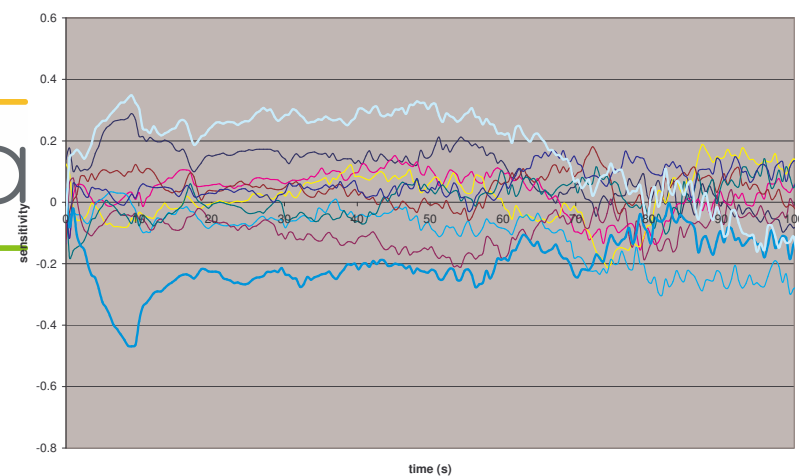


Évolution de la température maximale de la gaine du combustible

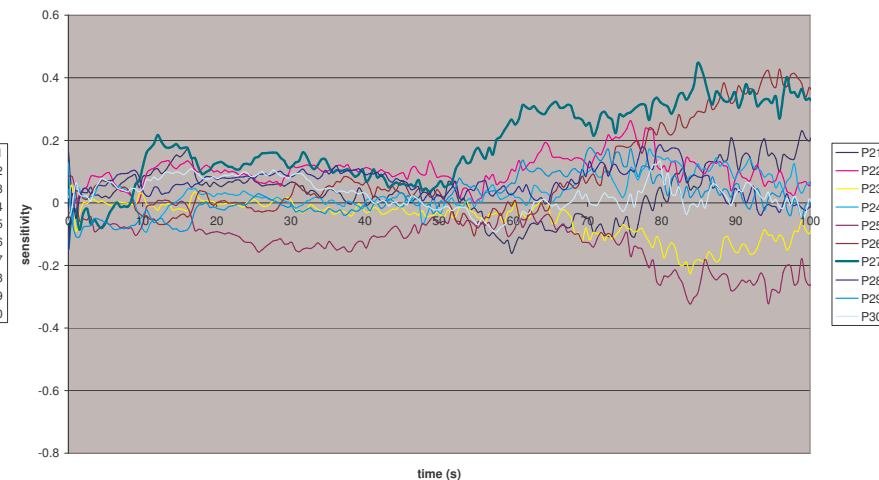
Analyse de sensibilité temporelle (SRRC)



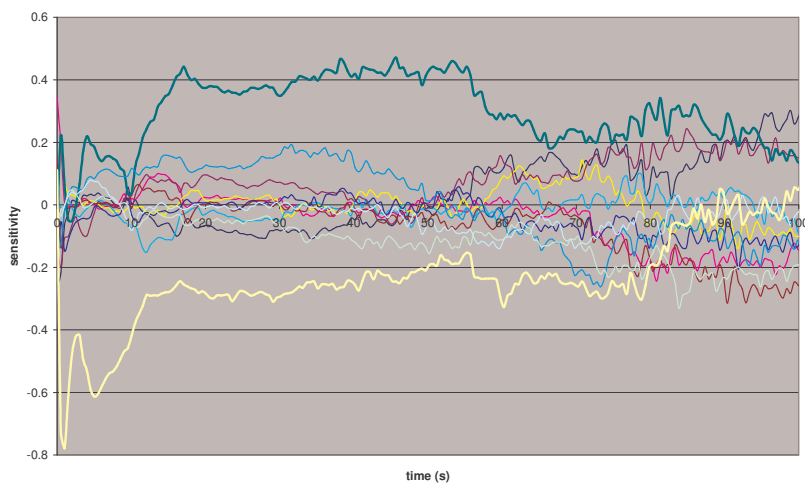
CEA: Max_TC: sensitivities of P11 to P20



