Ciência e Natura, Santa Maria, v. 36 Ed. Especial, 2014, p. 82–96
Revista do Centro de Ciências Naturais e Exatas - UFSM
ISSN impressa: 0100-8307 ISSN on-line: 2179-460X

CIÊNCIA e NATURA 35 anos

Formulação e solução de um Problema Inverso de Difusão Anômala com Técnicas Estocásticas

Formulation and solution of an Inverse Anomalous Diffusion Problem with Stochastic Techniques

Luciano Gonçalves da Silva[1], Diego Campos Knupp*[2],Luiz Bevilacqua[3], Augusto César Noronha Rodrigues Galeão[4] e Antônio José da Silva Neto[5]

[1] Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, Brasil
[2] Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil
[3] Laboratório Nacional de Computação Científica, LNCC, Petrópolis, Brasil

Resumo

Recentemente uma nova formulação analítica foi desenvolvida para a modelagem de fenômenos de difusão anômala, com aplicações em diversas áreas, incluindo difusão populacional, reações químicas, transferência de massa, escoamentos multifásicos em meios porosos, etc. Nesta nova formulação, um parâmetro, que controla a fração de partículas do sistema que se difunde, é introduzido, resultando numa equação com um termo diferencial de quarta ordem. A caracterização dos parâmetros do modelo, que além do coeficiente de difusão incluem mais dois parâmetros, coeficiente de retenção e fração de partículas capazes de se difundir, requer a análise através de formulação e solução de um problema inverso juntamente com técnicas experimentais específicas. Este artigo tem por objetivo a análise de sensibilidade dos parâmetros do modelo proposto, e a formulação e solução do problema inverso através de duas abordagens: o procedimento da máxima verossimilhança, que resulta numa função objetivo a ser minimizada; e a abordagem Bayesiana, onde a informação a priori disponível sobre os parâmetros pode ser naturalmente incorporada no modelo do problema inverso, o que é particularmente interessante neste trabalho, uma vez que, como será mostrado, não é possível estimar os três parâmetros simultaneamente, requerendo a caracterização em duas etapas. Na primeira abordagem são empregados os métodos estocásticos denominados algoritmo de colisão de partículas e evolução diferencial no problema de otimização. Na segunda abordagem é utilizado o método de Monte Carlo com cadeia de Markov.

***Palavras-chave**: Difusão anômala, Algoritmo de Colisão de Partículas, Evolução Diferencial, Inferência Bayesiana, Método de Monte Carlo com Cadeia de Markov.*

Abstract

Recently a new analytical formulation has been developed for modeling the phenomena of anomalous diffusion, with relevant applications in several areas such as population spread, chemical reactions, mass transfer, multiphase flow in porous media, etc. In the new formulation a control parameter which represents the fraction of particles able to diffuse is introduced, and the resulting governing equation requires a fourth order differential term. Specific experimental techniques, together with an inverse analysis, need to be settled to characterize the complementary parameters. The present work is aimed at performing a sensitivity analysis on the model parameters, and the inverse problem formulation and solution by means of two different approaches: the maximum likelihood procedure, yielding an objective function to be minimized; and the Bayesian inference, which allows for the incorporation of any available a priori information with respect to the parameters being estimated, and is quite interesting in the present work, since, it is not possible to estimate the three parameters simultaneously, and the characterization procedure must be tackled in two stages. In the first approach we use the stochastic methods known as PCA (Particle Collision Algorithm) and DE (Differential Evolution) in the optimization problem solution. In the second approach, the posterior probability density is explored via the Markov Chain Monte Carlo method.

***Keywords**: Anomalous diffusion, Particle Collision Algorithm, Differential Evolution, Bayesian Inference, Markov Chain Monte Carlo method.*

*diegoknupp@iprj.uerj.br
Recebido: 17/03/2014 Aceito: 24/03/2014

Nomenclatura

K_2	Coeficiente de difusão;
K_4	Coeficiente de retenção;
L	Comprimento espacial;
N_d	Número de dados experimentais;
$p(x,t)$	Concentração de partículas;
$\mathbf{p}^{\text{calc}}$	Valores calculados através do problema direto;
$\mathbf{p}^{\text{exp}}$	Valores para a variável observável medidos experimentalmente;
$Q(\mathbf{Z})$	Função objetivo, Eq. (7);
t	Tempo;
$\mathbf{W}$	Matriz de covariâncias;
$\bar{X}_{Z_j}$	Coeficiente de sensibilidade normalizado;
x	Posição;
$\mathbf{Z}$	Vetor de incógnitas para o problema inverso;

Letras gregas

$\pi(\mathbf{p}^{\text{exp}})$	Densidade marginal;
β	Fração de partículas capazes de se difundir;
α	Fração de partículas retidas;
$\pi(\mathbf{p}^{\text{exp}} \mid \mathbf{Z})$	Função da máxima verossimilhança;
$\pi_{post}(\mathbf{Z})$	Função densidade de probabilidades a posteriori;
$\pi_{pr}(\mathbf{Z})$	Função densidade de probabilidades a priori;
Δt	Passo de tempo;
Δx	Passo na discretização espacial da malha computacional;
μ	Média;
σ	Desvio padrão;
ξ	Fator de aceitação no método de Monte Carlo com cadeia de Markov;

1 Introdução

O processo de difusão de partículas é, em geral, modelado matematicamente pela clássica lei de Fick. No entanto, nesta formulação, alguns fenômenos que podem ocorrer juntamente com a difusão são completamente ignorados. Por exemplo, fenômenos de atraso de tempo e efeitos de retenção, que ocorrem devido a adsorção, aprisionamento mecânico, reações químicas e interações mecânicas com o meio de suporte (Brandani *et al*., 2000; Liu e Thompson, 2002; Atsumi, 2002; D'Angelo *et al*., 2003; Deleersnijder *et al*., 2006; Green, 1996; Muhammad, 2004; Derec et al., 2010). A maior parte dos casos observados na literatura lida com esta questão através da equação da difusão, modelando os efeitos de retenção através de derivadas fracionárias (Metzler e Klafter, 2000), ou modelando-se o coeficiente de difusão variando arbitrariamente com o tempo e/ou com a concentração (Joannès *et al*., 2014; Wu e Berland, 2008). Bevilacqua et al. (2011a, 2011b), a partir de uma abordagem discreta em dinâmica de população, propuseram um novo modelo para difusão anômala. Nesta nova formulação, um parâmetro de controle, que indica a fração de partículas que são capazes de se difundir, é levado em consideração. Através deste parâmetro, e de um termo diferencial de quarta ordem, que representa o efeito anômalo, é possível descrever não só o fenômeno de difusão pura, como também outros efeitos decorrentes de fenômenos que ocorrem simultaneamente com a difusão. Como por exemplo, a retenção parcial de partículas ou aceleração do processo difusivo.

Modelos de difusão anômala com derivadas fracionárias no tempo têm sido investigados sob o ponto de vista de problemas inversos em várias publicações recentes, demonstrando o interesse da comunidade científica nesta classe de problemas (Cheng e Fu, 2012; Wei e Zhang, 2013; Xiong *et al*., 2012; Zhang e Wei, 2013). Entretanto, o novo modelo de difusão anômala proposto por Bevilacqua et al. (2011a,b), devido à sua recente proposição e desenvolvimento, ainda não teve o seu correspondente problema inverso estudado na literatura específica. Os novos parâmetros que aparecem neste modelo, em adição ao coeficiente de difusão, que caracterizam o processo de difusão anômala, demandam técnicas experimentais específicas, que, juntamente com a análise por meio da formulação de um problema inverso, precisam ser utilizados para caracterização destes parâmetros.

O presente trabalho visa, portanto, a investigação do problema de difusão anômala, modelado de acordo com Bevilacqua *et al*. (2011a,b), sob a ótica de problemas inversos. Primeiramente o problema inverso é formulado por meio do procedimento da máxima verossimilhança, resultando num problema de otimização (Moura Neto e Silva Neto, 2012; Schwaab e Pinto, 2007), onde a função objetivo é minimizada com os métodos estocásticos PCA (Particle Collision Algorithm) (Sacco et al., 2006) e DE (Differential Evolution) (Storn e Price, 1997; Price *et al*.,

2006). Como ficará claro na análise de sensibilidade dos parâmetros, e na análise dos resultados para o caso teste selecionado, não é possível estimar os três parâmetros do modelo simultaneamente, e um procedimento de caracterização em duas etapas se faz necessário, onde na primeira etapa se força uma situação de difusão pura, em que o coeficiente de difusão possa ser estimado independentemente. Na segunda etapa, os parâmetros que caracterizam a difusão anômala são então estimados. Apenas para citar um exemplo prático envolvendo esta possibilidade, Derec *et al.* (2010) estudaram a difusão de ferrofluido em água, num escoamento em um microcanal, e observaram que na ausência de campo magnético o processo difusivo ocorre de acordo com a teoria clássica, enquanto na presença de campo magnético se observa o efeito da difusão anômala. Neste cenário, o presente trabalho também apresenta a formulação do problema inverso por meio de inferência Bayesiana (Kaipio e Somersalo, 2004; Orlande *et al.*, 2011), que permite a modelagem de maneira natural da informação a priori, obtida na primeira etapa, na formulação do problema inverso da segunda etapa. A solução do problema inverso na abordagem Bayesiana é realizada por meio do método de Monte Carlo com Cadeia de Markov (MCMC – Monte Carlo Markov Chain). Os resultados são criticamente analisados, e os procedimentos investigados neste trabalho contribuem para a construção de uma base teórica necessária na caracterização dos parâmetros deste novo modelo.

