

Optimisation sans dérivées pour le problème de calage d'historique dans le domaine pétrolier

Benjamin MARTEAU

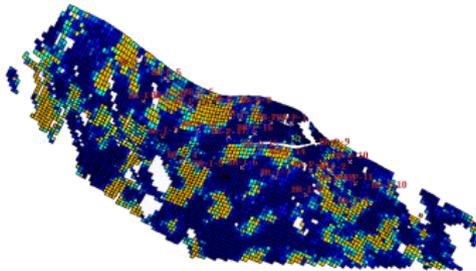
IFP Énergies Nouvelles
Université de Versailles Saint-Quentin

22 mai 2014

Plan de l'exposé

- 1 Problèmes inverses en ingénierie pétrolière
 - Contexte du calage d'historique
 - Caractéristiques du problème à résoudre
- 2 Méthodes de type région de confiance
 - Notions générales
 - Adaptation aux fonctions partiellement séparables
 - Résultats numériques
- 3 conclusion

Contexte

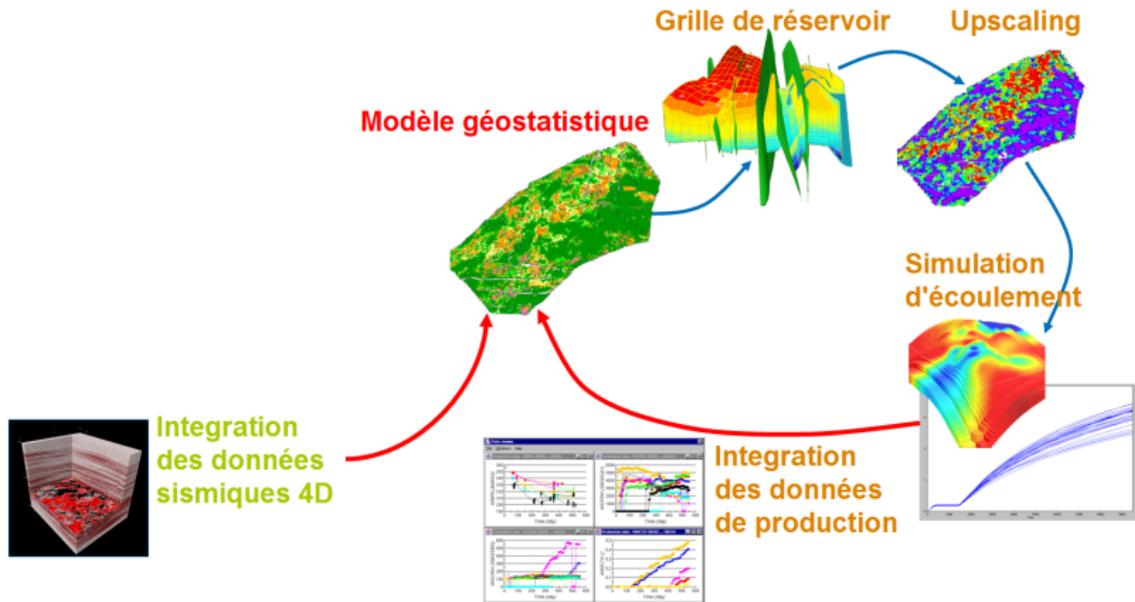


Caractéristiques d'un réservoir pétrolier :

- Grandes dimensions
- Nombreux puits
- Phénomènes physiques complexes

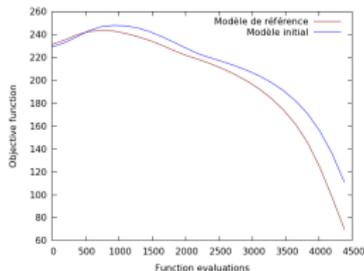
Disposer d'un bon modèle permet de prévoir son comportement, d'optimiser la production, de placer correctement de nouveaux puits...

