

# OPUS : Cas test CEA

## Propagation des incertitudes dans le problème COUPLEX-1

JM. Martinez, Ph. Montarnal, M. Lamoureux, O. Le Maître

6 avril 2010

### Table des matières

<b>1</b>	<b>Contexte</b>	<b>1</b>
<b>2</b>	<b>Modèle Couplex probabiliste</b>	<b>2</b>
<b>3</b>	<b>Code opuscouplex</b>	<b>3</b>
3.1	Compilation d'opuscouplex . . . . .	4
3.2	Réalisation d'une seule simulation . . . . .	5
3.3	Réalisation de plusieurs simulations . . . . .	5
3.4	Objectif du cas test . . . . .	6
<b>4</b>	<b>Résultats obtenus sur le cas CT_KT</b>	<b>7</b>
<b>5</b>	<b>Conclusions</b>	<b>7</b>

## 1 Contexte

Cette note décrit le cas test industriel proposé par le CEA dans le cadre du projet OPUS et le code **opuscouplex** téléchargeable permettant sa réalisation. Ce cas test est basé sur le bench Couplex ([http://www.gdrmommas.org/ex\\_qualifications.html](http://www.gdrmommas.org/ex_qualifications.html)) défini en collaboration avec l'ANDRA par le Groupement MoMas (Modélisations Mathématiques et Simulations Numériques liées aux problèmes de gestion des déchets nucléaires).

Couplex est un bench bidimensionnel sur la simulation des écoulements saturés et du transport des radionucléides dans un milieu géologique. C'est un cas simplifié mais représentatif des études de sûreté qui peuvent être menées dans le cadre d'un stockage profond des déchets nucléaires. Ce cas test est constitué d'un terme source volumique à 450 m de profondeur, représentatif de la zone de stockage. Ce terme source est situé dans une couche géologique fortement imperméable d'une centaine de mètres d'épaisseur, cette couche étant entourée de trois autres couches géologiques présentant des caractéristiques physiques (perméabilités, diffusion, ...) fortement différentes. La hauteur du domaine de simulation est de 695 m tandis que son extension horizontale est de 25 km. Le problème à résoudre est donc celui d'une équation de convection-diffusion avec diffusion dominante, sur des coefficients fortement hétérogènes et un domaine de calcul possédant un fort rapport d'aspect (695/25000).

Le bench Couplex initial, cas d'étude déterministe, a été proposé pour valider des modèles physiques, des méthodes et des schémas de résolution numérique. Les coefficients caractérisant la modélisation dans chaque couche géologique ont été homogénéisés et les conditions aux limites du domaine fixées. Les études pour le stockage profond des déchets nucléaires relevant de la sûreté, il s'est avéré nécessaire d'étendre ce bench pour prendre en compte les effets des incertitudes

sur les paramètres des modèles et sur ceux définissant les scénarii des conditions aux limites imposées sur le domaine.

## 2 Modèle Couplex probabiliste

Le domaine de calcul du bench Couplex est rectangulaire et s'étend sur une longueur horizontale de 25,000 (m) sur une hauteur de 695 (m). Il contient 4 couches géologiques superposées, dénotées D (Dogger), C (Clay), L (Limestone) et M (Marl) dans la suite. La géométrie et les dimensions des différentes couches sont représentées dans la figure 1. Le dépôt, de géométrie rectangulaire se trouve dans la couche C (Clay) et correspond au sous-domaine  $x \in [18440, 21680]$ ,  $z \in [244, 250]$ .



FIG. 1 – Géométrie du domaine de calcul (les échelles ne sont pas respectées).

L'écoulement est supposé stationnaire en milieu saturé isotrope. On note  $\mathbf{u}(\mathbf{x})$  la vitesse de Darcy reliée à la charge hydraulique  $H$  (m) par :

$$\mathbf{u} = -K_i \nabla H, \quad (1)$$

où  $K_i$  est la perméabilité de la couche  $i$ . La charge hydraulique est solution de l'équation elliptique

$$\nabla \cdot (-K_i \nabla H) = 0. \quad (2)$$

Les incertitudes porteront sur les perméabilités  $K_{D,C,L,M}$  des 4 couches. Pour la résolution de l'équation elliptique, on considérera des conditions aux limites pour la charge de type Neumann homogène et de Dirichlet. Les incertitudes portent sur ces charges hydrauliques. Elles ont été spécifiées de façon à assurer un sens d'écoulement constant, de la droite du domaine à la gauche.