2 Formulação e solução do problema direto

Considere o processo difusivo esquematicamente representado na figura 1. A redistribuição dos conteúdos de cada célula indica a fração αp_n que é retida na n-ésima célula, e o conteúdo restante que é igualmente dividido para as células vizinhas, $0{,}5\beta p_n$ para a célula da esquerda, (n-1)-ésima célula, e $0{,}5\beta p_n$ para a célula da direita, (n+1)-ésima célula, em cada passo temporal, onde β é a fração de partículas capazes de se difundir, ou seja, $\beta = (1-\alpha)$

Traduzindo algebricamente esta regra, e depois de algumas manipulações, é possível estabelecer em termos de diferenças finitas (Bevilacqua *et al.*, 2011a):

$$\frac{\Delta p_n^{t+\Delta t}}{\Delta t} = \beta\left\{\frac{1}{2}\frac{L_0^2}{T_0}\frac{\Delta^2 p_n}{\Delta x^2} + \frac{O\left(\Delta x^2\right)}{\Delta x^2} - (1-\beta)\frac{1}{4}\frac{L_1^4}{T_0}\frac{\Delta^4 p_n}{\Delta x^4}\right\}^{t-\Delta t} \quad (1)$$

onde L_0, L_1 e T_0 o fatores de escala, e Δt e Δx são, respectivamente, os tamanhos do passo de tempo e do incremento da malha.

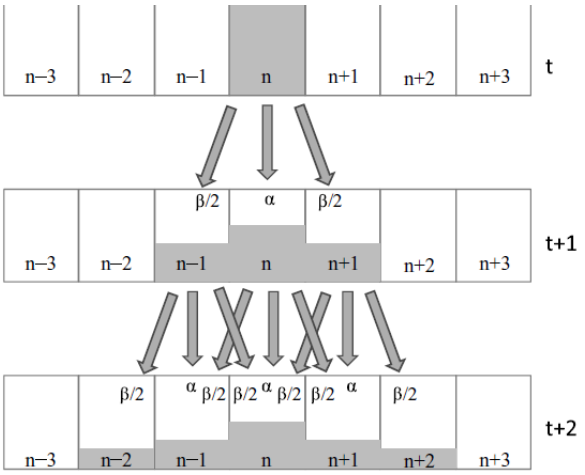


Figura 1: Representação esquemática da distribuição simétrica $\beta/2$ com retenção α.

Definindo $K_2 = L_0^2/2T_0$ e $K_4 = L_1^4/4T_0$, e tomando o limite $\Delta x \to 0$ e $\Delta t \to 0$, obtém-se:

$$\frac{\partial p}{\partial t} = \beta K_2 \frac{\partial^2 p}{\partial x^2} - \beta(1-\beta) K_4 \frac{\partial^4 p}{\partial x^4} \quad (2)$$

que é a equação que governa o fenômeno de difusão anômala para um meio unidimensional. Observa-se que para $\beta = 0$, tem-se uma solução estacionária, e para $\beta = 1$ tem-se a clássica formulação para a difusão pura de partículas (Bevilacqua *et al.*, 2011a).

Neste trabalho serão utilizados dois conjuntos de condição inicial e condições de contorno diferentes. Considerando o primeiro conjunto:

$$p(0,t) = 0, \quad p(L,t) = 0, \quad t > 0 \quad (3a,b)$$

$$\left.\frac{\partial^2 p(x,t)}{\partial x^2}\right|_{x=0} = \left.\frac{\partial^2 p(x,t)}{\partial x^2}\right|_{x=L} = 0, \; t > 0 \quad (3c,d)$$

$$p(x,0) = \operatorname{sen}\left(\frac{\pi x}{L}\right), 0 \le x \le L, \; t = 0 \quad (3e)$$

tem-se uma solução analítica exata disponível, na seguinte forma (Silva, 2013):

$$p(x,t) = \exp\left[-\beta K_2 \pi^2 \, \mathrm{t}\right] \times \exp\left[-\beta K_4 (1-\beta)\pi^4 \, \mathrm{t}\right] \sin\left(\frac{\pi x}{L}\right) \quad (4)$$

que será utilizada principalmente na validação do esquema de solução numérica por meio de diferenças finitas utilizada neste trabalho

O segundo conjunto de condição inicial e condições de contorno, apresentado abaixo, não apresenta solução analítica exata, mas consiste no caso de condições de contorno tradicionais do processo difusivo, com condições do primeiro e segundo tipo. Este caso será utilizado na investigação do problema inverso.

$$p(0,t)=1, \quad p(L,t)=1, \quad t>0 \tag{5a,b}$$

$$\left.\frac{\partial p(x,t)}{\partial x}\right|_{x=0} = \left.\frac{\partial p(x,t)}{\partial x}\right|_{x=L} = 0, \quad t>0 \tag{5c,d}$$

$$p(x,0)=\operatorname{sen}^{100}\left(\frac{\pi x}{L}\right)+1, \quad 0\le x\le L \tag{5e}$$

Para a solução numérica do modelo matemático, foi feito o uso de aproximações com um erro de truncamento de ordem $O\left(\Delta x^2\right)$ para as derivadas de segunda e quarta ordem em relação a x e um erro de truncamento de ordem $O\left(\Delta t\right)$ para a derivada de primeira ordem em relação a t. Usando expansões de Taylor, tem-se (Tannehill *et al.*, 1997):

$$\left.\frac{\partial^2 p}{\partial x^2}\right|_{i,n+1} = \frac{p_{i-1}^{n+1}-2p_i^{n+1}+p_{i+1}^{n+1}}{\Delta x^2} + O\left(\Delta x^2\right) \tag{6a}$$

$$\left.\frac{\partial^4 p}{\partial x^4}\right|_{i,n+1} = \frac{p_{i-2}^{n+1}-4p_{i-1}^{n+1}+p_i^{n+1}-4p_{i+1}^{n+1}+p_{i+2}^{n+1}}{\Delta x^4} + O\left(\Delta x^2\right) \tag{6b}$$

$$\left.\frac{\partial p}{\partial t}\right|_{n,i} = \frac{p_i^{n+1}-p_i^n}{\Delta t} + O\left(\Delta t\right) \tag{6c}$$

com $i=0, 1, ..., N\text{-}1, N$ e $n=0, 1, 2, ..., T$, onde $t_{\max} = T\Delta t$ representa o tempo total de observação. A partir das aproximações por diferenças dadas nas Eqs. (6a-c) é possível obter uma solução aproximada para o modelo de difusão anômala.

3 Formulação e solução do problema inverso

Neste trabalho são apresentadas duas abordagens na formulação e solução do problema inverso. A primeira consiste no procedimento da máxima verossimilhança, resultando num problema de otimização, onde a função objetivo, que caracteriza os desvios entre as previsões do modelo e as observações experimentais, deve ser minimizada. Na segunda abordagem é adotado o procedimento de inferência Bayesiana, que permite, de forma natural, a inclusão de informação a priori sobre os parâmetros.

Considere que no problema formulado matematicamente pela Eq. (2) estejam disponíveis alguns valores da variável observável, p, medidos experimentalmente, em uma ou mais posições, em tempos distintos, $\mathbf{p}^{\exp} = \{p_1^{\exp}, p_2^{\exp}, ..., p_i^{\exp}, ..., p_{N_d}^{\exp}\}$, onde N_d é o número de dados experimentais, e que os valores do conjunto de parâmetros $\mathbf{Z} = \{\beta, K_2, K_4\}$ sejam desconhecidos, e deseja-se estimá-los. Este é o problema inverso de interesse neste trabalho.

3.1 Procedimento da máxima verossimilhança

Dado que os erros experimentais são aditivos e podem ser modelados pela distribuição Normal, com média zero e desvio padrão constante, e que os erros não são correlacionados, a função objetivo a ser minimizada para produzir estimativas de máxima verossimilhança é dada pelo somatório dos resíduos quadrados entre as previsões do modelo e as observações experimentais (Beck et al., 1985; Schwaab e Pinto, 2007), ou seja:

$$Q(\mathbf{Z}) = \left(\mathbf{p}^{\exp} - \mathbf{p}^{\text{calc}}\right)^T \left(\mathbf{p}^{\exp} - \mathbf{p}^{\text{calc}}\right) = \sum_{i=1}^{N_d}\left(p_i^{\exp} - p_i^{\text{calc}}\right)^2 \tag{7}$$

onde $\mathbf{p}^{\text{calc}}$ é o vetor que contém os valores calculados para a variável $p(x,t)$ através da solução da Eq. (2), nas mesmas posições e instantes das observações experimentais, $\mathbf{p}^{\exp}$. Portanto buscam-se os valores para os parâmetros do conjunto $\mathbf{Z}$ de tal forma que a função objetivo dada pela Eq. (7) seja mínima.

Dentre as alternativas para minimização da função objetivo, destacam-se os métodos determinísticos, estocásticos e as hibridizações entre eles (Silva Neto e Becceneri, 2012). Neste trabalho, devido à robustez e relativa independência na escolha da estimativa inicial, serão usados métodos estocásticos, dentre os quais foram utilizados o algoritmo de colisões de partículas, PCA (*Particle Collision Algorithm*) (Sacco *et al.*, 2006) e o método de evolução diferencial, DE (*Differential Evolution*) (Storn e Price, 1997; Price *et al.*, 2006), escolhidos basicamente pela comprovada robustez e experiência dos autores em trabalhos anteriores (Silva Neto e Becceneri, 2012). Estes métodos são brevemente descritos a seguir.

3.1.1 Evolução diferencial (DE)

No método DE é gerada uma população inicial de forma aleatória, com N_{pop} indivíduos, fazendo:

$$Z_{i,j} = Z_{j,L} + r_{ij}\left(Z_{j,U} - Z_{j,L}\right), \quad i = 1,2....,\mathrm{N}_{pop}, \quad j = 1,2,...,N_p \tag{8a}$$

onde $Z_{j,L}$ e $Z_{j,U}$ são respectivamente os limites inferiores e superiores do j-ésimo elemento do vetor de parâmetros $\mathbf{Z}$, r_{ij} é um número aleatório de uma distribuição uniforme no intervalo [0, 1], e N_p é o número de parâmetros sendo estimados, ou seja, a dimensão do vetor $\mathbf{Z}$.

