

GML6402A : Géostatistique

Cours 10 : Problèmes inverses



**POLYTECHNIQUE
MONTREAL**

UNIVERSITÉ
D'INGÉNIERIE

Objectifs

- Comprendre la nature des problèmes inverses (déterministe et stochastique);
- Formuler un problème inverse en géoscience à partir des données observées, en définissant les inconnues, les équations directes, et les conditions aux limites;
- Utiliser des algorithmes pour la résolution de problèmes inverses;
 - MCMC-Metropolis Hasting, Déformation Graduelle, Échantillonnage Spatial Itératif, Filtre de Kalman d'Ensemble, FFTMA-SA, S-STBM, Point Pilote Voyageur, etc.
- Analyser la sensibilité et l'incertitudes des solutions.



Plan du cours

1. Introduction aux problèmes inverses;
2. Problèmes inverses déterministes;
3. Problèmes inverses stochastiques;
4. Optimisation stochastique;
5. Filtre de Kalman d'ensemble.

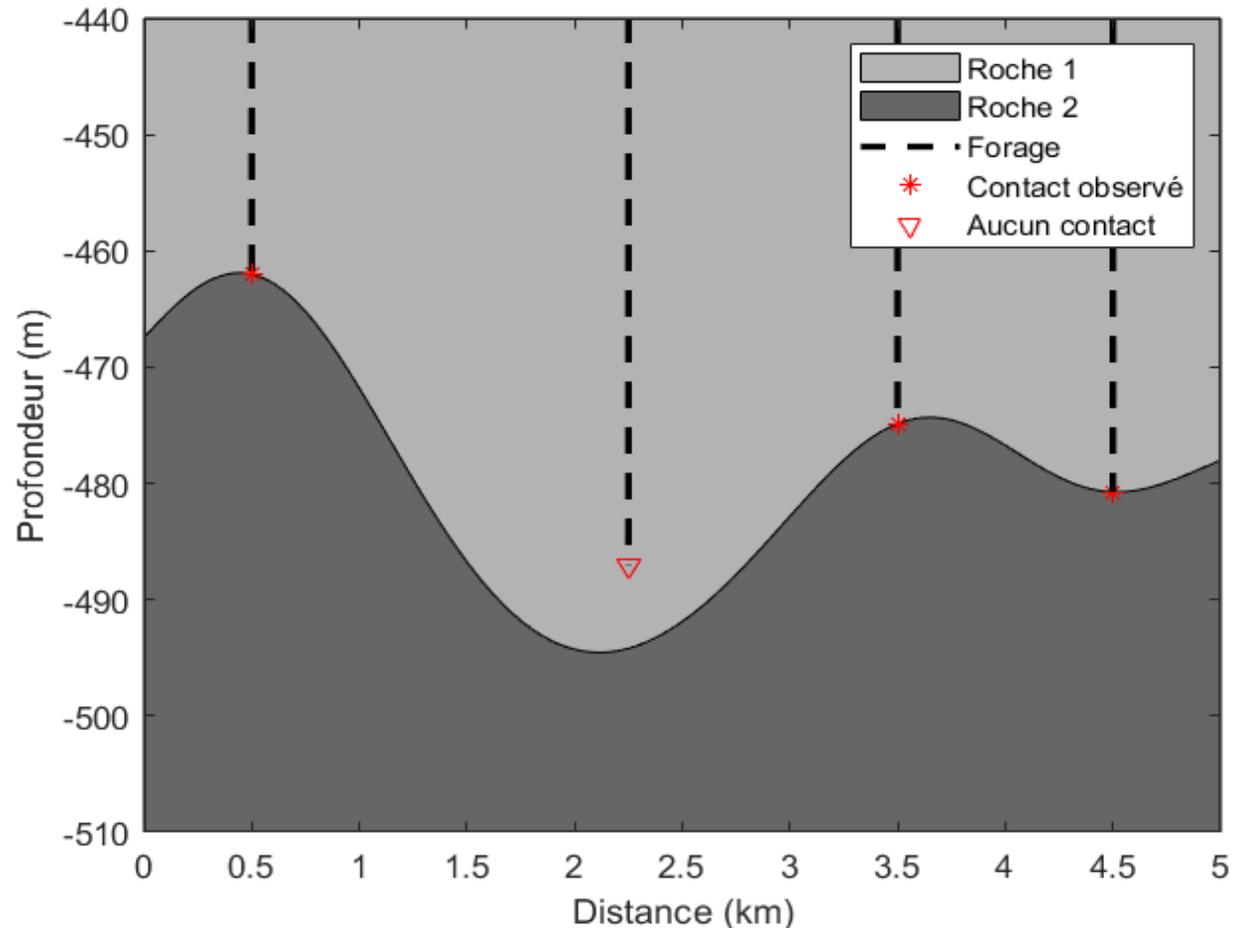


1. Introduction aux problèmes inverses

Situation :

Supposons que l'on cherche à estimer le contact géologique entre deux unités distinctes

- Contact observé dans 3 forages
- Un forage n'a pas observé le contact



1. Introduction aux problèmes inverses

Situation :

Supposons que l'on cherche à estimer le contact géologique entre deux unités distinctes

- Contact observé dans 3 forages
- Un forage n'a pas observé le contact

Solutions possibles :

Géostatistique non-linéaire

- Krigeage d'indicatrices
- Simulation géostatistiques conditionnelles
 - Échantillonnage
 - Recherche stochastique



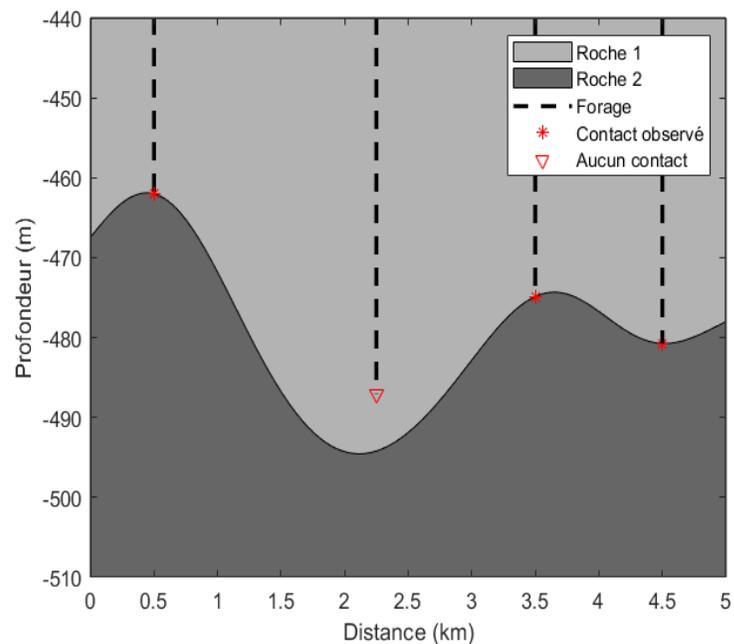
1. Introduction aux problèmes inverses

Situation :

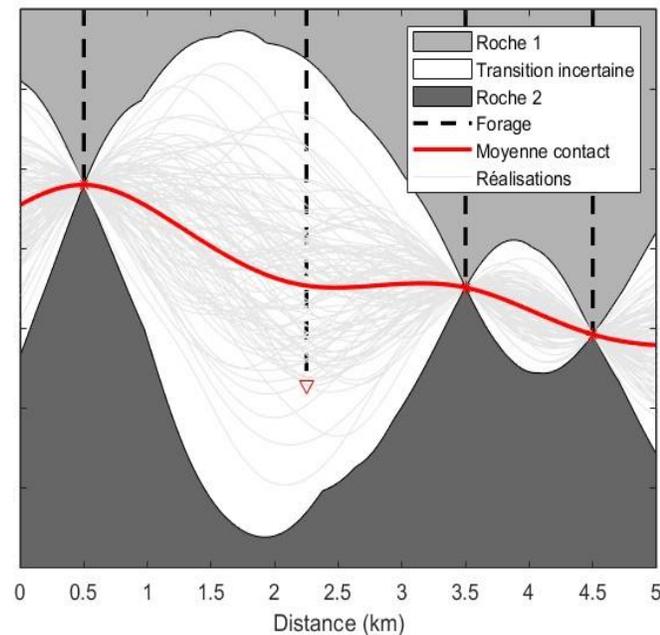
Supposons que l'on cherche à estimer le contact géologique entre deux unités distinctes

- Contact observé dans 3 forages
- Un forage n'a pas observé le contact

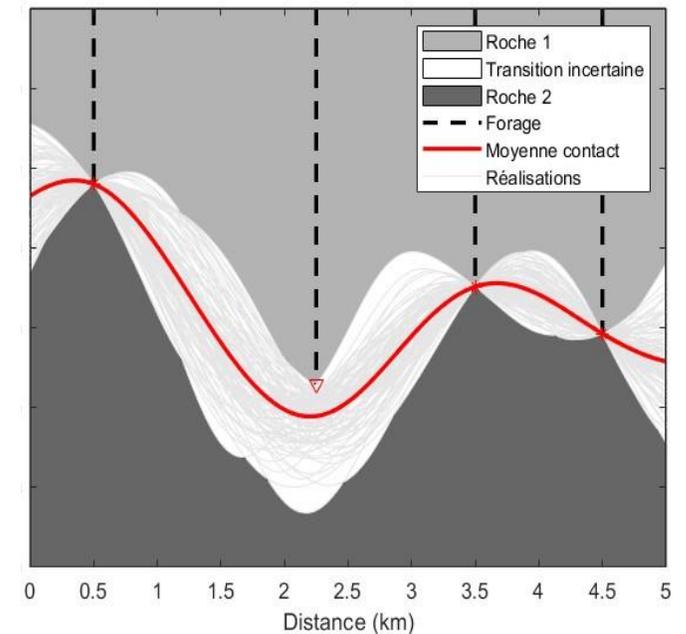
Référence



Réalisations non calées



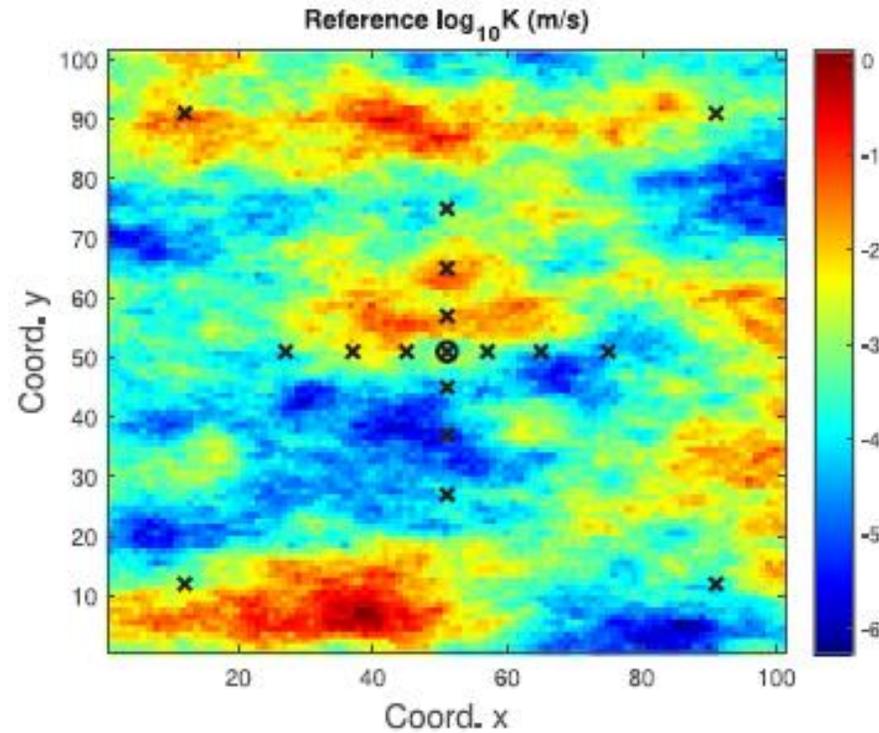
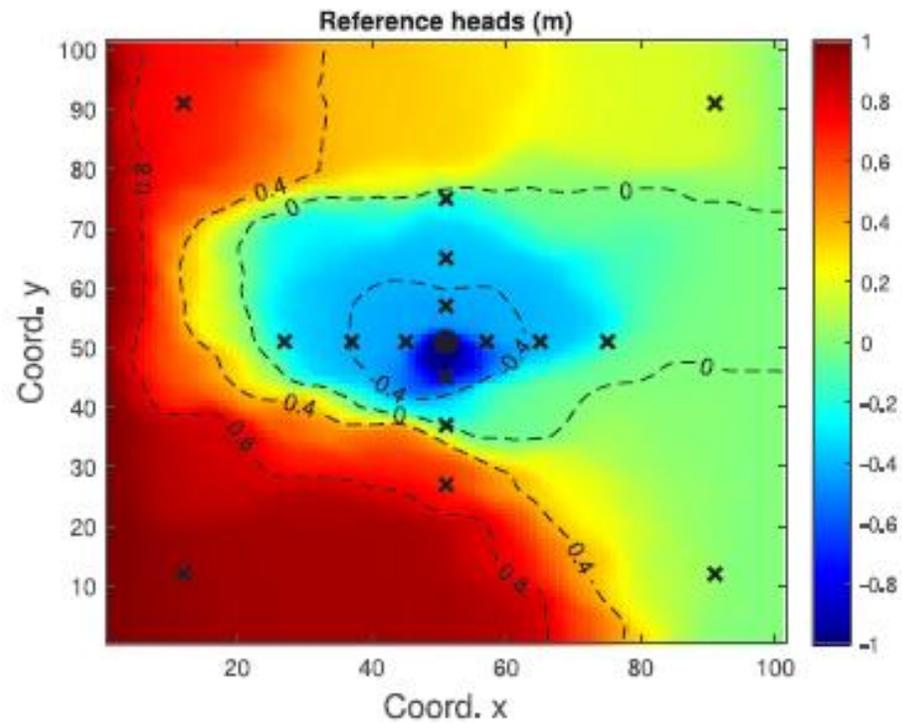
Réalisations calées



1. Introduction aux problèmes inverses

Contexte : Supposons que l'on cherche à estimer la conductivité hydraulique d'un aquifère confinée:

- Aucune observation de la conductivité hydraulique
- 16 données de rabattement en régime permanent acquies par un essai de pompage (Conditions frontières : W-1m, E-0m, N-NoFlow, S-NoFlow, Pompage : 120L/min)



1. Introduction aux problèmes inverses

Contexte :

Deux questions :

1- Comment déterminer les conductivités hydrauliques hétérogènes à partir des données de rabattements ? **Par modélisation inverse**

- Échantillonnage: Devient rapidement complexe et très lent.
- Recherche stochastique: Simple avec la bonne méthodologie!

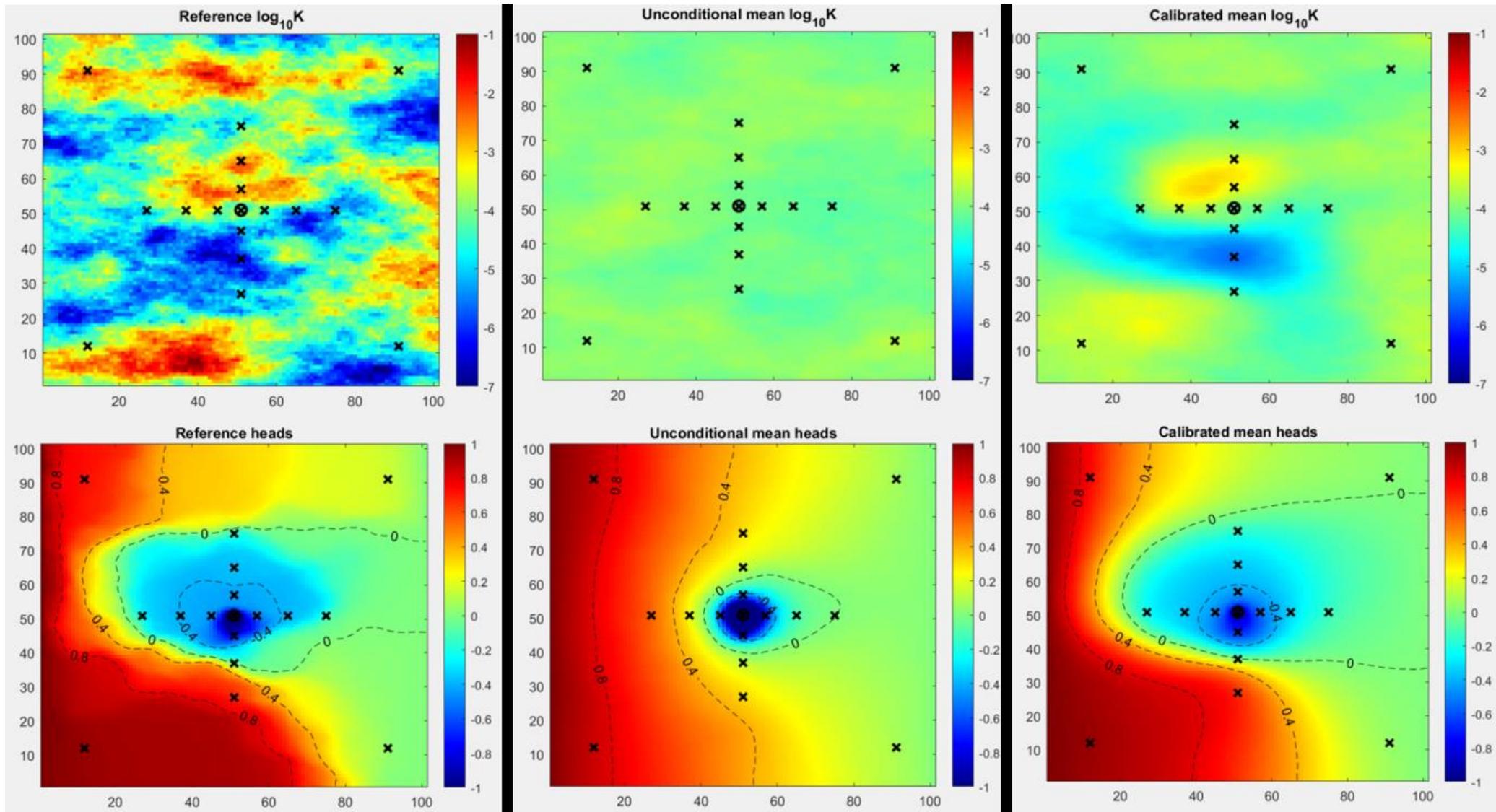
2- Est-ce que les données secondaires apportent de l'information supplémentaire lors de l'estimation, simulations, des conductivités hydrauliques?

Un peu comme le cokrigage. Oui, si l'information est intéressante.