Pour la partie transport de radionucléide, on s'intéresse à l'élément I-129. L'équation de convection dispersion de sa concentration notée  $C$  est

$$\omega \left( \frac{\partial C}{\partial t} + \lambda C \right) - \nabla \cdot (\mathbf{D} \nabla C + \mathbf{u} C) = f, \quad (3)$$

avec comme condition initiale  $C(\mathbf{x}, t = 0) = 0$ . Dans l'équation de convection dispersion  $\lambda$  est la demie vie de l'iode et  $\mathbf{D}$  est le tenseur de diffusion dispersion dont l'expression est

$$\mathbf{D} = d_e \mathbf{I} + |\mathbf{u}| [\alpha_l \mathbf{E}(\mathbf{u}) + \alpha_t (\mathbf{I} - \mathbf{E}(\mathbf{u}))], \quad E_{ij} = \frac{u_i u_j}{|\mathbf{u}|^2}, \quad (4)$$

avec  $\alpha_l$  (resp.  $\alpha_t$ ) la dispersivités longitudinale (resp. transversale) et  $d_e$  la diffusion effective,  $\omega$  la porosité effective du milieu, et  $f$  le terme source représentant le relâchement de l'iode dans le

milieu au niveau du dépôt. Nous supposons  $\omega$ ,  $f$  et  $\lambda$  déterministes. L'évolution temporelle du terme source  $f$  est tabulée, ses valeurs étant lues dans un fichier (iode.dat). Les conditions aux limites pour l'équation de convection dispersion sont de nouveaux de type Neumann et Dirichlet, cette fois toutes deux homogènes.

Les aléas pour le transport portent sur les dispersivités longitudinales et les rapports d'anisotropie  $\rho = \alpha_l/\alpha_t$  dans les couches D (Dogger) et L (Lime) et les diffusions effectives des couches C (Clay) et M (Marl). Ces aléas sont de nouveau pris indépendants.

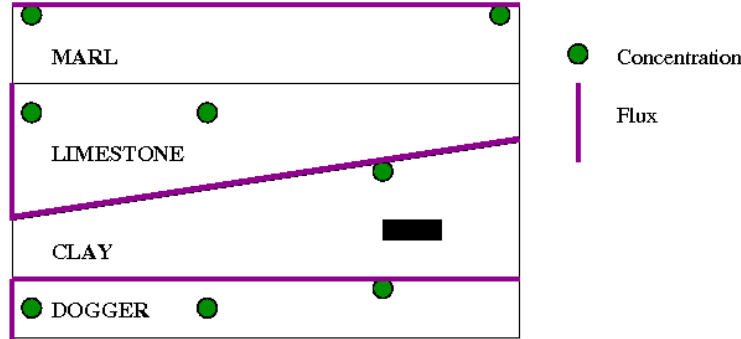


FIG. 2 – Variables d'intérêt : Concentration en Iode-129 en 8 points du domaine (vert) et Flux intégré en espace sur 5 frontières (violet) + sur le contour du domaine. Ces 14 variables dépendent du temps, elles ont été discrétisées sur 12 instants de  $10^3$  à  $10^7$  années.

Les variables d'intérêt sont représentées dans la Figure 2. Elles sont discrétisées aux 12 instants suivants exprimés en années :

$$T \in \{10^3, 5.10^3, 10^4, 2.10^4, 5.10^4, 10^5, 2.10^5, 5.10^5, 10^6, 2.10^6, 5.10^6, 10^7\} \quad (5)$$

L'objectif du cas test est le calcul, pour chacune de ces 14 variables, de leur moyenne, leur variance et leurs indices de Sobol (décomposition de la variance) pour les 12 instants retenus. Le cas test propose donc un total  $168 = 14 \times 12$  variables à analyser.

