

Utilização de Modelos Autoregressivos na Quantificação de Incertezas em Problemas de Transporte Linear

Thiago Jordem Pereira

Depto de Ciências Exatas, Biológicas e da Terra, INFES, UFF,
28470-000, Santo Antônio de Pádua, RJ
E-mail: tjordem@gmail.com,

Helio Pedro Amaral Souto

Depto de Modelagem Computacional, Instituto Politécnico, UERJ,
28601-970, Nova Friburgo, RJ
E-mail: helio@iprj.uerj.br.

Resumo: *O objetivo deste trabalho é o de utilizar modelos autoregressivos na quantificação de incertezas em problemas de escoamentos miscíveis em meios porosos heterogêneos empregando uma abordagem Bayesiana para a seleção dos campos de permeabilidades, baseada em um conjunto de medições da concentração do traçador em pontos específicos do meio poroso. O método da Soma Sucessiva de Campos Gaussianos Independentes (SSCGI) é utilizado na parametrização das incertezas contidas nos meios porosos heterogêneos. Na resolução do problema inverso, utiliza-se um método do tipo Monte Carlo via Cadeias de Markov a dois estágios. Através deste procedimento, gera-se uma cadeia Markov que converge para a distribuição estacionária, que neste caso é a distribuição a posteriori de interesse. Resultados numéricos são apresentados para um conjunto de realizações dos campos de permeabilidades.*

Palavras-chave: *Campos estocásticos Gaussianos, Escoamento em meios porosos, Métodos MCMC, Traçador*

1 Introdução

O estudo dos escoamentos multifásicos em meios porosos heterogêneos representa uma ferramenta útil em diversas áreas importantes para o desenvolvimento científico e tecnológico, veja, e.g, [1]-[8] e referências lá citadas. As técnicas que empregam a injeção de traçadores vêm se mostrando viáveis no estudo de problemas relacionados à simulação numérica de escoamentos miscíveis de fluidos incompressíveis em reservatórios de petróleo, bem como no transporte de contaminantes em reservatórios subterrâneos.

Matematicamente, o escoamento do traçador é modelado por um sistema de equações diferenciais parciais sujeito a certas condições de contorno e inicial. Este sistema de equações diferenciais é constituído de um sub-sistema elíptico, que determina a pressão e a velocidade da mistura, e de uma equação de transporte, que determina em cada instante de tempo o valor da concentração do fluido injetado (traçador). A obtenção de soluções numéricas deste sistema de equações diferenciais parciais não é uma tarefa fácil visto que elas são fortemente influenciadas pelas propriedades do meio poroso, tais como a porosidade e a permeabilidade, cuja caracterização precisa via a construção de campos de porosidades e permeabilidades ainda é um problema complexo. Estas propriedades são responsáveis pela introdução das incertezas contidas no modelo matemático, que ocorrem em múltiplas escalas, e que reduzem de forma significativa a confiabilidade das previsões obtidas mediante a resolução destas equações. Desta forma, uma grande importância vem sendo dada à aplicação de diferentes tipos de metodologias capazes de

quantificar e reduzir as incertezas dos modelos geológicos e, conseqüentemente, obter previsões acuradas para, por exemplo, a produção de óleo em reservatórios de petróleo e o monitoramento do transporte de contaminantes em formações subterrâneas [5].

A abordagem Bayesiana pode ser utilizada para a atualização estatística dos parâmetros físicos (permeabilidade ou porosidade), baseando-se no conhecimento prévio dos valores da concentração do traçador em alguns pontos específicos do meio poroso. A atualização da informação do parâmetro de interesse é feita através do Teorema de Bayes, levando-se em consideração a distribuição *a priori* e a distribuição amostral dos dados (expressa formalmente pela função de verosimilhança), obtendo-se a distribuição *a posteriori* do parâmetro de interesse. Na inferência Bayesiana, os procedimentos numéricos baseados em simulações estocásticas têm sido bastante utilizados para explorar a complexa distribuição *a posteriori*, dentre os quais pode-se destacar o método MCMC a dois estágios [8], abreviatura do inglês *Markov Chain Monte Carlo* (Monte Carlo via Cadeias de Markov). Este método gera cadeias de Markov ergódicas cuja distribuição estacionária é a distribuição de probabilidades *a posteriori* do problema de escoamento.

