

Méthodes statistiques et exécution de plusieurs simulations

Nathalie Moitrier

Content

1	Introduction	5
2	Spécificités	5
3	Fonctionnement général	6
3.1	Configuration et fonctionnement communs à ces divers traitements	7
3.1.1	Accès à la documentation du paquetage R utilisé pour la méthode sélectionnée	8
3.1.2	Sélection des paramètres d'entrées et variables de sorties	9
3.1.3	Temps et profondeurs	14
3.1.4	Modification du script R	18
3.1.5	Exécution en parallèle	19
3.1.6	Déportation des calculs sur des serveurs	19
3.1.7	Synthèse du paramétrage du traitement	19
3.1.8	Exportation du répertoire de sorties des statistiques ou de multi-simulation	20
4	Exécution multiple d'un modèle	23
4.1	Présentation générale	23
4.2	Fonctionnement global de l'exécution multiple	23
5	Analyse de sensibilité	28
5.1	Présentation générale	28
5.2	Fonctionnement global de l'analyse de sensibilité dans la plateforme . . .	28
5.2.1	Paramétrage de l'analyse de sensibilité	28
5.2.2	Sélection des paramètres d'entrées et configuration	31
5.2.3	Sélection des sorties	33
5.2.4	Exécution de l'analyse de sensibilité	33
5.2.5	Résultats de l'analyse de sensibilité	33
6	Estimations de paramètres	34
6.1	Présentation générale	34
6.2	Fonctionnement global de l'estimation de paramètres dans la plateforme .	34
6.3	Les méthodes d'estimation	35
6.3.1	Configuration et paramétrage des méthodes d'estimation	35

6.4	Les données observées	40
6.5	Estimation	41
6.5.1	Sorties et observations	41
6.5.2	Estimation simple	44
6.5.3	Estimation sur de la multi-configuration	44
6.5.4	Estimation des paramètres pour une ou toutes les configurations .	46
6.5.5	Sélection des sorties et de leur observations associées	47
6.5.6	Exécution de l'estimations de paramètres	47
6.5.7	Résultats de l'estimation de paramètres	47

List of Figures

1	Présentation générale traitement statistique/batch	6
2	Sélection du traitement, méthode	7
3	Paramètres globaux pour la méthode fast99	7
4	Accès à la documentation des méthodes statistiques	8
5	Liste des modules du modèle, avec symbolisation des modules présentant des paramètres ou des variables de sorties impliqués dans le traitement/méthode choisi.	9
6	Sélection d'un paramètre	9
7	Sélection pour un paramètre profilé	10
8	Sélection d'un paramètre vectorisé	10
9	Sélection des paramètres dans un vecteur	11
10	Valeurs du paramètres temps	12
11	Sélection d'une sortie dans la partie Initial outputs values	12
12	Sélection d'une sortie	13
13	Menu contextuel pour la sélection des valeurs initiales d'une sortie	13
14	Valeur initiale pour la profondeur 0 sélectionnée	13
15	Sélection d'une variable de sortie pour la multi-simulation	13
16	Sélection d'une variable de sortie "tagguée" pour la multi-simulation	14
17	Sélection des temps et des profondeurs	14
18	Sélection des temps souhaités pour les sorties	15
19	Paramétrage et génération automatique de temps	16
20	Paramétrage pour la génération automatique de temps	17
21	Sélection des points de grilles souhaités pour les sorties profilées	17
22	Erreur : Il manque une ou des profondeurs pour les sorties	18
23	Accès à la fonction de modification du script R	18
24	Synthèse du paramétrage du traitement statistique / multi-simulation	20
25	Message d'alerte pour l'exportation des résultats	21
26	Message d'information de l'exportation des résultats	21
27	Liste des fichiers résultats issue d'un export après un traitement statistique ou de multi-simulation	21
28	Paramètres liés à l'option Batch	23
29	Exemple fichier contenant les valeurs de paramètres pour la méthode "Batch"	24
30	Saisie de l'index de la colonne du fichier correspondant au paramètre cth0	25
31	Sauvegarde des fichiers de sorties dans un dossier	25
32	Contenu du répertoire de sauvegarde, les fichiers contenant les valeurs calculées des sorties sont encadrées en rouge.	26
33	Erreur d'index : chaque paramètre doit avoir un index de colonne du fichier attribué et distinct	26
34	Erreur d'index : l'index de la colonne doit exister dans le fichier.	27
35	Les méthodes d'analyse de sensibilité disponibles sont encadrées de rouge.	28
36	Aide contextuelle sur un paramètre de la méthode d'analyse de sensibilité fast99.	29
37	Paramétrage des entrées pour Morris	32
38	Paramétrage des entrées de type "int-choice" pour Sobol	33
39	Accès aux résultats de l'analyse de sensibilité	34

40	Configuration des paramètres généraux de la méthode Levenberg-Marquardt (LM)	35
41	Configuration des paramètres généraux de la méthode Shuffle Complex Evolution - University of Arizona (SCE-UA)	36
42	Configuration des paramètres généraux de la méthode	37
43	Configuration des paramètres généraux de la méthode	37
44	Configuration de certains paramètres généraux pour DREAM	38
45	Configuration de certains paramètres généraux pour DREAM	38
46	Configuration des sorties pour une estimation simple	42
47	Configuration des sorties pour une configuration multiple	43
48	Exécution du modèle pour une estimation avec des mesures moyennées .	44
49	Sélection de l'estimation en multi-configuration	45
50	Visualisation du code variable de la sortie soil temperature	45
51	Visualisation du contenu du répertoire pour l'estimation sur 3 fichiers de configuration et 2 sorties	46
52	Configuration pour que le paramètre soit estimé pour chaque configuration	47
53	Accès aux résultats de l'estimation de paramètres	48
54	Superposition des données observées et simulées	49

1 Introduction

La plateforme propose différentes méthodes pour réaliser : (1) des estimations de paramètres, (2) des analyses de sensibilité des modèles et (3) des exécutions successives (batch) d'un modèle en modifiant la valeur de certains paramètres. Ces méthodes sont accessibles dans les applications `vsoil_modules`, `vsoil_models` et `vsoil_player`.

La plateforme utilise pour (1) l'estimation de paramètres et (2) l'analyse de sensibilité des méthodes développées par des tiers. Ce document explique comment les utiliser dans la plateforme, mais ne précise pas :

- Quelle méthode utiliser en fonction de vos objectifs et contraintes.
- Le fonctionnement et les fondements de ces méthodes. Vous devez vous référer aux documentations « officielle ».

2 Spécificités

1. Sous le système WINDOWS : Si vous utilisez la plateforme sous le système Windows, vous pourrez préparer le traitement que vous souhaitez (choix de la méthode, spécification des paramètres, etc..) mais vous ne pourrez pas exécuter sur votre machine les calculs, que ce soit pour l'analyse de sensibilité ou la multi-simulation. En effet ces exécutions sont réalisées uniquement sous les environnements Linux. Toutefois, la plateforme offre la possibilité de préparer les calculs et de les déporter sur des machines distantes (cf. manuels "model usage", "user manual") puis de récupérer les résultats.
2. Visualisation des sorties : Dans le cas des traitements statistiques et de la multi-simulation, l'onglet "Plot" sera inaccessible. La visualisation des résultats sera accessible sous l'onglet "Run".

3 Fonctionnement général

Pour accéder et utiliser les différentes méthodes disponibles, la plateforme est « connectée » au langage R au sein duquel sont disponibles les méthodes proposées. Cette connexion est transparente pour l'utilisateur. La plateforme se charge de :

- créer le script R adapté au traitement sélectionné par l'utilisateur,
- d'installer les paquets R nécessaires,
- d'exécuter le script et de récupérer les résultats.

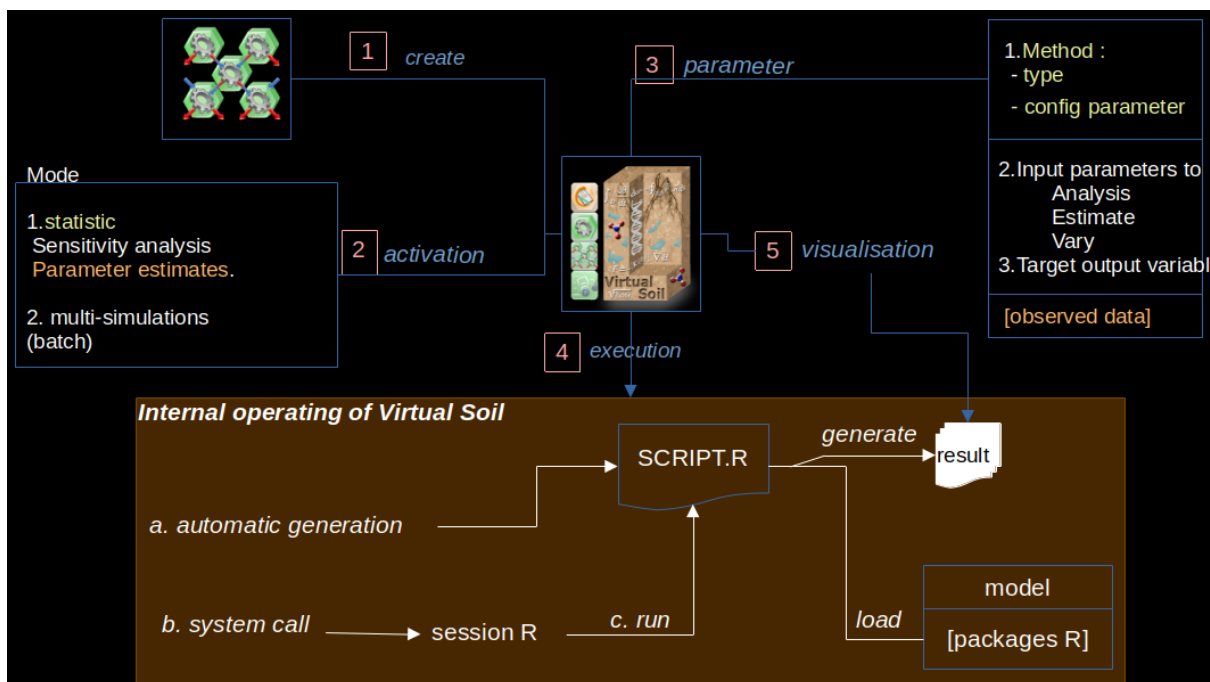


Figure 1: Présentation générale traitement statistique/batch

Les traitements proposés :

- exécution successive d'un modèle,
- analyse de sensibilité aux paramètres,
- estimation de paramètres.

Le choix de la méthode se fait dans l'onglet « Initialization. » puis dans l'onglet « Global » sous le titre « Statistical analysis » (partie droite de l'écran) :

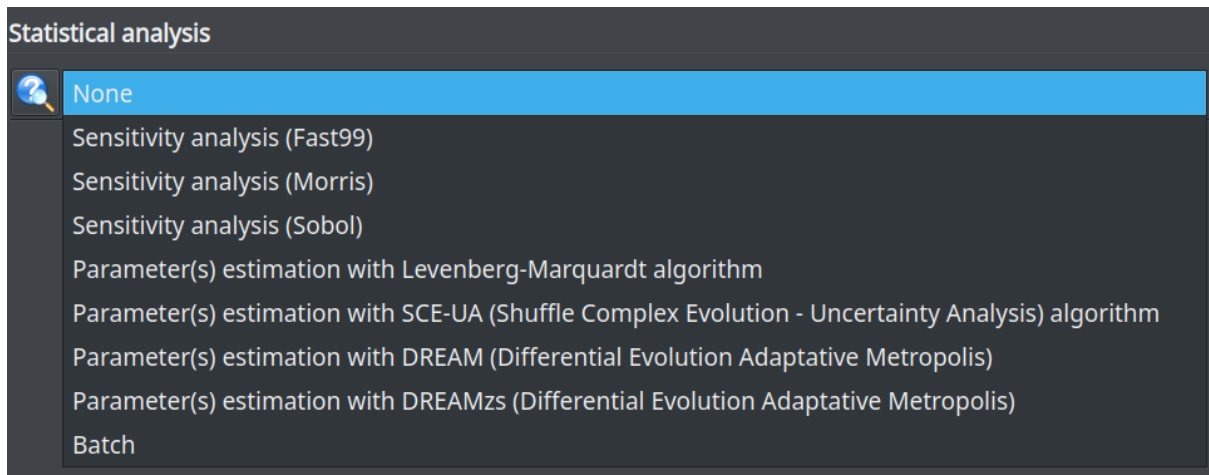


Figure 2: Sélection du traitement, méthode

Par défaut, l'option « None » est sélectionnée. Avec cette option par défaut, on a une exécution du modèle utilisant les paramètres et les situations initiales fournis. Si vous souhaitez :

- faire de l'exécution multiple pour le modèle sélectionné, choisissez l'option « Batch »,
- faire de l'analyse de sensibilité, sélectionnez une des méthodes : Fast99 ou Morris ou Sobol,
- faire de l'estimation de paramètres, sélectionnez une des méthodes : Levenberg-Marquard , Differential Evolution Adaptive Metropolis(DREAM ou DREAMzs) , Shuffle Complex Evolution.

3.1 Configuration et fonctionnement communs à ces divers traitements

Lorsque vous aurez sélectionné un traitement/méthode, la configuration de leurs paramètres généraux sera accessible en dessous de la zone de sélection du traitement/méthode.

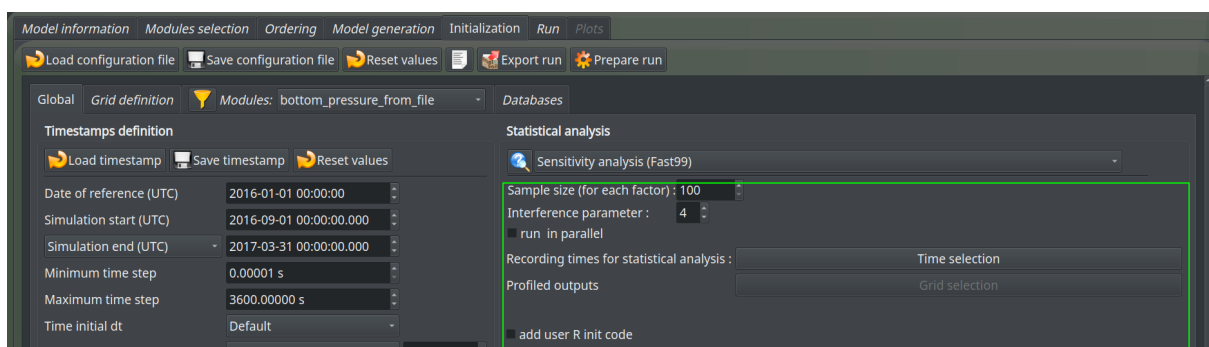


Figure 3: Paramètres globaux pour la méthode fast99

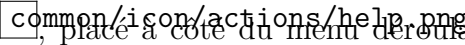
Dans ce chapitre nous allons aborder les fonctionnements communs quelque soit le traitement sélectionné.

Nous expliquerons comment

1. accéder à la documentation des paquages R des méthodes statistiques proposées par la plateforme.3.1.1
2. sélectionner et paramétrer les variables du modèle qui sont utilisées comme entrée ou sortie dans la méthode statistique ou traitement en multi-simulation. 3.1.2
3. sélectionner soit les temps d'exécutions du modèle pour le traitement statistique ou de multi-simulation soit les profondeurs souhaitées pour les variables de sorties.3.1.3
4. modifier le code de traitement statistique ou de multi-simulation avant son exécution.3.1.4
5. paralléliser l'exécution des traitements statistiques ou de multi-simulations.3.1.5
6. déporter les calculs sur un serveur/machine distante.3.1.6
7. interpréter la synthèse de la configuration de la méthode statistique avant son exécution.3.1.7
8. exporter les sorties produites lors des traitements statistiques ou de multi-simulations.45

3.1.1 Accès à la documentation du paquetage R utilisé pour la méthode sélectionnée

Pour accéder à la documentation du paquetage R qui implémente la méthode :

- Cliquez sur le bouton  , placé à côté du menu déroulant permettant de choisir une méthode.