Le Tableau (1) donne l'ensemble de la modélisation probabiliste des paramètres incertains. Les valeurs nominales représentent le cas test déterministe Couplex. Il y a donc 15 paramètres incertains. Pour évaluer les méthodes de façon graduelle sur différents niveaux de complexité stochastique, le cas test a été spécifié de façon à pouvoir être utilisé en ne considérant qu'un sous ensemble de paramètres incertains. Le Tableau (2) définit les cas d'utilisation en spécifiant les incertitudes d'abord sur les perméabilités (4), les charges (5), les coefficients de diffusion et de dispersivité (6) puis sur les différentes combinaisons de ces 3 groupes de variables. Au total le bench propose 7 cas d'utilisation.

### 3 Code opuscouplex

Le cas test Couplex a été réalisé de façon à pouvoir être utilisé très simplement en mode non intrusif. Pour cela nous avons réalisé le logiciel **opuscouplex** gérant automatiquement les pré et post traitements des données.

De façon générique, le modèle du code se met sous la forme *boîte noire*  $Y = f(X)$  où  $X$  est le vecteur regroupant les paramètres incertains et  $Y$  les sorties. L'utilisation du cas test est très simple. Le programme lit les valeurs de  $X$  dans un fichier, lance le calcul et recopie les valeurs correspondantes de  $Y$  dans un fichier. Les entrées et les sorties sur fichier sont sous forme ascii et donc facilement exploitables.

Les simulations sont réalisées par l'outil CASTEM (CAST3M) téléchargeable depuis son site :

<http://www-cast3m.cea.fr/cast3m/index.jsp>

	Noms	Paramètres	Lois	nominale	min	max
1	PerDog	Permeabilite_Dogger	Log Uniforme	25	5	125
2	PerCla	Permeabilite_Clay	Log Uniforme	$3 \cdot 10^{-6}$	$3 \cdot 10^{-7}$	$3 \cdot 10^{-5}$
3	PerLim	Permeabilite_Limestone	Log Uniforme	6	1.2	30
4	PerMar	Permeabilite_Marl	Log Uniforme	$3 \cdot 10^{-5}$	$10^{-5}$	$10^{-4}$
5	DifCla	Diffusion_Effective_Clay	Log Uniforme	$10^{-6}$	$2 \cdot 10^{-7}$	$5 \cdot 10^{-6}$
6	DifMar	Diffusion_Effective_Marl	Log Uniforme	$5 \cdot 10^{-4}$	$2.5 \cdot 10^{-4}$	$10^{-3}$
7	DiLoDo	Dispersivite_Long_Dogger	Log Uniforme	50	25	100
8	DiLoLi	Dispersivite_Long_Limestone	Log Uniforme	50	25	100
9	DiTrDo	Dispersivite_Trans_Dogger	Log Uniforme	50	25	100
10	DiTrLi	Dispersivite_Trans_Limestone	Log Uniforme	50	25	100
11	DeCH21	Delta_Charge_H21	Uniforme	21	16	26
12	DeCH31	Delta_Charge_H31	Uniforme	51	41	61
13	DeCh16	Delta_Charge_H16	Uniforme	3	1	5
14	DeCh25	Delta_Charge_H25	Uniforme	110	100	120
15	DeCh34	Delta_Charge_H34	Uniforme	160	140	180

TAB. 1 – Quantification des paramètres incertains : les perméabilités des 4 couches géologiques, les coefficients de diffusion effective (Clay, Marl), les coefficients de dispersivité longitudinale et transverse (Limestone, Dogger) et les différences de charges hydrauliques (conditions aux limites) spécifiées par rapport à la charge  $H_1$  sur la frontière *est* du Dogger.

Cas	Perméabilité	Charge Hydraulique	Dispersivité Diffusion	Dimension stochastique
CT-K	4			4
CT-H		5		5
CT-T			6	6
CT-KH	4	5		9
CT-KT	4		6	10
CT-HT		5	6	11
CT-KHT	4	5	6	15

TAB. 2 – Différents cas d'études de complexité croissante en dimension stochastique.