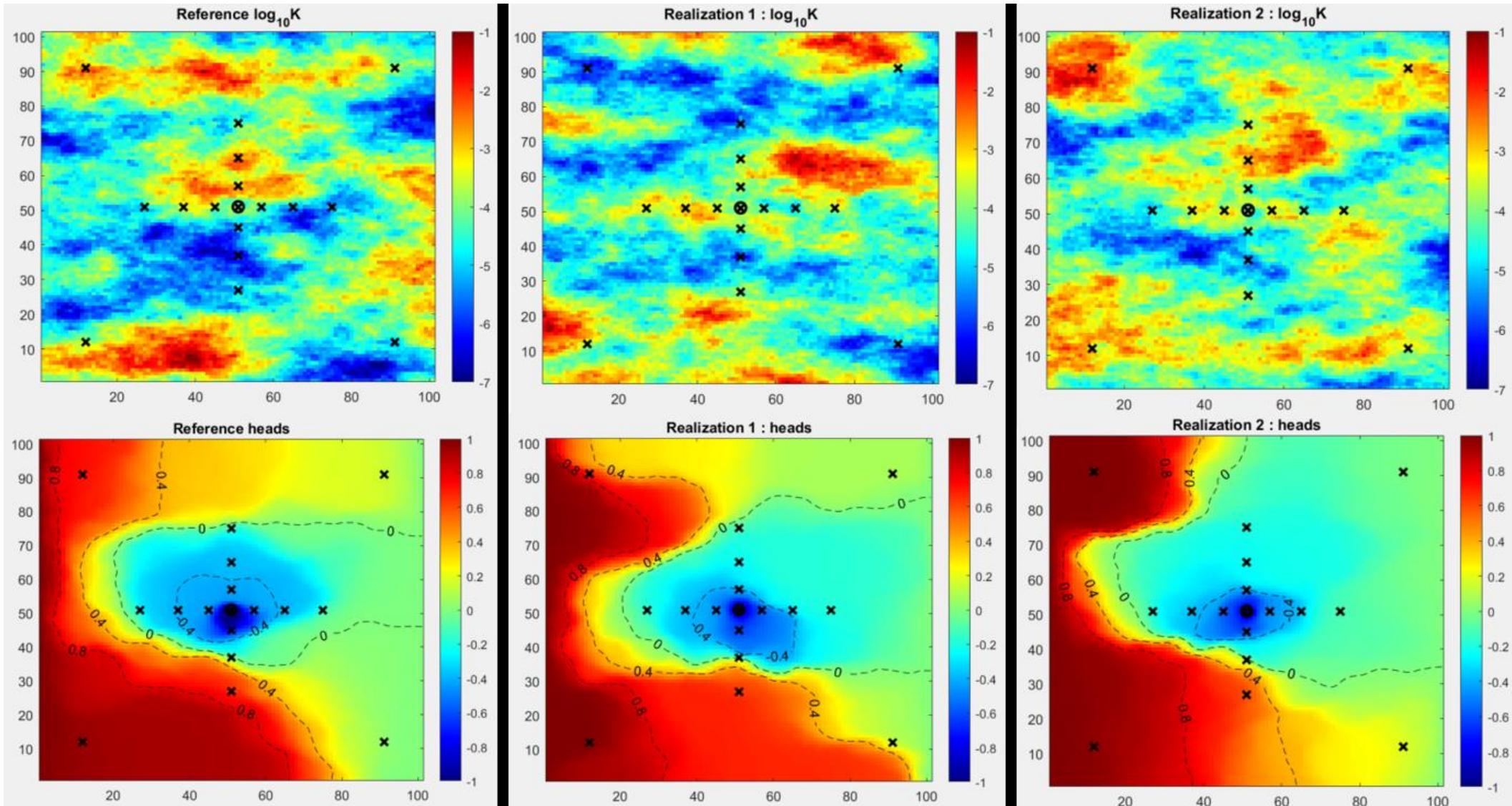
1. Introduction aux problèmes inverses

Contexte :



1. Introduction aux problèmes inverses

Contexte :

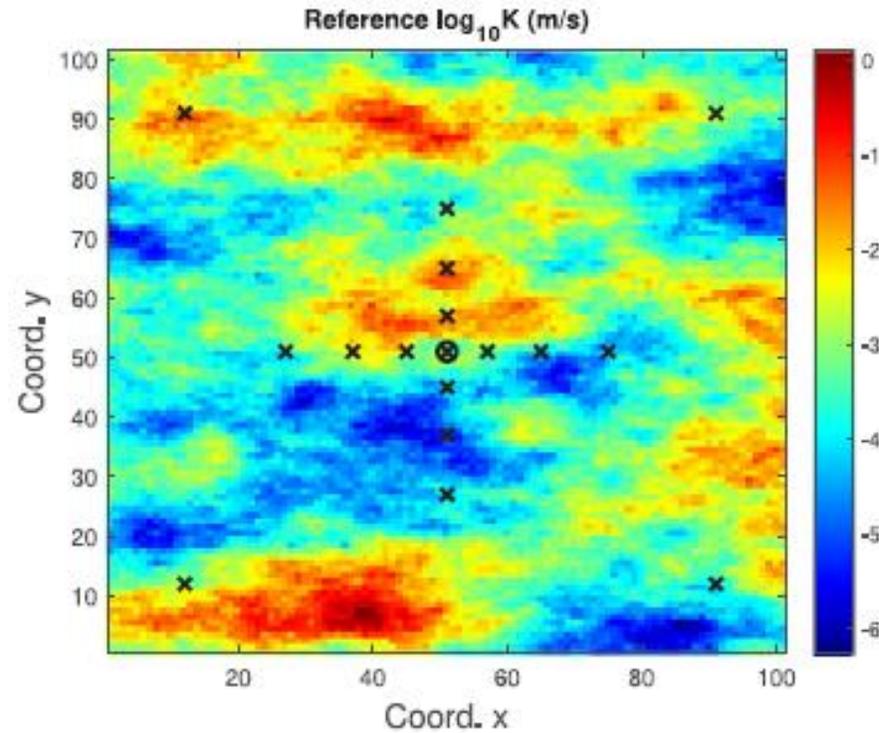
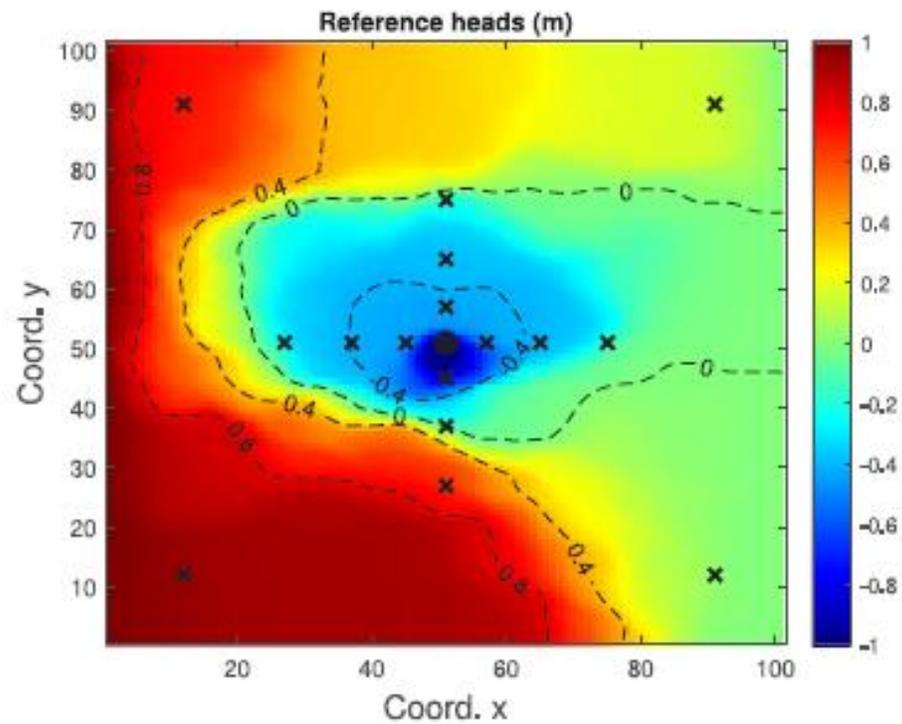


1. Introduction aux problèmes inverses

Contexte : Un problème inverse implique des **données observées** (d_{obs}) et un **modèle géologique** (m):

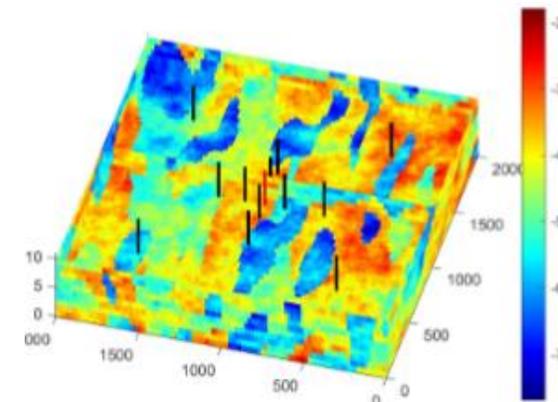
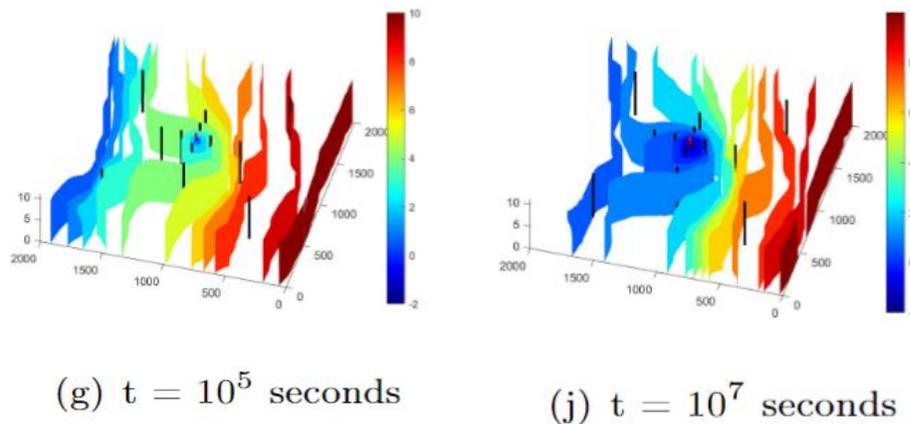
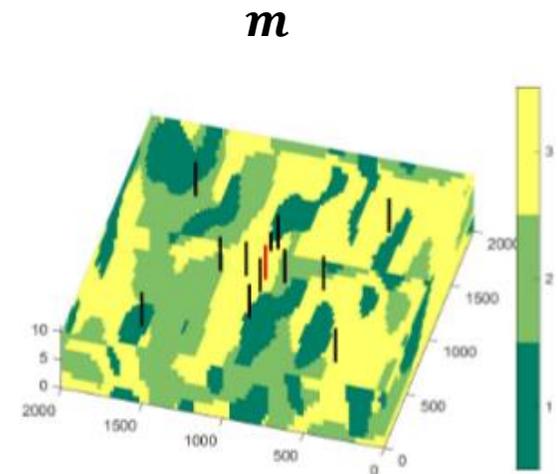
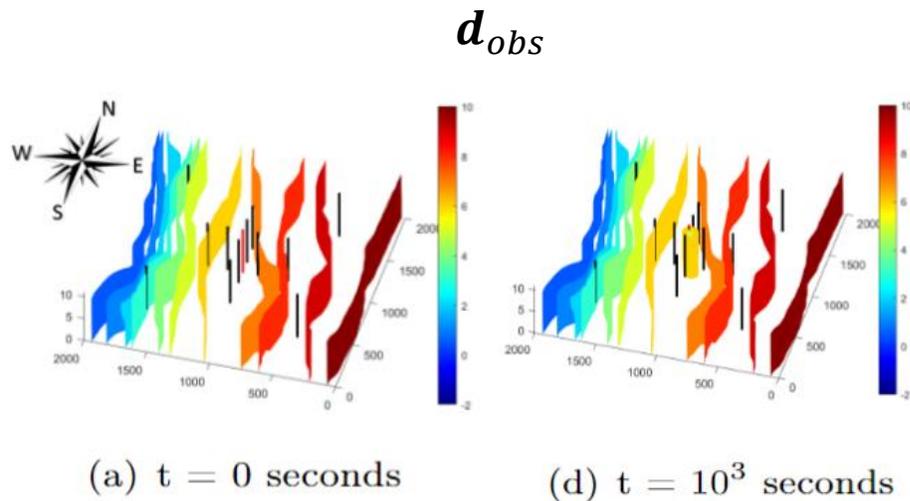
d_{obs}

m



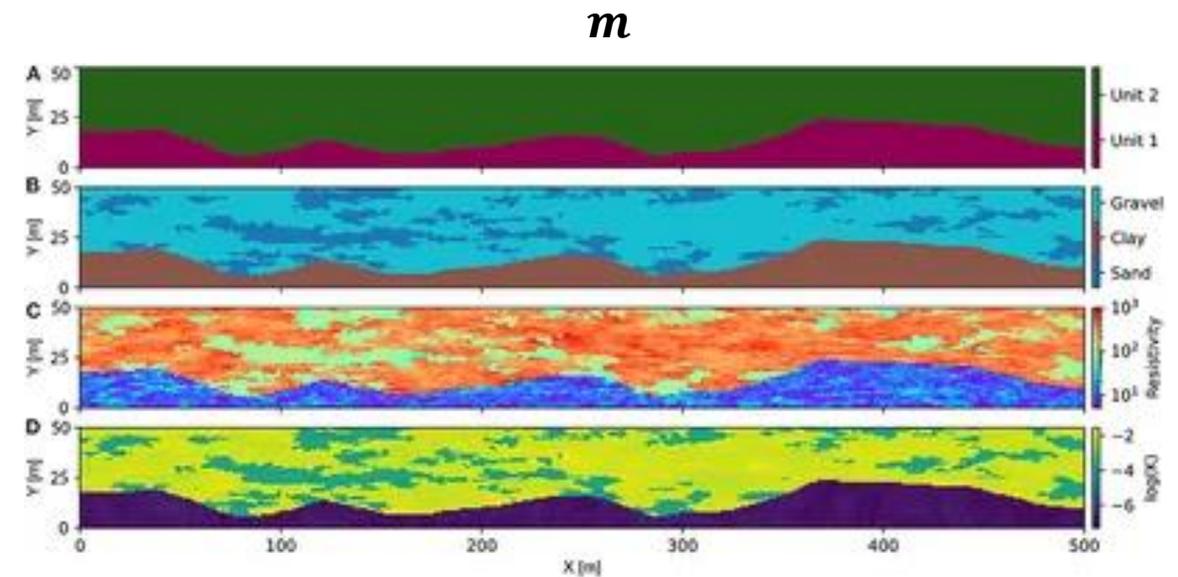
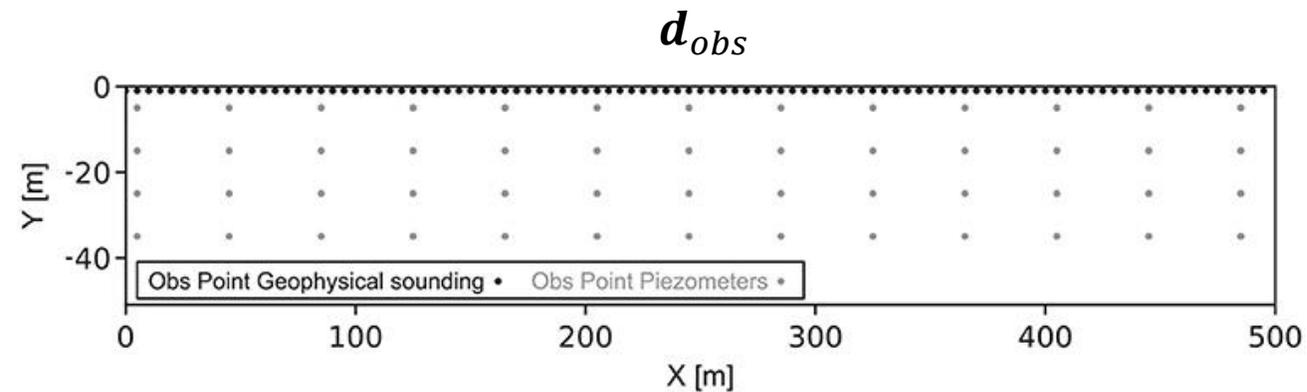
1. Introduction aux problèmes inverses

Résumé : Un problème inverse implique des **données observées** (d_{obs}) et un **modèle géologique** (m):



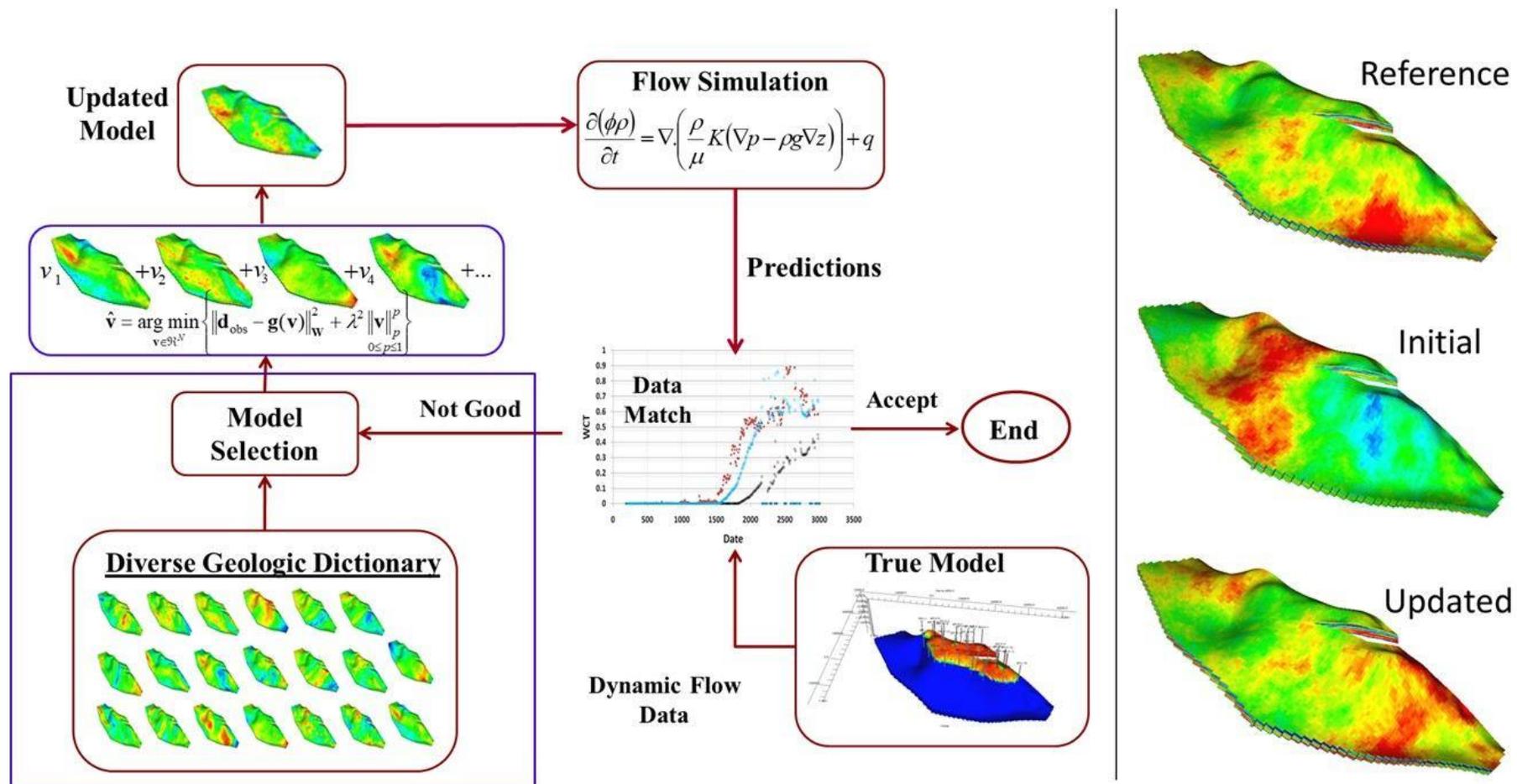
1. Introduction aux problèmes inverses

Contexte : Un problème inverse implique des **données observées** (d_{obs}) et un **modèle géologique** (m):



1. Introduction aux problèmes inverses

Contexte : Un problème inverse implique des **données observées** (d_{obs}) et un **modèle géologique** (m):