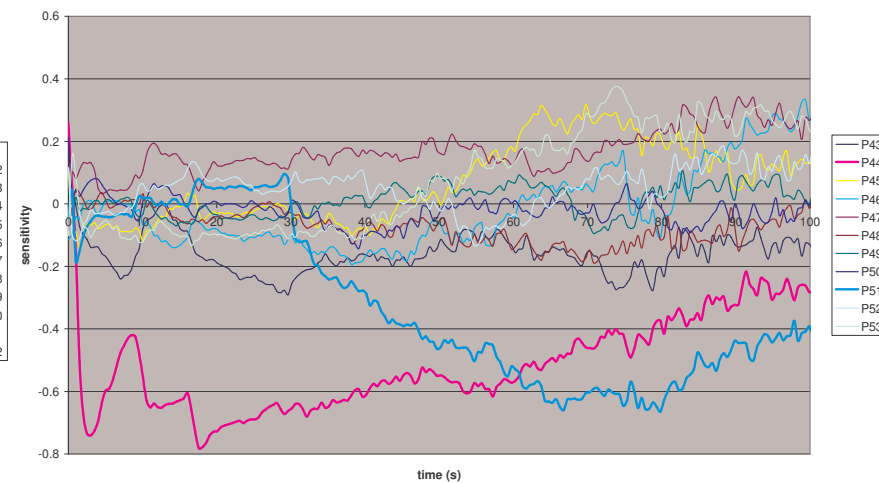
CEA: Max_TC: sensitivities of P21 to P30



CEA: Max_TC: sensitivities of P31 to P42



CEA: Max_TC: sensitivities of P43 to P53



Sorties fonctionnelles multidimensionnelles

- Concentration spatio-temporelle en ^{90}Sr sur un site pollué
- Objectif: Estimer l'impact de la contamination sur l'environnement



Étude CEA :

Développement d'un scénario de transport généralisé de ^{90}Sr sur le site entre 2002 et 2010 avec le logiciel MARTHE

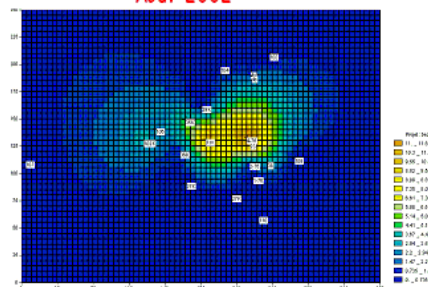
20 paramètres :

Perméabilité, dispersivité, K_d , intensités d'infiltration, ...

Sortie :

Cartes spatio-temporelles de concentrations

Carte de concentration initiale
Août 2002



Jeu de valeurs
pour les 20 paramètres d'entrée

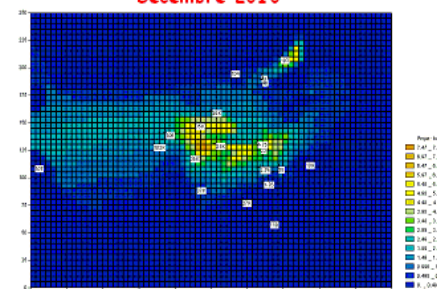
per1	perz3	dt1	kd3
per2	perz4	dt2	poros
per3	d1	dt3	i1
perz1	d2	kd1	i2
perz2	d3	kd2	i3

300 simulations LHS

MARTHE



Carte de concentration finale
Décembre 2010



Problématiques des réponses fonctionnelles

■ Utilisation de la discrétisation complète de la fonction ?!?!

- Analyse de sensibilité, propagation d'incertitudes, ou construction d'un métamodèle à réaliser en chaque point de discrétisation

⇒ Possible mais très coûteux !

⇒ Il faut **synthétiser l'information**

■ Remplacer la fonction par quelques paramètres d'intérêt (valeur finale, max, ...)

- Exploitation et interprétation réduites
- Choix des paramètres d'intérêt fortement lié à la problématique de départ

■ Décomposition dans des bases adaptées et interprétables

- Modélisation et études de sensibilité réalisées sur les coefficients de cette décomposition

■ Analyse de sensibilité : utilisation de l'ACP (Campbell, RESS, 2006)

Remarque : Préalablement aux analyses et aux choix des bases,
une **phase de classification de courbes** peut s'avérer nécessaire.

Problématiques liées au métamodèle



■ Création d'un métamodèle par régression fonctionnelle

- Étude et application des travaux sur les réponses fonctionnelles

■ Krigeage / Cokrigeage (Santner et al., 2003, Fang et al., 2006)

- Envisageable pour un nombre restreint de variables de sorties (ex : faible échelle de discrétisation, ...)
- Possible avec des réductions matricielles dans le cas de grilles de discrétisation uniformes