2 O Modelo de Escoamento do Traçador

O modelo matemático, na escala de laboratório ou de Darcy (escala fina), que governa o escoamento de traçadores em um meio poroso Ω é constituído por uma equação de transporte linear da concentração do fluido marcado

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \nabla c = 0, \tag{1}$$

e o campo de velocidades que é determinado pela lei de Darcy juntamente com a hipótese de incompressibilidade do fluido

$$\mathbf{u} = -\frac{k(\mathbf{x})}{\mu} \nabla p, \quad \nabla \cdot \mathbf{u} = 0, \tag{2}$$

onde k é um campo escalar de permeabilidades, com distribuição log-normal, a ser especificado, \mathbf{u} e p são as velocidades de Darcy e a pressão da mistura fluida. O coeficiente μ representa a viscosidade da mistura e c a concentração do fluido marcado. Neste modelo, supõe-se que o meio poroso é indeformável e a porosidade é igual a uma constante. Além disso, os efeitos da gravidade, da compressibilidade e da dispersão na escala dos poros foram negligenciados.

As Equações (1) e (2) são definidas em um domínio retangular $\Omega = [0, L_x] \times [0, L_y]$, com contorno regular $\partial\Omega$ e, um intervalo de tempo $0 \leq t \leq T$, com condições de contorno dadas por

$$\begin{aligned} \mathbf{u} \cdot \mathbf{n} &= -q, & \text{in } x &= 0, \\ p &= 0, & \text{in } x &= L_x, \\ \mathbf{u} \cdot \mathbf{n} &= 0, & \text{in } y &= 0 \text{ e } L_y, \end{aligned} \tag{3}$$

e condição inicial do tipo Riemann

$$c(x, 0) = \begin{cases} 1, & \text{se } x \leq 0, \\ 0, & \text{se } x > 0, \end{cases} \tag{4}$$

onde \mathbf{n} é o vetor normal unitário exterior a $\partial\Omega$ e x representa a variável espacial na direção do escoamento médio uniforme.

O método MCMC a dois estágios utiliza um modelo matemático na escala de campo (malha grossa) a fim de reduzir o esforço computacional e aumentar a taxa de aceitação das amostras propostas do campo de permeabilidades [6, 8]. Neste caso, deve-se utilizar uma técnica de *ups-caling* para modelar o escoamento do traçador na escala de campo, cuja estratégia utilizada neste processo de mudança de escala pode ser encontrada em [4]. A principal ideia desta técnica

consiste na derivação das equações na escala de campo e de um campo de permeabilidades equivalente, apropriado para a simulação numérica na escala de campo do problema do escoamento do traçador no interior do meio poroso [6, 8].

Na resolução do problema de pressão-velocidade (2), para as malhas fina e grossa, as equações diferenciais parciais são discretizadas pelo método dos elementos finitos mistos, que são adequados para o cálculo acurado dos fluxos nos campos de permeabilidades heterogêneos [1]. Já na resolução da equação de advecção que governa o transporte do traçador (1) emprega-se o método lagrangeano localmente conservativo *Forward Integral Tracking* (FIT) [2].