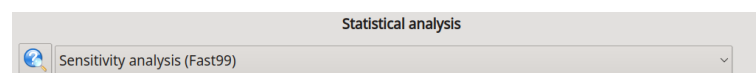


Figure 4: Accès à la documentation des méthodes statistiques

- La documentation est affichée via un outil externe d'affichage de document au format pdf)

Synthèse des méthodes disponibles et accès à la documentation de leur implémentation informatique :

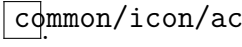
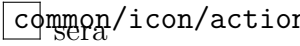
- Analyse de sensibilité Fast99, Morris et Sobol, paquet sensitivity
- Estimation de paramètres
 1. Levenberg-Marquardt , paquet minpack.lm
 2. Shuffle Complex Evolution , paquet rtop

3. Dream,
 - (a) Dream de Jasper Vrugt [aut] (Matlab original), Joseph Guillaume [aut, cre] (R implementation and added functionality), Felix Andrews [aut] (R implementation and added functionality)
Pour information, ce paquet n'est pas sur le CRAN et n'est plus maintenu.
 - (b) DreamZS, paquet BayesianTools

3.1.2 Sélection des paramètres d'entrées et variables de sorties

Dès qu'une méthode d'analyse de sensibilité, d'estimation de paramètres ou de multi-simulations est choisie, vous pourrez sélectionner les paramètres mis en jeu et les sorties du modèle sur lesquelles sont faites les traitements.

Leur sélection est identique quel que soit le traitement sélectionné. Ce qui diffère, c'est la nature des informations qui sont demandées. Cela dépendra du traitement sélectionné et de la méthode choisie.

L'icône suivant symbolise l'implication dans une des méthodes disponibles : . Lorsque un paramètre et/ou une sortie d'un module sont sélectionnés, l'icône  sera positionné avant le nom du module - dans la liste des modules contenus dans le modèle- :

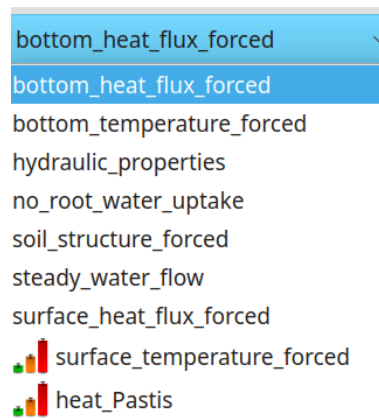


Figure 5: Liste des modules du modèle, avec symbolisation des modules présentant des paramètres ou des variables de sorties impliqués dans le traitement/ méthode choisi.

1. Les paramètres d'entrées

Des paramètres scalaires, vectorisés ou profilés sont sélectionnables pour les méthodes proposées. Les valeurs initiales des sorties peuvent aussi être sélectionnées, exceptées pour la méthode de multi-simulation. La sélection se fait comme suit :

Sélectionnez le module souhaité. Parmi les paramètres du module affiché dans l'interface graphique, un paramètre « sélectionnable » présente une case à cocher à sa gauche et l'icône de paramétrage à sa droite. La sélection d'un paramètre se fait en cochant la case qui le précède. L'icône de paramétrage est actif quand le paramètre est sélectionné.

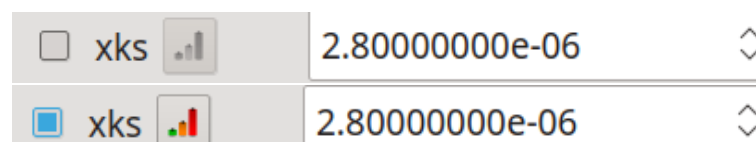


Figure 6: Sélection d'un paramètre

Cas particulier :

(a) Paramètres "profilés"

Si vous avez défini plusieurs couches, sélectionnez les paramètres dans les couches souhaitées :

- Sélectionnez le numéro de la couche (onglet horizontal).
- Sélectionnez comme précédemment les paramètres souhaités.

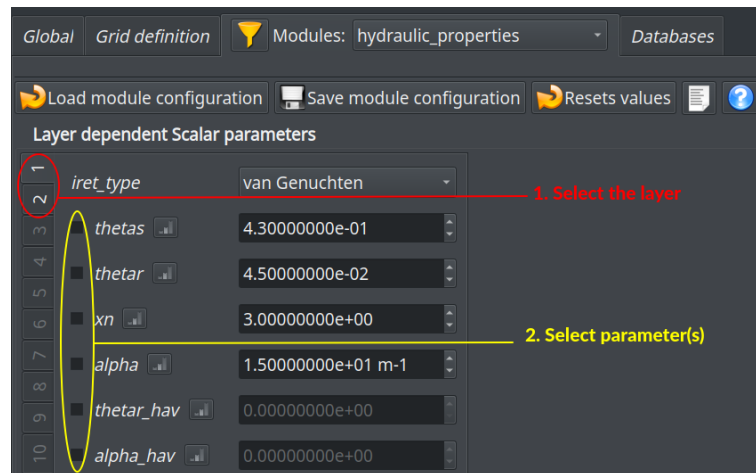


Figure 7: Sélection pour un paramètre profilé

(b) Paramètres vectorisés

Un paramètre peut être associé à plusieurs valeurs. Par exemple, dans le cas du module `surface_heat_flux_forced`, la valeur des paramètres flux dépendent du temps. Sélectionnez dans un premier temps le paramètre :

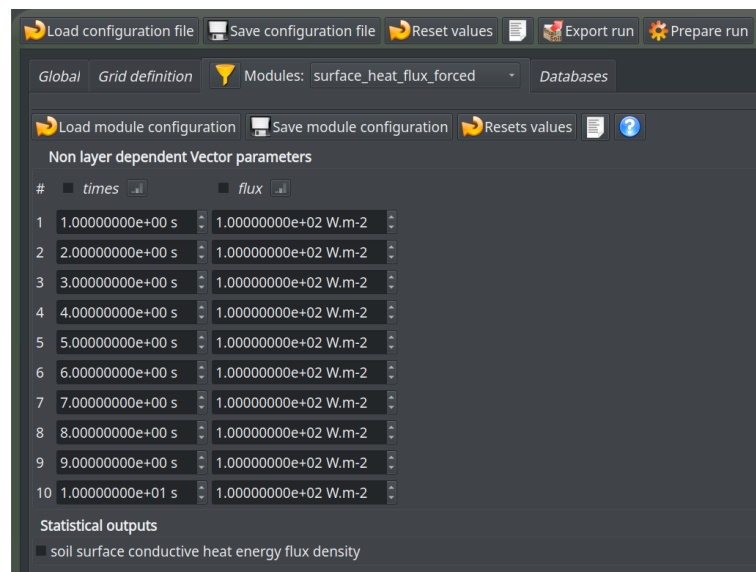



Figure 8: Sélection d'un paramètre vectorisé

Une fois le paramètre sélectionné, il faut ensuite identifier quel(s) composant(s) (quelle(s) valeur(s)) le traitement utilisera.

- Cliquez sur le bouton .
- Une fenêtre de sélection s'affiche :

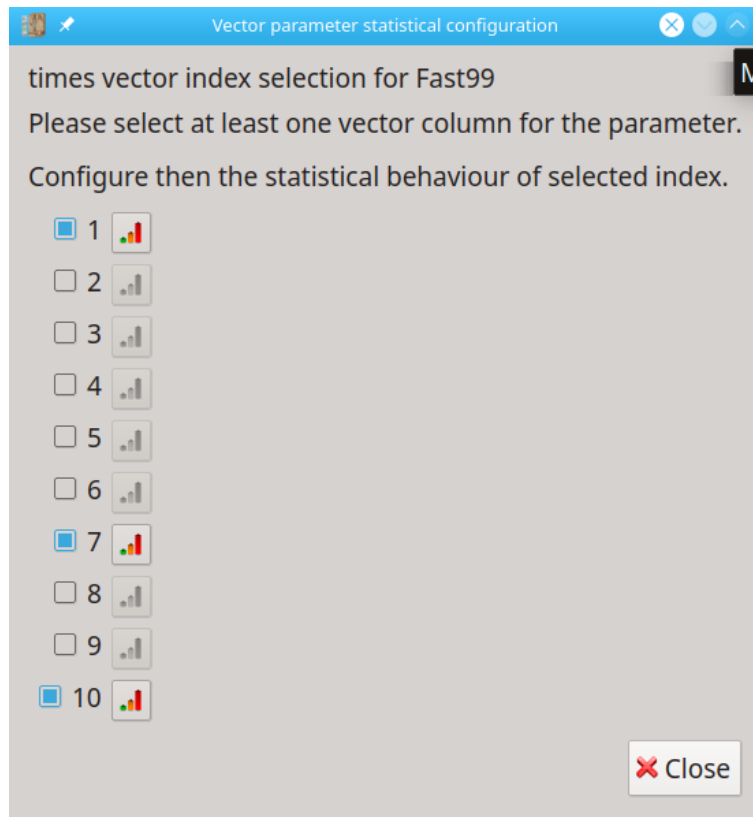


Figure 9: Sélection des paramètres dans un vecteur

- Sélectionnez les index des paramètres souhaités. Ces index correspondent aux valeurs du paramètres saisis dans la partie module :

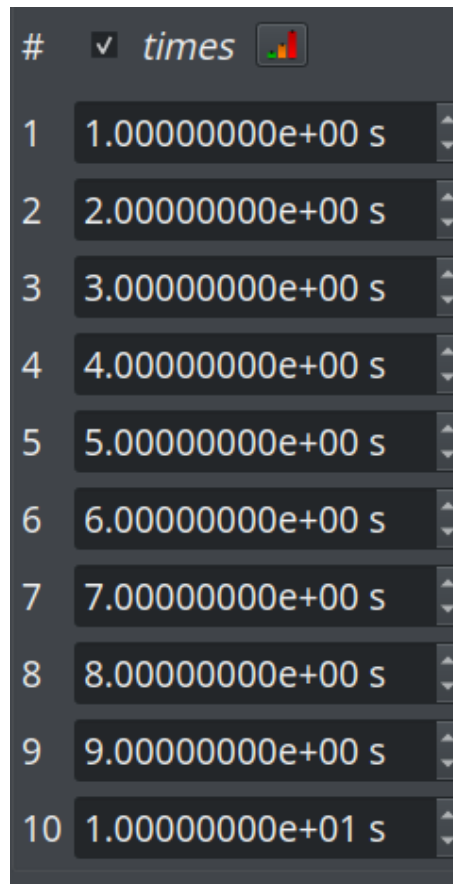


Figure 10: Valeurs du paramètres temps

- Puis pour chaque index sélectionné, paramétrer-le en cliquant sur .

(c) Valeur initiales des sorties

- Sélectionnez la variable de sortie dans la partie « Initial outputs values » en cochant la case qui la précède.

Par exemple, dans le module heat_Pastis, sélectionnez dans la partie “Initial output vector values”, sélectionnez soil temperature :

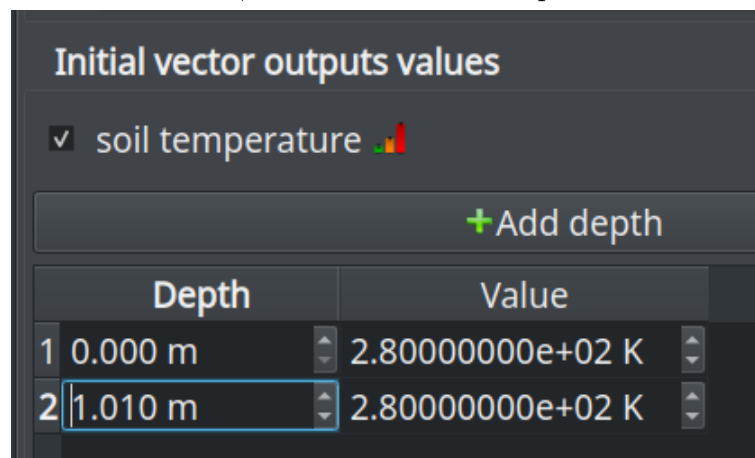



Figure 11: Sélection d’une sortie dans la partie Initial outputs values

L’icône  apparaît pour préciser que les valeurs initiales de cette variable

peuvent être utilisées comme paramètres dans un traitement. Dans ce cas précis, la sélection va se réaliser au cas par cas sur les valeurs initiales.

	Depth		Value	
1	0.000 m	◇	2.80000000e+02 K	◇
2	1.010 m	◇	2.80000000e+02 K	◇

Figure 12: Sélection d'une sortie

- En dessous de la variable, pour la profondeur désirée cliquez (bouton droit) dans la case de la valeur que vous souhaitez utiliser comme paramètre. Un menu contextuel s'affiche :

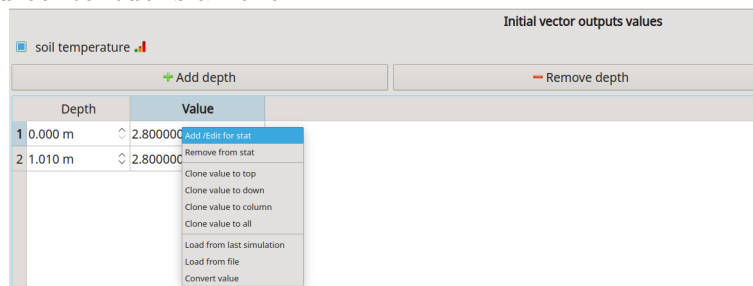


Figure 13: Menu contextuel pour la sélection des valeurs initiales d'une sortie

- Sélectionnez dans ce menu « Add/Edit » for stat.
- Saisissez les informations du paramètre en fonction du traitement/méthodes sélectionnées (cf chapitre dédié traitement/méthodes) et validez.
- La cellule de la valeur sélectionnée va être colorée pour indiquer que cette valeur sera utilisée comme paramètre.

	Depth		Value	
1	0.000 m	◇	2.80000000e+02 K	◇
2	1.010 m	◇	2.80000000e+02 K	◇

Figure 14: Valeur initiale pour la profondeur 0 sélectionnée

2. Sélection des variables de sorties

- Allez dans l'onglet du module calculant la variable.
- Dans le cas de traitements statistiques et multi-simulation, une nouvelle partie « Statistical outputs » est présente dans l'interface graphique pour chaque module. Cette partie se trouve après la partie de saisie des paramètres et des situations initiales. Elle présente la liste des variables pouvant être utilisées. Cochez la case à gauche du nom de la variable de sortie que vous souhaitez sélectionner. Vous pouvez sélectionner plusieurs variables dans plusieurs modules.

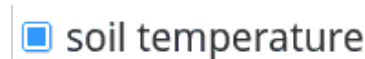


Figure 15: Sélection d'une variable de sortie pour la multi-simulation

Si la variable est “taggée” (“étiquetée”), sélectionnez les tag souhaités en sortie :



Figure 16: Sélection d’une variable de sortie “tagguée” pour la multi-simulation

3.1.3 Temps et profondeurs

Pour l’exécution multiple ou l’analyse de sensibilité, dans l’onglet « Initialization » puis « Global » et l’espace « Statistical analysis », vous pourrez sélectionner les temps et les profondeurs souhaitées. Cela signifie que l’analyse de sensibilité sera faite pour toutes les profondeurs et temps donnés. Il en va de même pour l’exécution multiple du module qui stockera les résultats pour les temps et profondeurs donnés.

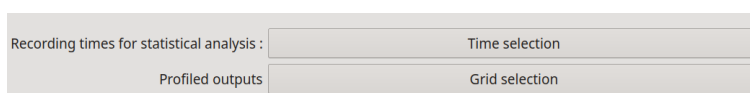


Figure 17: Sélection des temps et des profondeurs

L’option de sélection des points de grilles est accessible uniquement dans le cas où une des sorties sélectionnée est du type *profile*.

Dans le cas de l’estimation des paramètres, les temps et profondeurs seront ceux issus des fichiers d’observations.