Le programme **opuscouplex** gère automatiquement les entrées - sorties sans que l'utilisateur soit amené à modifier le fichier contenant le jeu de données de Castem. Les valeurs des paramètres caractérisant la modélisation dans les couches géologiques et celles des conditions aux limites du domaine sont lues dans un fichier au format ascii. Le jeu de données Castem est réactualisé automatiquement, le calcul Castem est ensuite lancé et enfin les valeurs des variables sur lesquelles portent l'analyse, sont recopiés dans un fichier au format ascii.

L'exécutable **opuscouplex** est issu d'un programme écrit en langage C et peut être recompilé très simplement sur tout type d'environnement (il ne nécessite pas de bibliothèques spécifiques). Le code peut être utilisé de 2 façons, pour réaliser une seule simulation ou bien pour réaliser un ensemble de simulations.

### 3.1 Compilation d'opuscouplex

Après avoir installé CAST3M, la procédure d'acquisition du code **opuscouplex** est la suivante :

1. Télécharger `toolsOpusCoupLex.tgz`
2. Dézipper et extraire les fichiers. Sous Linux faites `tar xzf toolsOpusCoupLex.tgz`
3. Le répertoire `toolsOpusCoupLex` a été créé. Ce répertoire contient 6 fichiers.
4. Ouvrir et modifier le fichier `opus_coupLex_env.h` afin d'adapter les variables d'environnement du code Castem

```
// Repertoire d'installation de Castem
char const ENVIRONNEMENT_CASTEM[] = "/home/user/Castem/castem_V09";
// Nom de l'exécutable Castem
char const BIN_CASTEM[] = "castem09";
// Repertoire d'installation d'opuscouplex
char const FICHIER_REFERENCE[] = "/home/user/toolsOpusCouples";
```

5. Ouvrir et modifier dans le fichier `makefile` la variable d'environnement `BIN` définissant le chemin d'accès où binaire **opuscouplex** sera installé.
6. Compiler en lançant, sous Linux, la commande `make`.

Un premier test peut être fait en tapant la commande `opuscouplex`. Vous devriez voir apparaître sur la sortie standard une documentation du code **opuscouplex**.

### 3.2 Réalisation d'une seule simulation

La réalisation d'un seul calcul se fait par la commande suivante

```
opuscouplex fileParam
```

où **fileParam** est un fichier contenant les valeurs des paramètres à modifier. Dans le fichier les valeurs sont précédées d'une chaîne de caractères précisant le nom du paramètre à modifier. Les noms des paramètres sont ceux qui figurent dans le Tableau (1).

Par exemple si le fichier **fileParam** contient :

```
PerDog 5.55
PerCla 3.33e-06
PerLim 6.66
PerMar 3.33e-05
```

la simulation sera réalisée en affectant aux perméabilités du Dogger, du Clay, du Lime et du Marl les valeurs respectives suivantes 5.55,  $3.33 \cdot 10^{-6}$ , 6.66,  $3.33 \cdot 10^{-5}$ . Les données peuvent être disposées sur une seule ligne ou sur plusieurs lignes comme dans l'exemple. Si le nom du paramètre ne figure pas dans la liste des paramètres figurant dans le Tableau (1), le programme signale l'erreur.

Après le calcul Castem, un post-traitement recopie les variables d'intérêt dans le fichier **resu\_post.dat**. Chaque ligne du fichier contient la valeur du temps et les valeurs correspondantes des 14 variables d'intérêt (8 concentrations et 6 flux). On obtiendra par exemple dans le fichier **resu\_post.dat** :

#TEMPS(annees)	CONC/P1	CONC/P2	...	FLUX/G_DOG	FLUX/H_MARL
1.000000e+03	0.000000e+00	0.000000e+00	...	0.000000e+00	0.000000e+00
5.000000e+03	6.783822e-31	0.000000e+00	...	-1.187698e-93	1.457795e-64
.....					
1.000000e+07	4.537000e-48	2.343000e-48	...	-1.140000e-48	7.880000e-47

Ce fichier peut donc aisément être exploité de façon à visualiser le profil temporel de chacune des 14 variables d'intérêt.

### 3.3 Réalisation de plusieurs simulations

La réalisation de plusieurs simulations se fait par la commande suivante :

```
opuscouplex filePlanExperience fileReponse
```

où **filePlanExperience** est un fichier contenant le plan d'expériences à réaliser et où les réponses associées seront sauvegardées dans le fichier **fileReponse**. L'entête du fichier **filePlanExperience** spécifie le nombre et le nom des paramètres variables et le nombre de simulations.

