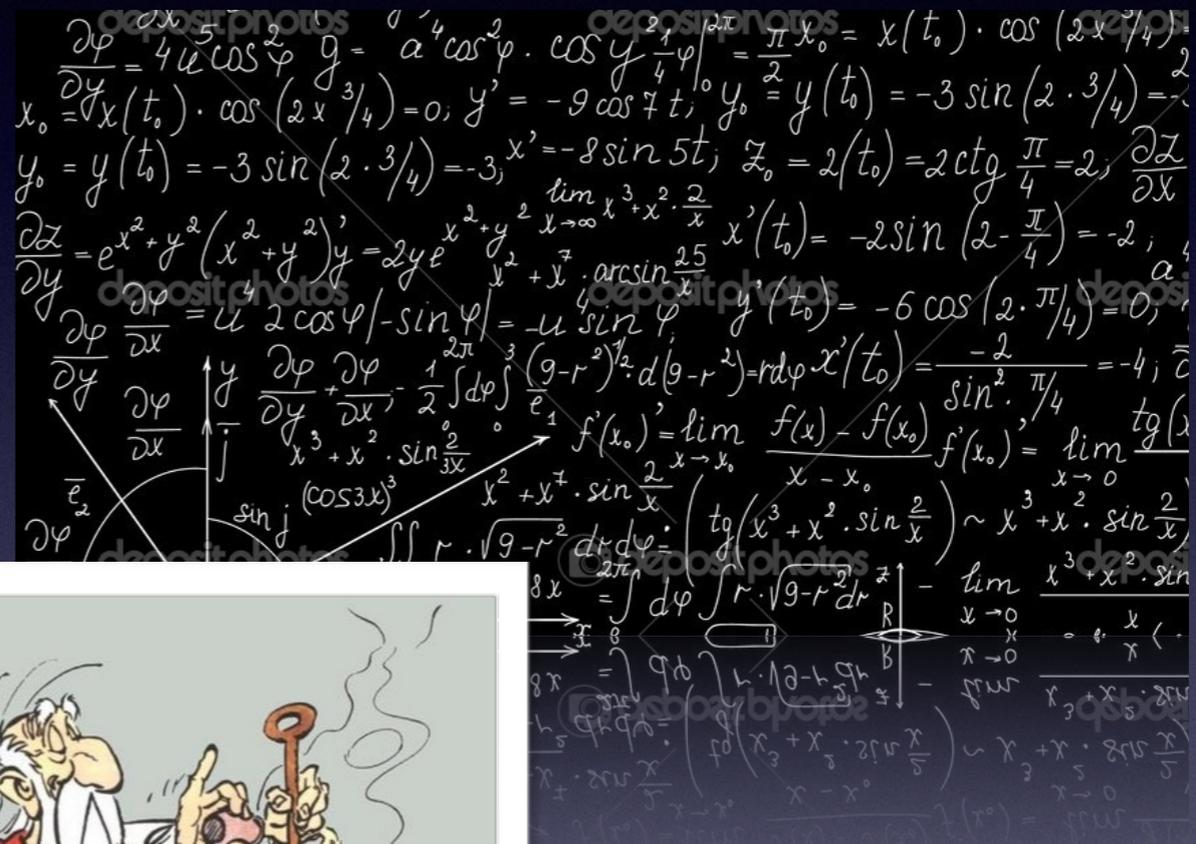
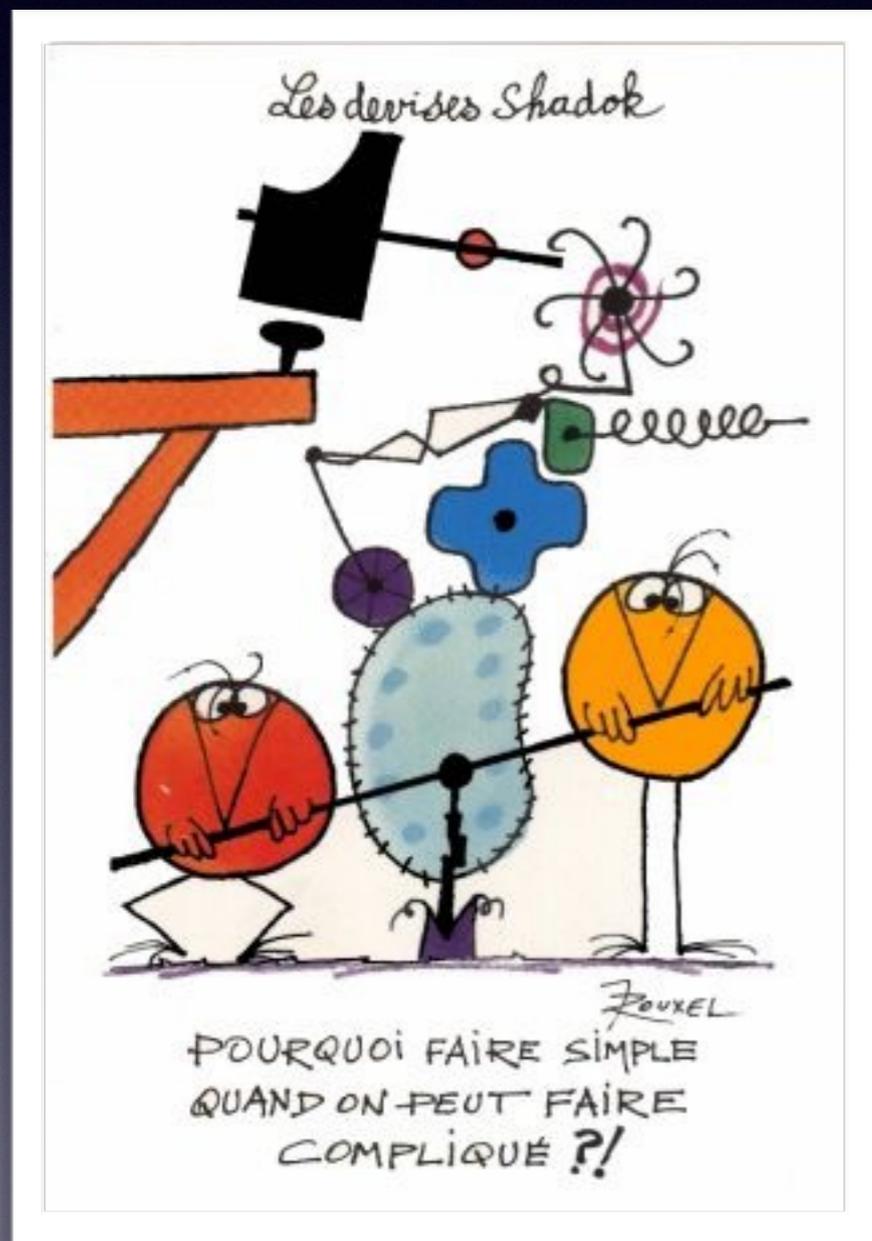
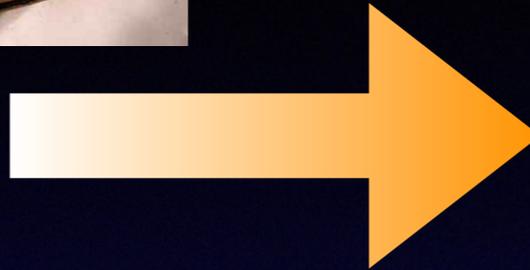
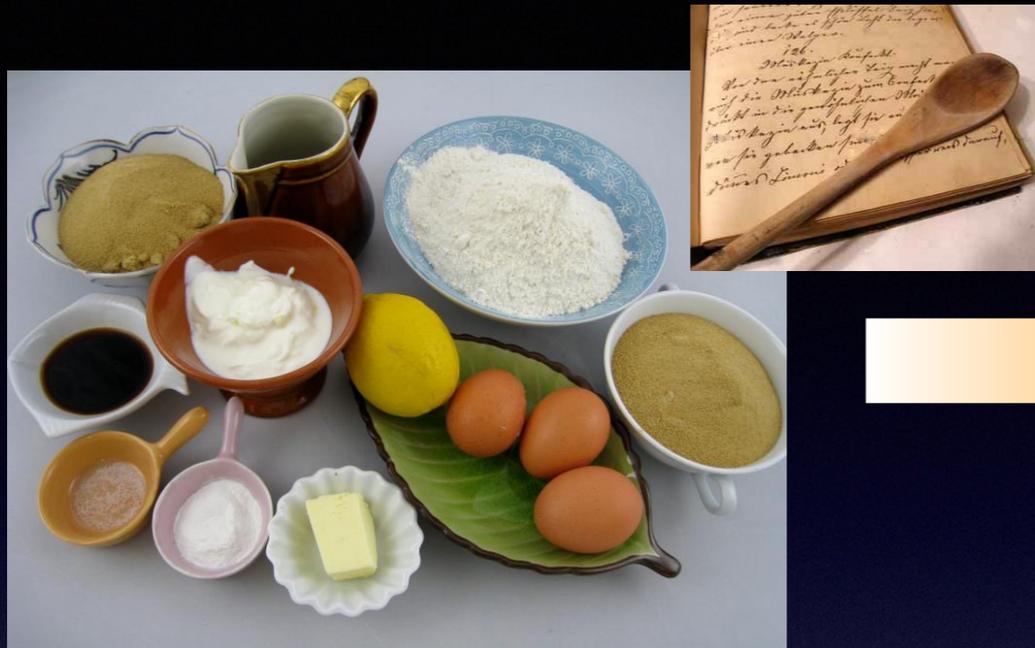


# Introduction au problème inverse



Le problème direct, c'est...



Le problème inverse, c'est...



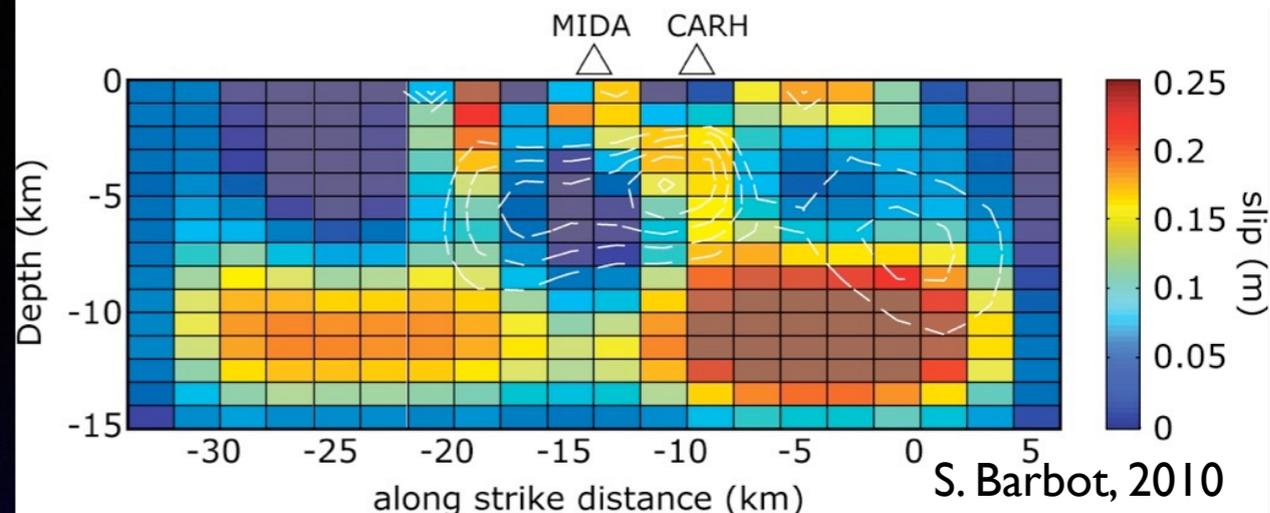
# Les méthodes inverses en géophysique

- Elles sont partout! Car on observe, et on en déduit (on propose) un modèle pour expliquer les observations.
  - La tomographie,
  - Le GPS,
  - La géochimie,
  - La géologie structurale...

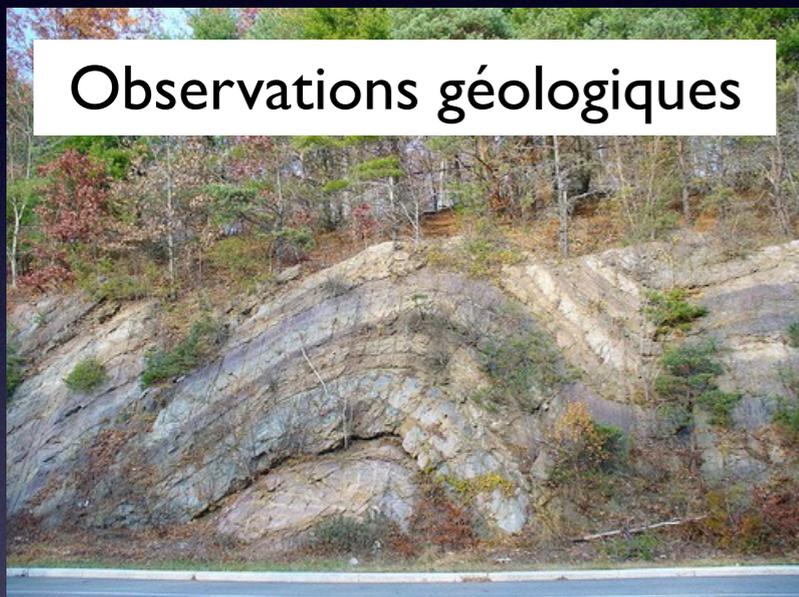
# Mesures de déformation en surface (GPS)