Após a geração da população inicial, o método DE realiza as operações de mutação e recombinação para a geração de uma nova população com N_{pop} indivíduos. O operador de mutação diferencial adiciona um vetor de referência, $\mathbf{Z}_{l_1}$, escolhido aleatoriamente na população, a um vetor diferença obtido a partir de outros dois vetores também escolhidos aleatoriamente na população, $\mathbf{Z}_{l_2}$ e $\mathbf{Z}_{l_3}$.

Um novo candidato à solução é então obtido fazendo

$$\mathbf{Z}_{l,novo} = \mathbf{Z}_{l_1} + t_p\left(\mathbf{Z}_{l_2} - \mathbf{Z}_{l_3}\right), l = 1,2,...,N_{pop} \qquad (8b)$$

onde t_p representa a taxa de perturbação, um número real que controla a magnitude do vetor diferença obtido em cada operação aritmética realizada.

Para complementar a operação de mutação, o método DE emprega o operador de cruzamento. Nesta operação o vetor $\mathbf{Z}_{l.novo}$ pode ou não ser aceito na população atual de acordo com

$$\mathbf{Z}_l = \begin{cases} \mathbf{Z}_{l,novo}, & \text{se } r_1 < p_c \text{ ou } j = j_{aleat} \\ \mathbf{Z}_{l,1}, & \text{caso contrário} \end{cases} \qquad (8c)$$

onde p_c é a probabilidade de cruzamento definida a priori. Se o número aleatório $r_l \sim U(0,1)$ é menor que a probabilidade de cruzamento, o novo candidato $\mathbf{Z}_{l,novo}$ é aceito. Caso contrário $\mathbf{Z}_{l,1}$ é mantido na população atual. Na sequência é então empregado o operador de seleção. Se o vetor $\mathbf{Z}_{l,novo}$ apresenta um valor menor do que o vetor $\mathbf{Z}_{l,1}$ para a função objetivo, Eq. (7), ele o substitui na geração seguinte:

$$\mathbf{Z}_{l_1}^{k+1} = \begin{cases} \mathbf{Z}_{l,novo}, & \text{se } Q(\mathbf{Z}_{l,novo}) \le Q(\mathbf{Z}_{l_1}) \\ \mathbf{Z}_{l_1}, & \text{caso contrário} \end{cases} \qquad (8d)$$

Uma vez completado o processo de atualização da população a ser considerada na geração seguinte, todo o processo aqui descrito é repetido até que um critério de parada seja satisfeito. O critério de parada mais usual consiste no estabelecimento prévio do número de gerações n_g que será utilizado.

3.1.2 Algoritmo de colisão de partículas (PCA)

A seguir, os passos para a implementação do PCA canônico para um problema de minimização da função objetivo são apresentados.

Passo 1. *Atribua 1 como o valor para a variável* ***contador****.*

Passo 2. *Gere uma solução inicial aleatória dentro do domínio de busca da solução. Chame esta solução de* $\mathbf{Z}_{ant}$.

$$Z_{ant} = Z_{j,L} + r_j\left(Z_{j,U} - Z_{j,L}\right), j = 1,2,...,N_p \qquad (9a)$$

onde $Z_{j,L}$ e $Z_{j,U}$ são respectivamente os limites inferiores e superiores do j-ésimo elemento do vetor de parâmetros $\mathbf{Z}$, r_j é um número aleatório de uma distribuição uniforme no intervalo [0, 1], e N_p é o número de parâmetros sendo estimados, ou seja, a dimensão do vetor $\mathbf{Z}$.

Passo 3. *Caso* $contador \ge n_{PCA}$ *encerre o PCA. Caso contrário, incremente o valor de* ***contador*** *em uma unidade e gere uma nova solução,* $\mathbf{Z}_{novo}$*, que é calculada a partir de uma perturbação estocástica na solução* $\mathbf{Z}_{ant}$ *da seguinte forma:*

$$\mathbf{Z}_{novo} = \mathbf{Z}_{ant} + [(\mathbf{Z}_U - \mathbf{Z}_{ant}) \cdot r - (\mathbf{Z}_{ant} - \mathbf{Z}_L) \cdot (1-r)] \qquad (9b)$$

Se o valor da função-objetivo para $\mathbf{Z}_{novo}$ for menor que o valor da função-objetivo para $\mathbf{Z}_{ant}$ então $\mathbf{Z}_{ant}$ recebe o valor de $\mathbf{Z}_{novo}$ e vá para o ***Passo 4***. Caso contrário, vá para o ***Passo 5***.

Obs: Mantenha uma variável $\mathbf{Z}_{best}$ que contém o valor da melhor solução até o momento. Sempre que uma nova solução for adotada para $\mathbf{Z}_{ant}$ deve-se verificar se esta solução é melhor que a $\mathbf{Z}_{best}$ e então, se for o caso, fazer $\mathbf{Z}_{best}$ receber o valor de $\mathbf{Z}_{ant}$.

Passo 4. *Faça o seguinte procedimento* $n_{Exploração}$ *vezes:*

$$\mathbf{Z}_{novo} = \mathbf{Z}_{ant} + [(\mathbf{Z}^*_U - \mathbf{Z}_{ant}) \cdot r - (\mathbf{Z}_{ant} - \mathbf{Z}^*_L) \cdot (1-r)] \qquad (9c)$$

onde:

$$\mathbf{Z}^*_U = (1+0,2r) \cdot \mathbf{Z}_{ant}, \quad \mathbf{Z}^*_L = (1-0,2r) \cdot \mathbf{Z}_{ant} \qquad (9d,e)$$

Caso o valor da função-objetivo para $\mathbf{Z}_{novo}$ seja menor que o valor da função-objetivo para $\mathbf{Z}_{ant}$ então $\mathbf{Z}_{ant}$ recebe o valor de $\mathbf{Z}_{novo}$. Sempre que uma nova solução for adotada como $\mathbf{Z}_{ant}$ deve-se verificar se esta solução é melhor que a $\mathbf{Z}_{best}$ e então, se for o caso, fazer $\mathbf{Z}_{best}$ receber o valor de $\mathbf{Z}_{ant}$.

Depois de repetido este procedimento $n_{Exploration}$ vezes, vá para o ***Passo 3***.

Passo 5. *Calcule a probabilidade de espalhamento,* $p_{scattering}$.

$$p_{scattering} = 1 - \frac{Q(\mathbf{Z}_{best})}{Q(\mathbf{Z}_{novo})} \qquad (9f)$$

Se $p_{scattering}$ for maior que $r \sim U(0,1)$, então $\mathbf{Z}_{ant}$ recebe o valor de $\mathbf{Z}_{novo}$ e vá para o ***Passo 4***. Caso contrário, vá para o ***Passo 2***.

Ao término do PCA, a estimativa do mínimo da função-objetivo é dada pela solução $\mathbf{Z}_{best}$.

3.2 Inferência Bayesiana

Na abordagem Bayesiana o problema inverso é formulado como um problema de inferência estatística e está baseado nos seguintes princípios (Kaipio e Somersalo, 2004): (i) todos os parâmetros do modelo são modelados como variáveis aleatórias; (ii) a aleatoriedade descreve o nível de informação a priori; (iii) o nível de informação a priori deve ser modelado dentro de uma distribuição de probabilidades; (iv) a solução do problema inverso é a distribuição a posteriori. Desta forma, toda informação a priori sobre os parâmetros do modelo podem e devem ser incorporadas na formulação do problema inverso. O teorema de Bayes em problemas inversos pode ser expresso da seguinte forma:

$$\pi_{post}(\mathbf{Z}) = \pi(\mathbf{Z} \mid \mathbf{p}^{\exp}) = \frac{\pi_{pr}(\mathbf{Z})\pi(\mathbf{p}^{\exp} \mid \mathbf{Z})}{\pi(\mathbf{p}^{\exp})} \tag{10}$$

onde $\pi_{post}(\mathbf{Z})$ é a função densidade de probabilidades a posteriori, $\pi_{pr}(\mathbf{Z})$ é a função densidade de probabilidades a priori; $\pi(\mathbf{p}^{\exp} \mid \mathbf{Z})$ é a função da máxima verossimilhança e $\pi(\mathbf{p}^{\exp})$ é a densidade marginal, que exerce o papel de constante de normalização, e, em geral, não é conhecida. Dado que os erros experimentais são aditivos e podem ser representados pela distribuição Normal, temos:

$$\begin{aligned}\pi(\mathbf{p}^{\exp} \mid \mathbf{Z}) = &\frac{1}{\sqrt{(2\pi)^{N_d}}}\frac{1}{\det(\mathbf{W})} \\ &\exp\left[-\frac{1}{2}\left(\mathbf{p}^{\exp}-\mathbf{p}^{\text{calc}}\right)^T \mathbf{W}^{-1}\left(\mathbf{p}^{\exp}-\mathbf{p}^{\text{calc}}\right)\right]\end{aligned} \tag{11}$$

onde $\mathbf{W}$ é a matriz de covariâncias relacionada aos dados experimentais $\mathbf{p}^{\exp}$.

Neste trabalho, para se obter uma aproximação da distribuição a posteriori, será utilizado o método de Monte Carlo com Cadeia de Markov (MCMC), simulando-se amostras de $\pi_{post}(\mathbf{Z})$. De modo a implementar a cadeia de Markov é necessária uma densidade de transição, $q(\mathbf{Z}^t, \mathbf{Z}^*)$, que dá a probabilidade de se escolher um candidato $\mathbf{Z}^*$ a partir do estado atual da cadeia, $\mathbf{Z}^t$, e neste trabalho o algoritmo de Metropolis-Hastings (Kaipio e Somersalo, 2004; Orlande *et al.*, 2011), como descrito a seguir, é utilizado na implementação.

