TP - Evaluation

Bertrand Iooss – biooss@yahoo.fr

Polytech Nice Sophia – Février 2025

Ce TP (à réaliser en binôme ou en solo, mais pas en trinôme !) est à rendre pour le 20 mars 2025. Vous devrez me fournir (par email à biooss@yahoo.fr) un rapport rédigé (dans un **format rédigé et agréable à lire**) des résultats des calculs, de leurs interprétations, des figures et tableaux étayant les interprétations, ainsi que le code R (dans un fichier unique) me permettant de tester vos commandes (initialisez la graine du générateur aléatoire par la fonction *set.seed()*).

La fonction « borehole() » modélise l'écoulement de l'eau à travers un forage. Sa simplicité et sa rapidité d'évaluation en font une fonction couramment utilisée pour tester une grande variété de méthodes. La sortie (c'est-à-dire la réponse de la fonction) est le débit d'eau, en m^3/an. Les 13 variables d'entrée de la fonction et leurs plages de variation habituelles sont :

$r_w \in [0.05, 0.15]$	Rayon du forage (m)
$r_{iw} \in [0.02, 0.08]$	Rayon interne du forage (m)
$r \in [100, 50000]$	Rayon d'influence (m)
$T_u \in [63100, 116000]$	Transmissivité de l'aquifère supérieur (m^2/an)
$H_u \in [1000, 1100]$	Hauteur potentiométrique de l'aquifère supérieur (m)
$T_{um} \in [6310, 11600]$	Transmissivité de l'aquifère moyen supérieur (m^2/an)
$H_{um} \in [900, 1000]$	Hauteur potentiométrique de l'aquifère moyen supérieur (m)
$T_l \in [631, 1160]$	Transmissivité de l'aquifère moyen inférieur (m^2/an)
$H_l \in [800, 900]$	Hauteur potentiométrique de l'aquifère moyen inférieur (m)
$T_l \in [63.1, 116]$	Transmissivité de l'aquifère inférieur (m^2/an)
$H_l \in [700, 800]$	Hauteur potentiométrique de l'aquifère inférieur (m)
$L \in [1120, 1680]$	Longueur du forage (m)
$K_w \in [3000, 12000]$	Conductivité hydraulique du forage (m/an)

Pour la quantification d'incertitudes, les distributions des variables d'entrée aléatoires sont :

$r_w \sim$ Normal(μ=0.10, σ=0.015)
$r_{iw} \sim$ Normal(μ=0.05, σ=0.01)
$r \sim$ Lognormal(μ=7.71, σ=1.0056)
$T_u \sim$ Uniform[63100, 116000]
$H_u \sim$ Uniform[1000, 1100]
$T_{um} \sim$ Uniform[6310, 11600]
$H_{um} \sim$ Uniform[900, 1000]
$T_{lm} \sim$ Uniform[631, 1160]
$H_{lm} \sim$ Uniform[800, 900]
$T_l \sim$ Uniform[63.1, 116]
$H_l \sim$ Uniform[700, 800]
$L \sim$ Uniform[1120, 1680]
$K_w \sim$ Uniform[3000, 12000]

avec Normal(μ, σ) la distribution normale de moyenne μ et variance σ^2 et Lognormal(μ, σ) la distribution lognormale telle que le logarithme de la variable a une distribution N(μ, σ).

1) Mise en place

Récupérez dans le fichier « borehole.R » le modèle de forage (fonction *borehole()*) permettant de calculer le débit d'eau en fonction des 13 entrées. Récupérez également la fonction *EchantBorehole()* qui permettra de générer le vecteur des entrées suivant leurs lois de probabilité.

Faites un premier test en évaluant la sortie aux valeurs minimales, moyennes et maximales des entrées. Donnez les résultats. Pour les valeurs minimales et maximales, prenez les valeurs des plages de variation habituelles. Pour la valeur moyenne de la loi lognormale, ne vous trompez pas et dites comment vous la calculez.

2) Propagation d'incertitudes par Monte Carlo

a) Calculez la moyenne et la variance de la sortie par échantillonnage Monte Carlo de taille 1000. Tracez aussi son histogramme.

b) En comparant au résultat obtenu en question 1), est-ce-que la moyenne de la fonction est égale à la fonction évaluée en la moyenne des entrées ? Qu'en déduisez-vous sur la fonction ?

c) Avec le même échantillon, calculez le quantile d'ordre 95% de la sortie. Calculez-en un intervalle de confiance à 95% en répétant l'estimation Monte Carlo un grand nombre de fois et en récupérant les quantiles à 2.5% et 97.5% du quantile d'ordre 95%.

d) Avec une erreur relative de 10%, calculez par Monte Carlo la probabilité que le débit du forage soit supérieur à 250 m^3/an en augmentant progressivement la taille de l'échantillon (en partant par exemple d'une taille de 2.10^6 avec un pas d'incrément de 2.10^6). Quelle a été la taille nécessaire pour l'échantillon ? Faîtes un graphe de convergence.

3) Analyse de sensibilité

a. Méthode de Morris

Réalisez un criblage des entrées de la fonction forage en utilisant la méthode de Morris (fonction *morris()* du package R « sensitivity ») et avec un coût inférieur à 100 évaluations de la fonction. Expliquez comment vous faîtes et interprétez les résultats. Quelles sont les variables qui peuvent être fixées dans la suite des analyses de sensibilité ?

b. Indices basés sur la corrélation/régression

Dans cette question, les variables identifiées comme non influentes par la méthode de Morris doivent être fixées à leur moyenne.

i. A partir d'un échantillon Monte Carlo de taille 100 sur lequel la sortie 'débit d'eau' est calculée, tracez les scatterplots entre la sortie et les entrées (conservées variables) sur le même graphique. Interprétez-les.

ii. Ajustez un modèle de régression linéaire pour la relation sortie/entrées. Interprétez avec les outils appropriés les résultats de la régression. Calculez les indices de sensibilité SRC^2 et interprétez-les. Est-il nécessaire d'aller plus loin (utiliser une autre méthode) dans l'analyse de sensibilité de la sortie ? N'avait-on pas déjà la réponse à cette question (cf. question 2) b)) ?

c. Indices de Sobol

On oublie les résultats obtenus en a) et b) de cette question 3). A l'aide du package « sensitivity », calculez les indices de Sobol de la sortie par rapport à toutes les entrées en utilisant une taille d'échantillon adaptée et en contrôlant la précision de vos estimations. Utilisez les outils appropriés pour interpréter les résultats.