

**Projet à réaliser en binôme – Data limite de rendu : 27 Octobre 2023**

L'objectif de ce projet est de vous familiariser avec trois méthodes d'analyse de sensibilité (OAT standard, Morris, FAST) et de déterminer, dans le cadre d'un exemple, les paramètres du modèle dont la variation entraîne le plus de variation des sorties du modèle. Vous devrez comprendre et expliciter le code R d'un modèle épidémiologique fourni, puis développer un script pour en réaliser l'analyse de sensibilité à partir des exemples fournis ou de vos propres recherches de solutions. N'oubliez pas de mobiliser les aides fournies avec les packages. Le package R sensitivity est une solution parmi d'autres. Vous pouvez également utiliser ChatGPT comme une aide pour l'élaboration de vos scripts, mais faites-le à bon escient.

Les valeurs par défaut des 15 paramètres sont les suivantes (dans le même ordre que la liste de paramètres dans le script du modèle) : `ValNominale = c(100, 0.5, 0.0014, 0.00029, 0.0019, 0.0019, 0.0082, 5, 1/365, 1/365, 0.3, 1/5, 1/20, 1/100, 0.001)`

Questions :

1. Faites une description du modèle dont le code est fourni :
  - a. Définissez le type de ce modèle (déterministe / stochastique ; à compartiments / individu-centré ; temps discret ou continu)
  - b. Quels sont les processus biologiques en jeu ?
  - c. Ecrivez les équations associées
  - d. Fournissez le schéma des transitions entre états concordant avec ces équations
  - e. Listez les hypothèses principales de ce modèle (quel type de force d'infection, population variable ou constante, etc)
  - f. Quelles sont les conditions initiales ?
  - g. Listez et définissez les sorties possibles du modèle
  - h. Listez les paramètres du modèle, donnez leurs définitions et leurs valeurs
  - i. Réalisez quelques figures (avec titres explicites et légendes) pour illustrer au mieux le comportement du modèle pour les valeurs nominales des paramètres (valeurs fournies)
  - j. Interprétez les figures (de manière concise)

*À noter : pour lancer la fonction `modAppli` (nom du modèle) sous R : (1) définissez les valeurs des paramètres (15 paramètres) sous forme de matrice (les 15 paramètres en colonnes, les scénarios en ligne). Par exemple avec seulement le scénario par défaut, on aurait : `PAR = matrix(ValNominale, nrow=1)` ; (2) lancer la fonction en associant ce qu'elle renvoie à une variable, par exemple : `Sorties = modAppli(PAR)`. « Sorties » possède 4 colonnes (les 4 sorties ponctuelles calculées par la fonction), et autant de lignes que de scénarios dans PAR.*

2. Faites une analyse de sensibilité du modèle en faisant varier les paramètres un par un (OAT = one-at-a-time) sur une gamme réaliste de valeurs (définissez 10 valeurs par paramètres) :
  - a. Justifiez les valeurs testées
  - b. Fournissez votre script
  - c. Réalisez des figures illustrant le comportement du modèle
  - d. Interprétez vos résultats biologiquement et expliciter les limites possibles

3. Faites maintenant une analyse de sensibilité en utilisant la méthode « Morris » : mêmes questions qu'au point 2, en utilisant moins de valeurs par paramètre (ex : levels = 6 et grid.jump = 3) mais un nombre assez élevé de trajectoires (100) pour limiter le temps de calcul à une durée raisonnable (plusieurs minutes néanmoins) tout en explorant l'espace de valeur des paramètres. Plus le nombre de trajectoires est élevé, plus les résultats sont fiables.
  - a. Fournissez une figure équivalente à celle donnée dans le cours et interprétez-la.
  - b. Qu'en concluez-vous si vous comparez les deux approches et les résultats obtenus ?
4. Faites enfin une analyse de sensibilité globale en utilisant la méthode FAST (qui peut s'avérer plus chronophage, vérifiez bien vos temps de simulation avant de les lancer. Le cas à 1000 scénarios par paramètre prend plus d'une heure sur un ordinateur relativement performant).
  - a. Identifiez et explicitez le package que vous utilisez pour cette analyse ; donnez les éléments les plus importants qui peuvent être modulés lors de cette analyse. Notamment, comment choisir le nombre de scénarios à réaliser ?
  - b. Réalisez un premier échantillonnage avec peu de scénarios par paramètre (100). Combien cela fait-il de simulations au total ?
  - c. Sauvegardez le tableau obtenu et réalisez un graphique illustrant les variations des paramètres obtenues.
  - d. Réalisez alors un second échantillonnage avec beaucoup de scénarios par paramètre (1000). De même, sauvegardez le tableau obtenu et illustrez graphiquement votre échantillonnage.
  - e. Comparez les deux échantillonnages et discutez des implications possibles.
  - f. Lancez le modèle sur ces deux échantillonnages en sauvegardant bien à chaque fois les valeurs des sorties dans chaque cas.
  - g. Visualisez la distribution des sorties obtenues : qu'en pensez-vous ? Quel impact dans le cadre d'une analyse de variance (telle que faite pour calculer les indices de sensibilité) ? Pouvez-vous donc interpréter toutes les sorties et pourquoi ?
  - h. Pour les sorties interprétables, calculez les indices de sensibilité principaux (effet principal) et d'ordre 1 (interactions deux-à-deux) et visualisez graphiquement les résultats.
  - i. Comparez les résultats des deux échantillonnages : qu'en concluez-vous ?
  - j. Interprétez biologiquement le cas à 1000 scénarios par paramètre.
5. Discussion : les résultats sont-ils cohérents entre méthodes ? Quelles implications possibles ? Que feriez-vous ensuite en termes d'analyse de ce modèle ?
6. Proposer une modification de votre choix du modèle (ajout d'un état, d'une ou plusieurs transitions)
  - a. Documenter cette modification (hypothèse sous-jacente, équations, schéma, pourquoi ce choix)
  - b. Réaliser les mêmes analyses (OAT, Morris et FAST)
  - c. Discutez les résultats obtenus, incluant une comparaison avec le modèle initial.

Vous pouvez rédiger votre rapport sous forme de *Rmarkdown* (en utilisant par exemple Jupiter), en explicitant de manière concise et précise les points listés ci-avant.

Les meilleurs rendus seront susceptibles d'être publiés sur le site web du master MODE et serviront d'exemples pour le cours en analyse de sensibilité de l'année prochaine.