



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

APLICAÇÃO DO MÉTODO DE MONTE CARLO HAMILTONIANO NA IDENTIFICAÇÃO DE PROPRIEDADES TÉRMICAS COM UM MODELO DE CAPACITÂNCIA GLOBAL VIA INFERÊNCIA BAYESIANA

Maicon de Paiva Torres - mptorres@iprj.uerj.br

Géssica Ramos da Silva - grsilva@iprj.uerj.br

Diego Campos Knupp - diegoknupp@iprj.uerj.br

Leonardo Tavares Stutz - ltstutz@iprj.uerj.br

Antônio J. Silva Neto - ajsneto@iprj.uerj.br

Universidade do Estado do Rio de Janeiro, Instituto Politécnico - Nova Friburgo, RJ, Brasil

Resumo. Neste trabalho, o método de Monte Carlo Hamiltoniano (HMC) é utilizado como ferramenta de amostragem da distribuição de probabilidade a posteriori, obtida via inferência Bayesiana, para a solução do problema inverso de determinação das propriedades de um problema de transferência de calor com capacitância global. O efeito do nível de ruído, da estimativa inicial dos parâmetros de interesse e do número de iterações necessárias para se atingir a distribuição de probabilidade estacionária são analisados e os resultados são comparados com aqueles obtidos via Metropolis-Hastings (MH). É visto que, com o HMC, é possível trabalhar com estimativas iniciais mais distantes dos valores alvo e a amostragem da distribuição a posteriori possui qualidade superior, devido à alta taxa de aceitação obtida. Além disso, conclui-se que é necessário um número menor de estados de aquecimento para se atingir a distribuição de equilíbrio das cadeias de Markov.

Palavras-chave: Capacitância global, Inferência Bayesiana, Metropolis-Hastings, Método de Monte Carlo Hamiltoniano

1. INTRODUÇÃO

A inferência Bayesiana tornou-se uma abordagem robusta e de papel inquestionável, nas últimas décadas, para a análise de dados em problemas inversos (Bartholomew e Martz, 1985; Wang e Beck, 1989; Wang e Zabarar, 2004; Naveira-Cotta et al., 2010; Choi et al., 2018). Tendo em mãos um modelo probabilístico para os dados observados, por meio da inferência Bayesiana é possível quantificar adequadamente as incertezas e ainda obter a estrutura global do espaço dos parâmetros.

Na maioria dos problemas de engenharia, a formulação decorrente do teorema de Bayes não pode ser resolvida analiticamente, necessitando para isso de eficientes técnicas de integração.

Estes métodos tendem a ser custosos do ponto de vista computacional e, portanto, são usados algoritmos de Monte Carlo com Cadeias de Markov (MCMC) para simular amostras da distribuição de probabilidade alvo.

O algoritmo de Metropolis-Hastings (Metropolis et al., 1953; Hastings, 1970), devido à sua simplicidade de implementação e ao baixo custo computacional, é o método MCMC mais utilizado até os tempos atuais. Contudo, por ser um método de caminhada aleatória, este é pouco eficiente na exploração do espaço dos parâmetros (Zhang et al., 2017). Neste trabalho, um método MCMC sem caminhada aleatória, conhecido como método de Monte Carlo Hamiltoniano (HMC), é estudado no problema inverso de identificação de propriedades de um modelo de capacitância global. O método HMC atenua o comportamento aleatório do movimento, conseguindo uma exploração mais efetiva e consistente na obtenção da distribuição de probabilidade *a posteriori* (Wang et al., 2017). Neste contexto, passa a ser possível tratar distribuições complexas e com um grande número de variáveis. Desta forma, são obtidas amostras com tamanho efetivo grande e a convergência para a distribuição alvo é acelerada.