3 Modelagem Estocástica da Geologia

Para a parametrização das incertezas contidas nos meios porosos heterogêneos, emprega-se campos de permeabilidades escalares $k(\mathbf{x})$ com distribuição log-normal da forma:

$$k(\mathbf{x}) = k_0 e^{\varphi Y(\mathbf{x})}, \tag{5}$$

onde $k_0 \in \mathbb{R}^+$ é uma constante, $\varphi \in \mathbb{R}^+$ é o fator que controla o grau de heterogeneidade do meio e $Y(\mathbf{x})$ é uma campo aleatório Gaussiano caracterizado por sua média $\langle Y(\mathbf{x}) \rangle = 0$ e sua função de covariância dada por uma lei de potência [3]

$$\mathcal{C}(\mathbf{x}, \mathbf{y}) = \langle Y(\mathbf{x})Y(\mathbf{y}) \rangle = |\mathbf{x} - \mathbf{y}|^{-\beta}, \quad \beta > 0, \tag{6}$$

onde $\langle \cdot \rangle$ indica a média do conjunto e β é o coeficiente de Hurst [8].

Gera-se os campos aleatórios $Y(\mathbf{x})$ utilizando o método das Soma Sucessiva de Campos Gaussianos Independentes (SSCGI) [3]. Este eficiente procedimento gera campos aleatórios Gaussianos baseando-se na superposição hierárquica de malhas uniformes. A sua versão modificada, adequada para a geração de campos aleatórios condicionados, necessários para a implementação dos métodos computacionais Bayesianos pode ser encontrada em [8].

4 O Método de Monte Carlo via Cadeias de Markov

Neste trabalho, as simulações de quantificação de incertezas são empregadas com o intuito de atualizar os campos de permeabilidades $k(\mathbf{x})$, baseado-se no conhecimento prévio de um conjunto de valores medidos da concentração do traçador em alguns pontos específicos do meio poroso $O(c_o)$. As informações referentes aos valores observados da concentração são introduzidos no modelo através da função de verossimilhança $P(O(c_o)|k(\mathbf{x}))$. Esta informação adicional é utilizada na seleção dos novos campos de permeabilidades através do Teorema de Bayes [5]

$$\pi(k) \propto P(O(c_o)|k(\mathbf{x})) P(k(\mathbf{x})), \tag{7}$$

onde $P(k(\mathbf{x}))$ e $\pi(k) = P(k(\mathbf{x})|O(c_o))$ são, respectivamente, as distribuições de probabilidade *a priori* e *a posteriori* do campo de permeabilidades $k(\mathbf{x})$. Supõe-se que a função de verossimilhança segue uma distribuição normal [5], isto é,

$$P(O(c_o)|k(\mathbf{x})) \propto \exp \left\{ -\frac{\sum_{j=1}^{N_t} \sum_{i=1}^{N_s} [c_s(\mathbf{x}_i, t_j) - c_o(\mathbf{x}_i, t_j)]^2}{2\sigma_f^2} \right\}, \tag{8}$$

onde $c_s(\mathbf{x}_i, t_j)$ e $c_o(\mathbf{x}_i, t_j)$ são a concentração simulada e observada do traçador, respectivamente. O parâmetro σ_f^2 é a precisão associada às medidas simuladas e observadas, N_t representa o número de vezes que a concentração é avaliada no tempo e N_s o número de sensores.

A ideia básica dos métodos MCMC é simular Cadeias de Markov ergódicas $\{k^{(t)}\}$ que têm como sua distribuição estacionária (ou de equilíbrio) a distribuição *a posteriori* de interesse $\pi(k)$. Para a construção destas Cadeias de Markov, os métodos MCMC fazem uso de algoritmos

específicos, nos quais pode-se destacar o algoritmo de Metropolis-Hastings (consultar [8] e as referências lá indicada). Para cada iteração, um campo de permeabilidades k é proposto usando uma distribuição instrumental $q(k|k^{(t)})$, onde o instante t representa o estado atual em que se encontra a cadeia de Markov. Em seguida, resolve-se os problemas de pressão-velocidade (2) e do transporte do traçador (1). Então, o candidato k será aceito como o próximo estado da cadeia com probabilidade dada por

$$\alpha(k|k^{(t)}) = \begin{cases} \min\left(\frac{\pi(k)q(k^{(t)}|k)}{\pi(k^{(t)})q(k|k^{(t)})}, 1\right) & \pi(k^{(t)})q(k|k^{(t)}) > 0, \\ 1 & \pi(k^{(t)})q(k|k^{(t)}) = 0, \end{cases} \quad (9)$$

tal que $k^{(t+1)} = k$ com probabilidade $\alpha(k|k^{(t)})$ e $k^{(t+1)} = k^{(t)}$ com probabilidade $1 - \alpha(k|k^{(t)})$.