Le calage d'historique

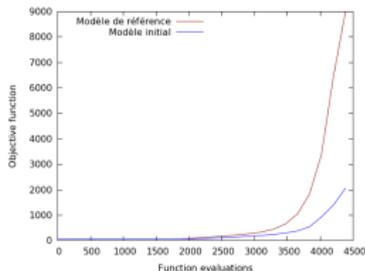


Calage des données de production

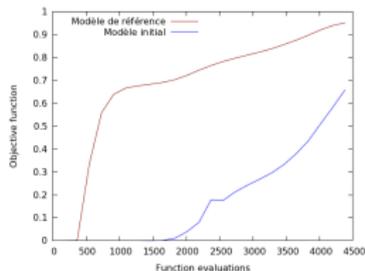
- Dans le cas testé, on choisit de caler les pressions en fond de puits (BHFP), les ratios huile/gaz (GOR) ainsi que la proportion d'eau produite (WC).
- Les données mesurées sont en rouge et celles simulées en bleu.



(a) BHFP



(b) GOR



(c) WC

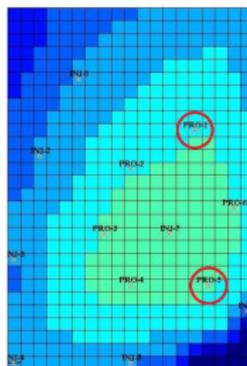
Caractéristiques du problème

- L'évaluation de la fonction objectif à optimiser nécessite la simulation d'écoulements complexes dans un modèle de réservoir : processus très coûteux.
- Les caractéristiques principales du problème à résoudre sont :
 - Dépendance en de nombreux paramètres
 - Évaluation coûteuse de la fonction objectif
 - Non disponibilité des dérivées
 - Fonction objectif partiellement séparable

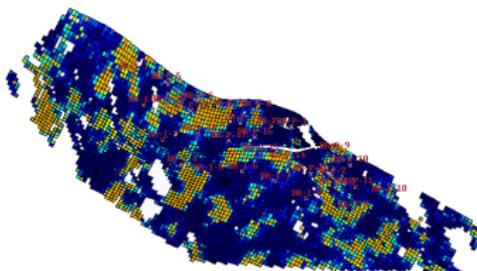
Optimisation pour le calage d'historique

On fait la supposition suivante :

Les variations d'un paramètre associé à un puits n'influera pas sur les données de production d'un puits distant



Séparabilité partielle



La fonction objectif peut s'écrire sous la forme :

$$\begin{aligned}
 F(x_1, \dots, x_n) &= \frac{1}{2} \sum_{i=1}^{n_1} \frac{\omega_i^P}{N_P(i)} \sum_{j=1}^{N_P(i)} \left(\frac{P_{i,j}^{obs}(x) - P_{i,j}^{sim}(x)}{\sigma_{i,j}^P} \right)^2 \\
 &= \sum_{i=1}^P f_i(x_1, \dots, x_n) \\
 &\approx \sum_{i=1}^P f_i(x_{1_i}, \dots, x_{n_i})
 \end{aligned}$$

avec $\forall i, n_i \leq n$

Type de méthode choisie

Quatre grande familles de méthodes d'optimisation sans dérivées sont disponibles :

- Méthodes directes de type simplexe (Nelder Mead)
- Méthodes statistiques (Recuit-simulé, algorithmes génétiques, CMA-ES,...)
- Méthodes de type surface de réponse (RBF, krigeage,...)
- **Méthodes de type région de confiance**

- 1 Problèmes inverses en ingénierie pétrolière
 - Contexte du calage d'historique
 - Caractéristiques du problème à résoudre
- 2 Méthodes de type région de confiance
 - Notions générales
 - Adaptation aux fonctions partiellement séparables
 - Résultats numériques
- 3 conclusion

Un algorithme simpliste

ALGORITHME SIMPLE

- 1 **Initialisation** Construire un modèle quadratique initial d'interpolation m_0 de la fonction sur une région de confiance
- 2 **Itération k**
 - 1 Calculer x_k^+ le minimum de m_k sur une région de confiance
 - 2 remplacer le point d'interpolation le plus éloigné du point courant par x_k^+
 - 3 si x_k^+ est "bon" alors $x_{k+1} := x_k^+$ sinon $x_{k+1} := x_k$ et réduire la région de confiance
 - 4 Mettre à jour le modèle m_k .
- 3 **Condition d'arrêt** S'arrêter lorsque le gradient du modèle devient trop petit

