Ajuste Automático ao Histórico
Simulação de Reservatórios

Flavio Dickstein IM/UFRJ
Paulo Goldfeld IM/UFRJ
Renan V. Pinto LabMAPetro/UFRJ

Novembro/2018

Simulação sob incerteza

(i) O reservatório é inacessível

(ii) Não se pode medir diretamente as propriedades (porosidade, permeabilidade absoluta) do campo

(iii) É possível extrair amostras de rocha dos locais onde poços são perfurados

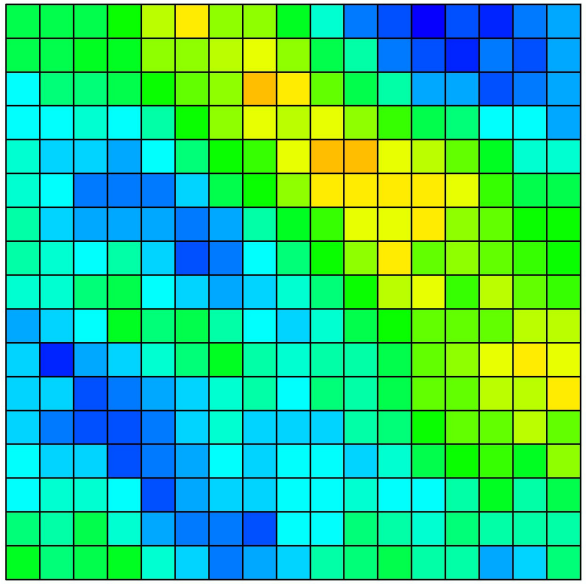
(iv) Medidas indiretas permitem obter informações sobre o campo

(v) Caracterizar o reservatório é um problema inverso, de natureza estatística, de difícil solução

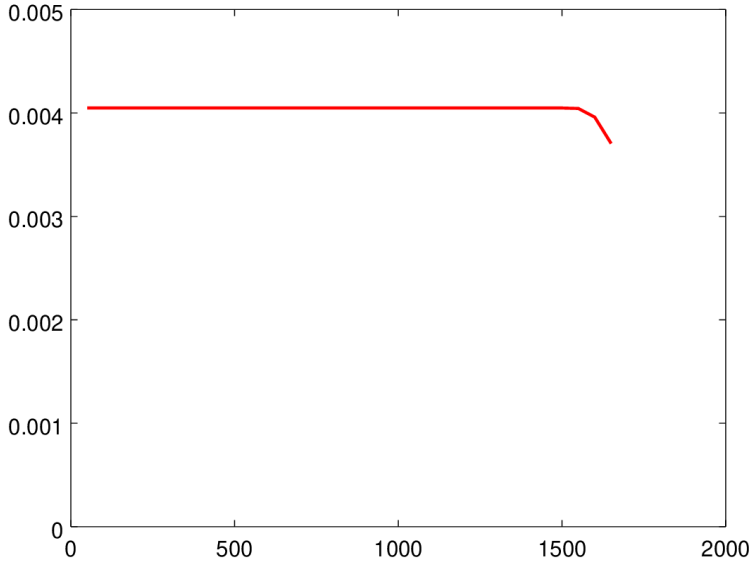
(vi) Ao final, há grande incerteza sobre o reservatório