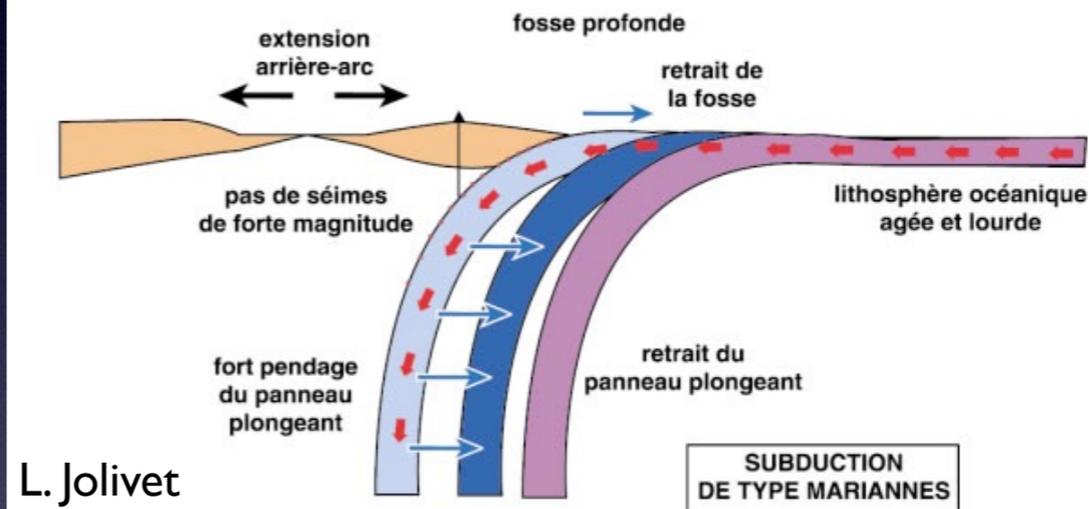
# Glissement cosismique sur une faille



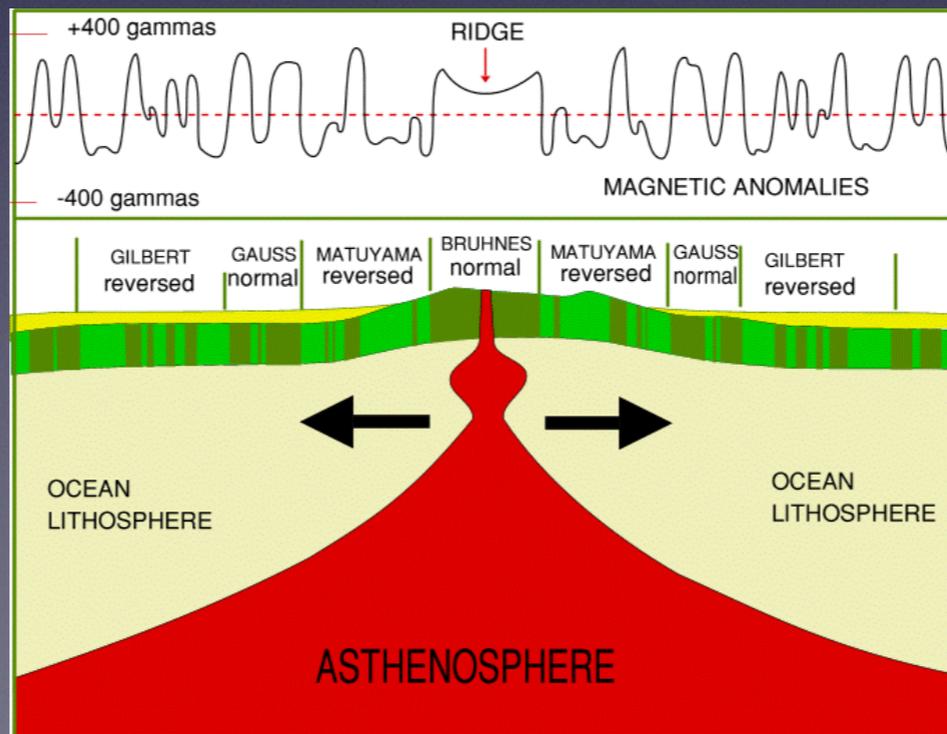
# Observations géologiques



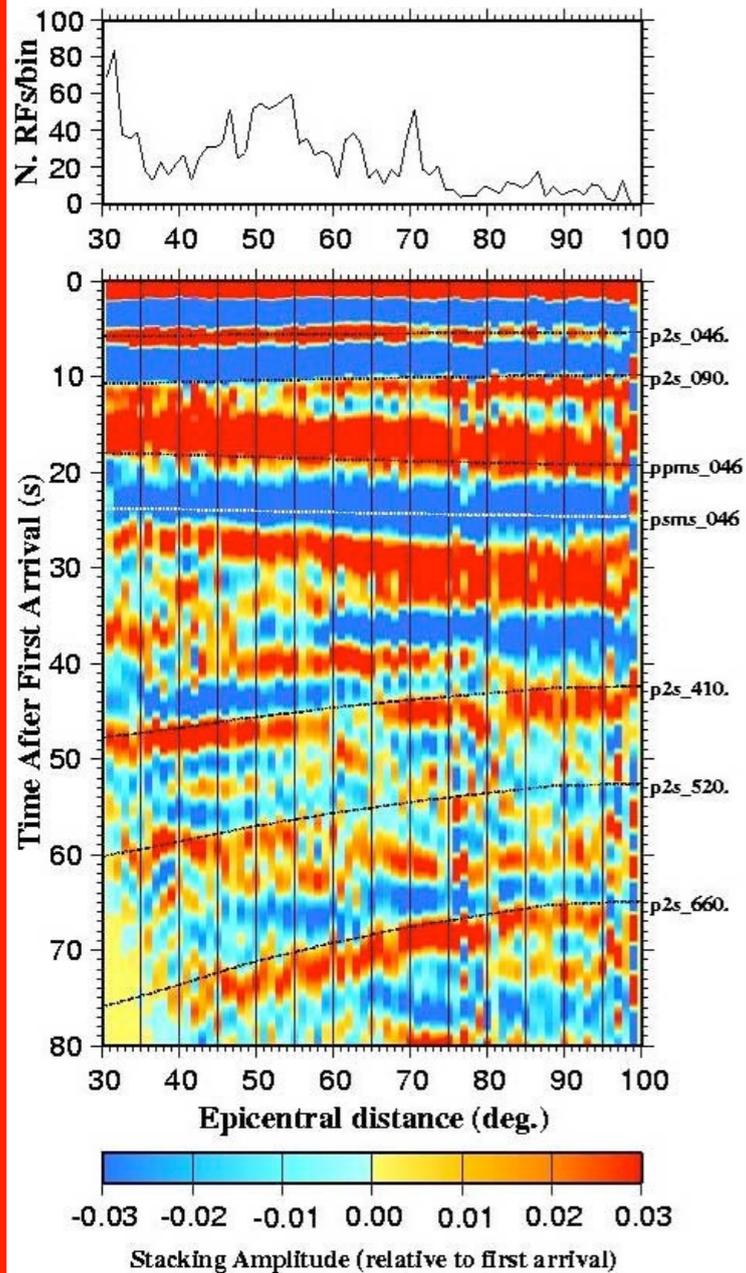
# modèle géodynamique



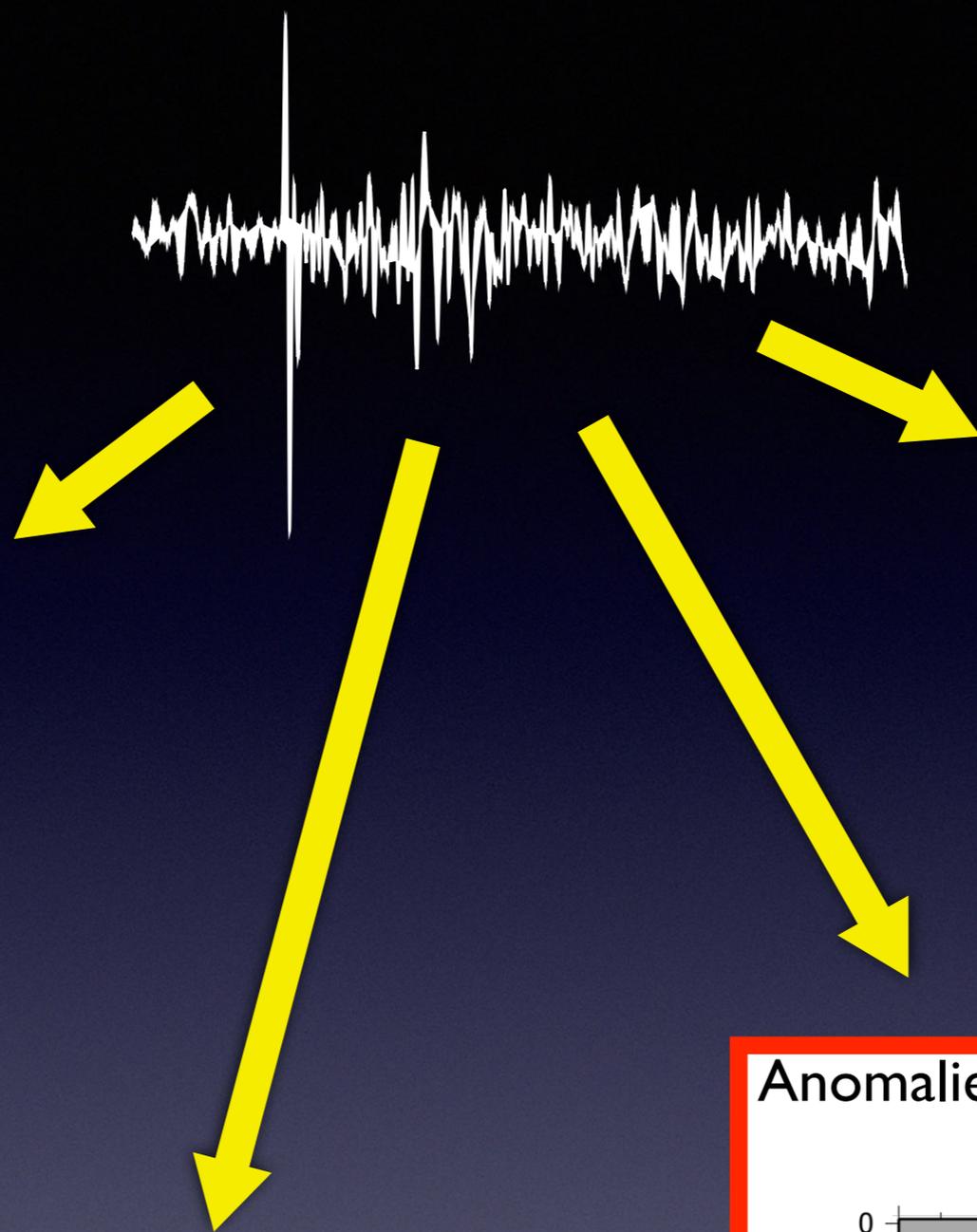
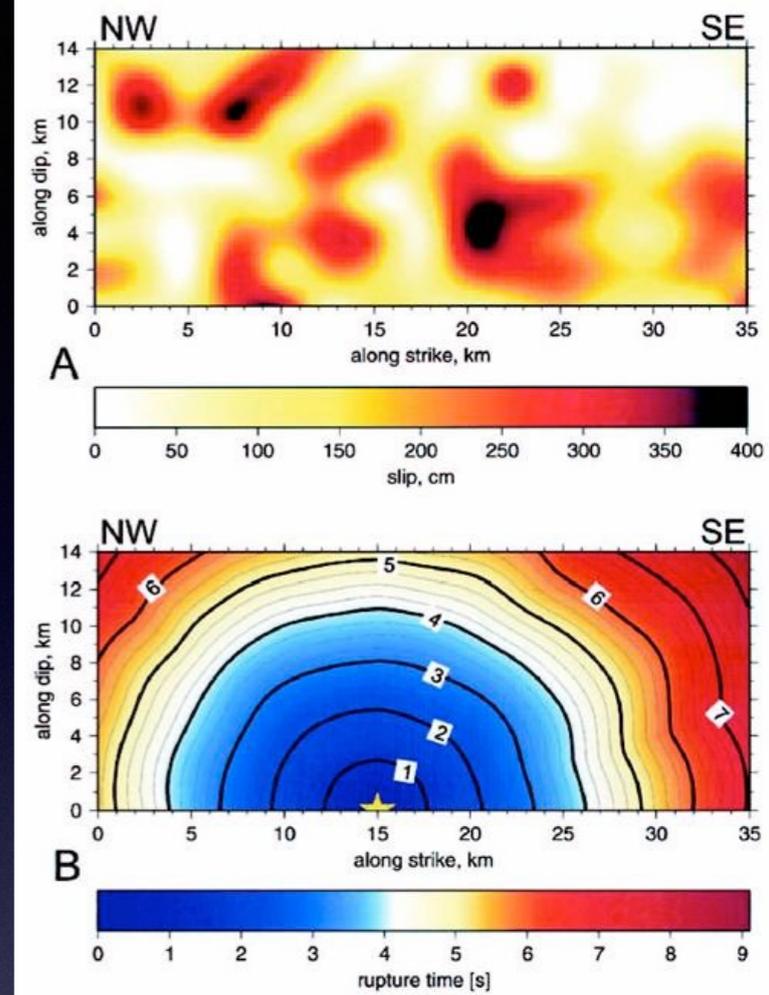
# Anomalies magnétiques



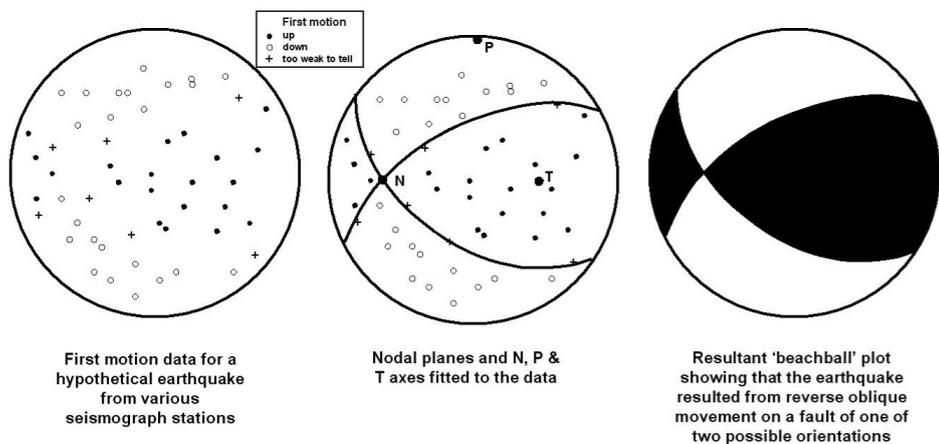
# Imager les interfaces



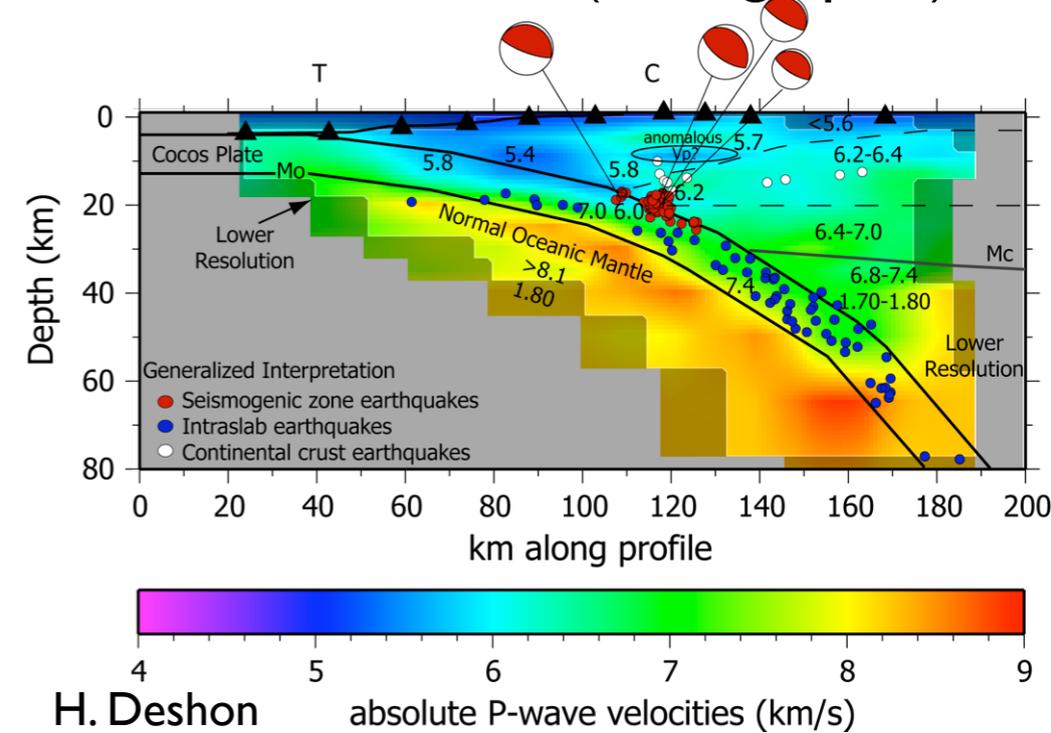
# Imager la rupture



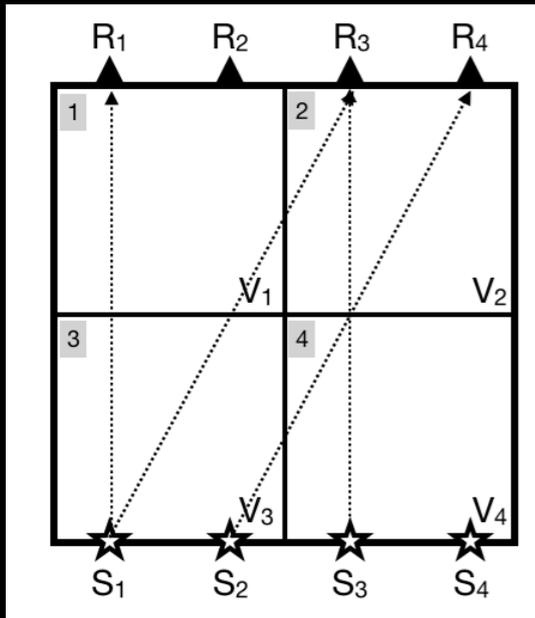
# Mécanisme au foyer



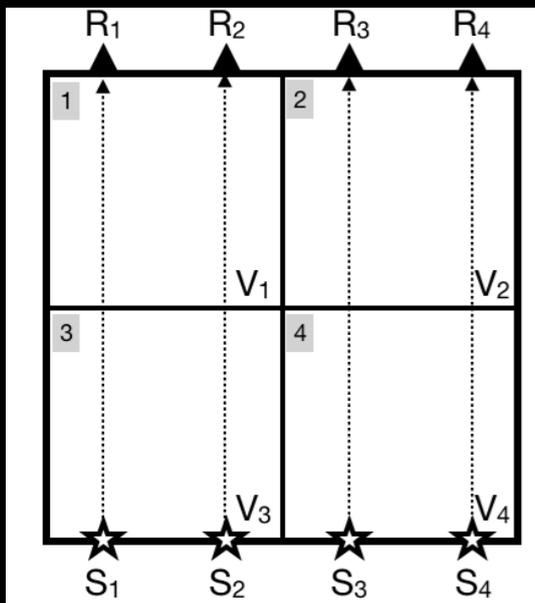
# Anomalies de vitesses (tomographie)



# Problème inverse = résolution de système d'équations



$$\left\{ \begin{array}{l} t_1 = \frac{d_{13}}{V_3} + \frac{d_{11}}{V_1} \\ t_2 = \frac{d_{23}}{V_3} + \frac{d_{21}}{V_1} + \frac{d_{22}}{V_2} \\ t_3 = \frac{d_{33}}{V_3} + \frac{d_{34}}{V_4} + \frac{d_{32}}{V_2} \\ t_4 = \frac{d_{44}}{V_4} + \frac{d_{42}}{V_2} \end{array} \right.$$

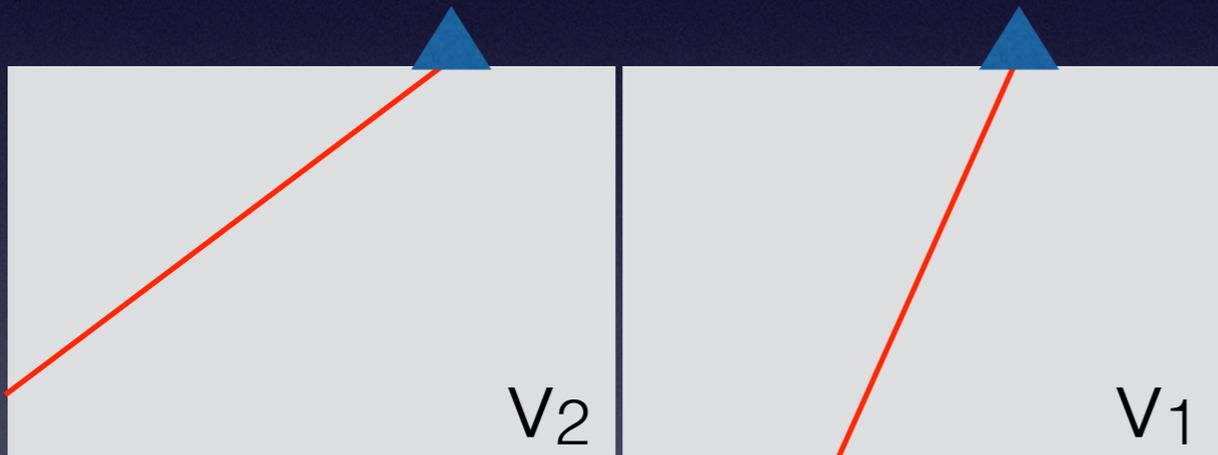


$$\left\{ \begin{array}{l} t_1 = t_2 = \frac{d_3}{V_3} + \frac{d_1}{V_1} \\ t_3 = t_4 = \frac{d_4}{V_4} + \frac{d_2}{V_2} \end{array} \right.$$