1. Sélection des temps

- cliquez sur le bouton « Time selection »

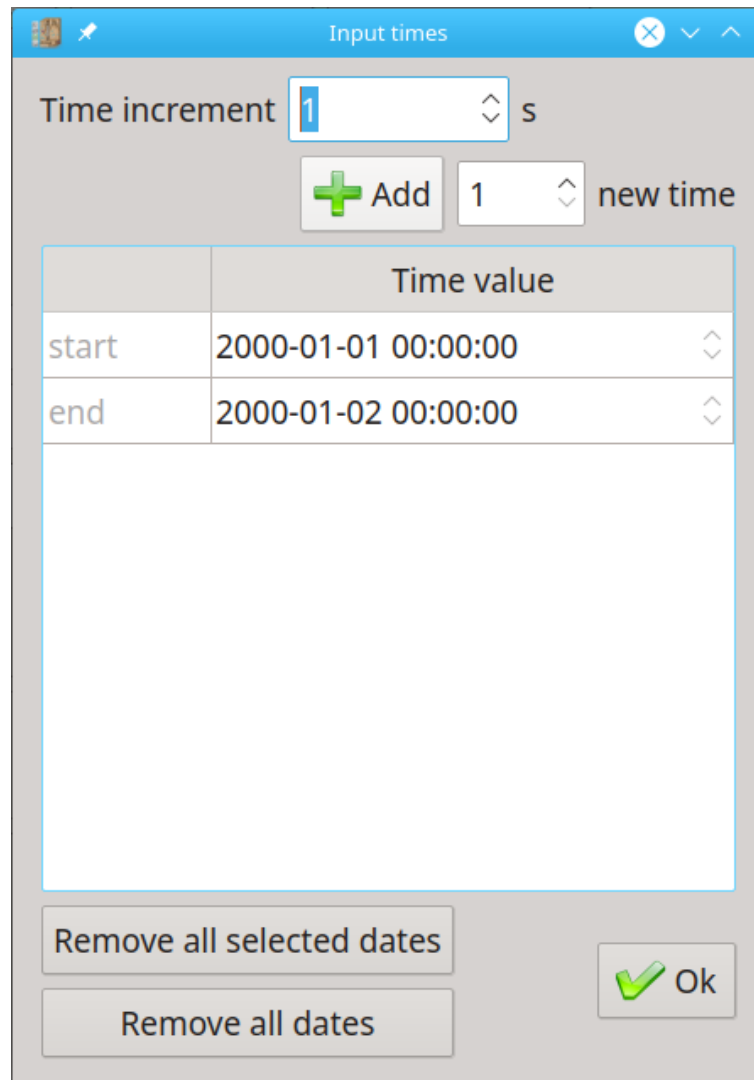


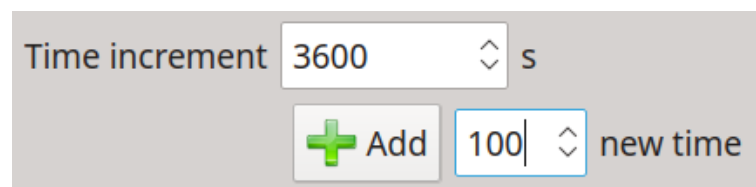
Figure 18: Sélection des temps souhaités pour les sorties

- saisissez les temps souhaités.

Par défaut deux temps sont positionnés et non modifiables : les temps de début et de fin de la simulation.

Pour accélérer la saisie des temps, il est possible de générer automatiquement le nombre de temps souhaités dans la période de simulation.

Par exemple pour générer 100 temps toutes les heures :



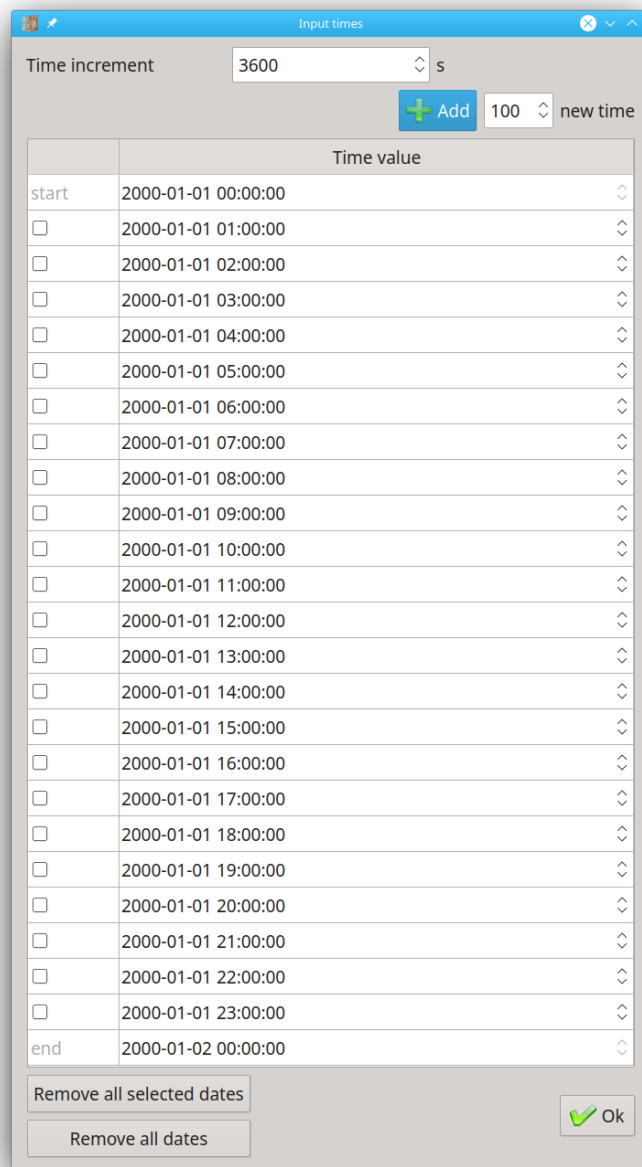


Figure 19: Paramétrage et génération automatique de temps

Si la date de fin de la simulation est dépassée alors un message d’alerte vous informera du nombre de temps réellement inséré :

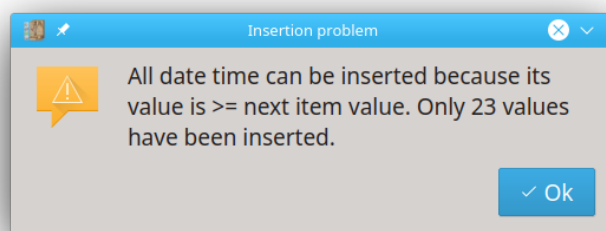


Figure 20: Paramétrage pour la génération automatique de temps

2. Sélection des profondeurs

- Cliquez sur le bouton « Grid selection » :

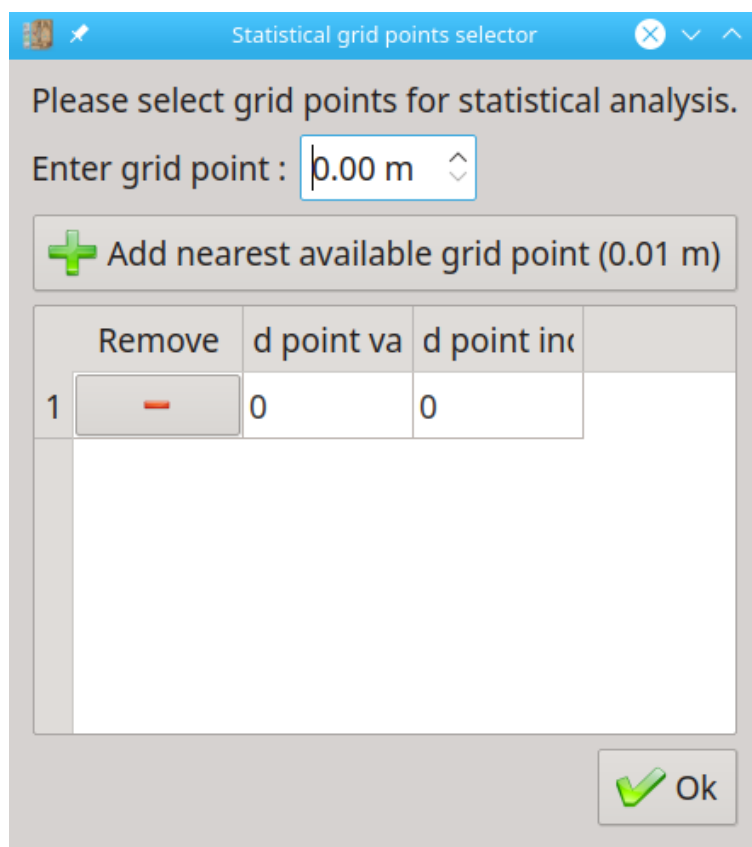


Figure 21: Sélection des points de grilles souhaités pour les sorties profilées

- Saisissez la/les profondeurs souhaitées.
Par défaut aucune profondeur n'est saisie. Dans le cas d'une sortie du type *profile* au moins une profondeur doit être saisie. Dans le cas contraire une erreur sera générée lors du « Prepare run ». Vous pouvez sélectionner autant de profondeurs que de points présents dans la la grille.

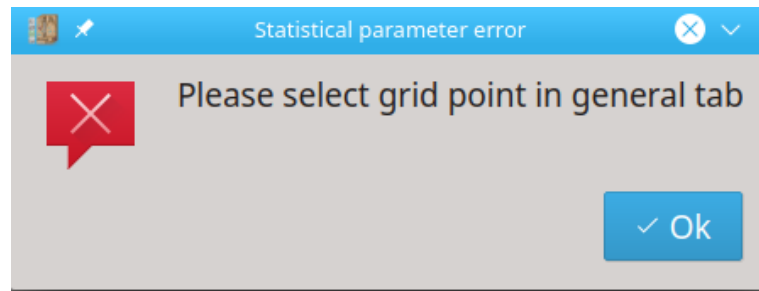


Figure 22: Erreur : Il manque une ou des profondeurs pour les sorties

3.1.4 Modification du script R

Actuellement il y a deux façons de modifier le script R. La première est l'ajout de code dans la partie d'initialisation du script R. C'est une modification qui est ciblée et sauvegardée. Elle est réalisée lors de la phase d'initialisation du modèle.

La seconde est la modification possible dans le code R script généré, donc réalisée après sa génération et donc non sauvegardé par la plateforme.

Il est à noter qu'il n'y a pas de vérification du code saisi, avant son exécution. Il est donc conseillé de tester les insctuction via R studio ou R.

1. Ajout de code R dans la partie initialisation Lors de la configuration de votre simulation, vous pouvez rajouter du code dans la partie initialisation du script R généré.Ce code sera exécuté dans la partie initialisation. Cette partie a été rajoutée pour pallier à un problème sur une des méthodes statistique d'un paquet R. L'objectif est de pouvoir charger une version modifiée de la fonction à la place de la méthode "boguée". exemple du code
2. Modification du code R du script généré Une fois le paramétrage du traitement réalisé, il est possible de modifier le script R et de l'exécuter dans la plateforme. Cette possibilité est à utiliser avec « modération » car aucun contrôle du script modifié n'est effectué. L'objectif est de pouvoir rajouter une écriture ou un élément de sortie pour expliciter un résultat... Si vous avez besoin d'une nouvelle méthode ou d'éléments de sortie supplémentaires, contactez l'équipe de développement (vsoil-support@inrae.fr).

Pour modifier le script R :

- Dans l'onglet « Run », cliquez sur le bouton : `common/icon/buttons/external_editing.p`

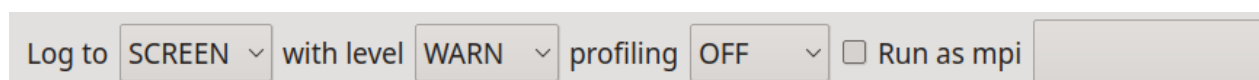


Figure 23: Accès à la fonction de modification du script R

- Une fenêtre s'ouvre contenant le script R. Modifiez le script. Sur la version actuelle, l'outil d'édition est très basique (pas de recherche, de coloration syntaxique).
- Sauvegardez et exécutez

ATTENTION, vos modifications dans le scrip seront **ECRASEES** à chaque fois que vous ferez un « **Prepar Run** ». Donc sauvegardez vos modifications dans un fichier externe à la plateforme.

3.1.5 Exécution en parallèle

Dans certains cas, il sera possible d'activer l'option « run in parallel » . Si vous cochez cette option, le traitement sera parallélisé. Le nombre de processeurs utilisé sera déterminé automatiquement par Vsoil en fonction du nombre de processeurs disponibles et de leur taux d'occupation.

Traitement/Méthodes	Parallélisation ?
Multi simulation	
Batch	oui
Analyse de sensibilité	
Fast99	oui
Morris	oui
Sobol	oui
Estimation de paramètres	
Simple simulation	
Levenberg Marquard	non
DREAM	non
DREAMzs	non
SCE-UA Shuffle Complex Evolution	non
Multi simulation	
Levenberg Marquard	oui
DREAM	oui
DREAMzs	oui
SCE-UA Shuffle Complex Evolution	non

3.1.6 Déportation des calculs sur des serveurs

Comme précisé précédemment, les traitements statistiques et la multi-simulation ne pourront s'exécuter que sur des environnements linux. Une fois votre traitement configuré (cf documentation vsoil-player, chapitre Running a model ou model usage), vous pouvez déporter votre calcul. Cette option est aussi utile pour des traitements statistiques dont les temps d'exécution peuvent être conséquents.

3.1.7 Synthèse du paramétrage du traitement

Après le « Prepar Run », dans l'onglet « Run » partie « Statistical result », une synthèse du paramétrage est disponible dans l'onglet « General ».

Cette synthèse regroupe les informations suivantes :

- la période de simulation,

- le traitement effectué (si statistique alors nom de la méthode, si multi-simulation alors le terme batch est précisé),
- les paramètres sélectionnés avec les informations concernant leur paramétrage. Ces informations sont dépendantes du traitement sélectionné.
- les sorties sélectionnées,
- le nombre de points de la grille et leur valeur,
- le nombre de temps sélectionné et leur valeur.

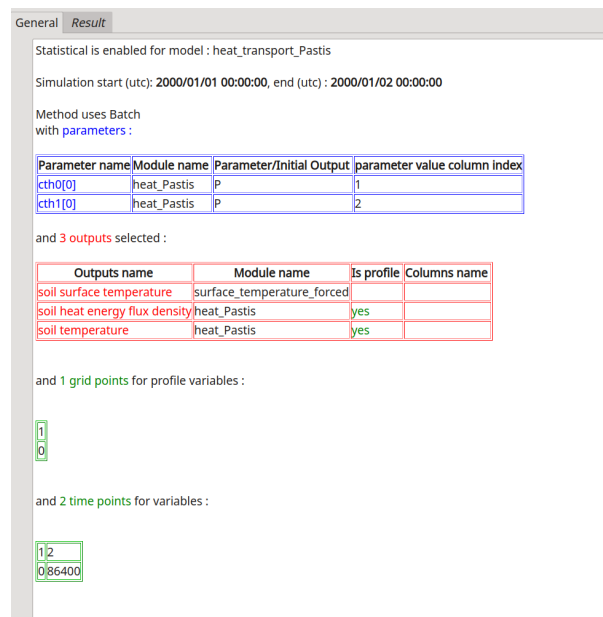
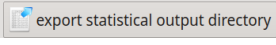


Figure 24: Synthèse du paramétrage du traitement statistique / multi-simulation

3.1.8 Exportation du répertoire de sorties des statistiques ou de multi-simulation

Après l'exécution d'une méthode d'analyse de sensibilité, d'estimation de paramètres ou d'une multi-simulation, vous avez la possibilité d'exporter vos résultats.

1. Export des résultats

- Dans l'onglet « Run », cliquez sur le bouton , positionné en bas à droite sous l'espace Statistical result.
- Sélectionnez le répertoire où les résultats seront sauvegardés.
Attention si le répertoire sélectionné contient des données (avant l'exportation) elles seront DETRUITES si vous confirmez dans la boîte d'alerte suivante :

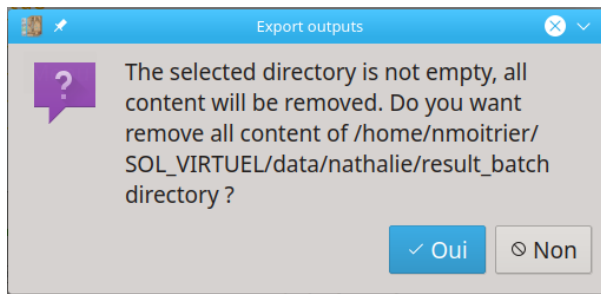


Figure 25: Message d'alerte pour l'exportation des résultats

- Confirmez, un message contenant le nom du répertoire est alors affiché :

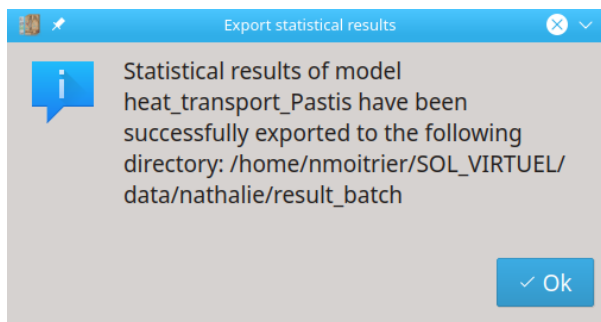


Figure 26: Message d'information de l'exportation des résultats

Dans le cas où vous ne confirmeriez pas l'exportation, vous revenez sur l'onglet "Run".