(vii) Mas isto não impede que o campo seja explorado

Abordagem determinística

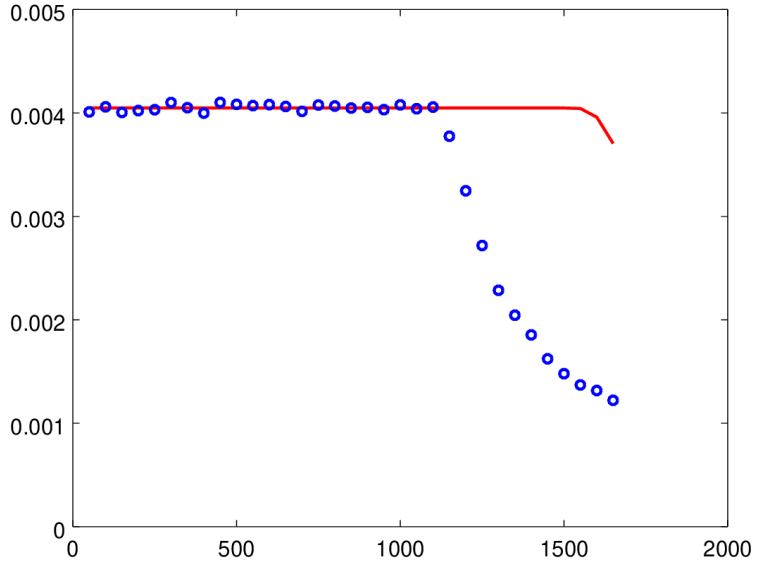


melhor modelo



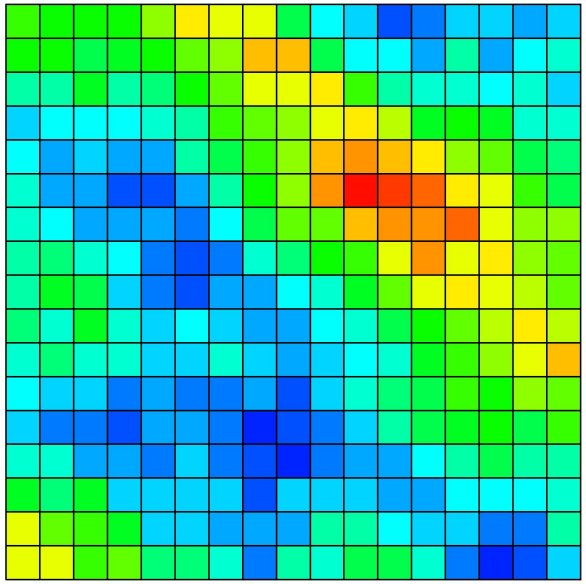
previsão

Abordagem determinística

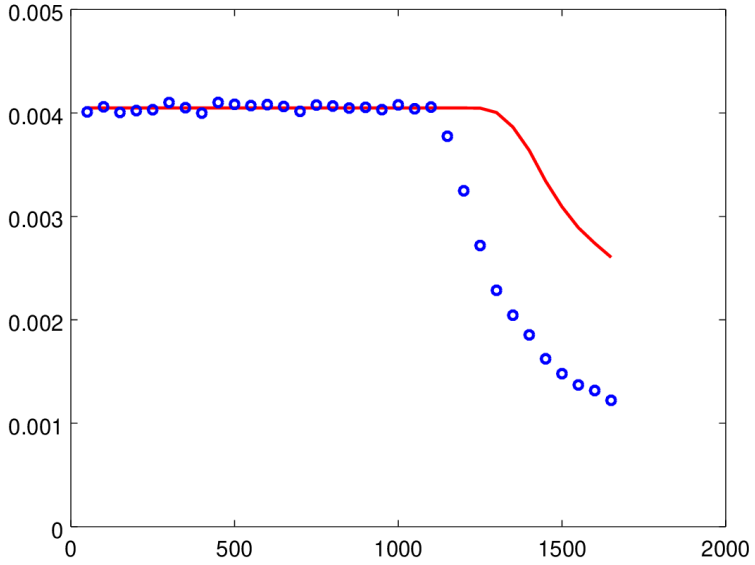


previsão $\times$ histórico

Ajuste ao Histórico determinístico

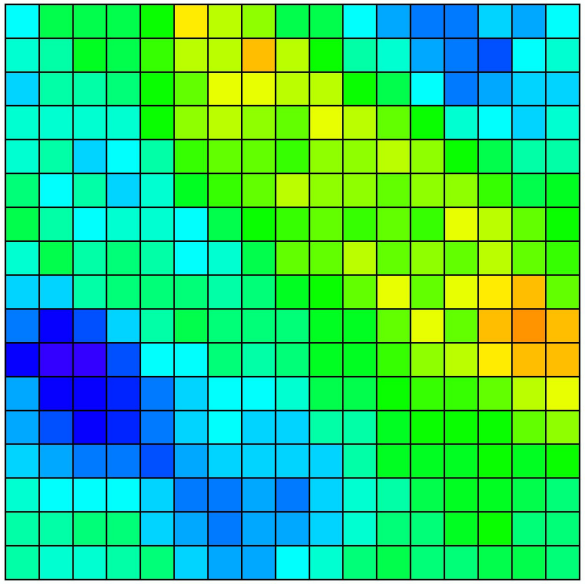


outro modelo

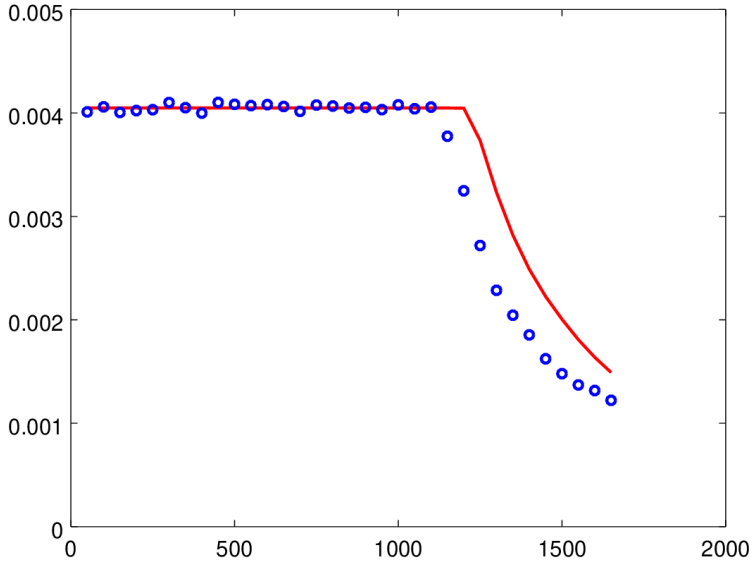


previsão $\times$ histórico

Ajuste ao Histórico determinístico

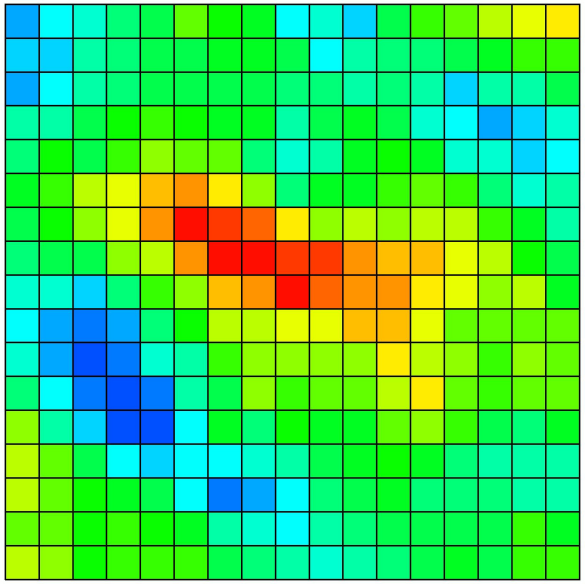


outro modelo

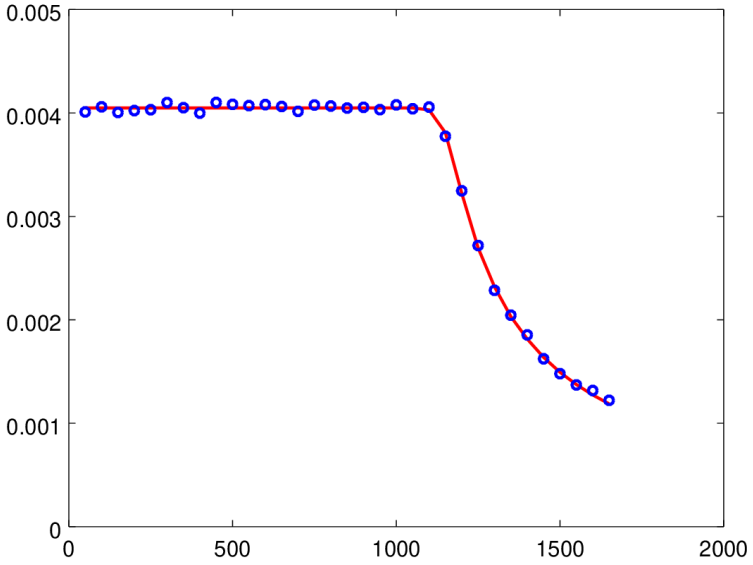


previsão $\times$ histórico

Ajuste ao Histórico determinístico

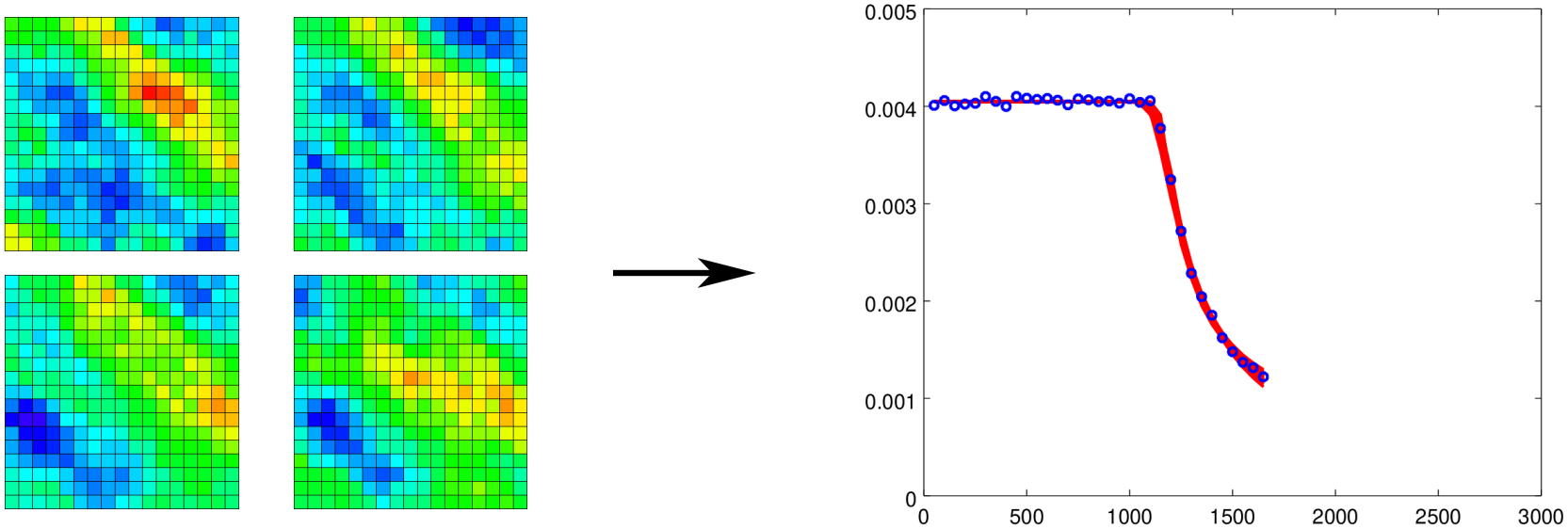


modelo ajustado



ajuste ao histórico

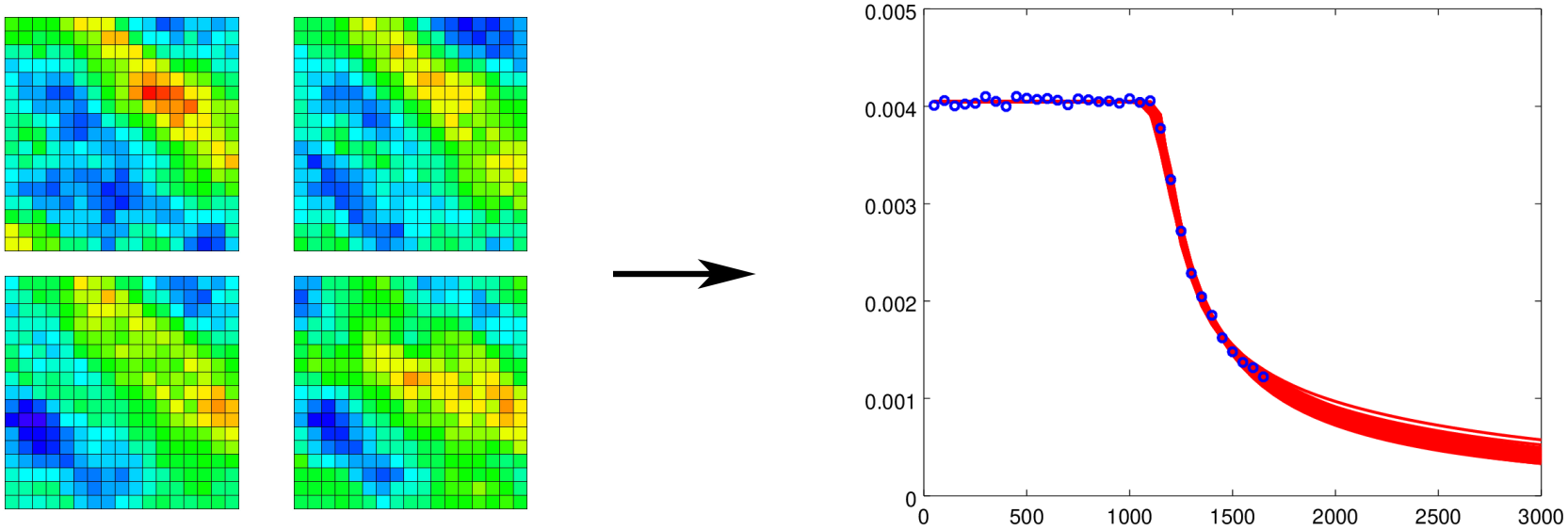
Múltiplas soluções



várias soluções

vários ajustes

Múltiplas soluções



várias soluções

previsões

A modelagem Bayesiana

Modelo $m \in \mathbb{R}^{N_M}$ descreve K, ϕ, ...

Assumimos um modelo *a priori* $m \sim \mathcal{N}(m_{\text{prior}}, C_M)$, com dp:

$$f_{\text{prior}}(m) \propto e^{-\frac{1}{2}\left\langle C_M^{-1}(m-m_{\text{prior}}),m-m_{\text{prior}}\right\rangle} = e^{-\frac{1}{2}\left\|m-m_{\text{prior}}\right\|^2_{C_M^{-1}}}.$$

A construção de C_M

Suponha dados n modelos $x_1, \ldots, x_n$. Sua média amostral é

$$\mu_a = \frac{1}{n} \sum_j x_j,$$

Considere $\overline{x}_j = x_j - \mu_a$ e $\overline{X}_{N_M \times n}$ a matriz

$$\overline{X} = [\overline{x}_1 \ \overline{x}_2 \ \ldots \overline{x}_n]$$

A covariância (amostral) de C_a é dada por

$$C_a = \frac{1}{n-1} \overline{X} \, \overline{X}^T.$$

Justificativa

Seja m uma variável aleatória de média μ e covariância C.

