



Calibration et optimisation de modèles

ULg - Arlon Campus Environnement

RISQ2002-1 - Analyse et Modélisation des systèmes naturels

Julien Minet, Bernard Tychon

Plan du cours

1. Introduction/enjeux
 1. Qu'est-ce qu'un modèle ?
 2. Terminologie
 3. Pourquoi calibrer ?
2. Comment calibrer un modèle ?
 1. Régression linéaire simple
 2. Calibration d'un modèle complexe
 3. Fonction objective
 4. Algorithmes d'optimisation
 5. Optimisation bayésienne

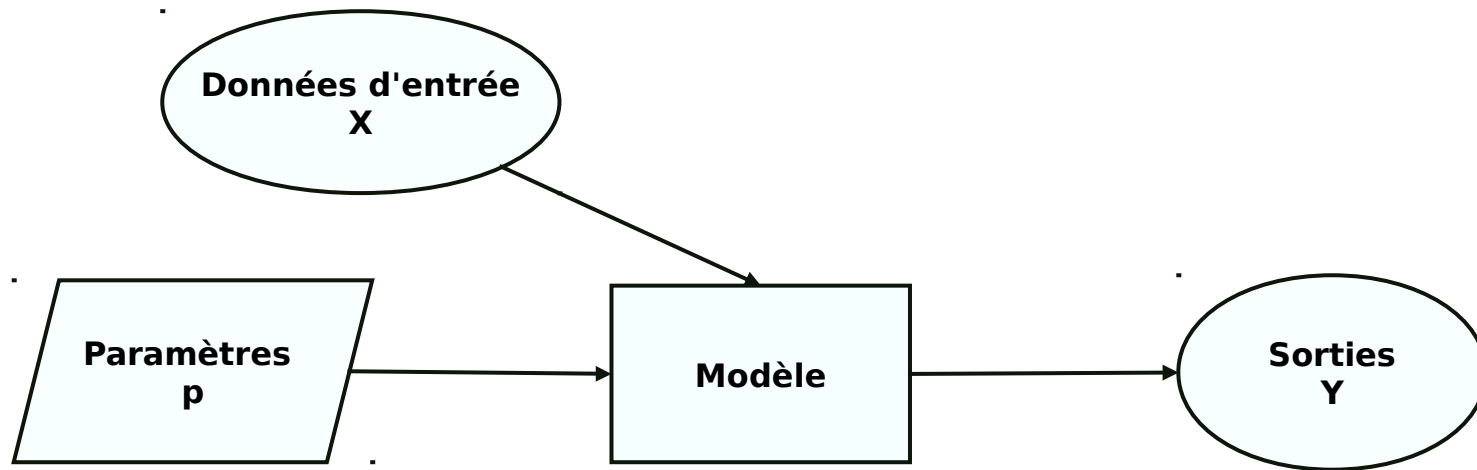
Qu'est-ce qu'un modèle

- Représentation de la réalité
- Un modèle (M) peut s'exprimer sous la forme mathématique suivante :

$$Y = M(X, p)$$

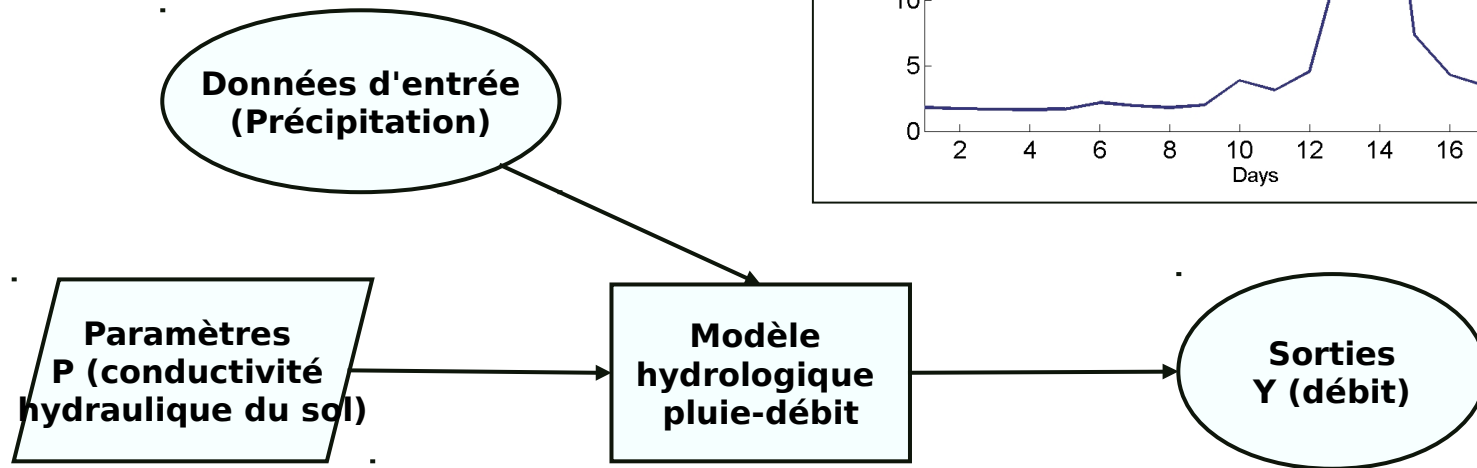
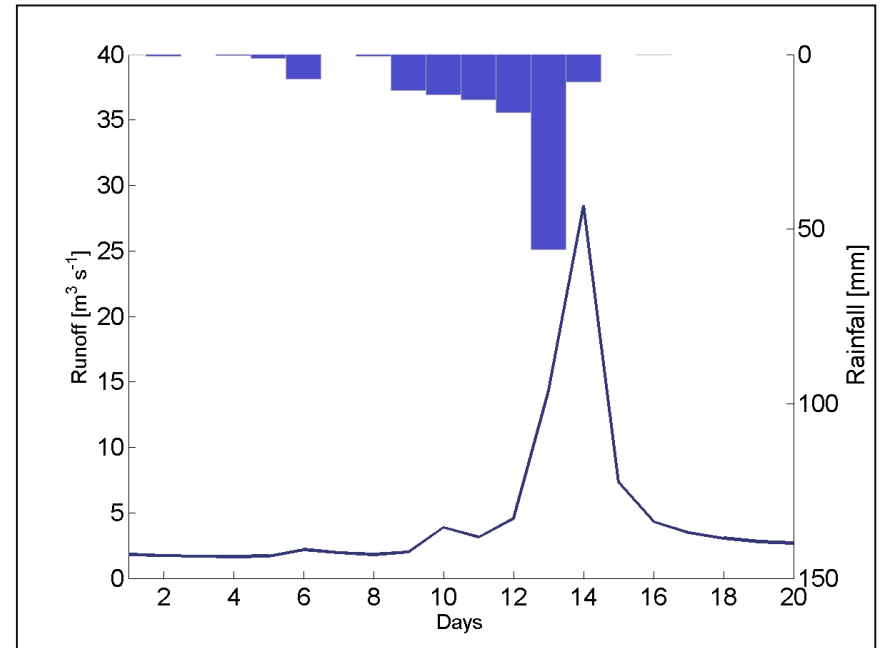
- Où Y est la (les) sortie(s) du modèle, la variable modélisée, la variable expliquée
- X est l'entrée (les entrées) du modèle, la variable explicative
- **p** sont les paramètres du modèles

Qu'est-ce qu'un modèle



$$Y = M(X, p)$$

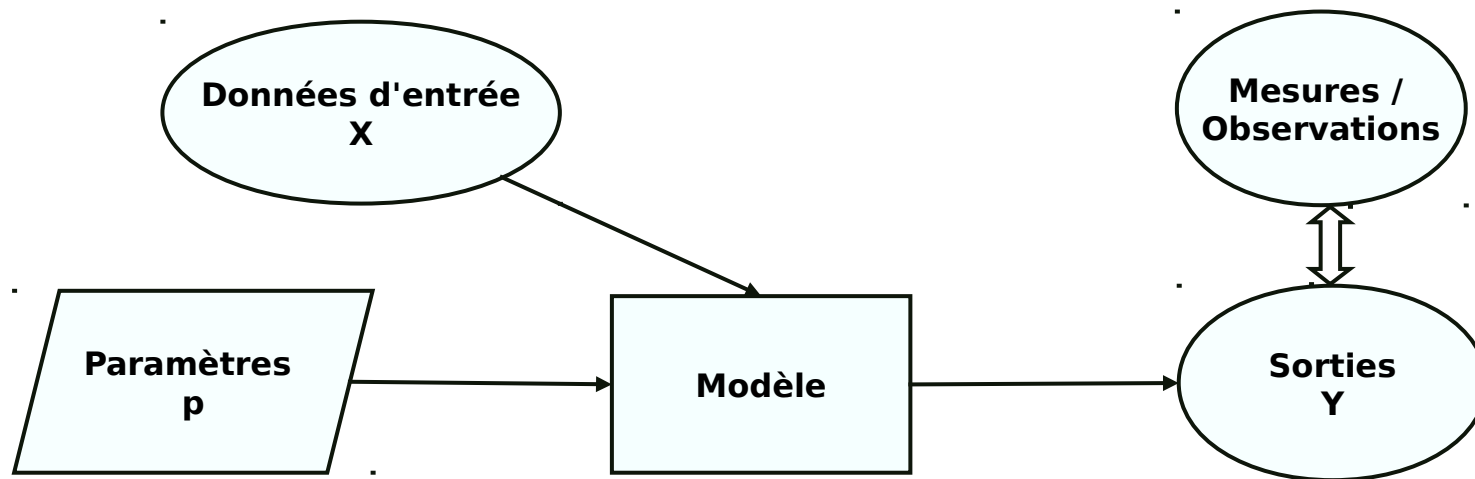
Qu'est-ce qu'un modèle



$$\text{Débit} = M(\text{Précipitations}, \text{paramètres})$$

Principe de la calibration

- Trouver les meilleurs paramètres p qui font correspondre les sorties Y du modèle à des observations



On a toujours besoin d'observations pour calibrer !

