



HAL
open science

Revue sur l'analyse de sensibilité globale de modèles numériques

Bertrand Iooss

► **To cite this version:**

Bertrand Iooss. Revue sur l'analyse de sensibilité globale de modèles numériques. 2010. hal-00503179v1

HAL Id: hal-00503179

<https://hal.science/hal-00503179v1>

Preprint submitted on 23 Jul 2010 (v1), last revised 7 Dec 2010 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

REVUE SUR L'ANALYSE DE SENSIBILITÉ GLOBALE DE MODÈLES NUMÉRIQUES

Bertrand Iooss

Résumé. — Cet article a pour objectif d'effectuer un survol rapide, mais dans un cadre méthodologique relativement complet, des différentes méthodes d'analyse de sensibilité globale d'un modèle numérique. Faisant appel de nombreux outils statistiques (régression, lissage, tests, apprentissage, techniques de Monte Carlo, ...), celles-ci permettent de déterminer quelles sont les variables d'entre d'un modèle qui contribuent le plus une quantité d'intérêt calculée à l'aide de ce modèle (par exemple la variance d'une variable de sortie). Trois grandes classes de méthodes sont ainsi distinguées : le criblage (tri grossier des entrées les plus influentes parmi un grand nombre), les mesures d'importance (indices quantitatifs donnant l'influence de chaque entrée) et les outils d'exploration du modèle (mesurant les effets des entrées sur tout leur domaine de variation). Une méthodologie progressive d'application de ces techniques est illustrée sur une application vocation pédagogique. Une synthèse est alors formulée afin de situer chaque méthode selon trois axes : coût en nombre d'évaluations du modèle, complexité du modèle et type d'information apportée.

Abstract (Review of global sensitivity analysis of numerical models)

This paper makes a brief review, in a relatively complete methodological framework, of various global sensitivity analysis methods of model output. Numerous statistical and probabilistic tools (regression, smoothing, tests, statistical learning, Monte Carlo, ...) aim at determining the model input variables which mostly contribute to an interest quantity depending of model output (as the variance of an output variable). Three kind of methods are distinguished: the screening (coarse sorting of the most influential inputs among a large number), the measures of importance (quantitative sensitivity indices) and the deep exploration of the model behaviour (measuring the effects of inputs on their all variation range). A progressive application methodology is illustrated on a scholar application. A synthesis is given to place every method according to three axes: cost in number of model evaluations, model complexity and nature of brought information.

Mots clefs. — Code de calcul ; Expérience numérique ; Incertitude ; Modèle ; Plan d'expériences.

Table des matières

1. Introduction.....	2
2. Mthodes de criblage.....	5
3. Dfinitions de mesures d'importance.....	10
4. Exploration du modle.....	16
5. Synthse et conclusion.....	20
6. Remerciements.....	24
Références.....	24

1. Introduction

Lors de la construction et de l'utilisation d'un modle numérique de simulation, les méthodes d'analyse de sensibilité (AS) sont des outils précieux. Elles permettent notamment de déterminer quelles sont les variables d'entre du modle qui contribuent le plus une quantité d'intért donne en sortie du modle, quelles sont celles qui n'ont pas d'influence et quelles sont celles qui interagissent au sein du modle. Pour un ingnieur, l'intrt est indniable car les rsultats d'une AS peuvent lui permettre de simplifier son modle, de mieux l'apprehender, voire de le vrifier. Au final, l'AS est une aide la validation d'un code de calcul, l'orientation des efforts de R&D, ou encore la justification en terme de sreté du dimensionnement d'un systme. Dans la suite, nous nommerons "entre" et "sortie" les variables d'entre et les rponses du modle numrique.

Le calcul des sensibilités d'un modle a longtemps t bas sur l'évaluation des répercussions sur les valeurs des sorties de petites perturbations des valeurs des entrées, et ce autour d'un point nominal. Cette approche déterministe consiste calculer ou estimer des indices basés sur les dérivées partielles du modle en un point précis. Grce aux approches adjointes permettant de traiter rapidement des modles grand nombre de variables d'entrée, ce type d'approches est couramment utilisé dans la résolution de gros systmes environnementaux (climatologie, océanographie, hydrogéologie, cf. Cacuci [5], Castaings [7]).

A partir de la fin des années 1980, pour pallier aux limites des méthodes locales (hypothses de linéarité et de normalité, variations locales), une nouvelle gnration de mthodes s'est dvelopp dans un cadre statistique. Par opposition aux méthodes locales, elles ont été dénommées "globales" car elles s'intéressent l'ensemble du domaine de variation possible des variables d'entrée (Saltelli et al. [50]). Celles-ci suscitent un intrt croissant de la part des modlisateurs et

des utilisateurs de modèles numériques, qui souhaitent pleinement profiter des outils de simulation numérique (cf. Helton [21], De Rocquigny et al. [12] et Volkova et al. [69] pour des exemples d'applications industrielles). Des premières synthèses sur le sujet ont ainsi récemment vu le jour (Kleijnen [30], Frey & Patil [18], Helton et al. [22], Badaea & Bolado [2]). Malheureusement, entre les heuristiques, les outils graphiques, la théorie des plans d'expériences, les techniques de type Monte Carlo, les méthodes basées sur un apprentissage statistique, etc., les novices du domaine peuvent se retrouver rapidement perdus quant au choix de la méthode la plus adaptée à leur problème. Dans cet article, nous souhaitons proposer une lecture différente des synthèses classiques en mettant en perspective les techniques d'AS au sein d'un cadre méthodologique applicatif. Cette classification des différentes méthodes a pour ambition de permettre une introduction plus aisée au sujet pour les lecteurs non spécialistes.

L'entrée du modèle est noté $\mathbf{X} = (X_1, \dots, X_d) \in \mathbb{R}^d$. Nous nous restreignons à l'étude d'une sortie scalaire $Y \in \mathbb{R}$ du modèle :

$$(1) \quad Y = f(\mathbf{X})$$

où $f(\cdot)$ représente le code de calcul. Dans le cadre probabiliste, \mathbf{X} est un vecteur aléatoire défini par une loi de probabilité et Y devient une variable aléatoire. La question de la dépendance entre les composantes de \mathbf{X} ne sera pas vue dans ce papier (cf. Kurowicka & Cooke [34] pour une introduction à cette question), c'est-à-dire que les entrées X_i ($i = 1 \dots d$) sont supposées indépendantes. Enfin, cette revue se focalise sur les AS relatives à la variabilité globale de la sortie, mesurée le plus souvent par sa variance. Si la quantité d'intérêt est différente, par exemple l'entropie de la sortie ou la probabilité de dépasser un seuil, d'autres techniques seront nécessaires (Saltelli et al. [50], Frey & Patil [18]).

Nous illustrons notre propos à l'aide d'un exemple simple concernant la simulation de la hauteur de crues d'une rivière, comparée à la hauteur d'une digue qui protège un site industriel (Fig. 1). Ce modèle, utilisé comme exemple pédagogique et benchmark (De Rocquigny [11], Arnaud et al. [1]), est ici légèrement remanié afin de mieux illustrer notre propos.

Le modèle de crues, simplification des équations hydrodynamiques de Saint-Venant en 1D avec un débit constant et uniforme, consiste en une équation qui fait intervenir les caractéristiques du tronçon de cours d'eau situé en amont du site industriel :

$$(2) \quad S = Z_v + H - H_d - C_b \quad \text{avec} \quad H = \left(\frac{Q}{BK_s \sqrt{\frac{Z_m - Z_v}{L}}} \right)^{0.6},$$

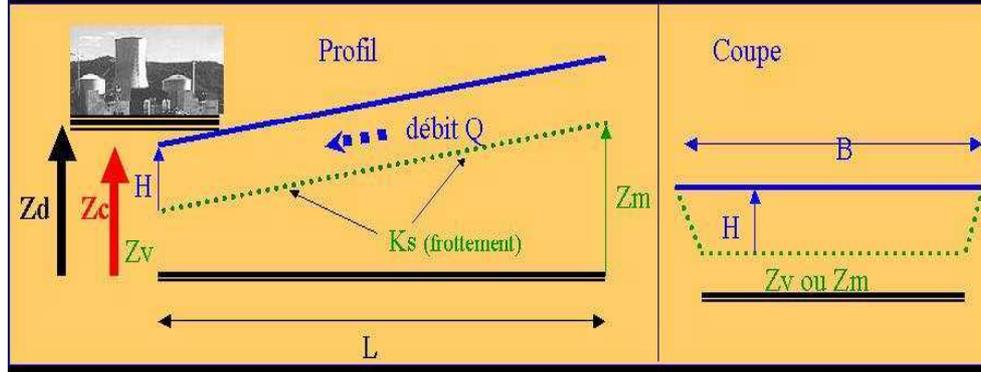


FIGURE 1. Exemple crues : modle simplifi d'une rivire et de la protection d'un site industriel l'aide d'une digue.

avec

- S la surverse (en m), sortie du modle ;
- H la hauteur d'eau maximale annuelle (en m) ;
- Q le dbit maximal annuel (en m^3/s), entre alatoire de loi Gumbel max $G(1013, 558)$ tronque infrieurement en 500 et suprieurement en 3000 ;
- K_s le coefficient de Strickler, entre alatoire de loi normale $N(30, 8)$ tronque infrieurement en 15 ;
- Z_v la cte en aval du fond du cours d'eau (en m), entre alatoire de loi $T(49, 50, 51)$, o $T(\min, \text{mode}, \max)$ est la loi triangulaire ;
- Z_m la cte en amont du fond du cours d'eau (en m), entre alatoire de loi $T(54, 55, 56)$;
- H_d la hauteur de la digue (en m), entre alatoire de loi $U(7, 9)$ o $U(\min, \max)$ est la loi uniforme ;
- C_b la cte de la berge (en m), entre alatoire de loi $T(55, 55.5, 56)$;
- L la longueur du tronon du cours d'eau (en m), entre alatoire de loi $T(4990, 5000, 5010)$;
- B la largeur du cours d'eau (en m), entre alatoire de loi $T(295, 300, 305)$;

Nous nous intrissons galement une autre sortie, le cot annuel d'entretien de la digue C_p (en Meuros), et dfinie par la formule suivante :

$$(3) \quad C_p = \mathbb{1}_{S>0} + \left\{ 0.2 + 0.8 \left[1 - \exp\left(-\frac{1000}{S^4}\right) \right] \right\} \mathbb{1}_{S \leq 0} + \frac{1}{20} (H_d \mathbb{1}_{H_d > 8} + 8 \mathbb{1}_{8 \leq H_d}) ,$$

avec $\mathbb{1}_A(x)$ la fonction indicatrice qui vaut 1 pour $x \in A$ et 0 sinon. Dans cette quation, le premier terme est le cot d une inondation ($S > 0$), le second un cot

d'entretien du lit de la riviere dans le cas o il n'y a pas d'inondation ($S \leq 0$), et le troisieme le cot d'investissement li la digue (constant pour une hauteur de digue infrieure 8 m et croissant au-del).

