

08 a 11 de Outubro de 2018  
Instituto Federal Fluminense  
Búzios - RJ

## COMPARAÇÃO ENTRE OS MÉTODOS METROPOLIS-HASTINGS E MONTE CARLO HAMILTONIANO NA IDENTIFICAÇÃO DE PROPRIEDADES TERMOFÍSICAS EM PROBLEMAS DE CONDUÇÃO DE CALOR

**Maicon de Paiva Torres** – mptorres@iprj.uerj.br

**Leonardo Tavares Stutz** – ltstutz@iprj.uerj.br

**Antônio J. da Silva Neto** – ajsneto@iprj.uerj.br

Universidade do Estado do Rio de Janeiro, Instituto Politécnico – Nova Friburgo, RJ, Brazil

**Resumo.** Neste trabalho é formulado e resolvido o problema inverso de transferência de calor por condução, para a identificação de propriedades termofísicas, i.e., a capacidade térmica volumétrica e a condutividade térmica de um material polimérico, via inferência Bayesiana. Para a solução do problema inverso formulado com base no teorema de Bayes são considerados dois métodos de Monte Carlo com cadeias de Markov: o método de Metropolis-Hastings (MH) e o método de Monte Carlo Hamiltoniano (HMC). O mesmo modelo matemático é usado na geração dos dados experimentais sintéticos e na solução do problema direto, contudo, diferentes técnicas de solução são consideradas de forma a atenuar o crime inverso. Para tanto, a técnica da transformada integral clássica (CITT) é usada para gerar os dados sintéticos enquanto o método das diferenças finitas (FDM) é empregado na solução do problema direto. Dos resultados analisados é possível concluir que o método HMC, por ter um mecanismo determinístico para a exploração do espaço dos parâmetros, é mais eficiente e consistente que o MH. As cadeias de Markov dos parâmetros obtidos via HMC possuem menor burn-in (região de aquecimento) e as amostras são menos correlacionadas.

**Palavras-chave:** Condução de Calor, Inferência Bayesiana, Método de Monte Carlo Hamiltoniano (HMC).

### 1. INTRODUÇÃO

O algoritmo de Metropolis-Hastings (MH), proposto inicialmente por Metropolis et al. (1953), é considerado o mais simples e o mais usado método de Monte Carlo com cadeias de Markov (MCMC). Embora seja efetivo ao explorar distribuições de baixa dimensão, ele demonstra ser ineficiente ao tratar distribuições complexas e de dimensão elevada (Hanson, 2001). Além disso, estados sucessivos podem apresentar alta correlação devido à natureza

aleatória do movimento no espaço dos parâmetros. Como resultado, o tamanho efetivo da amostra tende a ser pequeno e a convergência para a distribuição a ser lenta.

De forma a suprir a ineficiência dos métodos MCMC com caminhada aleatória, no trabalho de Duane et al. (1987) foi proposto o método de Monte Carlo Hamiltoniano (HMC), inicialmente conhecido com método de Monte Carlo Híbrido. O HMC possui um mecanismo determinístico inspirado na dinâmica Hamiltoniana para propor amostras para a distribuição de probabilidade alvo (Neal et al., 2011).

Este trabalho tem como objetivo a solução do problema inverso de identificação de propriedades termofísicas de um material polimérico. Para tanto, é considerado um problema de condução de calor transiente, no qual o modelo físico-matemático é baseado, e a inferência Bayesiana, por meio da qual a solução é tratada como uma inferência estatística que combina a nova informação (medidas experimentais) com a informação conhecida *a priori*. Neste contexto, um método MCMC de caminhada determinística, conhecido como método de Monte Carlo Hamiltoniano (HMC), é usado para gerar amostras da distribuição de probabilidade alvo.

De forma a comprovar sua eficácia, são realizadas comparações entre as soluções obtidas via MH e HMC considerando dados experimentais sintéticos com 5% de intensidade de ruído. Apesar deste ser um problema inverso de estimação de baixa dimensão, é possível avaliar a autocorrelação das cadeias de Markov, a fim de se analisar a qualidade das amostras obtidas.

## 2. FORMULAÇÃO E SOLUÇÃO DO PROBLEMA DIRETO

Considere uma placa termicamente fina, de seção retangular, posicionada na vertical. Uma de suas faces está sujeita a um fluxo de calor prescrito enquanto a face oposta troca calor por convecção natural com o ambiente na temperatura  $T_\infty$  (vide Figura 1). Se os gradientes de temperatura forem desprezados na direção da largura da placa (direção  $y$ ), o problema de condução de calor transiente pode ser modelado por

$$w \frac{\partial T_m(x,t)}{\partial t} + \Phi\{T_m(x,t)\} = P(x), \quad 0 < x < L_x, \quad t > 0, \quad (1a)$$