2. FORMULAÇÃO E SOLUÇÃO DO PROBLEMA DIRETO

Neste trabalho, um problema de resfriamento é considerado, sendo este representado por um modelo de capacitância global, i.e., os gradientes de temperatura são desprezados. Neste sistema, o vetor de medidas experimentais sintéticas depende apenas do histórico da temperatura em função do tempo. Para obter o histórico de temperatura são considerados a constante de tempo, τ_c , e a temperatura inicial, T_i , do sistema. O modelo matemático que descreve este fenômeno físico de transferência de calor é escrito por

$$C_p \frac{dT(t)}{dt} = -\frac{h}{L}[T(t) - T_\infty] \quad (1a)$$

$$T(0) = T_i \quad (1b)$$

onde C_p é a capacidade térmica volumétrica do sistema, h é o coeficiente de troca térmica convectivo do sistema com o ambiente na temperatura T_∞ , que é definida em 25 °C, e L é a espessura da placa. Este sistema possui solução analítica e sua expressão é dada por

$$T(t) = T_\infty - (T_\infty - T_i) \exp\left(\frac{-t}{\tau_c}\right) \quad (2)$$

onde τ_c é a constante de tempo do sistema, dada por $\tau_c = LC_p/h$.

Neste trabalho é considerado que os parâmetros T_i e τ_c não são conhecidos, mas dados experimentais de temperatura estão disponíveis. É empregada então a abordagem descrita na próxima seção, para a determinação dos parâmetros de interesse.

3. FORMULAÇÃO E SOLUÇÃO DO PROBLEMA INVERSO

Na abordagem Bayesiana, o problema inverso é formulado como um problema de inferência estatística e é baseado nos seguintes princípios (Kaipio e Sommersalo, 2006): (i) os parâmetros

de interesse do modelo são modeladas como variáveis aleatórias; (ii) a aleatoriedade descreve nosso grau de informação; (iii) o grau de informação é codificado em distribuições de probabilidade; e (iv) a solução do problema inverso é a distribuição de probabilidade a posteriori. Assim, na abordagem Bayesiana toda informação possível é incorporada ao modelo, visando reduzir o grau de incerteza presente no problema.

Considere que alguma informação *a priori* a respeito dos parâmetros $\mathbf{P} = [T_i, \tau_c]$ possa estar disponível. Assumindo que esta informação pode ser modelada como uma densidade de probabilidade $\pi_{pr}(\mathbf{P})$, o teorema de Bayes para problemas inversos pode ser expresso por (Kaipio e Sommersalo, 2006)

$$\pi_{post}(\mathbf{P}) = \pi(\mathbf{P}|\mathbf{Y}) = \frac{\pi_{pr}(\mathbf{P})\pi(\mathbf{Y}|\mathbf{P})}{\pi(\mathbf{Y})} \quad (3)$$

onde $\pi_{post}(\mathbf{P})$ é a densidade de probabilidade a posteriori, $\pi_{pr}(\mathbf{P})$ é a informação *a priori* dos parâmetros desconhecidos, modelada como uma distribuição de probabilidade, $\pi(\mathbf{Y}|\mathbf{P})$ é a função de verossimilhança e $\pi(\mathbf{Y})$ é a densidade marginal, que atua como uma constante de normalização. É importante destacar que o método estatístico produz uma distribuição que pode ser explorada de diferentes formas, usando diferentes métodos.

Considere que a informação *a priori* para os parâmetros pode ser modelada como uma distribuição normal. Logo, $\pi_{pr}(\mathbf{P})$ pode ser expressa por

$$\pi_{pr}(\mathbf{P}) = (2\pi)^{-N_p/2} |\mathbf{V}|^{1/2} \exp \left[-\frac{1}{2} (\mathbf{P} - \mu^T) \mathbf{V}^{-1} (\mathbf{P} - \mu) \right] \quad (4)$$

onde N_p é o número de parâmetros, \mathbf{V} e μ são, respectivamente, a matriz de covariância e a média para \mathbf{P} . Substituindo a Eq. (4) na Eq. (3) e usando a função de verossimilhança, obtém-se

$$\ln[\pi_{post}(\mathbf{P}|\mathbf{Y})] \propto -\frac{1}{2} [(N_p + N_d) \ln(2\pi) + \ln|\mathbf{W}^{-1}| + \ln|\mathbf{V}^{-1}| + S_{MAP}(\mathbf{P})] \quad (5)$$

onde \mathbf{W} é a matriz de covariância dos erros experimentais, que supostamente seguem uma distribuição normal, e

$$S_{MAP}(\mathbf{P}) = [\mathbf{Y} - \theta(\mathbf{P})]^T \mathbf{W}^{-1} [\mathbf{Y} - \theta(\mathbf{P})] + [\mu - \mathbf{P}]^T \mathbf{V}^{-1} [\mu - \mathbf{P}] \quad (6)$$