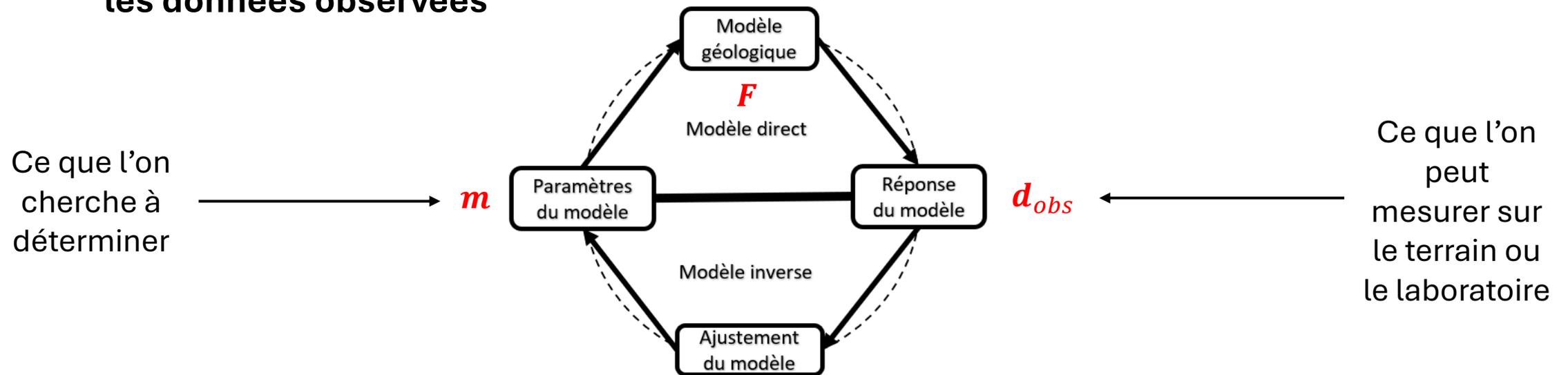


1. Introduction aux problèmes inverses

Contexte :

Un problème inverse est l'inverse du problème direct :

au lieu de déterminer les données produites par des paramètres de modèle spécifiques, **nous cherchons à identifier les paramètres du modèle qui génèrent les données observées**



Mathématiquement, on cherche les paramètres du modèle (m), tel que lorsque l'on applique le modèle direct (F), on retrouve les données observées (d_{obs}) :

$$d_{obs} = F(m)$$

2. Problèmes inverses déterministes

Problème linéaire :

Dans le cas d'un problème linéaire, les données simulées (\mathbf{d}_{sim}) peuvent être calculé par un modèle candidat (\mathbf{m}^{cad}) à partir d'une matrice \mathbf{F} :

$$\mathbf{F}\mathbf{m}^{cad} = \mathbf{d}_{sim} \rightarrow \text{Cas linéaire}$$

La matrice \mathbf{F} décrit la physique et la géométrie sous-jacentes du phénomène géologique (par exemple, en utilisant la méthode des éléments finis ou par une relation physique connue)



2. Problèmes inverses déterministes

Un exemple élémentaire :

Supposons que l'on cherche à déterminer la masse locale (qui est liée à la densité) de la roche (\mathbf{m}) dans le sous-sol à partir de mesure de l'accélération gravitationnelle locale (\mathbf{d}_{obs}) :

D'après notre compréhension de la gravité, nous savons que l'expression mathématique de la gravité est :

$$d = \frac{Gm}{r^2} = Fm, \quad F = G/r^2$$

Si nous avons trois observations, nous pouvons écrire un système d'équation :

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} \frac{G}{r_{11}^2} & \frac{G}{r_{12}^2} & \frac{G}{r_{13}^2} \\ \frac{G}{r_{21}^2} & \frac{G}{r_{22}^2} & \frac{G}{r_{23}^2} \\ \frac{G}{r_{31}^2} & \frac{G}{r_{32}^2} & \frac{G}{r_{33}^2} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}$$

2. Problèmes inverses déterministes

Un exemple élémentaire :

Un système avec trois équations et trois inconnues est une situation très spécifique :

- L'exemple est conçu pour aboutir à une réponse simple.
- En général, le nombre de données et d'inconnues est différent

Les problèmes inverses sont généralement mal posés lorsqu'ils sont formulés de manière déterministe :

1. La solution du problème inverse sera indéterminée (**solution non-unique**)
2. Les données observées peuvent être remplies de bruit (**solution non-existante**)
3. Le système d'équation n'est pas stable (**instabilité numérique**)

Lorsque le problème inverse est formulé de **manière probabiliste** (Tarantola, 2005), ces questions ne sont plus pertinentes.

2. Problèmes inverses déterministes

Un exemple élémentaire :

NOTE:

Les inversions déterministes ne génèrent **qu'un seul modèle** de propriété terrestre.

Si l'**inversion stochastique** génère **uniquement le modèle de propriété le mieux adapté** qu'elle trouve, il s'agit alors d'une **inversion déterministe**.



2. Problèmes inverses déterministes

Formulation Bayésienne :

Définitions : L'objectif est **d'obtenir la distribution de probabilité des valeurs des paramètres** qui soit compatible avec les variables d'état observées, la physique connue et les informations préalables sur les paramètres.

Théorème de Bayes :

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})} \cong p(\mathbf{d}|\mathbf{m})p(\mathbf{m})$$

$p(\mathbf{d}|\mathbf{m})$: Distribution de vraisemblance (*Likelihood*) – Résume les propriétés statistiques des résidus d'erreur entre les données observées et simulées – Il s'agit d'une distribution multivariable pouvant être évaluée en $\mathbf{d} = \mathbf{d}_{obs}$

$p(\mathbf{m})$: Distribution préalable (*prior*) – Avant même d'acquérir des données, nous avons une idée générale du système et des probabilités qui lui sont associées – Il s'agit d'une distribution multivariable ;

$p(\mathbf{m}|\mathbf{d})$: Distribution postérieure (*posterior*) – On utilise cette distribution pour déterminer la solution du problème inverse – Il s'agit d'une distribution multivariable.

$p(\mathbf{d})$: Important pour la sélection et le calcul de la moyenne du modèle, mais il peut être négligé lorsque le modèle est fixe.

2. Problèmes inverses déterministes

Vraisemblance : Formulations simplifiées

La fonction de vraisemblance est une distribution multivariée (n'est pas une fonction objectif)

Qu'elle distribution multivariée ? On considère deux simplifications

1) Le modèle direct

- $d = F(m)$
- Hypothèse du modèle (conditions frontières, équations d'état, géologie...)

2) Les erreurs

- Erreurs des données (généralement connu d'une certaine manière)
- Erreurs du modèle direct F (rarement connue, notre modèle n'est jamais parfait)
 - Les méthodes de résolution des équations d'état (éléments finis, ...)



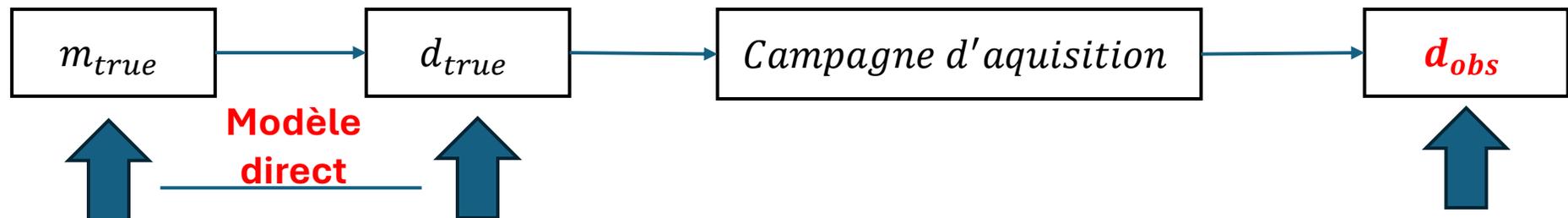
2. Problèmes inverses déterministes

Incertitudes des données

Idée générale : On **n'observe pas les données réelles**, comme définit par notre modèle direct (situation idéal), mais on **observe un phénomène différent**

Cette idée générale implique 'un modèle direct'. Les erreurs peuvent dépendre aussi du modèle direct, pas juste des erreurs de mesures qui sont indépendantes de la réalité.

En réalité on a:



2. Problèmes inverses déterministes

Rappel : Distribution normale multivariante

Dans le cas où Σ , la **matrice de covariance**, est définie positive, donc inversible, la loi normale multivariante admet la densité de probabilité suivante:

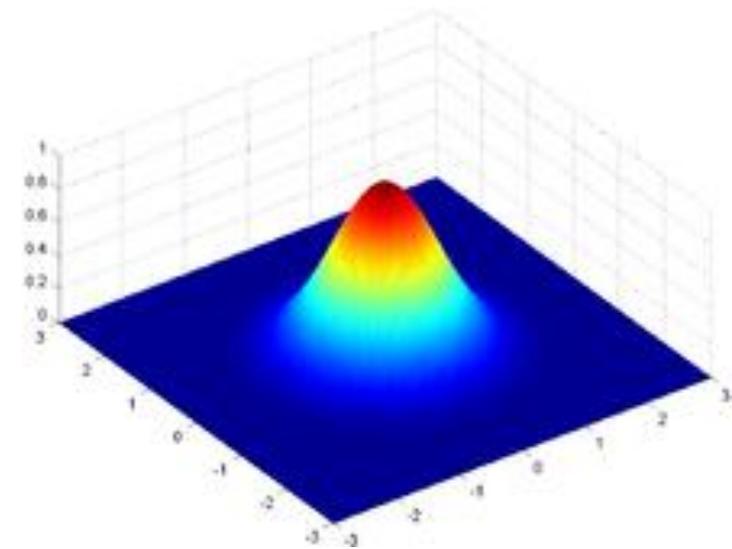
$$f_{\mu, \Sigma}(x) = \frac{1}{(2\pi)^{N/2} \det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu)\right)$$

Σ : la matrice de variance-covariance

μ : vecteur $\in \mathbb{R}^N$ représentant la moyenne

x : un vecteur $\in \mathbb{R}^N$

N : la dimension de la distribution



2. Problèmes inverses déterministes

Formulation probabiliste :

Distribution de vraisemblance : Résume les propriétés statistiques des résidus d'erreur entre les données observées et simulées

Le problème direct consiste à simuler la réponse \mathbf{d}_{sim} d'un modèle proposé \mathbf{m}^{cad} : $\mathbf{d}_{sim} = F(\mathbf{m}^{cad})$

Sous hypothèses gaussiennes, on peut écrire :

$$p(\mathbf{d}|\mathbf{m}) = \mathcal{L}(\mathbf{m}|\mathbf{d}) = \frac{1}{(2\pi)^{N/2} \det(\mathbf{C}_D)^{1/2}} \exp\left(-\frac{1}{2} (F(\mathbf{m}^{cad}) - \mathbf{d}_{obs})^T \mathbf{C}_D^{-1} (F(\mathbf{m}^{cad}) - \mathbf{d}_{obs})\right)$$

En assumant que $\mathbf{d}_{obs} = \mathbf{d}_{true} + \varepsilon$ et que les erreurs ne dépendent pas de la réalité, alors \mathbf{C}_D est la matrice de covariance des erreurs ($\mathbf{C}_D = \mathbf{C}_\varepsilon$)

2. Problèmes inverses déterministes

Formulation probabiliste :

Distribution préalable : Avant même d'acquérir des données, nous avons une idée générale du système et des probabilités qui lui sont associées

Le modèle géologique suit une distribution multigaussienne de moyenne \mathbf{m}_{prior} et de covariance \mathbf{C}_m

On peut donc écrire :

$$p(\mathbf{m}) = \frac{1}{(2\pi)^{M/2} \det(\mathbf{C}_m)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{m} - \mathbf{m}_{prior})^T \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{prior})\right)$$

Si le modèle géologique est représenté par un champ gaussien de moyenne nul, alors $\mathbf{m}_{prior} = 0$



2. Problèmes inverses déterministes

Formulation probabiliste :

Distribution postérieure (*posterior*) – On utilise cette distribution pour déterminer la solution du problème inverse

$$\begin{aligned} p(\mathbf{m}|\mathbf{d}) &\cong p(\mathbf{d}|\mathbf{m})p(\mathbf{m}) \\ &= \text{const.} * \exp\left(-\frac{1}{2}\left(F(\mathbf{m}^{cad}) - \mathbf{d}_{obs}\right)^T \mathbf{C}_D^{-1}\left(F(\mathbf{m}^{cad}) - \mathbf{d}_{obs}\right) + -\frac{1}{2}\left(\mathbf{m} - \mathbf{m}_{prior}\right)^T \mathbf{C}_m^{-1}\left(\mathbf{m} - \mathbf{m}_{prior}\right)\right) \end{aligned}$$



2. Problèmes inverses déterministes

Formulation probabiliste : Inversion linéaire

1. Le modèle direct est linéaire et exact : $d = F(m) \rightarrow d = Fm$
2. Les paramètres du modèle géologique sont multigaussienne de moyenne \mathbf{m}_{prior} et de covariance \mathbf{C}_m
3. Les erreurs sur les données sont multigaussienne de moyenne \mathbf{d}_{obs} et de covariance \mathbf{C}_d

On peut résoudre $p(\mathbf{m}|\mathbf{d})$ par la méthode des moindres carrés :

$$\mathbf{m}^* = \mathbf{m}_{prior} + \mathbf{C}_m \mathbf{F}^T (\mathbf{F} \mathbf{C}_m \mathbf{F}^T + \mathbf{C}_d)^{-1} (\mathbf{d}_{obs} - \mathbf{F} \mathbf{m}_{prior})$$

$$\mathbf{C}_m^* = \mathbf{C}_m - \mathbf{C}_m \mathbf{F}^T (\mathbf{F} \mathbf{C}_m \mathbf{F}^T + \mathbf{C}_d)^{-1} \mathbf{F} \mathbf{C}_m$$

On peut donc déterminer le modèle géologique (\mathbf{m}^*) le plus probable avec son niveau de précision (\mathbf{C}_m^*)



2. Problèmes inverses déterministes

Formulation probabiliste : Inversion linéaire

1. Prenons le négative du log de la distribution postérieur
2. Posons $\mathbf{C}_d = \mathbf{C}_m = \mathbf{I}$

On se retrouve avec :

$$\|\mathbf{F}\mathbf{m} - \mathbf{d}_{obs}\|_2^2 + \|\mathbf{m} - \mathbf{m}_{prior}\|_2^2$$

On peut ajouter un terme de régularisation (α) et obtenir un nouveau problème :

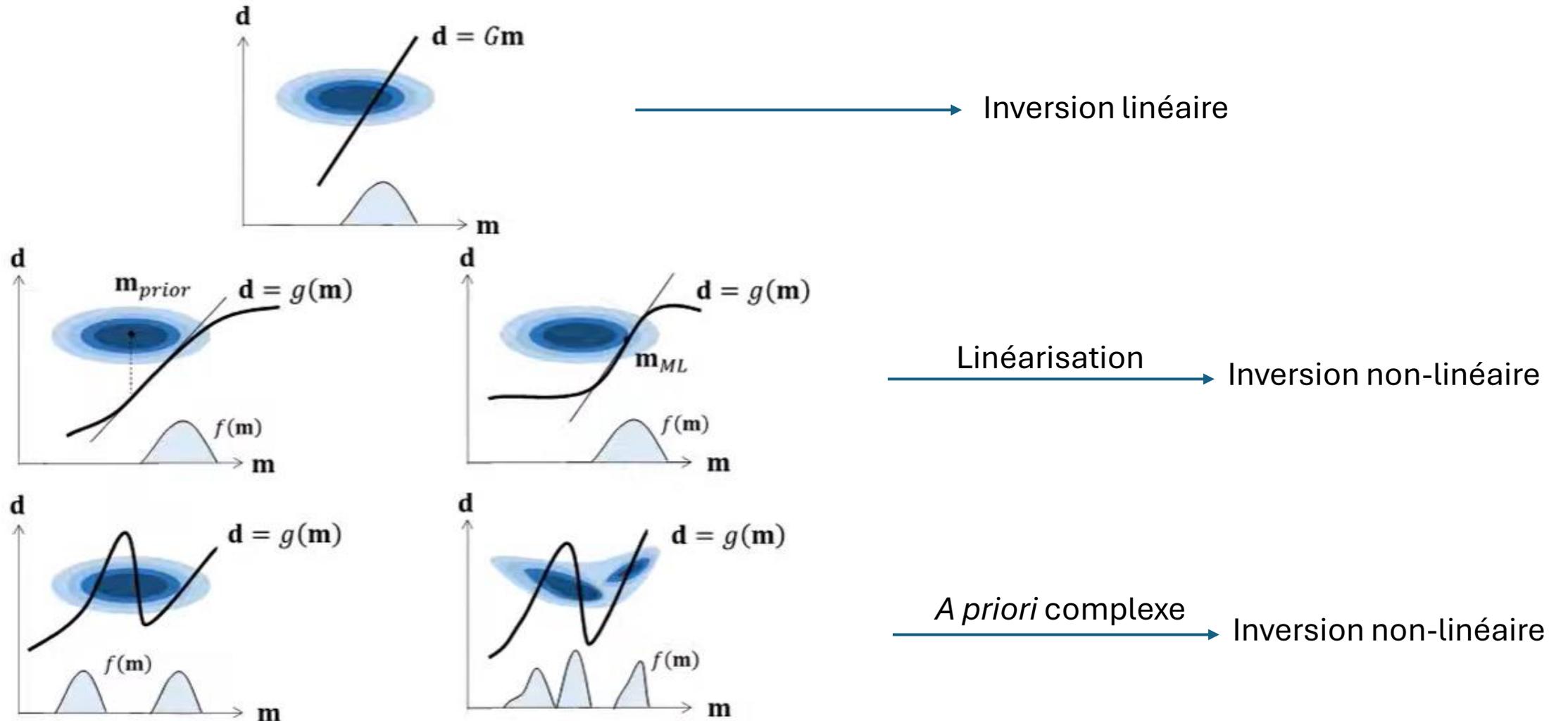
$$O(\mathbf{m}) = \|\mathbf{F}\mathbf{m} - \mathbf{d}_{obs}\|_2^2 + \alpha \|\mathbf{m} - \mathbf{m}_{prior}\|_2^2$$

L'inversion déterministique (forme linéaire) est :

$$\mathbf{m}^* = (\mathbf{F}^T \mathbf{F} + \alpha \mathbf{I})^{-1} (\mathbf{F}^T \mathbf{d}_{obs} + \alpha \mathbf{m}_{prior})$$