$$T_m(x, 0) = T_\infty, \quad 0 \leq x \leq L_x, \quad (1b)$$

$$-k \frac{\partial T(x,t)}{\partial x} \Big|_{x=0} = 0; \quad k \frac{\partial T(x,t)}{\partial x} \Big|_{x=L_x} = 0, \quad t > 0, \quad (1c,d)$$

em que a capacidade térmica volumétrica é considerada constante e é dada pela expressão  $w = \rho c_p$ , sendo  $\rho$  e  $c_p$  a massa específica e o calor específico, respectivamente, do material que compõe a placa. Neste trabalho, a condutividade térmica  $k$  também é considerada constante. O termo fonte  $P(x)$  e o operador linear  $\Phi$ , que aparecem na equação principal, são dados por

$$P(x) = \frac{1}{L_z} [h(x)T_\infty + q''(x)], \quad (2)$$

$$\Phi = -k \frac{\partial^2}{\partial x^2} (\cdot) + \frac{h(x)}{L_z} (\cdot). \quad (3)$$

Em  $z = 0$  e para a metade aquecida da superfície, o fluxo de calor é exatamente o fluxo dissipado pela resistência elétrica ao longo da espessura da placa (direção  $z$ ), enquanto que na metade inferior tem-se fluxo nulo devido ao isolamento térmico. O fluxo de calor pode ser então representado por uma função degrau do tipo

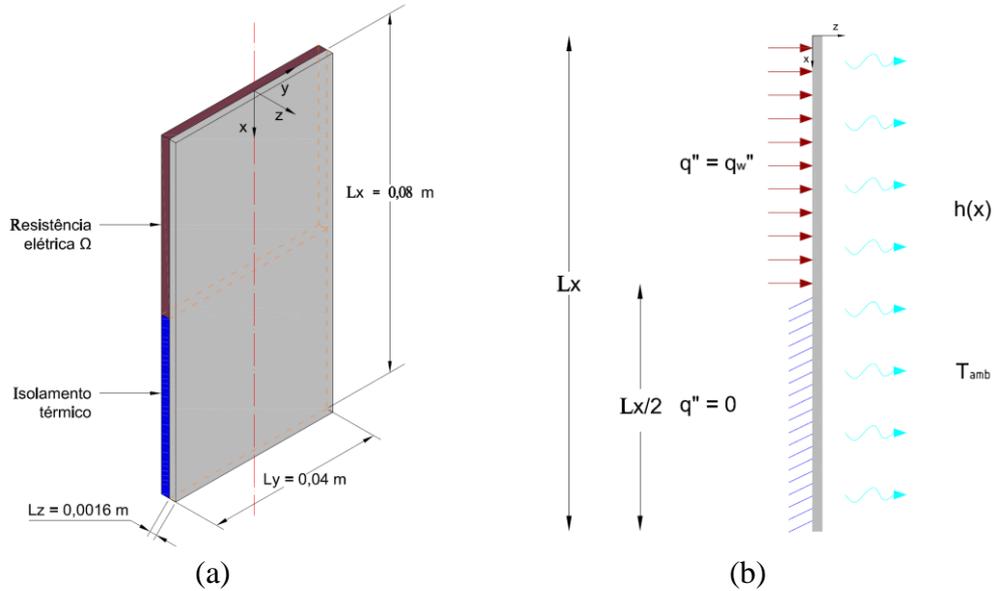


Figura 1 – (a) Configuração da placa de resina de poliéster e (b) representação esquemática do modelo físico-matemático considerado na solução do problema inverso.

$$q''(x) = \begin{cases} q''_{aq}, & 0 \leq x \leq L_x/2 \\ 0, & L_x/2 < x \leq L_x \end{cases} \quad (4)$$

sendo  $q''_{aq}$  o fluxo de calor dissipado pelo aquecedor. O coeficiente de troca térmica, que é dado por  $h(x)$ , assume valores distintos nas regiões com e sem aquecimento, sendo descrito por

$$h(x) = \begin{cases} h_{x0}, & 0 \leq x \leq L_x/2 \\ h_{xL}, & L_x/2 < x \leq L_x \end{cases} \quad (5)$$

## 2.1 Solução exata via técnica da transformada integral clássica

A solução formal exata do problema (1) é obtido por meio da transformada integral clássica, podendo esta ser escrita como (Mikhailov e Ozisik, 1984)

$$T_m(x, t) = T_\infty + \sum_{i=1}^{\infty} \tilde{\psi}_i(x) \int_0^t \tilde{g}_i(t') e^{-\mu_i^2(t-t')} dt', \quad (6)$$

em que os autovalores  $\mu_i$  e as autofunções  $\psi_i$  são obtidos do problema de autovalor de Sturm-Liouville (Cotta, 1993)

$$\Phi\{\psi(x)\} = \mu^2 w \psi(x), \quad (7a)$$

com as condições de contorno

$$-k \frac{\partial \psi(x)}{\partial x} \Big|_{x=0} = 0, \quad k \frac{\partial \psi(x)}{\partial x} \Big|_{x=L_x} = 0, \quad t > 0. \quad (7b,c)$$