é a função objetivo do *maximum a posteriori* (MAP). A minimização de $S_{MAP}(\mathbf{P})$ produz as estimativas de \mathbf{P} que maximizam a distribuição a posteriori $\pi_{post}(\mathbf{P}|\mathbf{Y})$. Neste trabalho, a função objetivo do MAP foi minimizada com o algoritmo estocástico da Evolução Diferencial (Storn e Price, 1997).

3.1 Métodos de Monte Carlo com Cadeias de Markov

O método de Monte Carlo com Cadeias de Markov (MCMC) foi inicialmente proposto por Metropolis et al. (1953), onde este foi empregado para simular estados de distribuição de moléculas idealizadas. Na inferência Bayesiana, o uso do método MCMC tem como objetivo calcular a densidade de probabilidade *a posteriori* frente a um conjunto de variáveis

aleatórias, quando a avaliação de integrais é intratável do ponto de vista analítico. Nestes casos, a solução do problema inverso é obtida considerando técnicas de amostragem baseadas no método MCMC. Neste trabalho, também foram utilizados o algoritmo de Metropolis-Hastings (MH) e o método de Monte Carlo Hamiltoniano (HMC) na solução do problema inverso.

O algoritmo de Metropolis-Hastings. 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 $\mathbf{Z}^* = \mathbf{Z}^{i-1} + \Delta\mathbf{Z}$, onde $\Delta\mathbf{Z}$ 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 θ 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 $\{\mathbf{Z}^1, \mathbf{Z}^2, \dots, \mathbf{Z}^n\}$.

Método de Monte Carlo Hamiltoniano. O método de Monte Carlo Hamiltoniano, ou método de Monte Carlo Híbrido, como é alternativamente conhecido, combina a amostragem de Gibbs com a função de aceitação do algoritmo de Metropolis-Hastings (Afshar e Sheehan, 2018). O método HMC possui um caráter determinístico inspirado na dinâmica Hamiltoniana para propor amostras seguindo a distribuição de probabilidade alvo. Deste modo, o comportamento referente à caminhada aleatória é atenuado para permitir a exploração mais efetiva e consistente do espaço de probabilidade, quando comparado com as técnicas de Gibbs e de Metropolis-Hastings.

O primeiro passo para construir o método MCMC com a dinâmica Hamiltoniana é definir uma função Hamiltoniana em termos da distribuição de probabilidade alvo que se deseja amostrar. Além das variáveis de interesse (variáveis de "posição"), é preciso introduzir variáveis auxiliares de "momento", que tipicamente possuem distribuições independentes e Gaussianas.

No método HMC, para cada parâmetro Z_i é introduzida uma variável de momento associada, p_i . O Hamiltoniano H é então construído considerando a soma de um termo de energia potencial com um termo de energia cinética, na forma

$$H(\mathbf{Z}, \mathbf{p}) = U(\mathbf{Z}) + K(\mathbf{p}) \quad (7)$$

onde $U(\mathbf{Z})$ é o negativo do logaritmo da distribuição de probabilidade *a posteriori*, dado por $U(\mathbf{Z}) = -\log(\pi_{post})$, e $K(\mathbf{p})$ é a energia cinética, definida por

$$K(\mathbf{p}) = \frac{\mathbf{p}^T \bar{M}^{-1} \mathbf{p}}{2} \quad (8)$$

onde \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

$$\frac{dZ_i}{dt} = \frac{\partial H}{\partial p_i}; \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial Z_i}, \quad i = 1, 2, \dots, N \quad (9a,b)$$