Construction des modèles quadratiques

- Pour construire un modèle quadratique entièrement déterminé, on a besoin de $p = \frac{(n+1)(n+2)}{2}$ points d'interpolation.
- Dans la pratique, on se contente de modèles sous déterminés qu'on obtient en résolvant un problème d'optimisation de la forme

$$\begin{cases} \min(\|\alpha_Q\|_F^2) \\ m_k(x_j) = F(x_j) \forall j \in \{1, \dots, m\} \end{cases}$$

Où les x_j sont les points d'interpolations et le modèle m_k s'écrit :

$$m_k(x) = F(x_k) + \langle x, \alpha_L \rangle + \frac{1}{2} \langle x, \alpha_Q x \rangle$$

Initialisation des modèles quadratiques

- Pour une fonction F dépendant de n paramètres, on initialise généralement le modèle m_0 avec $2n + 1$ points. Ainsi :
 - m_0 est un modèle entièrement linéaire.
 - la diagonale de la partie quadratique de m_0 est déterminée.
- On répartit au mieux les points d'interpolation initiaux
 - y_0 est au centre de la région de confiance
 - $\forall i \in \{0, 1, \dots, n - 1\}$

$$\begin{cases} y_{i+1} = y_0 + \Delta_0 e_i \\ y_{i+n+1} = y_0 - \Delta_0 e_i \end{cases}$$

Une bonne itération

- Un point d'essai x_k^+ est considéré "bon" si le modèle m_k a été suffisamment prédictif :

$$\frac{m_k(x_k^+) - m_k(x_k)}{F(x_k^+) - F(x_k)} \geq \rho$$

où $0 < \rho < 1$ est une constante fixée en début d'algorithme.

- Sans considération pour la qualité du modèle, il peut "dégénérer" au cours des itérations.

Pertinence des modèles quadratiques

l'obtention d'un mauvais point d'essai x_k^+ peut provenir de 2 causes distinctes :

- Le modèle n'est pas assez précis
- La région de confiance considérée est trop grande

Afin de distinguer ces 2 causes, on introduit un moyen de mesurer la qualité des modèles.

Qualité des modèles : Λ -Unisolvance

Soit $Y = \{y_1, \dots, y_m\}$ un ensemble de points d'interpolation dans une boule B . On souhaite avoir une mesure de la qualité du modèle obtenu avec les points de Y .

Si $\Lambda > 0$ et $B \subset \mathbb{R}^n$:

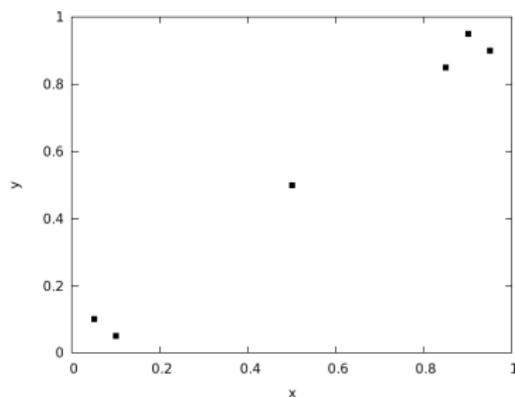
DÉFINITION

On dit qu'un ensemble de points Y est Λ -unisolvant dans B si :

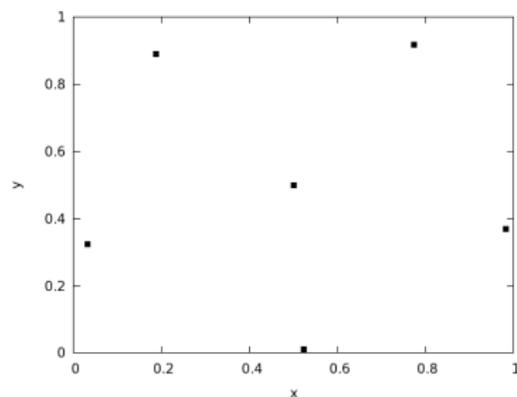
$$\max_i \max_{x \in B} |\lambda_i(x)| \leq \Lambda$$

Où $\forall i \in \{1, \dots, m\}$, λ_i est le polynôme de Lagrange associé à $y_i \in Y$.