2. Problèmes inverses déterministes

Formulation probabiliste : Inversion linéaire

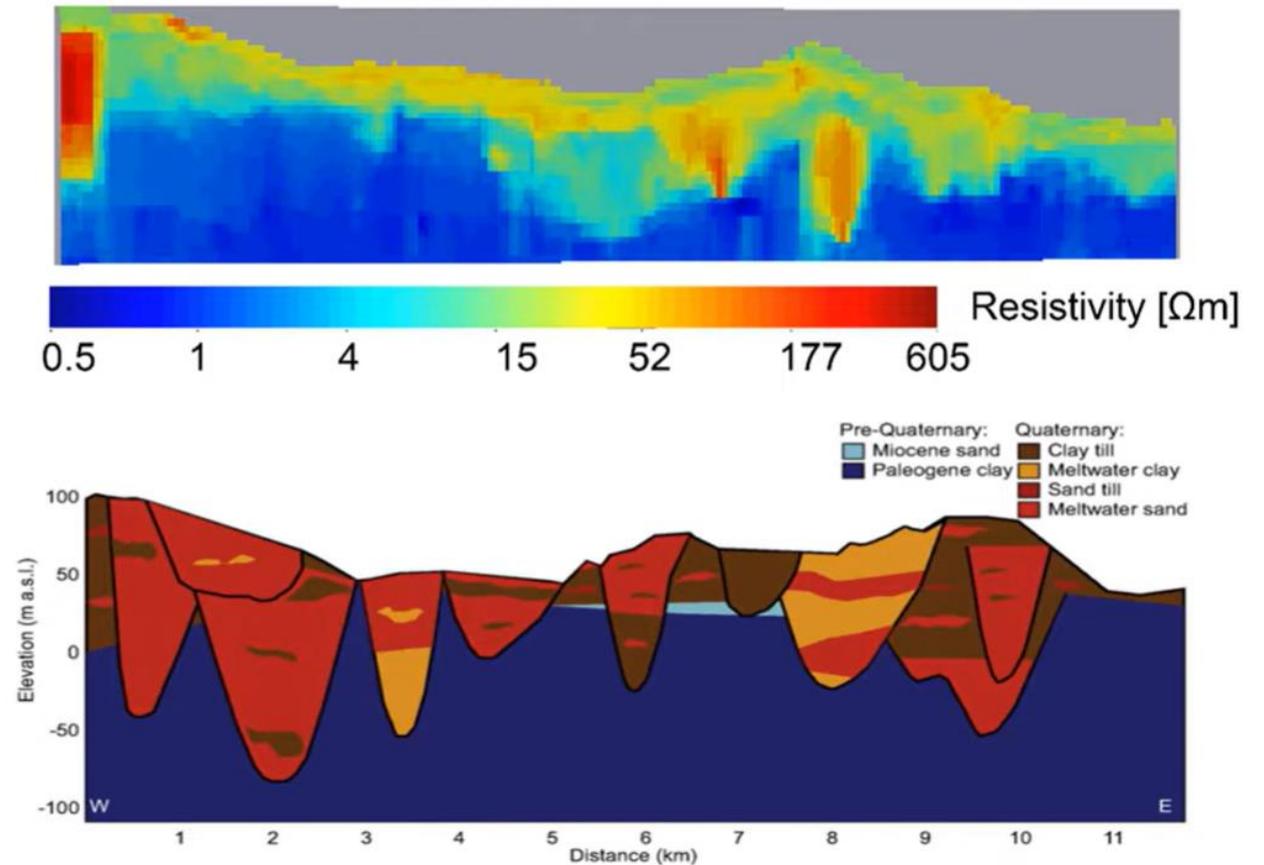


3. Problèmes inverses stochastiques

Modèle géologique *a priori*:

Une très grande liste d'informations à garder en tête

- **Modèle conceptuel;**
- Structures;
- Roches;
- Fluides;
- Formes,
- Relation d'âge;
- Stratigraphie,
- Faille et plissement;
- Incertitudes sur les structures géologique;
- Limite de résolution;

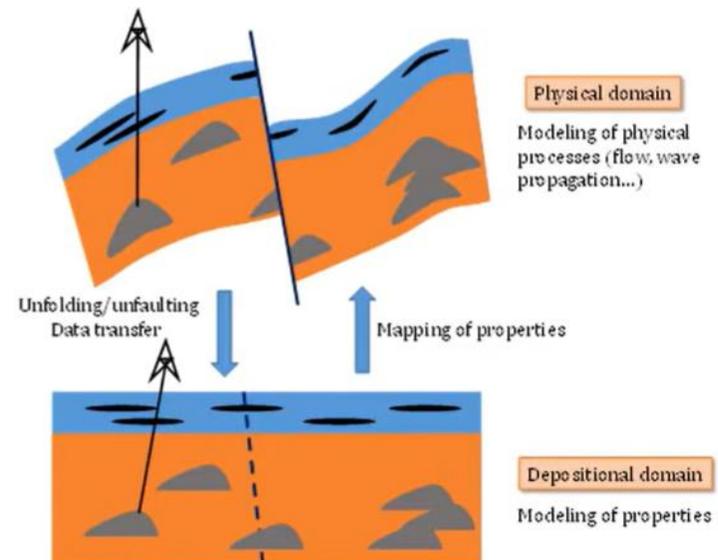
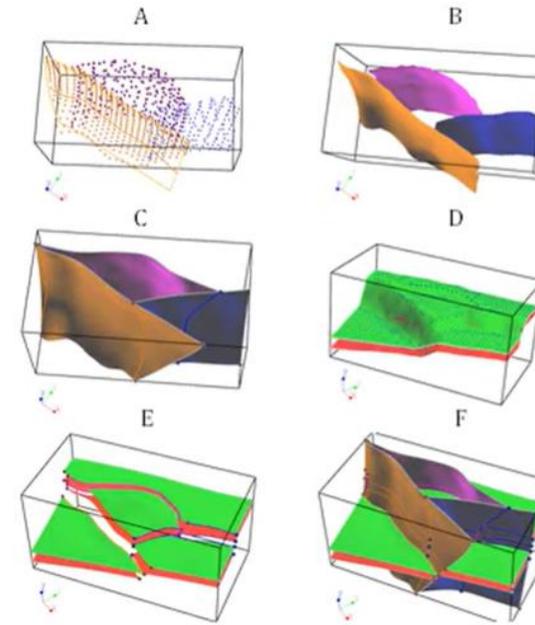


3. Problèmes inverses stochastiques

Modèle géologique *a priori*:

Une très grande liste d'informations à garder en tête

- Modèle conceptuel;
- **Structures;**
- **Roches;**
- **Fluides;**
- Formes,
- Relation d'âge;
- Stratigraphie,
- Faille et plissement;
- Incertitudes sur les structures géologique;
- Limite de résolution;

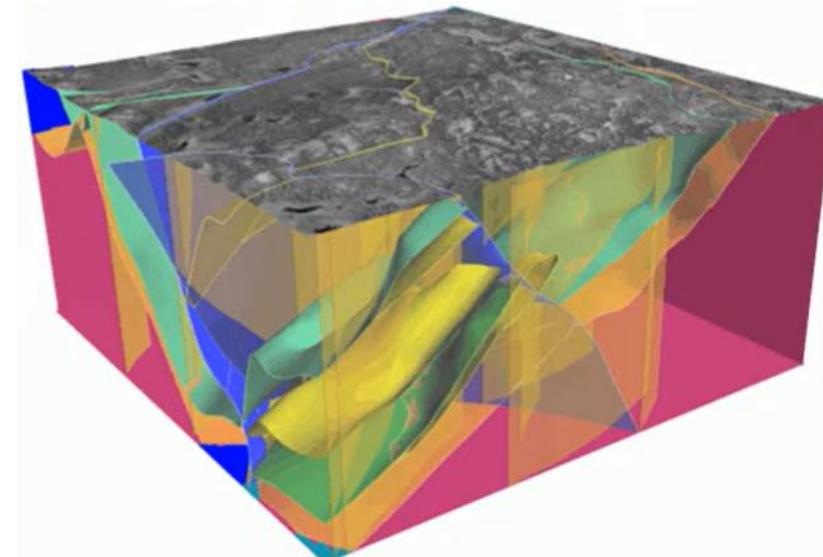
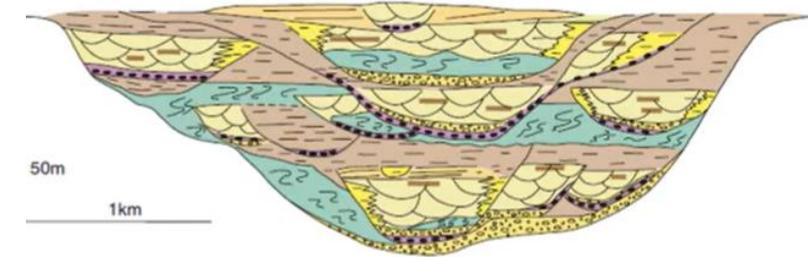
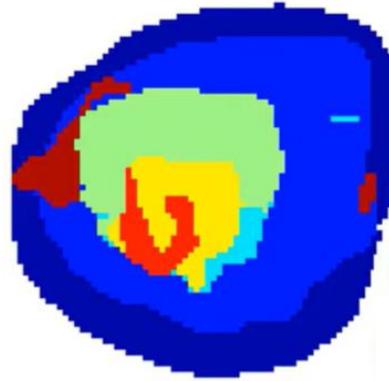


3. Problèmes inverses stochastiques

Modèle géologique *a priori*:

Une très grande liste d'informations à garder en tête

- Modèle conceptuel;
- Structures;
- Roches;
- Fluides;
- **Formes,**
- **Relation d'âge;**
- **Stratigraphie,**
- **Faïlle et plissement;**
- Incertitudes sur les structures géologique;
- Limite de résolution;

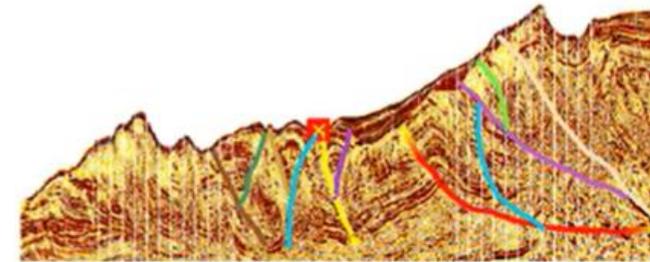
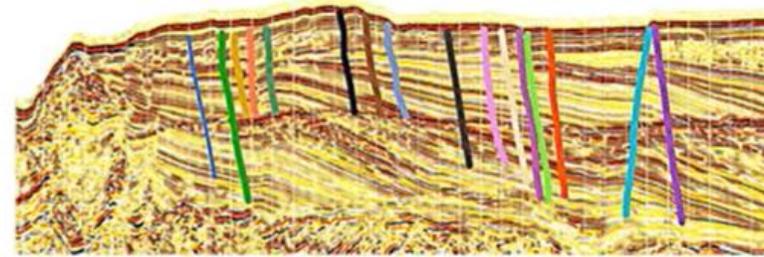


3. Problèmes inverses stochastiques

Modèle géologique *a priori*:

Une très grande liste d'informations à garder en tête

- Modèle conceptuel;
- Structures;
- Roches;
- Fluides;
- Formes,
- Relation d'âge;
- Stratigraphie,
- Faille et plissement;
- **Incertitudes sur les structures géologique;**
- Limite de résolution;

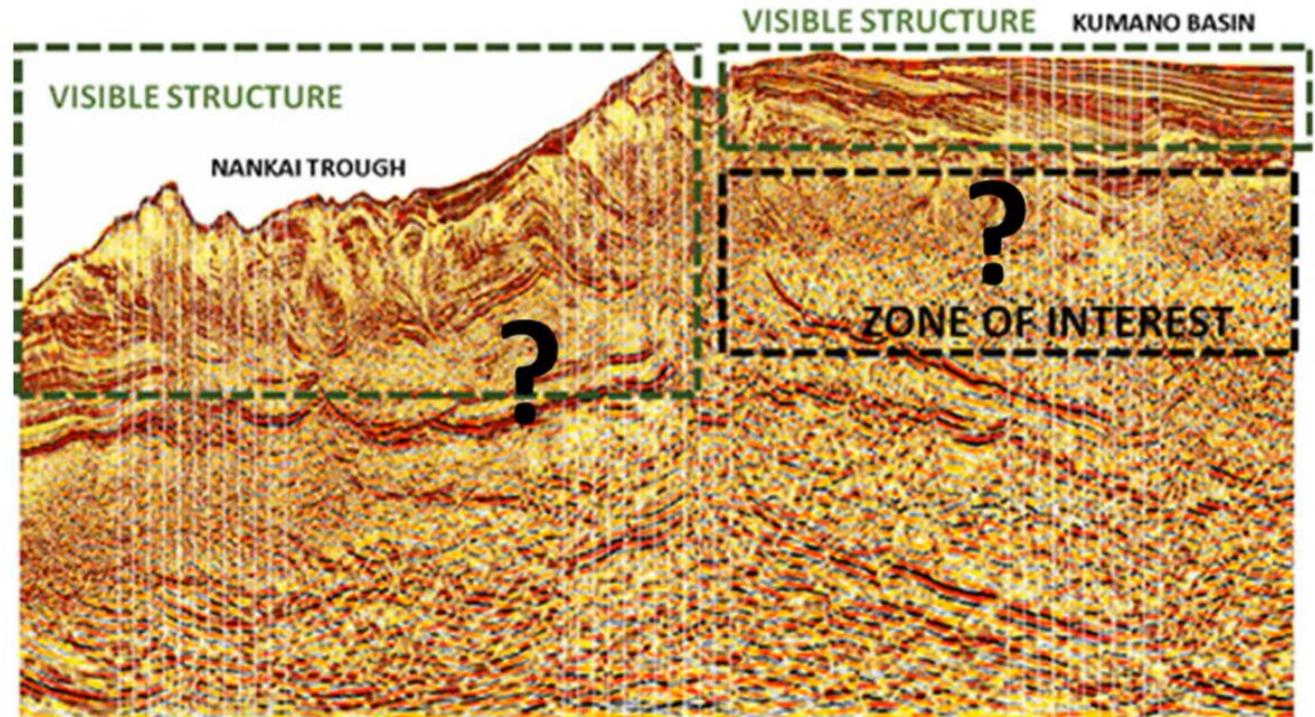


3. Problèmes inverses stochastiques

Modèle géologique *a priori*:

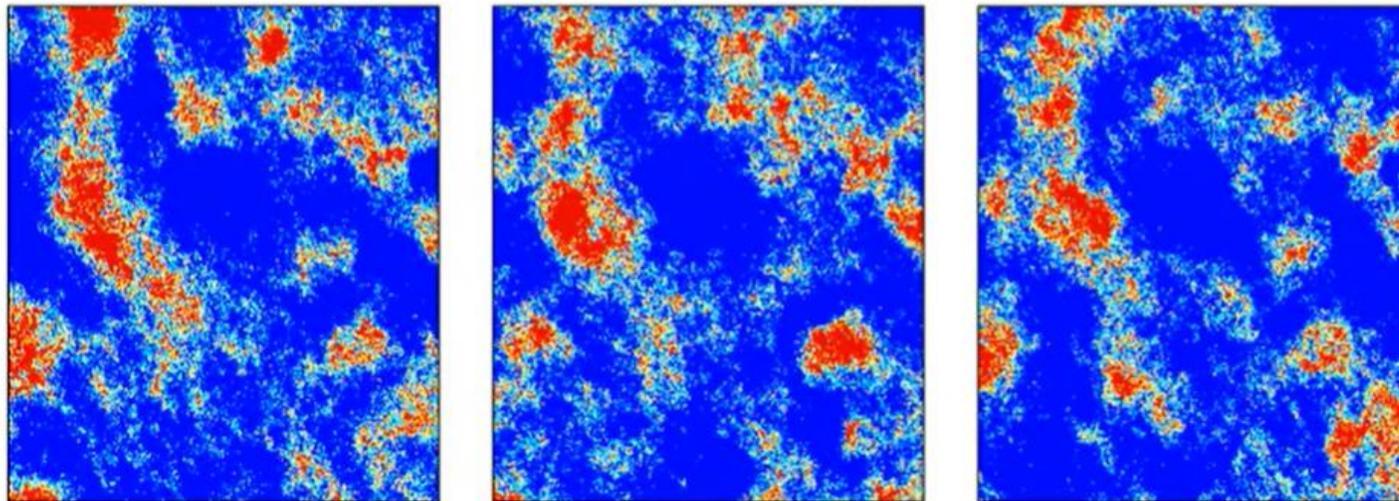
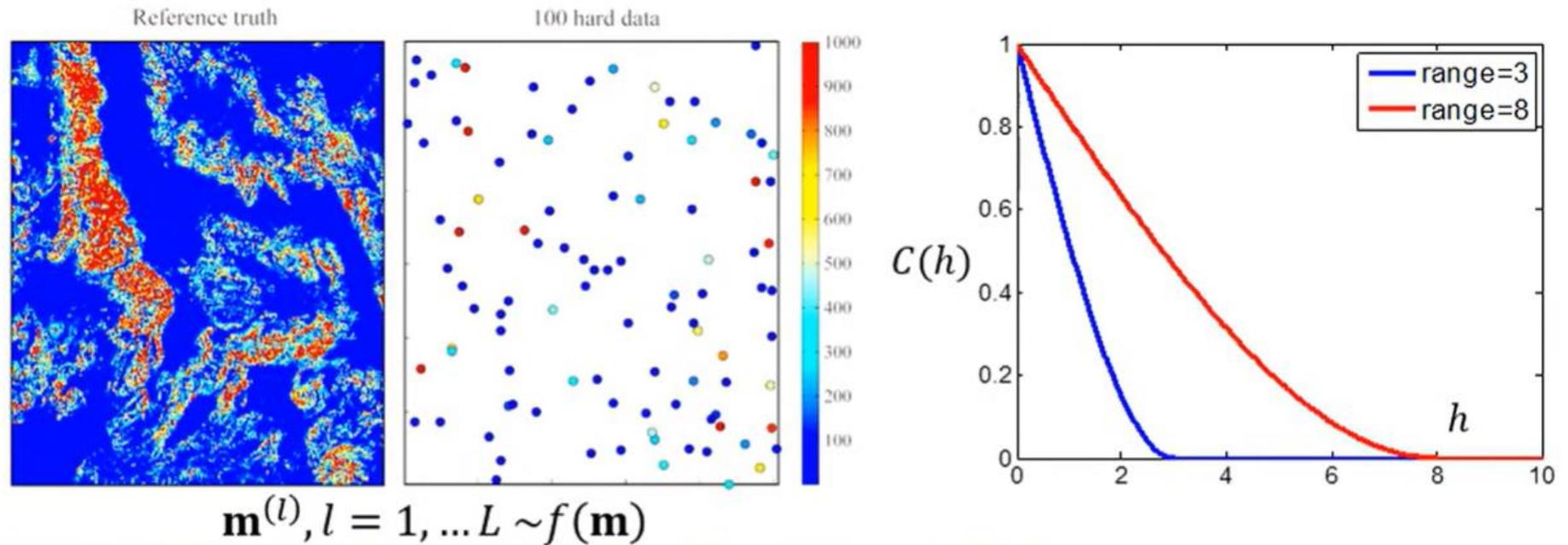
Une très grande liste d'informations à garder en tête

- Modèle conceptuel;
- Structures;
- Roches;
- Fluides;
- Formes,
- Relation d'âge;
- Stratigraphie,
- Faille et plissement;
- Incertitudes sur les structures géologique;
- **Limite de résolution;**