Passo 1. Amostrar um candidato $\mathbf{Z}^*$ a partir do estado atual da cadeia e da densidade de transição, $q(\mathbf{Z}^t, \mathbf{Z}^*)$;

Passo 2. Calcular:

$$\xi = \min\left[1, \frac{\pi(\mathbf{Z}^* \mid \mathbf{p}^{\exp})q(\mathbf{Z}^*, \mathbf{Z}^t)}{\pi(\mathbf{Z}^t \mid \mathbf{p}^{\exp})q(\mathbf{Z}^t, \mathbf{Z}^*)}\right] \tag{12a}$$

Passo 3. Se $r < \xi$, então:

$$\mathbf{Z}^{t+1} = \mathbf{Z}^* \tag{12b}$$

caso contrário:

$$\mathbf{Z}^{t+1} = \mathbf{Z}^t \tag{12c}$$

onde $r \sim U(0,1)$ denota uma amostra da distribuição uniforme entre 0 e 1;

Passo 4. Retornar ao **Passo 1** de modo a gerar a cadeia de Markov $\{\mathbf{Z}^1, \mathbf{Z}^2, ..., \mathbf{Z}^{N_{MCMC}}\}$. Ressalta-se que os estados iniciais da cadeia de Markov, conhecidos como amostras de aquecimento, até a cadeia entrar em equilíbrio (onde a auto-correlação da cadeia deve ser próxima de zero), devem ser descartados.

No presente trabalho é utilizada uma distribuição Normal como densidade de transição, que é simétrica, e nesse caso a Eq. (12a) se reduz a:

$$\xi = \min\left[1, \frac{\pi(\mathbf{Z}^* \mid \mathbf{p}^{\exp})}{\pi(\mathbf{Z}^t \mid \mathbf{p}^{\exp})}\right] \tag{12d}$$

Ressalta-se que tanto na avaliação da Eq. (12a) quanto na avaliação da Eq. (12d) a densidade marginal $\pi(\mathbf{p}^{\exp})$ não é necessária.

3.3 Análise de sensibilidade

A análise de sensibilidade contribui de forma significativa na formulação e solução de problemas inversos, permitindo, por exemplo, a escolha de dados experimentais que resultem em estimativas de melhor qualidade (Beck *et al.*, 1985).

Os coeficientes de sensibilidade normalizados são definidos como:

$$\overline{X}_{Z_j}(x,t) = Z_j \frac{\partial p(x,\mathrm{t})}{\partial Z_j}, \quad j = 1,2,3 \tag{13}$$

onde Z_j representa cada incógnita do problema inverso, sendo o número de parâmetros do modelo $N_p = 3$. É desejável que os coeficientes de sensibilidade possuam valores relativos elevados, significando que pequenas alterações nos parâmetros provocam significativas alterações na resposta do modelo. Quando dois ou mais parâmetros são estimados simultaneamente, os coeficientes

de sensibilidade em relação aos parâmetros não devem apresentar dependência linear entre eles, caso contrário não há unicidade na solução do problema inverso.

4 Resultados

4.1 Validação da solução numérica

Empregando o método de diferenças finitas, usando uma malha computacional para o domínio espacial com trezentos nós e um intervalo de tempo $\Delta t = 0,01$, a partir das Eqs. (6a-c), obteve-se uma solução numérica para a Eq. (2) com as condições inicial e de contorno dadas pelas Eqs. (3a-e), onde os valores apresentados na tabela 1 foram empregados para os parâmetros do modelo de difusão anômala. O caso teste selecionado, com valores apresentados na tabela 1, são um dos casos cuja solução do problema direto foi investigada por Silva (2013), em sua dissertação de mestrado. A figura 2 apresenta a comparação entre esta solução numérica e a solução analítica, dada pela Eq. (4), em três instantes de tempo distintos, t =10, 50 e 100, onde se observa que as curvas são essencialmente coincidentes. De fato, o erro relativo percentual máximo observado entre a solução numérica e a solução analítica foi de 0,0019%.

Tabela 1: Valores dos parâmetros utilizados no primeiro caso (solução analítica exata).

Parâmetro	Valor
β	0,2
K_2	1x10^{-3}
K_4	1x10^{-5}

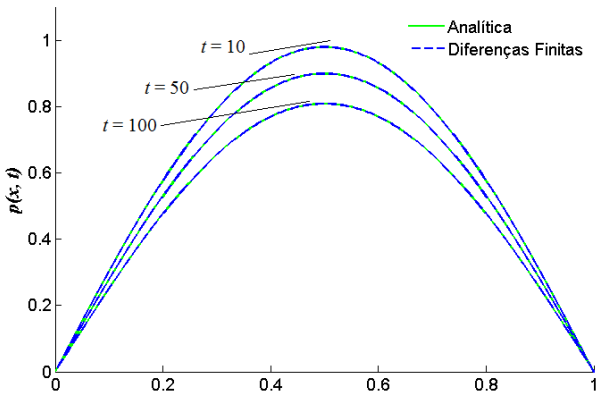


Figura 2: Comparação entre a solução analítica e a solução numérica.

4.2 Análise de sensibilidade

A inspeção direta da Eq. (2) mostra que nesta equação tem-se três parâmetros para definir apenas dois coeficientes dos termos diferenciais de segunda e quarta ordem, o que significa que infinitas combinações de valores para β, K_2, e K_4 levam exatamente à mesma formulação matemática, evidentemente inviabilizando a estimativa simultânea dos três parâmetros.

Inicialmente, investiga-se o caso dado pela Eq. (2), com condições inicial e de contorno dadas pelas Eqs. (3a-e), que possui solução analítica exata, o que permite o cálculo analítico dos coeficientes de sensibilidade. Definindo-se os coeficientes da Eq. (2) como:

$$t_1 = \beta K_2 \tag{14a}$$

$$t_2 = \beta(1-\beta)K_4 \tag{14b}$$

e diferenciando-se a solução dada pela Eq. (4) em relação a eles, resulta:

$$\frac{\partial p}{\partial t_1} = \pi^2 t \sin\left(\frac{\pi x_m}{L}\right)\left(-e^{-\pi^2 t \cdot t_1 - \pi^4 t \cdot t_2}\right) \tag{15a}$$

$$\frac{\partial p}{\partial t_2} = \pi^4 t \sin\left(\frac{\pi x_m}{L}\right)\left(-e^{-\pi^2 t \cdot t_1 - \pi^4 t \cdot t_2}\right) \tag{15b}$$

Ou seja, a razão $\dfrac{\partial p / \partial t_1}{\partial p / \partial t_2}$ é uma constante, $1/\pi^2$, o que evidencia uma dependência linear perfeita entre estes coeficientes de sensibilidade, ou seja, t_1 e t_2 não podem ser estimados simultaneamente neste caso, o que em relação aos parâmetros primitivos significa que apenas um dos três parâmetros pode ser estimado num procedimento de caracterização (os outros dois parâmetros devem ser conhecidos a priori).

Por outro lado, o caso dado pela Eq. (2) com condições inicial e de contorno dadas pelas Eqs. (5a-e), condições tradicionais de problemas difusivos, de primeiro e segundo tipo, e empregando-se os valores dos parâmetros apresentados na tabela 1, apresenta comportamento diferente, como mostra a figura 3, onde os coeficientes de sensibilidade normalizados, calculados numericamente com um esquema de diferenças centradas, em relação a cada um dos três parâmetros primitivos são apresentados. Nesta figura, a curva correspondente a $\bar{X}_{K_4} - \bar{X}_\beta$ também é apresentada, ficando clara a dependência linear entre $\bar{X}_{K_2}$ e uma combinação linear dos coeficientes de sensibilidade relacionados aos outros dois parâmetros. Entretanto, da análise de sensibilidade, verifica-se que a estimativa de dois dos três parâmetros pode ser possível.

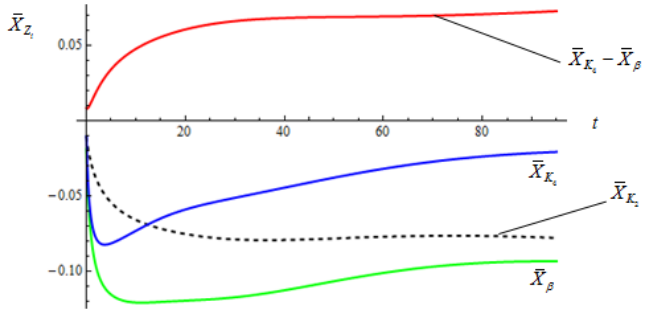


Figura 3: Coeficientes de sensibilidade normalizados para medidas transientes localizadas em $x = 0,5$.

Na figura 3 também é importante observar que para $t > 95$ os coeficientes de sensibilidade tornam-se essencialmente constantes, ou seja, a partir deste ponto as observações experimentais agregam muito pouca informação para a solução do problema inverso.

Dando continuidade à análise de sensibilidade para este caso, a figura 4 apresenta os coeficientes de sensibilidade normalizados em relação aos três parâmetros para um instante fixo, $t = 10$, ao longo da variável x. Este gráfico mostra que a posição $x = 0,5$ é onde os coeficientes de sensibilidade apresentam valores máximos e, portanto, deve ser preferida como localização do sensor para a observação experimental transiente.