Terminologie

- **Calibration** d'un modèle : Ajuster les paramètres du modèle pour améliorer la prédiction
- = « **Optimisation** (des paramètres) du modèle »
- **Modélisation inverse** : Calibrer un modèle en utilisant des données observées qui sont comparées à la sortie du modèle
- **Assimilation de données** : Calibration d'un modèle, avec un accent sur des données temporelles
- **Validation** : Vérifier l'exactitude d'un modèle calibré.

Pourquoi calibrer les modèles ?

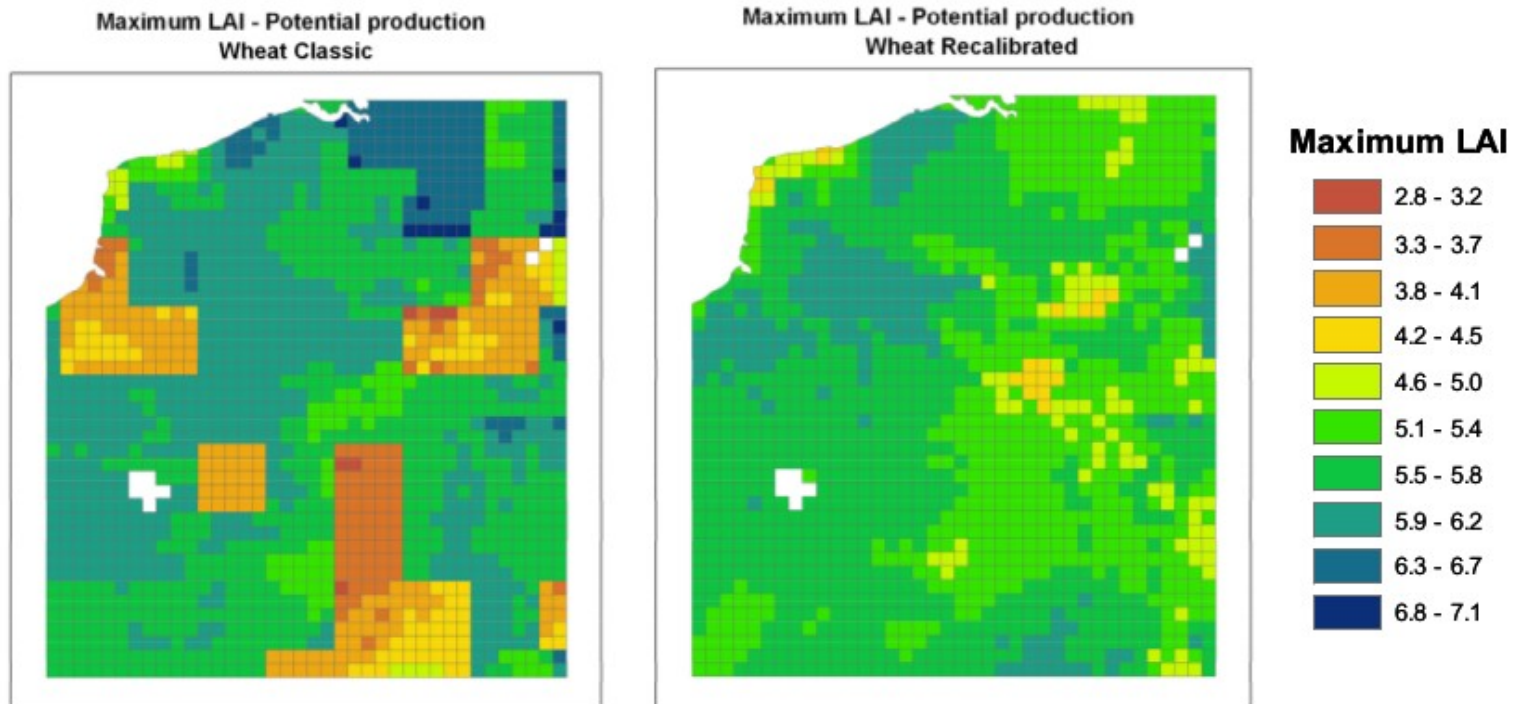
- Pour améliorer les prédictions
- Pour améliorer le modèle (dans le but d'avoir des applications opérationnelles)
- Pour retrouver des paramètres / variables d'intérêt

Pourquoi calibrer les modèles ?

- Pour améliorer les prédictions : $Y = M(X, p)$

Exemple : améliorer la prédiction du rendement des cultures sur une région agricole

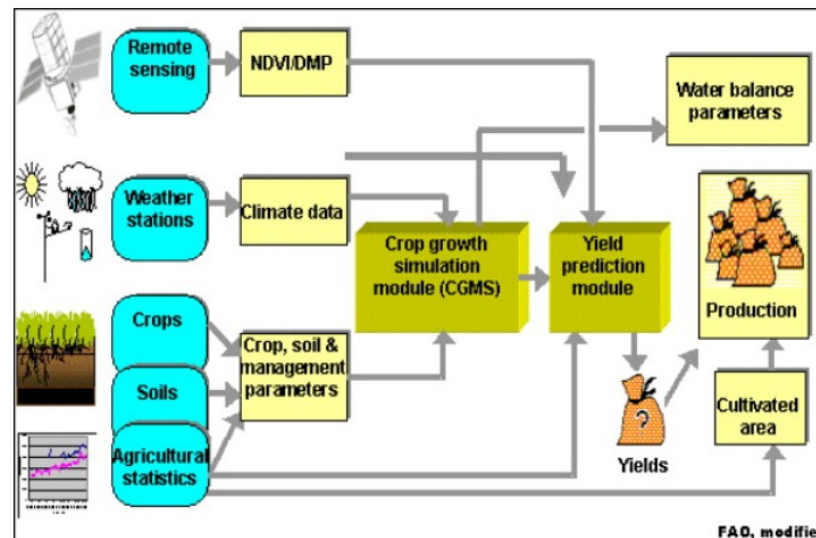
La prédiction est meilleure après calibration



Pourquoi calibrer les modèles ?

- Pour améliorer le modèle :
- Exemple : Améliorer le fonctionnement d'un modèle de prédiction de rendements des cultures en introduisant l'effet des maladies des plantes, en calibrant avec des données de terrain (El Jarroudi et al., Eur. J. of Agronomy, 2012)

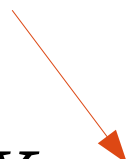
$$Y = M(X, p)$$



Pourquoi calibrer les modèles ?

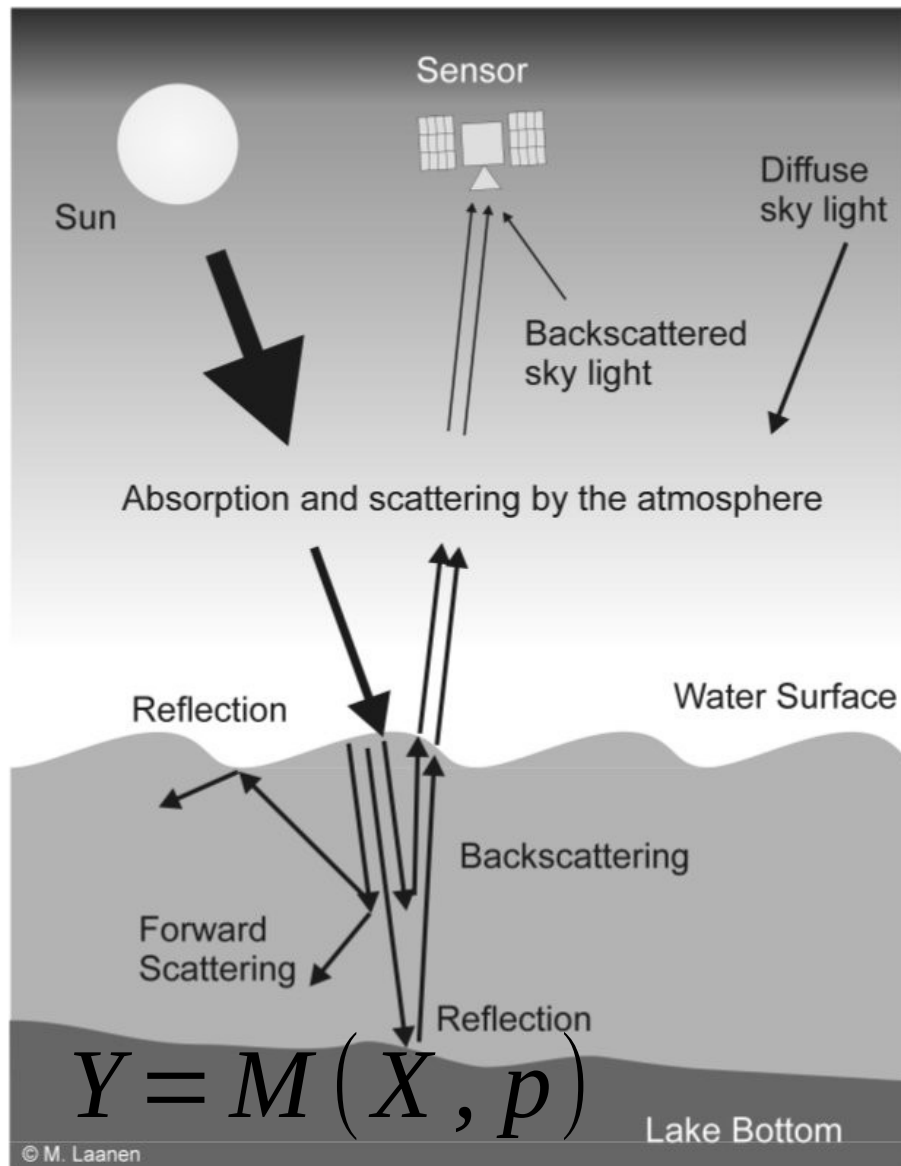
- Pour retrouver des paramètres / variables d'intérêt

Parfois, l'intérêt du modèle n'est pas dans la sortie du modèle mais dans la valeur réelle de ses paramètres. Par exemple, un modèle décrivant la propagation d'une onde radar entre un satellite et la surface du sol peut être calibré pour retrouver les propriétés de surface (sol, eau).

$$Y = M(X, p)$$


Pourquoi calibrer les modèles ?