# Résumé

Exemple : Déterminer la vitesse de chacun des blocs

$$V = D / t$$

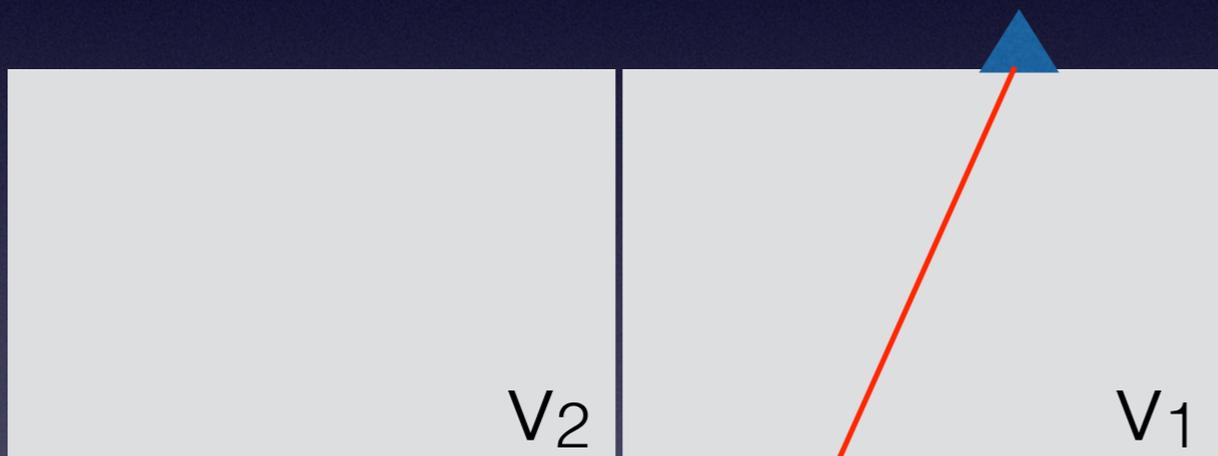


Problème  
**bien**  
**déterminé**

# Résumé

Exemple : Déterminer la vitesse de chacun des blocs

$$V = D / t$$

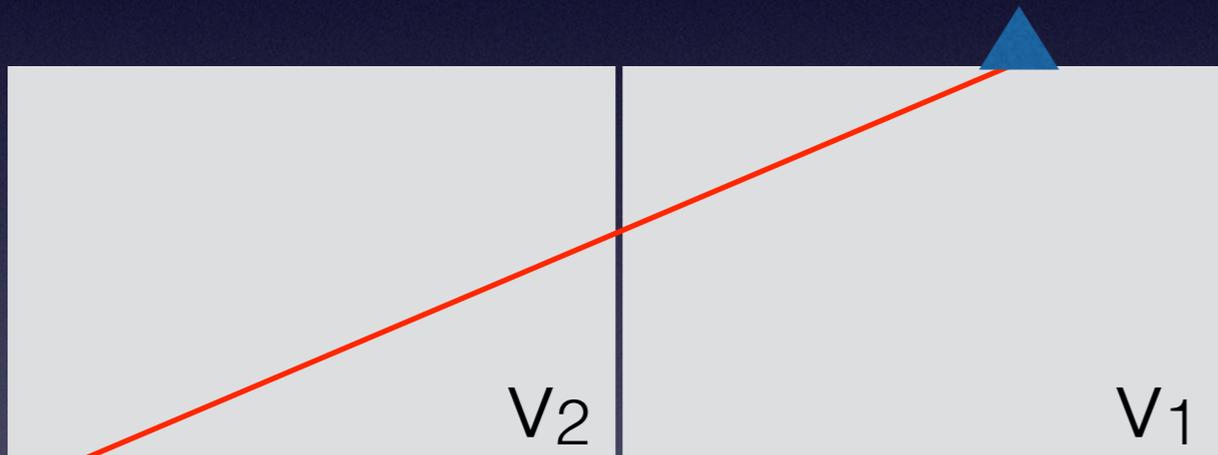


Problème  
**sous-**  
**déterminé**

# Résumé

Exemple : Déterminer la vitesse de chacun des blocs

$$V = D / t$$

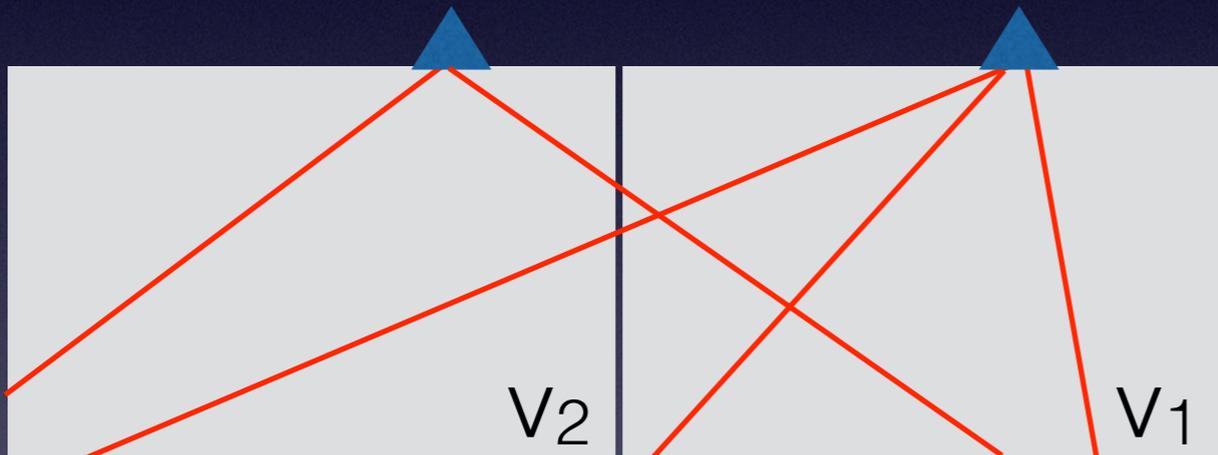


Problème  
**sous-**  
**déterminé**

# Résumé

Exemple : Déterminer la vitesse de chacun des blocs

$$V = D / t$$

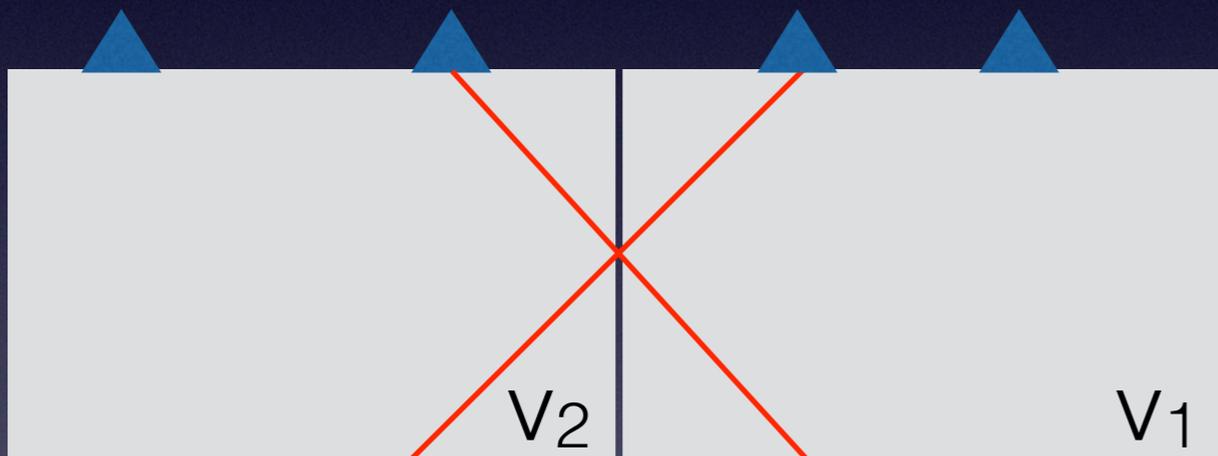


Problème  
**sur-**  
**déterminé**

# Résumé

Exemple : Déterminer la vitesse de chacun des blocs

$$V = D / t$$

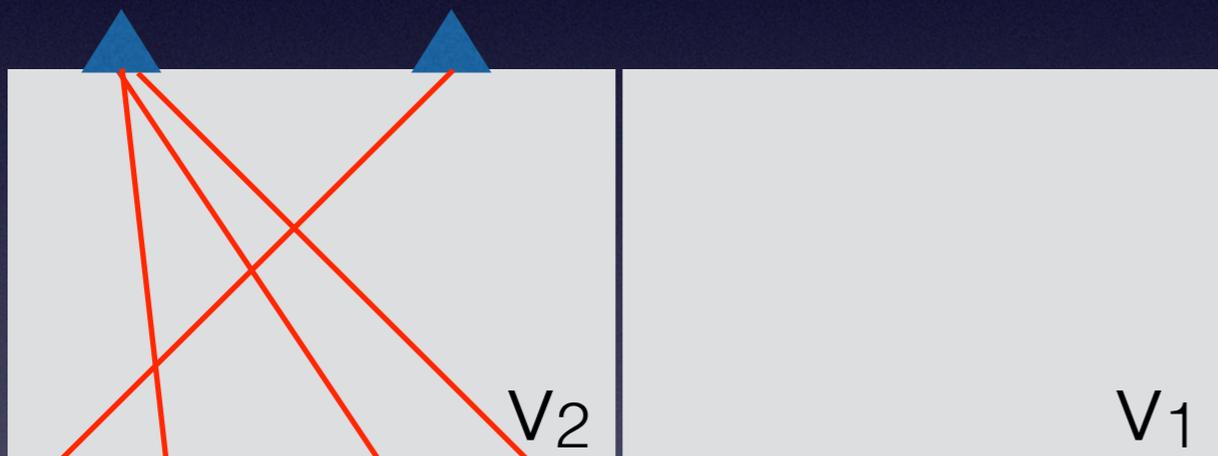


Problème  
**sous-**  
**déterminé**

# Résumé

Exemple : Déterminer la vitesse de chacun des blocs

$$V = D / t$$



Problème  
**mixte**

Souvent le cas en géophysique...

- On pallie le manque de données par des info *a priori*
- On résout le surplus de données par des approximations

# Définitions

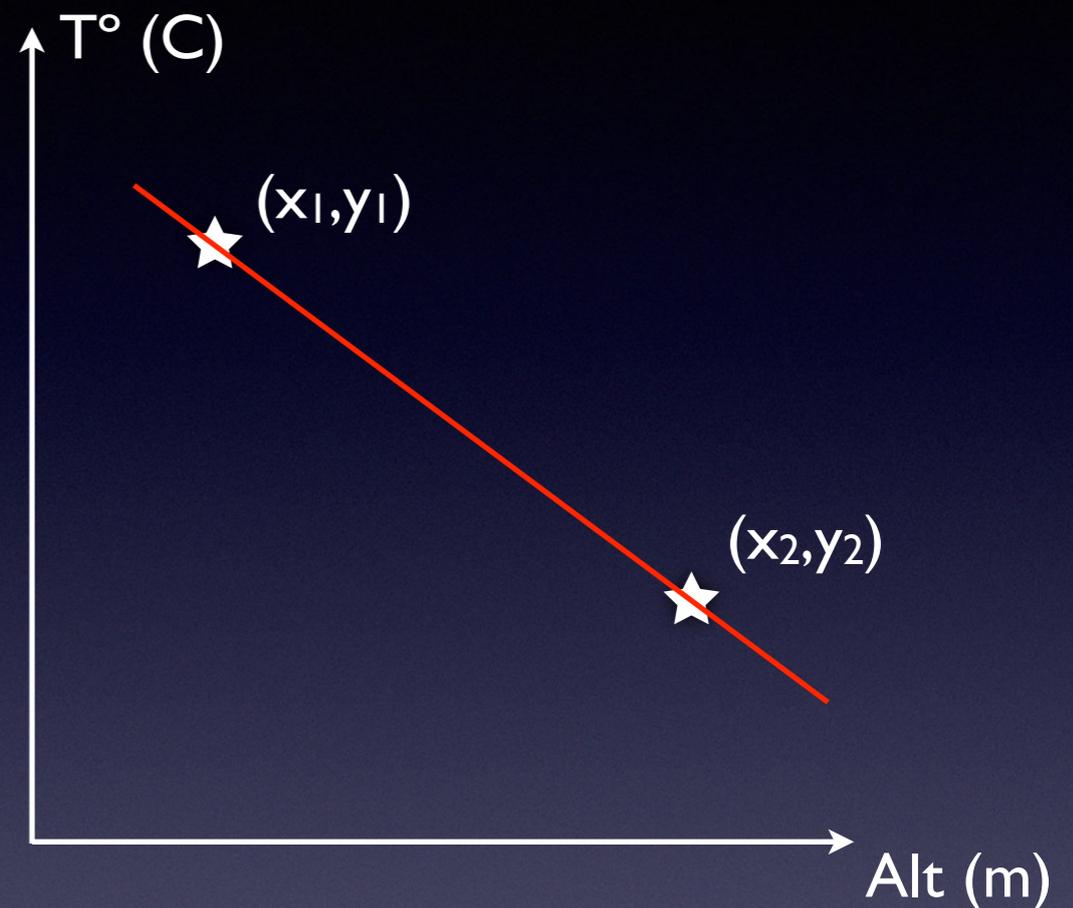
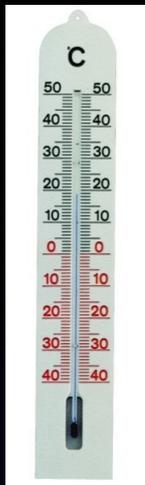
Problème **sur-déterminé**      Données  $>$  paramètres

Problème **sous-déterminé**      Données  $<$  paramètres

Problème **mixte**      Les deux précédents

Problème **bien déterminé**      Données  $=$  paramètres

# Autre exemple



Loi physique:  $y = a.x + b$

Quels sont les valeurs de  $a$  et  $b$  ?

$$y_1 = a.x_1 + b$$

$$y_2 = a.x_2 + b$$

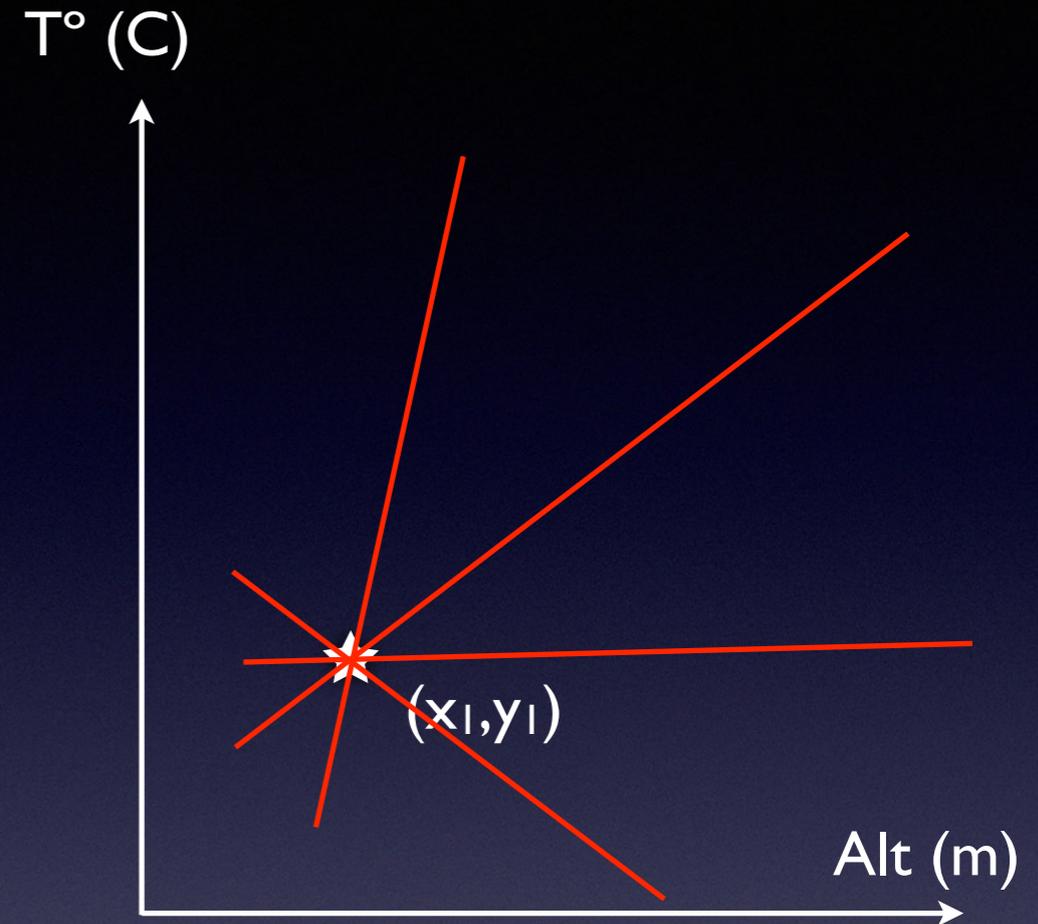


$$a = (y_1 - y_2) / (x_1 - x_2)$$

$$b = y_1 - (y_1 - y_2).x_1 / (x_1 - x_2)$$

Oui, mais...

# Autre exemple



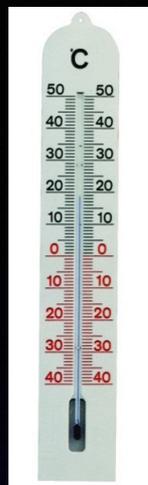
Quels sont les valeurs de  $a$  et  $b$  ?

$$y_1 = a.x_1 + b \quad \longrightarrow \quad \begin{aligned} a &= (y_1 - b) / x_1 \\ b &= y_1 - a.x_1 \end{aligned}$$

1 donnée, 2 paramètres...  $a$  et  $b$  ne peuvent pas être déterminés de façon unique. Infinité de solutions!