As variáveis que aparecem na solução exata (6) são computadas após a solução do problema de autovalor (7), sendo elas

$$\tilde{\psi}_i(x) = \frac{\psi_i(x)}{N_i^{1/2}}, \quad (8a)$$

$$N_i = \int_0^{L_x} w \psi_i^2(x) dx, \quad (8b)$$

$$\bar{g}_i(t) = \int_0^{L_x} P(x) \tilde{\psi}_i(x) dx. \quad (8c)$$

## 2.2 Solução via método das diferenças finitas

Na solução do problema (1) por diferenças finitas, o domínio contínuo  $(x, t)$  é substituído por uma malha computacional em que cada nó  $(x_i, t_n)$  é identificado pelo par de índices  $(i, n)$ . Considerando a discretização do domínio físico e as aproximações das derivadas por expansões em séries de Taylor, com formulação implícita, o problema (1) pode ser então aproximado por

$$T_{m_i}^{n+1} = T_{m_i}^n + \frac{\alpha \Delta t}{(\Delta x)^2} (T_{m_{i-1}}^n - 2T_{m_i}^n + T_{m_i}^n) - \frac{\Delta t}{w L_z} h_i T_{m_i}^n + \frac{\Delta t}{w} P_i, \quad (9a)$$

$$T_{m_i}^0 = T_\infty, \quad (9b)$$

$$T_{m_{-1}}^n = T_{m_1}^n, \quad T_{m_{I+1}}^n = T_{m_{I-1}}^n, \quad (9c,d)$$

$$P_i = \frac{1}{L_z} [h_i T_\infty + q''_i], \quad (9e)$$

em que  $\alpha = k/w$  é a difusividade térmica, considerando  $i = 1, 2, \dots, I$  pontos no espaço e  $n = 1, 2, \dots, M$  pontos no tempo. As condições de contorno 1(c,d) foram aproximadas considerando diferenças centradas, sendo necessário o uso de nós fictícios.

## 3. FORMULAÇÃO E SOLUÇÃO DO PROBLEMA INVERSO

Considerando o problema de condução de calor pelo qual o problema direto é formulado, o vetor de parâmetros a ser estimado pode ser escrito como

$$\mathbf{Z}^T = [w, k], \quad (10)$$

em que  $w$  é a capacidade térmica volumétrica [ $J/m^3K$ ] e  $k$  é a condutividade térmica [ $W/mK$ ]. Ambas formam as propriedades térmicas de interesse do material que constitui a placa, sendo consideradas constantes devido ao meio ser homogêneo.

### 3.1 Inferência Bayesiana

O teorema de Bayes pode ser formulado como (Beck & Arnold, 1977), (Kaipio & Somersalo, 2004)

$$\pi_{posterior}(\mathbf{Z}) = \pi(\mathbf{Z}|\mathbf{Y}) = \frac{\pi(\mathbf{Z})\pi(\mathbf{Y}|\mathbf{Z})}{\pi(\mathbf{Y})}, \quad (11)$$

em que  $\pi_{posterior}(\mathbf{Z})$  é a densidade de probabilidade *a posteriori*,  $\pi(\mathbf{Z})$  é a densidade de probabilidade *a priori*,  $\pi(\mathbf{Y}|\mathbf{Z})$  é a função de verossimilhança e  $\pi(\mathbf{Y})$  é a densidade de probabilidade marginal, que é uma constante de normalização.

Assumindo que os erros experimentais sintéticos são variáveis aleatórias com distribuição Gaussiana, com média nula e matriz de covariância  $W$  conhecida, além de serem aditivos e independentes dos parâmetros  $\mathbf{Z}$ , a função de verossimilhança pode ser escrita como

$$\pi(\mathbf{Z}|\mathbf{Y}) = (2\pi)^{-N_d/2} |W|^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{Y} - T(\mathbf{Z}))^T W^{-1} (\mathbf{Y} - T(\mathbf{Z})) \right], \quad (12)$$

em que  $T(\mathbf{Z})$  é a solução do problema direto e  $N_d$  é o número de medidas experimentais.

### 3.2 Métodos de Monte Carlo com cadeias de Markov

**O algoritmo de Metropolis-Hastings (MH).** Um dos mais simples algoritmos de Monte Carlo com cadeias de Markov é o Metropolis-Hastings. Neste algoritmo, é feita uma perturbação da posição atual no espaço de parâmetros pela seleção aleatória de um candidato de uma distribuição de probabilidade simétrica. O candidato é aceito ou rejeitado baseado na probabilidade da nova posição em relação à anterior.