- Pour retrouver des paramètres / variables d'intérêt

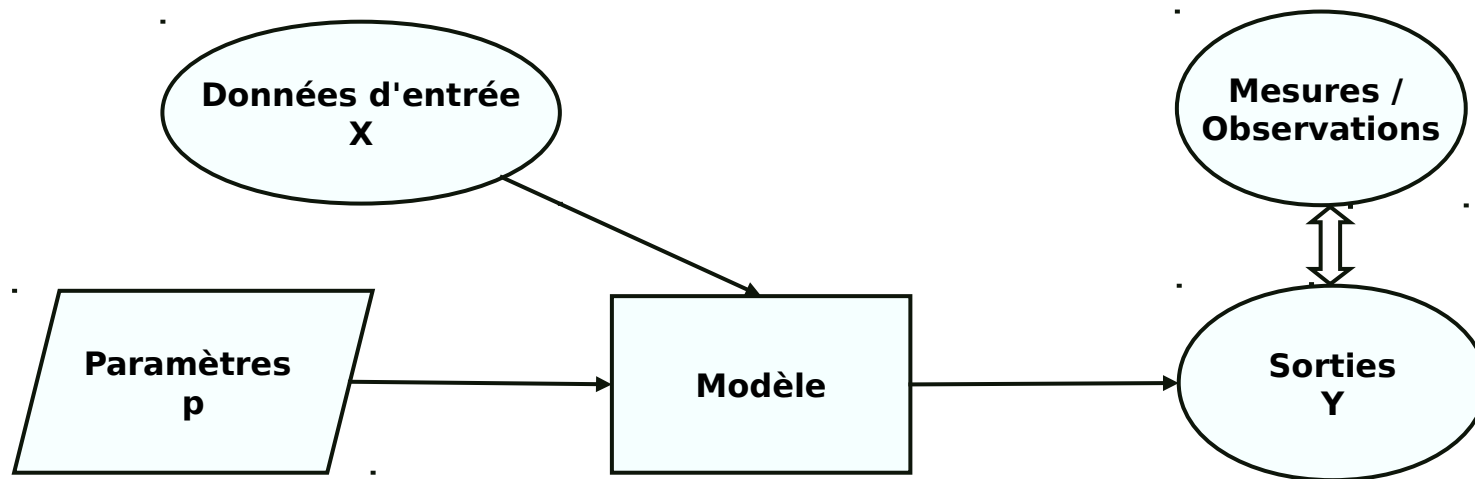


Exemple :

- Un modèle électromagnétique décrit la propagation de l'onde entre le satellite et la surface
- Calibrer ce modèle permet de retrouver la propriété de surface (couleur, humidité du sol, hauteur d'eau d'un lac, salinité, etc.)

Principe de la calibration

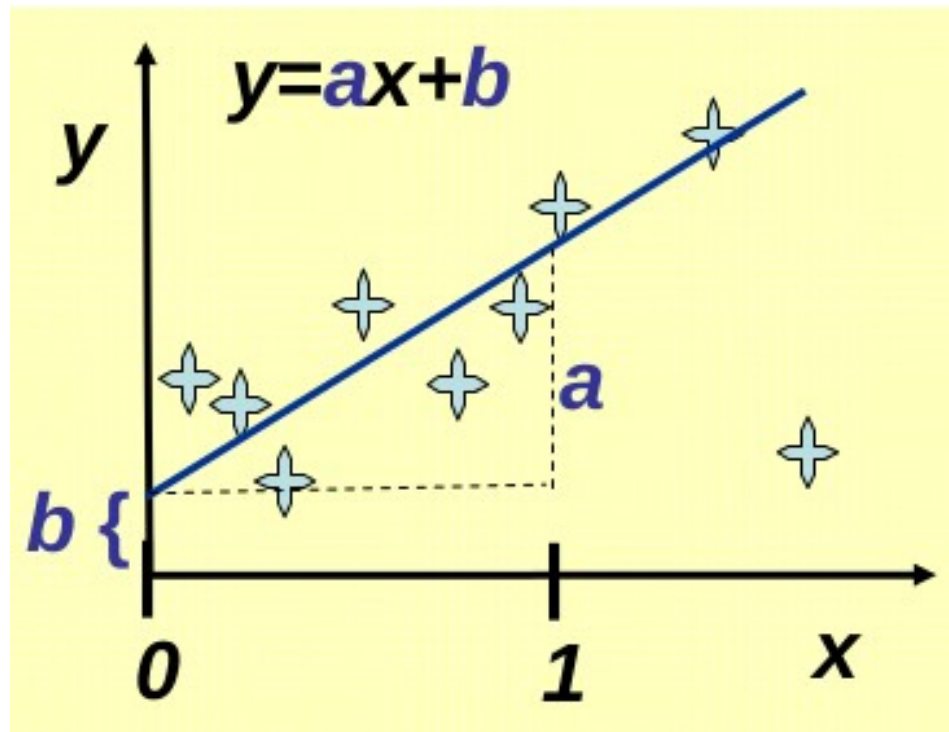
- Trouver les meilleurs paramètres p qui font correspondre les sorties Y du modèle à des observations



On a toujours besoin d'observations pour calibrer !

Régression linéaire simple

- Modèle simple (linéaire) $Y = aX + b$
- Retrouver les paramètres peut se faire par une régression linéaire simple

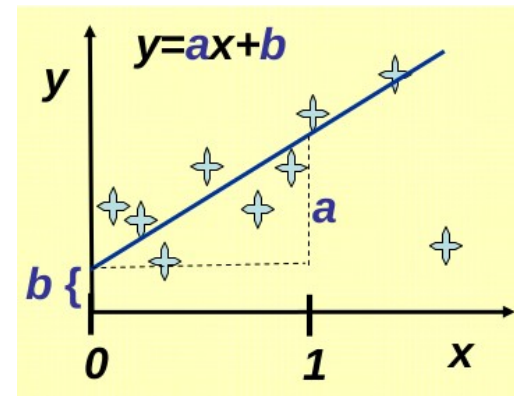


Régression linéaire simple

- Régression linéaire simple : les paramètres du modèle $Y = aX + b$ peuvent être retrouvés par la méthode des moindres carrés :
- Soit un vecteur d'observations Y (variable expliquée) lié par une fonction linéaire à un vecteur d'une variable indépendante X (variable explicative)

$$X = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

$$Y = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

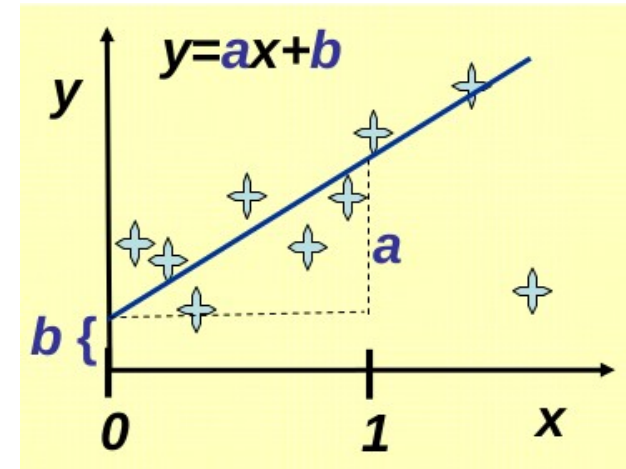


Régression linéaire simple

- Les paramètres a et b sont estimés par les formules suivantes :