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}) é aproximada utilizando-se a técnica de integração leapfrog (Anderson et al., 1984), que depende do número de passos, L , e do tamanho do passo, ϵ . Para $L = 1$ tem-se as seguintes etapas para a integração das equações de movimento

$$p_i \left(t + \frac{\epsilon}{2} \right) = p_i(t) - \frac{\epsilon}{2} \frac{\partial U}{\partial Z_i}(\mathbf{Z}(t)) \quad (10a)$$

$$Z_i(t + \epsilon) = Z_i(t) + \epsilon \frac{p_i \left(t + \frac{\epsilon}{2} \right)}{\bar{m}_i} \quad (10b)$$

$$p_i(t + \epsilon) = p_i \left(t + \frac{\epsilon}{2} \right) - \frac{\epsilon}{2} \frac{\partial U}{\partial Z_i}(\mathbf{Z}(t + \epsilon)) \quad (10c)$$

onde o gradiente da energia potencial é obtido, neste trabalho, por meio de diferenças finitas.

4. RESULTADOS E DISCUSSÕES

Na maioria dos resultados apresentados, é feita a comparação entre os métodos HMC e MH, considerando as seguintes vertentes: efeito do nível de ruído, efeito da estimativa inicial e o número de iterações necessárias para se atingir a distribuição de probabilidade estacionária. A distribuição estacionária, neste caso, se refere à distribuição de probabilidade alvo, i.e., a distribuição *a posteriori* da qual é feita as análises estatísticas para obter as estimativas para as variáveis a serem identificadas. Como são dois os parâmetros a serem determinados (temperatura inicial, T_i , e constante de tempo do sistema, τ_c), as distribuições de probabilidade consideradas neste trabalho são distribuições bivariadas.

Os dados sintéticos foram gerados considerando a solução do problema direto com a temperatura inicial $T_i = 101,7 \text{ }^\circ\text{C}$ e a constante de tempo $\tau_c = 100 \text{ s}$. A essa solução soma-se um sinal ruidoso gerado de uma distribuição normal com média zero e desvio padrão σ . Os parâmetros T_i (temperatura inicial) e τ_c (constante de tempo do sistema) possuem informação *a*

priori, seguindo uma distribuição Gaussiana, com médias 90 °C e 80 s e desvios padrões iguais a 10 °C e 20 s, respectivamente.

Os dados adotados para o leapfrog na solução do problema inverso pelo método HMC são $L = 2$ e $\epsilon = 0,1$. A matriz de massa fictícia do sistema Hamiltoniano é $\bar{M} = I$, onde I é a matriz identidade. Já para o MH, na distribuição de probabilidade auxiliar é considerado um desvio padrão para a geração das posições candidatas de 0,3% dos valores das estimativas iniciais.

A Tabela 1 mostra o efeito do nível de ruído dos dados experimentais sintéticos na determinação das propriedades do sistema. Conforme o nível do ruído aumenta, percebe-se o aumento do desvio padrão das estimativas obtidas via MH e HMC. As médias calculadas possuem valores próximos aos valores exatos dos parâmetros e seu comportamento, em relação ao nível de ruído, é similar ao comportamento da solução obtida via MAP.

Tabela 1 – Efeito do ruído na estimação dos parâmetros do sistema térmico. Os valores exatos dos parâmetros são $T_i = 101,7$ °C e $\tau_c = 100$ s.