Illustration de la Λ -Unisolvance



(d) $\Lambda = 440$



(e) $\Lambda = 1$

Amélioration des modèles

LEMME

Étant donné une boule $B(x, \Delta)$, Y un ensemble d'interpolation Λ -unisolvant et $\{\lambda_i\}$ les polynômes de Lagrange associés, alors :

$$\begin{cases} \|F(y) - m(y)\| \leq k_{ef} \sum_{j=1}^q \|y_j - y\|^2 |\lambda_j(y)| \\ \|\nabla F(y) - \nabla m(y)\| \leq k_{eg} \Lambda \Delta \end{cases}$$

LEMME

Soit Y un ensemble d'interpolation, B une boule de \mathbb{R}^n et $\{\lambda_i\}$ les polynômes de Lagrange associés aux points de Y . La procédure :

- trouver x et j tels que $\lambda_j(x) \geq \Lambda$
- Remplacer y_j par x dans Y

se termine en un nombre fini d'itération avec Y Λ -unisolvant.

Algorithme NEWUOA

NEWUOA

- 1 **Initialisation** Construire un modèle quadratique initial d'interpolation m_0 de la fonction
- 2 **Itération k**
 - 1 Calculer x_k^+ le minimum de m_k sur une région de confiance
 - 2 Si x_k^+ est "bon" remplacer le point d'interpolation le plus éloigné du point courant par x_k^+ et poser $x_{k+1} := x_k^+$
 - 3 Dans le cas contraire :
 - Si le modèle n'est pas Λ -unisolvant, améliorer la géométrie du modèle
 - Sinon réduire la taille de la région de confiance
 - 4 Mettre à jour le modèle m_k .
- 3 **Condition d'arrêt** S'arrêter lorsque le gradient du modèle devient trop petit

Auto-correction de la géométrie

Toint et Scheinberg (2009) ont proposé une propriété permettant de réduire le coût de l'étape d'amélioration de la géométrie

PROPRIÉTÉ

$\forall \Lambda > 1, \exists k_\Lambda$ tel que si :

- $\Delta \leq k_\Lambda \|\nabla m_k(x_k)\|,$
- $F_k := \{y_{k,j} \in Y_k / \|y_{k,j} - x_k\| > \beta \Delta_k \text{ et } l_{k,j}(x_k^+) \neq 0\} = \emptyset$

Alors :

$$C_k := \{y_{k,j} \in Y_k \setminus \{x_k\} / \|y_{k,j} - x_k\| \leq \beta \Delta_k \text{ et } l_{k,j}(x_k^+) \geq \Lambda\} \neq \emptyset$$

- 1 Problèmes inverses en ingénierie pétrolière
 - Contexte du calage d'historique
 - Caractéristiques du problème à résoudre
- 2 Méthodes de type région de confiance
 - Notions générales
 - Adaptation aux fonctions partiellement séparables
 - Résultats numériques
- 3 conclusion

Idée générale

On considère uniquement les fonctions de la forme :

$$F(x_1, \dots, x_n) = \sum_{i=1}^p f_i(x_{1_i}, \dots, x_{n_i})$$

avec $\forall i, n_i \leq n$

L'idée simple est de créer un modèle quadratique pour chaque sous fonction objectif, chaque sous modèle ayant alors son propre ensemble d'interpolation.

$$m_k(x) = \sum_{i=1}^p m_k^i(x)$$

Les gains espérés sont :