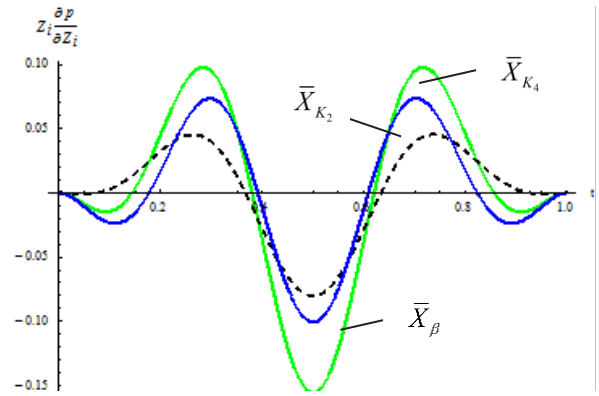


Figura 4: Coeficientes de sensibilidade normalizados ao longo da variável x para $t = 10$.

Na próxima Seção são apresentados e analisados resultados da solução do problema inverso para este caso, considerando a estimativa simultânea dos três parâmetros (β, K_2 e K_4), e apenas dois deles (β e K_4, com a hipótese de que o coeficiente de difusão, K_2, é conhecido a priori, por exemplo, obtido de outro procedimento experimental onde a difusão pura é forçada).

4.3 Problema inverso

Como evidenciado na análise de sensibilidade da Seção anterior, são aqui consideradas observações experimentais transientes em $x = 0,5$ até $t = 95$. Os dados experimentais empregados na simulação do problema inverso foram simulados adicionando-se ruído de uma distribuição Normal com média zero e desvio padrão constante, como apresentado a seguir:

$$p_i^{\exp} = p_i^{\text{calc}}(\mathbf{Z}_{\text{exato}}) + \varepsilon_i \text{, onde } \varepsilon \sim \quad (0, \sigma) \tag{16}$$

A figura 5 apresenta, juntamente com a curva de solução do modelo com os valores exatos dos parâmetros, os dados experimentais simulados empregando-se $\sigma_e = 0,02$ na Eq. (16), o que resulta em erros de até cerca de 4%.

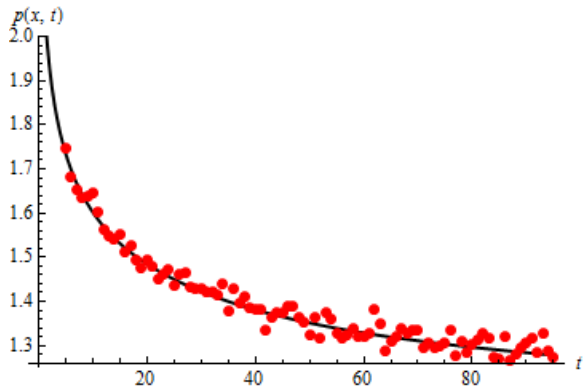


Figura 5: Dados experimentais simulados de observações transientes de um único sensor localizado em $x = 0,5$.

As tabelas 2(a) e 2(b) apresentam o resultado da minimização da função objetivo, Eq. (7), com os métodos estocásticos DE e PCA, respectivamente. Nestas tabelas são apresentados o resultado de dez execuções com cada método, onde em cada execução foi utilizado um conjunto de dados experimentais diferentes simulados com a Eq. (16), ou seja, simulando-se a realização de experimentos independentes para cada execução. Os dez conjuntos de dados experimentais empregados para a solução do problema inverso com o método DE foram também utilizados para o método PCA. A observação das tabelas 2(a) e 2(b) deixa claro que, de fato, não é possível a estimativa dos três parâmetros primitivos do modelo, como já havia sido evidenciado na análise de sensibilidade. Os valores estimados para β, K_2 e K_4 são muito distantes dos valores exatos, e apresentam uma grande dispersão. Entretanto, confirma-se que os parâmetros t_1 e t_2, definidos nas Eqs. (14a,b), que caracterizam o modelo, podem ser estimados simultaneamente. Os valores calculados para t_1 e t_2 empregando as estimativas para os parâmetros primitivos são bastante próximos dos valores exatos e apresentam uma dispersão muito

Tabela 2(a): Resultado de dez execuções com o método DE para minimização da função objetivo. Número de avaliações da função objetivo em cada execução: 5.000.

$Q(\mathbf{Z})$, Eq. (7)	β	K_2	K_4	t_1	t_2
$8{,}584347\times10^{-2}$	$4{,}453861\times10^{-4}$	$4{,}124811\times10^{-1}$	$3{,}261616\times10^{-3}$	$1{,}837134\times10^{-4}$	$1{,}452031\times10^{-6}$
$9{,}324333\times10^{-2}$	$3{,}034270\times10^{-4}$	$7{,}112865\times10^{-1}$	$4{,}525797\times10^{-3}$	$2{,}158235\times10^{-4}$	$1{,}372832\times10^{-6}$
$8{,}844045\times10^{-2}$	$1{,}120005\times10^{-3}$	$1{,}716378\times10^{-1}$	$1{,}385854\times10^{-3}$	$1{,}922352\times10^{-4}$	$1{,}550425\times10^{-6}$
$9{,}614516\times10^{-2}$	$3{,}703825\times10^{-4}$	$5{,}257681\times10^{-1}$	$4{,}440217\times10^{-3}$	$1{,}947353\times10^{-4}$	$1{,}643970\times10^{-6}$
$1{,}008961\times10^{-1}$	$1{,}363260\times10^{-3}$	$1{,}494757\times10^{-1}$	$1{,}177680\times10^{-3}$	$2{,}037742\times10^{-4}$	$1{,}603296\times10^{-6}$
$8{,}939197\times10^{-2}$	$1{,}390461\times10^{-3}$	$1{,}418104\times10^{-1}$	$1{,}080125\times10^{-3}$	$1{,}971819\times10^{-4}$	$1{,}499784\times10^{-6}$
$8{,}790108\times10^{-2}$	$3{,}387374\times10^{-4}$	$5{,}469459\times10^{-1}$	$5{,}322599\times10^{-3}$	$1{,}852711\times10^{-4}$	$1{,}802353\times10^{-6}$
$1{,}002546\times10^{-1}$	$3{,}108886\times10^{-4}$	$6{,}135466\times10^{-1}$	$5{,}124526\times10^{-3}$	$1{,}907446\times10^{-4}$	$1{,}592661\times10^{-6}$
$9{,}048612\times10^{-2}$	$3{,}543361\times10^{-4}$	$6{,}146952\times10^{-1}$	$4{,}421153\times10^{-3}$	$2{,}178087\times10^{-4}$	$1{,}566019\times10^{-6}$
9.632193×10^{-2}	$4{,}506204\times10^{-4}$	$4{,}553090\times10^{-1}$	$3{,}952143\times10^{-3}$	$2{,}051715\times10^{-4}$	$1{,}780114\times10^{-6}$
μ	$6{,}45\times10^{-4}$	$4{,}34\times10^{-1}$	$3{,}47\times10^{-3}$	$1{,}99\times10^{-4}$	$1{,}59\times10^{-6}$
σ	$4{,}31\times10^{-4}$	$2{,}00\times10^{-1}$	$1{,}57\times10^{-3}$	$1{,}12\times10^{-5}$	$1{,}27\times10^{-7}$
$\sigma/\mu\times100\%$	66,83%	46,00%	45,32%	5,64%	7,99%
Exato	$\mathbf{2\times10^{-1}}$	$\mathbf{1\times10^{-3}}$	$\mathbf{1\times10^{-5}}$	$\mathbf{2\times10^{-4}}$	$\mathbf{1{,}6\times10^{-6}}$

Tabela 2(b): Resultado de dez execuções com o método PCA para minimização da função objetivo. Número de avaliações da função objetivo em cada execução: 10.000.

$Q(\mathbf{Z})$, Eq. (7)	β	K_2	K_4	t_1	t_2
$8{,}584559\times10^{-2}$	$6{,}571396\times10^{-4}$	$2{,}797073\times10^{-1}$	$2{,}210363\times10^{-3}$	$1{,}838067\times10^{-4}$	$1{,}451563\times10^{-6}$
$9{,}324358\times10^{-2}$	$6{,}708121\times10^{-4}$	$3{,}216985\times10^{-1}$	$2{,}048131\times10^{-3}$	$2{,}157993\times10^{-4}$	$1{,}372989\times10^{-6}$
$8{,}844374\times10^{-2}$	$1{,}020589\times10^{-3}$	$1{,}885524\times10^{-1}$	$1{,}519088\times10^{-3}$	$1{,}924344\times10^{-4}$	$1{,}548782\times10^{-6}$
$9{,}614505\times10^{-2}$	$4{,}781834\times10^{-3}$	$4{,}070312\times10^{-2}$	$3{,}456255\times10^{-4}$	$1{,}946356\times10^{-4}$	$1{,}644821\times10^{-6}$
$3{,}088550\times10^{-1}$	$9{,}998760\times10^{-1}$	$3{,}392826\times10^{-4}$	$1{,}373229\times10^{-2}$	$3{,}392405\times10^{-4}$	$1{,}702070\times10^{-6}$
$8{,}939226\times10^{-2}$	$9{,}984766\times10^{-1}$	$1{,}973948\times10^{-4}$	$9{,}859282\times10^{-4}$	$1{,}970941\times10^{-4}$	$1{,}499652\times10^{-6}$
$8{,}790296\times10^{-2}$	$2{,}964068\times10^{-4}$	$6{,}258989\times10^{-1}$	$6{,}080997\times10^{-3}$	$1{,}855207\times10^{-4}$	$1{,}801915\times10^{-6}$
$1{,}002549\times10^{-1}$	$8{,}688597\times10^{-4}$	$2{,}196503\times10^{-1}$	$1{,}830960\times10^{-3}$	$1{,}908453\times10^{-4}$	$1{,}589465\times10^{-6}$
$9{,}048696\times10^{-2}$	$8{,}619339\times10^{-4}$	$2{,}529366\times10^{-1}$	$1{,}816846\times10^{-3}$	$2{,}180146\times10^{-4}$	$1{,}564651\times10^{-6}$
$9{,}632355\times10^{-2}$	$2{,}171329\times10^{-4}$	$9{,}449069\times10^{-1}$	$8{,}203482\times10^{-3}$	$2{,}051704\times10^{-4}$	$1{,}780859\times10^{-6}$
μ	$2{,}01\times10^{-1}$	$2{,}87\times10^{-1}$	$3{,}88\times10^{-3}$	$2{,}12\times10^{-4}$	$1{,}60\times10^{-6}$
σ	$3{,}99\times10^{-1}$	$2{,}81\times10^{-1}$	$4{,}02\times10^{-3}$	$4{,}37\times10^{-5}$	$1{,}32\times10^{-7}$
$\sigma/\mu\times100\%$	198,83%	97,74%	103,69%	20,61%	8,25%
Exato	$\mathbf{2\times10^{-1}}$	$\mathbf{1\times10^{-3}}$	$\mathbf{1\times10^{-5}}$	$\mathbf{2\times10^{-4}}$	$\mathbf{1{,}6\times10^{-6}}$

menor. Cabe ressaltar ainda que o método DE apresentou melhor desempenho do que o PCA, requerendo um esforço computacional menor (metade de avaliações da função objetivo em comparação com o PCA) para se atingir níveis semelhantes no valor da função objetivo.