Problème **sous-déterminé**

# Autre exemple



$$y_1 = a.x_1 + b$$

$$y_2 = a.x_2 + b$$

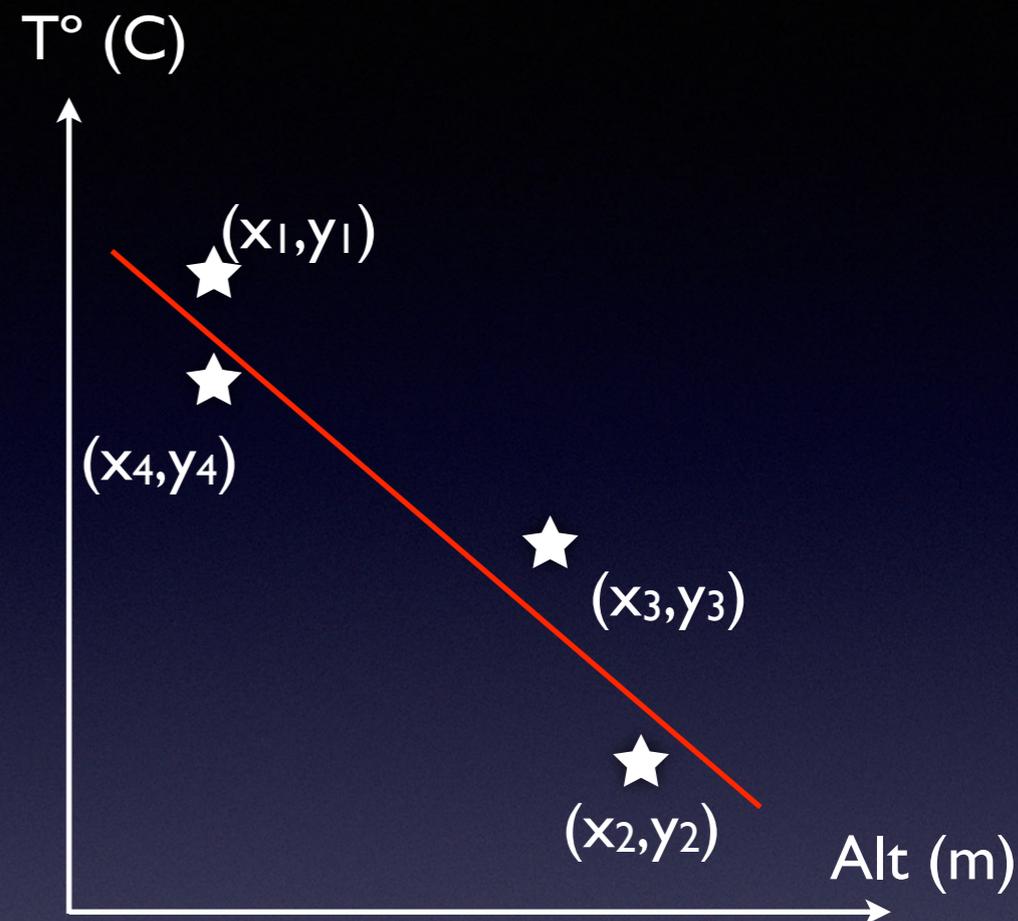
$$y_3 = a.x_3 + b$$

$$y_4 = a.x_4 + b$$



$$a = ??$$

$$b = ??$$

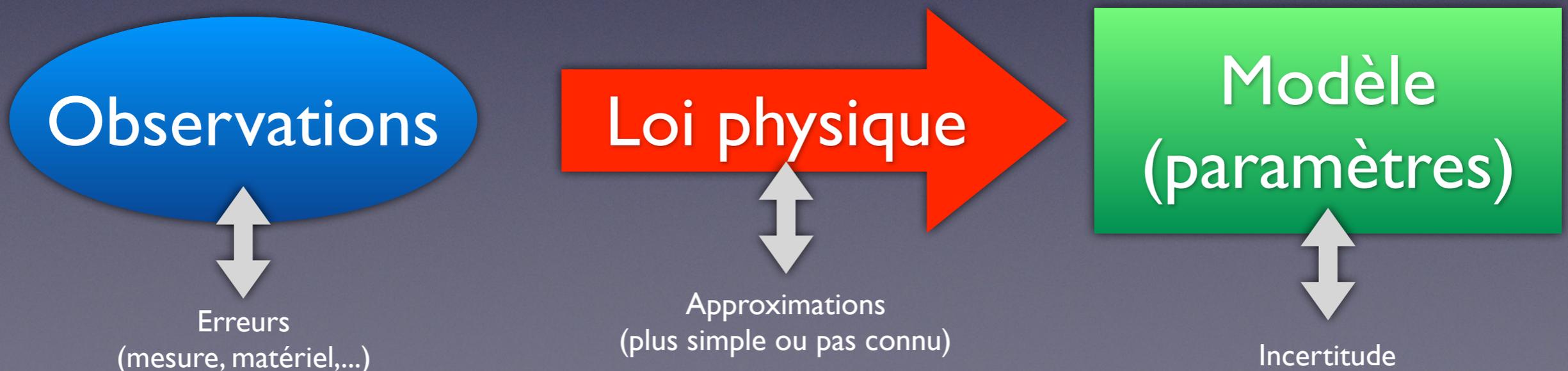


4 données, 2 paramètres... a et b ne peuvent être déterminés de façon unique que si les données sont cohérentes. Sinon, on doit approximer.

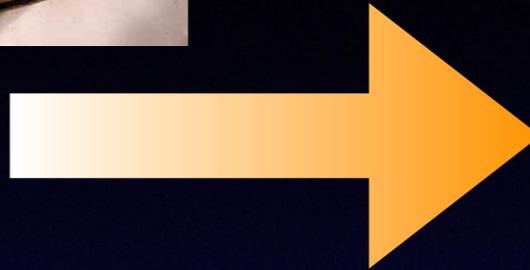
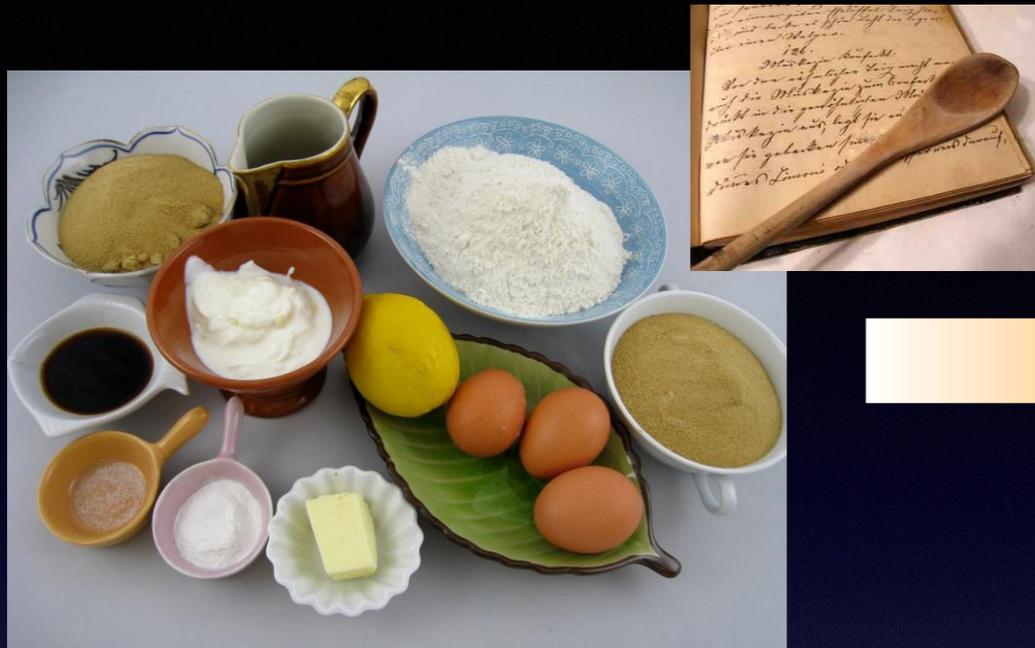
Problème **sur-déterminé**

# Problème inverse

- On récolte des observations (données)
- On en déduit un modèle via des lois (mathématiques, physiques,...)
- Ce modèle a une probabilité d'existence, une incertitude (qu'il faut estimer)



Le problème direct, c'est...

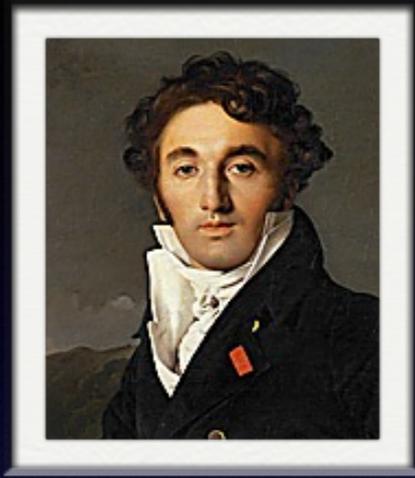


Le problème inverse, c'est...

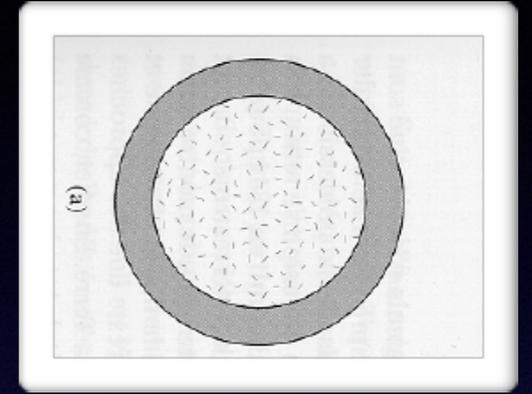


# Un seul modèle ?

1827



Cordier (1777-1861) : Gradient de température dans les mines =  $1^{\circ}\text{C} / 25 \text{ m}$   
Donc, à 50 km de profondeur, on atteint  $1600^{\circ}\text{C}$   
i.e. intérieur de la Terre = roches fondues !



gradient  $1^{\circ}/25\text{m}$

$f(T)$

liquide / solide

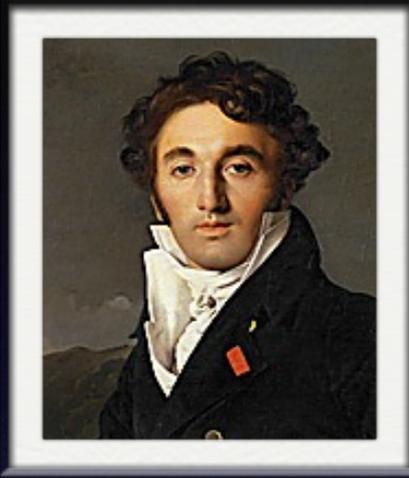
Observations

Loi physique

Modèle  
(paramètres)

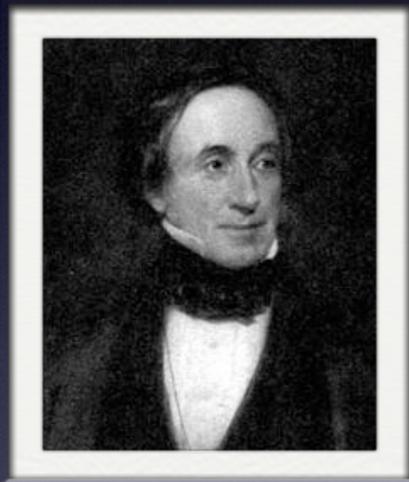
# Exemple I

1827



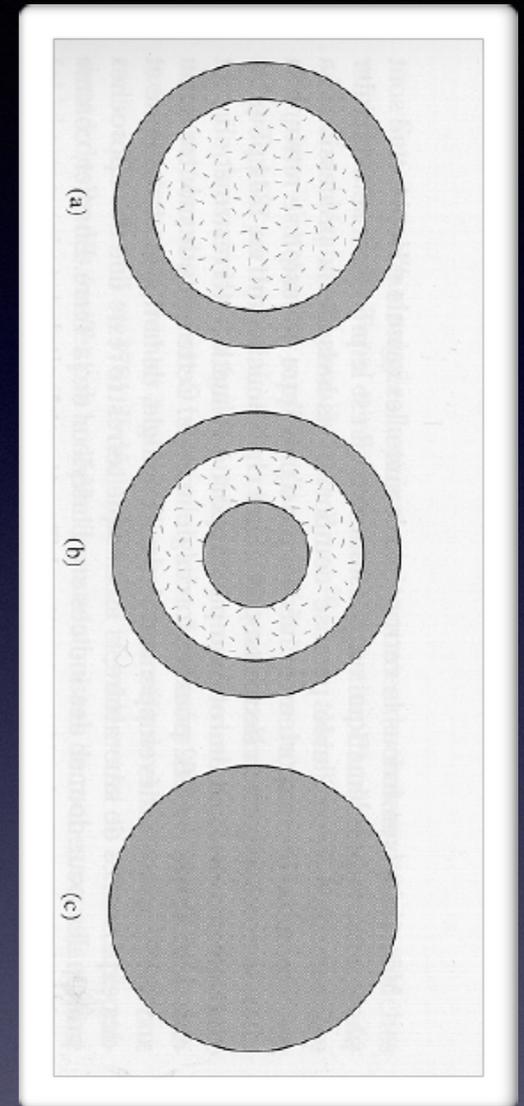
Cordier (1777-1861) : Gradient de température dans les mines =  $1^{\circ}\text{C} / 25 \text{ m}$

Donc, à 50 km de profondeur, on atteint  $1600^{\circ}\text{C}$  i.e. intérieur de la Terre = roches fondues !



Hopkins (1793-1866) : Etat des roches **dépend aussi de la pression...**

Expérience en laboratoire impossible, émet donc 3 hypothèses...



gradient  $1^{\circ}/25\text{m}$

$f(T,P)$

liquide / solide

Observations

Loi physique

Modèle  
(paramètres)

# Moralité...

Suivant les hypothèses de départ ou la théorie utilisée (approximations), les résultats diffèrent.

En inversion, on n'obtient toujours qu'un **modèle**, une représentation imagée de la réalité. Elle peut donc:

- être non-unique,
- être fausse,
- ne pas coller avec toutes les données.

Dans tous les cas, le(s) modèle(s) est dépendant des **données** et des **hypothèses théoriques**.

# Résoudre un pb inverse...

Revient à :

→ comparer les données (réelles)  
avec le modèle

→ minimiser l'erreur entre données  
et modèle

$$y = a.x + b$$

$(x_d, y_d)$

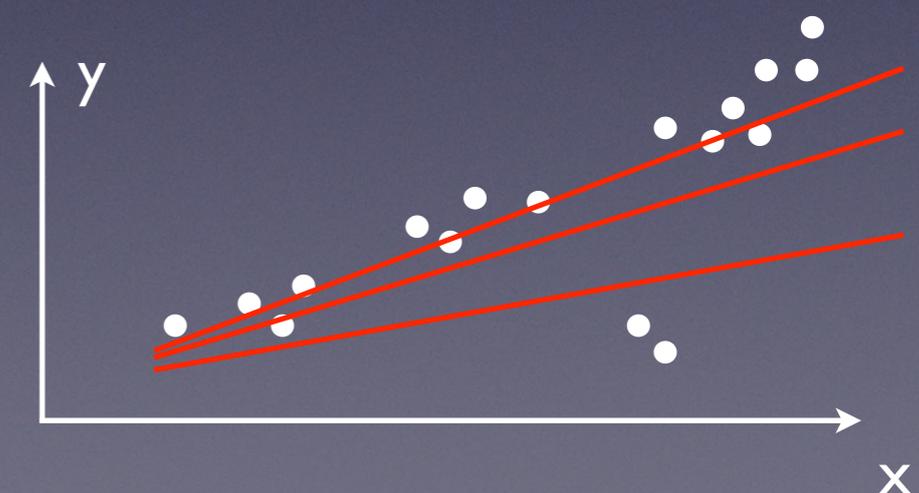
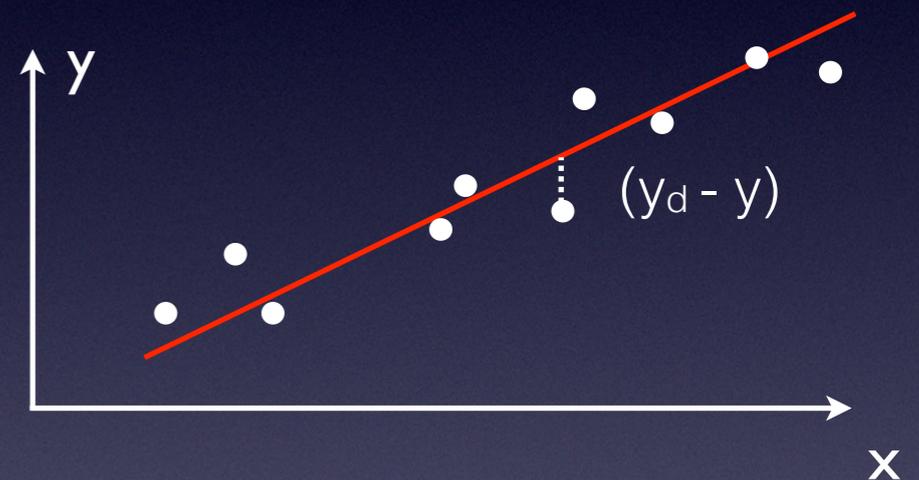
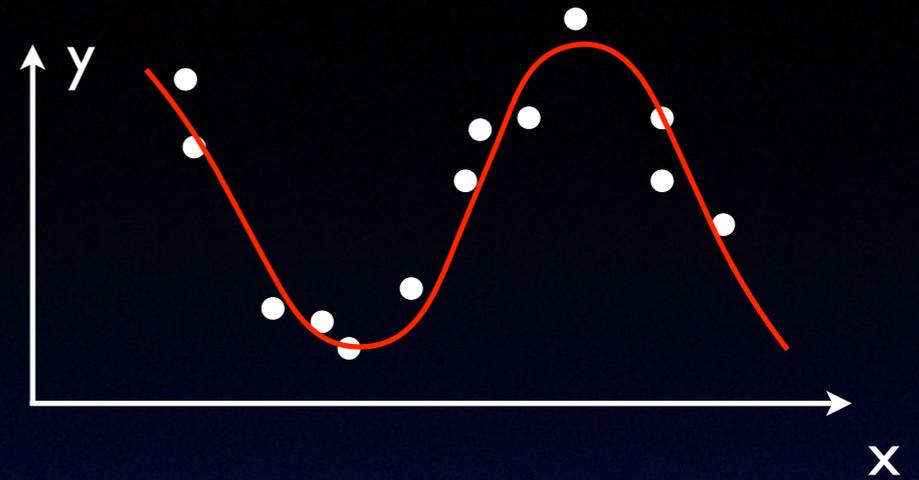
théorie  
données

Alors minimisons  $\sum(y_d - y)$

ou  $\sum(y_d - y)^2$

ou  $\sum(y_d - y)^3 \dots$

Bien choisir la norme!