2. Contenu de l'exportation

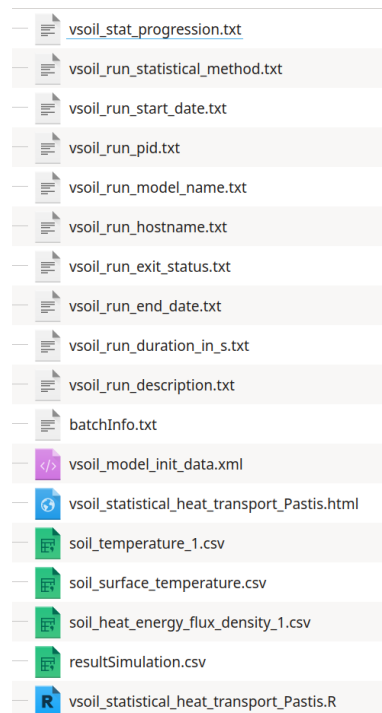


Figure 27: Liste des fichiers résultats issue d'un export après un traitement statistique ou de multi-simulation

- le script R (extension R) permettant de réaliser le traitement (statistique ou multi simulation) et de communiquer avec le modèle VSOIL,
- le fichier de configuration de la simulation (extension xml),
- les fichiers résultats (extension csv ,png ,pdf, RData) issus du traitement, par sorties et une synthèse (result_simulation),
- synthèse du traitement lancé 3.1.7 (informations des paramètres d'entrées, sorties..., extension HTML),
- différentes informations concernant la simulation (extension txt).

4 Exécution multiple d'un modèle

4.1 Présentation générale

Une option de la plateforme permet d'exécuter un modèle en changeant la valeur de certains paramètres et de stocker des variables de sortie pour faire des traitements qui seront à la charge de l'utilisateur. Les combinaisons de valeurs des paramètres pour lesquelles on veut faire tourner le modèle, sont fournies dans un fichier. Le modèle est alors exécuté autant de fois que combinaisons de valeurs des paramètres analysés sont données.

4.2 Fonctionnement global de l'exécution multiple

- Prérequis
 - Modèle vsoil fonctionnel.
 - Fichier de configuration (paramètres et situations initiales) des modules qui composent le modèle.
 - Fichier contenant les valeurs de paramètres qui varient. La structure de ce fichier doit respecter un formalisme défini par la plateforme (cf IFormat et structure du fichier des paramètres).
- Fonctionnement global de la multi simulation
 1. Configuration du modèle en mode multi-simulation. L'option « Batch » doit être sélectionnée.
 2. Exécution de la multi-simulation.
 3. Exportation des résultats.

Dans l'onglet « Initialization », sélectionnez dans la liste de choix de Statistical analysis la méthode : « Batch ».

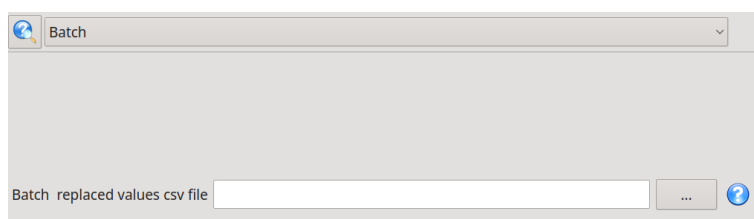


Figure 28: Paramètres liés à l'option Batch

Vous devez paramétrer les éléments suivants :

1. « Batch replaced values csv file » : sélectionnez l'emplacement et le nom du fichier contenant les valeurs des paramètres qui changent.
Le format et la structure du fichier doivent respecter les conditions suivantes:
 - Le fichier doit être au format csv avec comme séparateur des valeurs le « ; ».

- Chaque paramètre -dont la valeur va varier- doit correspondre à une colonne dans le fichier. Le nombre de colonnes dans le fichier doit être au minimum égal aux nombres de paramètres que vous sélectionnez dans l'interface graphique pour la méthode « Batch ».
- La première ligne du fichier ne sera pas traitée car elle correspond aux libellés des colonnes.
- Le modèle sera exécuté autant de fois qu'il y a de lignes dans le fichier csv.
- Le séparateur décimal doit être « . ».

batch1.csv

```

1 | fth0;cth1
2 | 0.01;0.1
3 | 0.2;0.2
4 | 0.3;0.3
5 | 0.45;0.4
6 | 0.089;0.5
7 | 0.5;0.5
8 | 0.001;0.01
9 | 0.002;0.02
10 | 0.003;0.03
11 | 0.004;0.04
12 | 0.005;0.05
13 | 0.006;0.06
14 | 0.007;0.07
15 | 0.008;0.08
16 | 0.009;0.09
17 | 0.010;0.10
18 | 0.011;0.11
19 | 0.012;0.12
20 | 0.013;0.13
21 | 0.014;0.14
22 | 0.015;0.15
23 | 0.016;0.16
24 | 0.017;0.17
25 | 0.018;0.18
26 | 0.019;0.19
27 | 0.020;0.20
28 | 0.021;0.21
29 | 0.022;0.22
30 | 0.023;0.23
31 | 0.024;0.24
32 | 0.025;0.25
33 | 0.026;0.26

```

Figure 29: Exemple fichier contenant les valeurs de paramètres pour la méthode "Batch"

2. Faire le lien entre les paramètres sélectionnés et les colonnes dans le fichier

Vous allez sélectionner les paramètres dont la valeur va changer (cf Sélection des paramètres d'entrées1):

Une boîte de dialogue contenant les informations liées au paramètre apparaît. Dans le cas de la multi-simulation (méthode "Batch"), précisez quelle colonne du fichier csv correspond au paramètre sélectionné :

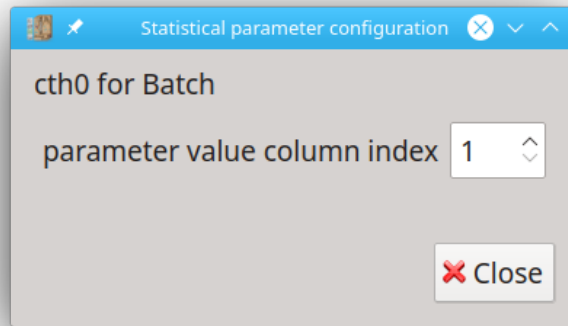


Figure 30: Saisie de l'index de la colonne du fichier correspondant au paramètre cth0

3. Sélectionnez la/les sortie(s) souhaité(es) (cf Sélection des variables de sorties2).
4. Exécutez « Prépare run »
Si aucune erreur n'est détectée, vous pouvez passer à l'onglet « Run » pour lancer l'exécution.
5. Dans l'onglet « Run » cliquez sur « Run stat ».
Si les simulations se sont bien exécutées, dans la partie « Statistical result », l'onglet « Result » permet de sauvegarder les sorties des simulations dans un répertoire utilisateur. Pour cela cliquer sur le bouton :

Save outputs result of multi-simulation

Figure 31: Sauvegarde des fichiers de sorties dans un dossier

Les sorties seront alors sauvées dans le répertoire de votre choix. Le répertoire contiendra :

- le fichier de simulation (extension xml),
- le fichier contenant les valeurs des paramètres (« vsoil_model_init_batch_values.csv »),
- les sorties :
 - (a) 1 fichier par variable de sortie. Dans le cas d'une variable de type "profile", il y a un fichier pour chaque profondeur sélectionnée. Le numéro associé au nom de la variable indique le noeud du maillage où est calculé la variable.
 - les x premières colonnes représentant les valeurs des paramètres ; x étant le nombre de paramètres sélectionnés
 - les z colonnes suivantes pour les valeurs de la sortie par temps ; z étant le nombre de temps sélectionnés.
 - (b) 1 fichier de synthèse (resultSimulation.csv), regroupant l'ensemble des valeurs de paramètres/sorties pour les temps calculés.
 - première colonne, temps calculé en seconde,

- x colonnes représentant les valeurs des paramètres ; x étant le nombre de paramètres sélectionnés.
- y colonnes pour les valeurs des sorties ; y étant le nombre de sorties sélectionnées.

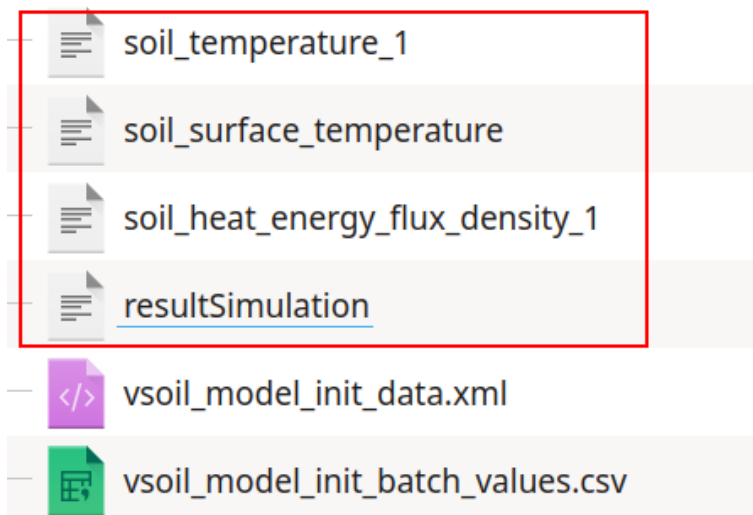


Figure 32: Contenu du répertoire de sauvegarde, les fichiers contenant les valeurs calculées des sorties sont encadrés en rouge.

Cas d'erreurs détectés

Lorsque vous cliquez sur « prepare run », des erreurs peuvent être détectées. Dans ce cas là, la multi-simulation ne pourra pas être exécutée. Le système vérifie que les paramètres ont des index de colonnes différents dans le fichier. Le système vérifie aussi que chaque index est inférieur ou égal au nombre de colonnes du fichier.

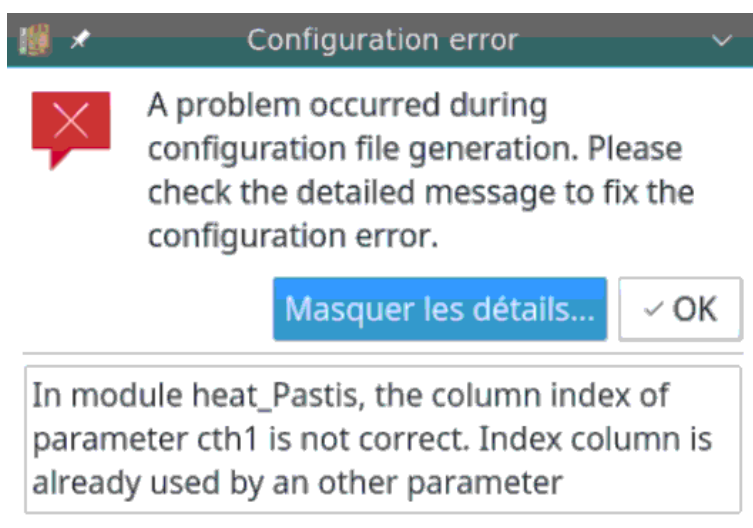


Figure 33: Erreur d'index : chaque paramètre doit avoir un index de colonne du fichier attribué et distinct

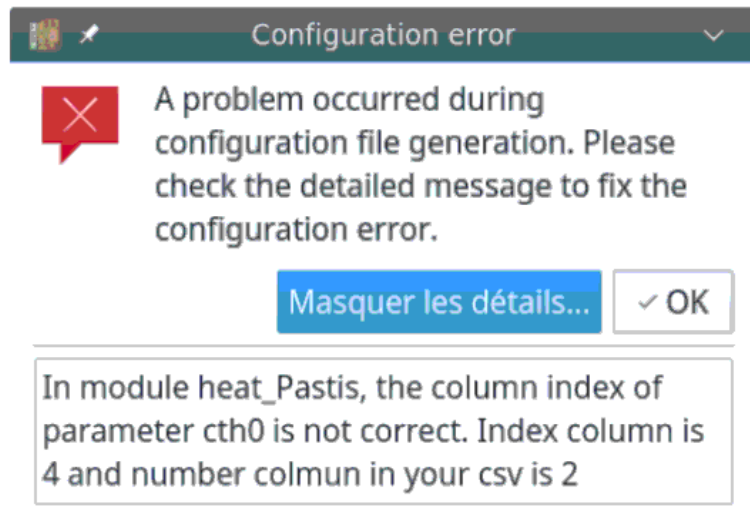


Figure 34: Erreur d'index : l'index de la colonne doit exister dans le fichier.

5 Analyse de sensibilité

5.1 Présentation générale

Trois méthodes ont été intégrées à la plateforme, issues du paquet R sensitivity : Fast99, Morris et Sobol.

5.2 Fonctionnement global de l'analyse de sensibilité dans la plateforme

- Prérequis
 - Modèle vsoil fonctionnel.
 - Fichier de configuration des modules.
- Fonctionnement global de l'analyse de sensibilité
 1. Configuration modèle en mode analyse de sensibilité.
 2. Exécution de l'analyse.
 3. Visualisation des résultats.

5.2.1 Paramétrage de l'analyse de sensibilité

Chargez votre modèle et chargez (ou saisissez) une configuration (onglet Initialization). Vous pouvez utiliser un fichier de configuration du modèle correspondant à une simulation habituelle. Il suffit que les paramètres et situations initiales correspondent à la situation que vous voulez traiter. La méthode d'analyse de sensibilité changera les paramètres dont on teste l'influence.

Dans l'onglet Global, sélectionnez la méthode d'analyse de sensibilité souhaitée :

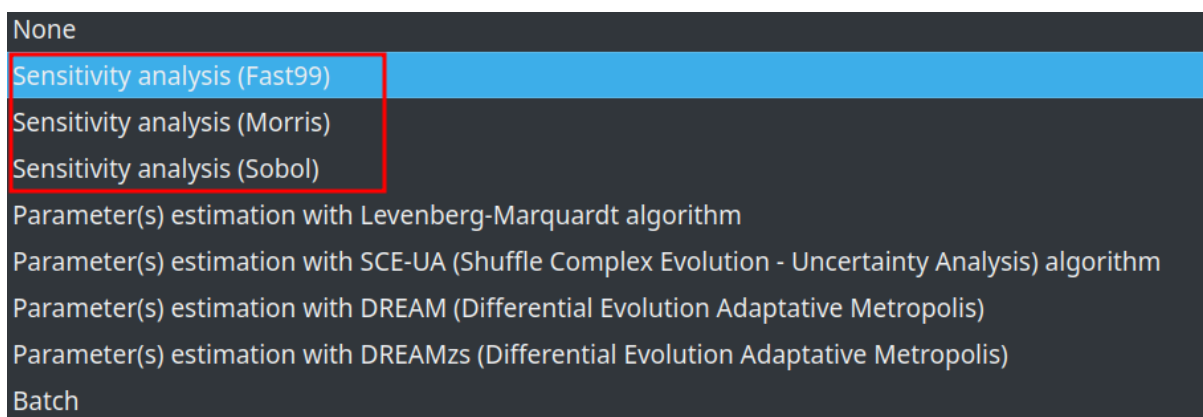


Figure 35: Les méthodes d'analyse de sensibilité disponibles sont encadrées de rouge.

L'implémentation informatique des méthodes Fast99, Morris et Sobol est issue du paquet R 'sensitivity'. Saisissez les paramètres généraux liés à la méthode sélectionnée. Pour chaque méthode, vous trouverez dans ce qui suit, le lien entre le paramètre proposé dans l'interface graphique de Vsoil et sa correspondance (son nom) dans le paquet R. Cette correspondance est accessible via l'interface : si vous positionnez le curseur de votre souris

sur le nom du paramètre un tooltip apparait avec la description et la correspondance dans le paquet R.