La section suivante traite des mthodes dites de criblage, mthodes qualitatives d'AS permettant de traiter des modles plusieurs dizaines de variables d'entre. Les mesures d'influence quantitatives les plus utilises dans la littrature sont dcrites dans la troisieme section, alors que la quatrieme section passe en revue les techniques avances d'exploration du comportement de la sortie d'un modle vis--vis de ses entres. La derniere section offre une classification de ces mthodes et évoque quelques problmes ouverts en AS.

2. Mthodes de criblage

Les mthodes de criblage ("screening") permettent d'explorer rapidement le comportement des sorties d'un code de calcul coteux en faisant varier un grand nombre de ses entres (typiquement plusieurs dizaines voire plusieurs centaines). Toutes sont bases sur la discratisation des entres en plusieurs valeurs nommes niveaux, ce qui en fait des mthodes dites dterministes qui n'utilisent pas de lois de probabilit sur les entres. Pour les entres de notre exemple crues dont les lois de probabilit ne sont pas support born, nous prenons les quantiles 5% et 95% comme valeurs minimale et maximale. Nous distinguons trois classes de mthodes.

2.1. Criblage trs grande dimension. — Certaines techniques issues de la pratique des plans d'experiences permettent de raliser un criblage en ralisant moins de calculs qu'il n'y a d'entres. Celles-ci supposent qu'il n'y a pas d'interaction entre les entres, que la variation de la sortie est monotone par rapport chaque entre et que le nombre des entres influentes est trs faible devant le nombre total d'entres (de l'ordre d'une sur dix). Il s'agit en premier lieu des *plans supersatur*s dvelopps dans le contexte de la planification d'experiences relles (Satterthwaite [56], Lin [37]). L'un des plans supersatur les plus connus rsulte de la division en deux parties (l'aide d'une colonne de branchement) d'une matrice d'Hadamart. Dans un contexte de criblage pour codes de calcul, Claeys-Bruno et al. [8] ont montr sur quelques applications que ce plan supersatur est l'un des plus robustes et qu'il faut au moins cinq fois plus de calculs que d'entres influentes pour les identifier toutes.

D'autres approches sont particulirement bien adaptes aux experiences numeriques car elles sont squentielles et adaptatives, c'est--dire qu'elles dfinissent

une nouvelle expérience réaliser en fonction des résultats des précédentes. On ne sait donc pas *a priori* combien elles vont nécessiter d'expériences. La technique du *criblage par groupe* (Dean & Lewis [13]) consiste à créer un certain nombre de groupes d'entres et à identifier les plus influents. En répétant progressivement l'opération en conservant les groupes influents, on extrait au final les entres influentes. Cette technique nécessite la connaissance du sens de variation de la sortie en fonction du sens de variation de chaque entre, connaissance qui n'est pas toujours disponible.

Enfin, la méthode des *bifurcations squentielles* de Bettonvil & Kleijnen [3] est une méthode de criblage par groupe avec deux groupes. C'est une approche dichotomique où on tente d'éliminer un groupe de variables à l'issue de chaque nouveau calcul (pour lequel les entres d'un groupe sont fixées à un niveau et les entres de l'autre groupe sont mises à un niveau différent). Comme pour le criblage par groupe, son coût dépend donc du nombre de variables influentes, mais aussi de la stratégie de classement, *i.e.* de notre capacité à suspecter quelles sont les entres influentes afin de les rassembler au sein d'un même groupe.

En appliquant cette technique en aveugle sur l'exemple crues (Eq. (2), modèle 8 entres), en supposant que le modèle est monotone par rapport à chaque entre et qu'il n'y a pas d'interaction, on obtient par exemple après 4 évaluations (permettant de créer au final 3 groupes) du modèle pour chaque sortie :

- sortie S : les groupes (Q, K_s) , (Z_v, Z_m) et (H_d, C_b, L, B) ont des influences peu près égales, mais la première étape de la méthode a permis de voir que le groupe (Q, K_s, Z_v, Z_m) est plus influent que (H_d, C_b, L, B) ;
- sortie C_p : (Z_v, Z_m) est le groupe le plus influent, (Q, K_s) est un groupe moyennement influent, (H_d, C_b, L, B) est un groupe d'influence négligeable.

Comme on le verra plus tard, ces conclusions sont fausses car les hypothèses inhérentes à la méthode (monotonie et non interaction) ne sont pas respectées. Cette technique donne donc des résultats grossiers et nécessite des connaissances et certaines hypothèses fortes vis-à-vis du modèle étudié. Elle a été illustrée ici sur un modèle à 8 entres mais est plutôt adaptée à des modèles possédant plusieurs dizaines d'entres.

Dans un contexte de criblage pour codes de calcul, Sergent et al. [58] comparent les plans supersaturés, le criblage par groupe et les bifurcations squentielles et en concluent que la technique des plans supersaturés est nettement plus risquée que les autres mais nécessite le moins d'hypothèses. En effet, il est nécessaire de connaître le sens de variation de la sortie par rapport à chaque entre pour pouvoir appliquer les méthodes des bifurcations squentielles et du criblage

par groupe. Par ailleurs, les méthodes adaptatives ne peuvent pas traiter plusieurs sorties la fois, contrairement aux plans supersaturés.

2.2. Plans d'expériences usuels. — La deuxième classe de méthodes concerne celle issue de la théorie classique des plans d'expériences (cf. par exemple Montgomery [41] ou Dreesbeke et al. [15]). Un *plan factoriel complet* consiste à valuer le code de calcul pour toutes les combinaisons des niveaux des entres, ce qui permet l'estimation de tous les effets des entres et de leurs interactions. En pratique, le nombre de simulations requis rend ce plan impraticable au del d'une dizaine d'entres. En effet, il nécessite 2^d calculs si on suppose que le modèle est monotone en travaillant avec deux niveaux pour chaque entre.

Si l'on veut estimer de manière non biaisée les effets du premier ordre (aussi appelés effets principaux) de chaque entre, il faut au minimum disposer d'évaluations du modèle correspondant $n \geq d + 1$ combinaisons des entres. Par définition, un *plan de résolution trois* (noté RIII) permet cette estimation non biaisée en supposant que les effets des interactions sont nuls, *i.e.* que le modèle est de la forme

$$(4) \quad Y = \sum_{j=0}^d \beta_j X_j + \epsilon ,$$

où $X_0 = 1$, $\boldsymbol{\beta} = (\beta_0, \dots, \beta_d)^t \in \mathbb{R}^{d+1}$ est le vecteur des effets des entres et $\epsilon \in \mathbb{R}$ est l'erreur du modèle. L'estimation de ces effets se fait par la méthode des moindres carrés ordinaires.

Le plan d'expériences le plus simple, encore très utilisé, est le plan nommé "*One At a Time*" (OAT), qui fait partie de la classe des plans RIII. Le plan OAT consiste à changer le niveau d'une entre à la fois, en utilisant deux ou trois niveaux par facteur (Kleijnen [31]). Avec deux niveaux, ce plan requiert donc exactement $n = d + 1$ calculs, mais ne permet pas de maîtriser la précision que l'on a sur les estimations des effets. Saltelli & Annoni [48] proposent une critique sévère du plan OAT. Il faut noter que l'utilisation de ce plan obéit à d'autres règles que celles de l'AS : le plan OAT est souvent utilisé au préalable par les utilisateurs pour une vérification systématique et manuelle de l'effet de chaque entre (détection de bugs ou d'anomalies).

Pour l'AS, une voie plus raisonnable qu'un simple plan OAT consiste à minimiser la variance des effets estimés, ce qui est l'un des objectifs de la théorie des plans d'expériences. Celle-ci se concentre sur les plans orthogonaux, c'est-à-dire

ceux qui satisfont

$$(5) \quad (\mathbf{X}_0^n)^t \mathbf{X}_0^n = n \mathbf{I}_{d+1},$$

où $\mathbf{X}_0^n = (X_j^{(i)})_{i=1..n, j=0..d}$ est la matrice du plan et \mathbf{I}_{d+1} est la matrice identité de dimension $d + 1$. Une classe bien connue de plans orthogonaux est celle des *plans factoriels fractionnaires*. Leur construction, qui fait appel à la notion d’alias, consiste à confondre des interactions que l’on soupçonne non actives avec des effets principaux (Droesbeke et al. [15]).

Il est parfois prudent de supposer que les interactions entre les entres peuvent avoir des effets importants. Par définition, un *plan de résolution quatre* (not RIV) permet une estimation non biaisée des effets principaux même si des interactions d’ordre deux sont présentes. Un plan RIV peut être construit en superposant un plan RIII avec son plan “miroir”. Pour un coût en terme de nombre de calculs de l’ordre de $2d$, un plan RIV permet donc d’identifier les effets principaux des entres pour des modèles avec interactions. Il existe de nombreux autres types de plans qui assouplissent les hypothèses des plans RIII tout en conservant un nombre de calculs raisonnable (Montgomery [41]).