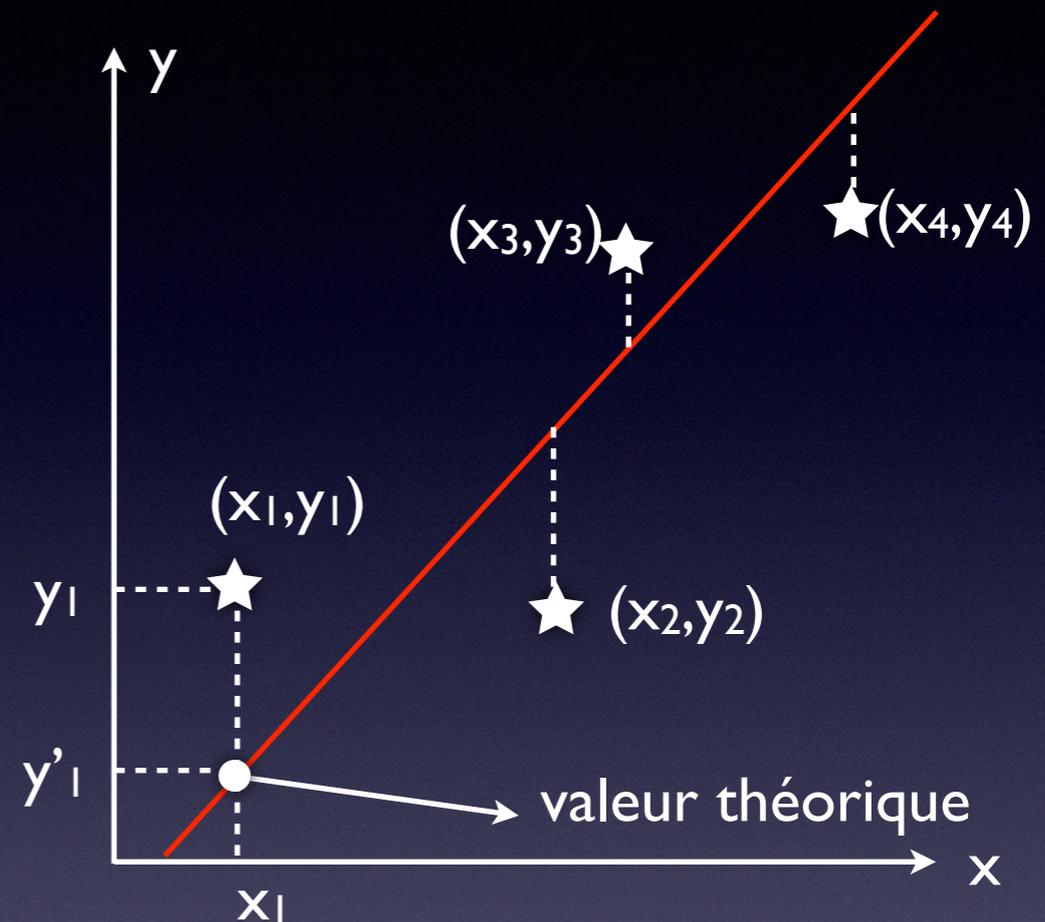


# Résoudre un pb inverse...

$$S = \sum_{i=1}^n \sqrt{(y_i - y'_i)^2} = \sqrt{(y_i - a.x_i - b)^2} = \min$$

Minimiser  $\sum [y - (ax + b)]^2$  signifie que ... sa dérivée est nulle !

$$\begin{cases} \frac{\partial J}{\partial a} = -2 \sum_{i=1}^n x_i (y_i - ax_i - b) = 0, \\ \frac{\partial J}{\partial b} = -2 \sum_{i=1}^n (y_i - ax_i - b) = 0. \end{cases}$$



$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \text{ et } b = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}.$$

Norme L2 permet de pondérer les points déviants

Solution des moindres carrés

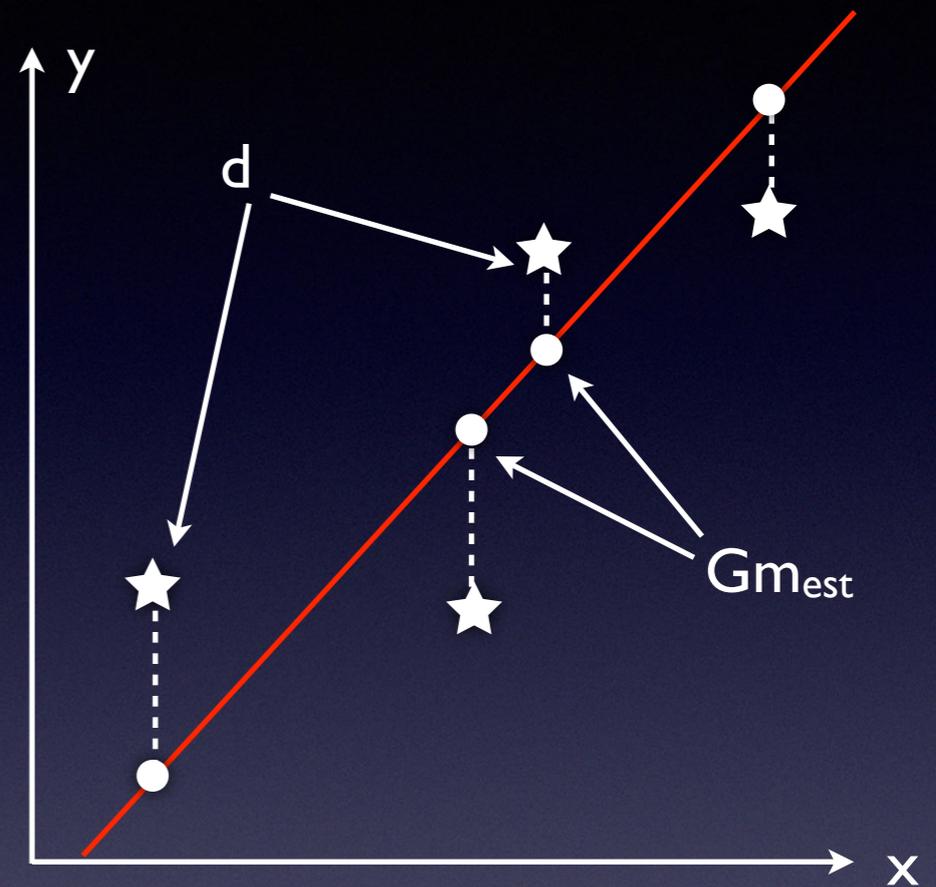


J. Gauss (1777-1855)

# Généralisation

- Ecriture matricielle
  - observables (données) :  $d$
  - paramètres (modèle) :  $m$
  - Loi physique reliant les 2 :  $G$

$$d = Gm \quad \text{Problème direct}$$



Habituellement, système sur-déterminé ( $d > m$ ),  
il n'existe donc pas de solution unique au  
problème  $d - Gm = 0$

$d - Gm_{est}$  doit être **minimum**

# Généralisation

$d - Gm_{\text{est}}$  doit être **minimum**

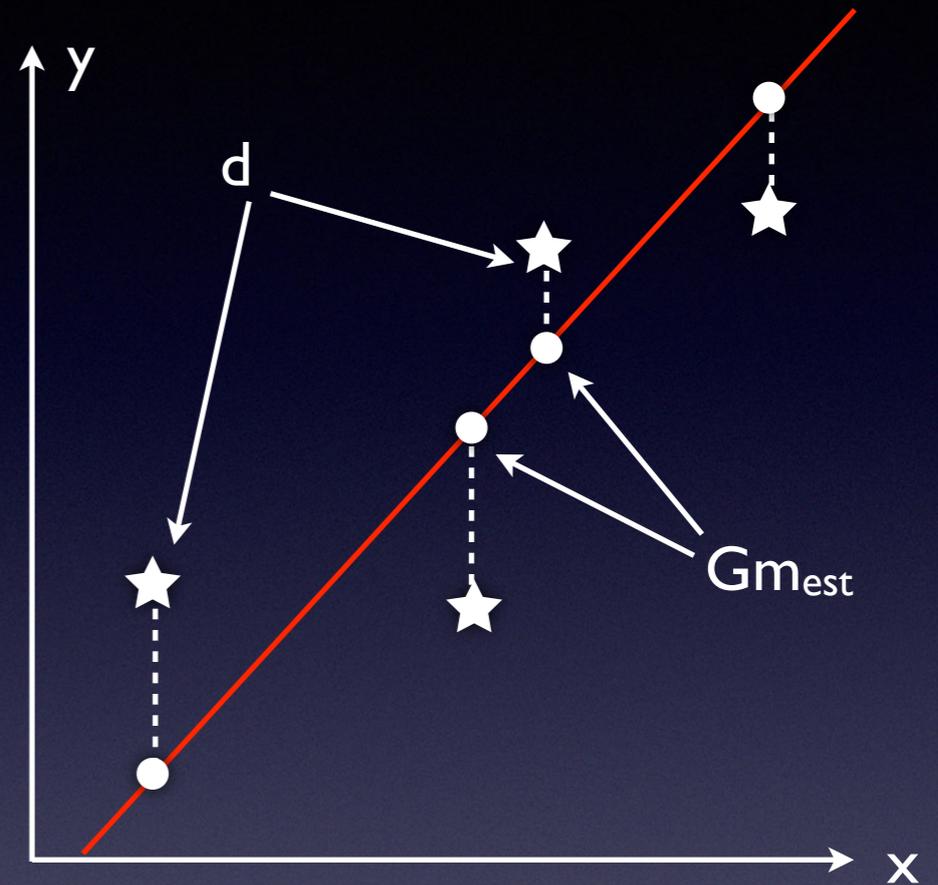
$E = (d - Gm_{\text{est}})^T (d - Gm_{\text{est}})$  doit être  
minimale

$$dE/dm = 0,$$

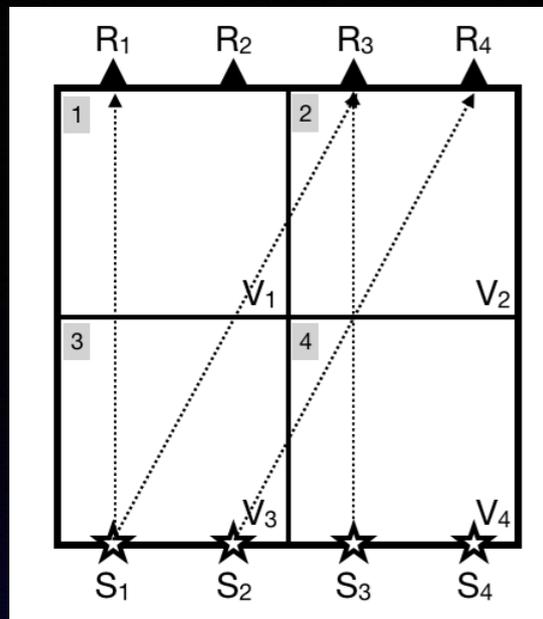
$$G^T G m_{\text{est}} - G^T d = 0$$

$(G^T G)^{-1} =$  inverse généralisé

$$m_{\text{est}} = (G^T G)^{-1} G^T d$$



Les moindres carrés



$$m_{\text{est}} = (G^T G)^{-1} G^T d$$

$$\begin{cases} t_1 = \frac{d_{13}}{V_3} + \frac{d_{11}}{V_1} \\ t_2 = \frac{d_{23}}{V_3} + \frac{d_{21}}{V_1} + \frac{d_{22}}{V_2} \\ t_3 = \frac{d_{33}}{V_3} + \frac{d_{34}}{V_4} + \frac{d_{32}}{V_2} \\ t_4 = \frac{d_{44}}{V_4} + \frac{d_{42}}{V_2} \end{cases}$$

$$m = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{pmatrix}$$

$$G = \begin{pmatrix} d_{11} & 0 & d_{13} & 0 \\ d_{21} & d_{22} & d_{23} & 0 \\ 0 & d_{32} & d_{33} & d_{34} \\ 0 & d_{42} & 0 & d_{44} \end{pmatrix}$$

$$d = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{pmatrix}$$

$$G^T = \begin{pmatrix} d_{11} & d_{21} & 0 & 0 \\ 0 & d_{22} & d_{32} & d_{42} \\ d_{13} & d_{23} & d_{33} & 0 \\ 0 & 0 & d_{34} & d_{44} \end{pmatrix}$$

# Généralisation

$$m_{\text{est}} = (G^T G)^{-1} G^T d$$

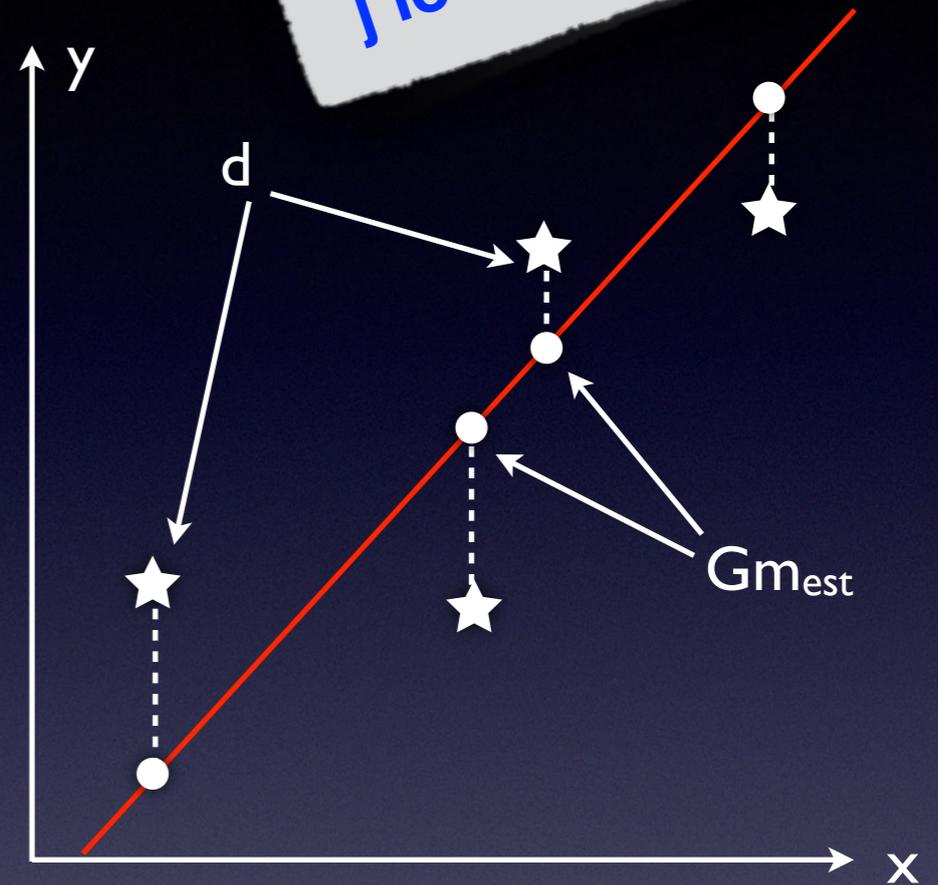
Pour améliorer les choses....

Introduction d'un **modèle initial**  
(ou *a priori*, ou de départ)

$$m_{\text{est}} = m_0 + (G^T G)^{-1} G^T (d - G m_0)$$

Permet de réduire les écarts (minimisation optimisée)  
Mais uniquement si "proche" de la solution ... (cf. TD2)  
(e.g. coefficient  $a > 0$ )

Calculs balaises, mais  
j'le fais qu'une fois...



Fin TD1

# Reprise problème inverse...

Les moindres carrés

$$m_{\text{est}} = (G^T G)^{-1} G^T d$$

Introduction d'un **modèle initial** (ou *a priori*, ou de départ)

$$m_{\text{est}} = m_0 + (G^T G)^{-1} G^T (d - G m_0)$$

Efficace si “proche” de la solution ...

Réduit les écarts si solution approximée

Peu d'effet si information complète (bon croisement de rais)

Introduit une information là où  $d=0$

# Reprise problème inverse...

Les moindres carrés

$$m_{\text{est}} = (G^T G)^{-1} G^T d$$

Introduction d'un **modèle initial** (ou *a priori*, ou de départ)

$$m_{\text{est}} = m_0 + (G^T G)^{-1} G^T (d - G m_0)$$

Quid des erreurs sur les données ?

- modifient les solutions
- peuvent se propager loin dans le modèle
- moins influentes si information bien répartie

# Reprise problème inverse...

Les moindres carrés

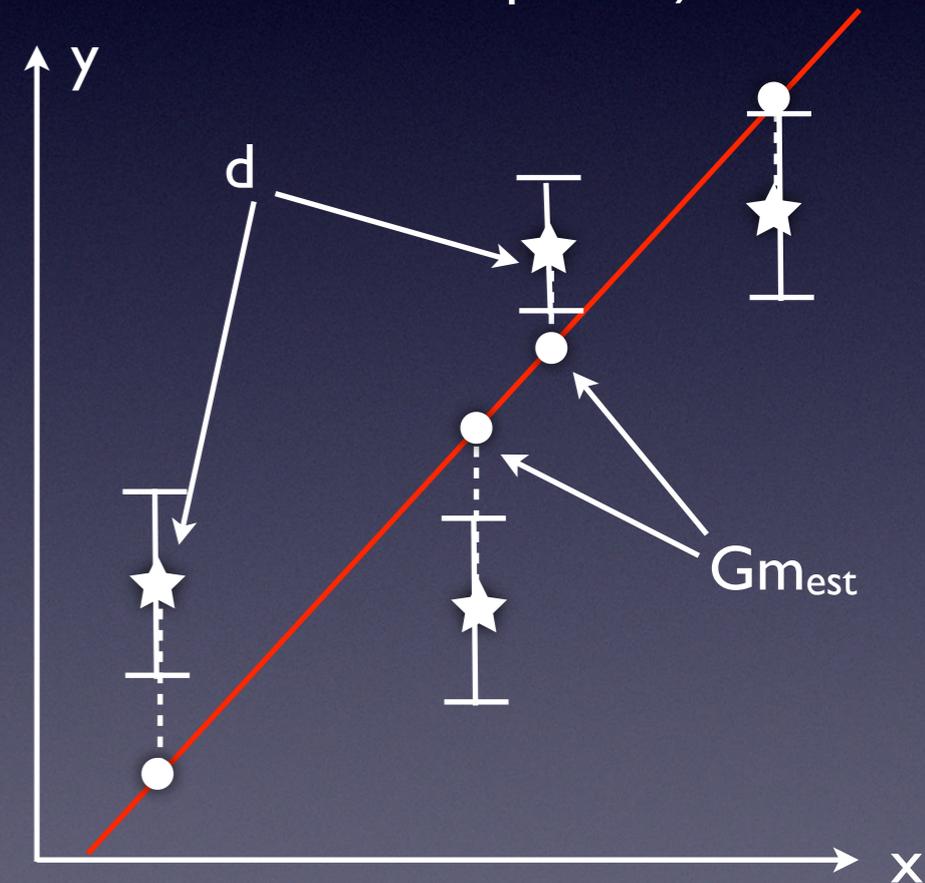
$$m_{\text{est}} = (G^T G)^{-1} G^T d$$

Introduction d'un **modèle initial** (ou *a priori*, ou de départ)

$$m_{\text{est}} = m_0 + (G^T G)^{-1} G^T (d - G m_0)$$

Quid des erreurs sur les données ?