Ficando clara a impossibilidade de se estimar simultaneamente os três parâmetros primitivos do modelo, Eq. (2), a sequência da análise se baseia num procedimento de estimativa de parâmetros em duas etapas, onde na primeira delas apenas o coeficiente de difusão K_2 é determinado. De fato, considerando-se uma situação experimental onde o processo de retenção de difusão não ocorre, ou seja, $\alpha = 0$ na figura 1, que corresponde a $\beta = 1$, o modelo se reduz ao problema clássico de difusão pura, e o coeficiente de difusão é então o único parâmetro a ser estimado. O novo problema inverso consiste então na busca de estimativas para os parâmetros do modelo de difusão anômala, β e K_4, considerando-se que K_2 é conhecido a priori (estimado na primeira etapa).

Os resultados do problema inverso, consideran-

do-se $K_2 = 1,0 \times 10^{-5}$, sendo estimados então β e K_4 são apresentados nas tabelas 3(a,b), obtidos com os métodos DE e PCA, respectivamente. Diferentemente dos resultados apresentados nas tabelas 2(a,b), é agora considerado que em cada uma das dez execuções dos métodos estocásticos o mesmo conjunto de dados experimentais foi empregado. Ou seja, o desvio padrão calculado refere-se à natureza estocástica dos métodos, e não às incertezas devidas ao ruído experimental. Nas tabelas 4(a,b) são apresentados resultados semelhantes, mas agora considerando-se em cada execução um conjunto diferente de dados experimentais, simulando-se experimentos independentes. Os dez conjuntos de dados experimentais empregados para a solução do problema inverso com o método DE foram também utilizados para o método PCA. Os novos desvios padrão calculados para β e K_4, refletem, portanto, adicionalmente à aleatoriedade dos resultados dos métodos estocásticos, o efeito da variação dos dados experimentais usados nas estimativas, ou seja, os desvios padrão calculados nas tabelas 4(a,b) possuem uma componente devido à natureza estocástica dos métodos, e outra componente devido às flutuações nos dados experimentais. Ressalta-se que os desvios padrão calculados na tabela 3 são pelo menos duas ordens de grandeza inferiores aos seus correspondentes, apresentados na tabela 4, indicando que os desvios padrão calculados na tabela 4 se referem principalmente à incerteza na estimativa dos parâmetros devido ao ruído experimental. Estes desvios padrão serão denotados por: $\sigma_{\beta,\exp}$ e $\sigma_{K_4,\exp}$.

Tabela 3(a): Resultado de dez execuções com o mesmo conjunto de dados experimentais com o método DE, fixando-se previamente o valor de $K_2 = 10^{-3}$

$Q(\mathbf{Z})$, Eq. (7)	β	K_4
9,749529x10^{-2}	2,039761x10^{-1}	9,882027x10^{-6}
9,749578x10^{-2}	2,038614x10^{-1}	9,887180x10^{-6}
9,749526x10^{-2}	2,039669x10^{-1}	9,882326x10^{-6}
9,749525x10^{-2}	2,039652x10^{-1}	9,882387x10^{-6}
9,749528x10^{-2}	2,039727x10^{-1}	9,882116x10^{-6}
9,749528x10^{-2}	2,039729x10^{-1}	9,882113x10^{-6}
9,749578x10^{-2}	2,038602x10^{-1}	9,887225x10^{-6}
9,749532x10^{-2}	2,039868x10^{-1}	9,881586x10^{-6}
9,749530x10^{-2}	2,039806x10^{-1}	9,881873x10^{-6}
9,749527x10^{-2}	2,039704x10^{-1}	9,882234x10^{-6}
μ	2,039513x10^{-1}	9,883107x10^{-6}
σ	4,812125x10^{-5}	2,170777x10^{-9}
$\sigma / \mu \times 100\%$	0,02%	0,02%
Exato	**2x10^{-1}**	**1x10^{-5}**

Tabela 3(b): Resultado de dez execuções com o mesmo conjunto de dados experimentais com o método PCA, fixando-se previamente o valor de $K_2 = 10^{-3}$.

$Q(\mathbf{Z})$, Eq. (7)	β	K_4
7,886666x10^{-2}	1,933833x10^{-1}	1,030744x10^{-5}
7,886948x10^{-2}	1,927115x10^{-1}	1,035465x10^{-5}
7,886511x10^{-2}	1,933360x10^{-1}	1,029984x10^{-5}
7,886653x10^{-2}	1,934446x10^{-1}	1,030071x10^{-5}
7,886600x10^{-2}	1,932088x10^{-1}	1,030600x10^{-5}
7,886490x10^{-2}	1,932299x10^{-1}	1,030376x10^{-5}
7,886953x10^{-2}	1,929439x10^{-1}	1,030756x10^{-5}
7,886657x10^{-2}	1,931346x10^{-1}	1,032135x10^{-5}
7,886513x10^{-2}	1,931367x10^{-1}	1,030878x10^{-5}
7,886589x10^{-2}	1,933180x10^{-1}	1,030693x10^{-5}
μ	1,931847x10^{-1}	1,031170x10^{-5}
σ	2,203511x10^{-4}	1,619796x10^{-8}
$\sigma / \mu \times 100\%$	0,11%	0,16%
Exato	**2x10^{-1}**	**1x10^{-5}**

Evidentemente há ainda de se considerar a propagação de incerteza referente ao valor de K_2, que foi considerado exatamente conhecido nos resultados apresentados nas tabelas 3 e 4, mas que possui alguma incerteza, inevitável do procedimento experimental, qualquer que seja, que interfere em sua estimativa. De modo a avaliar tal efeito, considere os coeficientes t_1 e t_2 definidos nas Eqs. 14(a,b), donde podemos escrever:

$$\beta = \frac{t_1}{K_2} \quad , \quad K_4 = \frac{(K_2^2 t_2)}{(K_2 - t_1)t_1} \tag{17a,b}$$

Considerando uma aproximação linear para a propagação do erro, definem-se os seguintes desvios padrão para β e K_4, resultantes da incerteza em K_2 (caracterizada por σ_{K_2}):

$$\sigma_{\beta,K_2} = \frac{\partial \beta}{\partial K_2}\sigma_{K_2} \quad , \quad \sigma_{K_4,K_2} = \frac{\partial K_4}{\partial K_2}\sigma_{K_2} \tag{18a,b}$$

Logo, os desvios padrão nas estimativas de β e K_2 devido às incertezas experimentais e devido à incerteza no valor de K_2, serão considerados como:

$$\sigma_\beta = \sigma_{\beta,\exp} + \sigma_{\beta,K_2} \tag{19a}$$

$$\sigma_{K_4} = \sigma_{K_4,\exp} + \sigma_{K_4,K_2} \tag{19b}$$

Tabela 4(a): Resultado de dez execuções, simulando-se experimentos independentes, com o método DE, fixando-se previamente o valor de $K_2 = 10^{-3}$.

$Q(\mathbf{Z})$, Eq. (7)	β	K_4
8,584339x10^{-2}	1,837785x10^{-1}	9,679795x10^{-6}
9,324325x10^{-2}	2,158869x10^{-1}	8,109730x10^{-6}
8,843953x10^{-2}	1,922899x10^{-1}	9,976707x10^{-6}
9,614470x10^{-2}	1,945853x10^{-1}	1,048813x10^{-5}
1,008963x10^{-1}	2,038272x10^{-1}	9,880190x10^{-6}
8,939152x10^{-2}	1,970769x10^{-1}	9,483118x10^{-6}
8,790099x10^{-2}	1,853000x10^{-1}	1,193963x10^{-5}
1,002538x10^{-1}	1,909061x10^{-1}	1,029269x10^{-5}
9,048609x10^{-2}	2,177839x10^{-1}	9,193903x10^{-6}
9,632174x10^{-2}	2,050785x10^{-1}	1,091935x10^{-5}
μ	1,986513x10^{-1}	9,996324x10^{-6}
σ	1,178579x10^{-2}	1,029492x10^{-6}
$\sigma / \mu \times 100\%$	5,93%	10,30%
Exato	**2x10^{-1}**	**1x10^{-5}**

Tabela 4(b): Resultado de dez execuções, simulando-se experimentos independentes, com o método PCA, fixando-se previamente o valor de $K_2 = 10^{-3}$.