Considera-se, em seguida, o caso especial do algoritmo de Metropolis-Hastings baseado em passeios aleatórios (*random walk*). O método SSCGI gera, numericamente, campos aleatórios Gaussianos a partir da superposição hierárquica de malhas uniformes [3], isto é,

$$Y(\mathbf{x}) \equiv \sum_{n=0}^{\infty} \underbrace{\left(\sum_{m=1}^{b^n} \xi_{nm}(\mathbf{x}) \right)}_{\xi_n(\mathbf{x})} \quad (10)$$

onde $b^n \times b^n$, com $n = 0, 1, 2, \dots$, é o tamanho de cada elemento e $m = 1, 2, \dots, b^n$ é o número de malhas em cada nível n . A técnica do passeio aleatório (*random walk*) é aplicada à variável $\xi_{nm}(\mathbf{x})$ ao invés da variável $\xi_n(\mathbf{x})$, ou $Y(\mathbf{x})$, a fim de manter-se a mesma estrutura de construção dos campos Gaussianos $Y(\mathbf{x})$. Assim, propõe-se $Y(\mathbf{x})$ a partir do novo valor de $\xi_{nm}(\mathbf{x})$ dado por

$$\xi_{nm}(\mathbf{x}) = \sqrt{1 - h_\xi} \xi_{nm}^{(t)}(\mathbf{x}) + \sqrt{h_\xi} z_\xi, \quad (11)$$

onde $n = 0, 1, 2, \dots$ e $m = 1, 2, \dots, b^n$. O parâmetro h_ξ determina o tamanho do salto em cada passo do algoritmo de Metropolis-Hastings e z_ξ é gerado a partir de uma distribuição normal com média 0 e variância 1, isto é, $z_\xi \sim N(0, 1)$. O processo de amostragem (11) é do tipo *random walk* onde emprega-se uma distribuição instrumental autoregressiva [7]. A utilização deste tipo de distribuição é bastante vantajosa, pois não se tem o problema de renormalização.

Geralmente, as simulações diretas do método MCMC apresentam a desvantagem de necessitar de um elevado esforço computacional, uma vez que um grande número de simulações é necessário para que obtenha-se a convergência da cadeia de Markov e uma boa taxa de aceitação. Portanto, a fim de minimizar este problema recorre-se a uma eficiente e rigorosa metodologia, conhecida como método MCMC a dois estágios [6], que pode ser descrito pelo Algoritmo 1.

5 Resultados Numéricos

As simulações foram realizadas com um código numérico serial considerando um tempo máximo de simulação de até 1.700 dias, em uma região física bidimensional com dimensões $L_x = 128$ m e $L_y = 128$ m. As simulações foram feitas empregando-se diferentes malhas na escala grossa: 16×16 , 32×32 e 64×64 elementos. A malha usada na escala fina foi mantida fixa e possui 128×128 elementos. O passo de tempo utilizado na resolução da equação de transporte linear foi de $\Delta t = 1$ dia. Em todas as simulações o traçador é injetado uniformemente a uma taxa constante de 0,2 volumes porosos a cada ano. O número inicial de pontos sobre a frente de avanço do traçador é de 256 pontos. Para evitar a influência da injeção nos resultados numéricos, no instante inicial especifica-se que o salto esteja localizado a 8 m de distância da fronteira esquerda do reservatório [8]. Na caracterização da rocha do reservatório utiliza-se campos Gaussianos aleatórios obtidos com covariância espacial fractal. O expoente de Hurst (β) é igual a 0,5 e o fator que controla o grau de heterogeneidade do meio (f_f) foi fixado em 0,4.