3. Problèmes inverses stochastiques

Modèle d'*a priori*: Fonction de covariance spatiale



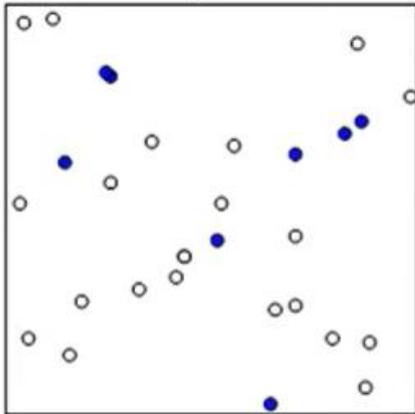
Conditional realizations

3. Problèmes inverses stochastiques

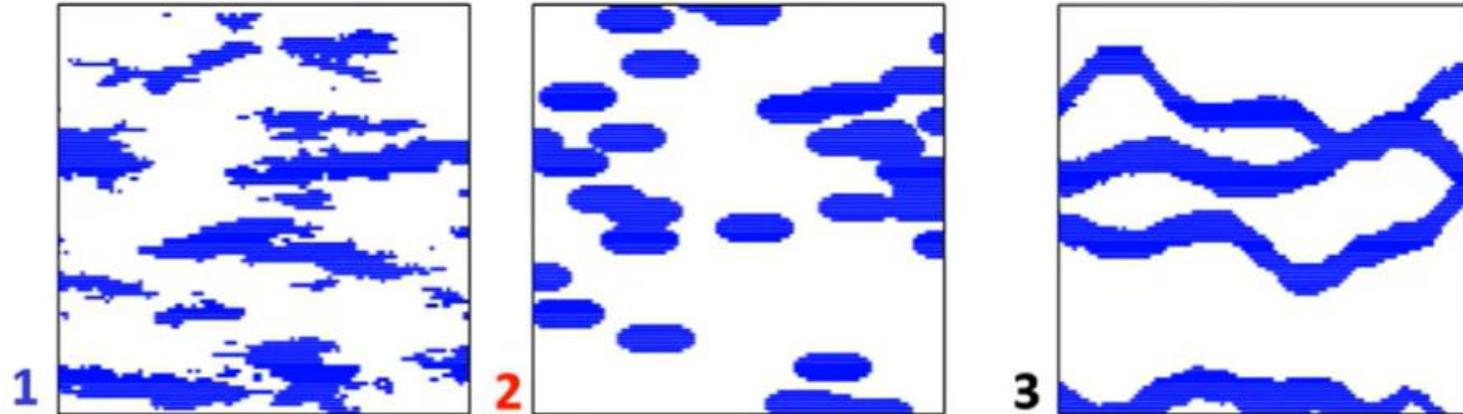
Modèle d'*a priori*: Fonction de covariance spatiale

Limitations

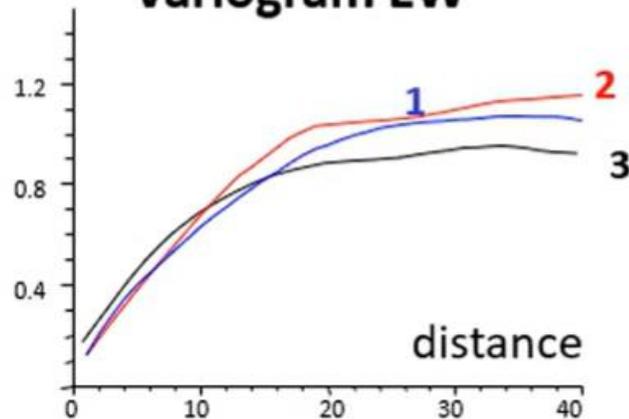
30 sample data



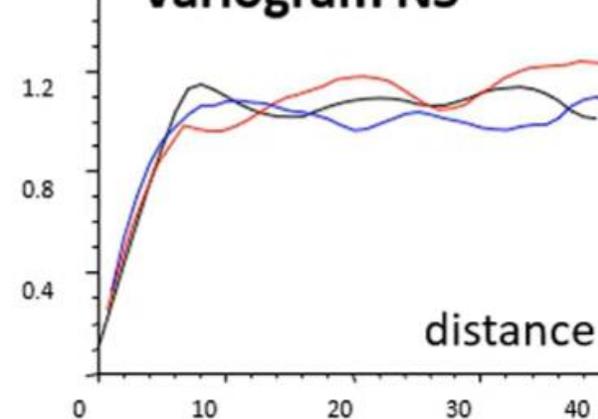
Different numerical models conditioned to the same data



Variogram EW



Variogram NS



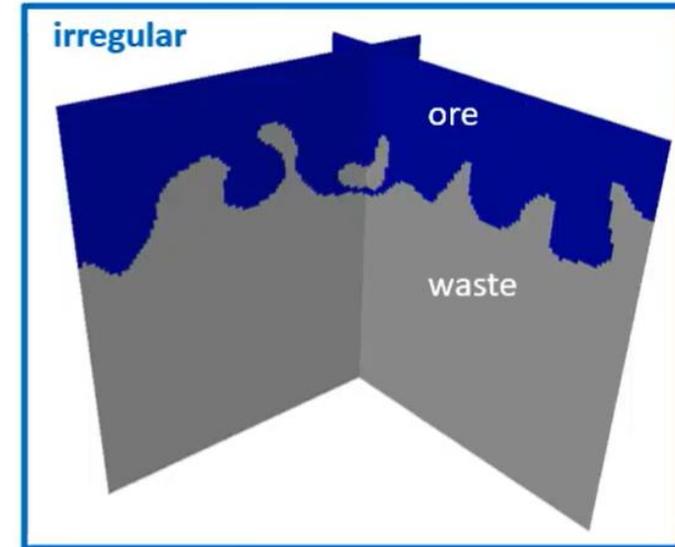
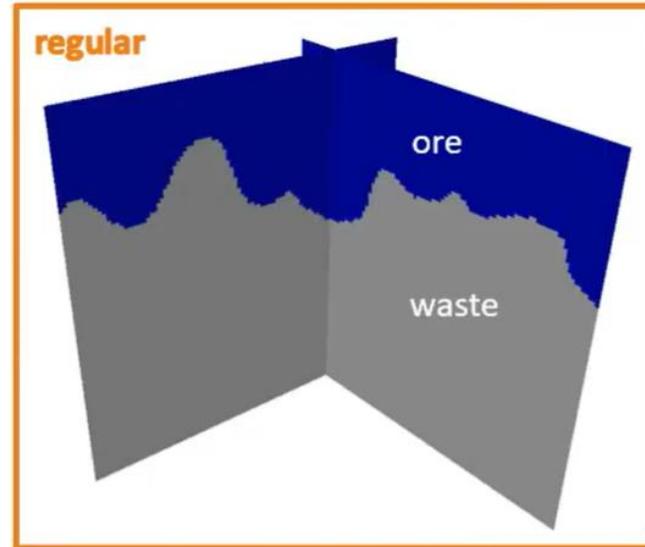
3. Problèmes inverses stochastiques

Modèle d'*a priori*: Fonction de covariance spatiale

Limitations

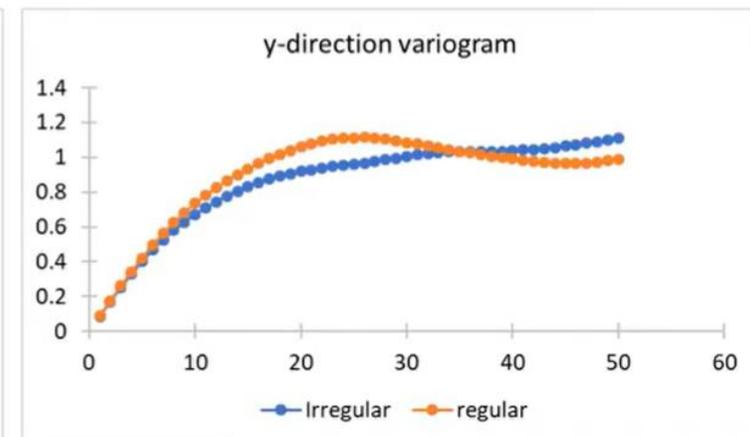
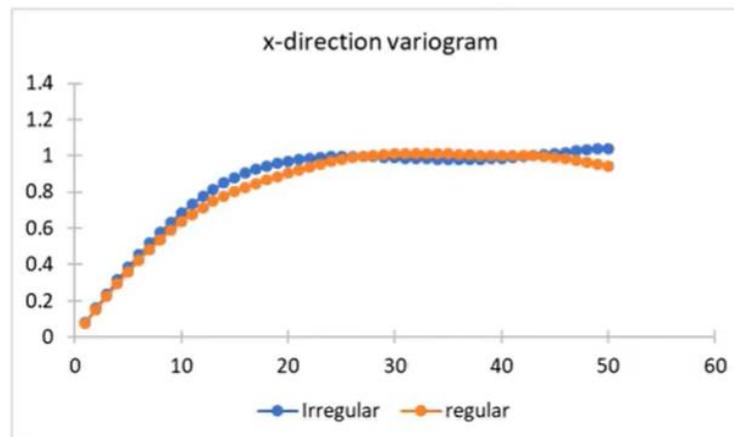
Autre exemple

Le variogramme ne peut détecter les différences topologiques



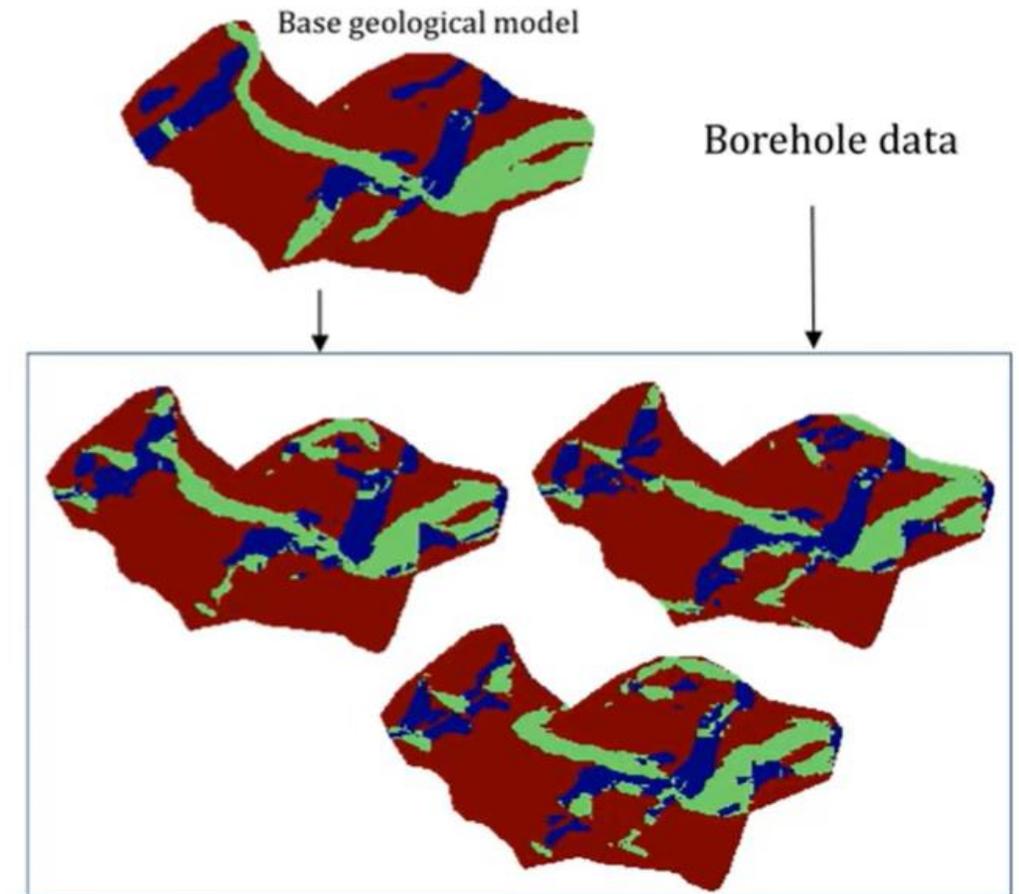
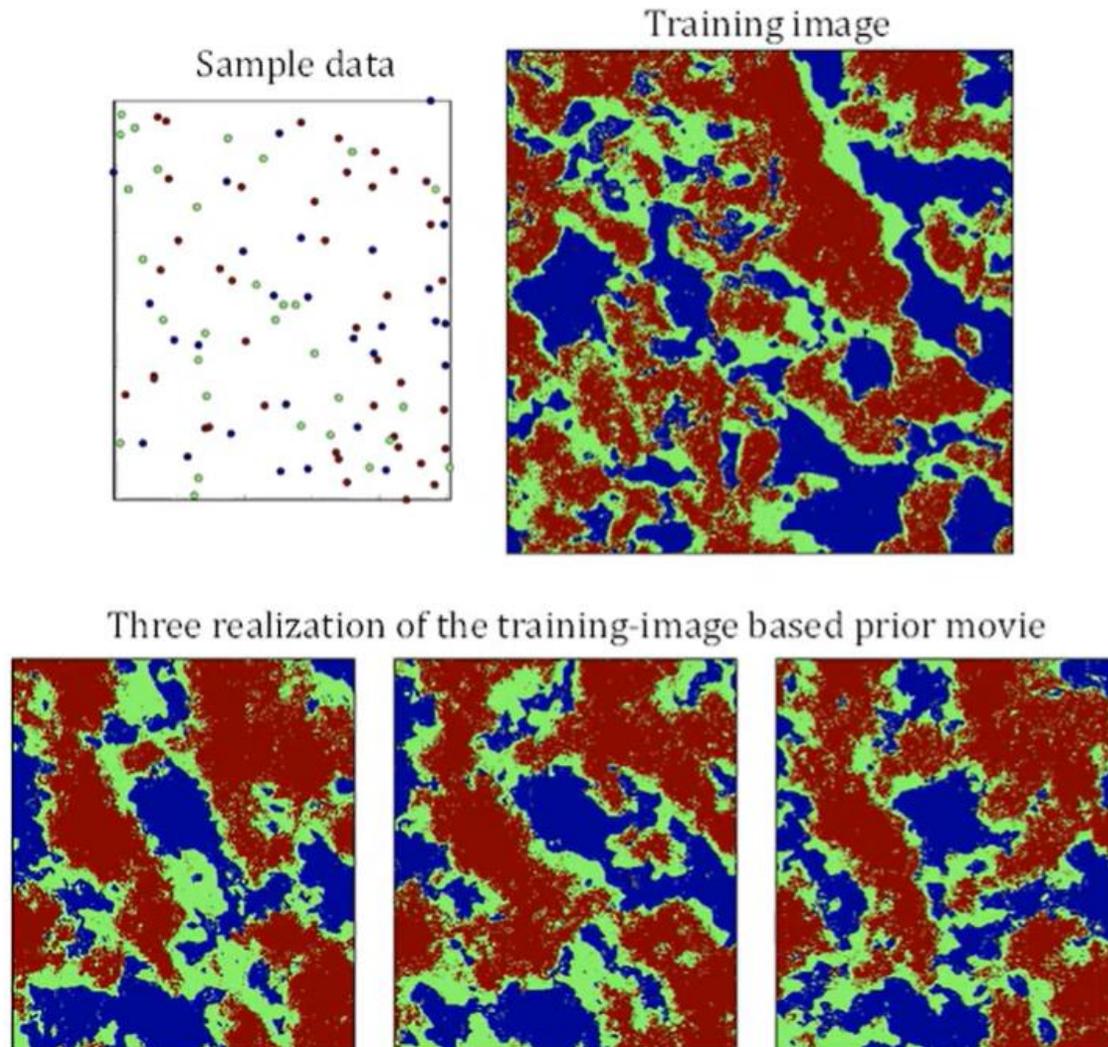
Différente topologie, même variogramme

Le variogramme est une ligne droite dans la direction z



3. Problèmes inverses stochastiques

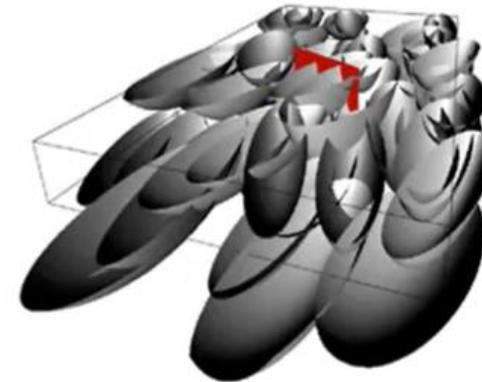
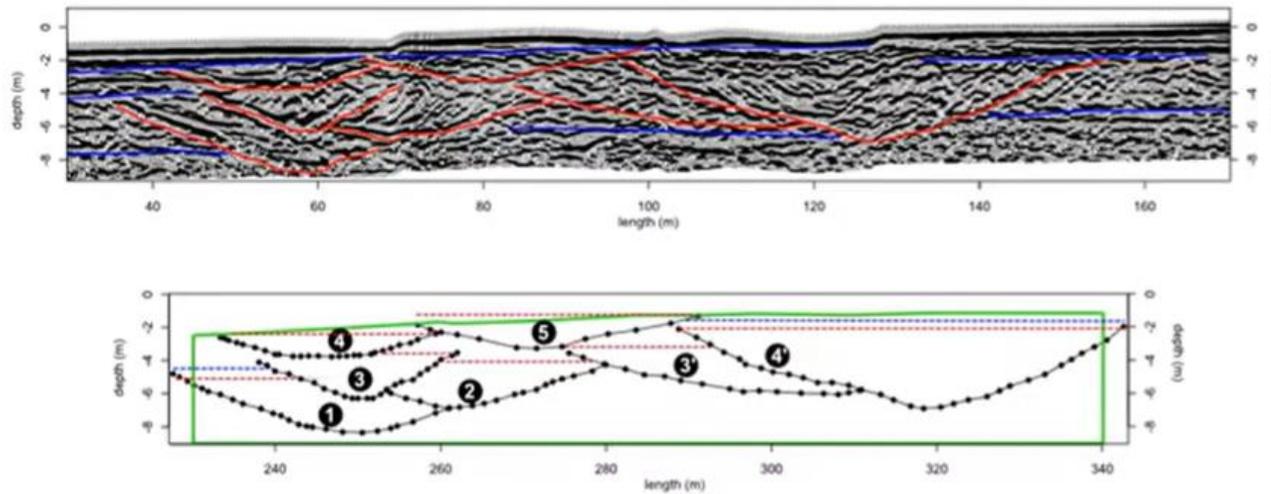
Modèle d'*a priori*: Image d'entraînement