O primeiro passo é definir o ponto de partida no vetor de espaço a ser amostrado,  $\mathbf{Z}^0$ . Então o seguinte algoritmo é repetido para se obter as posições  $\mathbf{Z}^{(i)}$  dos estados da cadeia de Markov:

- (1) Selecionar uma nova posição candidata  $\mathbf{Z}^* = \mathbf{Z}^{(i-1)} + \Delta\mathbf{Z}$ , em que  $\Delta\mathbf{Z}$  é escolhido de forma aleatória de uma distribuição auxiliar  $q(\mathbf{Z}^*|\mathbf{Z}^{(i-1)})$ ;
- (2) Calcular a razão de Hastings,  $\alpha = \min[1, \pi(\mathbf{Z}^*)/\pi(\mathbf{Z}^{(i-1)})]$ ;
- (3) Gerar um valor aleatório  $\theta$  de uma distribuição uniforme no intervalo (0,1);
- (4) Se  $\theta < \alpha$ , estabelecer  $\mathbf{Z}^{(i)} = \mathbf{Z}^*$ . Caso contrário, fazer  $\mathbf{Z}^{(i)} = \mathbf{Z}^{(i-1)}$ ;
- (5) Retornar ao passo 1 para gerar a sequência  $\{\mathbf{Z}^1, \mathbf{Z}^2, \dots, \mathbf{Z}^n\}$ .

**O método de Monte Carlo Hamiltoniano.** O método HMC foi proposto para ser análogo a um sistema físico (Duane et al., 1987). Para cada parâmetro  $Z_i$  é introduzida uma variável de momento associada,  $p_i$  (Neal et al., 2011). O Hamiltoniano  $H$  é então construído considerando a soma de um termo de energia potencial,  $U(\mathbf{Z}) = -\log(\pi_{posterior})$ , com um termo de energia cinética, na forma

$$H(\mathbf{Z}, \mathbf{p}) = U(\mathbf{Z}) + K(\mathbf{p}), \quad (13)$$

em que  $K(\mathbf{p})$  é a energia cinética, definida por

$$K(\mathbf{p}) = \mathbf{p}^T \bar{M}^{-1} \mathbf{p} / 2, \quad (14)$$

em que  $\bar{M}$  é a matriz de massa, simétrica e positiva-definida, que é tipicamente diagonal. Esta forma para  $K(\mathbf{p})$  corresponde ao negativo do logaritmo da densidade de probabilidade (mais uma constante) de uma distribuição Gaussiana com média zero e matriz de covariância  $\bar{M}$ .

As derivadas parciais do Hamiltoniano determinam como  $\mathbf{Z}$  e  $\mathbf{p}$  mudam ao longo do tempo,  $t$ , de acordo com as seguintes equações

$$\frac{dZ_i}{dt} = \frac{\partial H}{\partial p_i}; \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial Z_i}, \quad i = 1, 2, \dots, N, \quad (15a,b)$$

sendo  $N$  o número de variáveis (ou parâmetros). Para qualquer intervalo de tempo de duração  $s$ , estas equações definem um mapeamento,  $T_s$ , do estado atual, no instante  $t$ , para o estado seguinte, no instante  $t + s$ .

Cada iteração do algoritmo começa com uma amostragem de Gibbs para gerar novas variáveis de momento provenientes da distribuição Gaussiana dada pela energia cinética. Por conseguinte, a trajetória no espaço de estados ( $\mathbf{Z}, \mathbf{p}$ ) que mantém  $H$  constante é seguida usando a técnica de integração leapfrog, que consiste das etapas (neste caso para  $L = 1$ )

$$p_i\left(t + \frac{\epsilon}{2}\right) = p_i(t) - (\epsilon/2) \frac{\partial U}{\partial z_i}(\mathbf{Z}(t)), \quad (16a)$$

$$Z_i(t + \epsilon) = Z_i(t) + \epsilon \frac{p_i(t+\epsilon/2)}{\bar{m}_i}, \quad (16b)$$

$$p_i(t + \epsilon) = p_i(t + \epsilon/2) - (\epsilon/2) \frac{\partial U}{\partial z_i}(\mathbf{Z}(t + \epsilon)), \quad (16c)$$

em que  $\epsilon$  representa o passo de tempo usado na integração e o gradiente da energia potencial é obtido por meio de uma representação centrada de diferenças finitas.

#### 4. RESULTADOS E DISCUSSÕES

O material a ter suas propriedades termofísicas identificadas neste trabalho se trata de um polímero, mais especificamente, uma resina de poliéster. De acordo com a literatura, os valores para a capacidade térmica volumétrica ( $w$ ) e condutividade térmica ( $k$ ) podem variar nos intervalos  $[1,22 - 1,76] \times 10^6$  J/m<sup>3</sup>K e  $[0,15 - 0,17]$  W/mK, respectivamente (MARK et al., 2007). Neste trabalho, os valores exatos dos parâmetros  $w$  e  $k$  são considerados como sendo as médias destes intervalos,  $1,49 \times 10^6$  J/m<sup>3</sup>K e 0,16 W/mK.