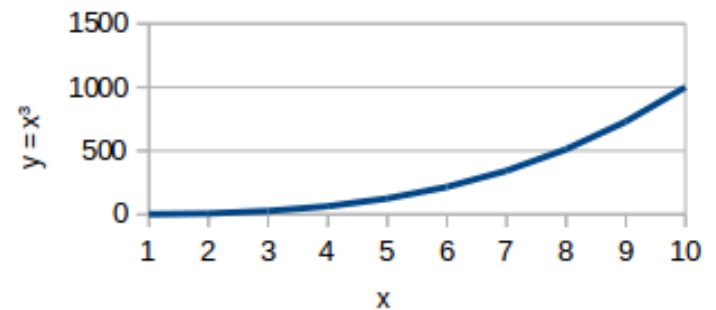
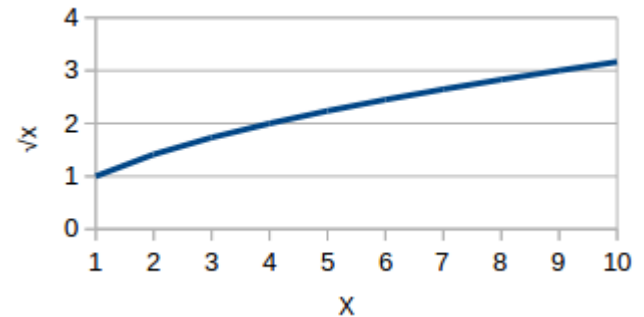
$$a = \sum_{i=1}^n \frac{(x_i - \bar{x})(y_i - \bar{y})}{(x_i - \bar{x})^2} = \frac{\text{cov}(X, Y)}{\text{var}(X)}$$

$$b = \bar{y} - a\bar{x}$$



Régression linéaire par transformation de variable

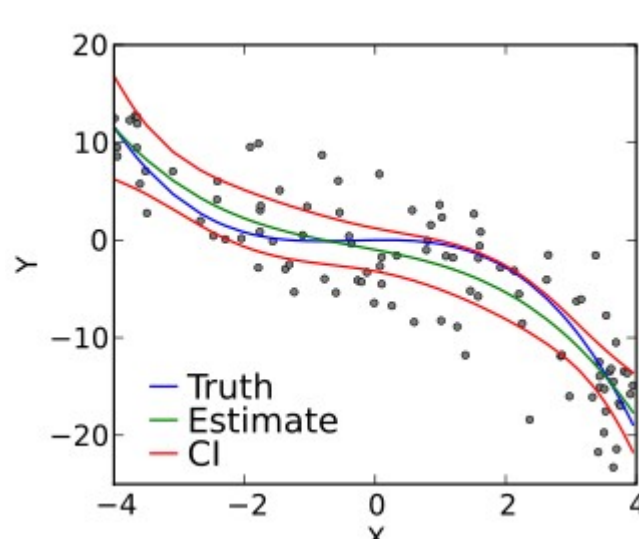
- Si le modèle n'est pas linéaire, il est parfois possible de le linéariser, c'est-à-dire de le rendre linéaire en faisant un changement de variable
- Exemples :
 - 1) Si $y = a.\sqrt{x}$, posons $x' = \sqrt{x}$, alors $y = a . x'$
 - 2) Si $y = a.x^n$, posons $\ln(y) = \ln(a) + n.\ln(x)$, alors $y' = a' . x'$



Régression linéaire par transformation de variable

- On peut aussi faire une régression polynomiale en utilisant différentes méthodes de lissage

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

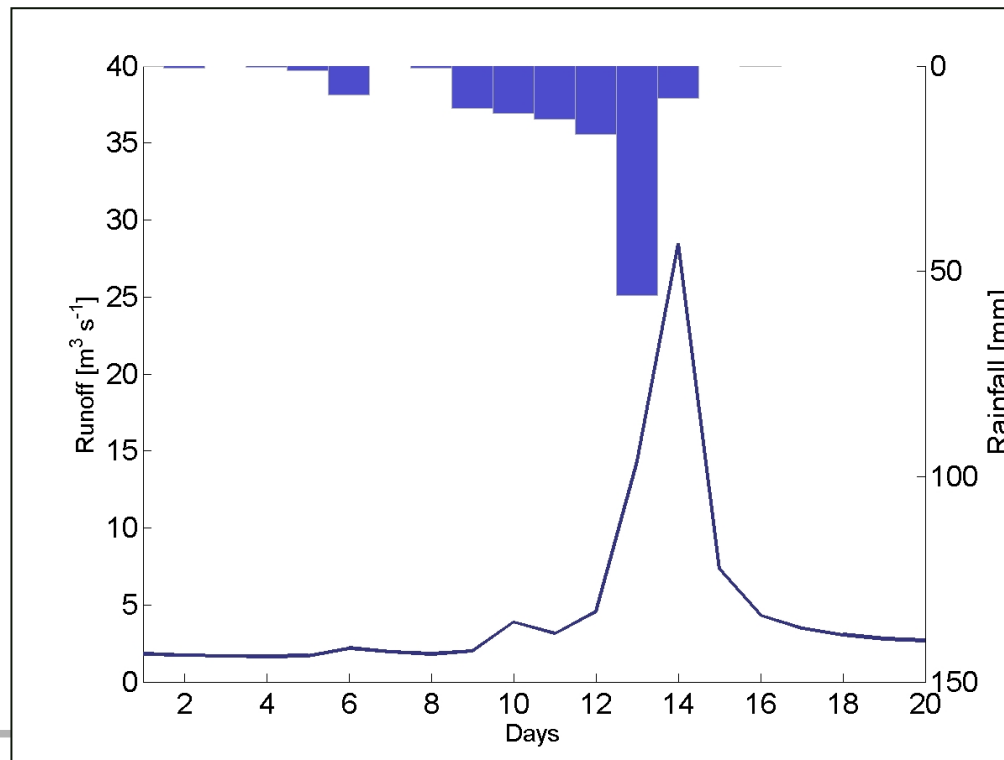


http://fr.wikipedia.org/wiki/R%C3%A9gression_polynomiale

Calibration d'un modèle complexe

Mais que se passe-t-il avec un modèle beaucoup plus complexe ?

Par ex. un modèle de relation pluie-débit est non-linéaire et n'a pas de solution analytique.



$$Y = aX + b$$

Calibration d'un modèle complexe

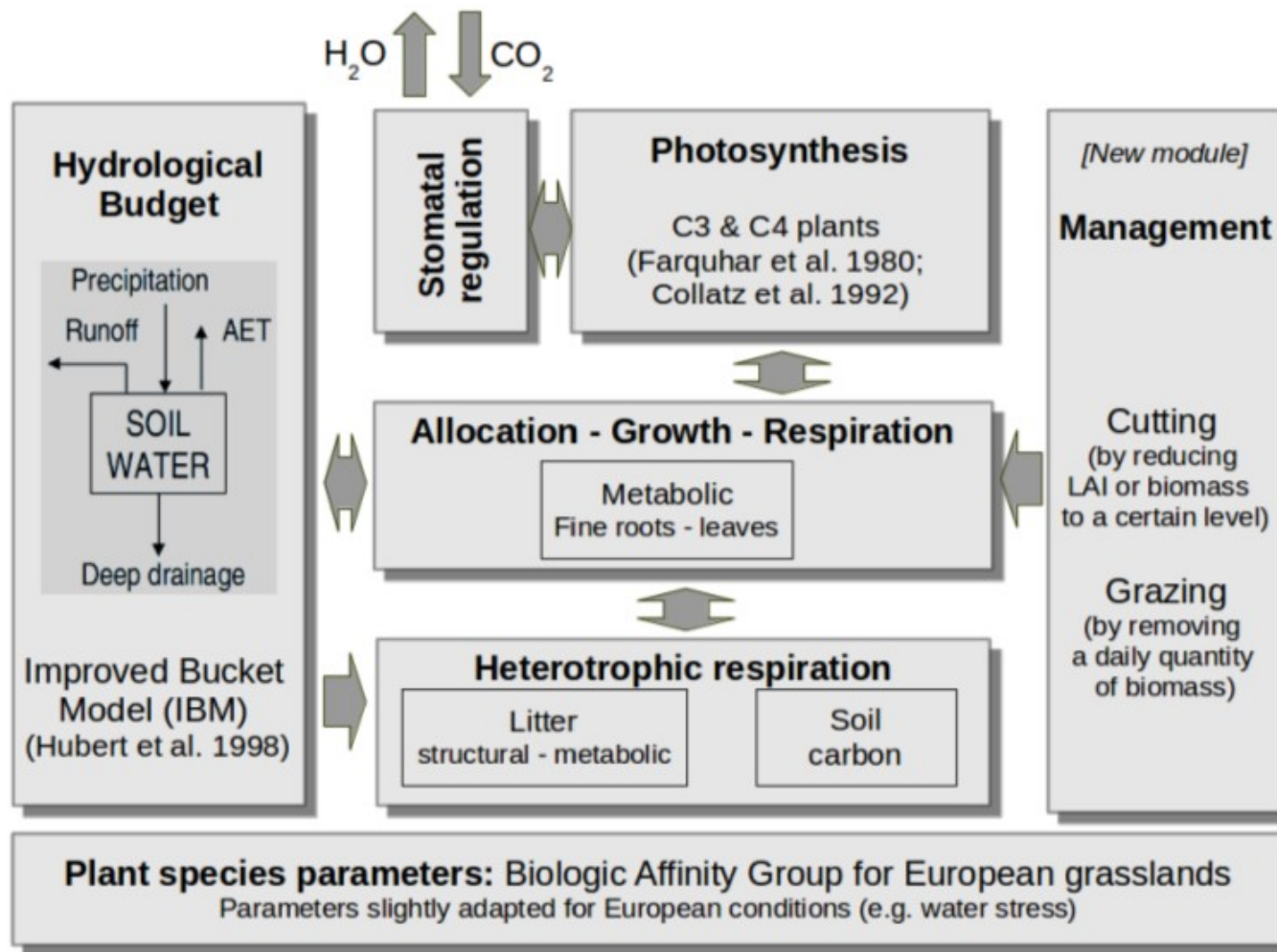
Beaucoup de modèles numériques n'ont pas de solutions analytique et il n'y a pas de modèles inverses :