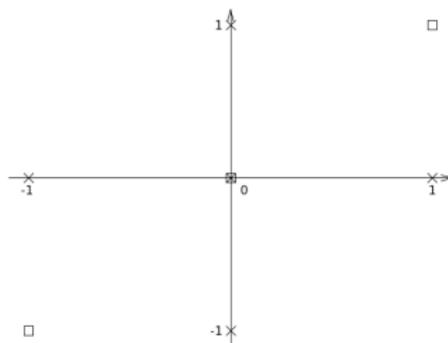
- Des modèles plus précis
- Un nombre de points nécessaires bien inférieur pour l'interpolation

Initialisation des modèles : un exemple simple

- Dans le cas où F s'écrit sous la forme :

$$F(x_1, x_2) = f_1(x_1) + f_2(x_2)$$

L'initialisation de l'algo peut se fait grâce aux points :

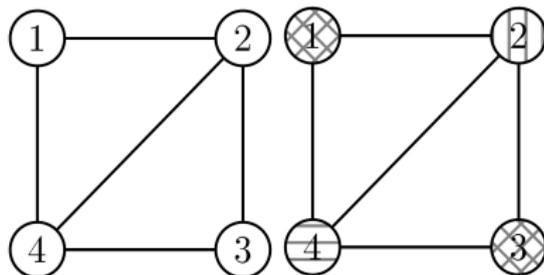


Coloration des variables : Exemple

- Dans le cas où F s'écrit sous la forme :

$$F(x_1, x_2, x_3, x_4) = f_1(x_1, x_2) + f_2(x_2, x_3) + f_3(x_3, x_4) + f_4(x_4, x_1) + f_5(x_2, x_4)$$

On peut créer puis colorer le graphe des dépendances des variables :



- 2 variables x_k et x_l ont la même couleur s'il n'existe pas de fonction f_i qui dépende à la fois de x_k et de x_l .

Initialisation des sous modèles

Pour initialiser chaque sous modèle avec $2n_i + 1$ points, on utilise le graphe coloré représentant la dépendance des variables

- Soit c le nombre de couleur du graphe et y_0 le point au centre de la région de confiance.
- $\forall i \in \{1, \dots, c\}$, soit c_i la i^{eme} couleur et I_i l'ensemble des indices des variables colorées par c_i .
- On construit alors les points d'interpolation $y_i \in Y$ avec

$$\begin{cases} y_{i+1} = y_0 + \Delta_0 \sum_{j \in I_i} e_j \\ y_{i+c+1} = y_0 - \Delta_0 \sum_{j \in I_i} e_j \end{cases}$$

Géométrie des sous modèles

- Un problème apparait pour l'étape d'amélioration des modèles : une simple adaptation de NEWUOA nécessiterait potentiellement une amélioration de tous les sous modèles à chaque itération.
- On introduit la notion de sous modèle dominant. A l'itération k , c'est le sous modèle maximisant la valeur $m_k^i(x_k^+) - m_k^i(x_k)$

Auto-correction de la géométrie des modèles

PROPRIÉTÉ

On note F_k^i et C_k^i les analogues de F_k et C_k de la propriété précédente. Pour tout $\Lambda > 1$, si

$$\left\{ \begin{array}{l} j \text{ est l'indice du sous modèle dominant} \\ \Delta_k \leq \min\left(\frac{1}{k_h}, \frac{(1-\nu)}{2k_{ef}(\beta+1)^2(p\Lambda+1)}\right) \|g_k\| = k_\Lambda \|g_k\| \\ \rho_k^j < \nu \\ F_k^j = \emptyset \end{array} \right.$$

Alors $C_k^j \neq \emptyset$

Un résultat de convergence

THÉORÈME

On suppose que :

- La fonction F est différentiable et ∇F est continue lipschitzienne sur un ensemble E comprenant toutes les itérations de l'algorithme.
- F est minorée sur E .