Considerando que a única informação *a priori* disponível para o material a ser identificado é que este pertence à classe dos polímeros, verifica-se que sua condutividade térmica pode variar entre 0,06 e 0,5 W/mK. Deste modo, a informação *a priori* para a condutividade térmica pode ser modelada como sendo uma distribuição uniforme neste intervalo. Já a capacidade térmica volumétrica pode assumir quaisquer valores entre  $0,7 \times 10^6$  e  $5,0 \times 10^6$  J/m<sup>3</sup>K. Sendo assim, também é conveniente modelar sua informação *a priori* como sendo a de uma distribuição uniforme neste outro intervalo.

Para obter a solução do problema direto, além dos valores dos parâmetros  $w$  e  $k$ , são considerados os valores presentes na Tabela 1. Estes são referentes ao coeficiente de troca térmica por convecção,  $h(x)$ , e ao fluxo de calor,  $q_{aq}''$ . Também são conhecidos os valores para a temperatura do ambiente que circunda a placa e a sua temperatura inicial. As dimensões da placa estão explicitadas no esquema proposto na Figura 1(a).

Tabela 1 – Parâmetros conhecidos do problema de condução de calor.

| Parâmetro                    | Símbolo    | Unidades           | Valor |
|------------------------------|------------|--------------------|-------|
| Coeficiente de troca térmica | $h_{x0}$   | W/m <sup>2</sup> K | 17    |
| Coeficiente de troca térmica | $h_{xL}$   | W/m <sup>2</sup> K | 4     |
| Fluxo de calor               | $q_{aq}''$ | W/m <sup>2</sup>   | 440,5 |
| Temperatura ambiente         | $T_\infty$ | °C                 | 24    |
| Temperatura inicial          | $T_i$      | °C                 | 24    |

Na Tabela 2 pode ser visto o comportamento da convergência das temperaturas para algumas posições e instantes de tempo selecionados. Pode-se observar que, com uma ordem de

truncamento na série da Eq. (6),  $N_{tr} = 55$ , foi possível obter uma convergência de cinco dígitos. Conclui-se então que os primeiros cinquenta e cinco autovalores são suficientes para garantir a convergência da solução analítica, sendo esta considerada a solução de referência.

Tabela 2 – Convergência da técnica da transformada integral clássica (CITT).

| $N_{tr}$ | $T_m(x, t)$ [°C] |              |              | $T_m(x, t)$ [°C] |              |              |
|----------|------------------|--------------|--------------|------------------|--------------|--------------|
|          | $t = 305$ s      |              |              | $t = 610$ s      |              |              |
|          | $x = 0,02$ m     | $x = 0,04$ m | $x = 0,06$ m | $x = 0,02$ m     | $x = 0,04$ m | $x = 0,06$ m |
| 35       | 46,9588          | 37,5207      | 24,0418      | 49,5340          | 40,3108      | 24,3987      |
| 40       | 46,9540          | 37,5199      | 24,0449      | 49,5298          | 40,3100      | 24,4012      |
| 45       | 46,9622          | 37,5204      | 24,0370      | 49,5368          | 40,3105      | 24,3944      |
| 50       | 46,9589          | 37,5200      | 24,0409      | 49,5340          | 40,3102      | 24,3978      |
| 55       | 46,9590          | 37,5203      | 24,0405      | 49,5340          | 40,3104      | 24,3974      |

Uma representação por diferenças finitas é dita consistente se o erro de truncamento tender a zero conforme a malha computacional é refinada. Considerando a solução de referência para o problema, é calculada a diferença com as soluções obtidas via FDM (Finite Difference Method) em diferentes posições e instantes de tempo. Pela Tabela 3 é possível notar que, conforme a malha espacial é refinada, a diferença entra as soluções analítica e numérica é reduzida. Esse efeito se destaca no centro da placa, em  $x = 0,04$  m, onde encontram-se as maiores diferenças. Conclui-se então que esta representação de diferenças finitas é consistente e estável.

Tabela 3 – Diferença entre as temperaturas analítica e numérica.

| $\Delta x$ [m] | $T_m(x, t) - Y_i$ |              |              |              |              |              |
|----------------|-------------------|--------------|--------------|--------------|--------------|--------------|
|                | $t = 305$ s       |              |              | $t = 610$ s  |              |              |
|                | $x = 0,02$ m      | $x = 0,04$ m | $x = 0,06$ m | $x = 0,02$ m | $x = 0,04$ m | $x = 0,06$ m |
| 1/250          | 0,0377            | 4,3291       | 0,0641       | 0,0159       | 3,9244       | 0,2140       |
| 1/500          | 0,0373            | 2,3849       | 0,0405       | 0,0109       | 2,1784       | 0,1009       |
| 1/1000         | 0,0064            | 1,2422       | 0,0099       | 0,0067       | 1,1387       | 0,0488       |
| 1/2000         | 0,0056            | 0,6331       | 0,0045       | 0,0041       | 0,5809       | 0,0224       |
| 1/4000         | 0,0051            | 0,3203       | 0,0021       | 0,0026       | 0,2935       | 0,0119       |

Para os resultados que seguem, foi considerada a CITT com  $N_{tr} = 55$  como solução de referência (analítica), enquanto que a solução numérica é obtida via FDM considerando uma malha computacional com  $\Delta x = 1/4000$  m e  $\Delta t = 0,2$  s.