- Modèles de prévisions météos
- Modèles climatiques
- Modèles de croissance des plantes/cultures
- Modèles hydrologiques
- Etc.

$$~~Y = aX + b~~$$

Calibration d'un modèle complexe

Modèle de croissance des plantes CARAIB (ULg)

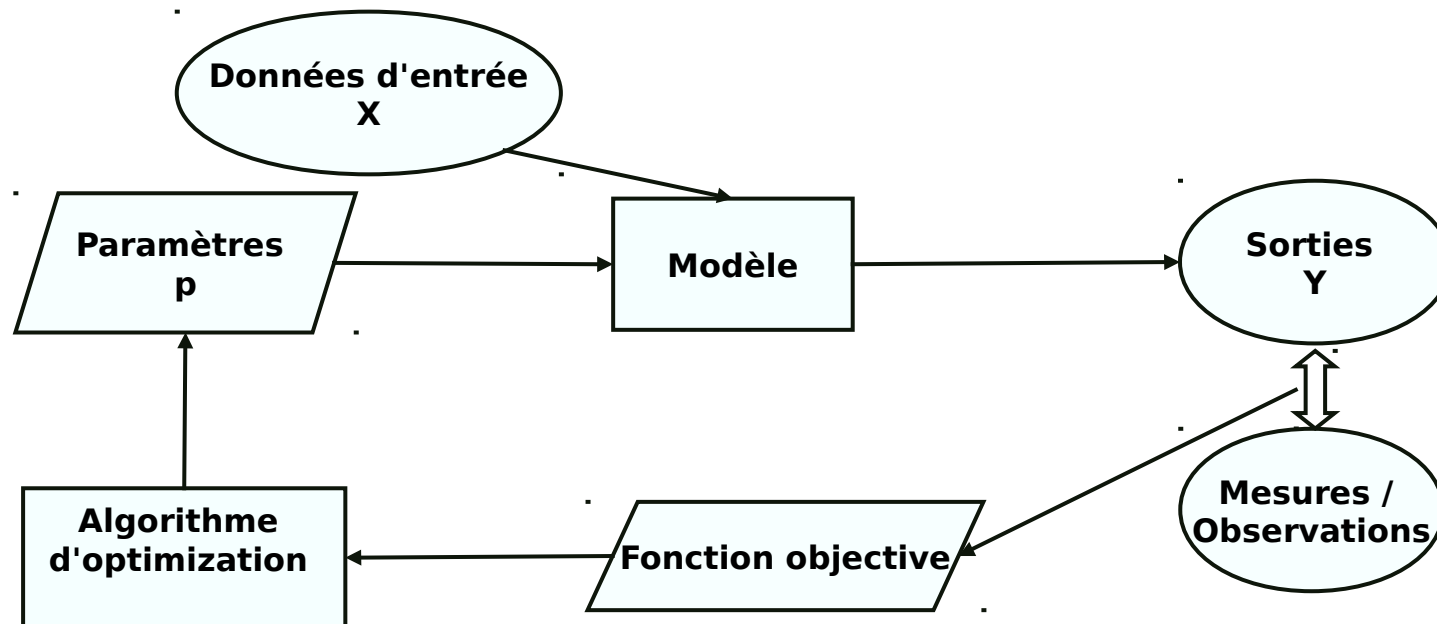


~~$$Y = aX + b$$~~

Calibration d'un modèle complexe

Calibration d'un modèle non-linéaire / sans solution analytique :

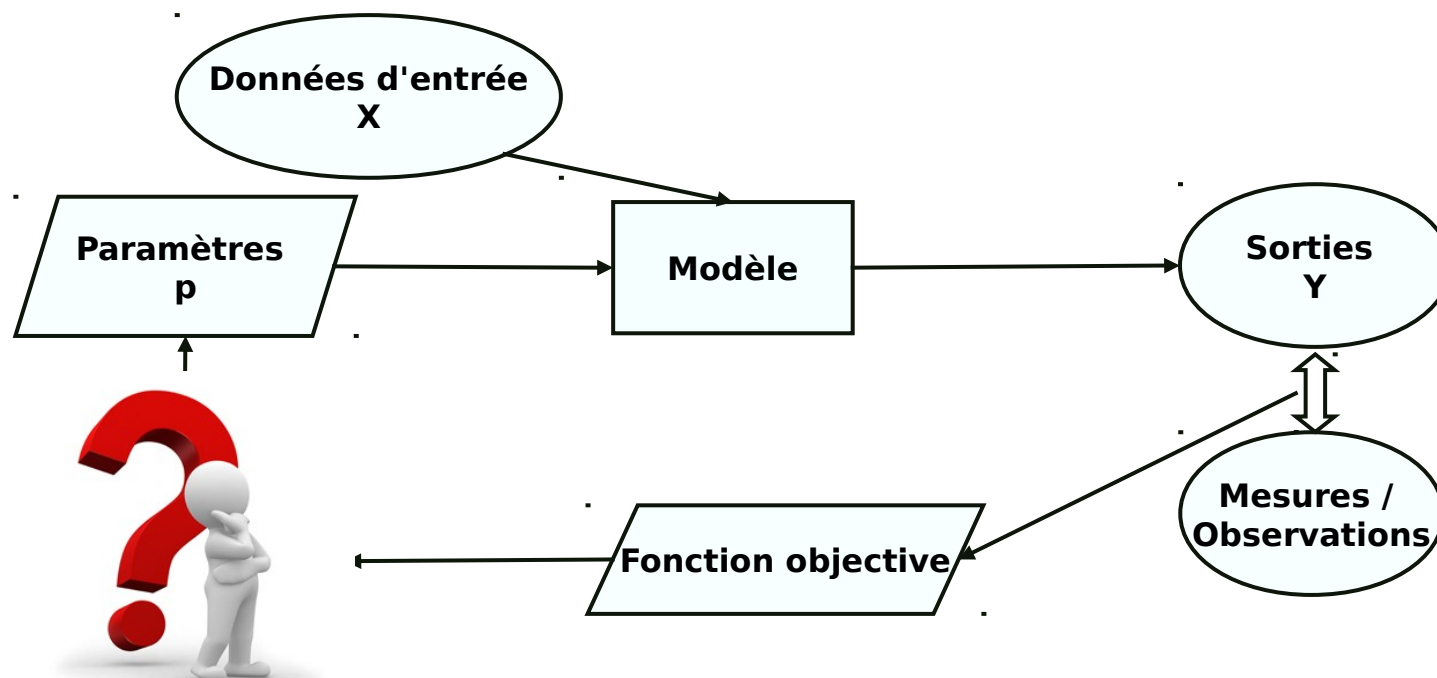
- Procédure **itérative**, avec un algorithme d'optimisation qui trouve le minimum d'une fonction objective



Calibration d'un modèle complexe

Calibration d'un modèle non-linéaire / sans solution analytique :

- Procéder par Essai-Erreur est très vite limitant si le nombre de paramètres est important



Calibration d'un modèle complexe

- Calibration d'un modèle de façon itérative avec un algorithme : exemple avec le modèle de végétation CARAIB

Fonction objective

- Synonyme : fonction de coût, fonction d'utilité,...
- Principe : La différence entre les prédictions du modèle et les observations doit être aussi petite que possible :

$$OF(p) = \sum (Obs - Y(X, p))^2$$

- Il s'agit donc de minimiser la fonction objective, de trouver son minimum en fonction des paramètres.
- Les paramètres optimaux sont ceux pour lesquels la fonction objective est minimale

$$p_{opt} = \operatorname{argmin}(OF(p))$$

Fonction objective

- Une fonction objective mesure la différence entre la prédiction du modèle et les observations (Obs). Elle peut prendre des formes différentes.
- Mais on prend très souvent la somme des différences au carré, parce que cela correspond à l'hypothèse que les erreurs de mesures sont normalement distribuées.

$$OF(p) = \sum (Obs - Y(X, p))^2$$

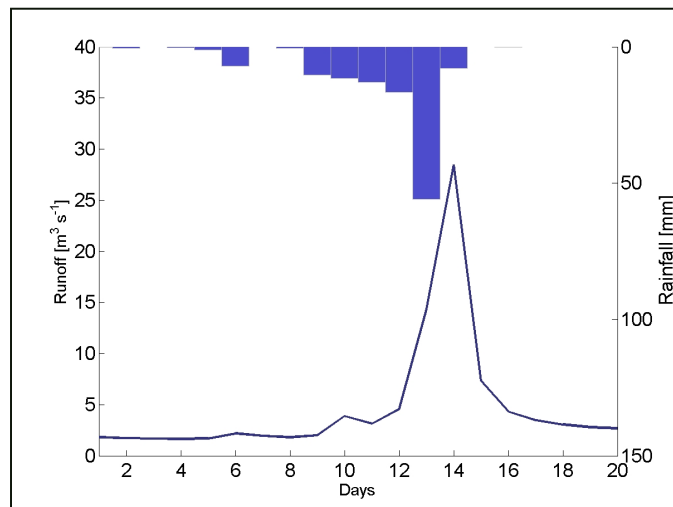
- Une fonction objective égale à la valeur absolue de la différence entre la prédiction du modèle et les observations correspond à des erreurs de mesures distribuées selon une double exponentielle.

$$OF(p) = \sum (|Obs - Y(X, p)|)$$

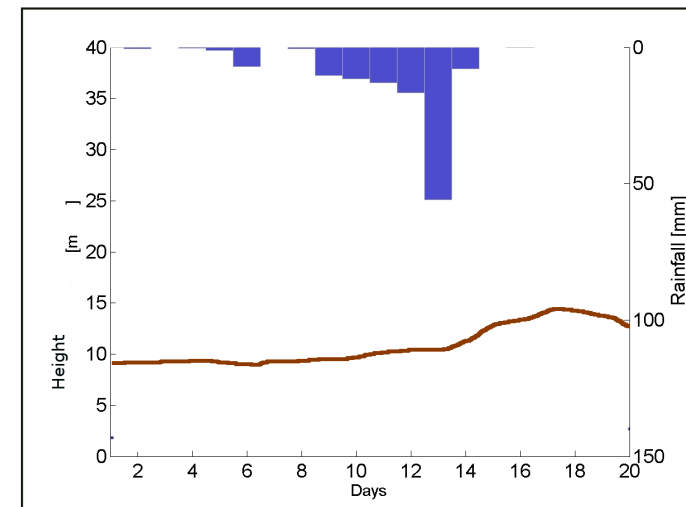
Fonction objective

- Fonction **multi-objectifs** : permet de prendre en compte plusieurs observations.
- Par exemple, pour un modèle pluie-débit, on prend en compte des mesures de débit + des mesures de hauteurs d'eau dans la nappe.