Ruído (%)	MH				HMC				MAP	
	T_i		τ_c		T_i		τ_c		T_i (°C)	τ_c (s)
	Média (°C)	Desv. Padrão (°C)	Média (s)	Desv. Padrão (s)	Média (°C)	Desv. Padrão (°C)	Média (s)	Desv. Padrão (s)		
1	101,73	0,02	100,08	0,07	101,64	0,05	100,03	0,12	101,73	99,97
2	101,76	0,07	99,93	0,18	101,78	0,16	99,77	0,19	101,63	100,08
3	101,93	0,13	99,64	0,32	101,88	0,13	99,55	0,31	101,66	100,10
4	101,61	0,15	100,15	0,33	101,76	0,17	99,80	0,38	101,59	100,35
5	101,75	0,21	100,55	0,54	101,43	0,20	100,22	0,51	101,77	99,75
10	102,43	0,38	98,33	0,85	101,20	0,40	101,76	0,95	101,41	100,74

Pela Tabela 2 é possível observar o efeito da estimativa inicial da temperatura inicial (T_i) na estimação dos parâmetros, enquanto que na Tabela 3 é mostrado o efeito da estimativa inicial da constante de tempo (τ_c). Das tabelas é visto que o algoritmo de MH não foi capaz de lidar com as estimativas iniciais de 20 e 40 °C para a temperatura inicial e de 20 e 40 s para a constante de tempo. Já o HMC não apresentou problemas para estimar as propriedades com estimativas iniciais mais distantes dos valores exatos.

Nas Figuras 1(a, b) estão as cadeias de Markov dos parâmetros T_i e τ_c , obtidas via HMC e MH. Foram considerados 1000 estados. Para o MH atingir a distribuição estacionária, são necessários cerca de 300 estados, enquanto que o HMC necessitou de apenas 100 estados. Mesmo com $L = 2$ fica evidente como a convergência do HMC é mais acelerada que a do MH.

Nas Figuras 2(a, b) estão os gráficos de dispersão e os histogramas para os parâmetros T_i e τ_c , obtidos via MH e HMC. 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 37% dos candidatos, enquanto que o HMC teve uma taxa de aceitação de 94,8%.

Tabela 2 – Efeito da estimativa inicial da temperatura inicial na estimação dos parâmetros do sistema térmico. Os valores exatos dos parâmetros são $T_i = 101,7 \text{ }^\circ\text{C}$ e $\tau_c = 100 \text{ s}$.

		MH				HMC			
		T_i		τ_c		T_i		τ_c	
Estim. Inicial T_i ($^\circ\text{C}$)	Estim. Inicial τ_c (s)	Média ($^\circ\text{C}$)	Desv. Padrão ($^\circ\text{C}$)	Média (s)	Desv. Padrão (s)	Média ($^\circ\text{C}$)	Desv. Padrão ($^\circ\text{C}$)	Média (s)	Desv. Padrão (s)
20	80	Incapaz de resolver				101,87	0,23	99,21	0,53
40	80	Incapaz de resolver				101,79	0,20	99,85	0,46
60	80	101,91	0,23	99,09	0,57	101,70	0,22	100,08	0,56
80	80	101,75	0,21	100,55	0,54	101,43	0,20	100,22	0,51
100	80	101,69	0,20	99,61	0,50	101,80	0,20	100,18	0,48
120	80	101,64	0,20	100,16	0,50	101,55	0,22	99,73	0,54
140	80	101,71	0,22	99,80	0,43	101,57	0,22	99,60	0,54

Tabela 3 – Efeito da estimativa inicial da constante de tempo na estimação dos parâmetros do sistema térmico. Os valores exatos dos parâmetros são $T_i = 101,7 \text{ }^\circ\text{C}$ e $\tau_c = 100 \text{ s}$.

		MH				HMC			
		T_i		τ_c		T_i		τ_c	
Estim. Inicial T_i ($^\circ\text{C}$)	Estim. Inicial τ_c (s)	Média ($^\circ\text{C}$)	Desv. Padrão ($^\circ\text{C}$)	Média (s)	Desv. Padrão (s)	Média ($^\circ\text{C}$)	Desv. Padrão ($^\circ\text{C}$)	Média (s)	Desv. Padrão (s)
80	20	Incapaz de resolver				101,77	0,22	99,67	0,55
80	40	Incapaz de resolver				101,82	0,22	99,68	0,49
80	60	101,51	0,37	100,78	0,42	101,85	0,20	99,51	0,48
80	80	101,75	0,21	100,55	0,54	101,43	0,20	100,22	0,51
80	100	101,71	0,23	99,64	0,52	101,54	0,19	100,11	0,45
80	120	101,42	0,21	100,45	0,49	101,66	0,23	100,61	0,56
80	160	101,70	0,21	100,40	0,49	101,56	0,21	99,85	0,51