Os dados experimentais sintéticos foram gerados pela adição de um sinal ruidoso com nível controlado na solução de referência, sendo geradas por (Knupp & Silva Neto, 2013)

$$Y_i = T_i(\mathbf{Z}_{exato}) + \sigma e_i, \quad i = 1, 2, \dots, M, \quad (17a)$$

em que  $M$  é o número de medições no tempo,  $e_i$  é um número aleatório gerado de uma distribuição normal com média zero e desvio padrão unitário e  $\sigma$  representa o desvio padrão dos dados experimentais. O nível de ruído pode então ser calculado por

$$\text{ruído}(\%) = \max \left| \frac{e_i \sigma}{T_i} \right| \times 100\%, \quad i = 1, 2, \dots, M, \quad (17b)$$

sendo o nível de ruído considerado como o maior valor obtido.

Na Figura 2 é apresentada a evolução temporal da temperatura experimental sintética avaliada na posição  $x = 0,045$  m. Empregando o desvio padrão para as medições de  $\sigma = 0,5$  °C, o maior valor para o ruído calculado pela Eq. (21b) não ultrapassou 5%.

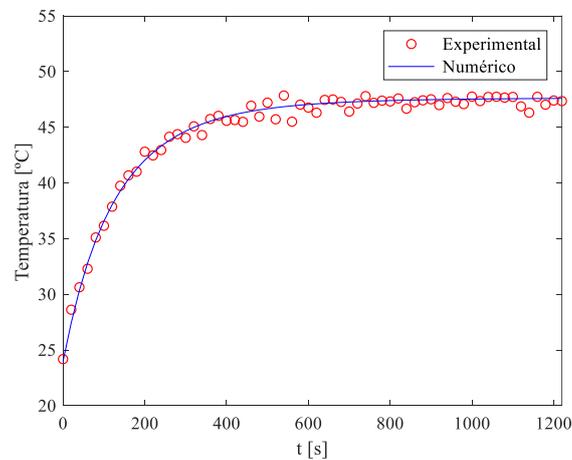


Figura 2 – Resposta experimental sintética gerada na posição  $x = 0,045$  m com 5% de ruído.

Para uma comparação justa entre as soluções via HMC e MH, é considerado apenas um passo de tempo para o método de integração numérica Leapfrog, *i.e.*,  $L = 1$ . Partindo da posição inicial,  $\mathbf{Z}^0 = [2,15 \times 10^6 \text{ J/m}^3\text{K}; 0,22 \text{ W/mK}]$ , nas Figuras 3(a,b) estão as cadeias de Markov dos parâmetros  $w$  e  $k$ , respectivamente, obtidas via MH e HMC. Foram considerados 1000 estados. Para o MH atingir a distribuição estacionária, são necessários cerca de 500 estados para as cadeias de  $w$  e de  $k$ . Para o HMC foram necessários apenas 200 estados. Mesmo com  $L = 1$  fica evidente como a convergência do HMC é mais acelerada que a do MH.

Do ponto de vista qualitativo, pode-se analisar as cadeias quanto a sua autocorrelação. Na Figura 4(a,b) estão os gráficos de autocorrelação para  $w$  obtidas pelos métodos MH e HMC, respectivamente. Os intervalos de confiança de 95% para a sequência de autocorrelação são as linhas horizontais. Enquanto no MH são necessários cerca de 300 estados para obter amostras não correlacionadas, no HMC são necessários 100 estados. Com isso, o número de amostras necessárias para a inferência das distribuições de probabilidade *a posteriori*, com a mesma significância estatística, é consideravelmente menor, para a cadeia de Markov obtida via HMC.