$Q(\mathbf{Z})$, Eq. (7)	β	K_4
8,584441x10^{-2}	1,837181x10^{-1}	9,683004x10^{-6}
9,324526x10^{-2}	2,156637x10^{-1}	8,121941x10^{-6}
8,844034x10^{-2}	1,921287x10^{-1}	9,984389x10^{-6}
9,614540x10^{-2}	1,946825x10^{-1}	1,046418x10^{-5}
1,008976x10^{-1}	2,033044x10^{-1}	9,953538x10^{-6}
8,939406x10^{-2}	1,970832x10^{-1}	9,459864x10^{-6}
8,847517x10^{-2}	1,783791x10^{-1}	1,249245x10^{-5}
1,002555x10^{-1}	1,912216x10^{-1}	1,027928x10^{-5}
9,048611x10^{-2}	2,178012x10^{-1}	9,193104x10^{-6}
9,632288x10^{-2}	2,053655x10^{-1}	1,090390x10^{-5}
μ	1,979348x10^{-1}	1,005357x10^{-5}
σ	1,277000x10^{-2}	1,148281x10^{-6}
$\sigma / \mu \times 100\%$	6,45%	11,42%
Exato	**2x10^{-1}**	**1x10^{-5}**

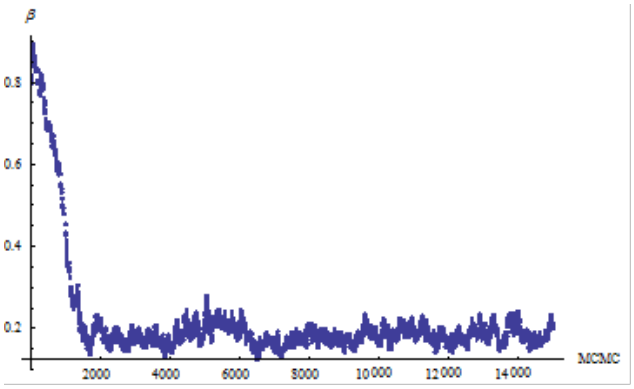
Como exemplo, consideremos $\sigma_{K_2} = K_2 \times 10\%$, ou seja, sendo K_2 uma variável aleatória com distribuição bem representada pela distribuição Normal com média 1×10^{-3}, espera-se com 95% de grau de confiança que o valor real de K_2 esteja no intervalo [0,8x10^{-3}, 1,2x10^{-3}] no exemplo aqui considerado. A tabela 5 apresenta os intervalos de confiança de 95% para β e K_4 com $\sigma_{\beta,\exp}$ e $\sigma_{K_4,\exp}$ calculados a partir dos métodos DE e PCA, nos resultados apresentados nas tabelas 4(a,b).

Na sequência são apresentados resultados empregando-se a abordagem Bayesiana, por meio do método de Monte Carlo com Cadeia de Markov (MCMC), descrito na Seção 3.2, utilizando-se prioris não informativas (distribuição uniforme) para β e K_2, e modelando-se a informação a priori de K_2 por uma distribuição Normal com média 1.0×10^{-3} e desvio padrão de 10% desta média, como foi realizado na propagação de incerteza de K_2 da análise anterior. Considerando uma cadeia com extensão total de 15 mil estados, obteve-se a evolução da cadeia de Markov apresentada nas figuras 6(a-c) para β, K_2 e K_4, respectivamente, onde fica claro que após os primeiros 2.500 estados iniciais a cadeia já está completamente convergida para a distribuição de equilíbrio. Descartando-se estes estados iniciais como amostras de aquecimento, as figuras (7a-c) apresentam os histogramas correspondentes às amostragens das distribuições a posteriori de cada um dos parâmetros, onde se chama atenção para o histograma da distribuição a posteriori de β, que não é simétrico em relação à média, ou seja, a distribuição Normal não deve ser uma boa escolha para se modelar a dispersão deste parâmetro, diferentemente de K_2 e K_4. De imediato, este fato já antecipa algum erro no cálculo nos intervalos de confiança da tabela 5, onde a distribuição Normal é uma hipótese implícita nos cálculos efetuados.

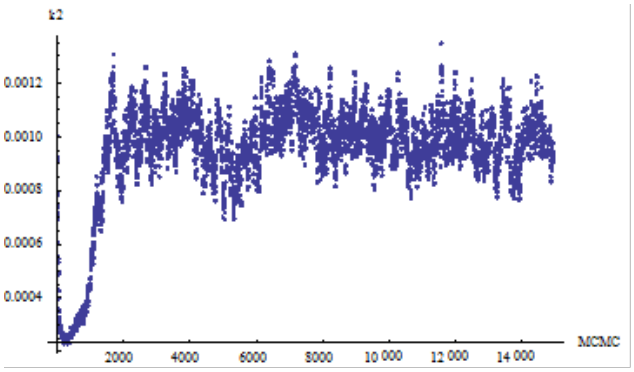
Tabela 5: Intervalos de confiança de 95% para as estimativas de β e K_4, considerando os métodos DE e PCA e propagação de incerteza de K_2.

	β	K_4
DE	[0,135, 0,262]	[0,644x10^{-5}, 1,35x10^{-5}]
PCA	[0,132, 0,263]	[0,626x10^{-5}, 1,38x10^{-5}]
Exato	**0,2**	**1,0x10^{-5}**

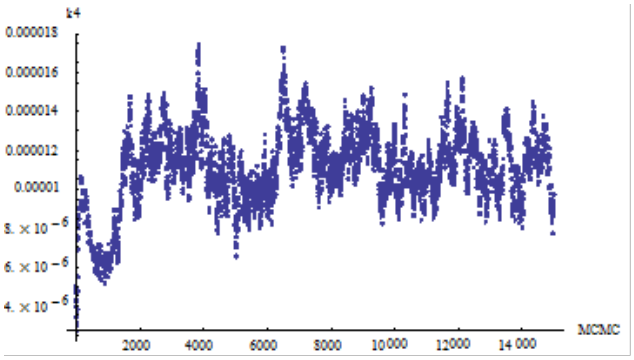
A tabela 6 apresenta os intervalos de confiança de 95% para cada uma das distribuições a posteriori amostradas com o MCMC. Os intervalos de confiança foram calculados simplesmente desprezando-se as caldas iniciais e finais de 2,5% das amostras ordenadas. Ressalta-se que o intervalo de confiança a posteriori para o parâmetro K_2 é essencialmente idêntico àquele da distribuição a priori fornecida, o que de fato era esperado uma vez que o procedimento de inversão com dados experimentais não contribui para sua estimativa, devido à dependência linear que seu respectivo coeficiente de sensibilidade apresenta com relação aos coeficientes de sensibilidade dos outros dois parâmetros. Na tabela 6, para facilitar a comparação, os intervalos de confiança apresentados na tabela 5 são repetidos. Nesta comparação percebe-se que os intervalos de confiança da distribuição a posteriori, estimados com o MCMC, são mais estreitos, indicando que a incerteza na estimativa obtida por meio da dis-



(a)

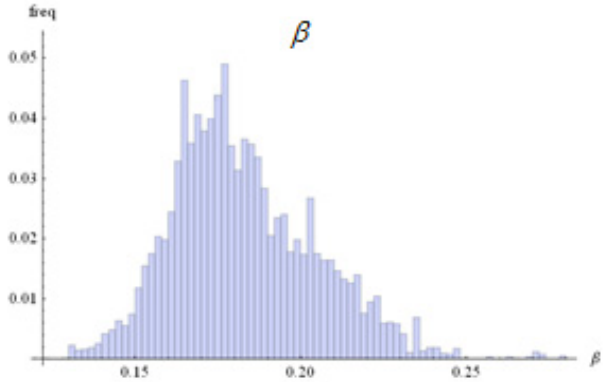


(b)

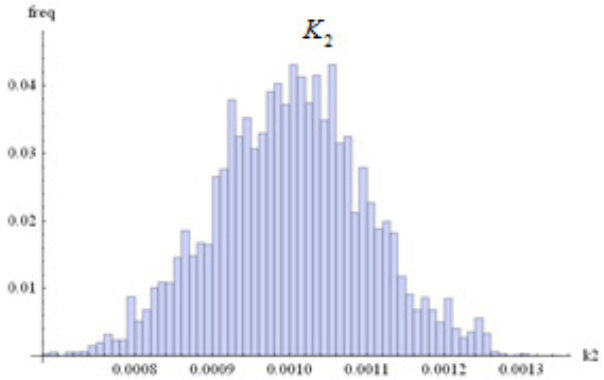


(c)

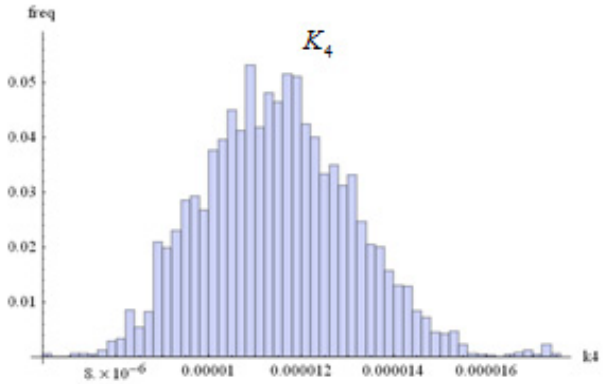
Figura 6: Evolução da cadeia de Markov para os parâmetros: (a) β, (b) K_2, e (c) K_4.



(a)



(b)



(c)

Figura 7: Histograma da distribuição a posteriori referente aos parâmetros: (a) β, (b) K_2, e (c) K_4.

persão das soluções dos métodos estocásticos, afetadas pelos erros experimentais, adicionada à propagação de incerteza do valor usado de K_2, Eqs. (19a,b) superestimou a incerteza das estimativas, provavelmente devido ao pequeno número de execuções (dez, que de fato é uma amostra pequena para se inferir o desvio padrão), além da própria variabilidade dos métodos (por exemplo, de fato se observa que os intervalos de confiança estimados com o PCA são mais largos que aqueles estimados com o DE, indicando seu pior desempenho relativo na busca pelo mínimo global da função objetivo).