Para a obtenção dos resultados numéricos foram propostas 70.000 amostras, nas quais selecionou-se conjuntos de realizações com 1.000 amostras aceitas de campos de permeabilidades. O parâmetro de perturbação h_ξ do processo de amostragem do tipo *random walk* foi igual a 0,0025 e o valor da precisão da medida observada foi $\sigma_g^2 = 4\sigma_f^2$, com $\sigma_f^2 = 1,0$. As localizações dos sensores $S_i = S_i(x, y)$, $i = 1, \dots, 6$, de monitoramento são: $S_1=(35,0, 32,0)$, $S_2=(35,0, 64,0)$, $S_3=(35,0, 96,0)$, $S_4=(75,0, 32,0)$, $S_5=(75,0, 64,0)$ e $S_6=(75,0, 96,0)$. O valor prescrito da permeabilidade em cada sensor é $k_0 \neq 0$ e a concentração $c(x_i, t_j)$ foi avaliada a cada 1 dia de simulação. Os valores de referência são obtidos gerando-se um campo de permeabilidades de referência e resolvendo-se acuradamente os problemas de pressão-velocidade (2) e de transporte do traçador (1), em uma malha fina.

Na Tabela 1 são apresentados as taxas de aceitação, para as 1.000 primeiras amostras aceitas de campos de permeabilidades, obtidas com as malhas grossas de 16×16 , 32×32 e 64×64 elementos. Observa-se que quando passa-se de uma malha de 16×16 elementos para uma malha de 32×32 e 64×64 a taxa de aceitação reduz 3,2% e 2,1%, respectivamente. Estes resultados são esperados, uma vez que a taxa de aceitação é fortemente influenciada pelos valores da precisão das medidas, σ_f e σ_g , e dos erros nas malhas grossa e fina.

Tabela 1: Taxas de aceitação para 1.000 primeiras amostras aceitas.

Método MCMC a dois estágios			
Malhas na escala grossa	16×16	32×32	64×64
Propostas	49.430	68.002	55.135
Aceitas: malha grossa	10.738	16.365	13.861
Aceitas: malha fina	1.000	1.000	1.000
Taxa de aceitação	9,31%	6,11%	7,21%

A Figura 1 apresenta a variação do erro na malha fina em função dos primeiros 1.000 campos de permeabilidades aceitos, obtidos com malhas na escala grossa de 16×16 , 32×32 e 64×64 elementos.

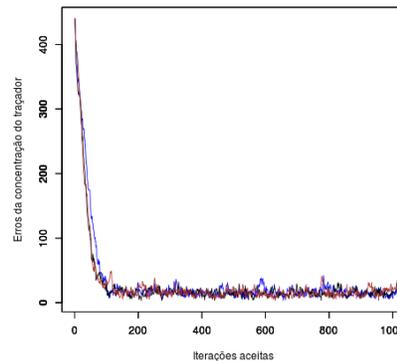


Figura 1: Erro na malha fina da concentração do traçador em função das amostras aceitas.

Pode-se observar que para todas as malhas as cadeias de Markov convergiram para a distribuição estacionária de interesse. Os erros decresceram de valores da ordem de aproximadamente $5 \cdot 10^2$ para $10 \cdot 10^0$. Os valores baixos destes erros são a garantia de que os campos aceitos possam fornecer, para a concentração simulada, valores próximos aos medidos nos sensores. Esse fato pode ser verificado observando-se a Figura 2, que contém os gráficos dos perfis médios de concentração simulada do traçador nas posições correspondentes à localização dos diferentes sensores instalados no meio poroso.