2.3. La méthode de Morris. — La *méthode de Morris* (Morris [42], Saltelli et al. [50]) consiste à répéter r fois ($r = 5 \dots 10$) un plan OAT aléatoirement dans l’espace des entres, en discrétisant chaque entre en un nombre convenable de niveaux (dépendant du nombre r de répétitions que l’on veut faire). “Aléatoirement” signifie que l’on tire aléatoirement le point de départ de l’expérience OAT et la suite de directions pour lesquelles on évalue séquentiellement les nouvelles expériences.

La méthode de Morris permet ainsi de s’extraire des hypothèses limitatives du plan OAT en classant les entres selon trois catégories :

- entres ayant des effets négligeables,
- entres ayant des effets linéaires et sans interaction,
- entres ayant des effets non linéaires et/ou avec interactions (sans distinction de ces deux types d’effets).

Chaque répétition i ($i = 1 \dots r$) permet d’évaluer un effet élémentaire $E_j^{(i)}$ (accroissement du modèle entre deux points successifs) par entre X_j . L’ensemble du plan d’expériences (r répétitions) fournit un r -chantillon des effets pour chaque entre X_j , dont sont issus les indices de sensibilité $\mu_j^* = \sum_{i=1}^r |E_j^{(i)}|$ (moyenne des valeurs absolues des effets) et σ_j (carte-type des effets). Ainsi, plus μ_j^* est importante, plus l’entre X_j contribue à la dispersion de la sortie. σ_j mesure quant à lui la linéarité du modèle étudié. En effet, si la sortie dépend linéairement de

X_j et que X_j n'interagit pas avec d'autres entres X_k ($k \neq j$), l'effet d'une perturbation linéaire de X_j est identique quelle que soit sa position dans l'espace des entres (donc aussi de la valeur des autres entres) : les r effets linéaires sont gaussiens et σ_j est alors égal à 0. Par conséquent, plus σ_j est élevé (par rapport à μ_j^*), moins l'hypothèse de linéarité et de non interaction est pertinente.

La méthode de Morris est appliquée sur l'exemple crués (Eq. (2)) avec $r = 5$ répétitions, ce qui nécessite $n = r(p + 1) = 45$ évaluations du modèle. La Figure 2 trace les résultats selon des coordonnées (μ_j^*, σ_j) . Cette visualisation est particulièrement adaptée pour interpréter les résultats :

- sortie S : K_s , Z_v , Q , C_b et H_d sont des entres influentes, alors que les autres entres ont peu d'influence. D'autre part, la sortie du modèle dépend des entres de manière linéaire et ces entres n'interagissent pas (car pour toutes ces entres j : $\sigma_j \ll \mu_j^*$).
- sortie C_p : H_d , Q , Z_v et K_s sont des entres fortement influentes avec des effets non linéaires et/ou d'interactions (σ_j et μ_j^* sont du même ordre de grandeur). C_b a une influence moyenne. Les autres entres n'ont que peu d'influence.

Au final, cette phase de criblage permet d'identifier les entres L , B et Z_m , qui n'ont que très peu d'influence sur les deux sorties étudiées. Ces 3 entres sont fixés dans la suite de ce papier à leur valeur nominale.

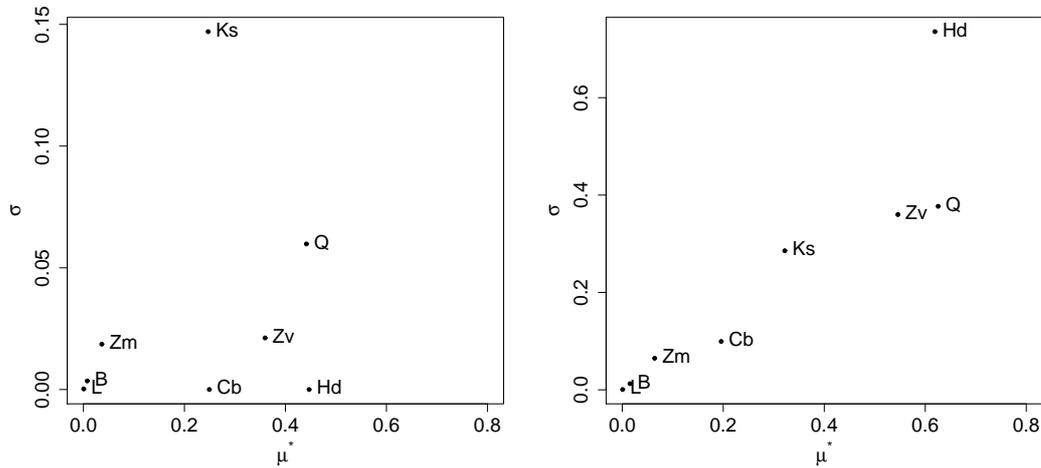


FIGURE 2. Résultats de la méthode de Morris ($r = 5$ avec 4 niveaux) : sorties S (gauche) et C_p (droite).

3. Dfinitions de mesures d'importance

S'il est possible d'obtenir un chantillon de simulations $(\mathbf{X}^n, \mathbf{Y}^n) = (X_1^{(i)}, \dots, X_d^{(i)}, Y_i)_{i=1..n}$ de taille suffisante, il est ais d'obtenir des indices de la sensibilit de la sortie par rapport aux entres en appliquant les techniques de rgression linaire, de rgression sur les rangs, des tests statistiques, voire de dcomposition de la variance fonctionnelle. Nous parlons prsent de mesures d'importance car ces techniques permettent une hirarchisation prcise de l'influence sur la sortie de toutes les entres, contrairement aux techniques de criblage qui ont pour but de dtecter les entres non influentes. Nous distinguons nouveau trois classes de mthodes.

3.1. Mthodes bases sur la rgression linaire. — Les principales mesures d'importance que l'on classe dans cette catgorie sont les suivantes :

◇ le *coefficient de corrlation linaire*, not $\rho(\cdot, \cdot)$, entre X_j et Y :

◇ le *coefficient de rgression standard* (not $\text{SRC}(\cdot, \cdot)$) :

$$(6) \quad \text{SRC}_j = \text{SRC}(X_j, Y) = \beta_j \sqrt{\frac{\text{Var}(X_j)}{\text{Var}(Y)}} ,$$

o β_j est le coefficient de rgression de X_j (Eq. (4)) ;

◇ le *coefficient de corrlation partielle* (not $\text{PCC}(\cdot, \cdot)$),

$$(7) \quad \text{PCC}_j = \text{PCC}(X_j, Y) = \rho(Y - \widehat{Y}, X_j - \widehat{X}_j) ,$$

avec \widehat{Y} la prvision du modle linaire dans lequel X_j n'est pas prsent et \widehat{X}_j la prvision du modle linaire qui exprime X_j en fonction des autres entres. Les PCC permettent d'liminer l'influence des autres variables et sont donc adapts au cas o les variables d'entre sont corrls (Saltelli et al. [50]) ;

◇ le *coefficient de corrlation sur les rangs* des variables (nomm coefficient de Spearman et not $\rho^S(\cdot, \cdot)$). Si $\mathbf{R}_X = (R_{X_1}, \dots, R_{X_d})$ est le vecteur des rangs des entres et R_Y est le rang de la sortie, on a :

$$(8) \quad \rho_j^S = \rho^S(X_j, Y) = \rho(R_{X_j}, R_Y) .$$

On calcule ces coefficients aprs avoir transform l'chantillon $(\mathbf{X}^n, \mathbf{Y}^n)$ en un chantillon $(\mathbf{R}_X^n, \mathbf{R}_Y^n)$ en remplaant les valeurs

par leur rang dans chaque colonne de la matrice (Saporta [55]) ;

◇ le *coefficient de rgression standard sur les rangs* (not $\text{SRRC}(\cdot, \cdot)$), pendant du SRC mais partir de l'échantillon $(\mathbf{R}_X^n, \mathbf{R}_Y^n)$:

$$(9) \quad \text{SRRC}_j = \text{SRRC}(X_j, Y) = \text{SRC}(R_{X_j}, R_Y) .$$

◇ le *coefficient de corrlation partielle sur les rangs* (not $\text{PRCC}(\cdot, \cdot)$), pendant du PCC mais partir de l'échantillon $(\mathbf{R}_X^n, \mathbf{R}_Y^n)$:

$$(10) \quad \text{PRCC}_j = \text{PRCC}(X_j, Y) = \text{PCC}(R_{X_j}, R_Y) .$$

En pratique, on effectue tout d'abord une rgression lineaire entre la sortie Y et les entres \mathbf{X} afin de savoir si leur relation est approximativement lineaire. Pour mesurer cette linarite, les outils statistiques classiques sont utilisés, comme par exemple le coefficient de dtermination R^2 (Saporta [55]) et le coefficient de prdictivité Q_2 du modle :

$$(11) \quad Q_2 = 1 - \frac{\sum_{i=1}^m [Y_i^p - \hat{Y}(\mathbf{X}^{p(i)})]^2}{\sum_{i=1}^m (\overline{Y^p} - Y_i^p)^2}$$

o $(\mathbf{X}^{p(i)}, Y_i^p)_{i=1..m}$ est un m-échantillon d'entres-sortie et $\overline{Y^p}$ est la moyenne de $(Y_i^p)_{i=1..m}$. Q_2 correspond au R^2 calcul sur une base de test. Si on juge l'hypothse de linarite acceptable (par exemple si $R^2 > 0.8$, ce qui signifie que plus de 80% de la variabilité de la sortie est expliquée par une relation lineaire), alors les indices de sensibilité Pearson, SRC et PCC sont utilisables. Par ailleurs, si les variables d'entre sont indépendantes, l'ensemble des SRC_j^2 forment une décomposition de la variance de la rponse : chaque SRC_j^2 exprime la part de variance de la rponse expliquée par le facteur X_j . Cette propriété en fait une mesure particulièrement appréciée.

Dans le cas o la relation entre \mathbf{X} et Y n'est pas lineaire mais monotone, les coefficients de corrlation et de rgression basés sur les rangs (Spearman, SRRC, PRCC) peuvent être utilisés. L'hypothse de monotonie doit bien sûr être valide, par exemple l'aide du coefficient de dtermination R^{2*} et du coefficient de prdictivité Q_2^* associé à la rgression lineaire sur les rangs.