Alors :

$$\liminf_{k \rightarrow \infty} \|\nabla F(x_k)\| = 0$$

Une itération du nouvel algorithme proposé

ITÉRATION k

- 1 Calculer x_k^+ le minimum de m_k sur une région de confiance
- 2 Traitement de chaque sous modèle m_k^i dans les 3 cas de figure :
 - Le modèle global m_k est bon, remplacer un point de Y_k^i par x_k^+
 - m_k n'est pas bon mais m_k^i l'est, remplacer un point de Y_k^i par x_k^+
 - Ni m_k ni m_k^i ne sont bons.
 - Si $F_k^i \neq \emptyset$, en remplacer un point par x_k^+
 - Si $C_k^i \neq \emptyset$, en remplacer un point par x_k^+
- 3 Si $\rho_k < \rho$ et $\rho_k^j > \rho$, améliorer la géométrie du moins bon modèle.
- 4 Gérer la région de confiance
- 5 Mettre à jour le modèle m_k .

Explication du cas de figure 4

A l'itération k , il est possible que :

- Le modèle m_k ne soit pas bon ($\rho_k < \rho$)
- Le sous modèle dominant m_k^j soit bon

Dans ce cas, on ne peut pas garantir qu'un des modèles ait vu sa géométrie améliorée.

Pour éviter un blocage de l'algorithme, il est nécessaire de prendre en compte ce cas de figure.

- 1 Problèmes inverses en ingénierie pétrolière
 - Contexte du calage d'historique
 - Caractéristiques du problème à résoudre
- 2 Méthodes de type région de confiance
 - Notions générales
 - Adaptation aux fonctions partiellement séparables
 - Résultats numériques
- 3 conclusion

Fonctions tests

Des tests de la nouvelle méthode ont été réalisés sur des fonctions analytiques partiellement séparables : pour $x = (x_1, \dots, x_n)$

$$DQDRTIC(x) = \sum_{i=1}^{p-2} (x_i^2 + 100x_{i+1}^2 + 100x_{i+2}^2)$$

$$LIARWHD(x) = \sum_{i=1}^p 4(x_i^2 - x_1)^2 + (x_i - 1)^2$$

$$BDQRTIC(x) = \sum_{i=1}^{p-4} (-4x_i + 3)^2 + x_i^2 + 2x_{i+1}^2 + 3x_{i+2}^2 + 4x_{i+3}^2 + 5x_n^2$$

$$ARWHEAD(x) = \sum_{i=1}^{p-1} ((x_i^2 + x_n^2)^2 - 4x_i + 3)$$

$$ROSENBROCK(x) = \sum_{i=1}^{p-1} (100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2)$$

Comparaisons des résultats avec NEWUOA 1/2

Le tableau suivant compare les optimisations des fonctions tests précédentes avec NEWUOA avec 10 parameters.

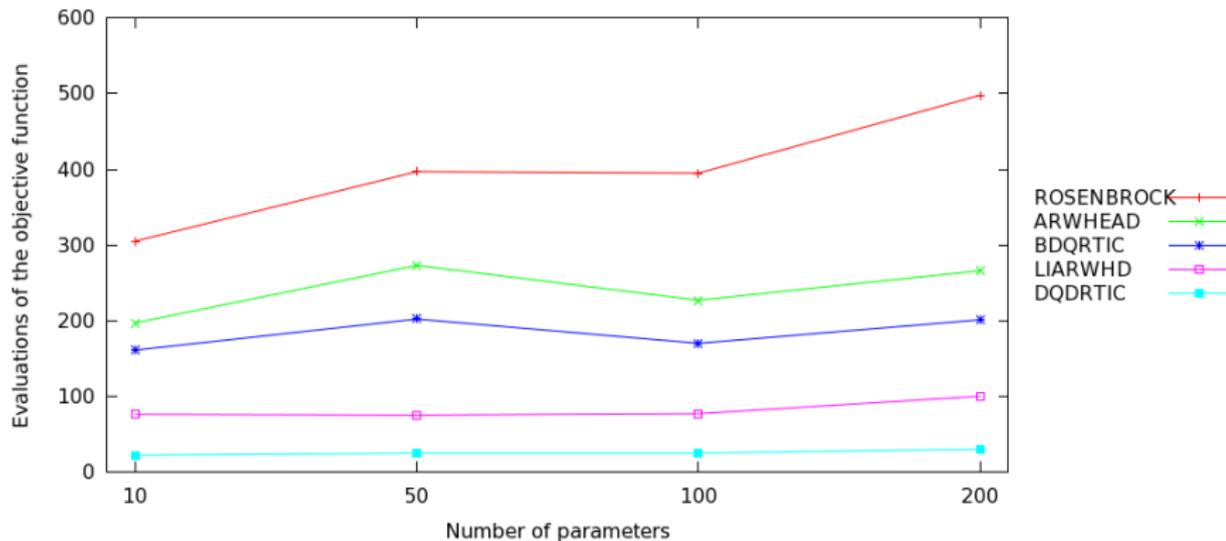
10 parameters	SQA		SQA-PSOF	
Function	Best point	Evaluations	Best point	Evaluations
DQDRTIC	$4,30E - 12$	839	$7,30E - 16$	19
LIARWHD	$4,90E - 13$	257	$1,26E - 09$	48
BDQRTIC	18,2812	534	18,2903	186
ARWHEAD	$2,09E - 13$	229	$1,19E - 09$	37
ROSENBROCK	$3,18E - 11$	753	$9,20E - 09$	312