L'entête est suivie des réalisations des paramètres variables. L'exemple ci-dessous illustre la définition d'un plan de 193 expériences numériques à réaliser sur les perméabilités des couches Dogger, Clay, Lime et Marl.

```
#Parametres 4
PerDog          PerCla          PerLim          PerMar
#Simulations 193
2.500000e+01    3.000000e-06    6.000000e+00    3.162278e-05
2.500000e+01    3.000000e-06    6.000000e+00    7.714332e-05
2.500000e+01    3.000000e-06    6.000000e+00    1.296288e-05
.....
```

La spécification d'un plan d'expériences suit donc la règle suivante :

1. le nombre de paramètres  $nx$ ,
2. la liste des noms des  $nx$  paramètres définis dans le Tableau (1)
3. le nombre  $np$  de simulations,
4. les valeurs des paramètres sous forme d'une matrice  $np \times nx$

Après le calcul Castem, un post-traitement recopie les variables d'intérêt dans le fichier **fileReponse** spécifié par l'utilisateur. Chaque ligne du fichier contient les valeurs des 168 variables d'intérêt (8 concentrations et 6 flux aux 12 instants retenus). Les 14 premières valeurs correspondent au premier instant ( $t(1) = 10^3$  années), les 14 suivantes au second instant ( $t(2) = 5.10^3$  années), et ainsi de suite jusqu'au dernier instant ( $t(12) = 10^7$  années). L'entête du fichier **fileReponse** indiquera alors le nombre de simulations et le nombre de réponses (dans notre cas égal à 168).

```
#NombreSimulations 81
#NombreSorties      168
..... 81 lignes de 168 valeurs ...
```

La valeur de la variable d'intérêt numéro  $i = 1, \dots, 14$  à l'instant  $t = 1, \dots, 12$  correspond donc à la valeur de rang  $i + (t - 1) * 14$ .

D'autre part, l'utilisateur peut, s'il le souhaite, rappeler dans le fichier **fileReponse** les valeurs des entrées (contenues dans le fichier **filePlanExperience**) par :

opuscouplex filePlanExperience fileReponse 1

et dans ce cas, l'entête du fichier **fileReponse** indiquera également le nombre d'entrées c'est à dire le nombre de paramètres ayant été modélisés par des variables aléatoires.

```
#NombreSimulations 81
#NombreEntrees      4
#NombreSorties      168
..... 81 lignes de 4 + 168 valeurs ...
```

Dans ce cas, la valeur de la variable d'intérêt numéro  $i = 1, \dots, 14$  à l'instant  $t = 1, \dots, 12$  correspond donc à la valeur de rang  $4 + i + (t - 1) * 14$ .

### 3.4 Objectif du cas test

Nous avons donné à ce cas test sa description la plus complète, c'est à dire celle du bench Couplex stochastique proposé par le groupe de travail *incertitudes* du Groupement MOMAS, bench visant à évaluer une large classe de méthodes de propagation et d'analyses d'incertitudes sur le cas Couplex.

Dans le cadre OPUS, pour réduire la complexité combinatoire du problème, nous nous sommes limités à l'analyse de 3 grandeurs d'intérêt : les concentrations aux 2 points proches de la surface du domaine, c'est à dire au points P7(x=50m, z=692m) et P8(x=24800m, z=692m) du Marl et le flux à travers la surface du domaine, c'est à dire en haut du Marl. Ces trois grandeurs ont respectivement comme rangs 7, 8 et 14 :

1. Concentration en P7 :  $\text{output}(7 + (t - 1) * 14), t = 1, \dots, 12$
2. Concentration en P8 :  $\text{output}(8 + (t - 1) * 14), t = 1, \dots, 12$ ,
3. Flux sur le haut du Marl :  $\text{output}(14 + (t - 1) * 14), t = 1, \dots, 12$

Compte tenu des plages de variation des valeurs de la concentration et du flux, les analyses d'incertitudes et de sensibilité ont été faites sur le **logarithme** (décimal) des valeurs.