Comme pour les plans d'expérience usuels (cf §2.2), ces méthodes basées sur la rgression lineaire nécessitent un échantillon de taille $n \geq d + 1$. En pratique, un échantillon de type Monte Carlo ou quasi-Monte Carlo est souvent utilisé lorsque l'utilisateur réalise, conjointement une AS, une propagation des incertitudes des entres (échantillonnées suivant leur loi de probabilité). L'estimation des indices de sensibilité est alors entachée de l'incertitude due à la taille limitée de l'échantillon.

Cette incertitude peut être estimée de manière analytique ou par des méthodes statistiques (*e.g.* par bootstrap).

Ces méthodes sont appliquées sur l'exemple ci-dessus (Eq. (2)) avec les $d = 5$ entrées identifiées influentes dans l'étape de criblage précédente. Un échantillon Monte Carlo d'entrée de taille $n = 100$ fournit 100 évaluations du modèle et on obtient :

- sortie S :
 $\text{SRC}^2(Q) = 0.28$; $\text{SRC}^2(K_s) = 0.12$; $\text{SRC}^2(Z_v) = 0.15$; $\text{SRC}^2(H_d) = 0.26$; $\text{SRC}^2(C_b) = 0.03$ avec $R^2 = 0.98$;
 $\text{SRRC}^2(Q) = 0.27$; $\text{SRRC}^2(K_s) = 0.12$; $\text{SRRC}^2(Z_v) = 0.13$;
 $\text{SRRC}^2(H_d) = 0.26$; $\text{SRRC}^2(C_b) = 0.02$ avec $R^{2*} = 0.95$;
- sortie C_p :
 $\text{SRC}^2(Q) = 0.25$; $\text{SRC}^2(K_s) = 0.16$; $\text{SRC}^2(Z_v) = 0.18$; $\text{SRC}^2(H_d) = 0.00$; $\text{SRC}^2(C_b) = 0.07$ avec $R^2 = 0.70$;
 $\text{SRRC}^2(Q) = 0.26$; $\text{SRRC}^2(K_s) = 0.19$; $\text{SRRC}^2(Z_v) = 0.18$;
 $\text{SRRC}^2(H_d) = 0.06$; $\text{SRRC}^2(C_b) = 0.03$ avec $R^{2*} = 0.73$.

Pour la sortie S , la linéarité du modèle est prouvée et les indices de sensibilité basés sur la variance sont donnés grâce aux indices SRC^2 . Pour la sortie C_p , le modèle n'est ni linéaire ni monotone. Les indices SRC^2 et SRRC^2 peuvent être utilisés en première approximation, sachant que 30% de sensibilités n'ont pas été expliquées. Par ailleurs, si un autre échantillon Monte Carlo de même taille est utilisé, on s'aperçoit que les valeurs obtenues peuvent être très différentes. Pour obtenir des indices de sensibilité plus précis, il faudrait augmenter significativement la taille de l'échantillon.

3.2. Méthodes basées sur des tests statistiques. — À partir d'un échantillon i.i.d (par exemple Monte Carlo), d'autres techniques d'AS peuvent être utilisées. Par exemple, pour chaque entrée, un découpage en classes de valeurs équiprobables permet d'obtenir plusieurs échantillons de données. Des tests statistiques peuvent alors être appliqués pour mesurer l'homogénéité des populations entre les classes : moyennes communes (CMN) basées sur un test de Fisher, médianes communes (CMD) basées sur un test de χ^2 , variances communes (CV) basées sur un test de Fisher, localisations communes (CL) basées sur le test de Kruskal-Wallis, ... (Kleijnen & Helton [32], Helton et al. [22]). Ces méthodes ne requièrent pas d'hypothèse sur la monotonie de la sortie en fonction des entrées mais présentent l'inconvénient d'être peu intuitives comparativement aux méthodes de régression.

Sur l'exemple crues et la sortie C_p , avec un dcoupage en 5 classes, les p -values fournies par le test CMN sont les suivantes $p(K_s) = 8.6e-07$, $p(Q) = 1.0e-05$, $p(H_d) = 6.7e-05$, $p(Z_v) = 4.8e-05$, $p(C_b) = 3.9e-01$. Ces rsultats identifient H_d comme une entre influente, ce qui n'avait pas t rvl par les indices SRC et SRRC, mais qui avait bien t fourni par la mthode de Morris. Ces mthodes bases sur des tests statistiques sont donc bien complmentaires aux mthodes bass sur la rgression linare.

3.3. Dcomposition de la variance fonctionnelle. — Dans le cadre gnral d'un modle non linare et non monotone, on peut estimer l'importance des entres sur la sortie du modle en utilisant la dcomposition de $f(\cdot)$ en somme de fonctions lmentaires (Hoeffding [23])

$$(12) \quad f(X_1, \dots, X_d) = f_0 + \sum_i^d f_i(X_i) + \sum_{i<j}^d f_{ij}(X_i, X_j) + \dots + f_{12..d}(X_1, \dots, X_d),$$

o $f(\cdot)$ est intgrable sur $\Omega = [0, 1]^d$, f_0 est une constante et les autres fonctions vrifient les conditions suivantes :

$$(13) \quad \int_0^1 f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}) dx_{i_k} = 0 \quad \forall k = 1, \dots, s, \quad \forall \{i_1, \dots, i_s\} \subseteq \{1, \dots, d\}.$$

Cette dcomposition a t introduite par Sobol [61] pour l'AS (d'o son appellation "dcomposition de Sobol" dans ce domaine). Celui-ci a notamment montr que les conditions (13) impliquent que la dcomposition est unique.

Si les X_i sont alatoires et mutuellement indpendants, l'quation (12) permet d'obtenir la dcomposition de la variance fonctionnelle (appele aussi representation ANOVA fonctionnelle) :

$$(14) \quad \text{Var}[Y] = \sum_{i=1}^d V_i(Y) + \sum_{i<j} V_{ij}(Y) + \sum_{i<j<k} V_{ijk}(Y) + \dots + V_{12..d}(Y),$$

o $V_i(Y) = \text{Var}[\mathbb{E}(Y|X_i)]$, $V_{ij}(Y) = \text{Var}[\mathbb{E}(Y|X_i, X_j)] - V_i(Y) - V_j(Y)$ et ainsi de suite. partir de (14), les indices de sensibilit s'obtiennent alors extrmement naturellement :

$$(15) \quad S_i = \frac{\text{Var}[\mathbb{E}(Y|X_i)]}{\text{Var}(Y)} = \frac{V_i(Y)}{\text{Var}(Y)}, \quad S_{ij} = \frac{V_{ij}(Y)}{\text{Var}(Y)}, \quad S_{ijk} = \frac{V_{ijk}(Y)}{\text{Var}(Y)}, \quad \dots$$

Ces coefficients sont nomms “mesures d’importance bases sur la variance” ou plus simplement “*indices de Sobol*”. Compris entre 0 et 1 et leur somme valant 1, les indices de Sobol sont particulièrement faciles interprter (en terme de pourcentage de la variance de la rponse explique), ce qui explique leur popularit. L’indice du second ordre S_{ij} exprime la sensibilit du modle l’interaction entre les variables X_i et X_j , et ainsi de suite pour les ordres suprieurs. Les indices d’ordre un sont gaux aux SRC² quand le modle $f(\cdot)$ est purement liniaire.

Lorsque le nombre de variables d’entre d augmente, le nombre d’indices de sensibilit crot exponentiellement (il vaut $2^d - 1$); l’estimation et l’interpration de tous ces indices deviennent vite impossibles. Homma & Saltelli [24] ont alors introduit la notion d’indice de sensibilit total pour exprimer tous les effets d’une variable d’entre sur la sortie :

$$(16) \quad S_{T_i} = S_i + \sum_{j \neq i} S_{ij} + \sum_{j \neq i, k \neq i, j < k} S_{ijk} + \dots = \sum_{l \in \#i} S_l,$$

o $\#i$ represente tous les sous-ensembles d’indices contenant l’indice i . Ainsi, $\sum_{l \in \#i} S_l$ est la somme de tous les indices de sensibilit faisant intervenir i . En pratique, quand d est grand (par exemple $d > 10$), on se contente souvent d’estimer et d’interprter les indices d’ordre un et les indices totaux.

Pour estimer les indices de Sobol, des mthodes bases sur des chantillons Monte Carlo ont t dveloppes (Sobol [61], Saltelli [47]). Malheureusement, pour obtenir des estimations prises des indices de sensibilit, ces mthodes sont extrmement coteuses en nombre d’valuations du modle (taux de convergence en \sqrt{N} o N est la taille de l’chantillon). Il n’est pas rare dans les applications que l’estimation d’un indice de Sobol requiert 10000 valuations de $f(\cdot)$ pour obtenir une prcision de 10%, et ce pour chaque variable d’entre. L’utilisation d’chantillons dterministes de type quasi Monte Carlo (par exemple les squences LP τ de Sobol) la place d’chantillons Monte Carlo permet de rduire d’un facteur 10 le cot de ces estimations (Saltelli et al. [51]). La mthode FAST (Cukier et al. [9]), base sur une transforme de Fourier multi-dimensionnelle de $f(\cdot)$, est une autre mthode d’estimation des indices, relativement fine et nettement moins coteuse que la mthode de Monte Carlo. Saltelli et al. [53] l’ont tendue au calcul des indices totaux. Celle-ci demeure nanmoins coteuse et supporte mal la monte en dimension des entres (Tissot [68]).

Pour illustrer l’estimation d’indices de Sobol sur l’exemple crues (Eq. (2)) avec $d = 5$ entres alatoires, nous utilisons la mthode de Saltelli [47] avec un chantillonnage Monte Carlo. Celle-ci a un cot en nombre d’valuations du modle

gal $N = n(d+2)$ o n est la taille de l'échantillon Monte Carlo que l'on gne. Ici, $n = 1e5$ et les estimations ont t rptes $r = 100$ fois afin d'obtenir des intervalles de confiance (sous forme de boxplots) sur chaque estimation d'indices. Il faut noter que ces rptitions peuvent tre vites en estimant les intervalles de confiance par bootstrap. La figure 3 fournit le rsultat de ces estimations, qui ont ncessit au final $N = 7e7$ valuations du modle.