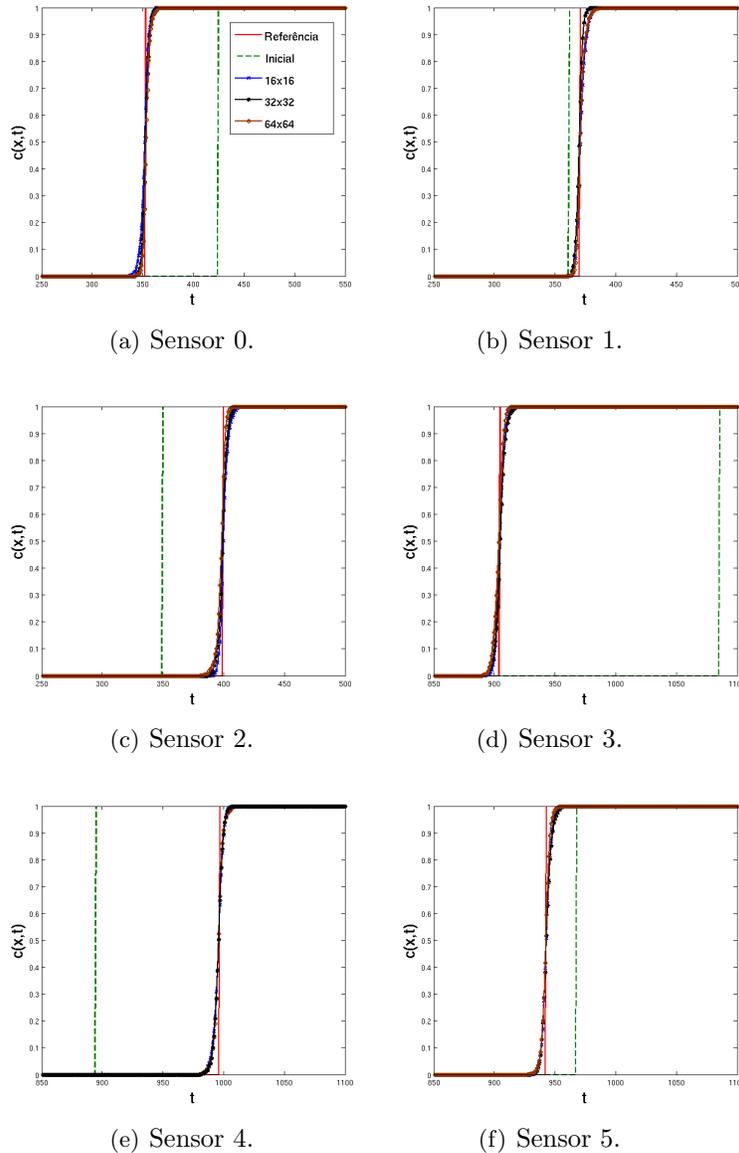


Figura 2: Valores da concentração do traçador em diferentes sensores.

Foi levado em consideração um conjunto de 800 amostras aceitas dos campos de permeabilidades, após o período de aquecimento de 200 campos aceitos. Como se pode observar, em todos os casos, existe uma boa concordância entre os perfis da concentração média simulada do traçador e os observados.

6 Conclusão

Neste trabalho foi utilizado uma abordagem Bayesiana para a atualização dos campos de permeabilidades, baseada no conhecimento prévio dos valores da concentração do traçador e da permeabilidade em alguns pontos específicos do meio poroso. O método da Soma Sucessiva de Campos Gaussianos Independentes foi usado para gerar as amostras dos campos de permeabilidades na solução do problema inverso estocástico associado ao escoamento do traçador em formações heterogêneas. Na solução do problema inverso complexo, foi utilizado o método Monte Carlo via Cadeias de Markov a dois estágios, cujo o processo de amostragem é do tipo *random walk* onde emprega-se uma distribuição instrumental autoregressiva. Os resultados numéricos obtidos indicam que esta abordagem é promissora para a caracterização dos campos de perme-

abilidades. Como perspectiva futura do presente trabalho cita-se, por exemplo, a utilização do método MCMC a dois estágios na quantificação de incertezas de problemas envolvendo escoamentos multifásicos.

Algoritmo 1: Método MCMC a dois estágios.