Ressalta-se que em cada execução do método DE, por exemplo, 5.000 avaliações da função objetivo foram realizadas (para o PCA é feito o dobro de avaliações) e o esforço computacional total na realização das dez execuções foi muito maior do que o esforço computacional na execução do MCMC, que exigiu 15.000 avaliações da função objetivo no total. A questão é que na solução do problema inverso com o procedimento da máxima verossimilhança apenas se está interessado no mínimo global

Tabela 6: Intervalos de confiança de 95% para as estimativas dos parâmetros.

	β	K_4	K_2
DE	[0,135, 0,262]	[0,644x10^{-5}, 1,35x10^{-5}]	[0,8x10^{-3}, 1,2x10^{-3}]*
PCA	[0,132, 0,263]	[0,626x10^{-5}, 1,38x10^{-5}]	[0,8x10^{-3}, 1,2x10^{-3}]*
MCMC	[0,146, 0,230]	[0,864x10^{-5}, 1,47x10^{-5}]	[0,804 x10^{-3}, 1,21x10^{-3}]**
Exato	**0,2**	**1,0x10^{-5}**	**1,0x10^{-3}**

* Estes intervalos de confiança são aqueles considerados da incerteza no conhecimento prévio de.
** Este intervalo de confiança refere-se à distribuição a posteriori de , que, como se vê, é essencialmente idêntica à distribuição a priori.

da função objetivo, enquanto na abordagem Bayesiana cada novo candidato contribui para a amostragem da distribuição a posteriori. Um meio de reduzir o custo computacional da primeira abordagem seria considerar a hibridização do método estocástico com um método determinístico (Silva Neto e Becceneri, 2012), o que possibilitaria um maior número de execuções simulando-se experimentos independentes com custo computacional reduzido, resultando numa melhor inferência amostral dos desvios padrão relacionados aos parâmetros devido às flutuações experimentais.

5 Conclusões

Recentemente uma nova formulação analítica foi desenvolvida para a modelagem de fenômenos de difusão anômala por meio da introdução de um parâmetro que controla a fração de partículas do sistema que se difunde, resultando numa equação com um termo diferencial de quarta ordem. A caracterização dos parâmetros deste modelo de difusão anômala, que além do coeficiente de difusão incluem mais dois parâmetros, requer a análise através de formulação e solução de um problema inverso juntamente com técnicas experimentais específicas.

Este artigo consolidou a formulação e solução de problemas inversos neste novo modelo por meio de duas abordagens: (i) o procedimento da máxima verossimilhança, que resulta num problema de otimização, neste trabalho resolvido com dois métodos estocásticos de otimização, o método de evolução diferencial (DE) e o algoritmo de colisões de partículas (PCA); e (ii) a abordagem Bayesiana, que permite a incorporação de qualquer informação a priori disponível sobre os parâmetros no modelo do problema inverso, que foi resolvido com o método de Monte Carlo com cadeia de Markov.

A característica de tratamento explícito das incertezas nesta segunda abordagem foi particularmente interessante neste trabalho, uma vez que foi demonstrado não ser possível estimar os três parâmetros do modelo simultaneamente, requerendo a caracterização em duas etapas, e a análise de incertezas nas estimativas obtidas, resultantes dos erros experimentais e da incerteza da estimativa obtida na primeira etapa.

A análise de incerteza para o procedimento de máxima verossimilhança resultou em estimativas mais pobres dos intervalos de confiança principalmente devido ao pequeno número de execuções realizadas simulando-se diferentes conjuntos de dados experimentais, o que naturalmente resultou numa inferência amostral mais pobre para os desvios padrão dos parâmetros relacionados às flutuações experimentais.

Em trabalhos futuros sugere-se a utilização do método da máxima verossimilhança com a hibridização de métodos estocásticos com métodos determinísticos de otimização, possibilitando a redução do custo computacional e consequentemente um maior número de execuções, simulando-se diferentes conjuntos de dados experimentais, e considerando o método de Monte Carlo para propagação da incerteza associada ao coeficiente de difusão.

Agradecimentos

Os autores agradecem o suporte financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, e da Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, FAPERJ. Os autores agradecem também a UFSM pelo convite para envio deste artigo para a revista Ciência e Natura.

Referências

Atsumi, H., 2002. Hydrogen bulk retention in graphite and kinetics of diffusion. J Nuc Mat 307/311, pp. 1466-1470.

Beck, J. V., Blackwell, B., e St. Clair Jr., C. R., 1985. Inverse Heat Conduction: Ill- Posed Problems. John Wiley & Sons, Nova York.

Bevilacqua, L., Galeão, A. C. N. R., e Costa, F. P.,

2011a. A new analytical formulation of retention effects on particle diffusion process. Annals of the Brazilian Academy of Sciences, 83 (4), pp. 1443-1464.

Bevilacqua, L., Galeão, A. C. N. R., e Costa, F. P., 2011b. On the significance of higher order differential terms in diffusion processes. Journal of the Brazilian Society of Mechanical Sciences and Engineering 34, pp. 166-175.

Brandani, S., Jama, M., e Ruthven, D., 2000. Diffusion, self-diffusion and counter-diffusion of benzene and p-xylene in silicalite. Micropor Mesopor Mat 35/36, pp. 283–300.

Cheng, H., e Fu, C. L., 2012. An iteration regularization for a time-fractional inverse diffusion problem. Applied Mathematical Modelling 36, pp. 5642-5649.

D'Angelo, M. V., Fontana, E., Chertcoff, R., e Rosen, M., 2003. Retention phenomena in non-Newtonian fluid flows. Physics A 327, pp. 44-48.

Deleersnijder, E., Beckers, J.-M., e Delhez, E. M., 2006. The residence time of setting in the surface mixed layer. Environ Fluid Mech 6, pp. 25-42.

Derec, C., Smerlak, M., Servais, J., e Bacri, J.-C., 2010. Anomalouos diffusion in microchannel under magnetic field. Physics Procedia 9, pp. 109-112.

Green, P. F., 1996. Translational dynamics of macromolecules in metals. In: Neogi, P. (ed.), Diffusion in Polymers. Marcel Dekker Inc.

Joannès, S., Mazé, L., e Bunsell, A. R., 2014. A concentration-dependent diffusion coefficient model for water sorption in composite. Composite Structures 108, pp. 111-118.

Kaipio, J., Somersalo, E., 2004. Statistical and Computational Inverse Problems. Springer Verlag.

Liu, H., e Thompson, K. E., 2002. Numerical modeling of reactive polymer flow in porous media. Comput Chem Eng 26, pp. 1595-1610.

Metzler, R., e Klafter, J., 2000. The random walk's guide to anomalous diffusion: a fractional dynamics approach. Phys. Rep. 339, pp. 1–77.

Moura Neto, F. D., e Silva Neto, A. J., 2012. An Introduction to Inverse Problems with Applications. Springer, Heidelberg.

Muhammad, N., 2004. Hydraulic, diffusion, and retention characteristics of inorganic chemicals in bentonite. Ph.D. Thesis. Department of Civil and Environmental Engineering, College of Engineering, University of South Florida. Florida.

Orlande, H. R. B., Colaço, M. J., Cotta, C. P. N., Guimarães, G., e Borges, V. L., 2011. Problemas Inversos em Transferência de Calor. Sociedade Brasileira de Matemática Aplicada e Computacional – SBMAC (Notas em Matemática Aplicada, v. 51), São Carlos.

Price, K., Storn, R., Lampinem, J. A., 2006. Differential Evolution: A Practical Approach to Global Optimization. Springer-Verlag.

Sacco, W. F., Oliveira, C. R. E., e Pereira, C. M. N. A., 2006. Two sthocastic optimization algorithms applied to nuclear reactor core design. Progress in Nuclear Energy 48, pp. 525-539.

Schwaab, M., e Pinto, J. C., 2007. Análise de Dados Experimentais I: Fundamentos de Estatística e Estimação de Parâmetros. E-papers, Rio de Janeiro.

Silva, L. G., 2013. Problemas Inversos em Processos Difusivos com Retenção. Dissertação de Mestrado em Modelagem Computacional. Instituto Politécnico do Rio de Janeiro. Universidade do Estado do Rio de Janeiro, Nova Friburgo.

Silva Neto, A. J., e Becceneri, J. C., 2012. Técnicas de Inteligência Computacional Inspiradas na Natureza – Aplicação em Problemas Inversos em Transferência Radiativa, 2ª Edição. Sociedade Brasileira de Matemática Aplicada e Computacional – SBMAC (Notas em Matemática Aplicada, v. 41), São Carlos.

Storn, R., e Price, K., 1997. Differential Evolution: a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 11, pp. 331-359.

Tannehill, J.C., Anderson, D.A., e Pletcher, R.H., 1997. Computational Fluid Mechanics and Heat Transfer, 2a Edição. Taylor & Francis.

Wei, T., e Zhang, Z. Q., 2013. Reconstruction of a time-dependent source term in a time-fractional diffusion equation. Engineering Analysis with Boundary Elements 37, pp. 23-31.

Wu, J., e Berland, K. M., 2008. Propagators and time-dependent diffusion coefficients for anomalous diffusion. Biophysical Journal 95, pp. 2049-2052.

Xiong, X., Guo, H., e Liu, X., 2012. An inverse problem for a fractional diffusion equation. Journal of Computation and Applied Mathematics 236, pp. 4474-4484.

Zhang, Z. Q., e Wei, T., 2013. Identifying an unknown source in time-fractional diffusion equation by a truncation method. Applied Mathematics and Computation 219, pp. 5972-5983.