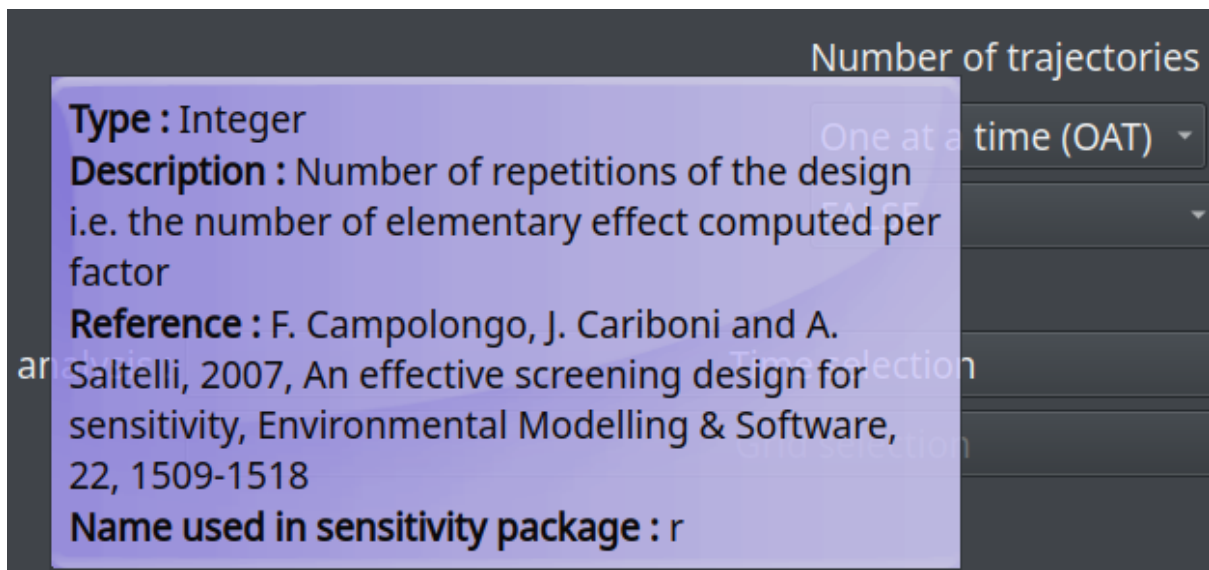
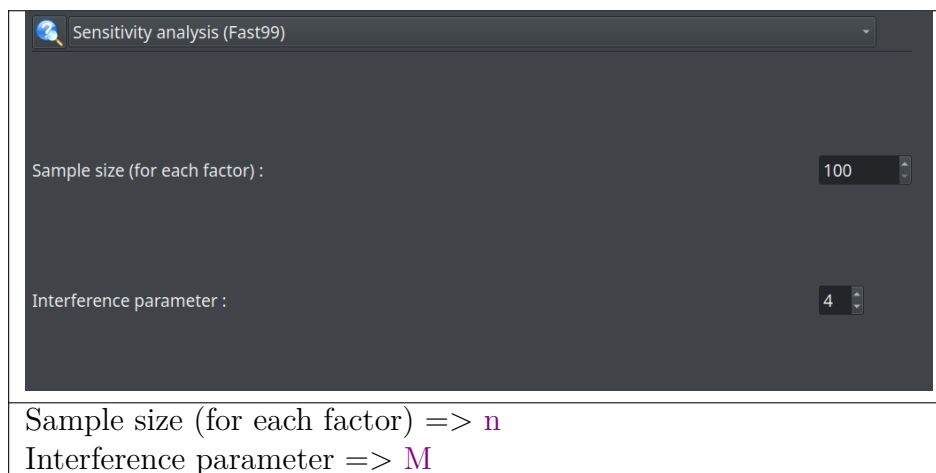
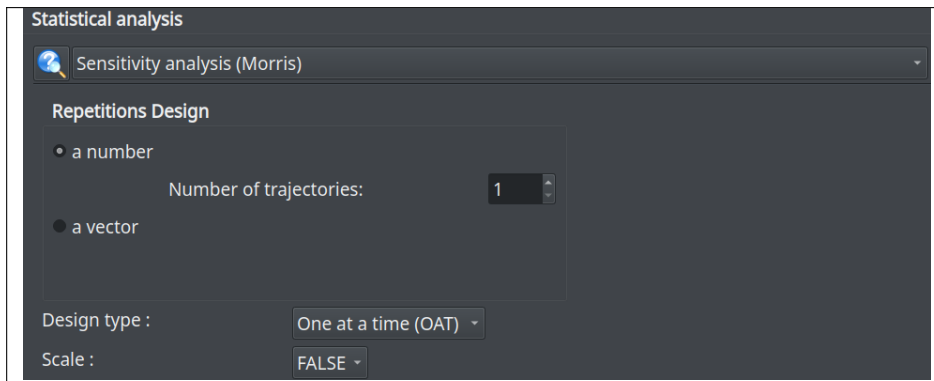


Figure 36: Aide contextuelle sur un paramètre de la méthode d'analyse de sensibilité fast99.

- Fast99 (Extended Fourier Amplitude Sensitivity Test)



- Morris




Repetition design : soit un nombre (valeur par défaut), soit un vecteur
 Si un nombre : Number of trajectories -> **r**
 Si un vecteur : The wanted design size -> **r1**, the size of the population-> **r2**
 Design type => **design**
 if simplex selected, **scale.factor**
 Scale => **scale**

- Sobol



Pour sobol, vous pouvez sélectionner la méthode souhaitée (sobol-jansen ou sobolsalt)
 Sample size (for each factor) => **n**
 The number of bootstrap replicates => **nboot**
 The confidence level for bootstrap confidence intervals: => **conf**
 Saltelli's scheme used (seulement pour sobolSalt) :=> **scheme**

5.2.2 Sélection des paramètres d'entrées et configuration

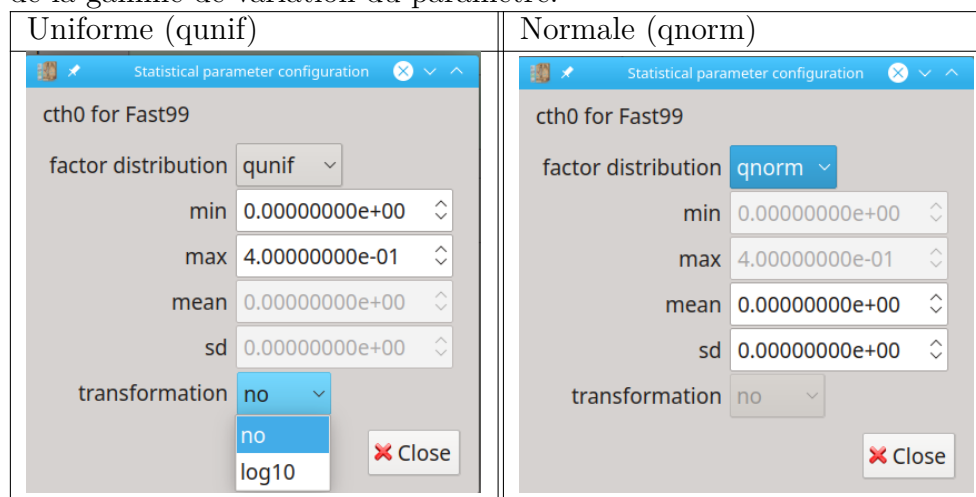
- Sélectionnez les paramètres qui vont être testés dans l'analyse. (1).
- Pour chaque paramètre choisi, utilisez l'icône  pour accéder aux paramètres de la méthode d'analyse de sensibilité.

Les noms des paramètres de l'analyse de sensibilité pour lesquels vous devez saisir des valeurs ou faire des choix, correspondent aux noms des arguments utilisés dans les méthodes du paquet R.

Référez-vous à la documentation de Fast99, Morris et Sobol du package R sensitivity pour avoir la signification précise de ces paramètres; même si une description brève est donnée ci-dessous.

1. FAST99

Vous pouvez sélectionner soit une distribution uniforme soit normale qui seront utilisées pour générer le plan d'expérience. En fonction de votre choix, vous devrez fournir les valeurs minimum et maximum du paramètre pour la distribution uniforme ou alors les valeurs de la moyenne et de l'écart-type pour la distribution normale. Vous avez aussi accès à une transformation éventuelle de la gamme de variation du paramètre.



2. MORRIS

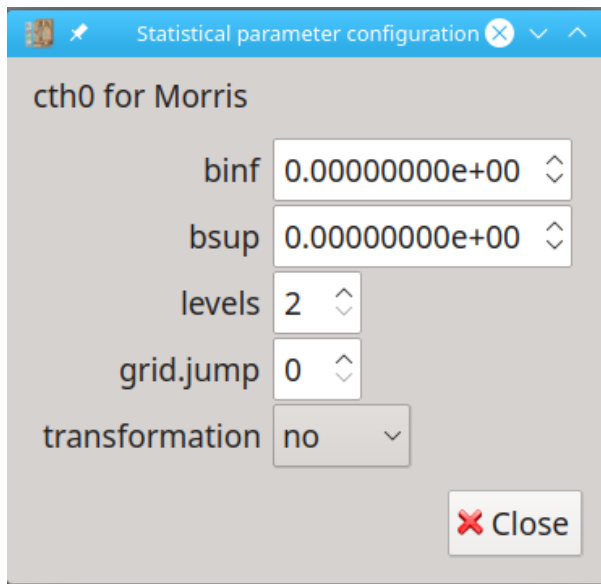


Figure 37: Paramétrage des entrées pour Morris

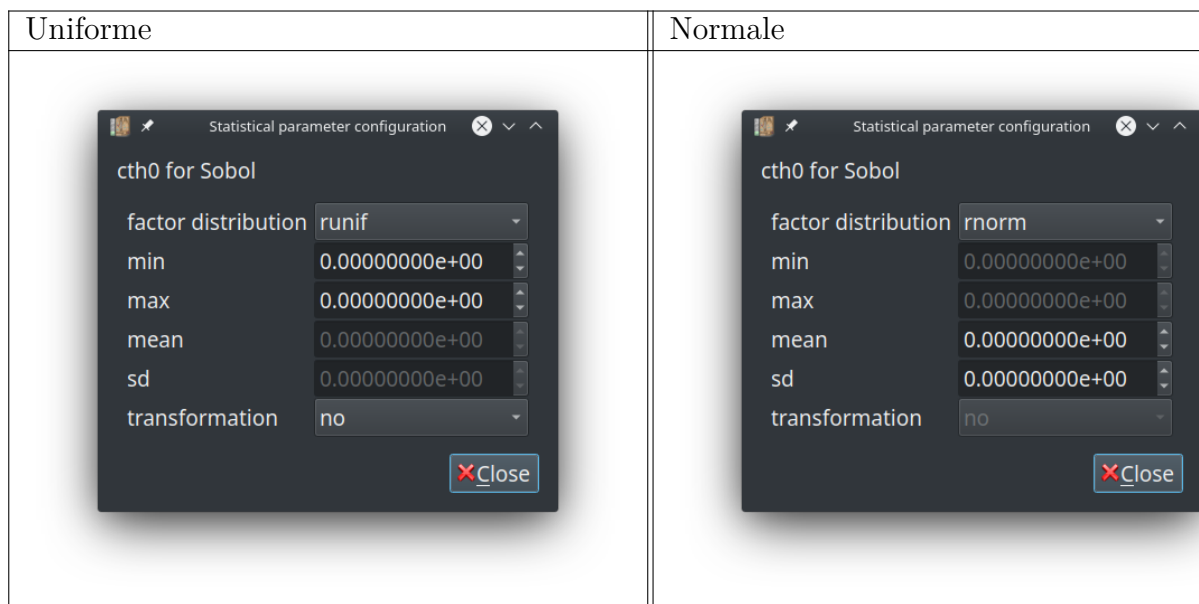
binf et bsup correspondent aux valeurs minimales et maximale du facteur/paramètre. Cet intervalle de variation va être découpé en niveaux : Levels. Les différentes trajectoires seront échantillonnées avec un saut aléatoire grid.jump. "Il est conseillé de choisir grid.jump de l'ordre de levels/2' car si le saut est trop petit, on n'a des informations que sur des variations locales."

3. SOBOLE

Pour sobol, en fonction du type du paramètre sélectionné, les informations demandées seront différentes.

Pour les paramètres de type réel vous pourrez comme pour FAST99, sélectionner la loi de distribution à appliquer.

– Paramètre réel



- Liste de choix (int-choice)

Dans le cas des paramètres de type liste de choix, il n’y a pas d’informations supplémentaires à saisir. Les valeurs seront prises (au hasard) dans la liste affichée dans la boîte de dialogue qui correspond aux index des différents choix possibles.

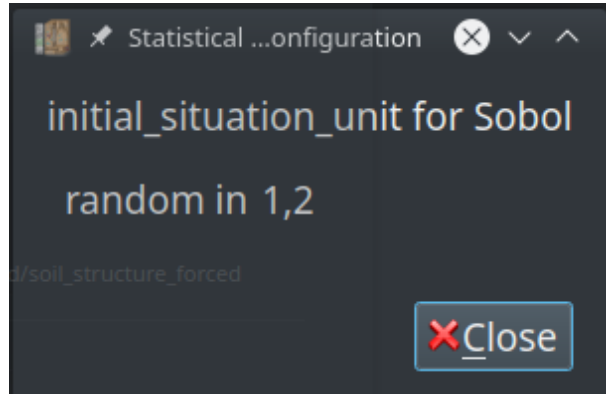


Figure 38: Paramétrage des entrées de type "int-choice" pour Sobol

Dans le cas des paramètres de type réel, vous pouvez appliquer une transformation à votre paramètre : actuellement la transformation \log_{10} est proposée pour les trois méthodes. La transformation s’appliquera sur les paramètres pour le traitement statistique (Pas clair ce que tu veux dire. Je crois que ce qu’il faut dire c’est que pour des paramètres qui varient dans une large gamme de valeur, il peut être judicieux de faire une transformation \log sur les bornes avant de générer la distribution des paramètres, puis d’appliquer la transformation inverse avant de lancer les calculs. La valeur du paramètre "non" transformée sera utilisée par le modèle. [Est ce qu’on explique pourquoi et quand faire ? une transformation]. (Je pense qu’on peut mieux expliquer comment ça marche et donner un exemple. Tu dois avoir un exemple quelque part.)

5.2.3 Sélection des sorties

Sélectionnez les sorties souhaitées 15.

5.2.4 Exécution de l’analyse de sensibilité

Une fois le paramétrage et la sélection des sorties réalisés,

- Cliquez sur « Prepare run ».
- Dans l’onglet « Initialization. », cliquez sur « Run stat »

5.2.5 Résultats de l’analyse de sensibilité

Les résultats de l’analyse de sensibilité seront disponible dans de nouveaux onglets dans la partie "Statistical result" de l’onglet "Run". Comme déjà précisé, en mode statistique, l’accès à l’onglet "Plot" n’est pas possible. Les résultat sont issus des traitements statistiques, ils seront accessibles par sortie (un onglet par sortie, après l’onglet "General" dans la partie "Statistical result" :

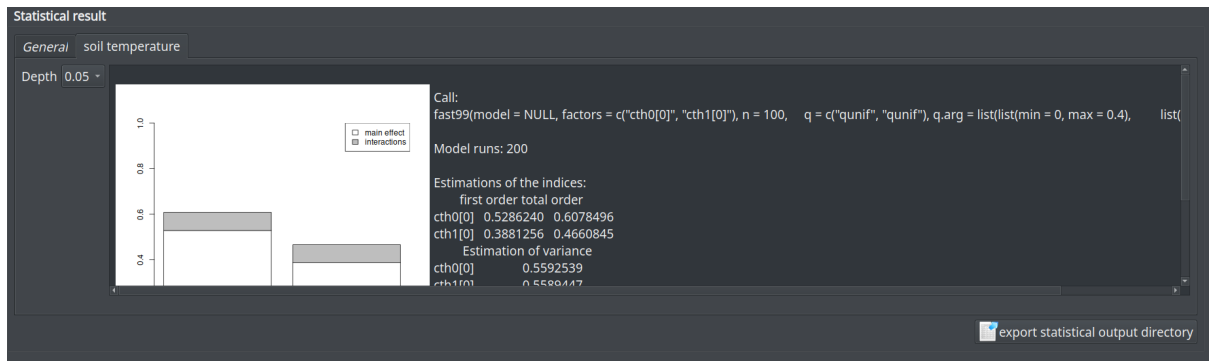


Figure 39: Accès aux résultats de l'analyse de sensibilité

Si la sortie est "profilée" alors une liste des profondeurs disponibles est proposée. Sélectionnez la sortie souhaitée, et si elle est de type "profile", la profondeur désirée. Les graphiques sont issus du paquet R de la méthode utilisée.