Débit



Hauteur de la nappe



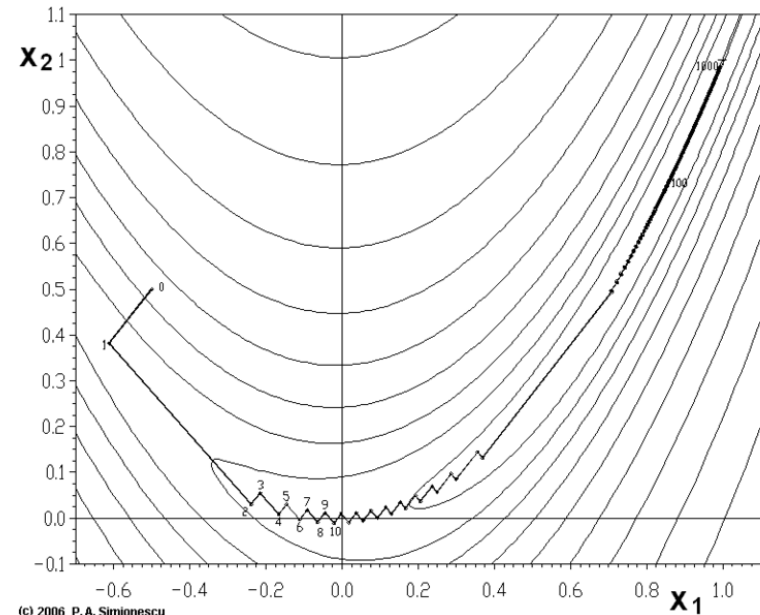
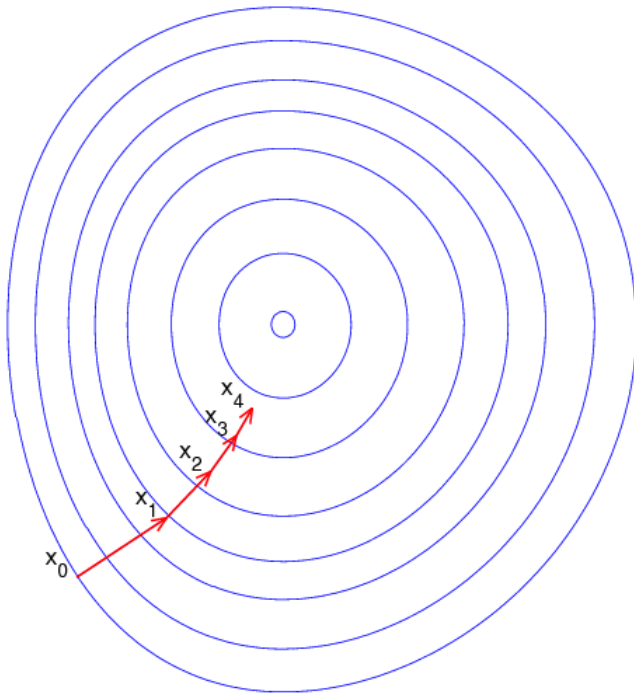
Algorithmes d'optimisation

- Il existe deux grandes catégories : algorithmes locaux et globaux
- Un algorithme local cherche le minimum de la fonction objective autour de ses valeurs précédentes. Il dépend du point de départ (paramètre initial)
- Un algorithme global cherche le minimum de la fonction objective « partout » et ne dépend pas du point de départ (s'il fonctionne bien)
- Algorithme d'optimisation = cherche le minimum d'une fonction

Algorithmes d'optimisation

Exemples d'algorithmes locaux :

- Algorithme de la plus forte pente (gradient)
- Gauss-Newton
- Levenberg-Marquardt

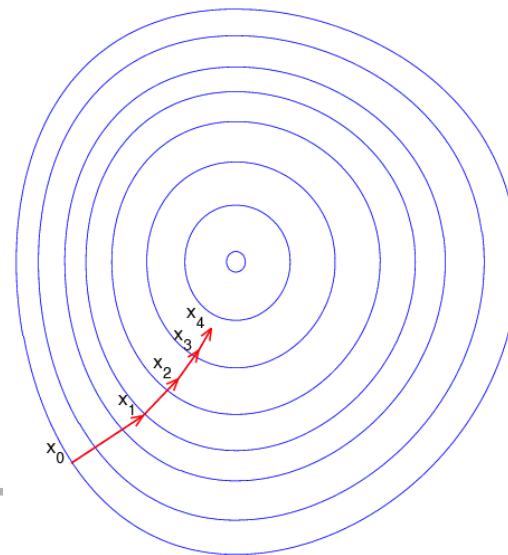


(c) 2006 P. A. Simionescu

Algorithmes d'optimisation

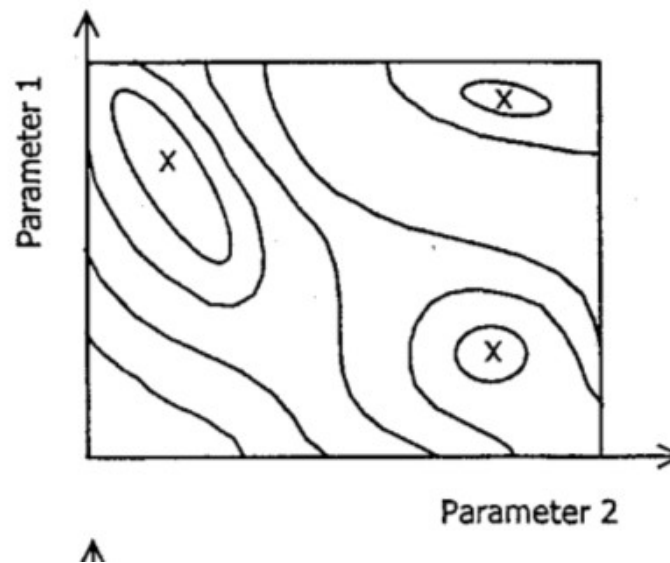
Algorithme de la plus forte pente (gradient) :

- On démarre en un point aléatoire de l'espace paramétrique (ici, en 2 dimensions)
- On calcule le gradient de la fonction (càd la pente) en ce point
- Puis on se déplace dans le sens de la plus forte pente par un petit pas.



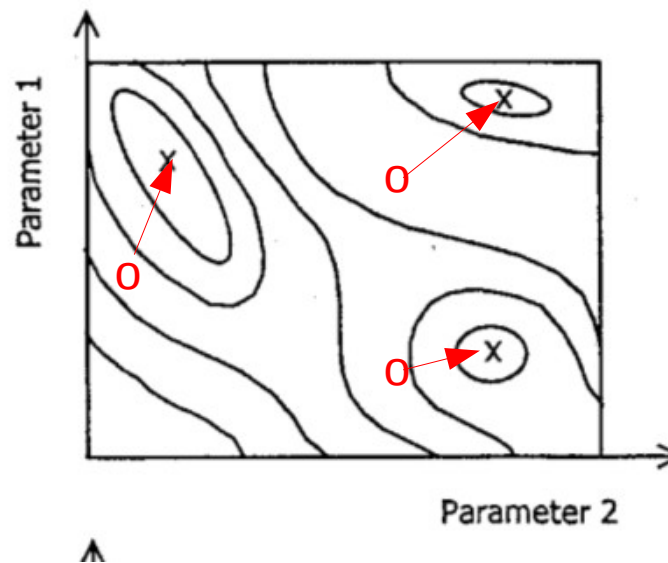
Algorithmes d'optimisation

- Algorithme d'optimisation = cherche le minimum d'une fonction
- Un algorithme local peut se « tromper » et tomber dans un minimum local, en fonction de son point de départ.