$$\mathbf{m}^{(l)}, l = 1, \dots, L \sim f(\mathbf{m})$$

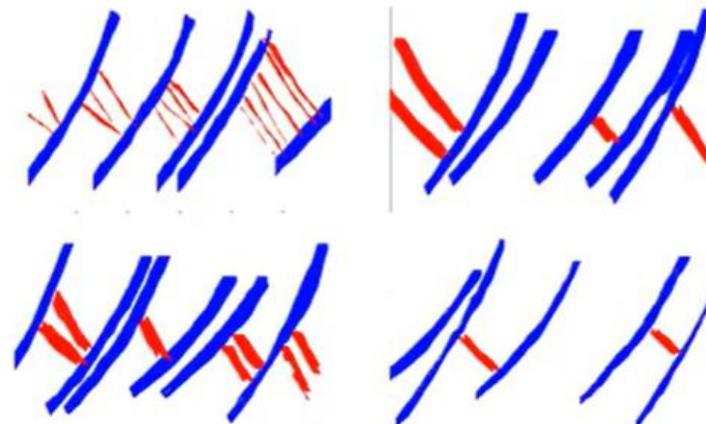
3. Problèmes inverses stochastiques

Modèle d'*a priori*: Modèle par objet



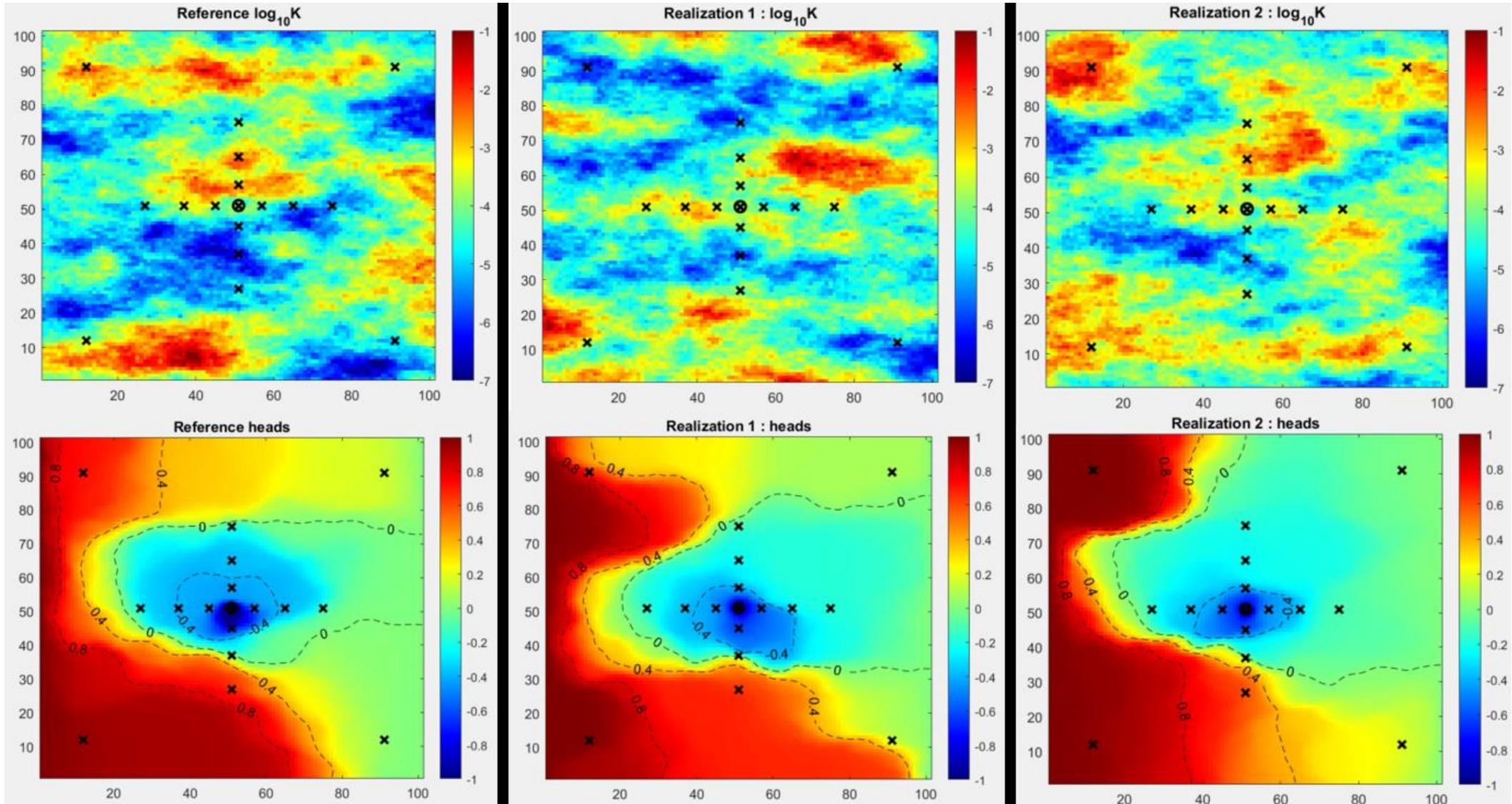
Prior

$$\mathbf{m}^{(l)}, l = 1, \dots, L \sim f(\mathbf{m})$$



3. Problèmes inverses stochastiques

Inversion bayésienne avec des modèles géologiques *a priori*



3. Problèmes inverses stochastiques

Inversion bayésienne avec des modèles géologiques *a priori* :

Méthode Monte Carlo par chaînes de Markov: Échantillonnage de Metropolis

Principe de base

Objectif : Générer une chaîne de Markov où les états suivent la distribution cible $p(\mathbf{m})$.

Idée clé :

1. Proposer une transition vers un nouvel état \mathbf{m}' à partir d'un état actuel \mathbf{m} .
2. Accepter ou rejeter cette transition selon une probabilité spécifique.

Étapes de l'algorithme

1. Choisir un état initial x_0 .

2. Pour $i = 1, 2, \dots, N$:

- I. Proposer un nouvel état \mathbf{m}^* à partir d'une distribution de proposition $\mathbf{m}^* \sim p(\mathbf{m} | \mathbf{m}^{i-1})$
- II. Calculer la probabilité d'acceptation :

$$\alpha = \min \left(1, \frac{\mathcal{L}(\mathbf{m}^* | \mathbf{d})}{\mathcal{L}(\mathbf{m}^{i-1} | \mathbf{d})} \right)$$

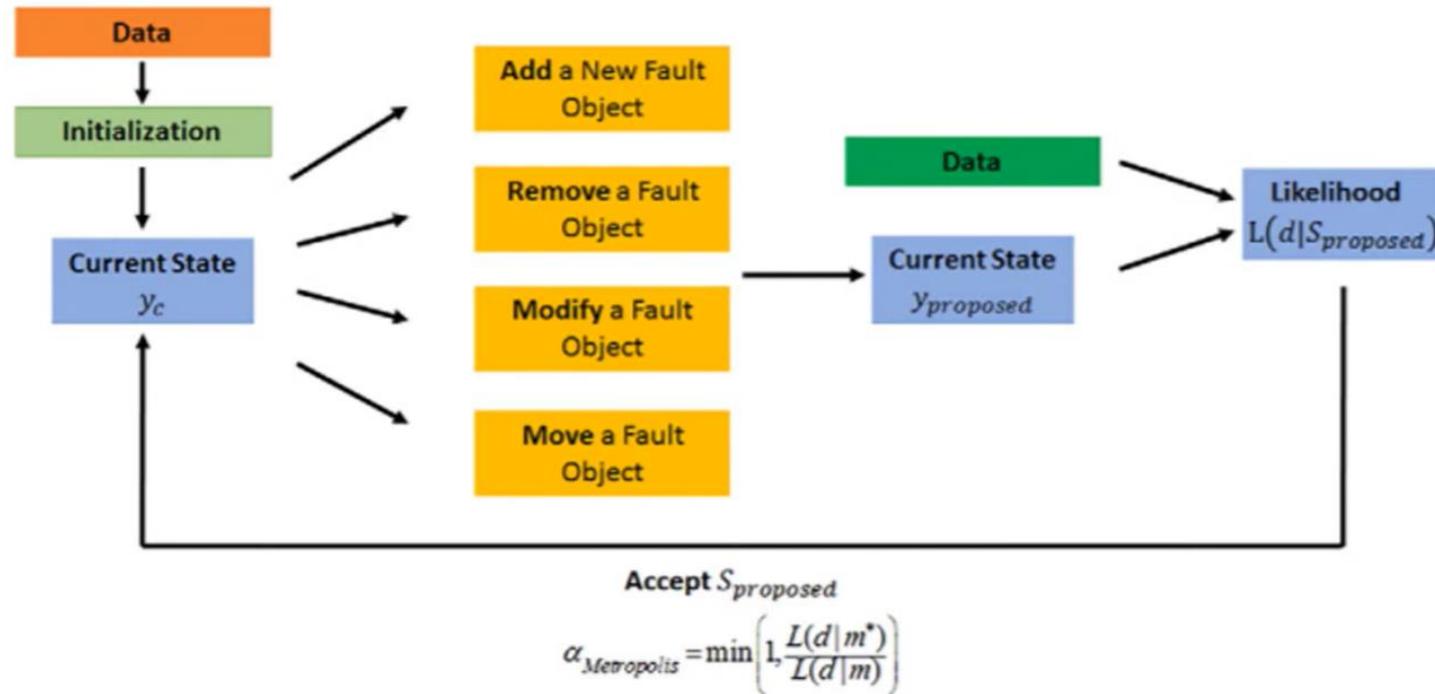
III. Accepter ou rejeter \mathbf{m}^* :

- Si $u \sim U(0,1)$ et $u \leq \alpha$, accepter \mathbf{m}^* (ajouter à la chaîne).
- Sinon, rejeter \mathbf{m}^* (répéter pour trouver un \mathbf{m}^*).

Pour Métropolis, la distribution doit être symétrique $p(\mathbf{m}^i | \mathbf{m}^j) = p(\mathbf{m}^j | \mathbf{m}^i)$

3. Problèmes inverses stochastiques

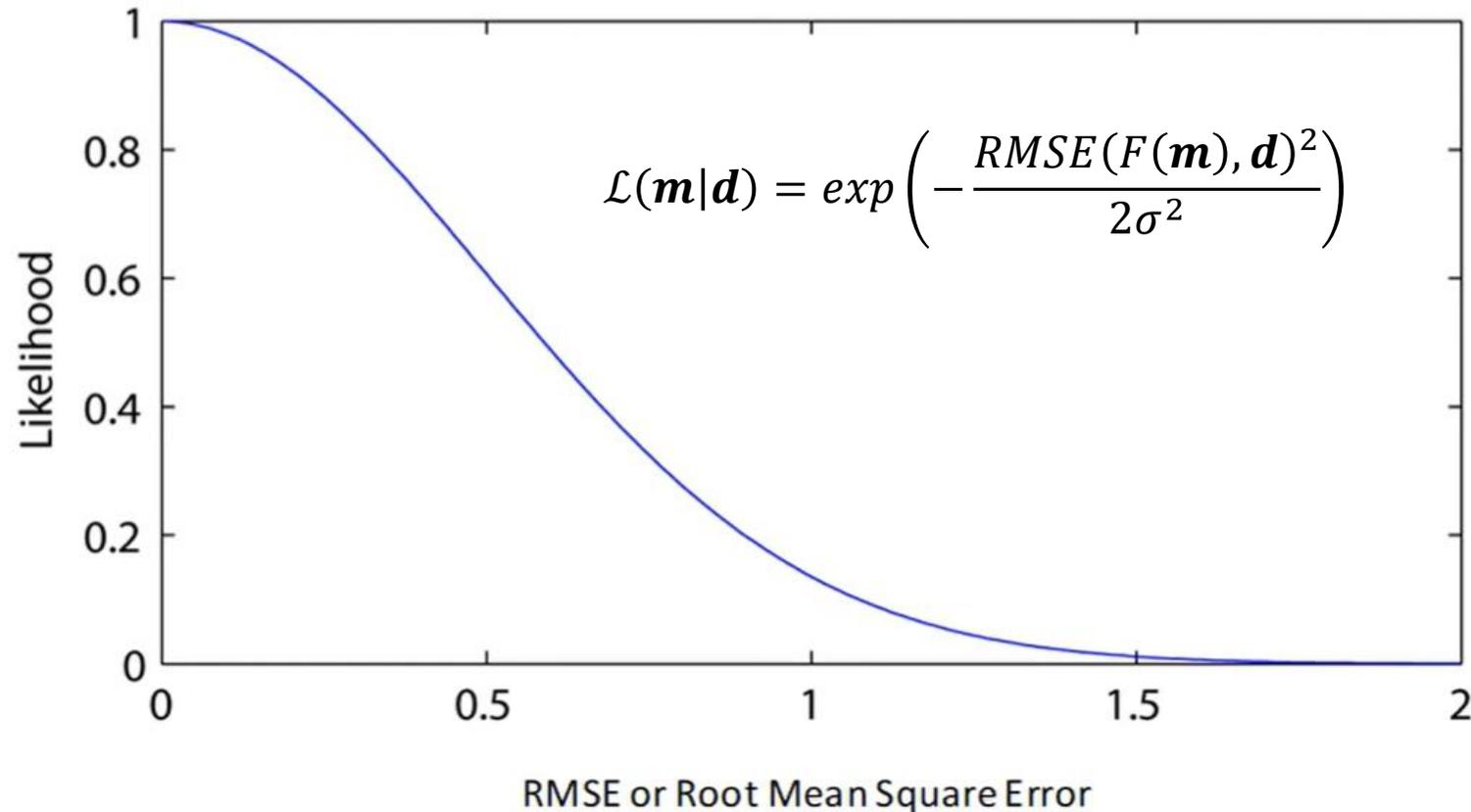
Inversion bayésienne avec des modèles géologiques *a priori* :
Méthode Monte Carlo par chaînes de Markov: Échantillonnage de Metropolis



3. Problèmes inverses stochastiques

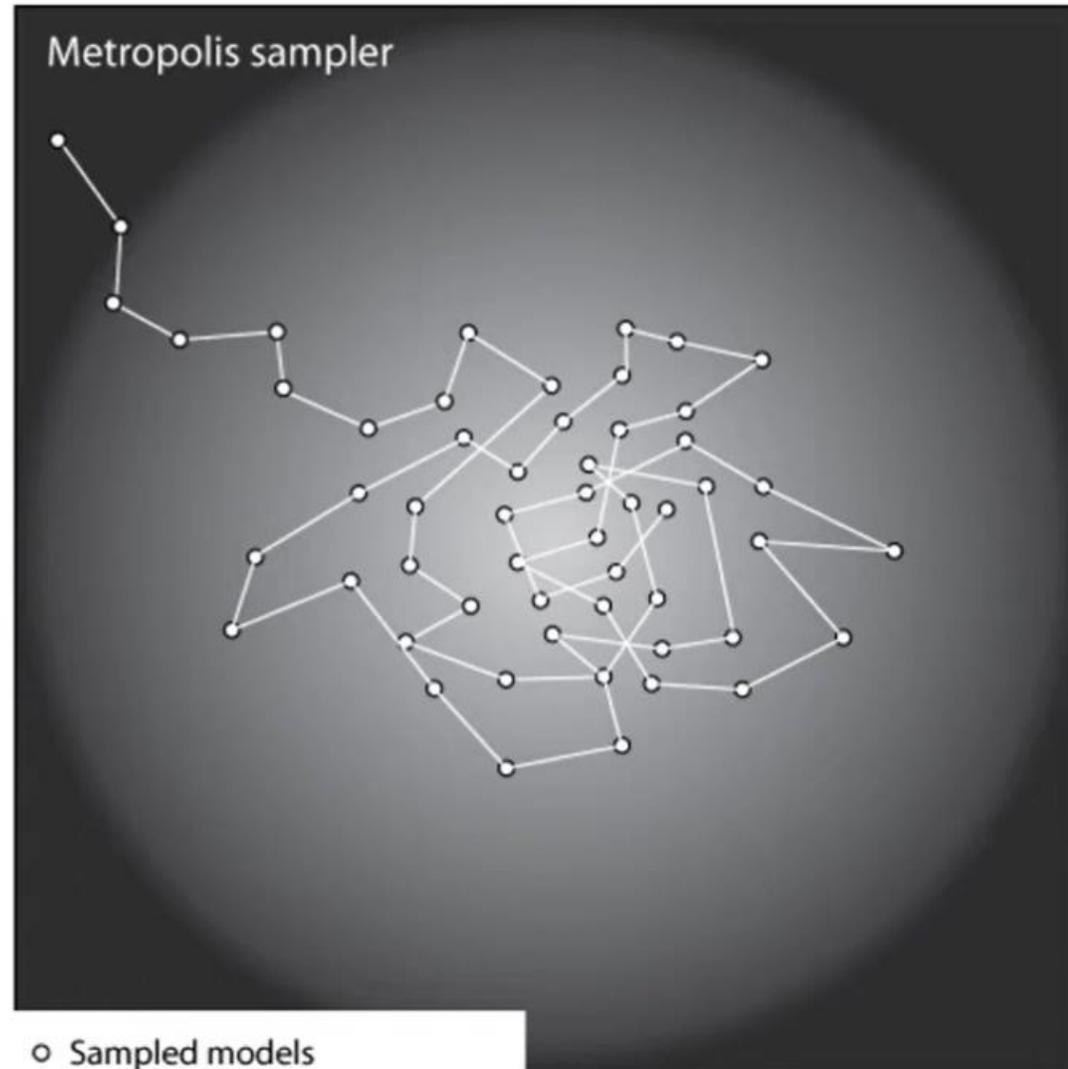
Inversion bayésienne avec des modèles géologiques *a priori* :
Méthode Monte Carlo par chaînes de Markov: Échantillonnage de Metropolis

On peut transformer une mesure d'erreur en une fonction de vraisemblance



3. Problèmes inverses stochastiques

Inversion bayésienne avec des modèles géologiques *a priori* :
Méthode Monte Carlo par chaînes de Markov: Échantillonnage de Metropolis

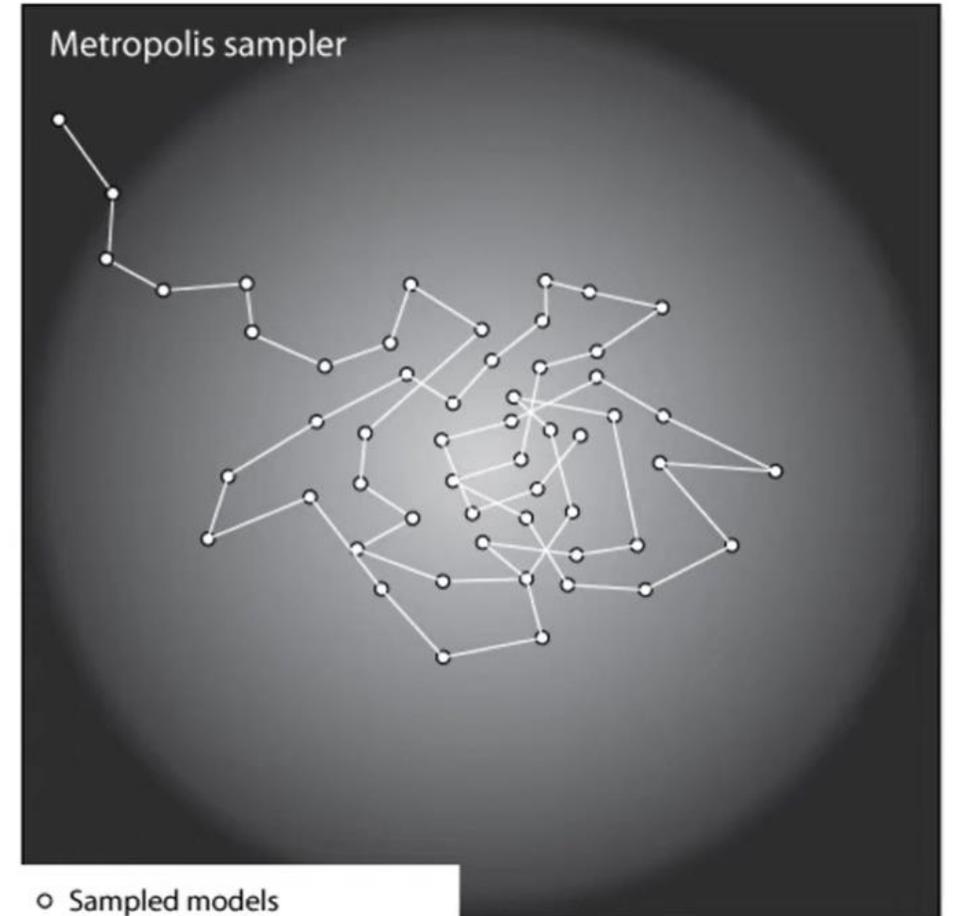


3. Problèmes inverses stochastiques

Inversion bayésienne avec des modèles géologiques *a priori* :
Méthode Monte Carlo par chaînes de Markov: Échantillonnage de Metropolis

Convergence

- 1) Les itérations doivent être grandes. Sinon, la solution sera très similaire au modèle *a priori* (***Burn-In***)
- 2) Des corrélations peuvent exister entre les modèles échantillonnés, il faut donc augmenter le nombre d'itération avant de rééchantillonner (***Sample discard interval***)



4. Optimisation stochastique

Définition:

Qu'est-ce que le calage :

Consiste à introduire des informations non linéaires obtenues sur le terrain à un modèle géologique par la **perturbation** de celui-ci.

L'impact des perturbations est quantifié par une **fonction objectif** (FO) constituée des paramètres d'intérêts.

On veut **minimiser** la FO.

Plus il y a de termes dans la fonction objectif et plus le calage devient complexe et possiblement long en termes d'exécution!

$$FO = \sum_{i=1}^N \lambda_i \text{Erreur}_i(Y_{sim}, Y_{ref})$$

4. Optimisation stochastique

Bien définir la fonction objectif :

Une mauvaise définition de la fonction objectif peut mener à des incohérences majeures au niveau de la géologie ou des statistiques (spatiales) du champ simulé.

