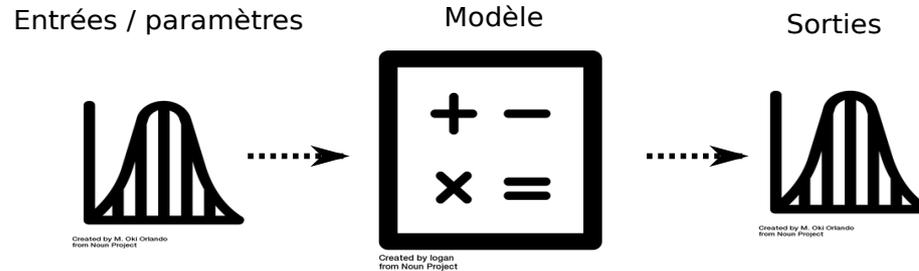


La notion d'incertitude et les outils pour l'analyser

Ronan Trépos, MIAT, INRAe
séminaire RMT Data Science & Modélisation
06/01/2023

Modèles et entrées / sorties



Météo :

conditions initiales

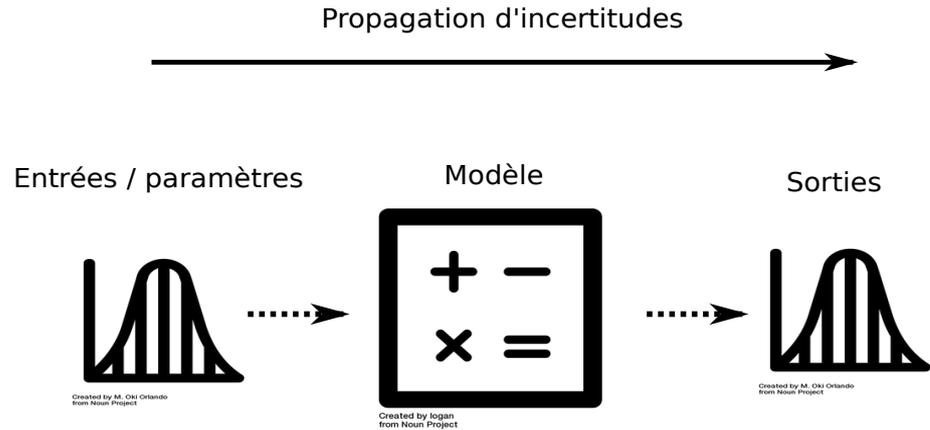
séries temporelles
(température, pression atmosphérique)

Agro :