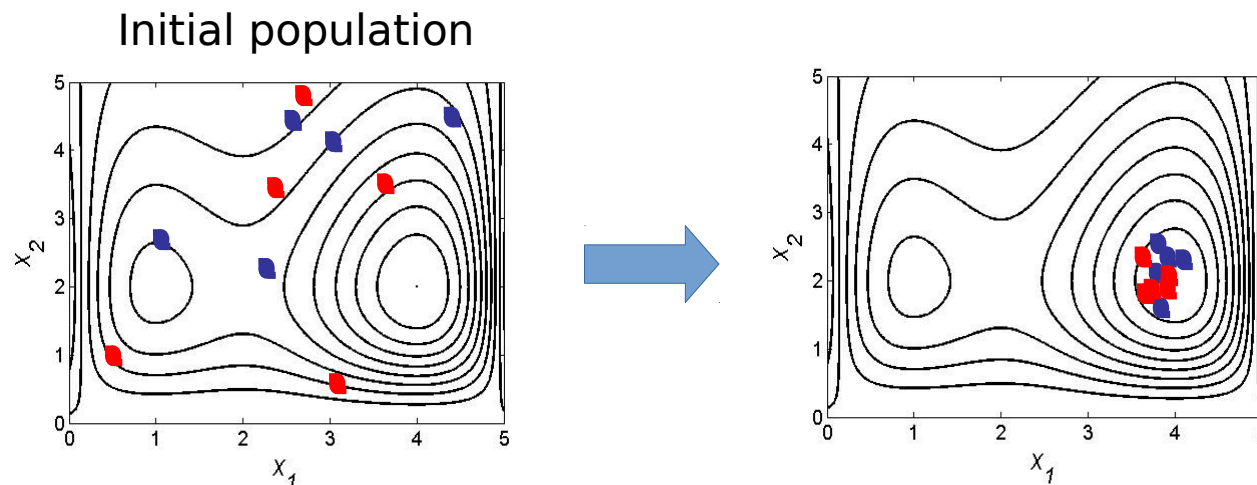
Algorithmes d'optimisation

- Algorithme d'optimisation = cherche le minimum d'une fonction
- Un algorithme local peut se « tromper » et tomber dans un minimum local, en fonction de son point de départ.



Algorithmes d'optimisation

- Algorithme global : Il cherche globalement dans l'ensemble de l'espace paramétrique. Habituellement, il démarre avec plusieurs points de départ différents et cherche la meilleure pente.
- Avantages par rapport aux algorithmes locaux : il risque moins de tomber dans un minimum local
- Inconvénient : demande beaucoup plus de ressources de calculs.



Algorithmes d'optimisation

Progression d'un algorithme global sur une fonction objective avec plusieurs minima :

[video]

Optimisation bayésienne

- Jusqu'à présent, les paramètres calibrés n'ont pas d'intervalles de confiance ou une autre estimation de leur incertitude.
- Or, il est possible d'estimer l'incertitude avec lesquelles les paramètres sont optimisés avec des méthodes bayésiennes de Markov Chain Monte Carlo (MCMC)
- Au lieu d'obtenir une valeur de paramètre, une **distribution** des valeurs du paramètre est obtenue.
- Cette distribution donne une estimation de **l'incertitude** du paramètre

Optimisation bayésienne

Optimisation « classique »

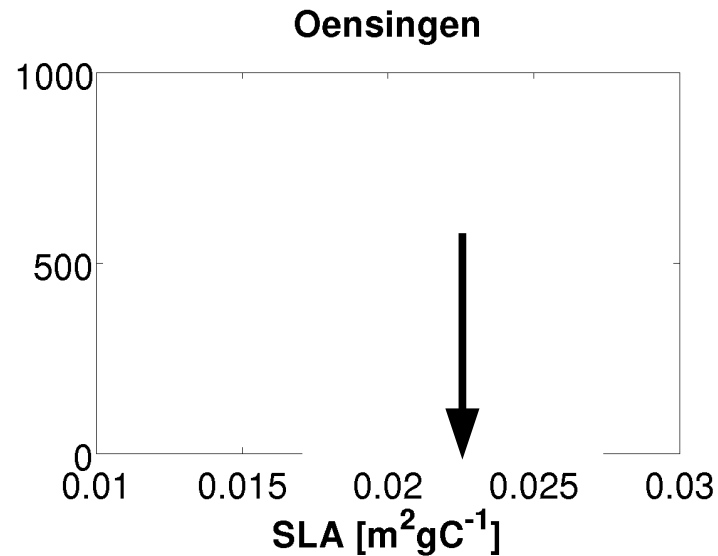
VS

Optimisation bayésienne

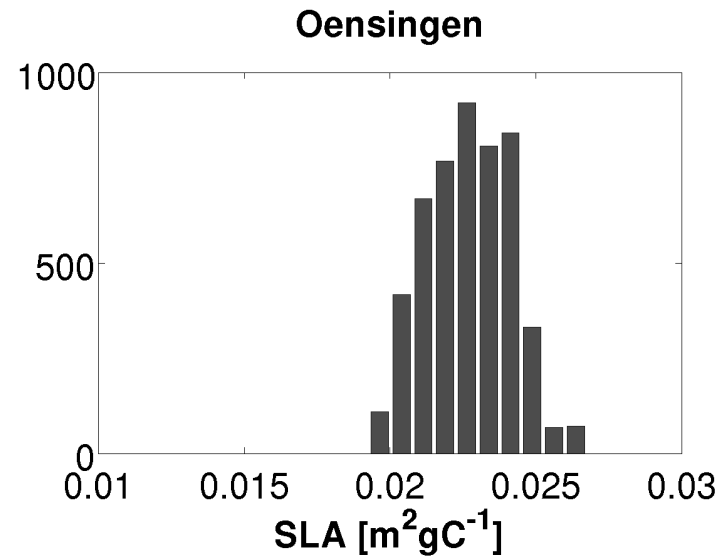
Valeur unique

VS

Distribution complète



Intervalle initial de paramètres

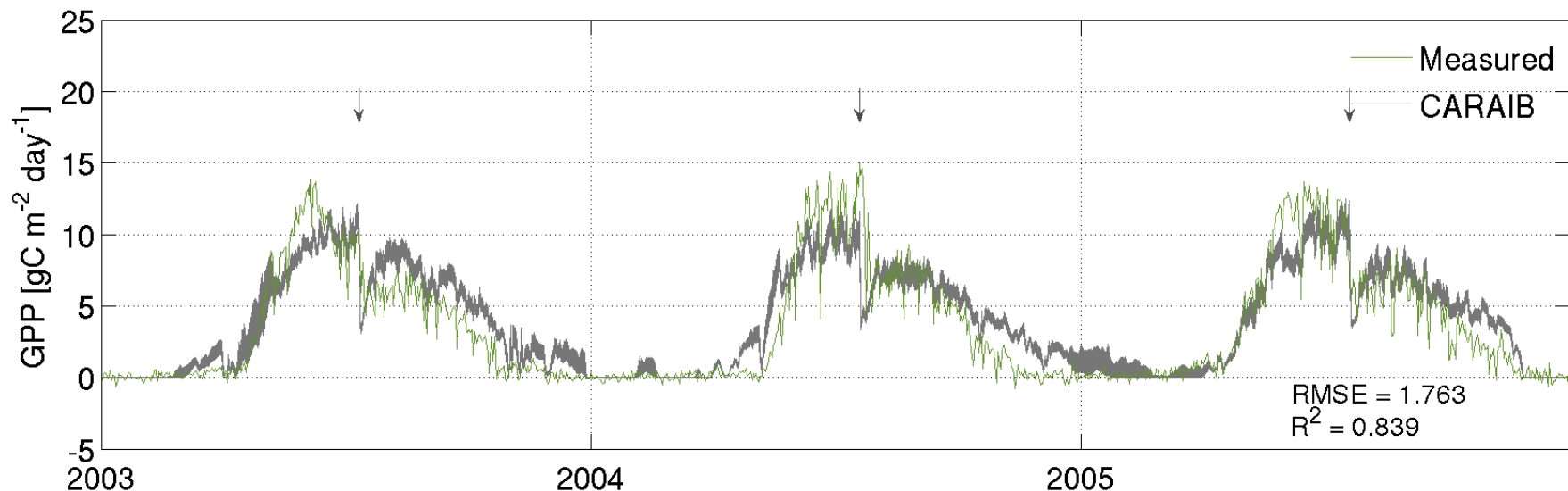


Intervalle initial de paramètres

Optimisation bayésienne

- Les sorties (optimisées) du modèle ne sont donc plus uniques mais il existe un ensemble de données, ce qui permet d'évaluer l'incertitude sur la prédiction du modèle

Measured VS modeled gross primary productivity (GPP), after calibration, Monte-Bondone



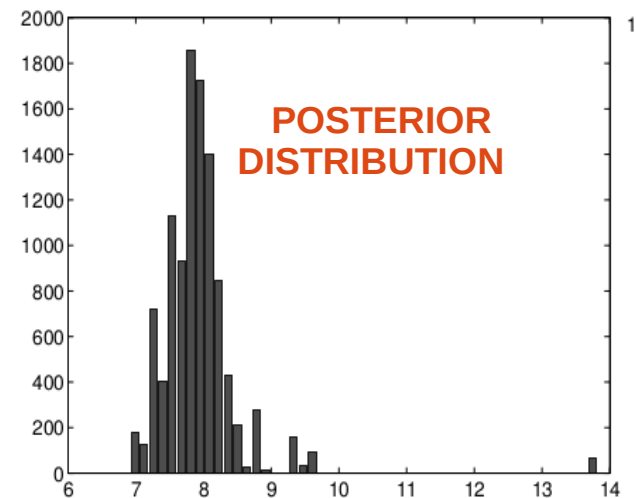
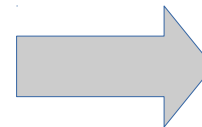
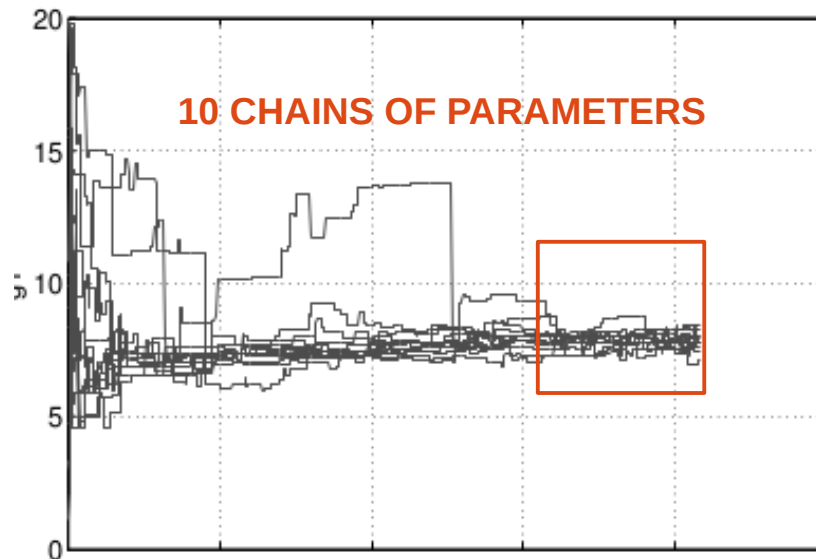
Optimisation bayésienne

Mais de quoi dépend l'incertitude des paramètres et des sorties optimisés du modèle ?