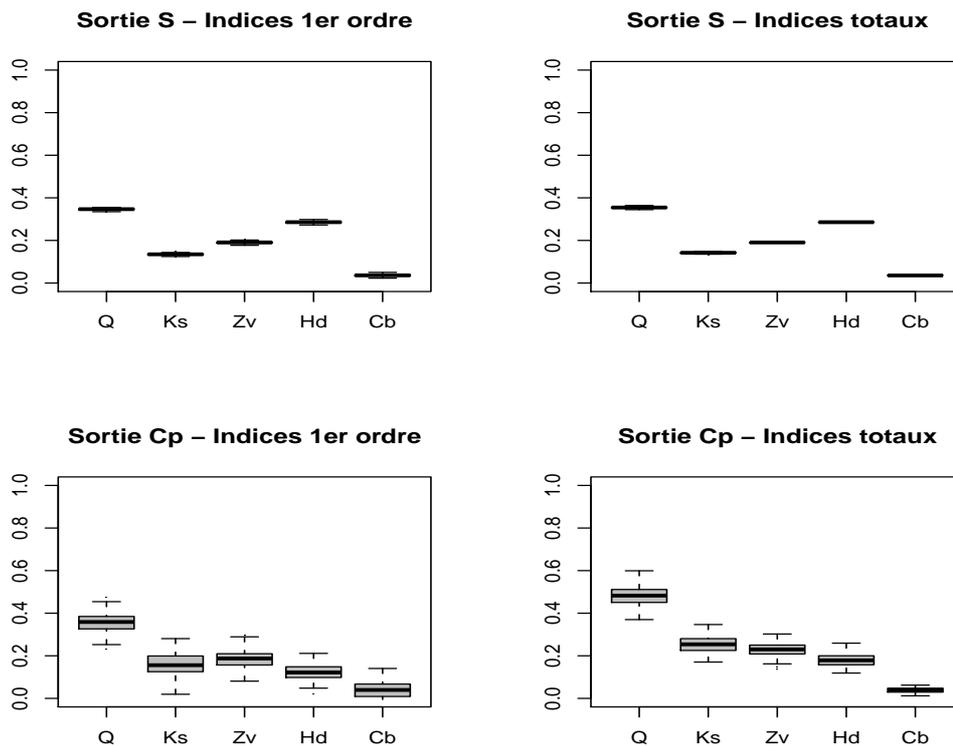


FIGURE 3. Indices de Sobol estimés sur l'exemple crues. Chaque box-plot correspond un échantillon de 100 estimations indépendantes.

Pour la sortie S , les indices du premier ordre sont quasiment égaux aux indices totaux, et les résultats semblent en tout point similaires ceux des SRC^2 . Le modèle ayant t prouv linéaire, l'estimation d'indices de Sobol est bien inutile dans ce cas. Pour la sortie C_p , des informations différentes de celles apportées par les SRC^2 et $SRRC^2$ sont obtenues : l'effet total de Q est de l'ordre de 50% (le double de son SRC^2), l'effet de H_d est conséquent (de l'ordre de 20%), alors que Q et K_s ont des effets d'interactions non négligeables. En estimant les indices de Sobol d'ordre deux, on estime que l'effet de l'interaction entre Q et K_s est de 6%.

4. Exploration du modle

En dehors de fonctions analytiques simples, de cas avec un trs faible nombre d'entres ou de codes de calcul demandant peu de ressources en temps, les cots d'estimation des indices de Sobol, mme avec des mthodes d'chantillonnage raffinés, sont souvent inatteignables. Cette section prsente sommairement une catgorie de mthodes d'approximation du modle numrique, qui permet d'estimer faible cot les indices de Sobol, et qui fournit galement une visualisation plus profonde des effets du modle.

4.1. Mthodes de lissage. — Au del des indices de Sobol qui ne donnent qu'une valeur scalaire pour l'effet d'une variable d'entre X_i sur la sortie Y , on peut tre intress par connatre l'influence sur Y de X_i le long de son domaine de variation. Dans la littrature, on parle souvent d'effets principaux, mais pour viter toute confusion avec les indices du premier ordre, il est prfrable de parler de visualisation (ou graphe) des effets principaux. L'outil graphique des scatterplots (visualisation du nuage de points d'un chantillon quelconque de simulations $(\mathbf{X}^n, \mathbf{Y}^n)$ l'aide des d graphes Y vs. X_i , $i = 1, \dots, d$) remplit cet objectif mais uniquement de manire visuelle, donc quelque peu subjective. Ceci est illustr la figure 4 qui reprend l'exemple crues et l'chantillon utilis au §3.1.

Bases sur des mthodes de rgression non paramtrique (Hastie & Tibshirani [19]), les techniques de lissage ont pour objectif, quant elles, d'estimer les moments conditionnels de Y d'ordre un ou plus. En AS, on se limite souvent l'esprance conditionnelle et aux ordres un et deux (Santner et al. [54]) pour obtenir :

- ▶ les graphes des effets principaux, entre X_i et $\mathbb{E}(Y|X_i) - \mathbb{E}(Y)$ sur tout le domaine de variation de X_i pour $i = 1, \dots, d$;
- ▶ les graphes des effets des interactions, entre (X_i, X_j) et $\mathbb{E}(Y|X_i X_j) - \mathbb{E}(Y|X_i) - \mathbb{E}(Y|X_j) - \mathbb{E}(Y)$ sur tout le domaine de variation de (X_i, X_j) pour $i = 1, \dots, d$ et $j = i + 1, \dots, d$.

Storlie & Helton [64] ont effectu une revue relativement complte des mthodes de lissage que l'on peut utiliser pour l'AS : moyennes mobiles, mthodes noyaux, polynmes locaux, splines de lissage, etc. Sur la figure 4, le lisseur par polynômes locaux est trace pour chaque nuage de points, ce qui permet d'identifier nettement la tendance moyenne de la sortie par rapport chaque entre.

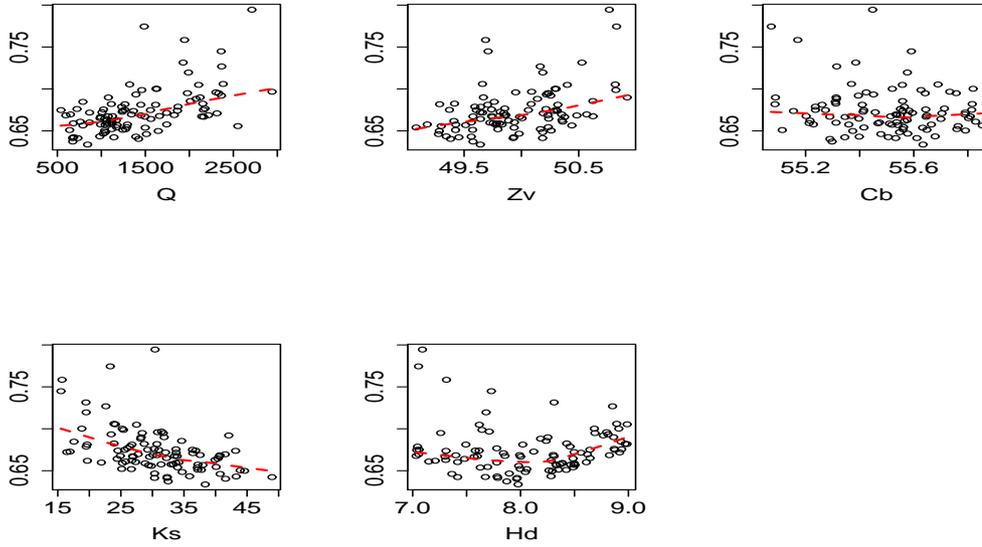


FIGURE 4. Scatterplots sur l'exemple crues entre les 5 entres Q , Z_v , C_b , K_s , H_d et la sortie C_p . La courbe en pointill est un lisseur par polynmes locaux.

Une fois que ces esperances conditionnelles sont modlises, il est alors facile par chantillonnage de quantifier leur variance, et ainsi d'estimer les indices de Sobol (cf. Eq. 15) d'ordre un, deux, voire d'ordres suprieurs. Da Veiga et al. [10] discutent des propriets thoriques des estimateurs par polynmes locaux de l'esperance et de la variance conditionnelles et en dduisent les propriets thoriques des estimateurs des indices de Sobol par polynmes locaux (cf. Cette approche leur permet de rsoudre le probleme des entres correes d'une maniere nettement moins coteuse que par les techniques usuelles. Storlie & Helton [64] discutent galement des modes additifs et des arbres de rgression pour estimer de maniere non paramtrique $\mathbb{E}(Y|X_1, \dots, X_d)$, ce qui revient construire un mode approch du mode numrique. Ce mode approch est alors appel "mtamodle".

4.2. Mtamodles. — Le terme "mtamodle" est un synonyme (quoiqu'un peu plus gnral) de l'appellation "surface de rponse", outil connu depuis bien longtemps dans le domaine de la planification d'experiences. La methode des surfaces de rponse a pour objectif de construire une fonction qui simule le comportement d'un phnomne physique ou chimique dans le domaine de variation des variables influentes, et ce, partir d'un certain nombre d'experiences

(Box & Draper [4]). Des généralisations ultérieures ont amené cette méthode à être utilisée pour construire des modèles simplifiés se substituant à l'exécution de codes de calcul nécessitant trop de temps d'exécution ou de ressources (Downing et al. [14], Sacks et al. [46], Fang et al. [16]).