Pour chacune de ces 3 variables transformées, on calcule la moyenne, la variance, et les indices de Sobol du premier ordre et totaux pour les 12 instants définis dans l'Equation (5).

## 4 Résultats obtenus sur le cas CT\_KT

Les résultats présentés dans les Figures (3, 4) sont ceux du cas CT\_KT (10 paramètres incertains) en utilisant les polynômes de chaos via la librairie NISP. Le nombre de variables aléatoires est 10. Le degré maximal a été fixé à 5. Le plan d'expériences numériques spécifié par la librairie SmolPack intégrée dans Nisp nécessite 19105 simulations. Cela nous a pris 4 jours de calculs (1 calcul Castem dure environ 20s).

Les Figures (3) donnent l'évolution de la moyenne avec un intervalle de confiance (approché) à un sigma du logarithme décimal du flux en haut du domaine (couche géologique Marl), la première extraite d'un polynôme de degré 4, la seconde d'un polynôme de degré 5. Les résultats de la moyenne sont très proches. Ceux de la variance sont proches dans la plage temporelle centrée autour de  $10^6$  années correspondant aux valeurs maximales du flux. Les écarts de variance en début et fin de la période de simulation ne sont pas significatifs car ils correspondent à des valeurs du flux extrêmement petites bien en deça du nombre d'Avogadro ( $\simeq 6.10^{23}$ ).

La Figure (4) donne l'évolution temporelle des indices de Sobol du premier ordre et les indices totaux pour le polynôme de degré 5. Les paramètres les plus influents sont la perméabilité et la diffusion effective du Clay. On remarque qu'en début de stockage, c'est la diffusion effective qui est le paramètre le plus influent puis à partir de  $5.10^5$  années c'est la perméabilité qui devient le facteur le plus influent.

Entre les paramètres les plus influents (perméabilité et diffusion effective du Clay), le passage de l'indice du premier ordre à l'indice total fait apparaître que le paramètre de diffusion effective présente un niveau d'interaction avec les autres paramètres plus important que le paramètre de perméabilité.

## 5 Conclusions

L'utilisation des polynômes de chaos sur ce cas test *industriel* s'est avérée aussi simple que celle sur les cas académiques que nous avons commencé à utiliser à partir de la rentrée 2009 dans nos différents enseignements (Master, Ecoles). Pour les cas tests industriels, le seul problème est celui de la supervision des calculs. Nous l'avons ici facilité en proposant le gestionnaire de calculs **opuscouplex** permettant de gérer les pré et post-traitements de données et téléchargeable depuis le site OPUS.

Cet étude souligne la nécessité de développer des méthodes fondées sur la réduction, le contrôle ou l'adaptation du développement fonctionnel des interactions entre variables. On retrouve le problème classique de la complexité combinatoire variant de façon exponentielle avec le nombre de variables.

Ce cas test peut donc également être utilisé pour évaluer les méthodes de construction de méta-modèles ou d'estimation de quantiles extrêmes. L'objectif étant de fournir des estimateurs de faible variance en réduisant le nombre d'appels au code de calcul.