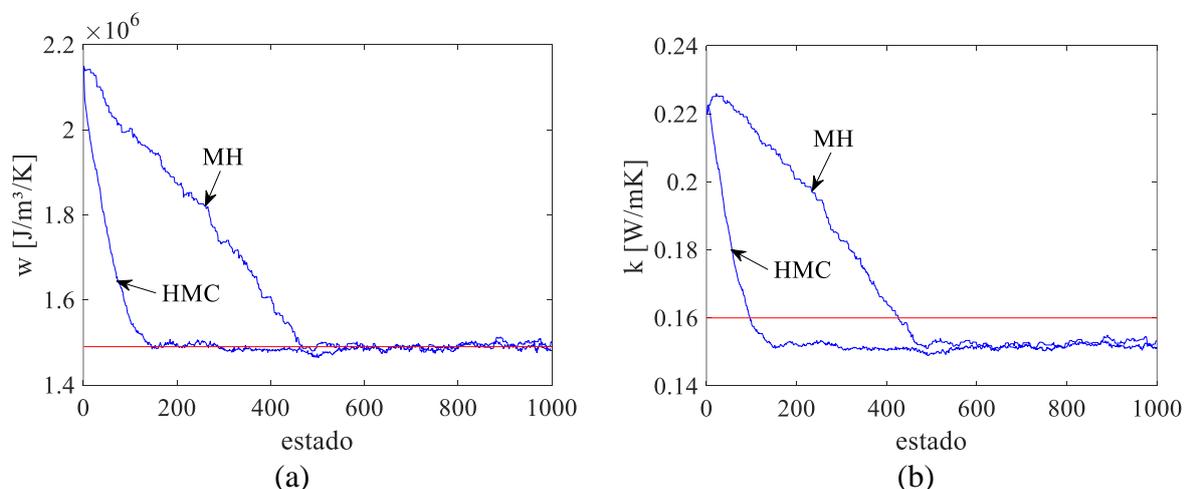


Figura 3 – Cadeias de Markov para os parâmetros estimados (a)  $w$  e (b)  $k$ .

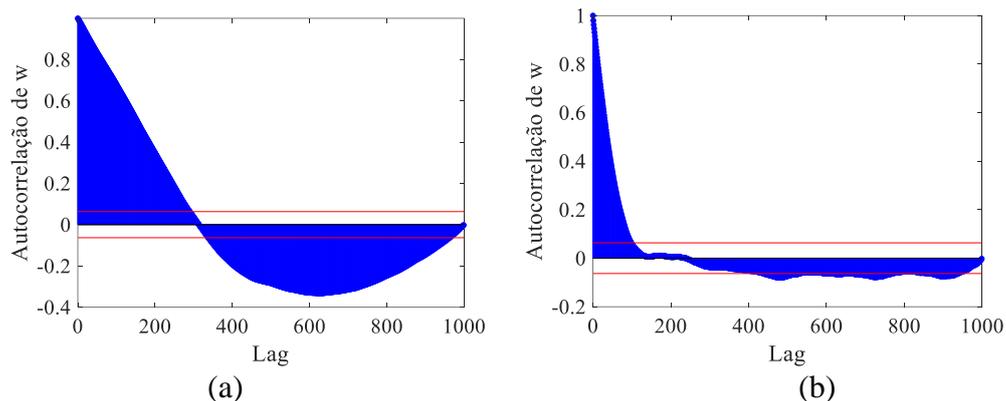


Figura 4 – Função de autocorrelação para as cadeias do parâmetro  $w$  identificado (a) via MH e (b) via HMC.

O mesmo comportamento é observado nos gráficos de autocorrelação para  $k$ . Também são necessários 300 estados para obter amostras não correlacionadas com o MH, enquanto que com o HMC são necessários apenas 100 estados.

Nas Figuras 5(a,b) estão os gráficos de dispersão  $w \times k$  e seus respectivos histogramas, obtidos via MH e HMC, respectivamente. Apesar dos histogramas serem similares, fica evidente que a amostragem gerada pelo HMC é qualitativamente superior, pois possui um número de amostras distintas superior à obtida com o MH. Isso ocorre principalmente pela taxa de aceitação das duas soluções. No MH foram aceitos 39,9% dos candidatos, enquanto que o HMC teve uma taxa de aceitação de 99,8%, que é consideravelmente maior.

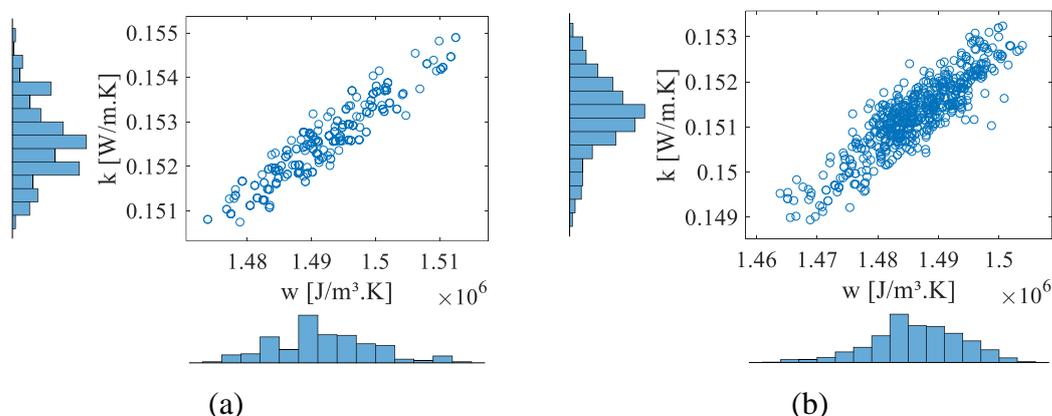


Figura 5 – Dispersão e histogramas para os parâmetros identificados via (a) MH e (b) HMC.