Introduction d'écart-types dans les moindres carrés...



# Reprise problème inverse...

Les moindres carrés

$$m_{\text{est}} = (G^T G)^{-1} G^T d$$

Introduction d'un **modèle initial** (ou *a priori*, ou de départ)

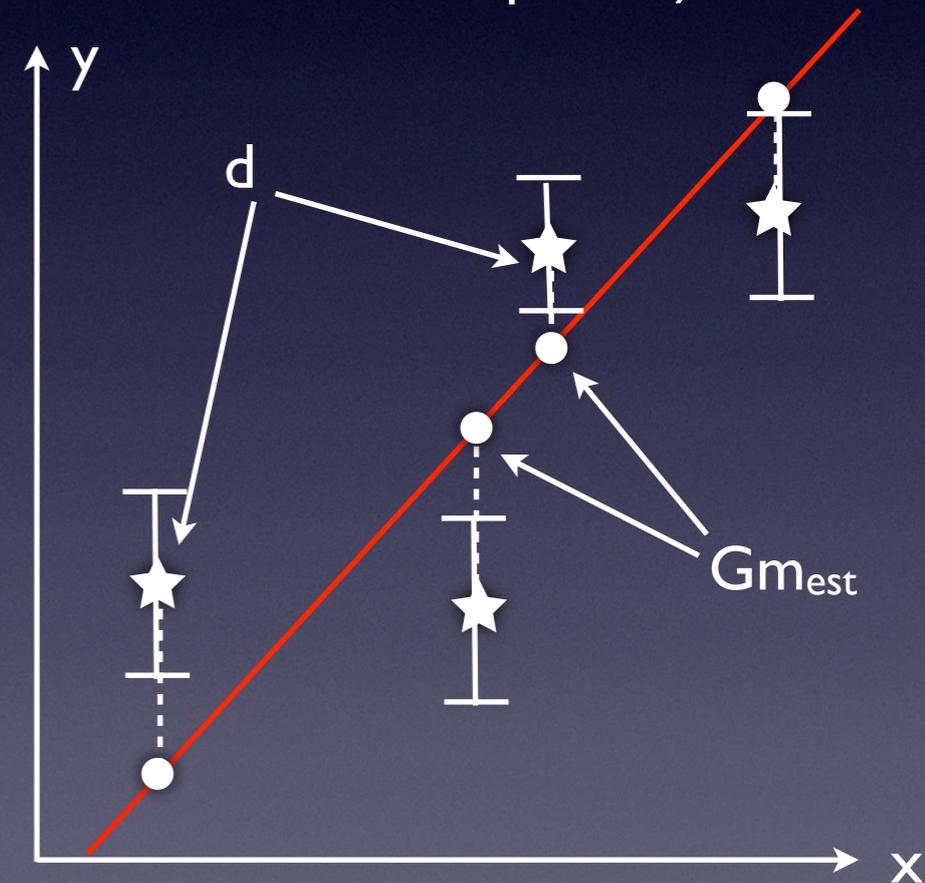
$$m_{\text{est}} = m_0 + (G^T G)^{-1} G^T (d - G m_0)$$

Moindres carrés amortis :

Incertitude des **données**  
(matrice de covariance)

Incertitude du **modèle**  
(matrice de covariance)

$$m_{\text{est}} = m_0 + (G^T C_d^{-1} G + C_m^{-1})^{-1} G^T C_d^{-1} (d - G m_0)$$



# Généralisation

avec  $L = (G^T C_d^{-1} G + C_m^{-1})^{-1} G^T C_d^{-1}$

$C_d = \sigma_d^2 \cdot I$  et  $C_m = \sigma_m^2 \cdot I$

Amortissement:

$\epsilon^2 = \frac{\sigma_d^2}{\sigma_m^2}$

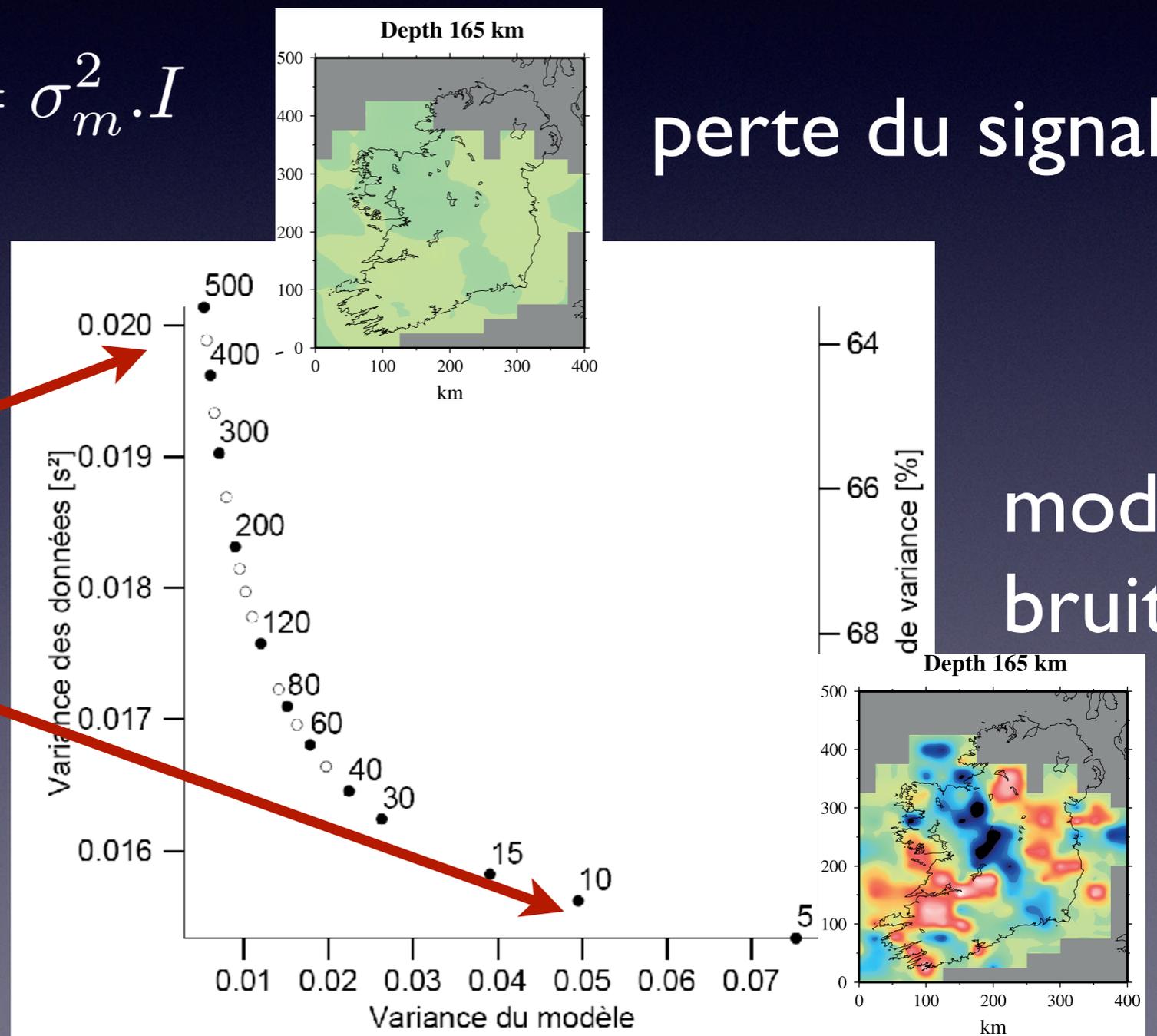
Trop fort

Trop faible



perte du signal

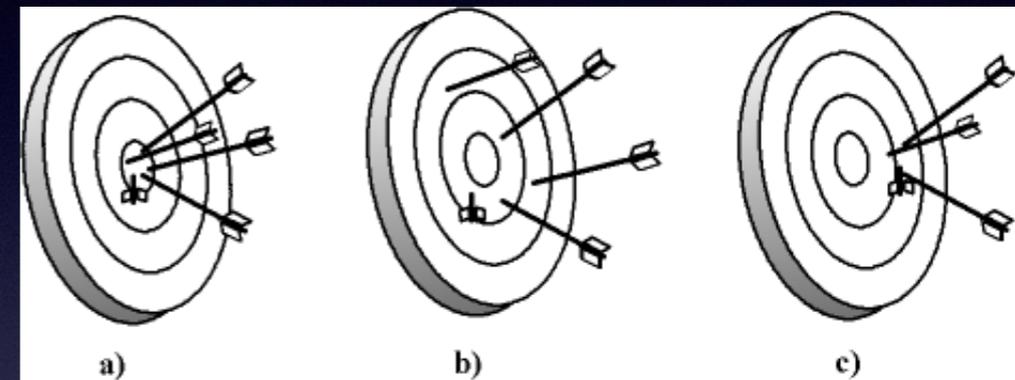
modèle  
bruité



# Les *a priori*...

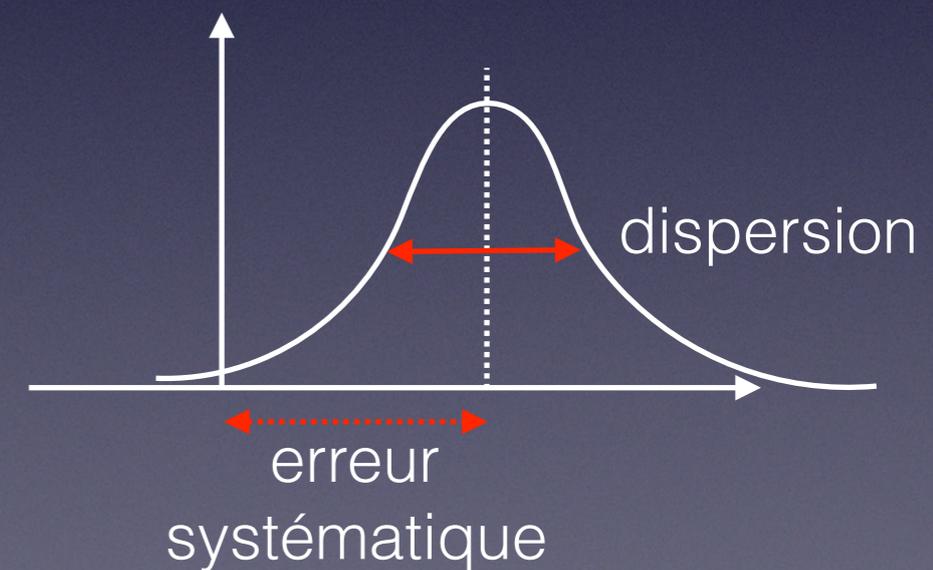
## Sur les données ...

- exactitude (*resolution*)
- dispersion (*precision*)
- erreur systématique (*accuracy*)



## Sur les paramètres ...

- compris entre 2 extrema
- $> 0$
- pas trop éloigné du voisin
- proche d'une valeur connue
- ...



# Conclusion

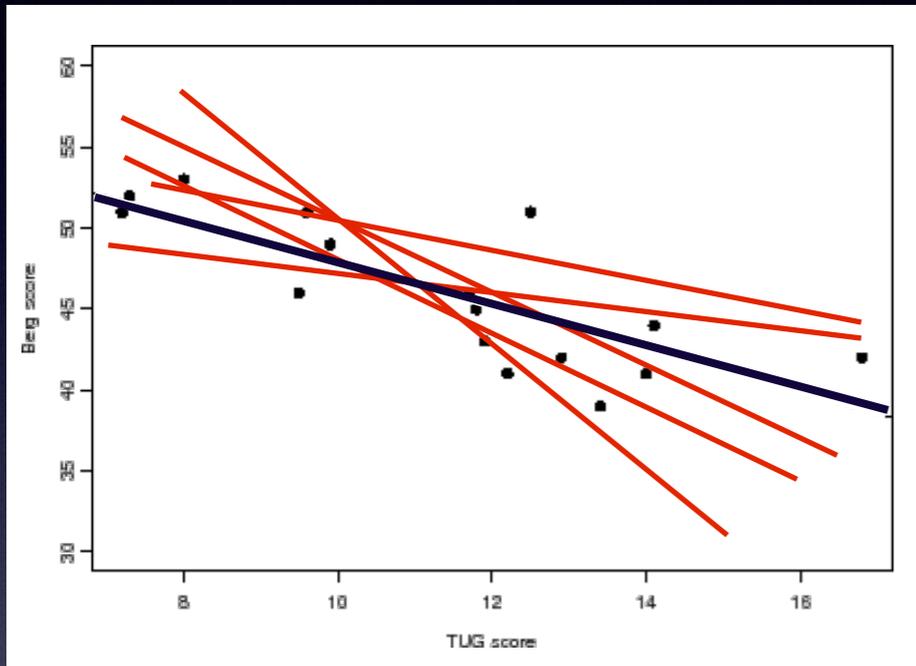
Inversion par moindres carrés (amortis)

- Si et seulement si  $G^T G$  est inversible
- Calcul matriciel un peu lourd (optimisation)
- Mal adapté pour les pb non-linéaires
- Tient compte des *a priori* sur les données+paramètres ( $C_d$  et  $C_m$ )
- Permet de “lisser” le modèle (Attention!!)
- Obtention d’un seul modèle ?
  - au sens des moindres carrés, oui
  - mais en jouant sur les  $C_d$ ,  $C_m$  et lissage, non!

La non-unicité est toujours sous-jacente!!

# Résoudre un pb inverse...

(Autre approche)



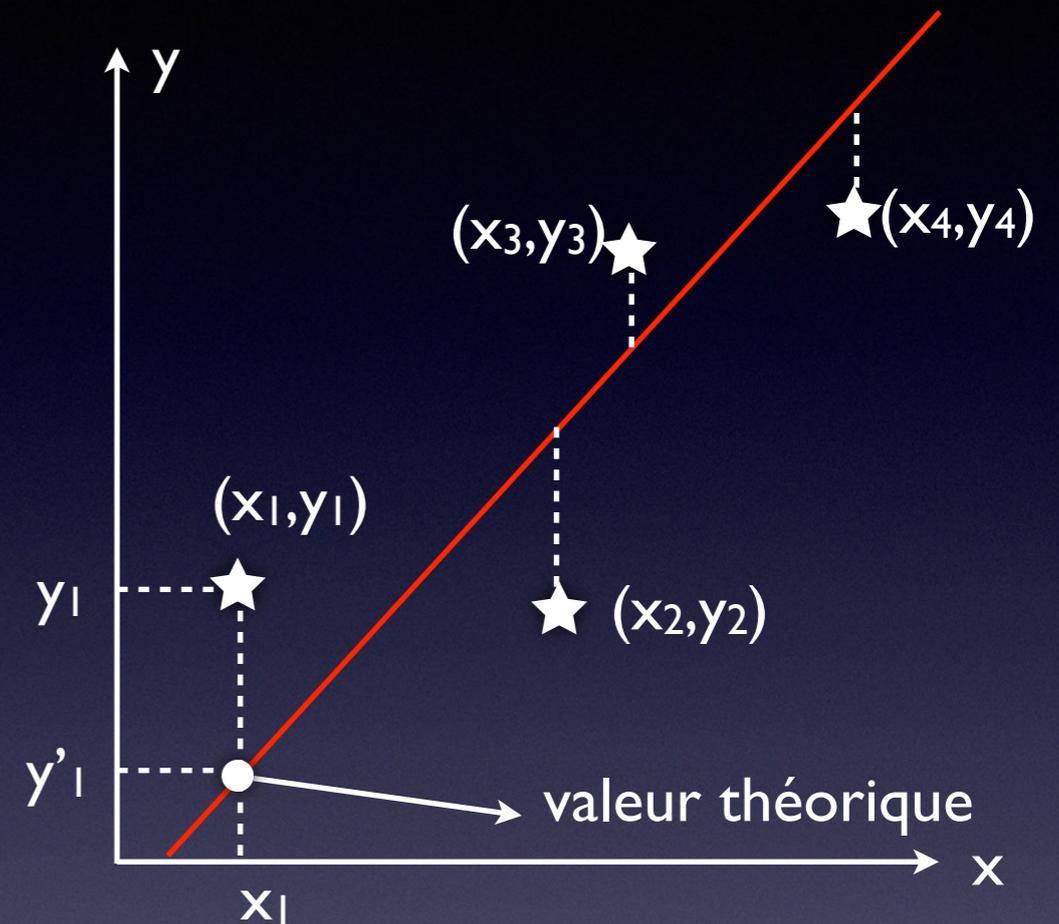
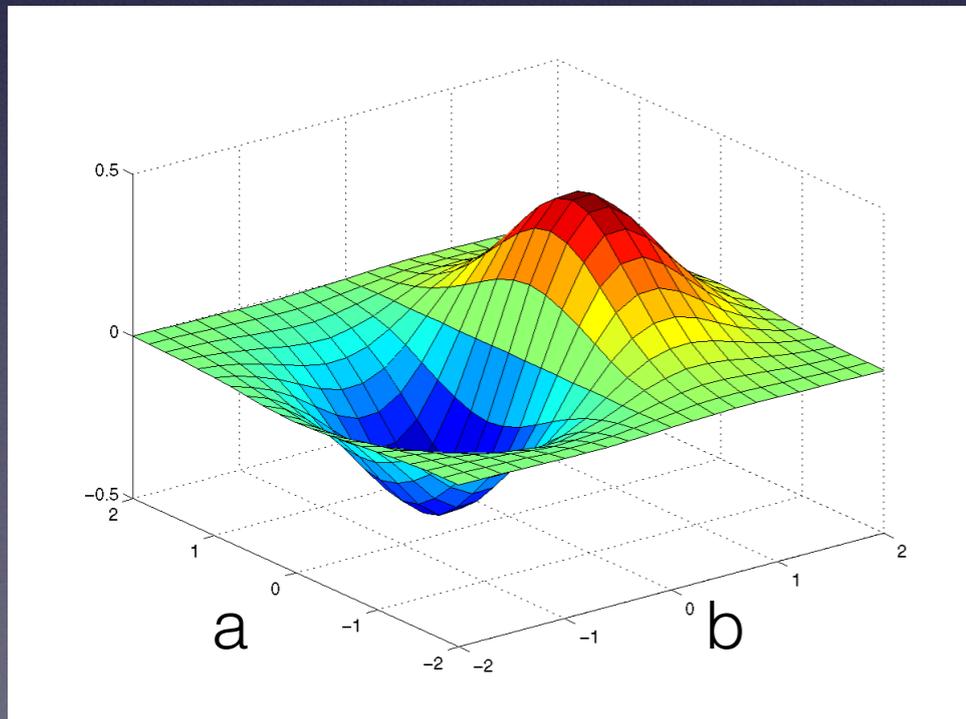
Trouvez  $a$  et  $b$  tels que  $y = ax + b$   
On peut tester **tous** les  $a$  et  $b$  et choisir ceux qui expliquent le mieux les données...