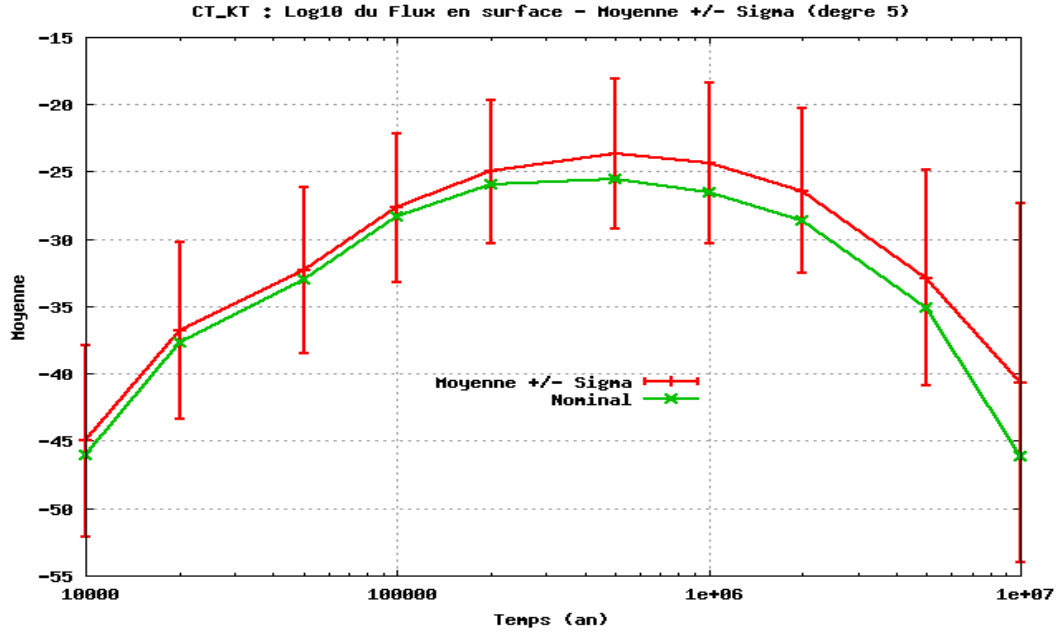
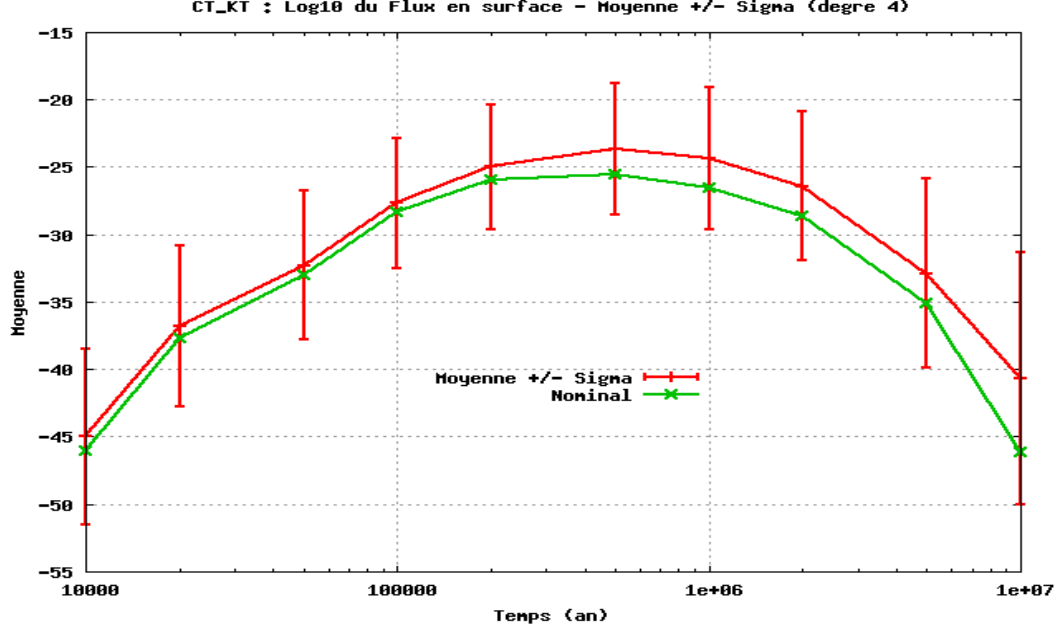


FIG. 3 – Cas CT\_KT : Calcul de l'évolution de la moyenne +/- un écart-type du logarithme décimal du flux en surface du Marl (mol/an). Degré maximal 5  $\rightarrow$  19105 simulations. Figure du haut polynôme de degré 4, celle du bas degré 5. Résultats sont proches autour des valeurs maximales centrées sur  $10^6$  années. Des écarts apparaissant sur les bornes de l'intervalle, *visuels* mais peu représentatifs compte tenu des valeurs du flux (rappelons le nombre d'Avogadro  $\simeq 6.10^{23}$  atomes/mol).