- 1 **Passo:** Faça $\tau = 0$ e especifique um valor inicial $k^{(0)}$ tal que $\pi(k^{(0)}) > 0$.
 - 2 **Passo:** Gere um candidato $\xi_{n_m} \sim q(\xi_{n_m} | \xi_{n_m}^{(\tau)})$. Então gere um campo de permeabilidades $k(\mathbf{x})$ dado por $\ln(k(\mathbf{x})) = Y(\mathbf{x})$.
 - 3 **Passo:** Resolva na malha grossa os problemas de pressão-velocidade e do transporte.
 - 4 **Passo:** Calcule a probabilidade de aceitação na malha grossa $\bar{\alpha}(k|k^{(\tau)})$.
 - 5 **Passo:** Gere uma variável aleatória u com distribuição uniforme $\mathcal{U}(0,1)$ e teste se o campo proposto será aceito ou não, isto é, $k^{(\tau+1)} = k$ com probabilidade $\bar{\alpha}(k|k^{(\tau)})$, e $k^{(\tau+1)} = k^{(\tau)}$ com probabilidade $1 - \bar{\alpha}(k|k^{(\tau)})$. Se k for rejeitado retorne ao **Passo 2**.
 - 6 **Passo:** Resolva na malha fina os problemas de pressão-velocidade e do transporte.
 - 7 **Passo:** Calcule a probabilidade de aceitação na malha fina $\alpha(k|k^{(\tau)})$.
 - 8 **Passo:** Gere uma variável aleatória u com distribuição uniforme $\mathcal{U}(0,1)$ e teste se o campo proposto será aceito ou não, isto é, $k^{(\tau+1)} = k$ com probabilidade $\alpha(k|k^{(\tau)})$, e $k^{(\tau+1)} = k^{(\tau)}$ com probabilidade $1 - \alpha(k|k^{(\tau)})$. Se o campo proposto k for rejeitado retorne ao **Passo 2**.
 - 9 **Passo:** Incremente $\tau = \tau + 1$, retorne ao **Passo 2** e continue o procedimento até atingir a convergência.
-

Referências

- [1] Aquino, J., Francisco, A. S., Pereira, F. & Amaral Souto, H. P., 2007. Numerical simulation of transient water infiltration in heterogeneous soils combining central schemes and mixed finite elements. *Communications in Numerical Methods in Eng.*, vol. 23, pp. 491–505.
- [2] Aquino, J., Francisco, A. S., Pereira, F. & Amaral Souto, H. P., 2011. A hybrid method for the simulation of radionuclide contaminant plumes in heterogeneous unsaturated formations *Progress in Nuclear Energy*, vol. 53, pp. 1159–1166.
- [3] Borges, M. R., Pereira, F. & Amaral Souto, H. P., 2010. Efficient generation of multi-scale random fields: A hierarchical approach. *International Journal for Numerical Methods in Biomedical Engineering*, vol. 26, pp. 176–189.
- [4] Durlafsky, L. J., 1998. Coarse Scale Models of two Phase in Heterogenous Reservoirs: Volume Averaged Equations and Their Relationship to the Existing Upscaling Techniques. *Computational Geosciences*, vol. 2, n. 2, pp. 73–92.
- [5] Douglas, C., Efendiev, Y., Ewing, R., Ginting, V. & Lazarov, R., 2006. Dynamic Data Drive Simulations in Stochastic Environments. *Computing*, vol. 77, pp. 321–333.
- [6] Efendiev, Y., Hou, T. & Lou, W., 2006. Preconditioning Markov Chain Monte Carlo Simulations Using Coarse-Scale Models. *SIAM J. Sci. Comput.*, vol. 2, n. 2, pp. 776–803.
- [7] Gamerman, D. & Lopes, H. F., 2006. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. *Texts in Statistical Science*, ed.2.
- [8] Pereira, T. J., 2012. *Estudo da quantificação de incertezas para o problema de contaminação de meios porosos heterogêneos*. Tese de Doutorado (Modelagem Computacional), Universidade do Estado do Rio de Janeiro, Nova Friburgo, Rio de Janeiro, Brasil. Disponível em: <http://www.bdtd.uerj.br/>.