6 Estimations de paramètres

6.1 Présentation générale

La plateforme donne accès à différentes méthodes d'estimations de paramètres.

L'estimation d'un jeu de paramètres est effectuée soit sur UNE configurations du modèle, soit sur PLUSIEURS configurations. Dans ce dernier cas les paramètres estimés peuvent l'être sur l'ensemble des configurations ou certains être spécifiques à chaque configuration (cas des paramètres fixes et variables).

6.2 Fonctionnement global de l'estimation de paramètres dans la plateforme

- Prérequis
 - Modèle vsoil fonctionnel.
 - Fichier(s) de configuration des modules.
 - Fichiers d'observations au format compatible à VSoil, pour les variables sur lesquelles se fait l'estimation.
- Fonctionnement global de l'estimation
 1. Configuration du modèle en mode estimation.
 2. Exécution de l'estimation.
 3. Visualisation des résultats.
 4. Selon la méthode d'estimation (possible uniquement pour LM en simulation simple et multi configuration c'est à dire une seule configuration): Pas clair ce que tu veux dire.
 - (a) Sauvegarde de la ou des configuration(s) avec les paramètres estimés
 - (b) Configuration du modèle avec les paramètres estimés

- (c) Exécution
- (d) Visualisation sorties simulées / observées

Actuellement la plateforme intègre 3 méthodes (dont 1 avec 2 implémentations différentes).

6.3 Les méthodes d'estimation

- Une méthode basée sur l'algorithme de Levenberg-Marquardt (LM), qui est une amélioration de la méthode classique de Gauss-Newton dans la minimisation de la somme des carrés pour des problèmes non-linéaires.
- Une méthode Differential Evolution Adaptive Metropolis (DREAM), cette méthode est proposée par deux implémentations R différentes.
- Shuffle Complex Evolution (SCE) - University of Arizona (Duan et al., 1992) une méthode d'optimisation globale qui "combine les points forts de la procédure simple de Nelder et Mead (1965) avec les concepts de recherche aléatoire contrôlée (Price, 1987), d'évolution compétitive (Holland, 1975)" avec le concept de brassage complexe, développé par Duan et al. (1992).

6.3.1 Configuration et paramétrage des méthodes d'estimation

1. Configuration

- LM

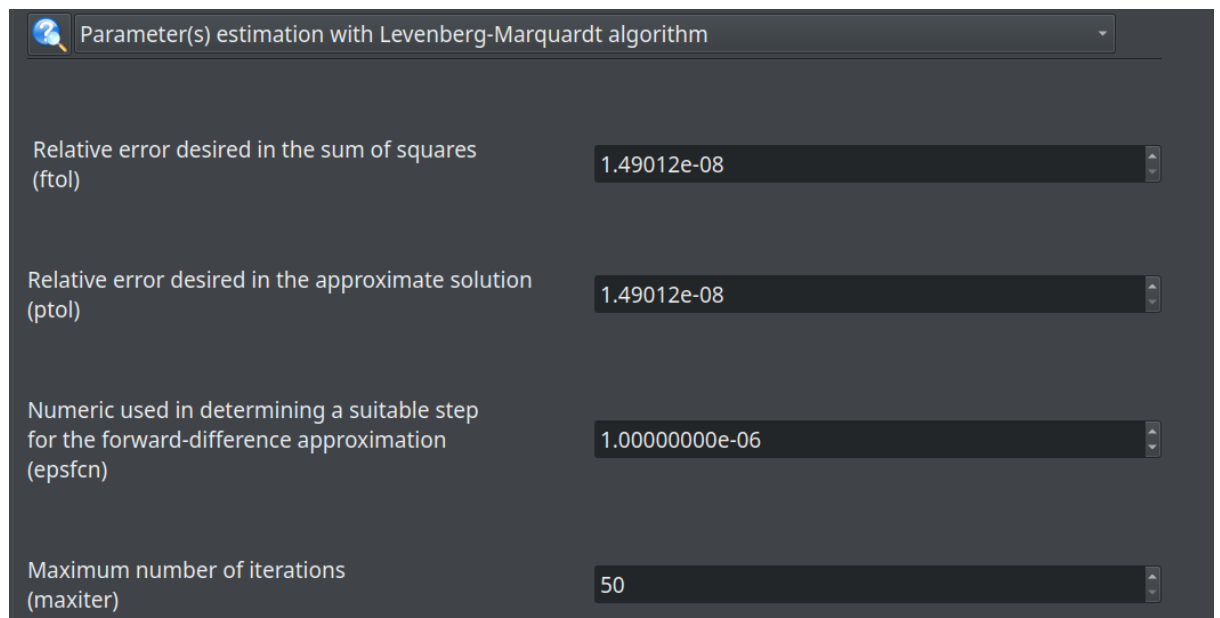


Figure 40: Configuration des paramètres généraux de la méthode Levenberg-Marquardt (LM)

Il faudrait dire que les valeurs indiquées sont des valeurs par défaut. Une description précises des paramètres de la méthode est disponible dans R en cliquant sur le bouton

- SCE-UA (Remarque : ça a changé ça !!!! Changer ce que tu dis ci-dessous. La figure n'est plus bonne. On a des paramètres accessibles par l'interface et des paramètres accessibles par le update control settings. Pourquoi cette différence. On peut pas tout mettre pareil. Ce serait bon d'avoir une petite définition popup des paramètres, même si il faut des fois aller dans la doc R. il y a deux paramètres avec le même nom.)// ***** Actuellement, seules les valeurs des paramètres (maxn, kstop, pcento, ngs, npg, nps, nspl) peuvent être modifiées. Par défaut les paramètres (npg, nps, nspl) sont à NA, dans ce cas les valeurs des ces paramètres seront recalculées dans la fonction sceua. il n'est pas possible de modifier via une interface graphique (évoluée) les paramètres généraux pour cette méthode, ce sont leurs valeurs par défaut qui seront utilisées. Mais si vous avez besoin d'en changer les valeurs par défaut, une interface simplifiée a été mise en place. Pour y accéder, sélectionnez "Update control settings" : une liste des paramètres utilisés par la méthode est alors affichée respectant un format précis (nom_parametre = valeur et chaque paramètres est séparé par le séparateur ",". Vous pouvez modifier la valeur des différents paramètres. Mais attention aucune vérification des valeurs saisies n'est effectuée et ces informations seront "passées" directement à la méthode (dans le script R). I

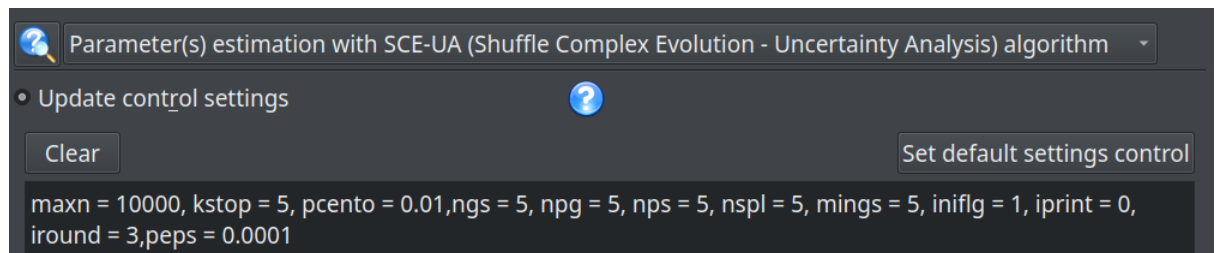


Figure 41: Configuration des paramètres généraux de la méthode Shuffle Complex Evolution - University of Arizona (SCE-UA)

- DREAM

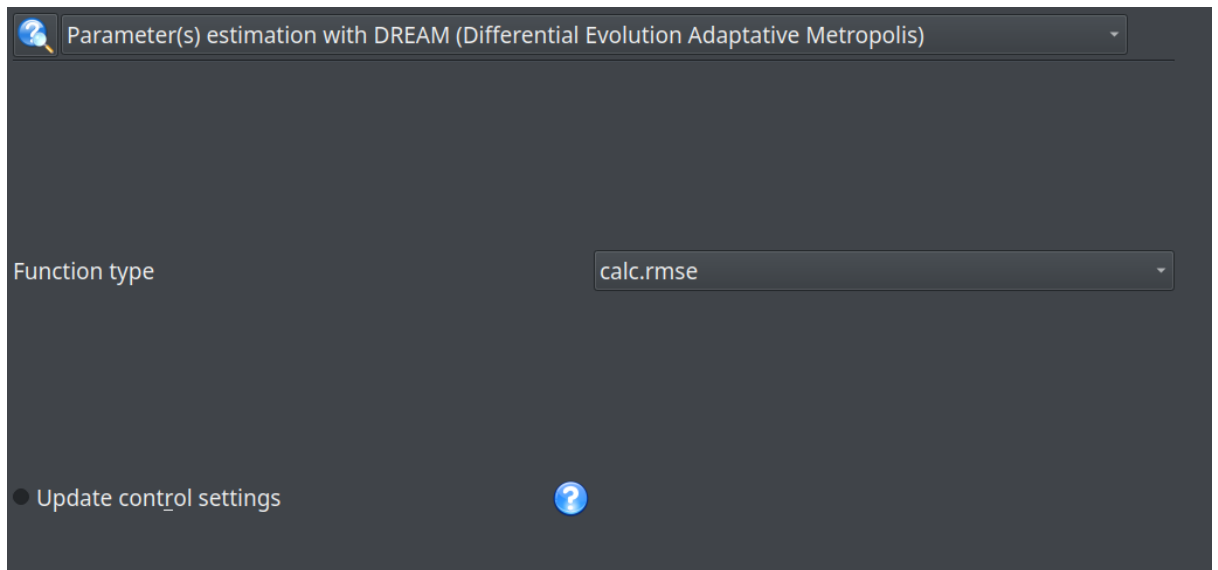


Figure 42: Configuration des paramètres généraux de la méthode

- DREAM-ZS

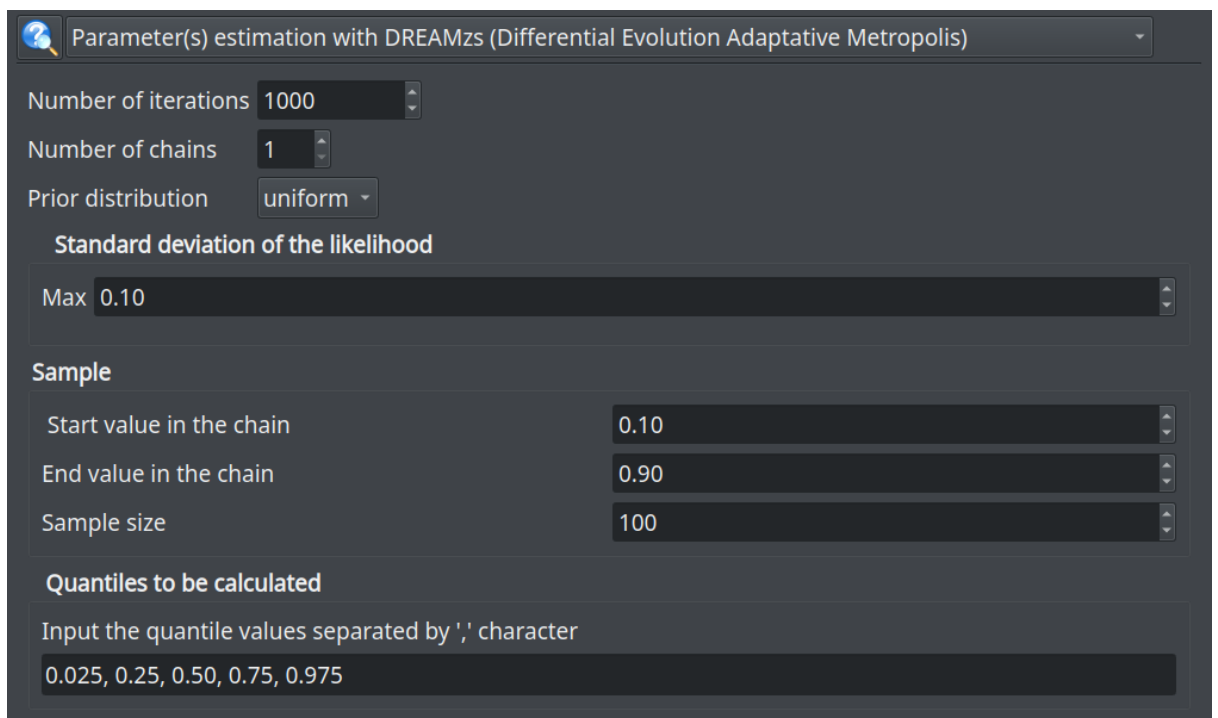


Figure 43: Configuration des paramètres généraux de la méthode

Il a été fait le choix de ne pas interfacer l'accès à certains paramètres généraux, car les valeurs par défaut proposées par les concepteurs des packages sont pertinentes. Toutefois il est possible de les modifier en sélectionnant "update control settings". Cette fonctionnalité est proposée pour les deux implémentations de DREAM. En activant cette option, vous pouvez accéder à la liste de l'ensemble des paramètres généraux de la méthode et d'en modifier leur valeur. (Remarque : ce n'est pas le cas pour la deuxième!!!!)

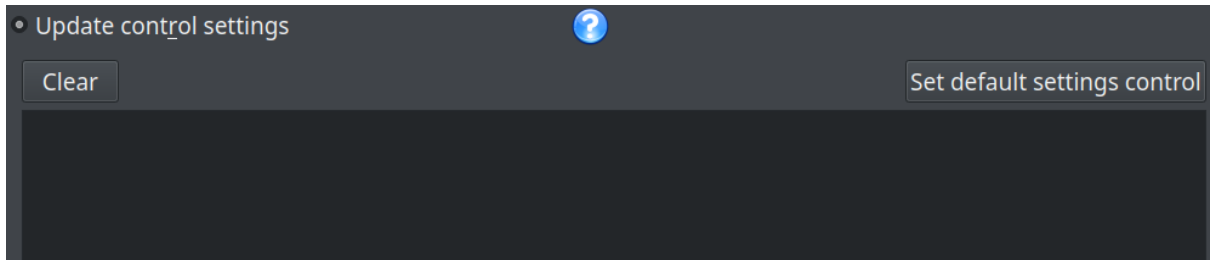


Figure 44: Configuration de certains paramètres généraux pour DREAM

Pour accéder à la liste cliquez sur "Set default settings control", la liste des paramètres s'affichent sous forme de liste. Vous pouvez alors modifier la valeur de paramètres souhaités mais attention de ne pas modifier la syntaxe (séparateur , =).