Comparaisons des résultats avec NEWUOA 2/2

Le tableau suivant compare les optimisations des fonctions tests précédentes avec NEWUOA avec 50 parameters.

50 parameters	SQA		SQA-PSOF	
Function	Best point	Evaluations	Best point	Evaluations
DQRDRTIC	$3,05E - 10$	4352	$7,62E - 16$	18
LIARWHD	$1,60E - 12$	598	$1,90E - 09$	72
BDQRDRTIC	178,489	2672	178,489	210
ARWHEAD	$1,12E - 11$	3597	$6,70E - 07$	48
ROSENBROCK	$3,80E - 10$	4138	$1,32E - 08$	383

Nombre d'évaluations en fonction du nombre de paramètres



Une application industrielle dans le domaine pétrolier : Réservoir PUNQ

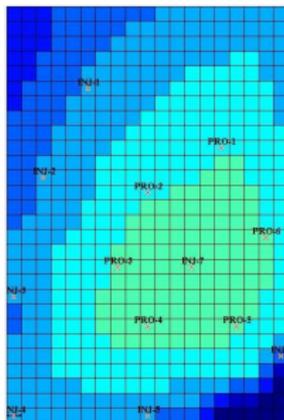
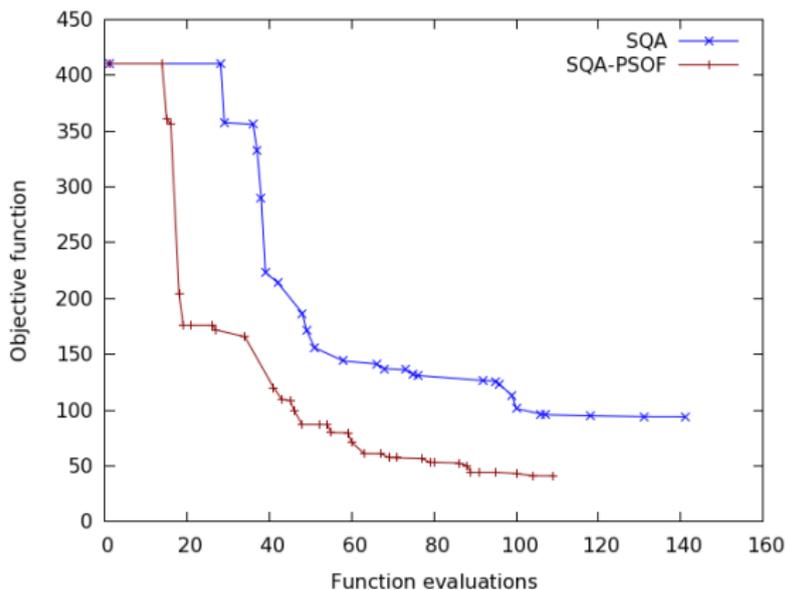