Definindo-se μ_a e C_a como no slide anterior, pelo Teorema Central do Limite,

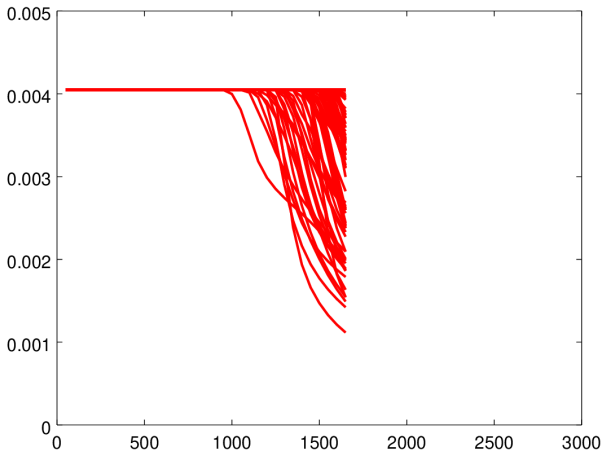
$$\mu_a \to \mu \ \text{ e } \ C_a \to C$$

quando $n \to \infty$.

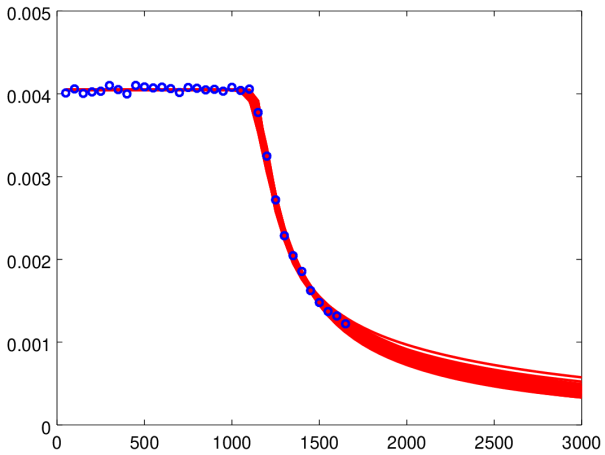
A distribuição condicionada

Ao longo do tempo, o histórico de dados de produção do campo é observado. Queremos encontrar a dp condicionada a esse histórico.

Seja $\delta = d_{\text{obs}} - g(m)$ a discrepância entre o histórico observado d_{obs} e o histórico associado a m, $g(m)$. Então queremos determinar a distribuição *a posteriori* $f_{\text{post}}(m|\delta)$.



priori



posteriori

Formulação Bayesiana

Regra de Bayes:

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$$

$$P(A|B) \propto P(B|A)P(A)$$

Analogamente, no caso contínuo,

$$f_{\text{post}}(m) = f_{\text{post}}(m|\delta) \propto \underbrace{f(\delta|m)}_{\text{verossimilhança}} f_{\text{prior}}(m).$$

O erro de medição

Note que, dado m, δ é simplesmente o erro de medição e.

Supondo que e tenha distribuição $\mathcal{N}_{0,C_D}$

$$f_\varepsilon(e) \propto e^{-\frac{1}{2}\left\langle C_D^{-1} e, e\right\rangle} = e^{-\frac{1}{2}\|e\|^2_{C_D^{-1}}},$$

então

$$f(\delta|m) \propto e^{-\frac{1}{2}\|g(m)-d_{\text{obs}}\|^2_{C_D^{-1}}}$$

(Tomamos C_D diagonal).

A distribuição *a posteriori*

Lembrando que $f_{\text{post}}(m) \propto f(\delta|m) f_{\text{prior}}(m)$, obtemos

$$f_{\text{post}}(m) \propto e^{-\frac{1}{2}\left(\|m - m_{\text{prior}}\|^2_{C_M^{-1}} + \|g(m) - d_{\text{obs}}\|^2_{C_D^{-1}}\right)}$$

Não basta encontrar um modelo que ajusta bem os dados. Devemos encontrar todos os **cenários razoáveis**. Mais exatamente, devemos amostrar m_{post}.

O MAP

Amostrar variáveis aleatórias não é fácil. O método mais conhecido é o de Metropolis-Hastings, que é um esquema Markov Chain Monte Carlo (MCMC).

Um objetivo mais simples é o de determinar m_{map}, o modelo de máximo *a posteriori* (MAP) que maximiza $f_{\text{post}}(m)$, isto é,

$$m_{\text{map}} = \arg\min J(m)$$

onde

$$J(m) = \|m - m_{\text{prior}}\|^2_{C_M^{-1}} + \|g(m) - d_{\text{obs}}\|^2_{C_D^{-1}}.$$

O esquema RML

No esquema Randomized Maximum Likelihood (RML), reescrevemos

$$J(m, m_{\text{prior}}, d_{\text{obs}}) = \|m - m_{\text{prior}}\|^2_{C_M^{-1}} + \|g(m) - d_{\text{obs}}\|^2_{C_D^{-1}}.$$

O RML amostra m_{prior} e d_{obs}, calculando

$$m_{\text{map}}(m_{\text{prior}}, d_{\text{obs}}) = \arg\min J(m, m_{\text{prior}}, d_{\text{obs}})$$

Isto nos conduz a resolver um número grande de problemas de Mínimos Quadrados Não-Linear (PMQNL).

O esquema RML

Quando $g(m)$ é linear, RML amostra corretamente a distribuição f_{post}.

Além disso,

$$J(m_{\text{map}}, m_{\text{prior}}, d_{\text{obs}}) = \|m_{\text{map}} - m_{\text{prior}}\|^2_{C_M^{-1}} + \|g(m_{\text{map}}) - d_{\text{obs}}\|^2_{C_D^{-1}}.$$

é uma v.a. χ^2 de média N_D e variância $2N_D$. Isto nos dá uma orientação para avaliar a qualidade dos resultados numéricos obtidos.