Finalmente, na Tabela 4, são apresentados os resultados numéricos das simulações, sendo eles a média e o desvio padrão, obtidos da distribuição *a posteriori* via MH e HMC. Conforme pode ser visto, não há diferenças significativas entre as soluções obtidas via MH e HMC, pois no problema tratado a distribuição *a posteriori* é relativamente simples de se amostrar.

Tabela 4 – Parâmetros estimados para as propriedades termofísicas.

| Parâmetro                | MH                     |                      | HMC                    |                      | Referência           |
|--------------------------|------------------------|----------------------|------------------------|----------------------|----------------------|
|                          | $\bar{\mu}$            | $\sigma_z$           | $\bar{\mu}$            | $\sigma_z$           |                      |
| $w$ [J/m <sup>3</sup> K] | $1,4919 \times 10^6$   | $7,7 \times 10^3$    | $1,4862 \times 10^6$   | $7,3 \times 10^3$    | $1,49 \times 10^6$   |
| $k$ [W/mK]               | $1,525 \times 10^{-1}$ | $8,4 \times 10^{-4}$ | $1,512 \times 10^{-1}$ | $8,0 \times 10^{-4}$ | $1,6 \times 10^{-1}$ |

## 5. CONCLUSÕES

Foi visto que o método HMC possibilita obter uma distribuição *a posteriori* com maior qualidade quando comparado com o MH. Contudo, para uma distribuição de probabilidade de baixa ordem, como o caso visto neste trabalho, chega-se à conclusão de que não há diferença significativa entre estes métodos. Segundo Hanson e Cunningham (1998), a eficiência do MH diminui com o aumento do número de variáveis do problema, enquanto a eficiência do HMC se mantém quase que constante. Portanto, a sugestão para trabalhos futuros é avaliar esta metodologia em meios heterogêneos, *i.e.*, as propriedades termofísicas variem com a posição.

### *Agradecimentos*

Os autores agradecem às agências de fomento FAPERJ, CNPq e CAPES.

## REFERENCIAS

- BECK, V; ARNOLD, K. J. **Parameter estimation in engineering and science**. James Beck, 1977.
- COTTA, R. M.. Integral transforms in computational heat and fluid flow. CRC Press, 1993.
- DUANE, S. et al. Hybrid Monte Carlo. Physics letters B, v. 195, n. 2, p. 216-222, 1987.
- HANSON, K. M.; CUNNINGHAM, G. S. Posterior sampling with improved efficiency. In: **Medical Imaging 1998: Image Processing**. International Society for Optics and Photonics, 1998. p. 371-383.
- HANSON, K. M. Markov Chain Monte Carlo posterior sampling with the Hamiltonian method. In: **Medical Imaging 2001: Image Processing**. International Society for Optics and Photonics, 2001. p. 456-468.
- KAPIO, J. P.; SOMERSALO, E. Computational and statistical methods for inverse problems. **Applied mathematical sciences**, v. 160, 2004.
- KNUPP, D. C.; NETO, A. J. S. Solution of the inverse radiative transfer problem of simultaneous identification of the optical thickness and space-dependent Albedo using Bayesian inference. **CMES: Computer Modeling in Engineering & Sciences**, v. 96, n. 5, p. 339-360, 2013.
- METROPOLIS, N. et al. Equation of state calculations by fast computing machines. The journal of chemical physics, v. 21, n. 6, p. 1087-1092, 1953.
- MIKHAILOV, M. D.; OZISIK, M. N. Unified analysis and solutions of heat and mass diffusion. 1984.
- MARK, J. E. et al. (Ed.). **Physical properties of polymers handbook**. New York: Springer, 2007.
- NEAL, R. M. et al. MCMC using Hamiltonian dynamics. Handbook of MCMC, v. 2, n. 11, 2011.

## COMPARISON BETWEEN THE METROPOLIS-HASTINGS AND HAMILTONIAN MONTE CARLO METHODS FOR THE THERMOPHYSICAL PROPERTIES IDENTIFICATION IN HEAT CONDUCTION PROBLEMS

**Abstract.** *In this work, the heat conduction inverse problem is formulated and resolved with the aim of the thermophysical properties identification, i.e., the volumetric heat capacity and heat conductivity of a polymer material via Bayesian inference. For the solution of the Bayes theorem it is considered two Markov Chain Monte carlo methods: the Metropolis-Hastings method and the Hamiltonian Monte Carlo method (HMC). The same mathematical model is used for the surrogate data creation and for the forward problem solution. Nonetheless, different resolution techniques are employed in order to mitigate the inverse crime. For this purpose, the integral classical transform technique (CITT) is employed for the surrogate temperature generation whereas the finite difference method (FDM) is used in the direct problem solution. In view of the findings, it is possible to conclude that the HMC is more efficient and consistent when compared to MH. The parameters Markov Chains were able to achieve less correlated samples with a smaller burn-in.*

**Keywords:** Heat conduction, Bayesian inference, Hamiltonian Monte Carlo (HMC)