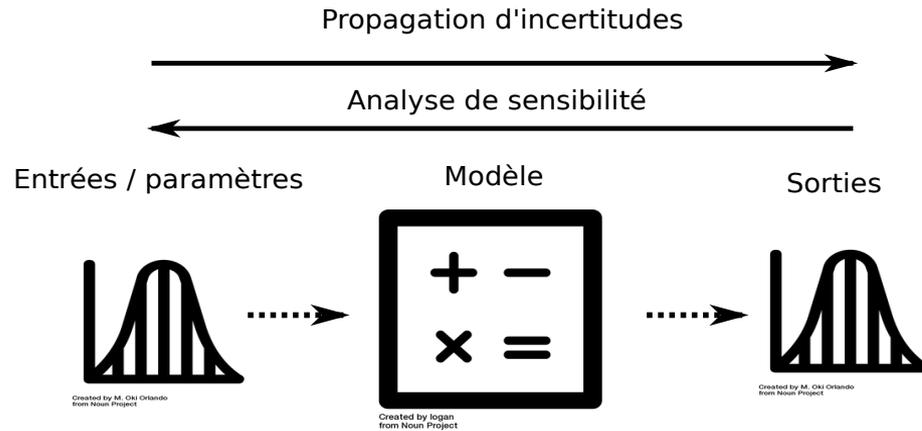
données météo
profondeur de sol

teneur en eau du sol
rendement

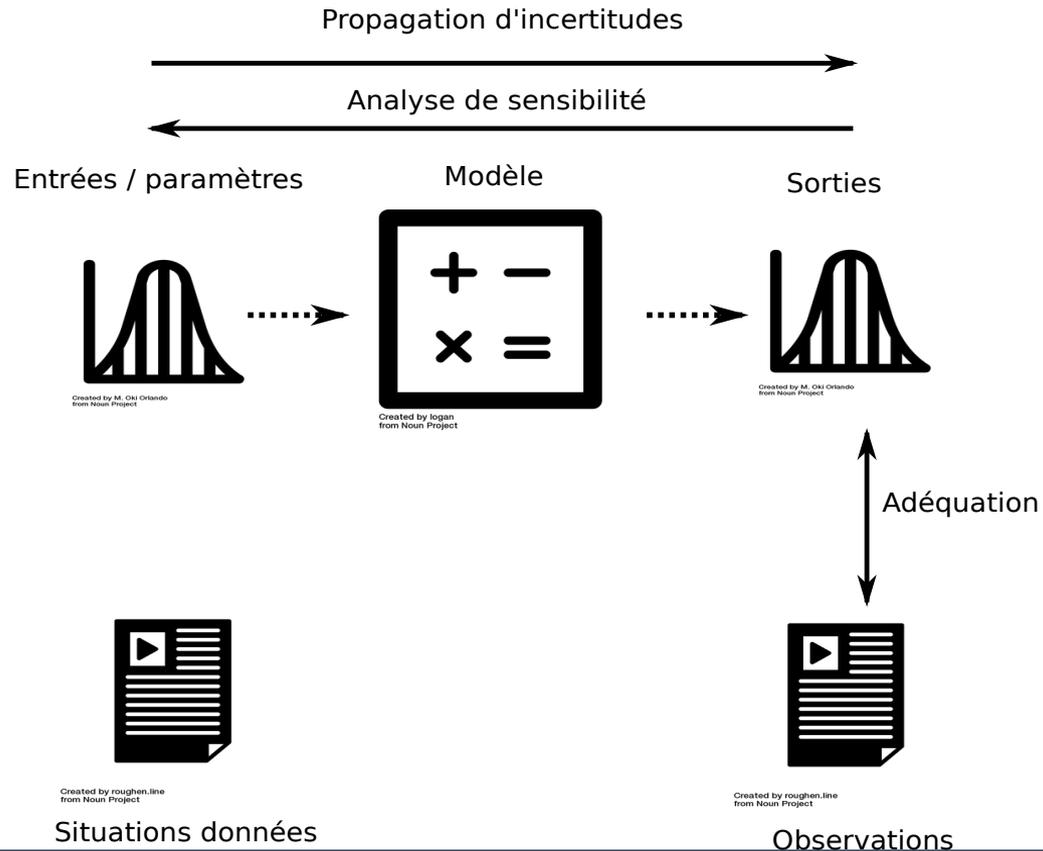
Propagation d'incertitudes



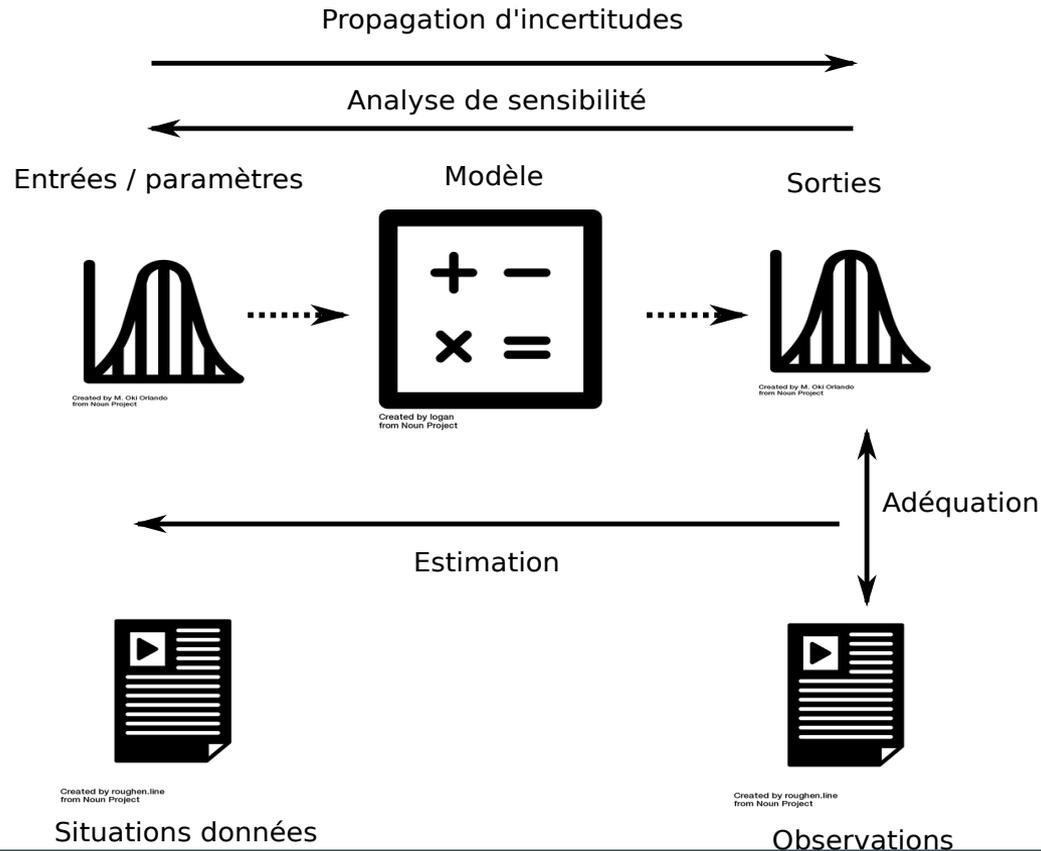
L'analyse de sensibilité



Confrontation aux données



Estimation de paramètres/entrées

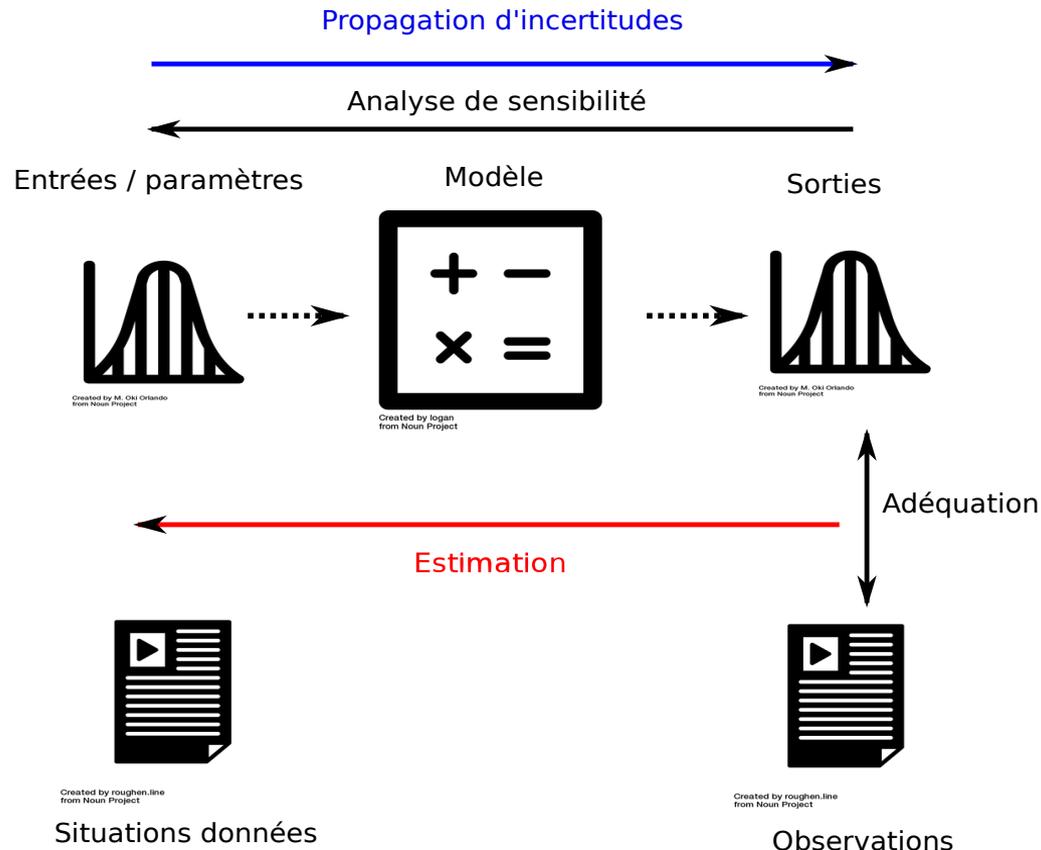


Sources d'incertitudes impactant la sortie d'intérêt

- incertitude sur les entrées/paramètres
 - ex. : capacité de rétention de l'eau est un paramètre impactant le rendement