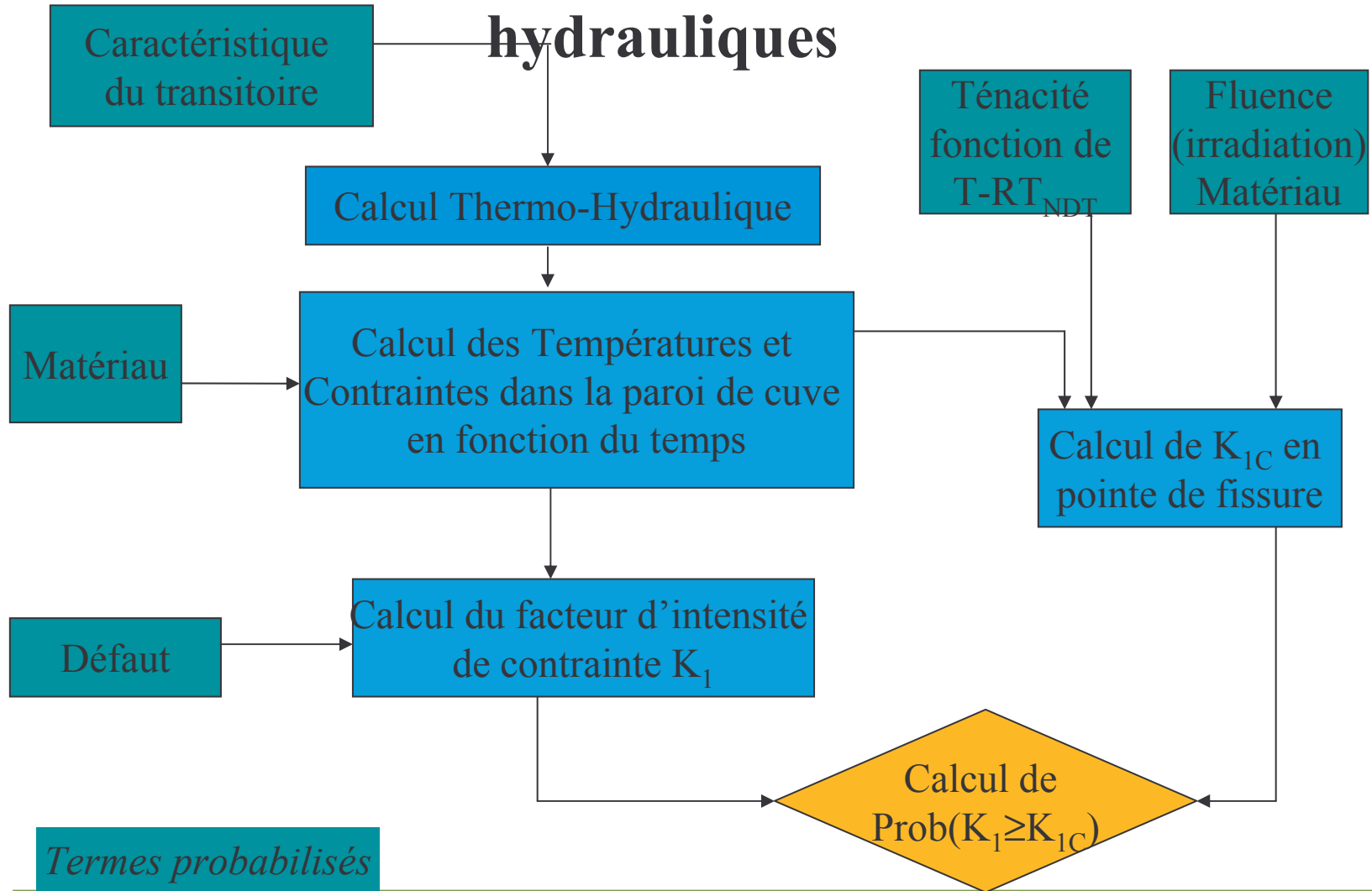
■ Planification adaptative ?

■ Cas de plusieurs réponses fonctionnelles non indépendantes?



ANNEXES

Nouvelle approche avec prise en compte des incertitudes dans les transitoires thermo-



La modélisation jointe par 2 GLM interdépendants

[Nelder (1989), Smyth (1999), Zabalza (2000)]



moyenne

$$\begin{cases} E(Y_i) = \mu_i, \eta_i := g(\mu_i) = \sum_j x_{ij} \beta_j \\ \text{Var}(Y_i) = \phi_i V(\mu_i) \end{cases}$$

dispersion

$$\begin{cases} E(d_i) = \phi_i, \zeta_i := h(\phi_i) = \sum_j u_{ij} \gamma_j \\ \text{Var}(d_i) = \tau V_D(\phi_i) \end{cases}$$

Remarques :

- ϕ n'est plus supposée constante, mais est un GLM ;
- pour la dispersion d , on prend la contribution à la déviance ;
- les variables explicatives u_j sont prises parmi les variables x_j ;
- on choisit $h(\phi) = \log(\phi)$ (fonction lien positive pour la dispersion) ;
- $\text{Var}(d) = 2\phi^2$ car d est proche d'une distribution *Gamma*.

Avantages / modélisation duale :

- dans un contexte hétéroscédastique, il est important de modéliser simultanément la moyenne $E(Y)$ et la variance $\text{Var}(Y)$;
- le plan d'expériences ne nécessite pas de répétitions (bcp - coûteux)

Scénario schématique de l'accident réacteur « grosse brèche »