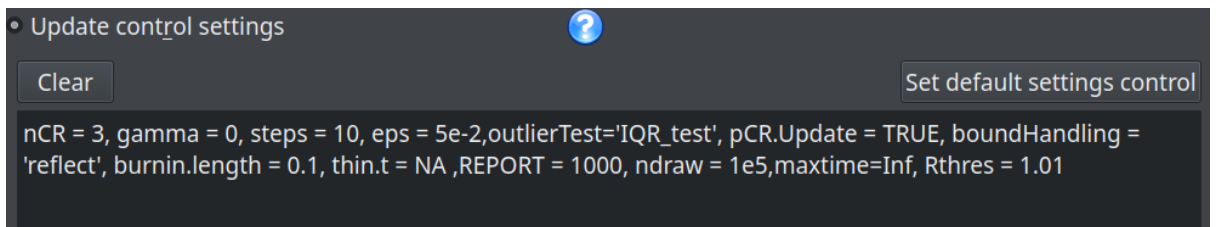
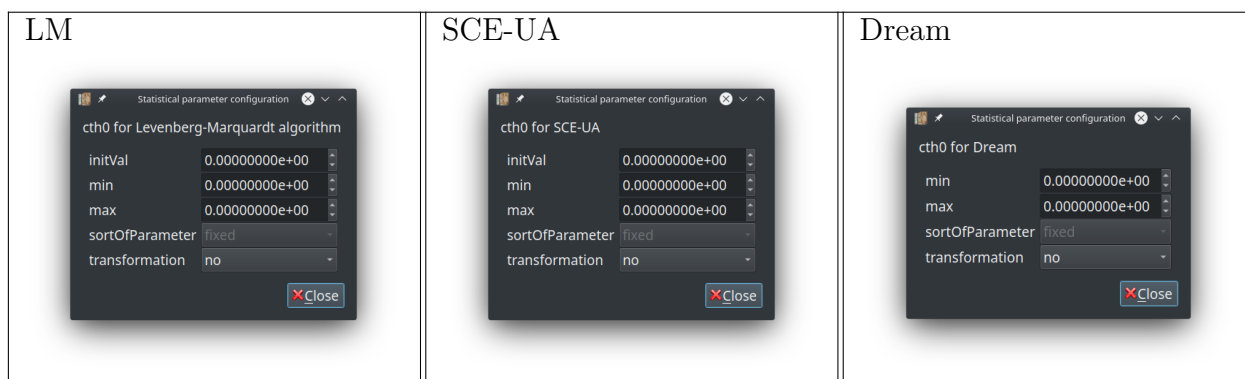


Figure 45: Configuration de certains paramètres généraux pour DREAM

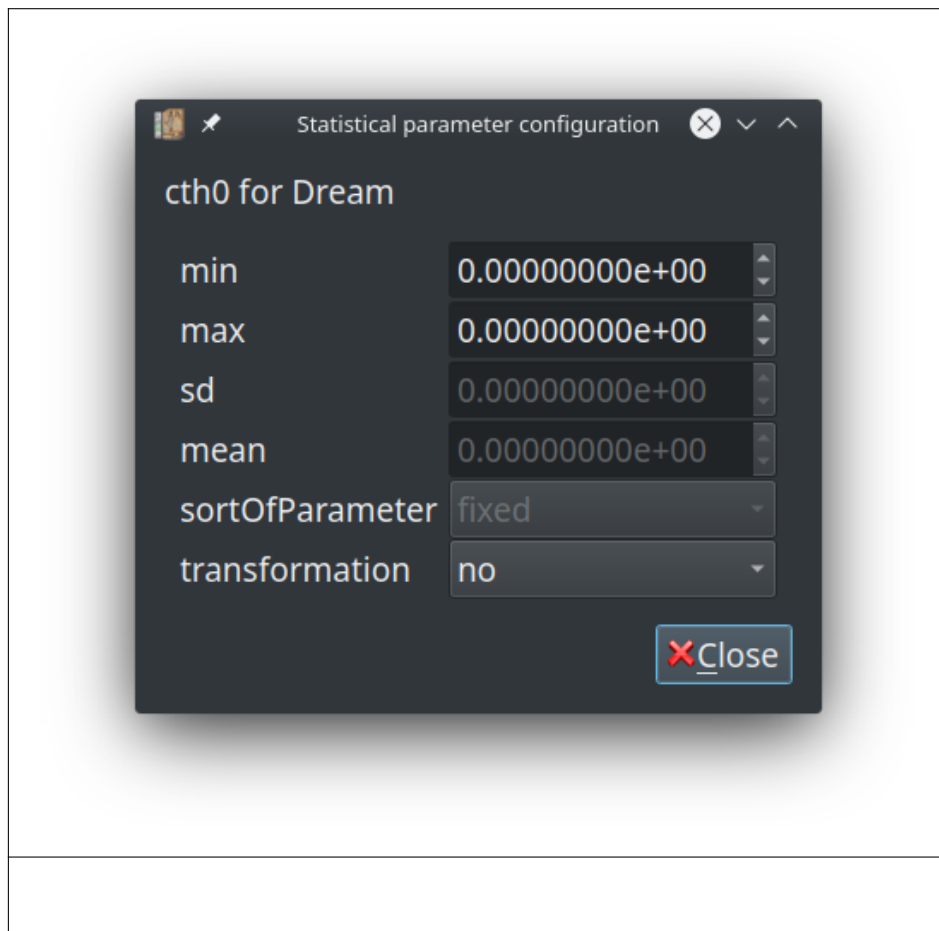
2. Sélectionner et configurer les variables d'entrées du modèle utilisé comme paramètres de la méthode d'estimation

Sélectionnez les variables du modèle et leur gamme de variation. Quelque soit la méthode d'estimation sélectionnée, vous devez sélectionner la valeur initiale de la variable et ses bornes de variation. La méthode d'estimation DreamZS, a une option supplémentaire qui permet de définir différemment la gamme de variation de la variable. En fonction de la distribution sélectionnée lors du paramétrage général de la méthode DreamZS (loi de distribution uniforme ou normale) des paramètres différents seront accessibles(cf.).

- test



- DREAMZS



(Remarque : là aussi la grosse figure ne me plait pas !!!)

6.4 Les données observées

L'estimation de paramètres se fait à partir de l'« ajustement » du modèle aux données observées pour les sorties sélectionnées.

Donc chaque sortie sélectionnée pour l'estimation de paramètres a un fichier contenant ses valeurs observées à des temps déterminés.

Format des fichiers contenant les données observées :

- Le fichier d'observation doit être au format csv. Les valeurs sont séparées par « ; ».
- Le format des temps et la position des valeurs observées sont paramétrables via la plateforme (cf. 2)
- La première ligne des fichiers d'observations contient, pour information, les en-têtes des colonnes :
 1. La première colonne des données observées, correspond automatiquement au temps et donc son nom d'en-tête est « time ». (Remarque : C'est obligatoire le "time" ??)
 2. Les colonnes suivantes :
 - Si la sortie est de type « profile » c'est à dire dépend de la profondeur, le nom de la colonne est une donnée numérique correspondant à la profondeur. (Remarque : profondeur obligatoire ???)
 - Sinon c'est une chaîne précisant le nom de la variable.
- Les lignes suivantes :
 - Colonne 1 : valeur numérique du temps au format défini dans la plateforme. Si le temps est en seconde alors cela doit être un entier.
 - Colonne [2,n] : valeur numérique pour la profondeur si la variable est de type "profile" ou de la variable scalaire.

IMPORTANT : les valeurs des temps et [profondeurs] des fichiers d'observations « forceront » le modèle. C'est à dire que le modèle calculera les sorties pour les profondeurs et les temps lus dans le fichier d'observations. (Remarque : Ce que tu veux dire c'est que le maillage est modifié pour tenir compte des profondeurs d'observations??? Si c'est ça il faut le préciser. Ce que tu dis juste après me laisse un peu perplexe.)

Si les variables de sorties ne sont pas de type "profile" alors les calculs se feront sur la base des profondeurs de la grille saisie lors de la phase d'initialisation.

L'unité utilisée pour les valeurs observées des variables doit être identique à l'unité utilisée dans la plateforme et donc doit être en SI.

(Remarque : Pour moi ce qui suit n'est pas clair. Variables de type "tagged"?? On ne voit pas ce qui change par rapport à ce qui précède. On voit cela ensemble si tu veux)

Cas des variables vectorisées : chaque variable doit avoir un fichier d'observation associé.

Si la variable est non profilée : une seule colonne après celle du temps. Si la variable est

profilée : autant de colonnes que de profondeurs souhaitées.

Remarques :

- La valeur peut être une valeur moyennée (cf. sélection des valeurs moyennées).
- Si la valeur est manquante, préciser NA.

Time	Nom variable ou valeur profondeur 1 en mètre (p1)	[valeur profondeur 2 en mètre (p2)]
Valeur temps au format sélectionné	Valeur de la variable à [p1] t1	Valeur de la variable à p2 t1

Pour associer les observations aux sorties cf. 6.5.1

6.5 Estimation

L'estimation de paramètres peut être réalisée soit :

- A partir d'un seul fichier de configuration c'est ce que l'on nommera estimation simple (cf. 6.5.2).
- A partir d'un ensemble de configurations c'est ce que l'on nomera estimation sur de la multi-configuration (cf. 6.5.3).

6.5.1 Sorties et observations

Une fois la méthode d'estimation et les paramètres à estimer configurés, les sorties sur lesquelles se fait l'estimation sélectionnées, et le type d'estimation choisi (simple ou multi-configuration) faite "prepare run". Une boîte de dialogue apparait pour configurer les variables de sorties. Le contenu de cette boîte s'adaptera en fonction du type d'estimation : estimation simple ou multi-configuration. Dans les deux cas, pour chaque variable utilisée pour l'estimation de paramètre, il faudra sélectionner :

- Quand plusieurs variables sont utilisées pour calculer l'écart entre le modèle et l'expérience, on peut être amené à donner un poids particulier à chaque variable dans le calcul de la fonction à minimiser. La colonne "Coefficient" permet d'affecter un facteur dans l'intervalle]0,1] à appliquer à cette sortie. Si on a une seule sortie alors ce coefficient sera égal à 1. Quand il y a plusieurs variables, la somme de tous les coefficients doit être égale à 1.
- Si les données observées ont été traitées via une transformation (logarithmique), alors vous devez sélectionner cette transformation afin qu'elle soit aussi réalisée sur les données calculées.

La colonne "Select file" permet de choisir le fichier d'observations correspondant à la variable. Toutefois, la sélection des fichiers d'observations diffère en fonction d'une simple ou d'une multi-configuration :

- Dans le cas d'une estimation sur une simulation (simple), un fichier d'observation sera associé à une variable de sortie.

- Dans le cas d'une multi-configuration, l'association des données observées sera réalisée avant le "prepare run" dans la partie configuration générale de la méthode d'estimation(cf 6.5.3).

Estimation simple

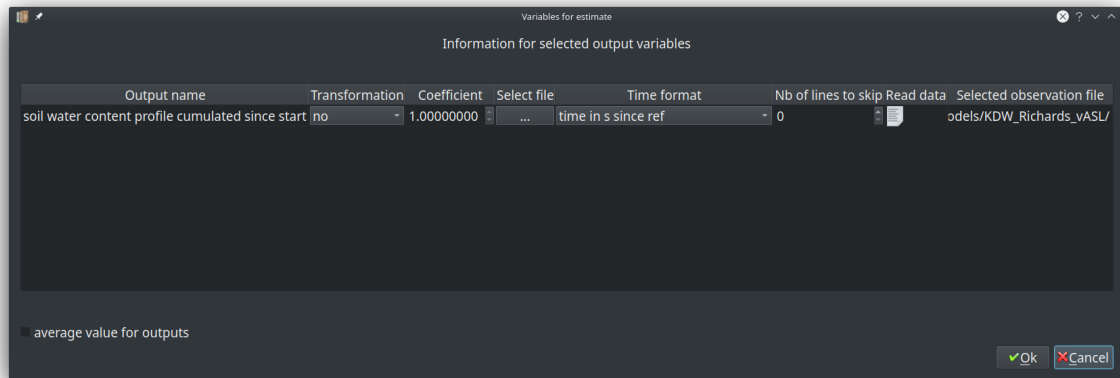


Figure 46: Configuration des sorties pour une estimation simple

Dans ce cas pour chaque variable utilisée pour faire l'ajustement, vous pouvez configurer les paramètres suivants pour la lecture des données observées :

1. Le chemin et le nom du fichier contenant les données observées pour la sortie (cf. structure du fichier contenant les données observées 6.4).
2. Le format du temps contenu dans la première colonne du fichier. Par défaut le temps est en seconde depuis le temps de référence. Mais vous pouvez sélectionner aussi le temps calculé à partir du temps de début de la simulation (tstart), ou le temps exprimé dans deux autres formats.
3. Le nombre de lignes situées en début de fichier et qui ne contiennent pas d'observations. Elles correspondent en général à l'entête des colonnes ou à des informations et commentaires sur le contenu du fichier.

Estimation avec des multi-configuration

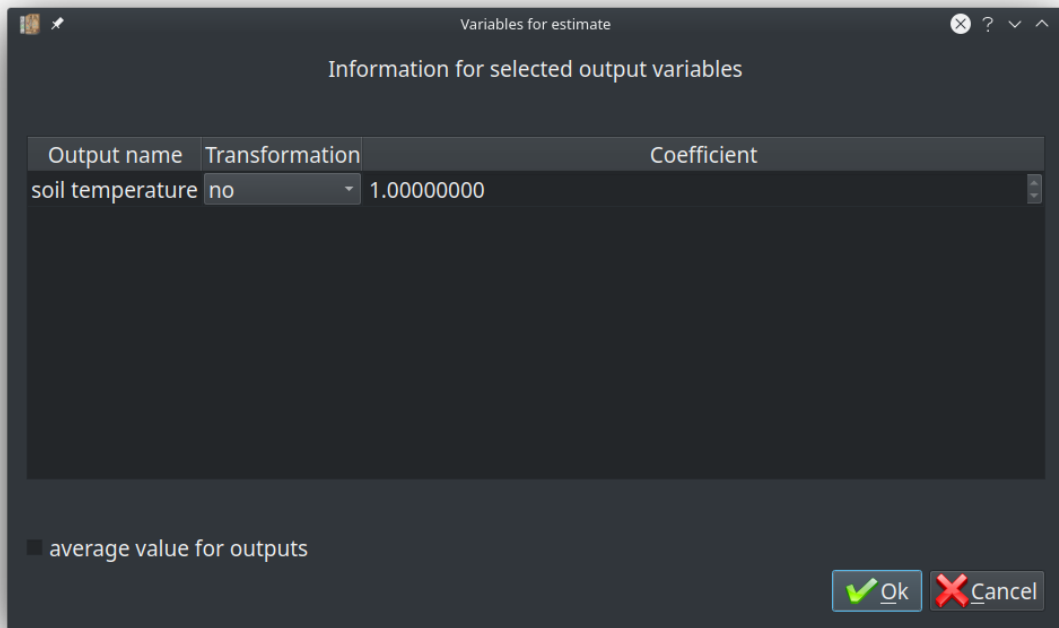


Figure 47: Configuration des sorties pour une configuration multiple

(Remarque : Là on ne sait pas ce qu'il faut faire. Il faudrait dire comment on fournit les multi config ???)

Données observées moyennées

Le fichier d'observation peut contenir des données moyennées. Les mesures effectuées ont subies un traitement. La plateforme prévoit les cas suivant dans le fichier d'observation. pour la valeur des temps : soit la valeur correspond au 1er temps de la série de mesures. soit la valeur correspond au dernier temps de la série de mesures. soit la moyenne des temps.

pour la valeur des mesures ! la valeur des mesures moyennées Les données observées utilisées pour les estimations peuvent être des valeurs moyennées sur une période de temps (Rem: C'est bien ça ???).

(Remarque : Pas clair. On ne sait plus faire la distinction entre la moyenne et la mesure?? Il faut être clair. J'ai repris, mais est-ce bien ça ??!!))

Le fichier d'observation doit avoir la structure définie à 6.4.

Pour ce cas particulier, les temps données dans le fichier d'observations, correspondent au temps de la dernière mesure effectuée avant le calcul de la moyenne en utilisant un nombre donné de mesures précédentes. Il faut donc préciser l'intervalle de temps en secondes entre deux mesures successives et le nombre de mesures utilisées pour faire la moyenne. Ces informations permettront d'exécuter le modèle pour calculer les variables aux temps exacts des mesures et les valeurs moyennes utilisées pour l'ajustement des paramètres.

La valeur des temps doit être supérieure à $t_{start} + (\text{nombre d'observations} * \text{intervalle de temps entre 2 observations})$. (Rem: Je ne comprend pas cette contrainte. A clarifier)

Exemple : On suppose que on a des mesures qui sont réalisées par paquet de quatre (m1, m2, m3, m4), avec un intervalle de temps de t secondes entre chaque mesure. Si

Tobs est le temps de la mesure m4, le fichier d'observation contiendra les temps Tobs et les moyennes $(m1+m2+m3+m4)/4$

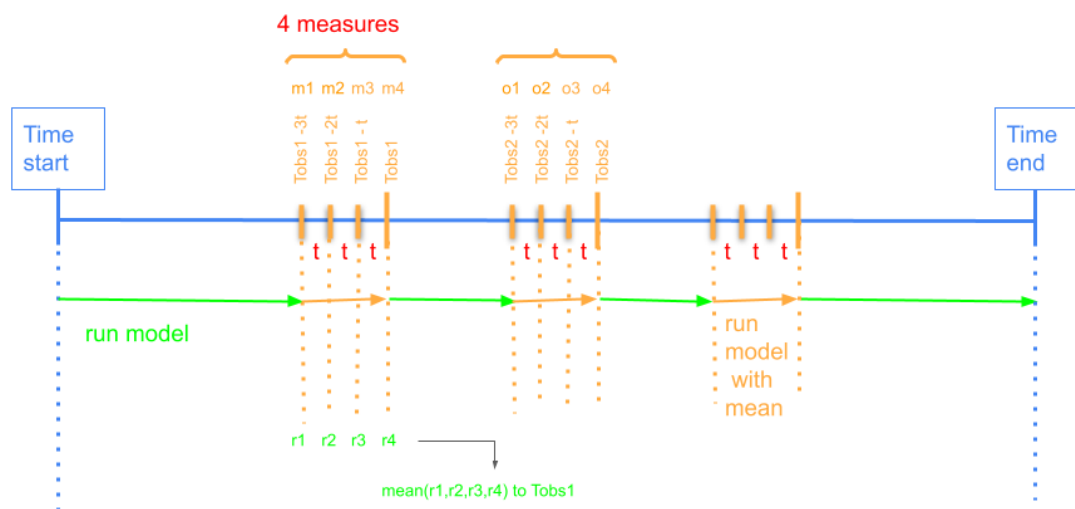


Figure 48: Exécution du modèle pour une estimation avec des mesures moyennées

6.5.2 Estimation simple

L'estimation des paramètres est réalisée à partir d'un seul fichier de configuration. Après avoir sélectionné la méthode d'estimation et l'avoir paramétrée (cf 4.1), sélectionnez « prepare run ».