On utilise souvent en hydrogéologie les FO suivantes (*MAE* , *RMSE*) :

$$FO = \frac{1}{N} \sum_{i=1}^N |h_{i,obs} - h_{i,sim}| , \quad FO = \sqrt{\frac{1}{N} \sum_{i=1}^N (h_{i,obs} - h_{i,sim})^2}$$

Par exemple : les h_i sont des charges hydrauliques observées et simulées mesurées en mètre. N est le nombre d'observations.

4. Optimisation stochastique

Bien définir la fonction objectif :

On peut regrouper les algorithmes de calage basé sur le variogramme en cinq grandes catégories.

- Propagation
- Point par point (séquentielle)
- Plusieurs points à la fois
- Fusion (combinaison)
- Constructif



4. Optimisation stochastique

Propagation: Points Pilotes (Certes and de Marsily, 1991, Doherty, 2003)

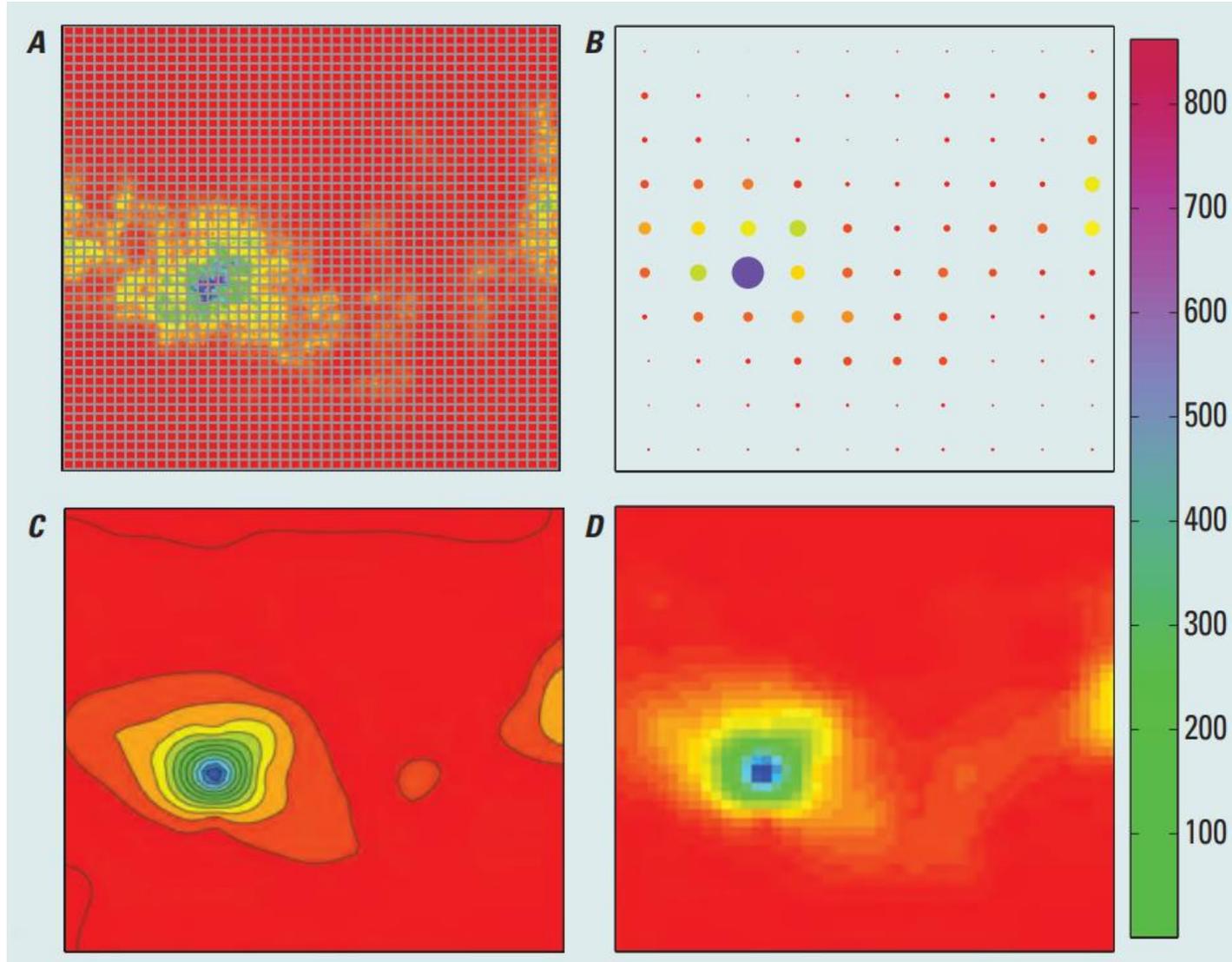
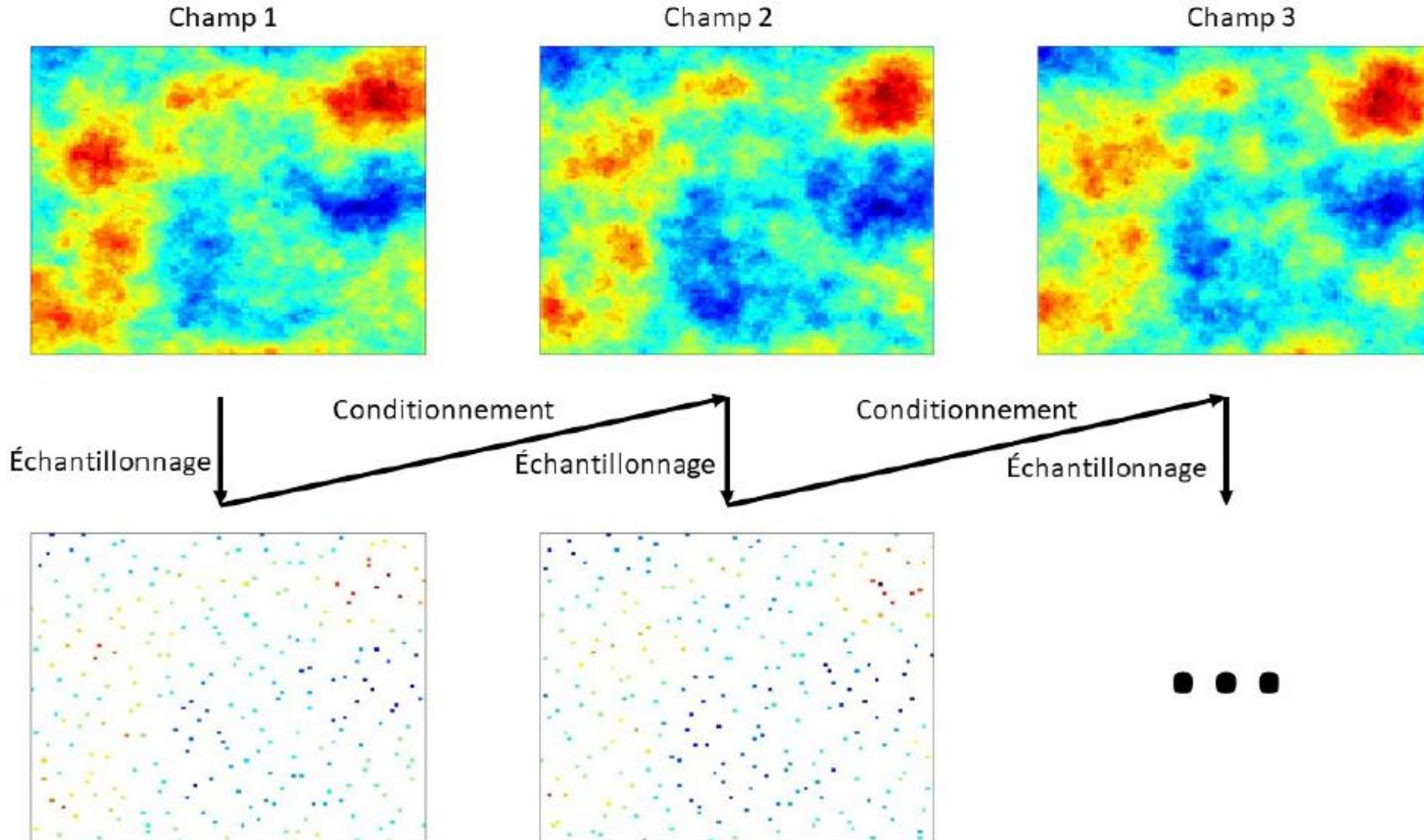


Figure 1. Conceptual overview of representing complex hydrogeologic conditions using pilot points. Panel (A) shows the inherent property value overlain by the model grid in gray. Panel (B) is a representation of the true property values by a grid of pilot points in which symbol size indicates value. Panel (C) shows an interpolated representation of panel (B) on an arbitrarily fine grid scale. Panel (D) shows the value from the pilot points interpolated to the computational-grid scale. Interpolation in all cases was performed using ordinary Kriging. The same color scale, in arbitrary units, applies to all four panels.

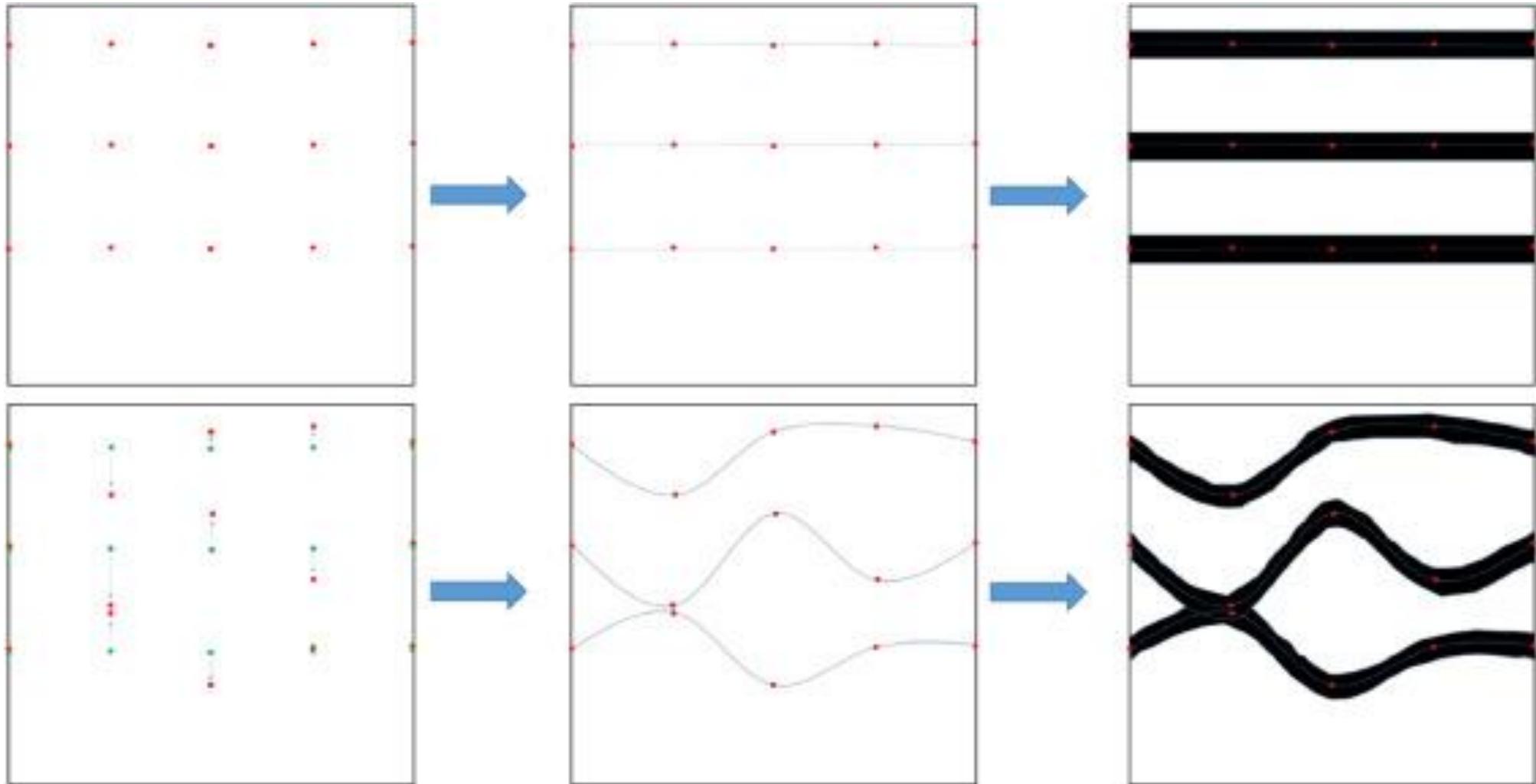
4. Optimisation stochastique

Propagation: Rééchantillonnage spatial itératif (Mariethoz, 2010)



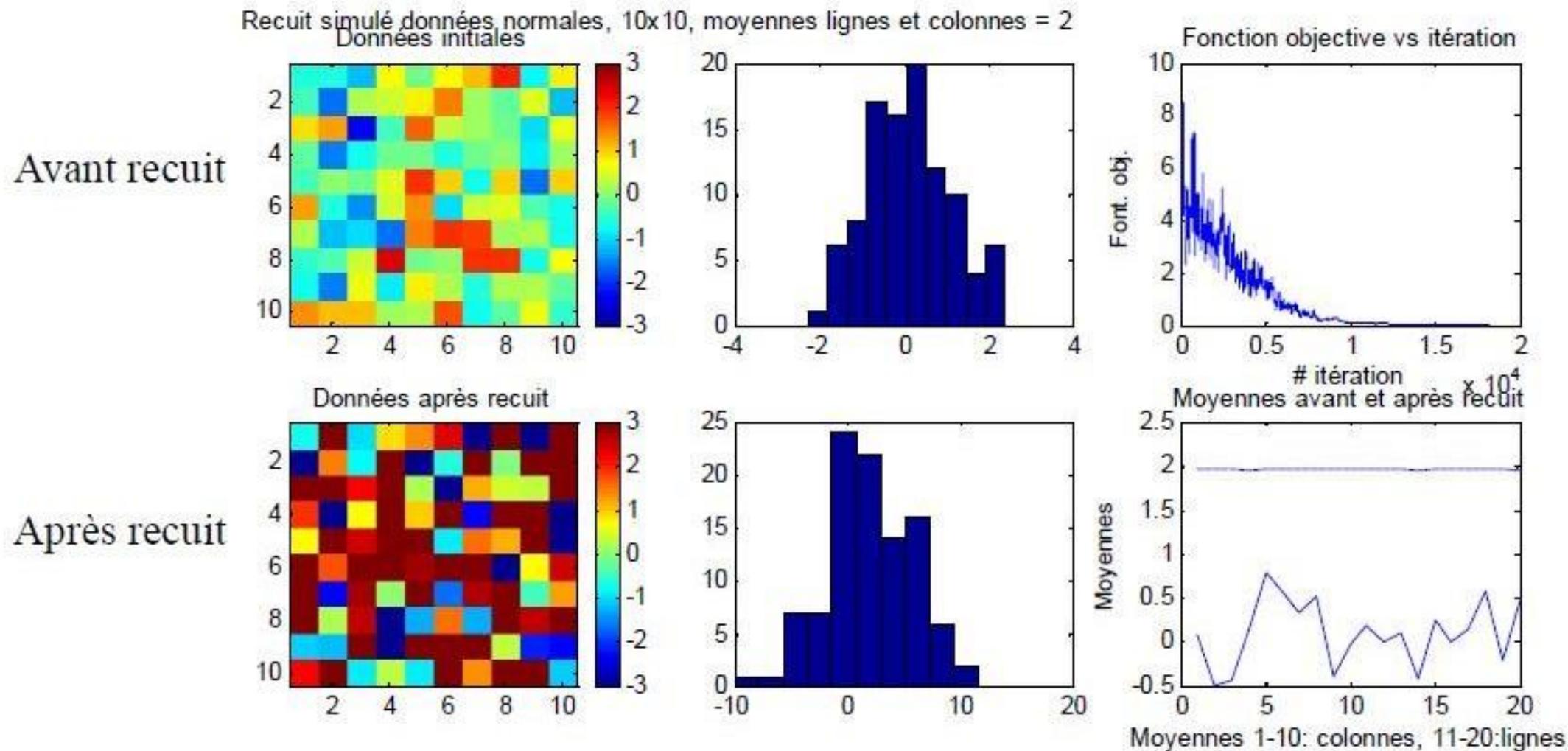
4. Optimisation stochastique

Propagation: Point pilot voyageur (Khambhammettu et al., 2020)



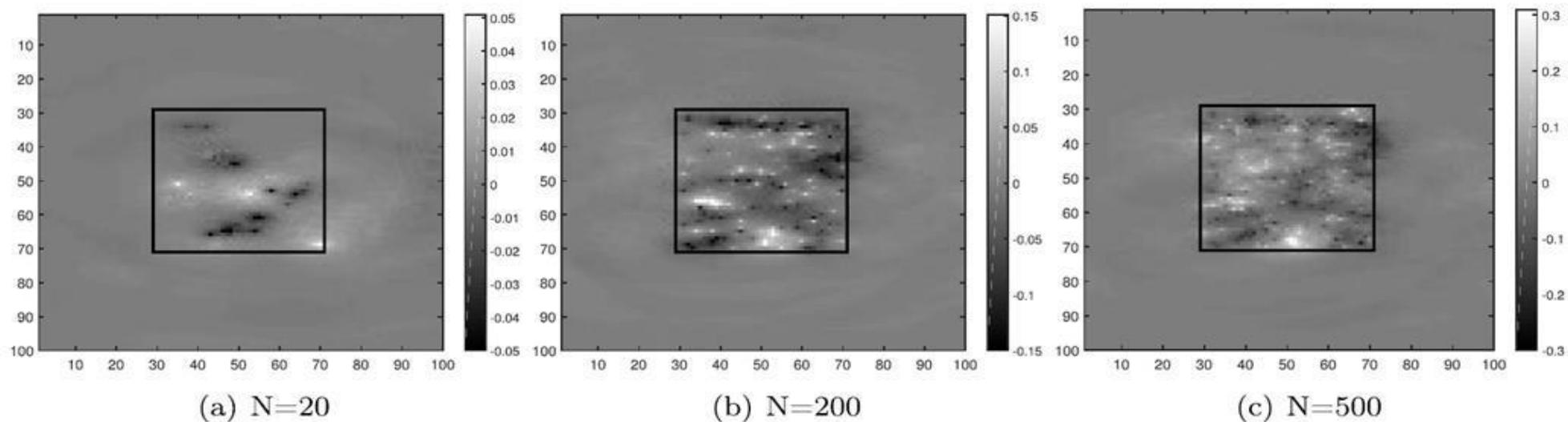
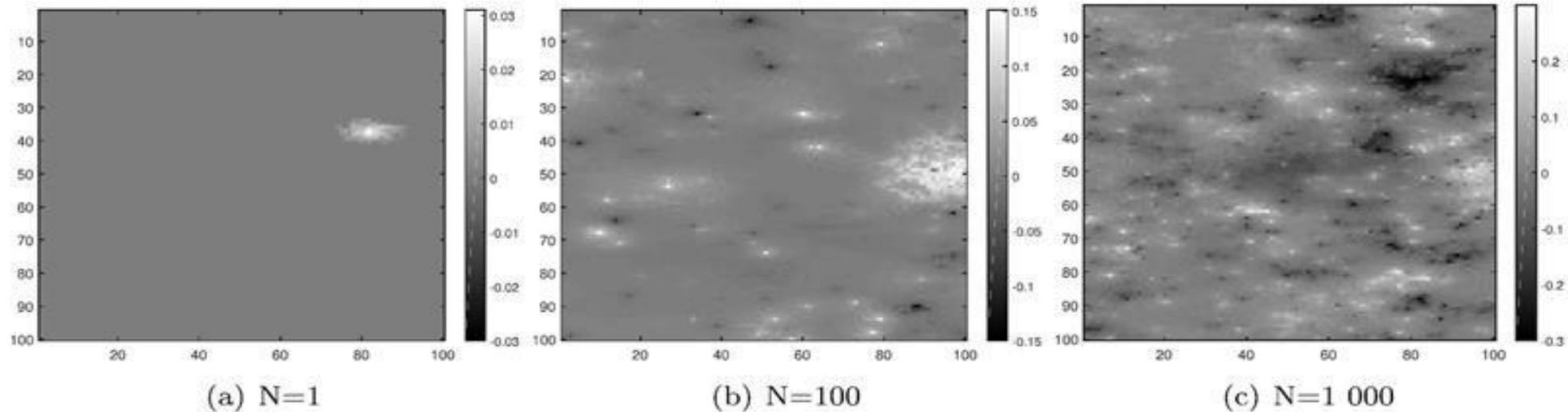
4. Optimisation stochastique