Resolução do PMQNL

Via Gauss-Newton: na iteração k, definimos

$$G_k = g'(m_k)$$

$$g_k(m) = g(m_k) + G_k(m - m_k)$$

$$J_k(m) = \|m - m_{\text{prior}}\|^2_{C_M^{-1}} + \|g_k(m) - d_{\text{obs}}\|^2_{C_D^{-1}}$$

Resolvemos o Problema de Mínimos Quadrados Linear (PMQL)

$$\widetilde{m}_k = \arg\min J_k(m)$$

e fazemos uma busca de linha

$$\min_{\alpha} J(m_k + \alpha(\widetilde{m}_k - m_k)).$$

Precondicionamento

Ocorre que o PMQL é **mal-condicionado**, porque

$$J_k(m) = \|m - m_{\text{prior}}\|_{C_M^{-1}} + \|g_k(m) - d_{\text{obs}}\|_{C_D^{-1}}$$

e C_M é mal-condicionada. Usamos então duas técnicas de precondicionamento.

(i) mudança de variáveis (Mahalanobis)

(ii) truncamento SVD (TSVD) (Componentes Principais)

Precondicionamento de Mahalanobis

Seja $C_M = XX^T$, de modo que $C_M^{-1} = X^{-T}X^{-1}$. Então,

$$\|m - m_{\text{prior}}\|_{C_M^{-1}} = \left\langle C_M^{-1}(m - m_{\text{prior}}), m - m_{\text{prior}} \right\rangle = \|X^{-1}(m - m_{\text{prior}})\|^2$$

Seja

$$\theta = X^{-1}(m - m_{\text{prior}}) \qquad m(\theta) = m_{\text{prior}} + X\theta$$

a transformação de Mahalanobis. Resolvemos a cada iteração

$$\min_\theta \|\theta\|^2 + \|g_k(m(\theta)) - d_{\text{obs}}\|^2_{C_D^{-1}} = \min_\theta \|\theta\|^2 + \|\underbrace{C_D^{-1/2} G_k X}_{A}\theta - b\|^2.$$

Equações normais: $(I + A^T A)\theta = c$.

Observação 1: A decomposição $C_M = XX^T$ pode ser evitada.

Observação 2: O problema continua mal-condicionado.

TSVD-θ

Solução para o mal-condicionamento: SVD truncada (TSVD).

Se $A = U\Sigma V^T$, aproximamos

$$A \approx A_j = U_j \Sigma_j V_j^T$$

e resolvemos

$$\min_{(\theta-\theta_k)\in \text{range}(V_j)} \|\theta\|^2 + \|A\theta - b\|^2,$$

cuja solução é

$$\theta = \theta_k + \sum_{i=1}^{j} \frac{\left\langle v_i, A^T(b - A\theta_k) - \theta_k \right\rangle}{\sigma_i^2 + 1} v_i.$$

O Esquema TCG

O Esquema de **Gradientes Conjungados Truncado** (TCG) para resolver o PMQNL

$$\min J(m) = \|m - m_{\text{prior}}\|^2_{C_M^{-1}} + \|g(m) - d_{\text{obs}}\|^2_{C_D^{-1}}$$

(i) Na iteração k, considera o PMQL $\min J_k(m)$.

(ii) Resolve $\min J_k(m(\theta))$, onde $m(\theta) = X\theta + m_{\text{prior}}$.
Equação normal: $(I + X^T G_k^T C_D^{-1} G_k X)\theta = *$

(iii) Resolve a equação acima via CG truncado.

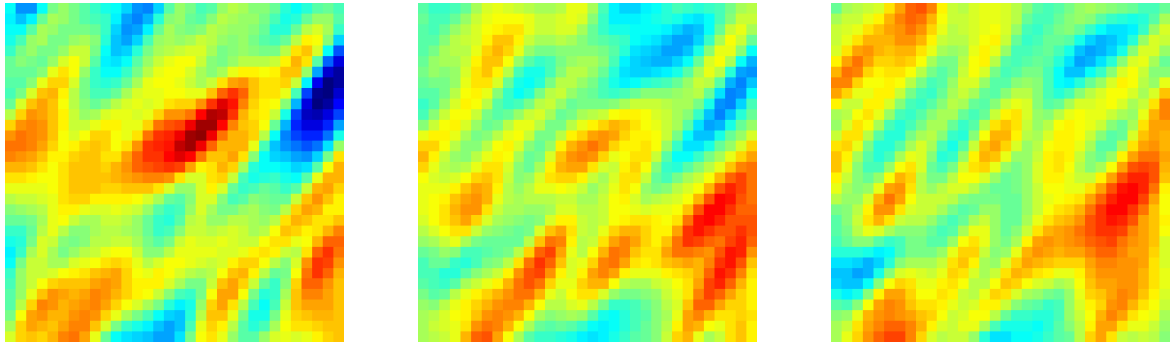
(iv) Encontrada uma direção $m(\theta)$, executa uma busca de linha para encontrar m_{k+1}.

Obs.: Usamos RML para amostrar f_{post}.

Resultados Numéricos

Caso 1 - “Paralelepípedo”

5-spot (1 inj., 4 prods.), $32 \times 32 \times 3$
$m =$ permeabilidade horizontal