No método HMC, a cada estado das cadeias de Markov é realizada a amostragem de Gibbs para as variáveis de momento e, considerando a evolução da dinâmica Hamiltoniana, uma nova posição candidata é gerada, para finalmente ser testada. O espaço de fases do Hamiltoniano, considerado neste trabalho, possui quatro dimensões, sendo duas dimensões para as variáveis de posição e duas dimensões para as variáveis de momento. Contudo, pode-se analisar os estados em gráficos separados, de posição x momento, conforme apresentado na Fig. 3.

Observa-se uma acelerada convergência, desde a estimativa inicial até a distribuição alvo. Nota-se também que, conforme o sistema se aproxima da solução, a distância entre uma posição e a imediatamente anterior diminui. Conforme a distribuição se aproxima da distribuição alvo, há uma diminuição da energia potencial, o que acarreta no decremento da energia do sistema Hamiltoniano. Isso leva à distribuição estacionária da cadeia.

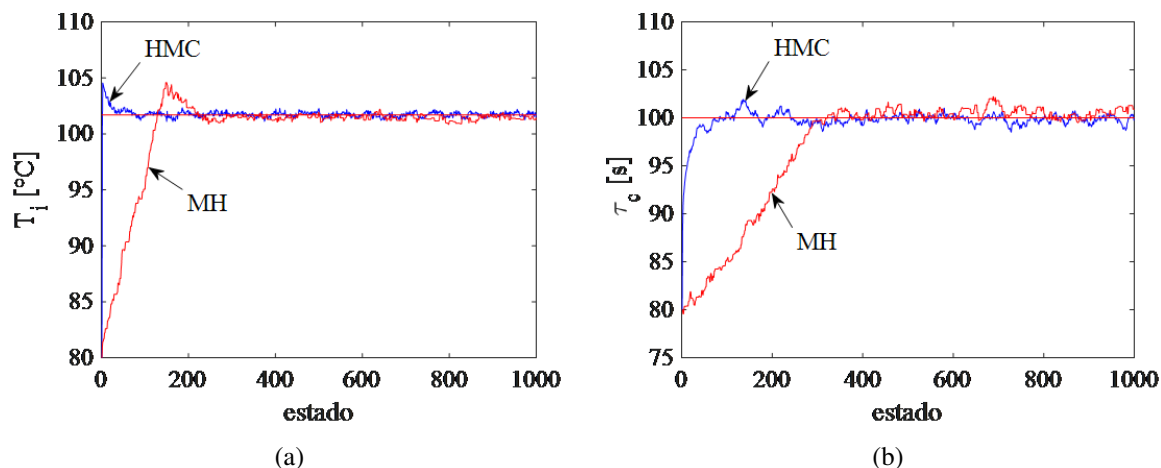


Figura 1- Evolução (a) da temperatura inicial e (b) da constante de tempo obtidas via HMC e MH. A estimativa inicial considerada é $T_i = 80^\circ C$ e $\tau_c = 80s$. Para o HMC foram considerados $L = 2$ e $\epsilon = 0, 1$.