Construire un métamodèle a ainsi pour objectif d'obtenir un modèle mathématique représentatif du code étudié en terme de qualité d'approximation, ayant de bonnes capacités de prédiction, et dont le temps de calcul pour valuer une réponse est négligeable. Ce métamodèle est construit et ajusté à partir de quelques simulations du code (correspondant à différents jeux de valeurs des paramètres). Le nombre de simulations nécessaires dépend de la complexité du code et du scénario qu'il modélise, du nombre de variables d'entrée et de la qualité d'approximation souhaitée. Ce métamodèle peut alors être substitué ou associé au code pour réaliser différents objectifs :

- prédiction rapide de nouvelles réponses,
- AS et exploration du modèle pour une meilleure compréhension de son comportement, des effets de ses variables d'entrée et de leurs interactions,
- résolution de problèmes d'optimisation de la réponse ou de calibration de paramètres inconnus du modèle numérique,
- estimation de la probabilité d'occurrence d'événements rares en sortie du modèle numérique,
- participation aux phases de validation et de qualification du modèle numérique.

La construction du métamodèle, basée la plupart du temps sur des techniques de moindres carrés, est évidemment réalisée en accord avec son utilisation future qui peut lui imposer des contraintes. La mise à disposition d'un métamodèle est également extrêmement utile si on étudie un système sans bien connaître les incertitudes sur ses variables d'entrée. Si un métamodèle est construit et validé dans un domaine de variation des entrées suffisamment large, différentes études pourront être réalisées en faisant varier les incertitudes des entrées.

Dans la pratique, on s'intéresse à trois principales questions lors de la construction d'un métamodèle :

- ▷ le choix du métamodèle qui peut être issu de tout modèle de régression linéaire, non linéaire, paramétrique ou non paramétrique (Hastie et al. [20]). Parmi les modèles les plus utilisés pour ajuster les réponses de codes de calcul, on peut citer les polynômes, splines, modèles linéaires généralisés, modèles additifs généralisés, krigeage, technique MARS, réseaux de neurones, SVM, boosting d'arbres de régression (Simpson et al. [60], Fang et al. [16]). Le choix du métamodèle est un problème

en soi, certains tant plus adaptés que d'autres différents types de situation. Une première stratégie est de privilégier la simplicité, donc de se satisfaire du modèle le plus simple possible en adéquation avec les objectifs de l'étude ;

- ▷ la planification des calculs. Les principales qualités requises pour un plan d'expériences sont sa robustesse (capacité d'analyser différents modèles), son efficacité (minimisation d'un critère), la répartition de ses points (remplissage uniforme de l'espace d'échantillonnage) et un coût faible pour sa construction (Santner et al. [54], Fang et al. [16]). Plusieurs travaux proposent des études numériques fines permettant d'étudier les qualités de différents types de plans d'expériences vis-à-vis de la prédictivité d'un modèle (Simpson et al. [59], Franco [17], Marrel [38]) ;
- ▷ la validation du modèle. Dans le domaine des plans d'expériences classiques, la validation correcte d'une surface de réponse est un aspect crucial et soigner particulièrement (Droesbecke et al. [15]). Dans le domaine des expériences numériques, peu de travaux s'attardent sur ce problème, la pratique usuelle tant d'estimer des critères globaux (erreur quadratique moyenne, erreur en valeur absolue, . . .) sur une base de test, par validation croisée ou par bootstrap (Kleijnen & Sargent [33], Fang et al. [16]). Lorsque le nombre de calculs est peu important et pour s'affranchir des problèmes induits par la validation croisée, Iooss et al. [25] se sont récemment intéressés à la minimisation de la taille de la base de test, tout en conservant une bonne estimation du coefficient de prédictivité du modèle.

Certains modèles permettent d'obtenir directement les indices de sensibilité. Par exemple, Sudret [66] a montré que les indices de Sobol' découlent directement de la décomposition en polynômes de chaos. La formulation du modèle du krigeage est également particulièrement intéressante car elle permet d'obtenir les indices de sensibilité de manière analytique (Oakley & O'Hagan [43], Marrel et al. [40]), en y associant les incertitudes dues à l'impact de l'approximation du modèle par le modèle. Malheureusement, la mise en application de cette méthode est plutôt ardue d'un point de vue numérique (Marrel [38]). C'est pourquoi, on préfère la plupart du temps appliquer une technique d'échantillonnage intensif (cf. §3.3) directement sur le modèle pour estimer les indices de Sobol' (Santner et al. [54], Iooss et al. [27]). La part de variance non expliquée

par le mtamodle (calculé par $1 - Q_2$, cf. Eq. (11)) nous permet de connaître ce que l'on perd en utilisant le mtamodle (Sobol [62], Jacques [28]). Storlie et al. [65] proposent une méthode de bootstrap permettant d'estimer l'erreur sur chaque indice de Sobol estim, du fait de l'utilisation du mtamodle la place du vrai modle.

De la même manière, une fois le mtamodle construit et validé, il est aisé de visualiser les effets principaux (Schonlau & Welch [57]). Ceux-ci sont donnés soit directement par le mtamodle (c'est le cas avec les méthodes polynômes de chaos, krigeage, GAM), soit en calculant les espérances conditionnelles $\mathbb{E}(Y|X_i)$ par simulation.

Pour illustrer notre propos sur l'exemple crues (Eq. (2)), un mtamodle de krigeage (Sacks et al. [46]) est construit sur un échantillon Monte Carlo d'entrées-sorties (entrées Q, K_s, Z_v, H_d, C_b et sortie C_p), de taille $N' = 100$. Le mtamodle de krigeage utilisé consiste en un terme déterministe issu d'une régression linéaire, et en un terme correctif stochastique modélisé par un processus stationnaire gaussien de fonction de covariance de type exponentielle généralisée (cf. Santner et al [54] pour plus de détails). La technique d'estimation des paramètres de ce mtamodle est décrite dans Roustant et al. [45]. Le coefficient de prédictivité estimé par leave-one-out vaut $Q_2 = 99\%$, comparé avec $Q_2 = 75\%$ obtenu avec un simple modèle linéaire. Le mtamodle du krigeage obtenu est alors utilisé pour estimer les indices de Sobol de la même manière qu'au §3.3 : méthode de Saltelli, échantillonnage Monte Carlo, $n = 1e5$, $r = 100$ répétitions, ce qui demande $N = 7e7$ prédictions du mtamodle. Dans le tableau 1, nous pouvons comparer les indices de Sobol (moyennés sur les 100 répétitions) obtenus avec le mtamodle ceux obtenus avec le "vrai" modèle crues (Eq. (2)). Les erreurs entre ces deux estimations sont relativement faibles : avec seulement 100 simulations du vrai modèle, on a réussi à obtenir des estimations correctes (erreurs $< 15\%$) des indices de Sobol du premier ordre et totaux.

5. Synthèse et conclusion

Même si toutes les méthodes d'AS n'ont pu être listées, cette revue a permis d'illustrer la grande variété de méthodes d'AS disponibles, que l'on a cherché à positionner clairement, en termes d'hypothèses requises et de réponses apportées. La Figure 5 présente une synthèse des principales méthodes d'AS relative à la variabilité de la sortie d'un modèle numérique. Cette liste n'est bien entendu pas exhaustive et ne tient pas compte d'éventuelles améliorations apportées aux

TABLE 1. Indices de Sobol estimés par échantillonnage Monte Carlo (utilisant $N = 7e7$ évaluations) en utilisant le modèle crues et en utilisant un mtamodèle construit sur $N' = 100$ simulations du modèle crues.

Indices (en %)	Q	K_s	Z_v	H_d	C_b
S_i modle	35.5	15.9	18.3	12.5	3.8
S_i mtamodle	38.9	16.8	18.8	13.9	3.7
S_{T_i} modle	48.2	25.3	22.9	18.1	3.8
S_{T_i} mtamodle	45.5	21.0	21.3	16.8	4.3

différentes méthodes. C'est le cas par exemple pour la méthode des bifurcations séquentielles (Kleijnen [31]) et celle de Morris (Pujol [44]). Cette figure a un lien direct avec les différentes sections de ce papier : le criblage trs grande dimension (§2.1) rsum sous le vocable “super criblage”, les plans d'expériences dits “classiques” (§2.2), la mthode de Morris (§2.3), les mthodes de dcomposition de la variance par rgression linéaire (§3.1), par indices de Sobol (§3.3) et par mtamodle (§4.2).

La figure 5 possde ainsi plusieurs niveaux de lecture :

- distinction entre mthodes de criblage (identification des entres non influentes parmi un grand nombre) et mthodes de hirarchisation précise de l'influence des entres ;
- positionnement des mthodes en fonction du cot requis en nombre d'évaluations du modle, qui dpend de manire plus ou moins linéaire du nombre de variables d'entre ;
- positionnement des mthodes en fonction de leurs hypothses sur la complexité et rgularité du modle ;
- visualisation d'une progression dans l'application de ces mthodes. En effet, une approche méthodologique raisonne consiste utiliser la méthode la plus simple adaptée au problme posé, fonction de l'objectif de l'étude, du nombre d'évaluations du modle que l'on peut réaliser et de la connaissance que l'on a sur la complexité du modle. La validation *a posteriori* de la méthode utilisée permet de savoir s'il est nécessaire d'utiliser une méthode plus performante, en réalisant ou non de nouvelles simulations.