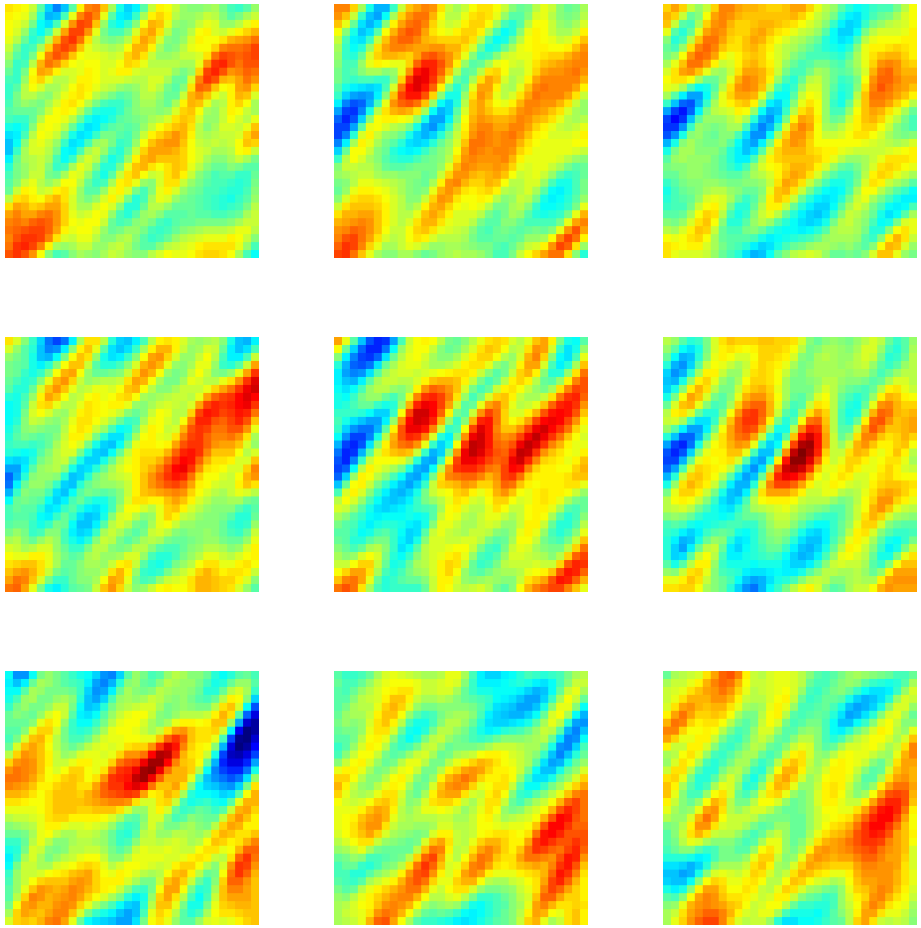


Campo Recuperado

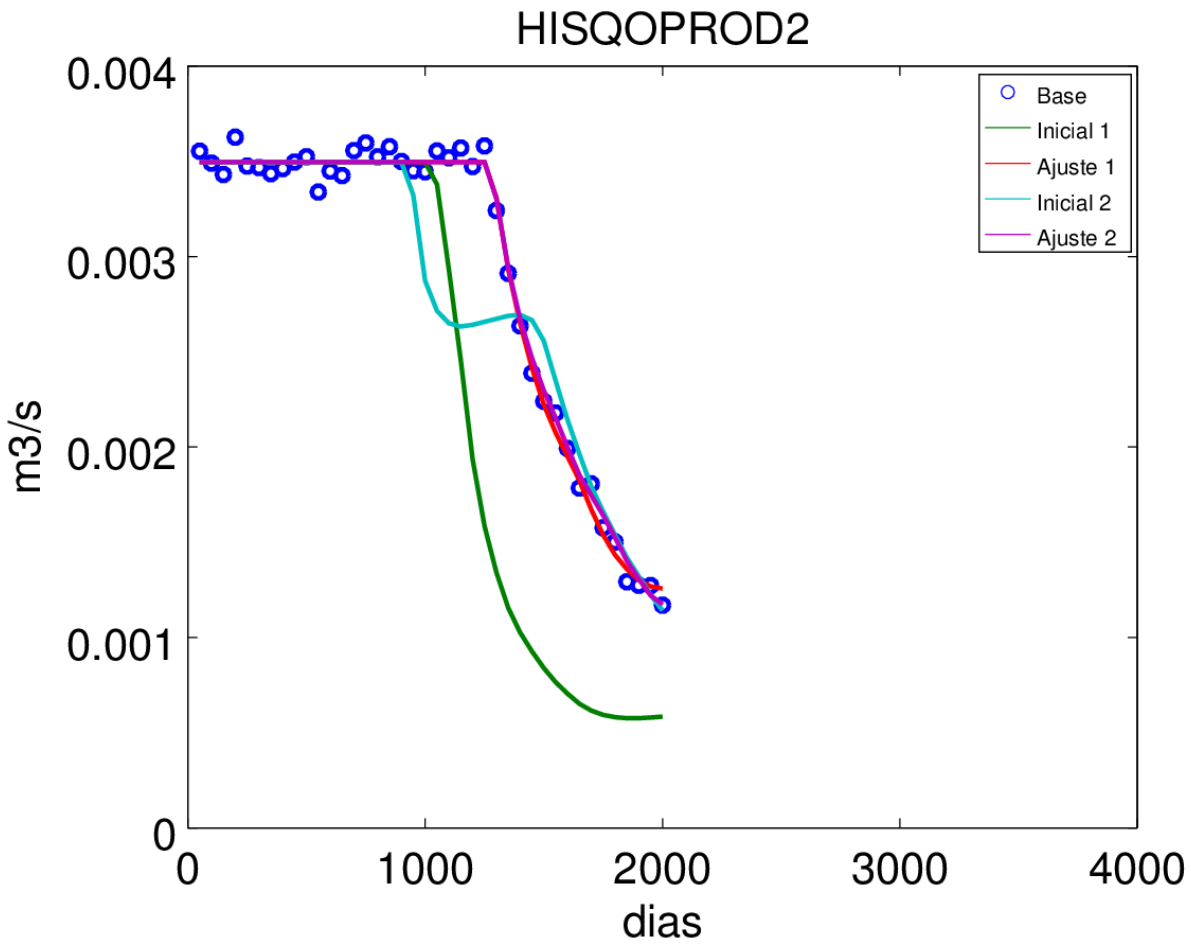
inicial

ajustado

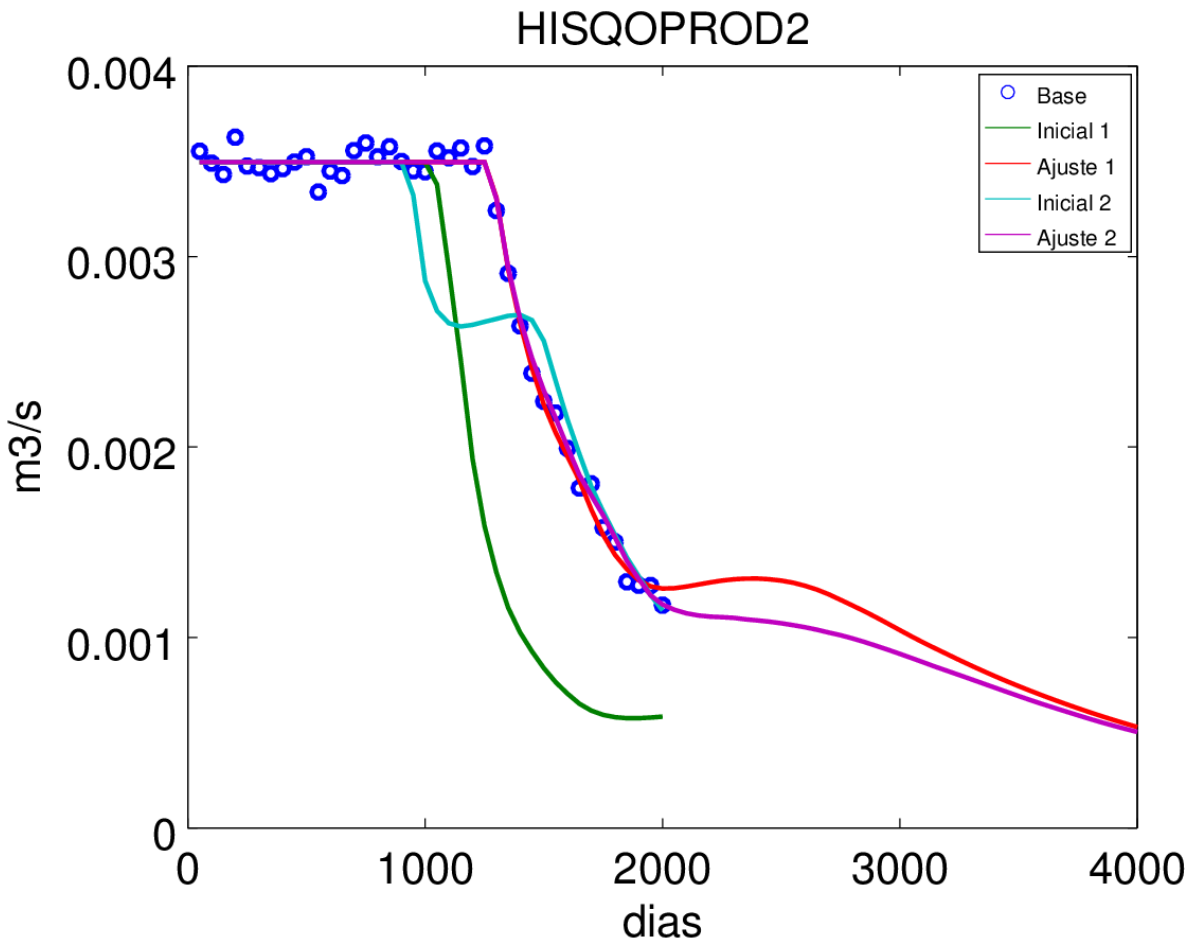
true



Ajuste

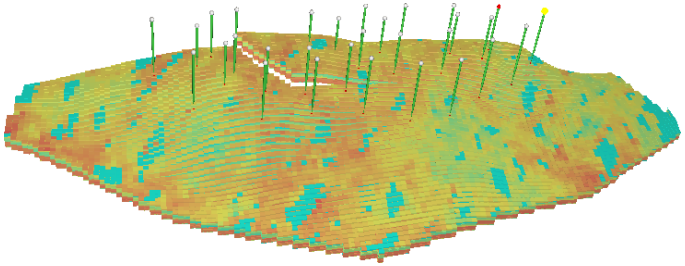

HISQOPROD2
0.004
0.003
0.002
0.001
0
m3/s
0
1000
2000
3000
4000
dias
Base
Inicial 1
Ajuste 1
Inicial 2
Ajuste 2