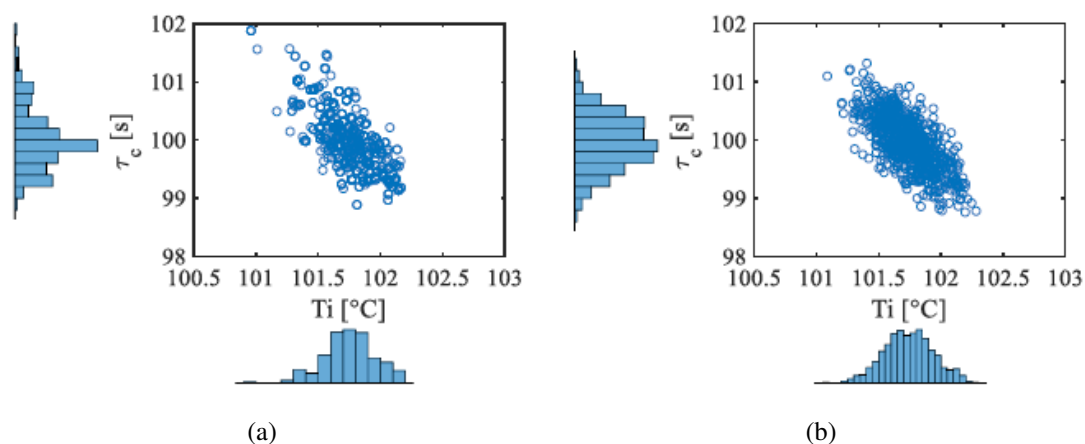


Figura 2- Dispersão e histogramas para os parâmetros identificados via (a) HMC e (b) MH. A estimativa inicial considerada é $T_i = 80^\circ C$ e $\tau_c = 80s$. Para o HMC foram considerados $L = 2$ e $\epsilon = 0, 1$.

5. CONCLUSÕES

No presente trabalho, o método de Monte Carlo Hamiltoniano (HMC) foi utilizado na identificação de propriedades de um problema de resfriamento representado com um modelo de capacitância global. O efeito do nível de ruído, das estimativas iniciais dos parâmetros de interesse (T_i , temperatura inicial, e τ_c , constante de tempo do sistema) e do número de iterações necessárias para se atingir a distribuição de probabilidade estacionária foram analisados e comparados com os resultados obtidos pelo algoritmo de Metropolis-Hastings (MH).

Na análise do efeito do nível de ruído na obtenção da solução do problema inverso, observou-se um aumento do desvio padrão das estimativas obtidas com o nível de ruído para ambos os

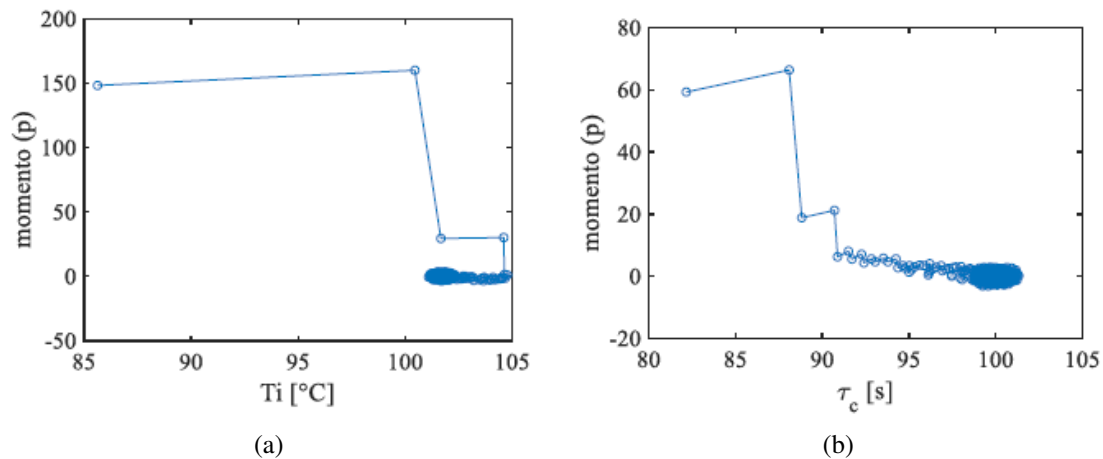


Figura 3- Evolução do espaço de fases Hamiltoniano para (a) a temperatura inicial e (b) a constante de tempo, obtidas via HMC. A estimativa inicial considerada é $T_i = 80^\circ C$ e $\tau_c = 80s$. Foram considerados $L = 2$ e $\epsilon = 0, 1$.

métodos (HMC e MH). Já em relação ao efeito da estimativa inicial, o MH apresentou problemas (não conseguiu identificar) para estimar propriedades com estimativas iniciais distantes dos valores exatos, diferentemente do HMC. Por último, observou-se que o HMC possui uma convergência bem mais acelerada que a do MH (foram necessários apenas 100 estados para se atingir a distribuição de probabilidade estacionária, contra 300 do MH).