- Comment choisir des paramètres pas trop *idiots* ?
- Comment choisir les meilleurs paramètres ?
- Quand s'arrêter ?

# Résoudre un pb inverse...

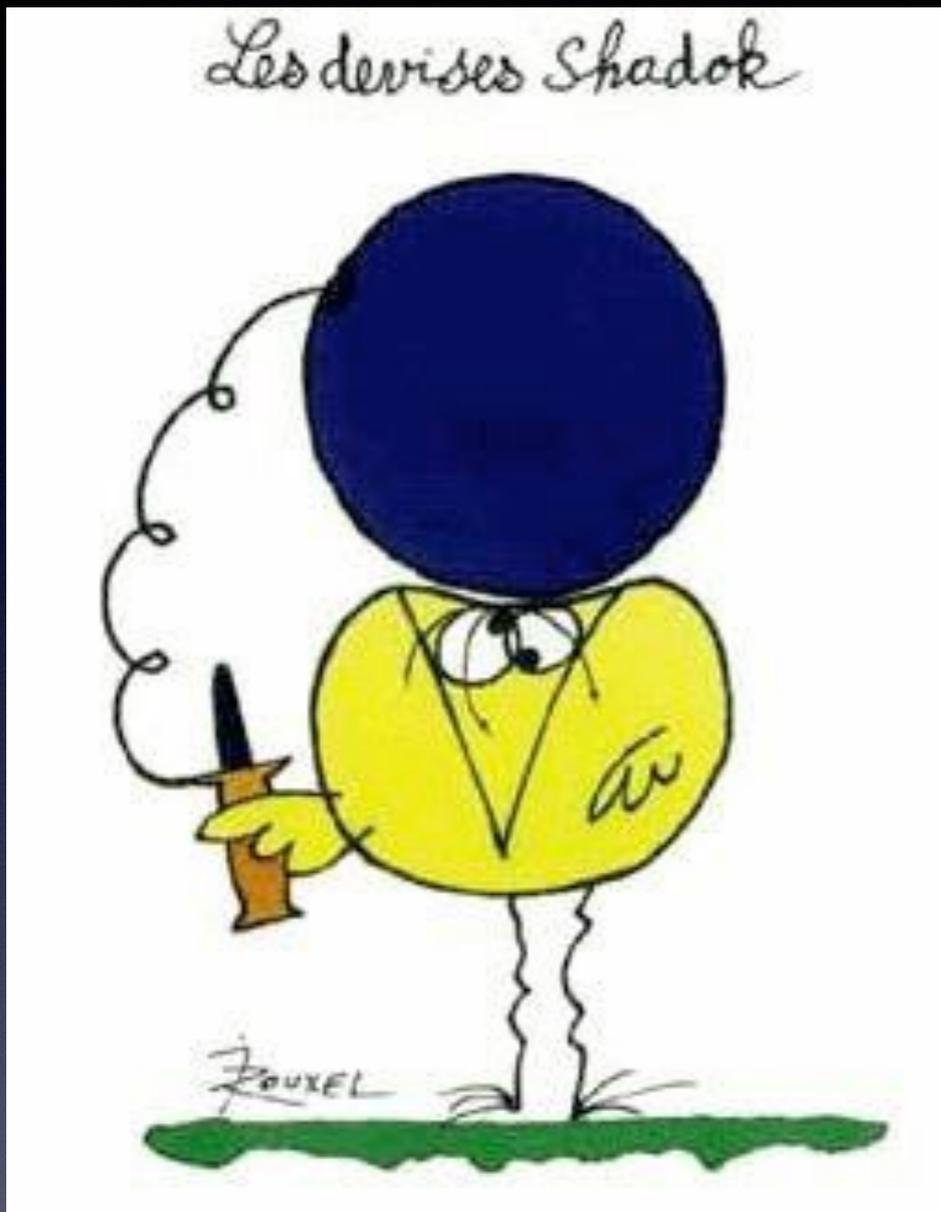
(Autre approche)

- On essaie *tous* les couples  $(a,b)$
- on calcule  $\sum [y - (ax + b)]$
- on choisit la plus petite



Inversion stochastique

# Méthodes stochastiques



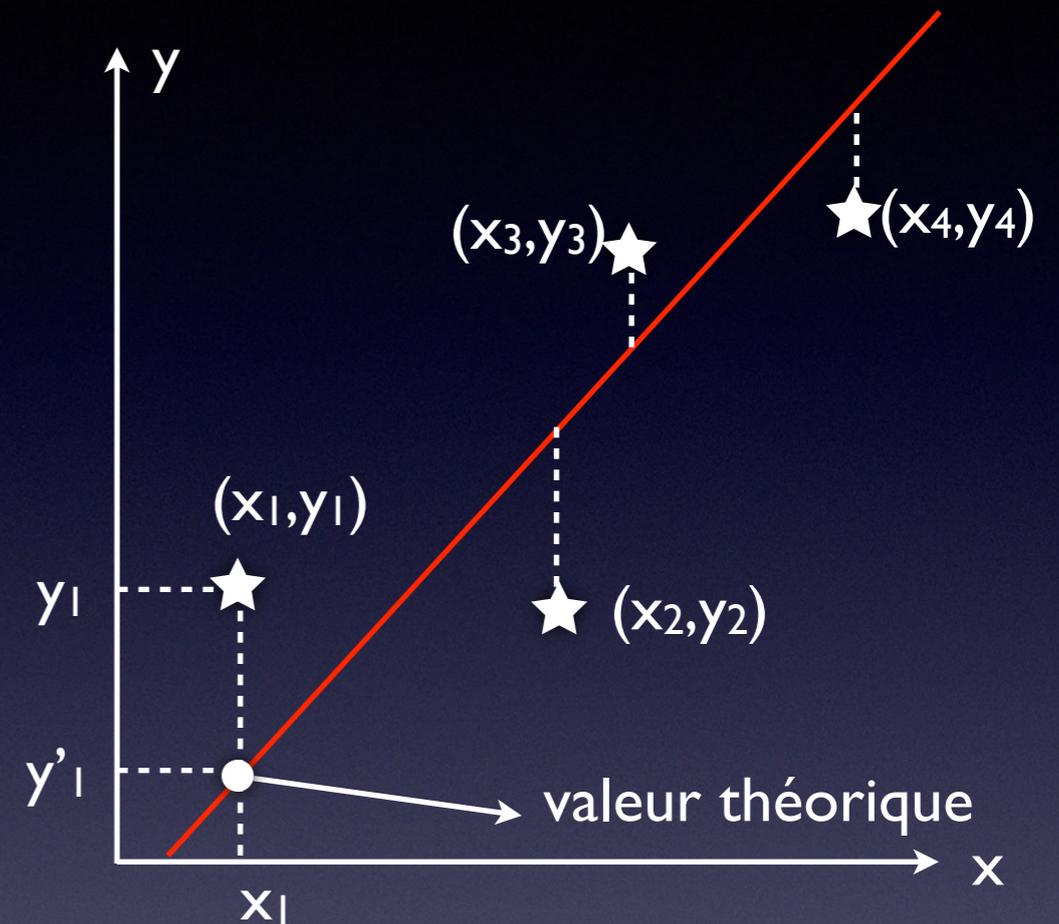
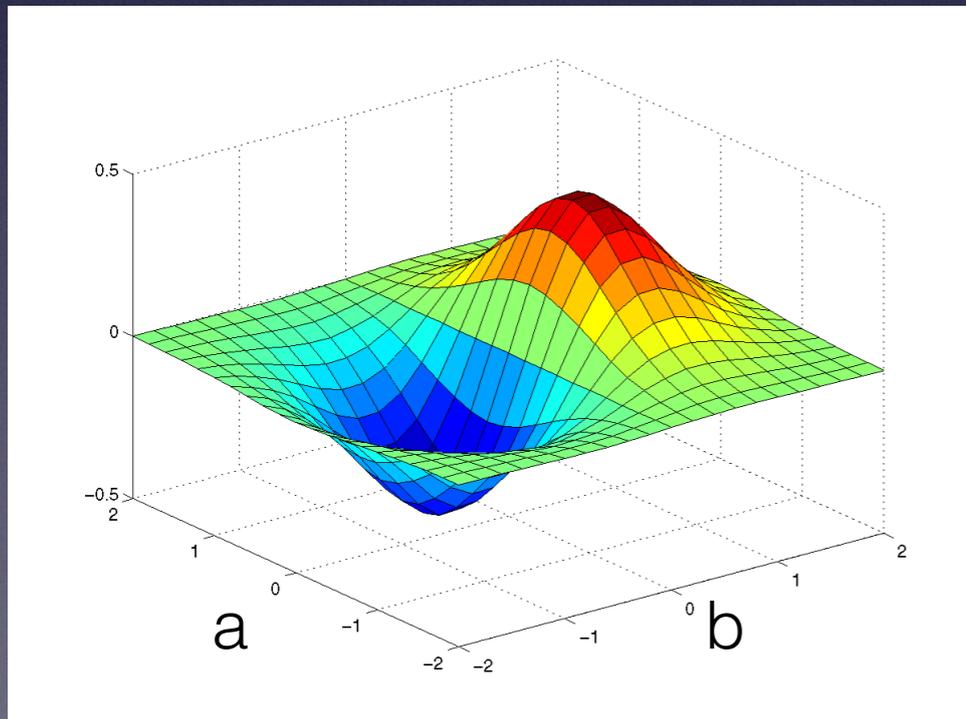
“En essayant continuellement on finit par réussir. Donc : plus ça rate, plus on a de chance que ça marche”.

**calculs faciles, mais  
j'prends mon temps...**

“La probabilité de réussir la mise sur orbite d'une fusée est d'une chance sur un million. Dépêchons-nous de rater 999 999 lancements !”

# Résoudre un pb inverse...

- On essaie *tous* les couples  $(a,b)$
- on calcule  $\sum [y - (ax + b)]$
- on choisit la plus petite



Inversion stochastique

$$d = Gm$$

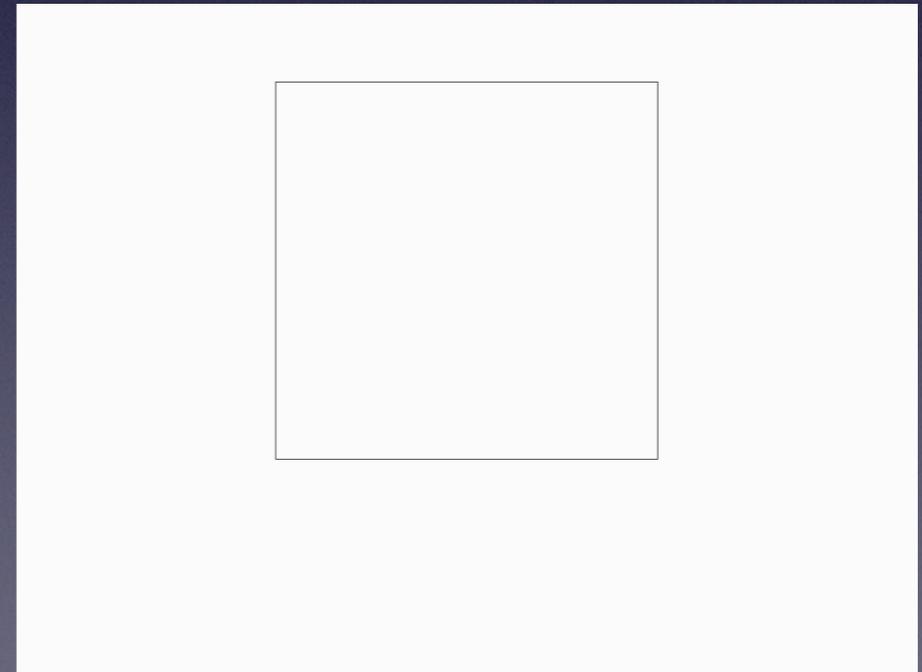
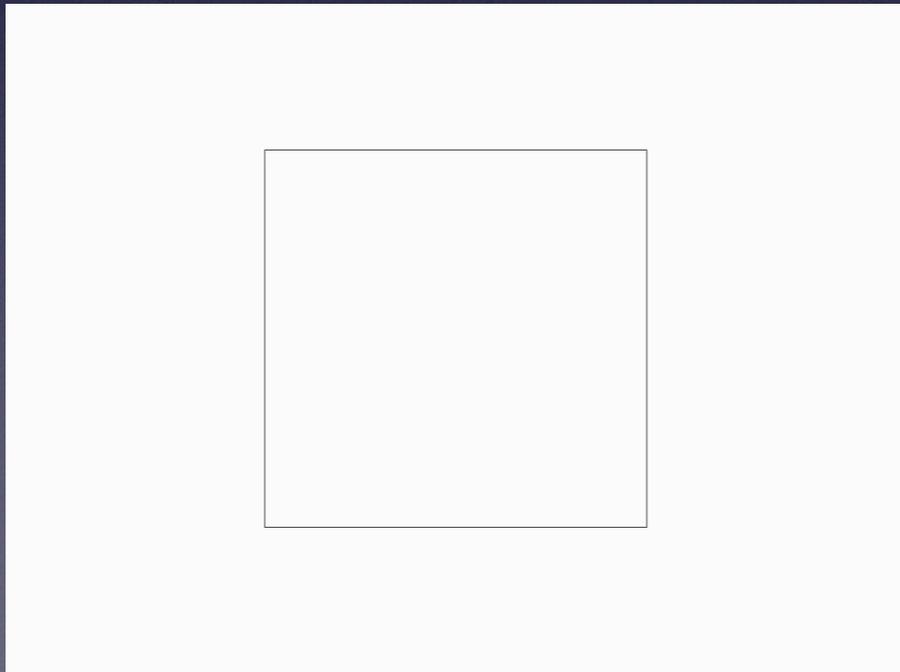
# Méthodes stochastiques

- Essor de ces techniques car:
  - Développement de la théorie des probabilités et des outils informatiques
  - Calcul direct + rapide et maîtrisé
  - Plus efficace pour les problèmes non-linéaires
- Différentes algorithmes (différentes stratégies)
  - recuit simulé
  - algorithme génétique
  - des gradients
  - Mais toutes basées sur la technique de Monte Carlo