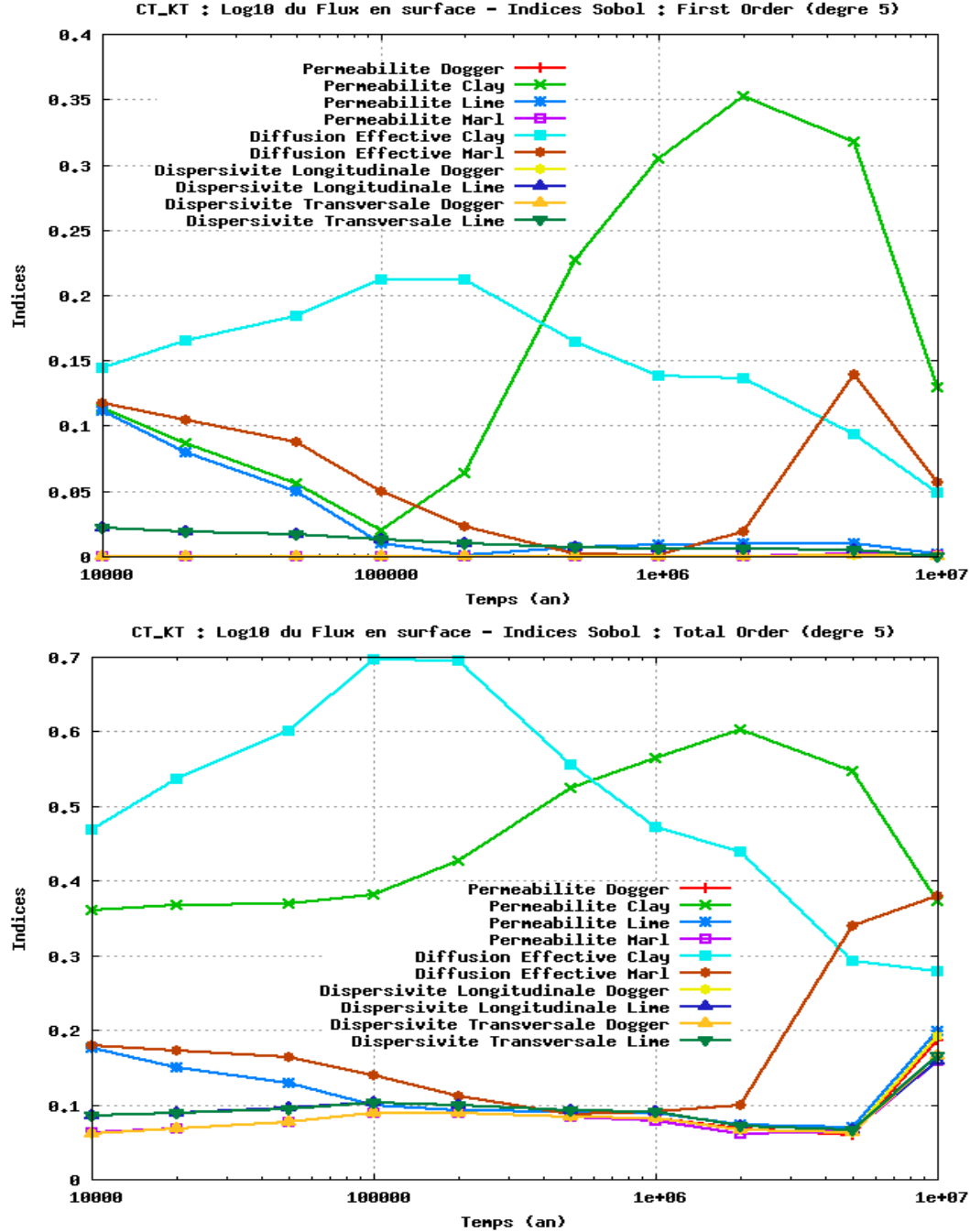


FIG. 4 – Cas CT\_KT : Calcul de l'évolution des indices de Sobol du premier ordre (haut) et totaux (bas) du logarithme décimal du flux en surface du Marl (mol/an). La décomposition fonctionnelle de la variance a été obtenue sur un chaos de degré 5 basé sur les polynômes de Legendre. La dimension stochastique est 10. Le nombre de coefficients du chaos polynomial est 3003. La construction de Smolayk utilisée pour le calcul des coefficients via NISP, a nécessité 19105 simulations.