Agradecimentos

Os autores agradecem o apoio financeiro concedido pelas agências de fomento CAPES, CNPq e FAPERJ.

REFERÊNCIAS

- AFSHAR, S.; SHEEHAN, M.. CFD and infrared thermography of particle curtains undergoing convection heat transfer: Image analysis and edge prediction. *Powder Technology*, v. 325, p. 167-179, 2018.
- ANDERSON, D. A.; TANNEHILL, J. C.; PLETCHER, R. H. *Computational fluid dynamics and heat transfer*. Washington: Hemisphere, 1984.
- BARTHOLOMEW, R. J.; MARTZ, H. F. *Bayesian analysis of heat pipe life test data for reliability demonstration testing*. Los Alamos National Lab., NM (USA), 1985.
- CHOI, W. et al. Bayesian inference for thermal response test parameter estimation and uncertainty assessment. *Applied Energy*, v. 209, p. 306-321, 2018.
- HASTINGS, W. K. *Monte Carlo sampling methods using Markov chains and their applications*. 1970.
- KAIPIO, J.; SOMERSALO, E. *Statistical and computational inverse problems*. Springer Science & Business Media, 2006.
- 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.
- NAVEIRA-COTTA, C. P.; ORLANDE, H. R. B.; COTTA, R. M. Integral transforms and Bayesian inference in the identification of variable thermal conductivity in two-phase dispersed systems. *Numerical Heat Transfer, Part B: Fundamentals*, v. 57, n. 3, p. 173-202, 2010.
- ORLANDE, H. R. B. Inverse problems in heat transfer: new trends on solution methodologies and applications. *Journal of Heat Transfer*, v. 134, n. 3, p. 031011, 2012.
- STORN, R.; PRICE, K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, v. 11, n. 4, p. 341-359, 1997.

- WANG, J.; ZABARAS, N. A Bayesian inference approach to the inverse heat conduction problem. *International Journal of Heat and Mass Transfer*, v. 47, n. 17-18, p. 3927-3941, 2004.
- WANG, K.; BECK, A. E. An inverse approach to heat flow study in hydrologically active areas. *Geophysical Journal International*, v. 98, n. 1, p. 69-84, 1989.
- WANG, Z.; BROCCARDO, M.; SONG, J. Hamiltonian Monte Carlo Methods for Subset Simulation in Reliability Analysis. arXiv preprint arXiv:1706.01435, 2017.
- ZHANG, C.; SHAHBABA, B.; ZHAO, H. Precomputing strategy for Hamiltonian Monte Carlo method based on regularity in parameter space. *Computational Statistics*, v. 32, n. 1, p. 253-279, 2017.

APPLICATION OF THE HAMILTONIAN MONTE CARLO METHOD IN THE THERMAL PROPERTIES IDENTIFICATION WITH A GLOBAL CAPACITANCE MODEL SYSTEM VIA BAYESIAN INFERENCE

Abstract. *In this work, the Hamiltonian Monte Carlo method (HMC) is used as sampling tool for the posterior probability distribution, obtained through Bayesian inference, for the solution of the inverse problem of determining the properties of a heat transfer problem with global capacitance. The effect of the noise level, the initial estimate of the parameters of interest and the number of iterations necessary to reach the steady-state probability distribution are analyzed and the results are compared with those obtained through Metropolis-Hastings (MH). It is seen that, with the HMC, it is possible to work with initial estimates that are more distant from the target values and the sampling of the posterior distribution has superior quality due to the high acceptance rate obtained. In addition, it is concluded that a lower number of heating states is required to achieve the equilibrium distribution of the Markov chains.*

Keywords: *Global capacitance, Bayesian inference, Metropolis-Hastings, Hamiltonian Monte Carlo method*