# Méthodes stochastiques

## Monte Carlo

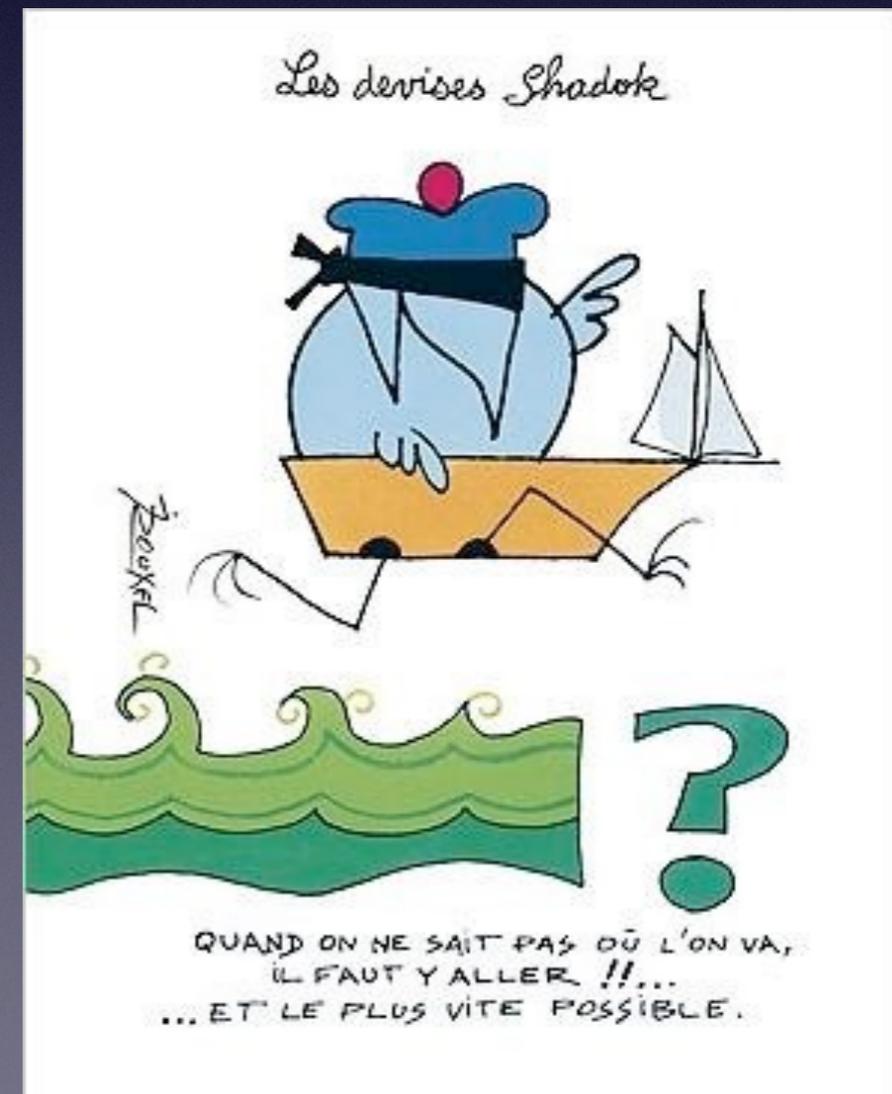
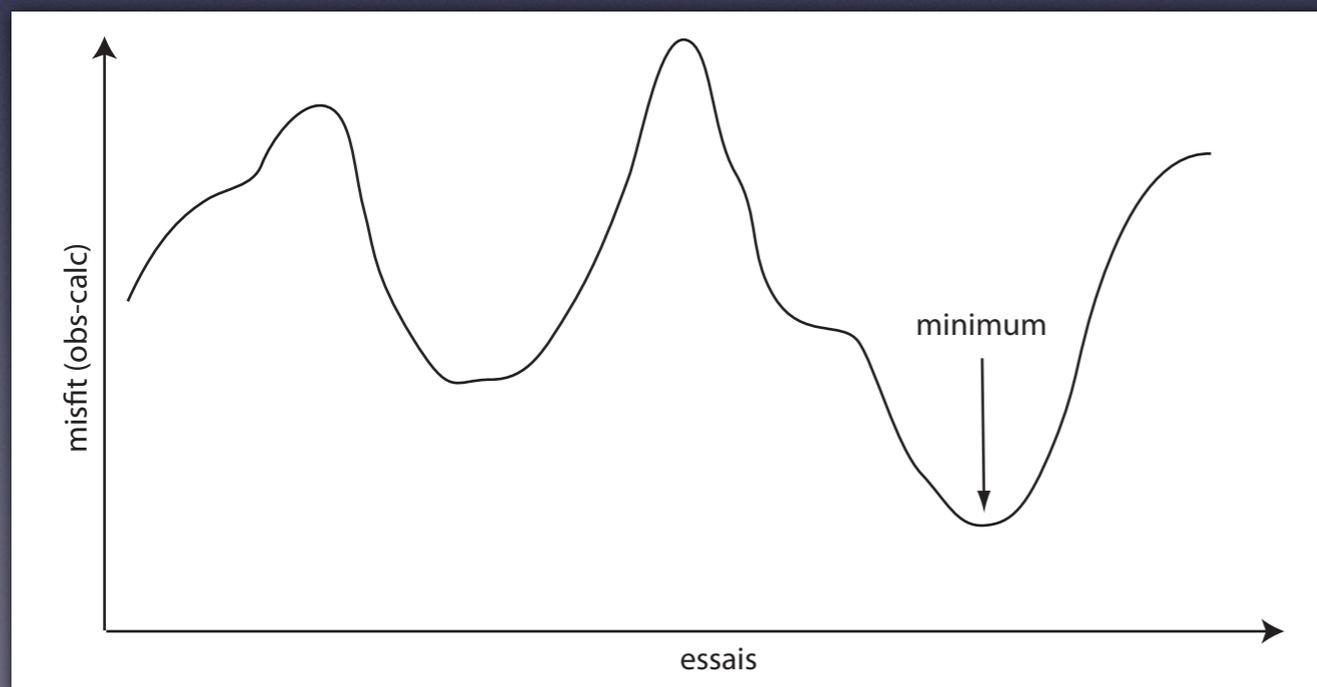
- Outil informatique permet de faire des millions de calculs très rapidement. L'exploration de l'espace des paramètres devient alors possible.
- Mais pas de façon systématique... Il faut explorer les différentes solutions de façon intelligente ou "pseudo-aléatoire".



si on explore suffisamment (mais pas totalement) l'espace, la probabilité de trouver la bonne solution est grande.

# Méthodes stochastiques

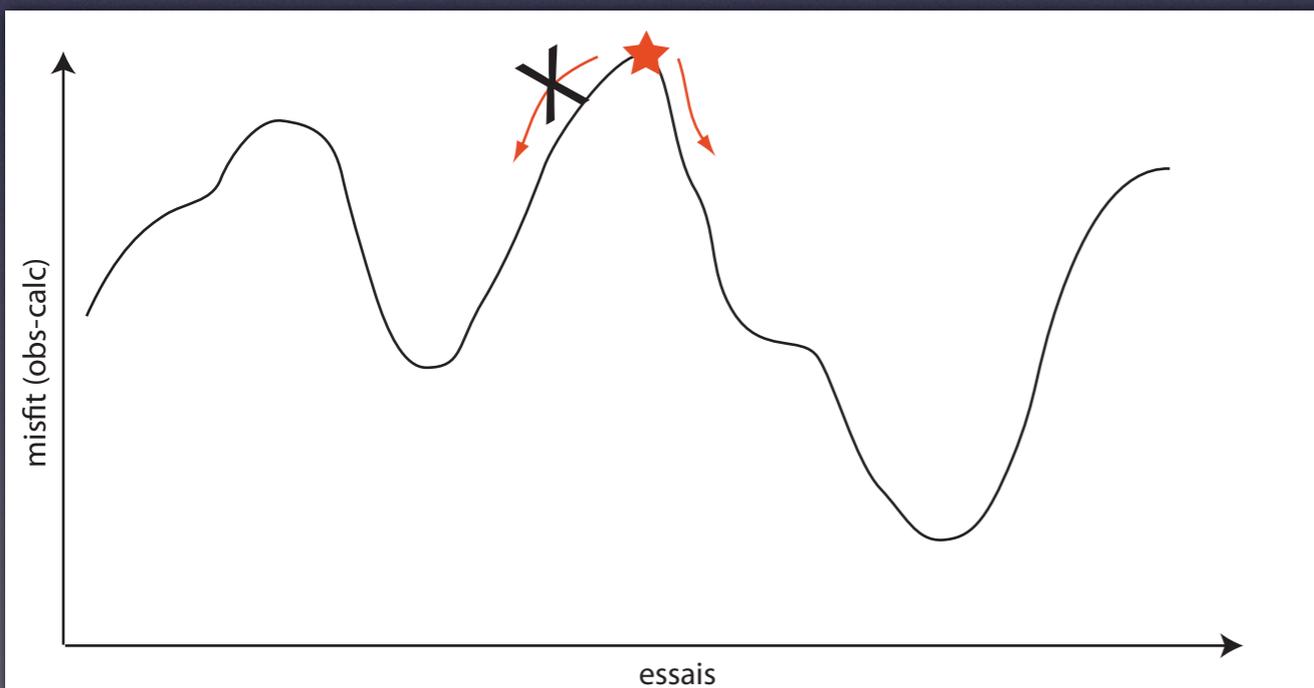
- On ne veut pas explorer tout l'espace, mais on a quand même besoin de milliers d'essais (recherche statistiquement significative)
- Il faut optimiser la recherche, i.e. tenter de converger le plus rapidement possible vers le minimum de misfit (obs-calc).
- Plusieurs algorithmes se sont développés.



# Méthodes stochastiques

La recherche par la technique des gradients

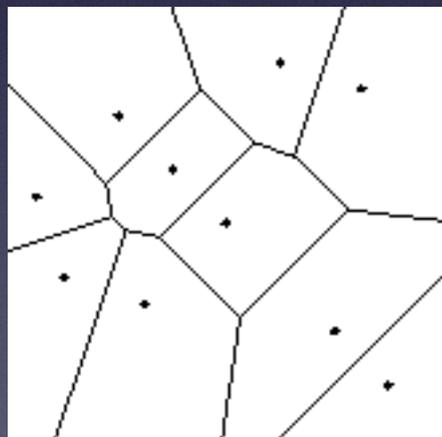
- OK si peu de minimum secondaires et bonne convergence
- bof si convergence difficile (gradient faible) ou si minimum secondaires existent
- s'assurer d'un bon modèle de départ!



# Méthodes stochastiques

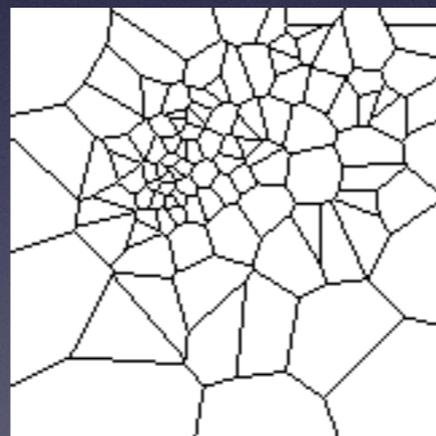
La recherche par l'algorithme de voisinage

- OK si peu de paramètres
- évite les minimum secondaires
- donne estimation de tous les modèles testés (i.e. gourmand en mémoire).



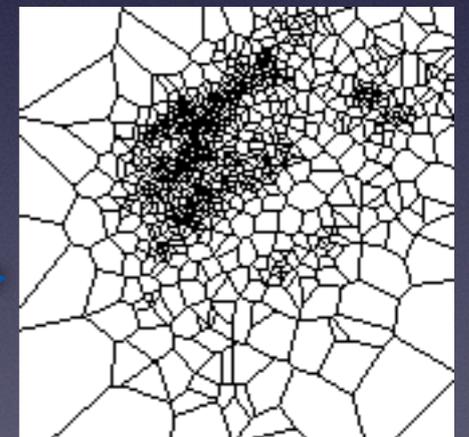
10 modèles  
choisis au hasard  
avec “zones”  
d’influence

calcul du  
misfit



echantillonnage  
plus resserré  
autour des  
meilleurs

...



Choix du  
“meilleur” parmi  
ceux testés

Sambridge, 2000

# Méthodes stochastiques

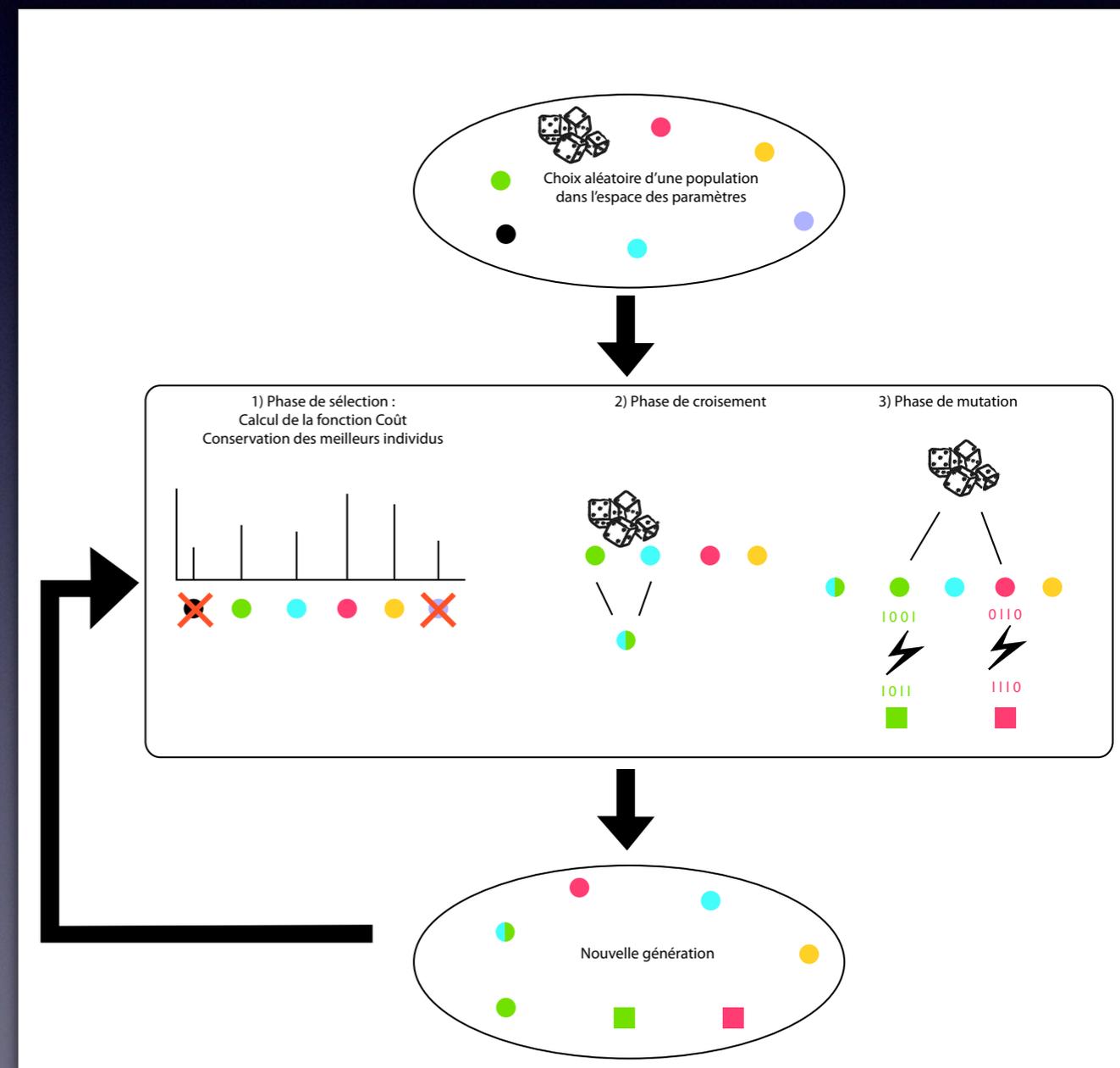
La recherche par l'algorithme génétique

Génération de modèles qui vont pouvoir:

- se croiser
- muter
- être sélectionnés.

pour arriver à la sélection du meilleur

Exploration très complète de l'espace des modèles pour ne pas oublier des minima possibles



# Méthodes stochastiques

- Beaucoup de méthodes différentes suivant leur application
  - favorise la rapidité d'exécution,
  - favorise l'exploration de l'espace
  - traitement des minima secondaires
- Elles ne proposent toutes qu'un ensemble de solution et sont les seules à tenir compte de la non-unicité des solutions.
- Importance des modèles initiaux!!

# Comparaison...

## Inversion par matrice

- Peu de mémoire
- Problème souvent approximé, amorti, lissé
- Une seule solution ( $m_{est}$ )
- Faire **très** attention à la validité de la solution!!!

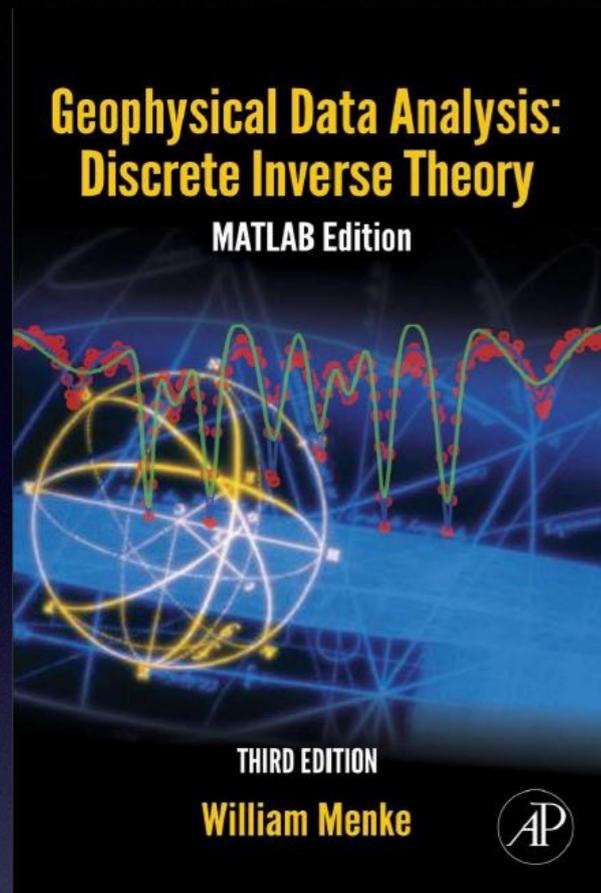
## Inversion stochastique

- Bcp de mémoire, bcp d'essais (>1000)
- très efficace pour non-linéaire
- Plusieurs solutions proposées
- La représentativité de la solution dépend de la stratégie de recherche

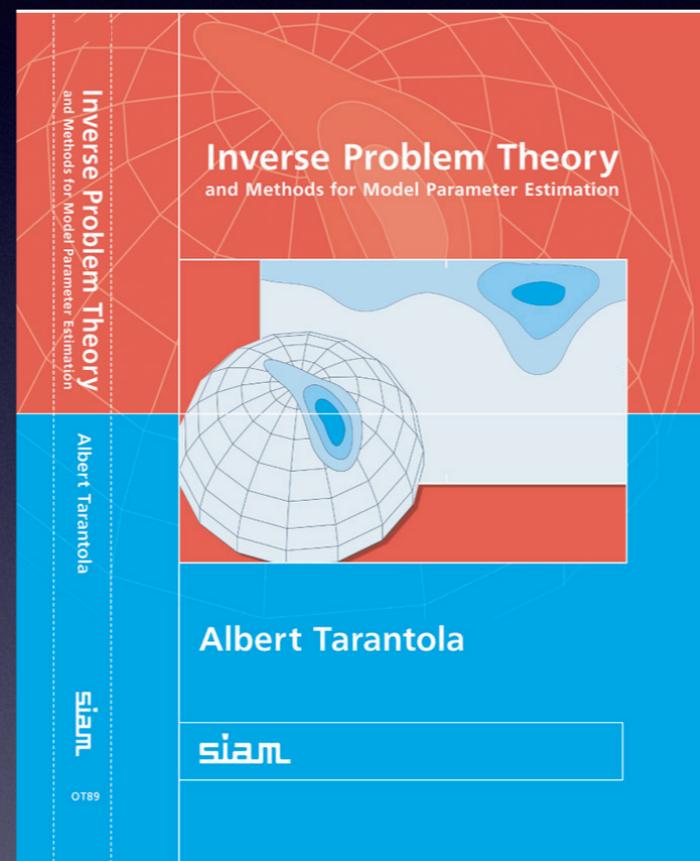
# Pour résumer

- Dans tous les cas, le modèle obtenu est un concept qui tente d'expliquer les observations de terrain.
  - il est non-unique
  - il est soumis à des hypothèses (loi physique, a priori, )
  - il est parfois soumis à des approximations (mathématiques, lissage, amortissement...)
  - il doit toujours être associé à une erreur ou probabilité d'existence.
- **Il faut toujours être capable de juger de sa pertinence (ses limites)**

# Coin lecture...



<http://www.ldeo.columbia.edu/users/menke/index.php>



<http://www.ipgp.fr/~tarantola/>

<http://www.ipgp.fr/~gibert/Downloads.html>

<https://mesoscopic.mines.edu/~jscales/gp605/main.pdf>