6.5.3 Estimation sur de la multi-configuration

L'objectif est d'optimiser les paramètres sélectionnés sur un ensemble de configurations. Un ensemble de fichiers de configuration (fichier VSoil) et d'observations (format csv) vont être utilisés pour l'estimation. Un fichier de configuration général va permettre de configurer la méthode d'estimation et de la paramétrer. C'est ce fichier de configuration qui sera lu lors de l'exécution. (Remarque : On comprend pas bien qu'il y ai un fichier de conf et aussi plein d'autres. Il faut préciser comment ça marche. A quoi sert le fichier de conf général. Les fichiers de conf pour chaque manip qui indiquent les spécificités de la manip : temps, paramètres non ajustés, situation initiale, etc..)

L'estimation multi-configuration n'est disponible qu'avec les méthodes : Dream/-DREAMZs - LM.

La plateforme propose

1. soit d'optimiser un seul jeu de paramètres pour toutes les configurations,
2. soit d'avoir une situation « mixte », afin de prendre en compte le fait que certains paramètres sont liés à UNE configuration et d'autres paramètres sont liés à TOUTES les configurations. Cette deuxième option est actuellement disponible pour la méthode DREAM.

Sélectionnez la méthode d'estimation.
 Cochez "multi-config"

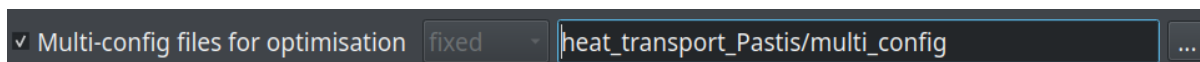


Figure 49: Sélection de l'estimation en multi-configuration

Vous allez saisir le chemin du répertoire contenant les fichiers de configurations et d'observations . Les fichiers de configuration et d'observations doivent être només de la façon suivante :

	Préfixe *	[Corps*]	Suffixe *	Extension *
Fichier configuration	Config_	Ce que vous voulez		.xml
Fichier d'observation	Obs_		Si code variable name Code variable name Si vectorisé Code variable name_nom de la colonne	.csv

Dans le répertoire le nombre de fichiers d'observations doit être égal au nombre de configurations multiplié par le nombre de variables observées.

La partie "corps" permet de distinguer les différentes configurations. Cette partie doit être identique pour les fichiers d'observations et de configurations. Pour le suffixe, c'est le nom de la variable ("code variable name") dans le code du module qui la calcule qui doit être utilisée. Pour obtenir le "code variable name" mettez le curseur de la souris sur le nom de la variable de sortie. Par exemple sur la variable de sortie "soil temperature", pour avoir son code variable, positionnez le curseur de la souris sur le nom pour afficher l'aide contextuelle :

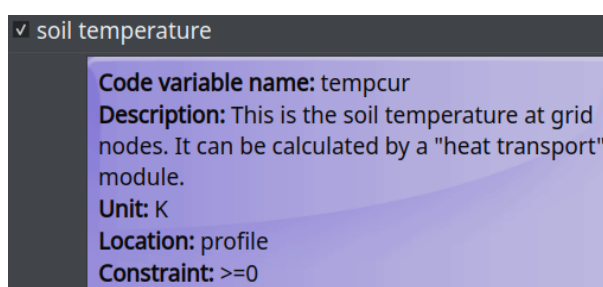


Figure 50: Visualisation du code variable de la sortie soil temperature

(Remarque : Pas clair et on ne comprend pas où il faut faire ce que tu indiques pour avoir le code name)

Le nom minimum des fichiers doit être :

1. Pour les fichiers d'observations : obs_corps1_code_variable_name.csv
 Le contenu du fichier d'observations doit répondre à la structure (cf 6.4).

2. Pour les fichiers de configuration : config_corps1.xml

Attention, lors d'une exécution en multi-configuration, les fichiers de configuration contenus dans le répertoire traité, ne sont pas vérifiés par la plateforme avant d'être exécutés. C'est à dire, qu'on ne vérifie pas si leur structure et format sont conformes à la version de la plateforme utilisée. Il est donc conseillé que ces fichiers de configuration soient testés un par un, en simple simulation, avant d'être utilisés dans la méthode d'estimation de paramètres en multi configuration.

Exemple d'une multi-estimation sur deux sorties. On utilise les deux variables suivantes définies dans la plateforme : "c_co2 organic matter pools production rate profile time cumulate" et "soil organic matter pools net nitrogen flux time cumulated", dont les noms dans le code sont : c_co2_mass_cum et net_n_flux_cumulated). On utilise dans cet exemple trois fichiers de configurations correspondant chacun à une expérience ayant ses propres caractéristiques. Le répertoire contient donc :

1. 3 fichiers de configurations : suffixe Config + "corps" + extension ".xml" (fichier de configuration Vsoil)
2. 6 fichiers d'observations car on a deux variables et trois configurations : suffixe Obs + "corps", identique au fichier de configuration + code de la variable+ extension ".csv"

(Remarque: Le jaune c'est pas très lisible!!)

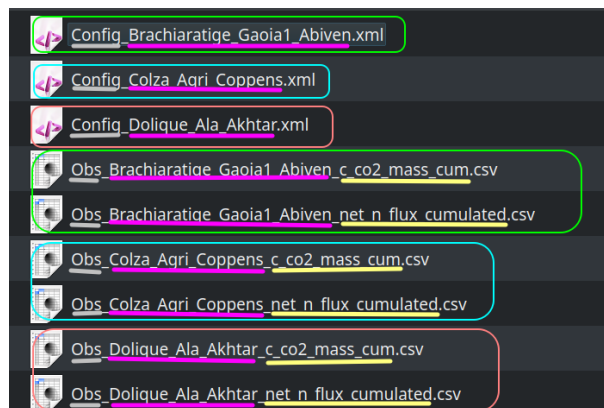


Figure 51: Visualisation du contenu du répertoire pour l'estimation sur 3 fichiers de configuration et 2 sorties

6.5.4 Estimation des paramètres pour une ou toutes les configurations

Pour la méthode DREAM, il est possible de préciser si un paramètre sera estimé pour l'ensemble des configurations ou pour chacune. Quand vous cochez la case "Multi configuration", une liste de choix apparaît avec les options : "fixed" ou "variable". Par défaut le choix est "fixed". Ceci signifie que chaque paramètre est estimé en utilisant toutes les données. Si on veut que certains paramètres soient estimés pour chaque configuration, choisissez "variable". Si "variable" est choisi, vous pouvez alors choisir pour chaque paramètre, si il est variable d'une configuration à l'autre ou s'il est le même pour toutes les configurations. Dans la boîte de définitions des caractéristiques pour chaque paramètre, la caractéristique "sortofparameter" permet de choisir entre "fixed" et "variable".

1. Si, quelque soit le paramètre, l'estimation a lieu sur l'ensemble des configurations, la valeur par défaut est : "fixed".
2. Si, pour certains paramètres, l'estimation varie d'une configuration à l'autre, sélectionnez "variable". Pour chaque paramètre concerné, sélectionnez pour "sortOfParameter" la valeur : "variable" dans la boîte de paramétrage des statistiques par paramètres (cf. 1)

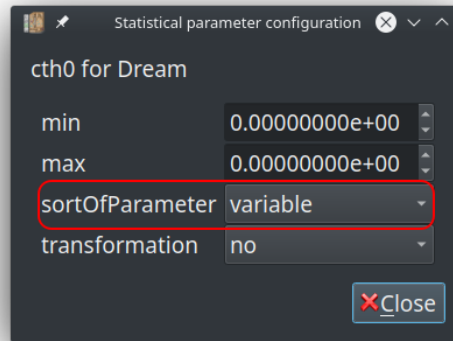


Figure 52: Configuration pour que le paramètre soit estimé pour chaque configuration

Le nombre de paramètres à estimer est lié au nombre de fichiers de configuration. Par exemple si vous avez 10 fichiers de configuration et que vous sélectionnez "variable" pour le paramètre P1 alors vous aurez 10 P1 estimés => P1_configuration1,..., P1_configuration10.

6.5.5 Sélection des sorties et de leur observations associées

Sélectionnez les sorties souhaitées comme pour les autres méthodes 15.

6.5.6 Exécution de l'estimations de paramètres

Une fois le paramétrage et la sélection des sorties réalisés,

- Cliquez sur « Prepar run ».
- Dans l'onglet « Initialization. », cliquez sur « Run stat »

6.5.7 Résultats de l'estimation de paramètres

Les résultats de l'estimation de paramètres seront disponibles dans un nouvel onglet "Result" dans la partie "Statistical result" de l'onglet "Run". Comme déjà précisé, en mode statistique, l'accès à l'onglet "Plot" n'est pas possible. Les résultat sont issus des traitements statistiques de la méthode choisie.

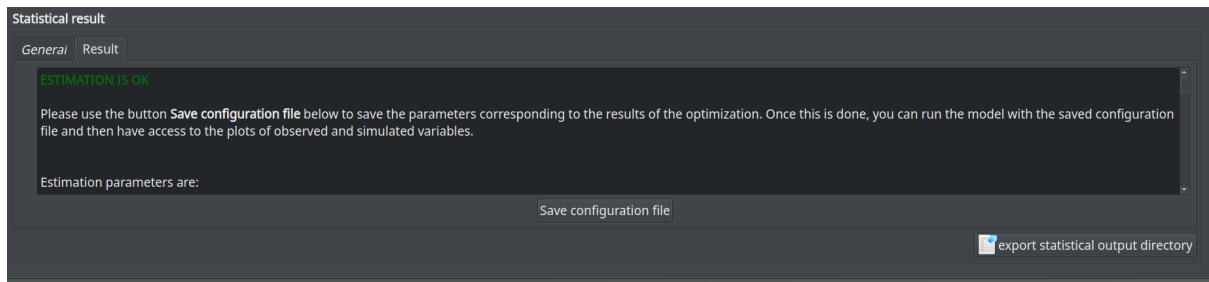


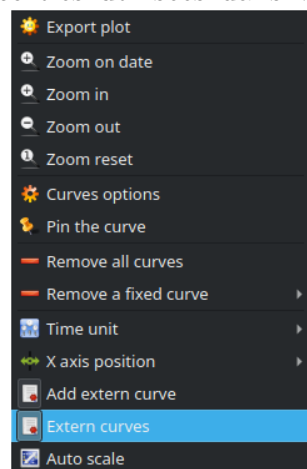
Figure 53: Accès aux résultats de l'estimation de paramètres

Si l'estimation est correctement réalisée vous pouvez sauver les paramètres estimés afin de pouvoir relancer une simulation et comparer observés et simulés. Sélectionnez le bouton "Save configuration file" puis sauver la configuration dans la localisation de votre choix.

Dans le cas de la multi-configuration, tous les fichiers de configurations initiaux seront modifiés avec les valeurs des paramètres estimés. Dans ce cas vous sélectionnerez un répertoire, et l'ensemble des fichiers de configurations modifiés y seront copiés.

Exécution du résultat de votre estimation : Chargez le fichier de configuration issu de l'exécution de votre estimation (onglet "Initialization") puis exécuter. Dans l'onglet "Plot", vous pouvez voir les résultats simulés et les données observées sur un même graphique.

Sélectionnez une des sorties utilisées dans votre estimation, puis sur le graphique, clic



droit "Extern curve":

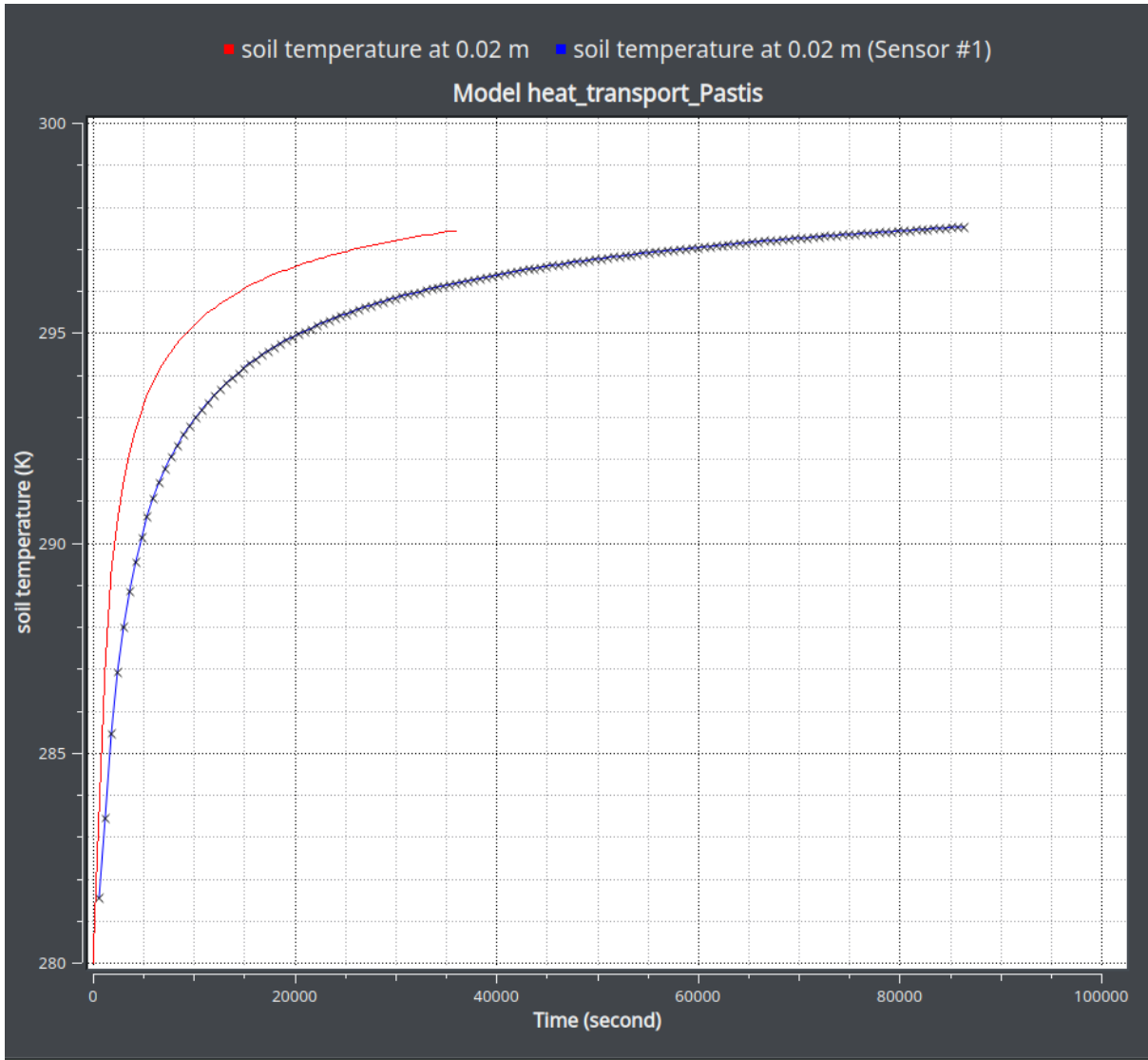


Figure 54: Superposition des données observées et simulées