Un positionnement plus dtaillé des différentes mthodes d'AS dans ce type de tableau synthtique est fourni par la figure 6 qui offre galement deux niveaux de lecture supplémentaires. Le type d'informations en terme d'AS apport par chaque mthode est retranscrit pas une couleur : certaines mthodes fournissent

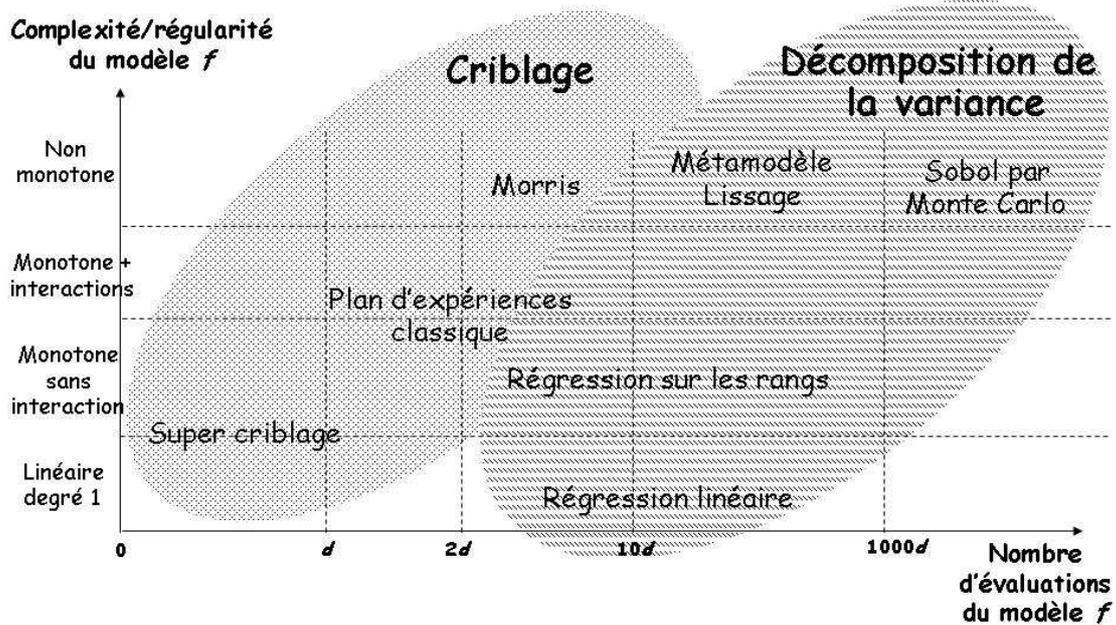


FIGURE 5. Synthèse des méthodes d'AS placées dans un diagramme (cot en nombre d'évaluations du modle vs. complexité du modle).

seulement des indices de sensibilité au premier ordre, alors que d'autres donnent des résultats plus riches (interactions, indices totaux, etc.). Les cases grises indiquent les méthodes nécessitant de posséder certaines informations sur le modèle : le sens de variation de la sortie par rapport aux entrées pour les bifurcations squelettiques et la connaissance d'interactions qui n'ont pas d'effet pour les plans factoriels fractionnaires.

En définitive, l'approche statistique développée depuis une vingtaine d'années propose un cadre méthodologique utile et performant pour l'AS de modèles numériques complexes. Cependant, de nombreux problèmes restent ouverts :

- ▷ La question de l'AS pour des entrées non indépendantes a été abordée par plusieurs auteurs (Saltelli & Tarantola [52], Jacques et al. [29], Xu

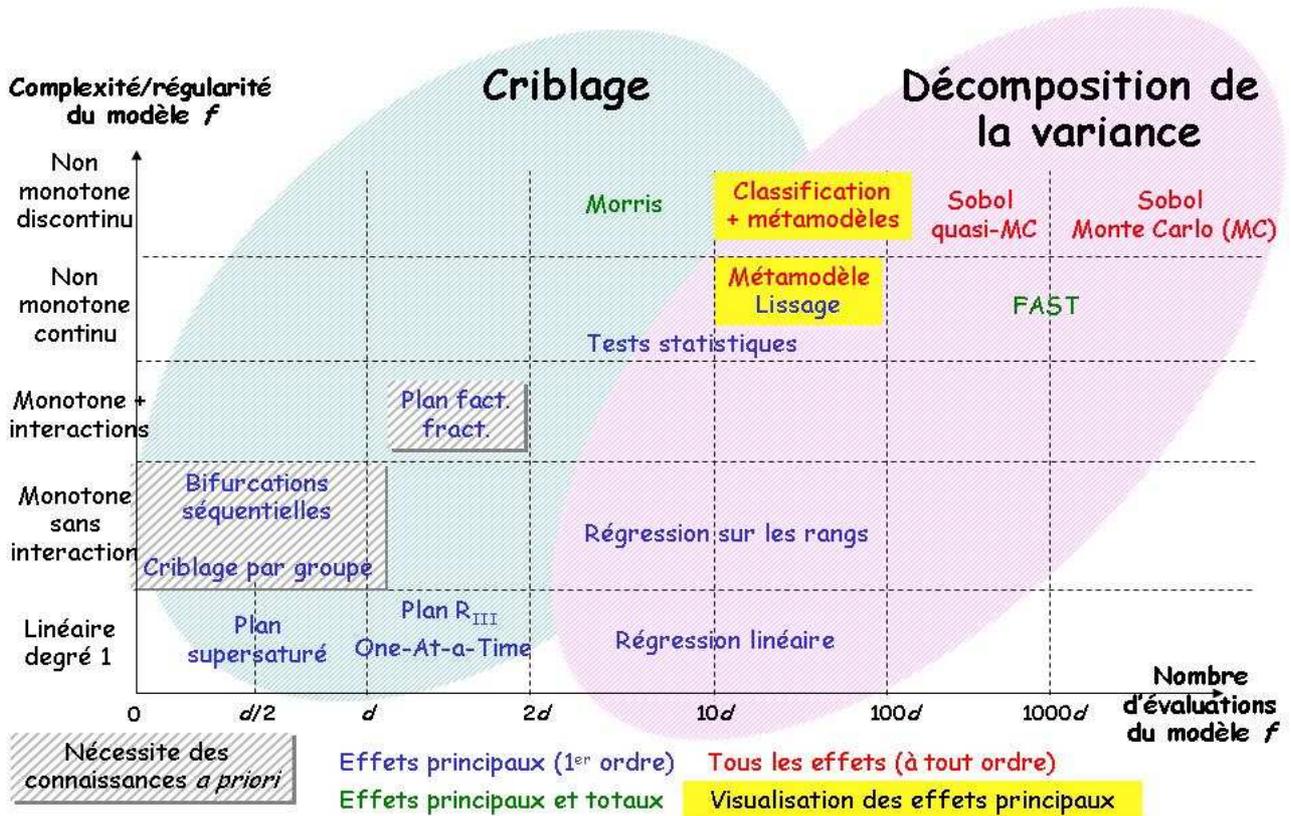


FIGURE 6. Synthèse détaillée des méthodes d'AS.

& Gertner [70], Da Veiga et al. [10]) mais reste toujours d'actualité ;

- ▷ Les recherches actuelles pour estimer les indices de Sobol se portent sur le développement d'algorithmes qui permettent d'estimer tous les indices du premier ordre avec un coût indépendant de d . Par exemple Tarantola et al. [67] utilisent une technique dite de "Random Balance Design" couplée avec la méthode FAST.
- ▷ L'estimation à moindre coût des indices de Sobol totaux est toujours un axe de recherche de première importance dans les applications

(cf. Saltelli et al. [49] pour une revue récente sur le sujet).

- ▷ Ce papier s’est focalisé sur l’AS par rapport à la variabilité de la sortie d’un modèle. En pratique, les résultats d’une AS peuvent être demandés vis-à-vis d’autres quantités d’intrt, comme par exemple la probabilité de dépasser un seuil (Saltelli et al. [50], De Rocquigny et al. [12]). De nombreux travaux restent à réaliser pour bien baliser ce type de questions.
- ▷ Dans de nombreuses applications, les sorties du modèle étudié sont multiples, voire fonctionnelles (cf. Campbell et al. [6], Lamboni et al. [35] et Marrel et al. [39] pour de premiers travaux). Le cas des entrées fonctionnelles reçoit également un intrt croissant (Iooss & Ribatet [26], Lilburne & Tarantola [36]), mais son traitement dans un cadre statistique fonctionnel reste à faire.
- ▷ Enfin, les méthodes d’AS statistiques et quantitatives sont limitées des modèles avec peu de variables aléatoires (quelques dizaines). Les méthodes déterministes, de type adjoint, sont bien adaptées quant à elle aux modèles avec un grand nombre de variables d’entrées. Une idée naturelle est de tirer profit des avantages des deux méthodes (cf. Sobol & Kucherenko [63] pour de premières idées).

6. Remerciements

Ces travaux ont été en partie financés par l’Agence Nationale de la Recherche (ANR) au travers du programme COSINUS (projet COSTA-BRAVA n° ANR-09-COSI-015). L’auteur remercie Aurélie Arnaud, Gilles Pujol et Hervé Monod pour de fructueuses discussions. Les calculs statistiques ont été réalisés dans l’environnement R, à l’aide des packages “sensitivity” et “DiceKriging”, ainsi que du logiciel SMURFER.

Références

- [1] A. Arnaud, E. De Rocquigny, and D. Lagrange. *Projet “Incertitudes” - Lot “Rfrentiel méthodologique” - Panorama pédagogique global des méthodes de maîtrise des incertitudes travers un cas d’cole*. Rapport EDF R&D H-T56-2007-01082-FR, 2007.
- [2] A. Badea and R. Bolado. *Review of sensitivity analysis methods and experience*. PAMINA 6th FPEC Project, European Commission, 2008. <http://www.ip-pamina.eu/downloads/pamina.m2.1.d.4.pdf>.