- inadéquation du modèle : aucun modèle n'est parfait (hypothèses simplistes)
 - ex : avec ou sans impact de caractéristiques variétales
- variabilité résiduelle : stochasticité
 - ex : le rendement est X plus ou moins Y
- variabilité paramétrique : impossibilité de réaliser les entrées
 - ex : le travail du sol ne peut se faire.
- erreur d'observation : la mesure est idéalement attachée d'une erreur
 - ex : échantillon de mesure au champs
- algorithmique/code : problèmes numériques

Kennedy, O'Hagan (2001).
"Bayesian calibration of computer models".
Journal of the Royal Statistical Society.

Propagation *avant* des incertitudes/ évaluation inverse de l'incertitude du modèle



Les méthodes Monte Carlo pour la propagation d'incertitudes

Distribution des entrées / paramètres

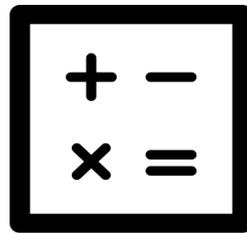


Created by M. Oki Orlando from Noun Project

Echantillon des entrées / paramètres

$$\left\{ \begin{array}{l} x_1^1, x_2^1, \dots, x_n^1 \\ x_1^2, x_2^2, \dots, x_n^2 \\ \dots \\ x_1^M, x_2^M, \dots, x_n^M \end{array} \right\}$$