FIGURE : Profondeur du réservoir PUNQ

Réservoir synthétique comprenant 13 puits (7 injecteurs et 6 producteurs)

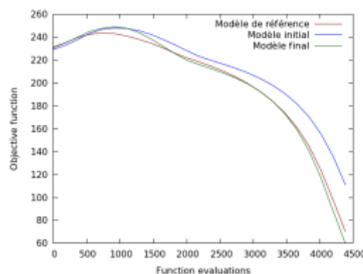
On dispose d'un historique de 12 ans de production

Évolution de la fonction objectif pour le cas PUNQ

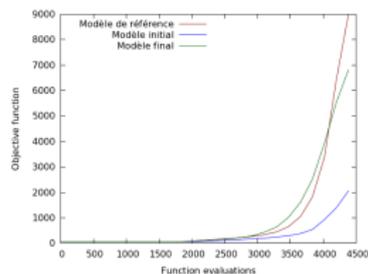


Calage des données de production

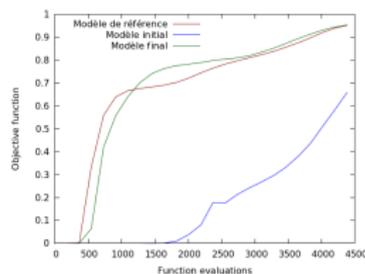
- Données de production après le processus d'optimisation.
- Les données mesurées sont en rouge et celles simulées en bleu.



(a) BHFP



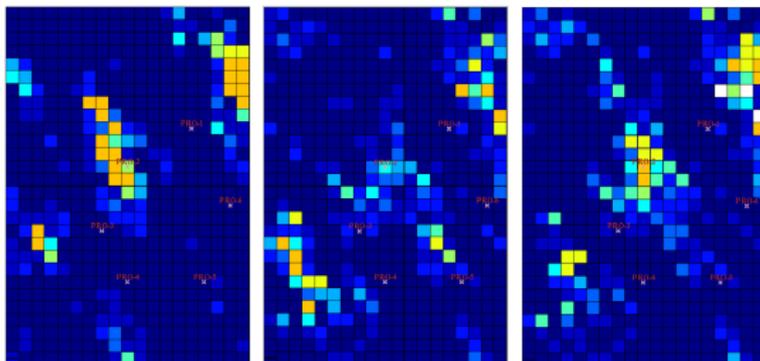
(b) GOR



(c) WC

Perméabilité du réservoir

- Cartes de perméabilités du réservoir



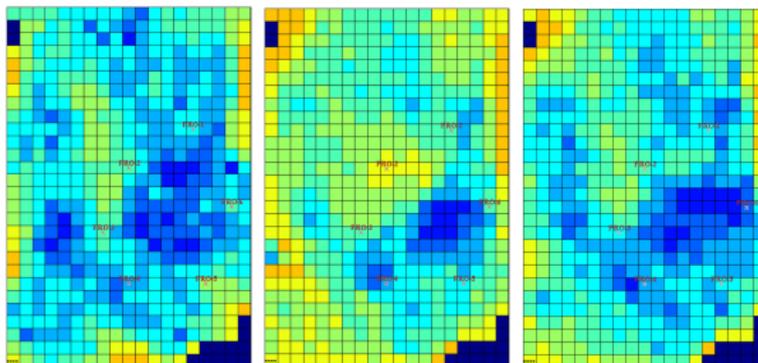
(d) Référence

(e) Initial

(f) Optimal

Saturation en huile du réservoir

- Cartes de saturation en huile du réservoir



(g) Référence

(h) Initial

(i) Optimal

Conclusion

- On propose une nouvelle méthode d'optimisation sans dérivée pour résoudre des problèmes inverses dans le domaine pétrolier.
 - Méthode sans dérivée
 - Adaptation aux fonctions partiellement séparables
 - Réduction du coût des modèles via exploitation d'une propriété d'auto-correction
- Cette méthode présente de très bons résultats, notamment dans le cas où un grand nombre de variables est à optimiser.