Point par point : Recuit simulé (Kirkpatrick, 1983, Deutsch, 1992)



4. Optimisation stochastique

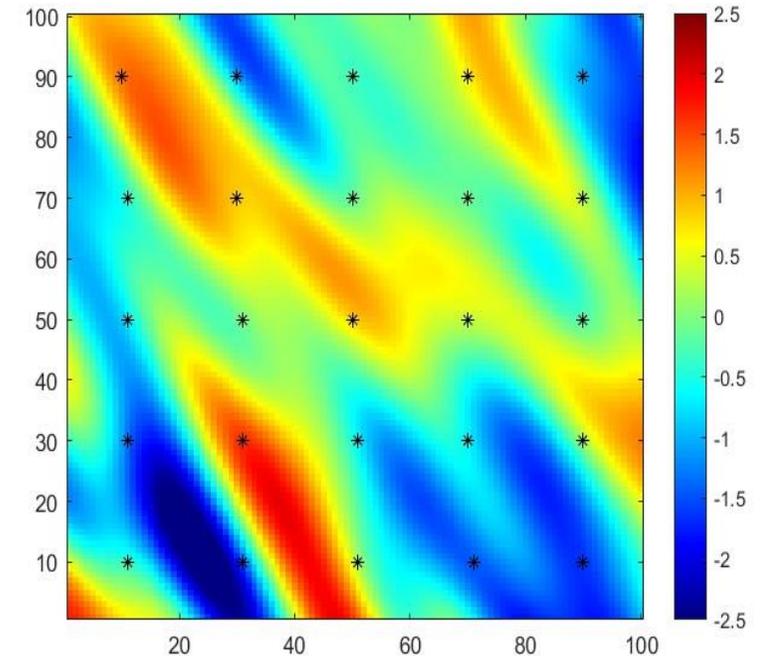
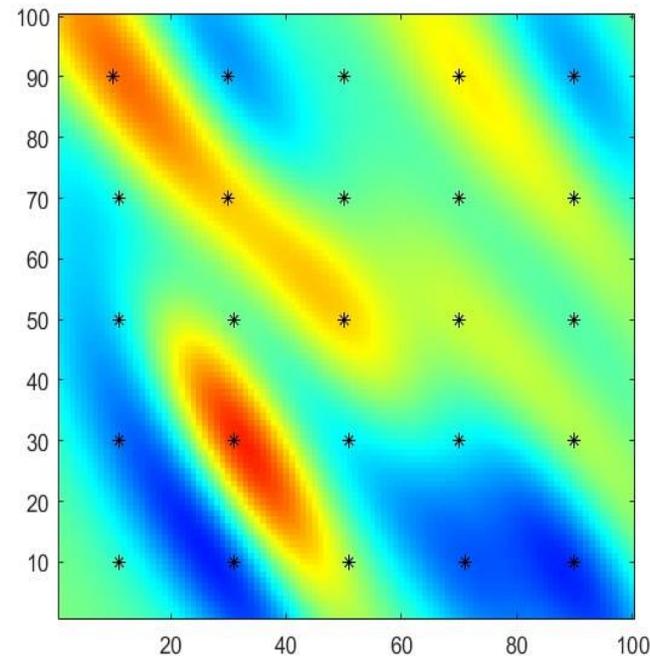
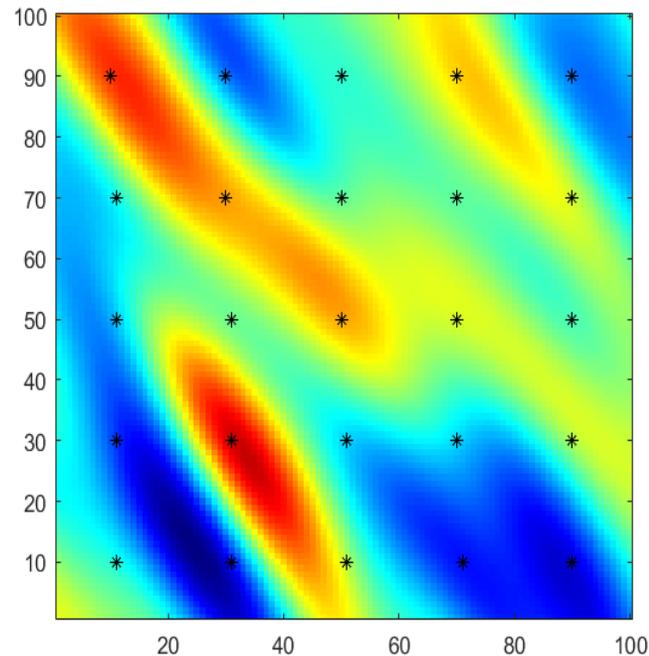
Plusieurs points simultanément: FFTMA-SA (Lauzon et Marcotte, 2019)



4. Optimisation stochastique

Fusion (combinaison) : Déformation graduelle (Hu,2000)

$$Y_i = Y_{i-1} \cos(t) + Y_n \sin(t)$$



4. Optimisation stochastique

Constructif : Bandes tournantes spectrales séquentielles (Lauzon et Marcotte, 2020)

Principe : Réduire le processus de simulation en n-D, à plusieurs simulations rapides en 1-D

(similaire à la méthode des bandes tournantes, mais cette fois-ci dans le domaine spectral)

$$Z_i(x) = \sqrt{\frac{1}{i}} Z_{i-1}(x) + \sqrt{\frac{i-1}{i}} \left(\sqrt{2} \cos(\underbrace{\langle v_i, x \rangle}_{\text{Projection des vecteurs fréquences aux vecteurs coordonnés. Passage de 1-D vers n-D}} + 2\pi u_i \right)$$

Champ calé à l'itération i

Champ calé à l'itération i-1

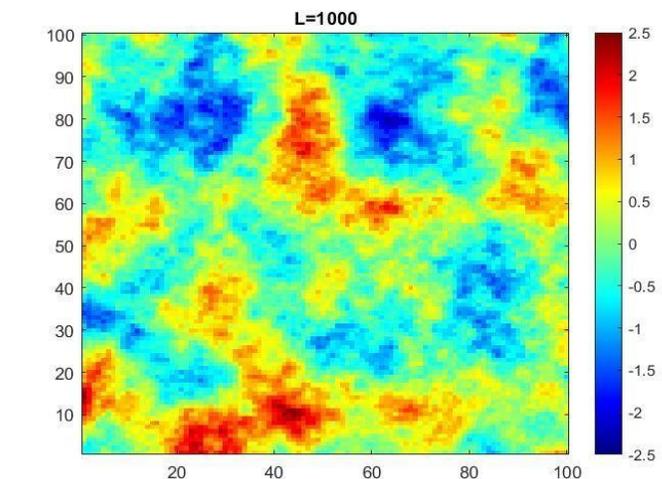
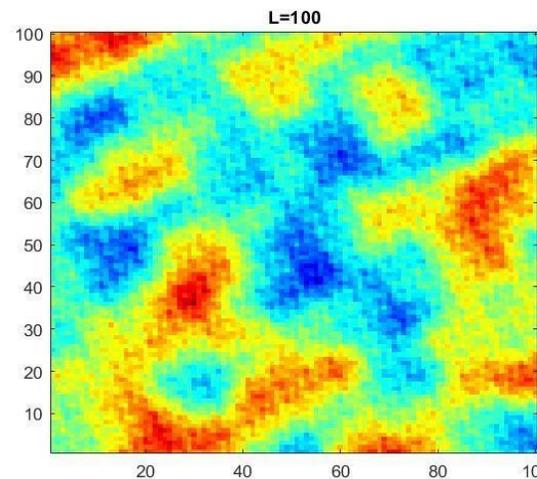
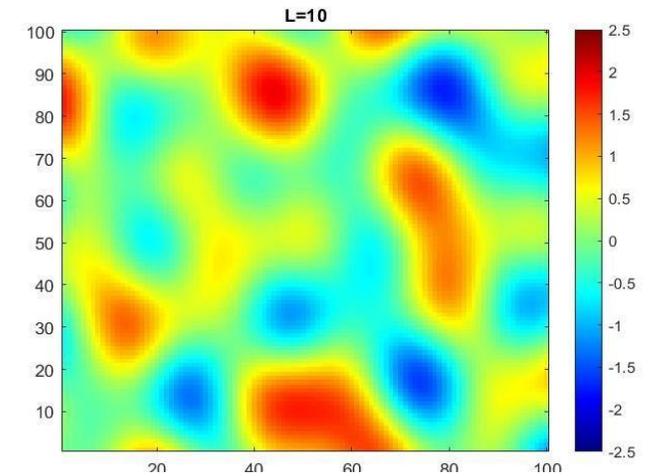
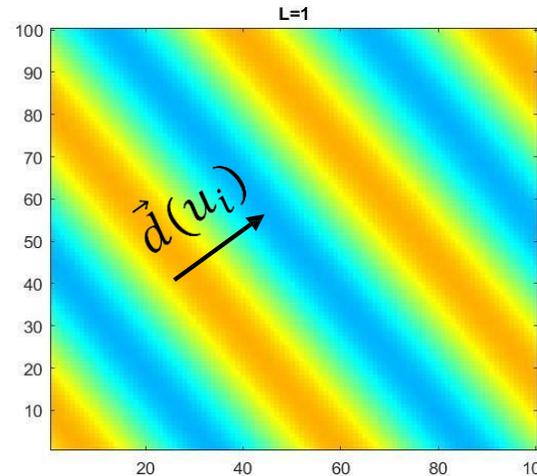
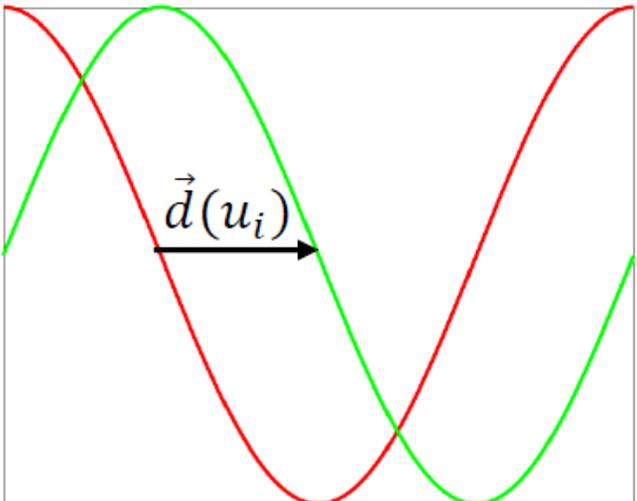
Vecteurs fréquences (contiens l'Information sur la continuité spatiale)

Phases suivant une loi uniforme [0,1]. À optimiser.

4. Optimisation stochastique

Constructif : Bandes tournantes spectrales séquentielles (Lauzon et Marcotte, 2020)

Idée :
Optimisation des phases!
(c.-à-d. perturber les phases
d'un vecteur(u_i))



4. Optimisation stochastique

Que retenir du calage

1— Type de **perturbation**

- Par propagation (points pilotes, rééchantillonnage spatial itératif);
- Point par point (recuit simulé);
- Plusieurs points à la fois (recuit simulé, recuit de phase, FFTMA-SA);
- Par combinaison de plusieurs simulations (déformation graduelle et S-STBM).

2— Bien définir la **fonction objectif**

- Éviter les problèmes d'incohérence entre le calage et l'objectif visé

3— Il faut un **lien physique** entre les paramètres simulés et les données secondaires

- Simulateur d'écoulement; relation pétrophysique; équations d'états.

4— **Bien choisir l'algorithme** de calage en fonction des objectifs

- Calage local versus globale;
- Grille irrégulière versus régulière;
- Propriétés de convergences;
- Taille des simulations et espace mémoire.



5. Filtre de Kalman d'ensemble (EnKF)

Problématique : Données temporelles

Les systèmes dynamiques complexes, comme les prévisions météorologiques ou les modèles de réservoirs, présentent des **incertitudes significatives**. La relation temporelle n'est pas toujours évidente à prendre en considération.

Le Filtre de Kalman classique suppose :

- Une dynamique linéaire.
- Une gaussianité stricte des erreurs.
- Limitation : Inadapté pour des modèles non linéaires ou de grande dimension.

Solution : EnKF

Le EnKF est une extension stochastique :

- Adapte le Filtre de Kalman pour des systèmes non linéaires.
- Utilise un ensemble de réalisations pour représenter la distribution des états du système.



5. Filtre de Kalman d'ensemble (EnKF)

Idée Clé : Approche par Ensembles

Représentation stochastique

X = Ensemble d'états (N membres)

Mise à jour des états et des incertitudes basées sur :

La dynamique du modèle (prévision).

L'intégration des observations (analyse).

Modèle → Ensemble prédit → Mise à jour avec observations → Ensemble corrigé.



5. Filtre de Kalman d'ensemble (EnKF)

Idée Clé : Approche par Ensembles (EnFK) – Everson (2008)

L'EnKF se compose d'une étape d'initialisation suivie d'étapes répétées de prévision et de mise à jour comme suit :

- 1. Étape d'initialisation.** Un ensemble initial (\mathbf{m}^0) de N réalisations est généré à l'aide de méthodes géostatistiques en intégrant autant de connaissances préalables que possible
- 2. Étape de prévision.** Le modèle direct (F) est résolu, à partir du temps zéro, pour chaque réalisation i , afin d'obtenir des prédictions du modèle ($\mathbf{d}_{i,t}$) au pas de temps (t) en utilisant la dernière mise à jour des ensembles ($\mathbf{m}_{i,t-1}$)

$$\mathbf{d}_{i,t} = F(\mathbf{m}_{i,t-1} | \mathbf{d}_{i,t-1}), \quad \forall i = 1, \dots, N$$

- 3. Étape de mise à jour.** Le vecteur des modèles est mis à jour ($\mathbf{m}_{i,t}^u$) en fonction des écarts entre les données prédites ($\mathbf{d}_{i,t}^f$) et observées (\mathbf{d}_t^0).

$$\begin{aligned} \mathbf{m}_{i,t}^u &= \mathbf{m}_{i,t-1}^u + \mathbf{K}_t (\mathbf{d}_t^0 + \boldsymbol{\varepsilon}_{i,t}^0 - \mathbf{d}_{i,t}^u) \\ \mathbf{K}_t &= \mathbf{C}_{MD}^t (\mathbf{C}_{DD}^t + \mathbf{R}_t)^{-1} \end{aligned}$$

- 4. Retour à l'étape de prévision.**

$$\begin{aligned} \mathbf{C}_{MD} &= \frac{1}{N-1} (\mathbf{m} - \bar{\mathbf{m}})(\mathbf{d} - \bar{\mathbf{d}})^T \\ \mathbf{C}_{DD} &= \frac{1}{N-1} (\mathbf{d} - \bar{\mathbf{d}})(\mathbf{d} - \bar{\mathbf{d}})^T \end{aligned}$$

5. Filtre de Kalman d'ensemble (EnKF)

Idée Clé : Lissage d'ensemble (ES) – Skjervheim (2011)

L'ES-MDA se compose d'une étape d'initialisation suivie d'étapes répétées de prévision et de mise à jour comme suit :

- 1. Étape d'initialisation.** Un ensemble initial (\mathbf{m}^0) de N réalisations est généré à l'aide de méthodes géostatistiques en intégrant autant de connaissances préalables que possible
- 2. Étape de prévision.** Le modèle direct (F) est résolu complètement pour chaque réalisation i , afin d'obtenir des prédictions du modèle (\mathbf{d}_i) en utilisant la dernière mise à jour des ensembles (\mathbf{m}_{i-1})

$$\mathbf{d}_i = F(\mathbf{m}_{i-1}), \quad \forall i = 1, \dots, N$$

- 3. Étape de mise à jour.** Le vecteur des modèles est mis à jour (\mathbf{m}_i) en fonction des écarts entre les données prédites (\mathbf{d}_i) et observées (\mathbf{d}^0).

$$\begin{aligned} \mathbf{m}_i &= \mathbf{m}_{i-1} + \mathbf{K}(\mathbf{d}^0 - \mathbf{d}_i) \\ \mathbf{K} &= \mathbf{C}_{MD}(\mathbf{C}_{DD} + \mathbf{R})^{-1} \end{aligned}$$

- 4. Retour à l'étape de prévision.**

$$\begin{aligned} \mathbf{C}_{MD} &= \frac{1}{N-1} (\mathbf{m} - \bar{\mathbf{m}})(\mathbf{d} - \bar{\mathbf{d}})^T \\ \mathbf{C}_{DD} &= \frac{1}{N-1} (\mathbf{d} - \bar{\mathbf{d}})(\mathbf{d} - \bar{\mathbf{d}})^T \end{aligned}$$

5. Filtre de Kalman d'ensemble (EnKF)

Idée Clé : Lissage d'ensemble avec assimilation de données multiples (ES-MDA) – Emerick and Reynolds (2013)

L'ES-MDA se compose d'une étape d'initialisation suivie d'étapes répétées de prévision et de mise à jour comme suit :

- 1. Étape d'initialisation.** Un ensemble initial (\mathbf{m}^0) de N réalisations est généré à l'aide de méthodes géostatistiques en intégrant autant de connaissances préalables que possible
- 2. Étape de prévision.** Le modèle direct (F) est résolu complètement pour chaque réalisation i , afin d'obtenir des prédictions du modèle (\mathbf{d}_i) en utilisant la dernière mise à jour des ensembles (\mathbf{m}_{i-1})

$$\mathbf{d}_i = F(\mathbf{m}_{i-1}), \quad \forall i = 1, \dots, N$$

- 3. Étape de mise à jour.** Le vecteur des modèles est mis à jour (\mathbf{m}_i) en fonction des écarts entre les données prédites (\mathbf{d}_i) et observées (\mathbf{d}^0).

$$\mathbf{m}_i = \mathbf{m}_{i-1} + \mathbf{K}_t (\mathbf{d}^0 + \sqrt{\alpha} \boldsymbol{\varepsilon}_i^0 - \mathbf{d}_i)$$
$$\mathbf{K}_t = \mathbf{C}_{MD} (\mathbf{C}_{DD} + \alpha \mathbf{R})^{-1}$$

- 4. Retour à l'étape de prévision.**

$$\mathbf{C}_{MD} = \frac{1}{N-1} (\mathbf{m} - \bar{\mathbf{m}})(\mathbf{d} - \bar{\mathbf{d}})^T$$
$$\mathbf{C}_{DD} = \frac{1}{N-1} (\mathbf{d} - \bar{\mathbf{d}})(\mathbf{d} - \bar{\mathbf{d}})^T$$