- Erreurs du modèle (= représentation partielle de la réalité)
- Erreurs des observations qui servent à optimiser les paramètres du modèle
- L'incertitude sur le paramètre sera aussi plus grande si celui-ci est peu sensible dans le modèle

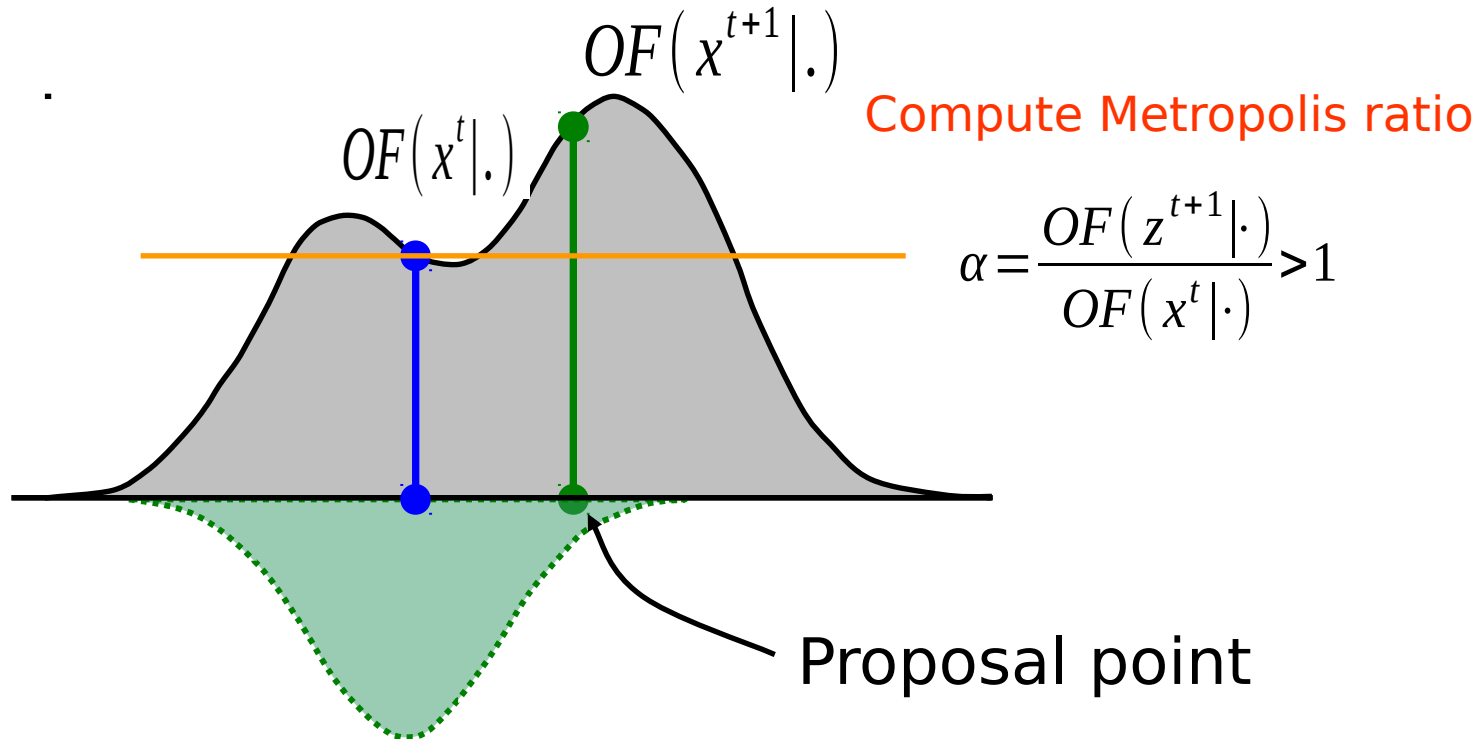
Optimisation bayésienne

Markov Chain Monte Carlo et algorithme de Metropolis-Hastings



Optimisation bayésienne

- Comment ça marche ?
- Algorithme de Metropolis-Hastings
 - 1^{er} cas : Si le nouveau point a une plus grande valeur, la chaîne bouge vers ce point

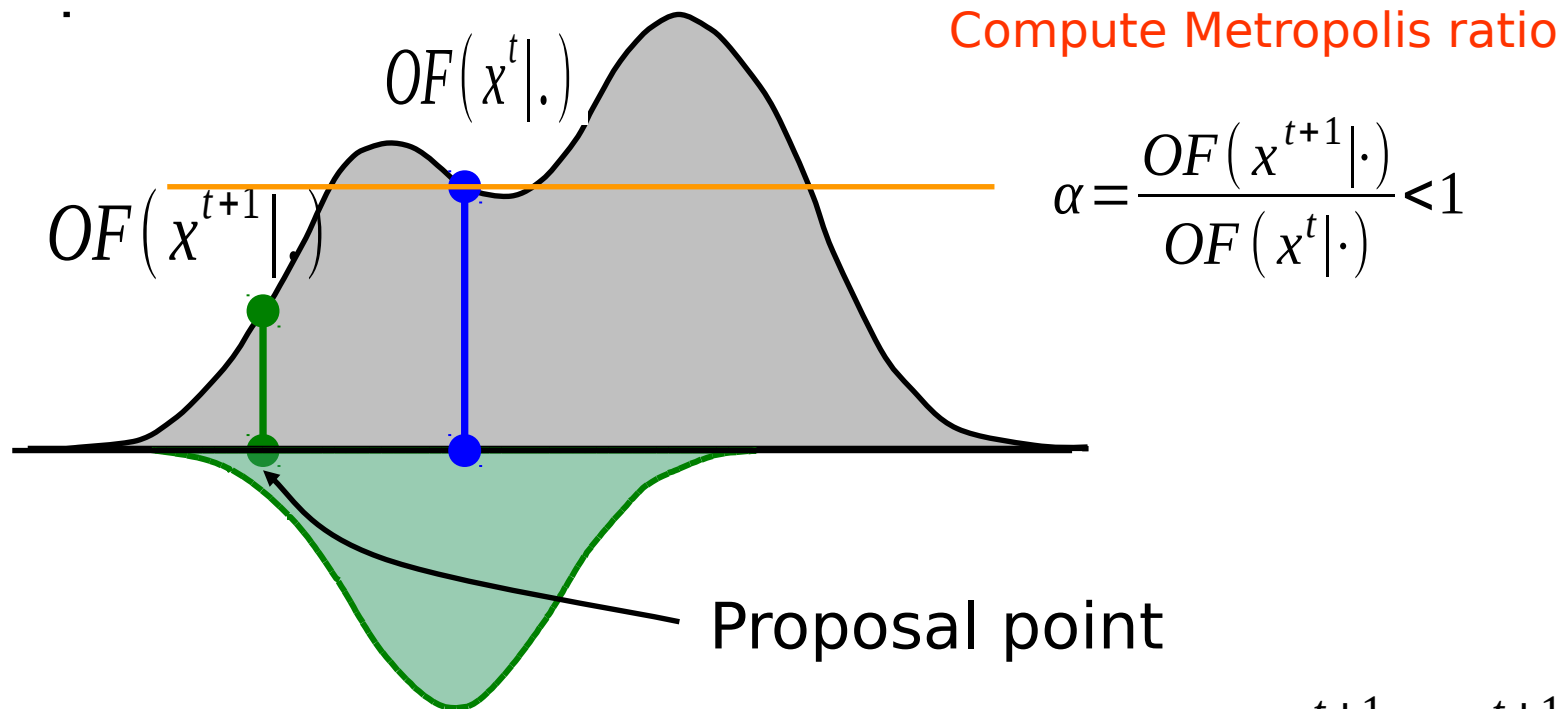


ACCEPT AND MOVE CHAIN TO NEW LOCATION

➔ $x^{t+1} = x^{t+1}$

Calibration d'un modèle complexe

- 2^{eme} cas : Si le nouveau point a une plus petite valeur, la chaine a 50 % de chances de bouger vers ce point et ce en fonction de la valeur de la fonction objective en ce point



ACCEPT IF $\alpha >$ uniform draw between 0 and 1 $\longrightarrow x^{t+1} = x^{t+1}$

OTHERWISE $\longrightarrow x^{t+1} = x^t$