Modèle

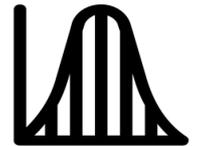


Created by logan from Noun Project

Echantillon des sorties

$$\left\{ \begin{array}{l} y^1 \\ y^2 \\ \dots \\ y^M \end{array} \right\}$$

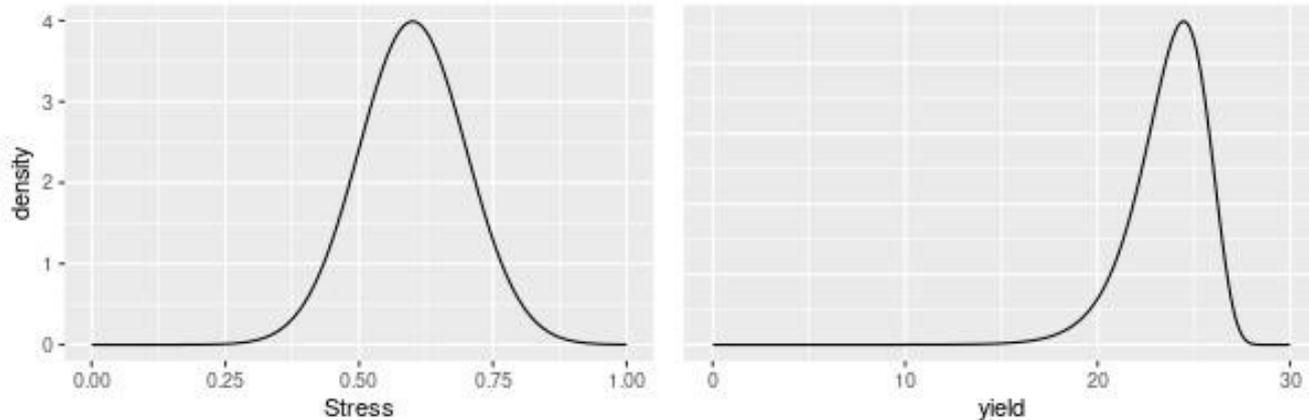
Distribution des sorties



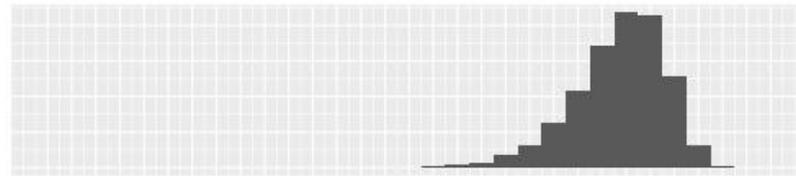
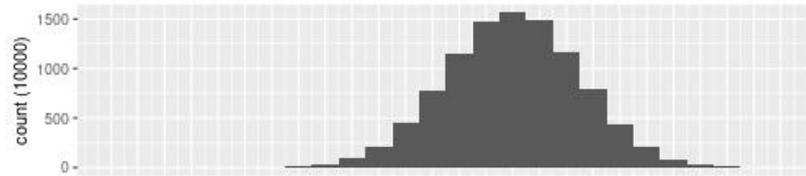
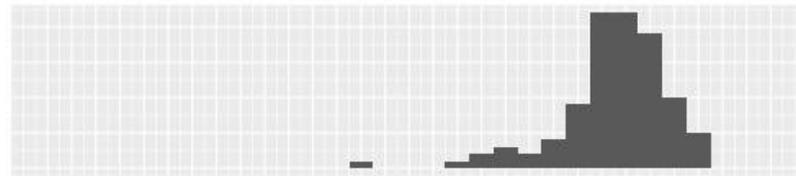
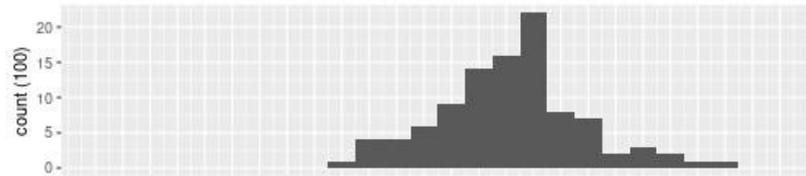
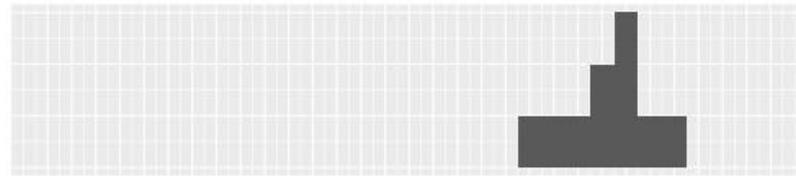
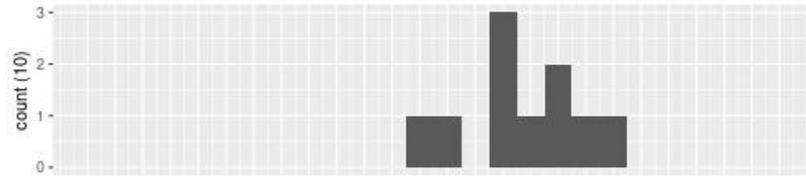
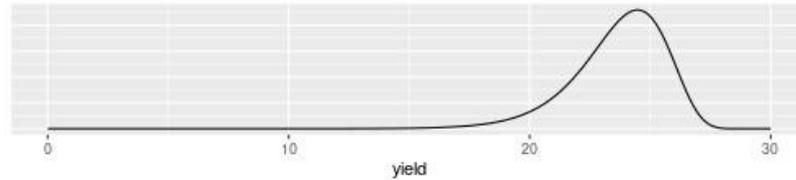
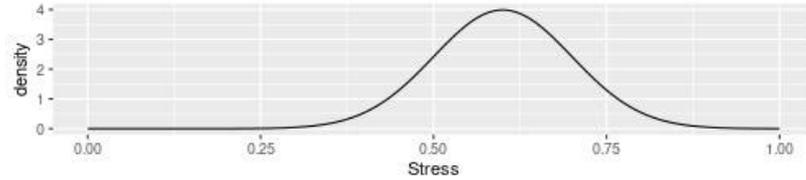
Created by M. Oki Orlando from Noun Project

Exemple

- mon modèle : $\text{yield} = 30 - \exp(3 * \text{Stress})$
- paramètre Stress incertain : $\text{Stress} \sim \text{Normal}(0.6, 0.01)$
- sortie incertaine : $\text{yield} \sim 30 - \text{LogNormal}(1.8, 0.09)$