- [3] B. Bettonvil and J.P.C. Kleijnen. Searching for important factors in simulation models with many factors : Sequential bifurcation. *European Journal of Operational Research*, 96 :180–194, 1996.
- [4] G.E. Box and N.R. Draper. *Empirical model building and response surfaces*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1987.
- [5] D.G. Cacuci. Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach. *Journal of Mathematical Physics*, 22 :2794, 1981.
- [6] K. Campbell, M.D. McKay, and B.J. Williams. Sensitivity analysis when model outputs are functions. *Reliability Engineering and System Safety*, 91 :1468–1472, 2006.
- [7] W. Castaings. *Analyse de sensibilit et estimation de paramtres pour la modlisation hydrologique : potentiel et limitations des mthodes variationnelles*. Thse de l'Universit Joseph Fourier, Grenoble 1, 2007.
- [8] M. Claeys-Bruno, M. Dobrijevic, R. Cela, R. Phan-Tan-Luu, and M. Sergent. Supersaturated design : construction, comparison and interpretation. In *VI Colloquium Chemiometricum Mediterraneum*, Saint Maximin La Sainte Baume, France, september 2007.
- [9] H. Cukier, R.I. Levine, and K. Shuler. Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26 :1–42, 1978.
- [10] S. Da Veiga, F. Wahl, and F. Gamboa. Local polynomial estimation for sensitivity analysis on models with correlated inputs. *Technometrics*, 51(4) :452–463, 2009.
- [11] E. De Rocquigny. La maîtrise des incertitudes dans un contexte industriel - 1ère partie : une approche méthodologique globale basée sur des exemples. *Journal de la Société Française de Statistique*, 147(3) :33–71, 2006.
- [12] E. De Rocquigny, N. Devictor, and S. Tarantola, editors. *Uncertainty in industrial practice*. Wiley, 2008.
- [13] A. Dean and S. Lewis, editors. *Screening - Methods for experimentation in industry, drug discovery and genetics*. Springer, 2006.
- [14] D.J. Downing, R.H. Gardner, and F.O. Hoffman. An examination of response surface methodologies for uncertainty analysis in assessment models. *Technometrics*, 27(2) :151–163, 1985.
- [15] J-J. Dreesbecke, J. Fine, and G. Saporta, editors. *Plans d'expriences (Applications l'entreprise)*. Technip, Paris, 1998.
- [16] K-T. Fang, R. Li, and A. Sudjianto. *Design and modeling for computer experiments*. Chapman & Hall/CRC, 2006.
- [17] J. Franco. *Planification d'expriences numriques en phase exploratoire pour la simulation des phnomnes complexes*. Thse de l'Ecole Nationale Suprieure des Mines de Saint-Etienne, 2008.
- [18] H.C. Frey and S.R. Patil. Identification and review of sensitivity analysis methods. *Risk Analysis*, 22.
- [19] T. Hastie and R. Tibshirani. *Generalized additive models*. Chapman and Hall, London, 1990.

- [20] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer, 2002.
- [21] J.C. Helton. Uncertainty and sensitivity analysis techniques for use in performance assesment for radioactive waste disposal. *Reliability Engineering and System Safety*, 42 :327–367, 1993.
- [22] J.C. Helton, J.D. Johnson, C.J. Salaberry, and C.B. Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering and System Safety*, 91 :1175–1209, 2006.
- [23] W. Hoeffding. A class of statistics with asymptotically normal distributions. *Annals of Mathematical Statistics*, 19 :293–325, 1948.
- [24] T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of non linear models. *Reliability Engineering and System Safety*, 52 :1–17, 1996.
- [25] B. Iooss, L. Boussouf, V. Feuillard, and A. Marrel. Numerical studies of the metamodel fitting and validation processes. *International Journal of Advances in Systems and Measurements*, 3, in press, 2010.
- [26] B. Iooss and M. Ribatet. Global sensitivity analysis of computer models with functional inputs. *Reliability Engineering and System Safety*, 94 :1194–1204, 2009.
- [27] B. Iooss, F. Van Dorpe, and N. Devictor. Response surfaces and sensitivity analyses for an environmental model of dose calculations. *Reliability Engineering and System Safety*, 91 :1241–1251, 2006.
- [28] J. Jacques. *Contributions l’analyse de sensibilit et l’analyse discriminante gnralise*. Thse de l’Universit Joseph Fourier, Grenoble 1, 2005.
- [29] J. Jacques, C. Lavergne, and N. Devictor. Sensitivity analysis in presence of model uncertainty and correlated inputs. *Reliability Engineering and System Safety*, 91 :1126–1134, 2006.
- [30] J.P.C. Kleijnen. Sensitivity analysis and related analyses : a review of some statistical techniques. *Journal of Statistical Computation and Simulation*, 57 :111–142, 1997.
- [31] J.P.C. Kleijnen. *Design and analysis of simulation experiments*. Springer, 2008.
- [32] J.P.C. Kleijnen and J.C. Helton. Statistical analyses of scatterplots to identify important factors in large-scale simulations, 1 : Review and comparison of techniques. *Reliability Engineering and System Safety*, 65 :147–185, 1999.
- [33] J.P.C. Kleijnen and R.G. Sargent. A methodology for fitting and validating metamodels in simulation. *European Journal of Operational Research*, 120 :14–29, 2000.
- [34] D. Kurowicka and R. Cooke. *Uncertainty analysis with high dimensional dependence modelling*. Wiley, 2006.
- [35] M. Lamboni, H. Monod, and D. Makowski. Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. *Reliability Engineering and System Safety*, submitted, 2010.
- [36] L. Lilburne and S. Tarantola. Sensitivity analysis of spatial models. *International Journal of Geographical Information Science*, 23 :151–168, 2009.

- [37] D.K.J. Lin. A new class of supersaturated design. *Technometrics*, 35 :28–31, 1993.
- [38] A. Marrel. *Mise en oeuvre et exploitation du métamodèle processus gaussien pour l'analyse de modèles numériques - Application un code de transport hydrogéologique*. Thèse de l'INSA Toulouse, 2008.
- [39] A. Marrel, B. Iooss, M. Jullien, B. Laurent, and E. Volkova. Global sensitivity analysis for models with spatially dependent outputs. *Environmetrics*, submitted, 2010.
- [40] A. Marrel, B. Iooss, B. Laurent, and O. Roustant. Calculations of the Sobol indices for the Gaussian process metamodel. *Reliability Engineering and System Safety*, 94 :742–751, 2009.
- [41] D.C. Montgomery. *Design and analysis of experiments*. John Wiley & Sons, 2004.
- [42] M. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33 :161–174, 1991.
- [43] J.E. Oakley and A. O'Hagan. Probabilistic sensitivity analysis of complex models : A Bayesian approach. *Journal of the Royal Statistical Society, Series B*, 66 :751–769, 2004.
- [44] G. Pujol. Simplex-based screening designs for estimating metamodels. *Reliability Engineering and System Safety*, 94 :1156–1160, 2009.
- [45] O. Roustant, D. Ginsbourger, and Y. Deville. DiceKriging, DiceOptim : Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization. *Journal of Statistical Software*, submitted.
- [46] J. Sacks, W.J. Welch, T.J. Mitchell, and H.P. Wynn. Design and analysis of computer experiments. *Statistical Science*, 4 :409–435, 1989.
- [47] A. Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communication*, 145 :280–297, 2002.
- [48] A. Saltelli and P. Annoni. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling and Software*, 2010, in press.
- [49] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communication*, 181 :259–270, 2010.
- [50] A. Saltelli, K. Chan, and E.M. Scott, editors. *Sensitivity analysis*. Wiley Series in Probability and Statistics. Wiley, 2000.
- [51] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Salasana, and S. Tarantola. *Global sensitivity analysis - The primer*. Wiley, 2008.
- [52] A. Saltelli and S. Tarantola. On the relative importance of input factors in mathematical models : Safety assessment for nuclear waste disposal. *Journal of American Statistical Association*, 97 :702–709, 2002.
- [53] A. Saltelli, S. Tarantola, and K. Chan. A quantitative, model-independent method for global sensitivity analysis of model output. *Technometrics*, 41 :39–56, 1999.
- [54] T. Santner, B. Williams, and W. Notz. *The design and analysis of computer experiments*. Springer, 2003.

- [55] G. Saporta. *Probabilités, analyse de données et statistique*. éditions Technip, 2me edition, 2006.
- [56] F. Satterthwaite. Random balance experimentation. *Technometrics*, 1 :111–137, 1959.
- [57] M. Schonlau and W.J. Welch. Screening the input variables to a computer model. In A. Dean and S. Lewis, editors, *Screening - Methods for experimentation in industry, drug discovery and genetics*. Springer, 2006.
- [58] M. Sergent, B. Corre, and D. Dupuy. Comparison of different screening methods. In *VI Colloquium Chemiometricum Mediterraneum*, Saint Maximin La Sainte Baume, France, september 2007.
- [59] T.W. Simpson, D.K.J. Lin, and W. Chen. Sampling strategies for computer experiments : Design and analysis. *International Journal of Reliability and Applications*, 2 :209–240, 2001.
- [60] T.W. Simpson, J.D. Peplinski, P.N. Kock, and J.K. Allen. Metamodel for computer-based engineering designs : Survey and recommandations. *Engineering with Computers*, 17 :129–150, 2001.
- [61] I.M. Sobol. Sensitivity estimates for non linear mathematical models. *Mathematical Modelling and Computational Experiments*, 1 :407–414, 1993.
- [62] I.M. Sobol. Theorems and examples on high dimensional model representation. *Reliability Engineering and System Safety*, 79 :187–193, 2003.
- [63] I.M. Sobol and S. Kucherenko. Derivative based global sensitivity measures and their links with global sensitivity indices. *Mathematics and Computers in Simulation*, 79 :3009–3017, 2009.
- [64] C.B. Storlie and J.C. Helton. Multiple predictor smoothing methods for sensitivity analysis : Description of techniques. *Reliability Engineering and System Safety*, 93 :28–54, 2008.
- [65] C.B. Storlie, L.P. Swiler, J.C. Helton, and C.J. Salaberry. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering and System Safety*, 94 :1735–1763, 2009.
- [66] B. Sudret. Global sensitivity analysis using polynomial chaos expansion. *Reliability Engineering and System Safety*, 93 :964–979, 2008.
- [67] S. Tarantola, D. Gatelli, and T. Mara. Random balance designs for the estimation of first order global sensitivity indices. *Reliability Engineering and System Safety*, 91 :717–727, 2006.
- [68] J-Y. Tissot. Transparents de presentation - Estimation des indices d'ordre 2 par la méthode FAST. Projet ANR COSTA BRAVA, May 2010. http://www.math.univ-toulouse.fr/COSTA_BRAVA/document/JYTissot.pdf.
- [69] E. Volkova, B. Iooss, and F. Van Dorpe. Global sensitivity analysis for a numerical model of radionuclide migration from the RRC "Kurchatov Institute" radwaste disposal site. *Stochastic Environmental Research and Risk Assesment*, 22 :17–31, 2008.

- [70] C. Xu and G. Gertner. Extending a global sensitivity analysis technique to models with correlated parameters. *Computational Statistics and Data Analysis*, 51 :5579–5590, 2007.

BERTRAND IOOSS, EDF R&D, 6 Quai Watier, 78401, Chatou
E-mail : `bertrand.iooss@edf.fr`