Extrapolação

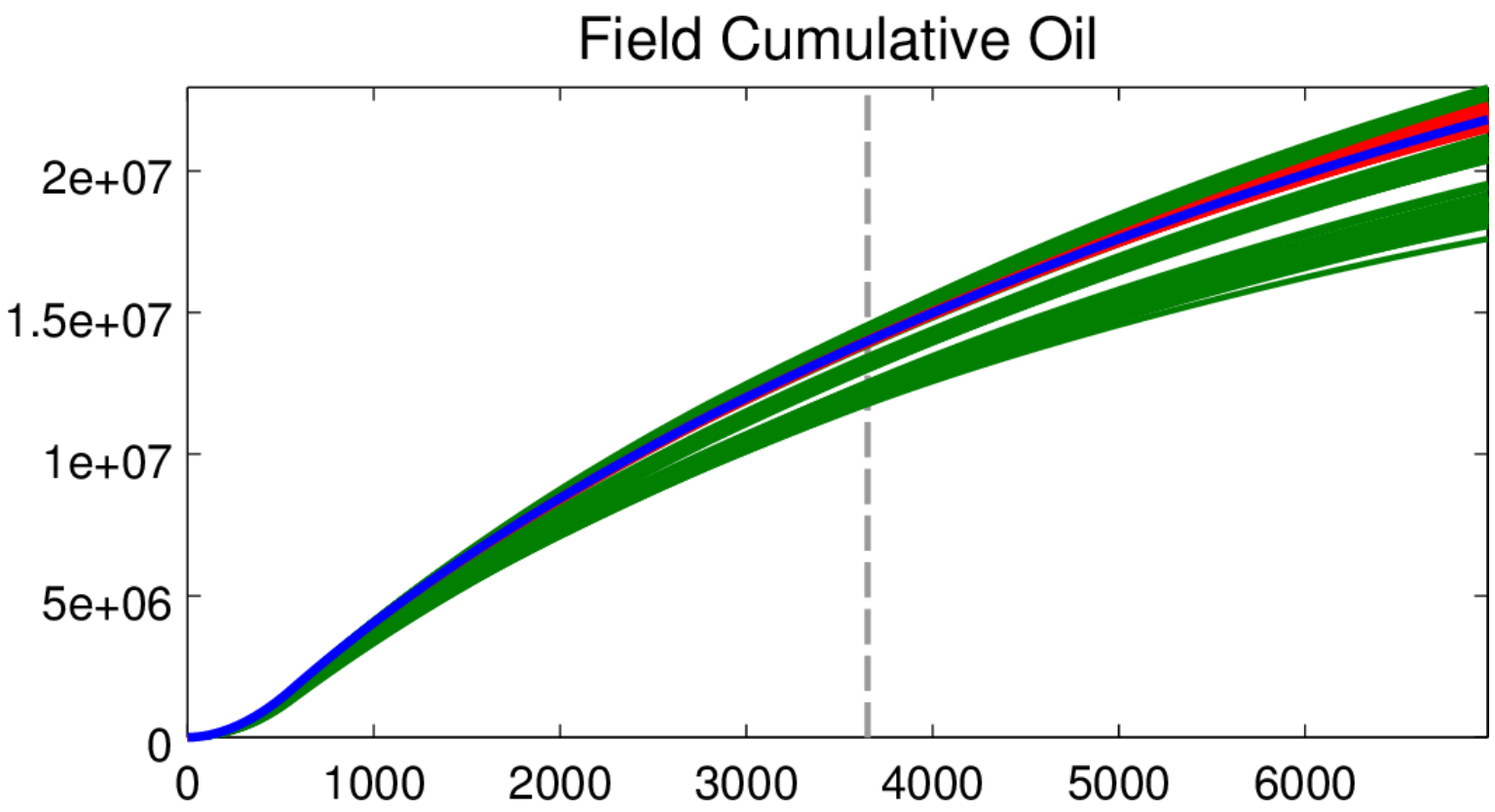


Brugge

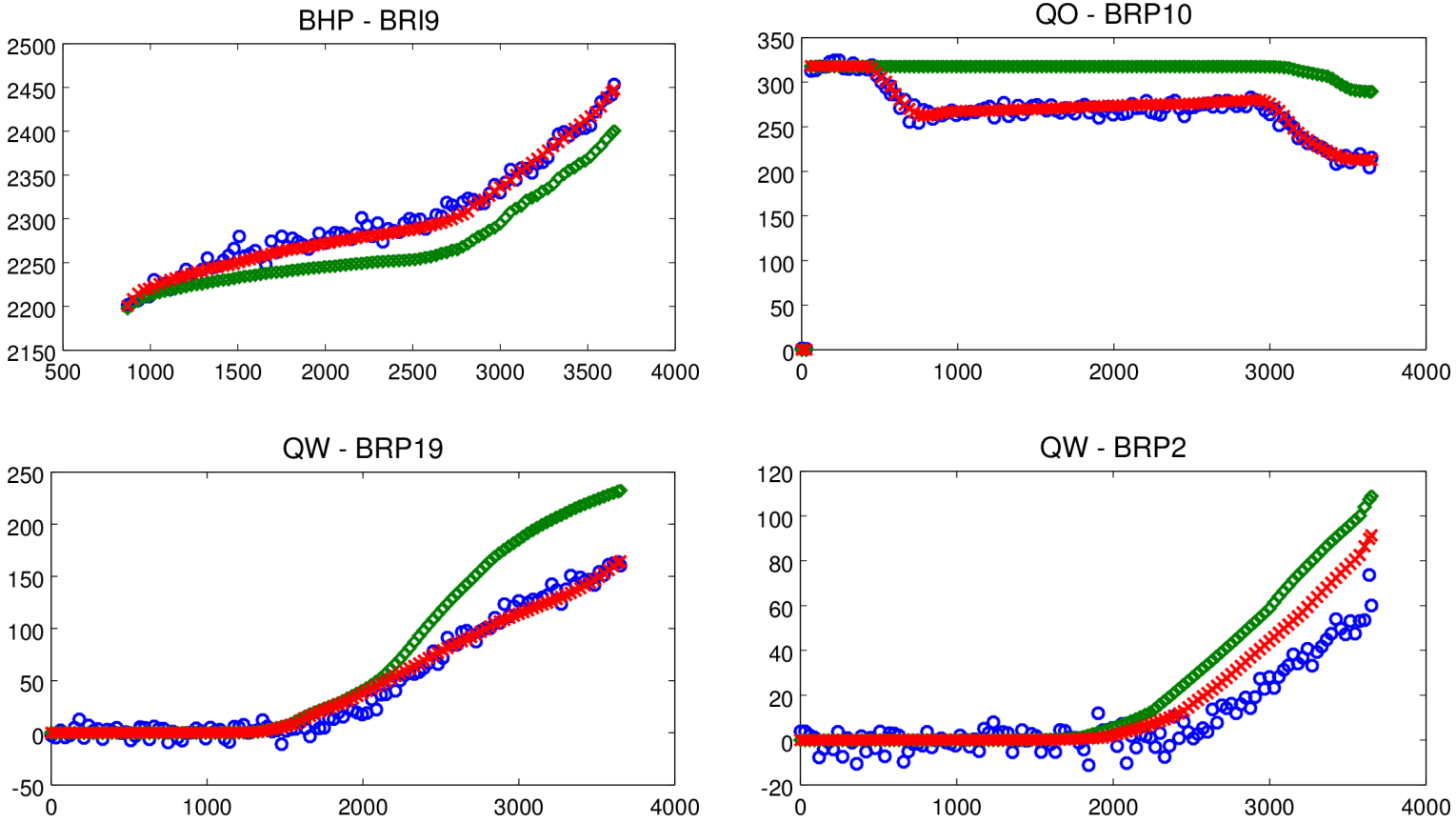
- Caso *benchmark*.
- 10 km $\times$ 3 km $\times$ 60 m
- malha 139 $\times$ 48 $\times$ 9
- 20 produtores (vazão de líquido / pressão prescrita)
- 10 injetores (vazão de água / pressão prescrita)



Brugge


Field Cumulative Oil
2e+07
1.5e+07
1e+07
5e+06
0
0
1000
2000
3000
4000
5000
6000

Brugge


BHP - BRI9
2500
2450
2400
2350
2300
2250
2200
2150
500
1000
1500
2000
2500
3000
3500
4000
QO - BRP10
350
300
250
200
150
100
50
0
0
1000
2000
3000
4000
QW - BRP19
250
200
150
100
50
0
-50
0
1000
2000
3000
4000
QW - BRP2
120
100
80
60
40
20
0
-20
0
1000
2000
3000
4000

Conclusão

- Uma melhor caracterização do campo pode ter um alto impacto na eficiência da recuperação. O Ajuste ao Histórico é um importante problema em Simulação de Reservatórios.

- Muitos progressos recentes foram feitos na área.

- No entanto, restam sérios desafios operacionais e tecnológicos para a resolução do problema.

- Há também sérios desafios científicos.

Conclusão

- Métodos geoestatísticos para melhor caracterização *a priori*.

- Métodos mais eficientes para a resolução do problema direto. Álgebra linear computacional.

- Correta amostragem da distribuição *a posteriori*.

- Esquemas de ajuste que aproximam derivadas.

- Esquemas que não usam derivadas (e.g., EnKF).