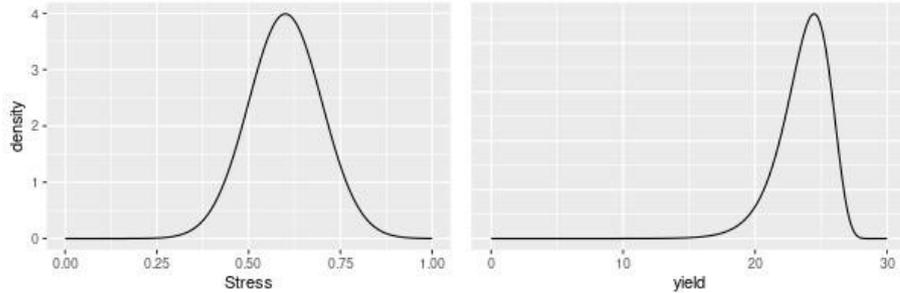
Effet d'échantillonnage



Extraire des statistiques de la sortie

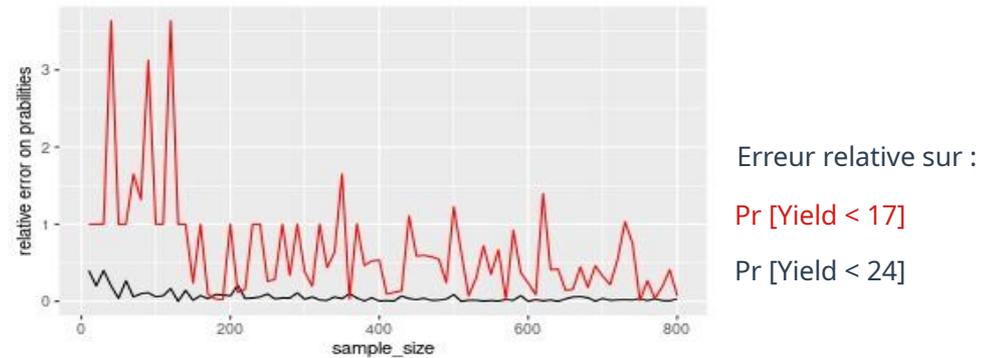
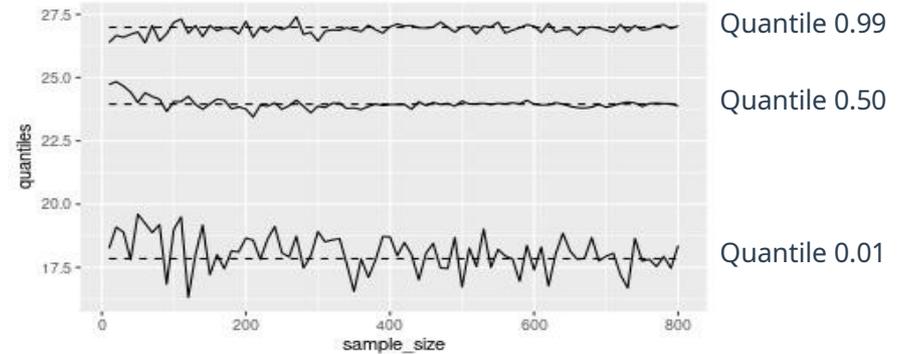
- **Extraire la moyenne et la variance (moments d'ordre faible)**
 - cela peut souvent suffire à caractériser la distribution entièrement dans un cadre paramétrique
- **Étudier la fiabilité du système.**
 - Par exemple, un rendement inférieur à 17 est considéré être une faille. Quelle est sa probabilité ?
- **Décrire la distribution entièrement**
 - approche paramétrique (choix d'une distribution)
 - ou non paramétrique (méthodes à noyau, choix d'hyper paramètres)

Fiabilité de système



Traiter les événements rares requiert un tirage Monte Carlo plus important. Des solutions existent pour limiter le budget :

- utilisation de tirages biaisés
- utilisation de *bootstrap*



Conclusion

- **De nombreuses sources d'incertitudes en lien avec l'utilisation d'un modèle**
- **Effet d'échantillonnage sur les méthodes MC pour la propagation d'incertitudes**
 - attention au budget nécessaire pour extraire les statistiques d'intérêt
- **Et puis..**
 - interpréter la variabilité de la sortie en fonction de l'incertitude modélisée en entrées.