

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Calibração de um Modelo de Crescimento Tumoral Avascular

Anna Claudia M. Resende¹

Heber L. Rocha²

Regina C. Almeida³

Renato S. Silva⁴

Laboratório Nacional de Computação Científica (LNCC), Petrópolis/RJ, Brazil

Ernesto A. B. F. Lima⁵

Institute for Computational Engineering and Sciences (ICES), Austin/TX, USA

Resumo. Neste trabalho utilizamos um modelo matemático simples para representar o crescimento de tumores avasculares. Através de dados da literatura realizamos a calibração do modelo proposto tendo como base a abordagem Bayesiana. Nesta abordagem, as distribuições *a posteriori* dos parâmetros são obtidas utilizando duas técnicas: amostragem via algoritmo de Metropolis-Hastings, um método de Monte Carlo via cadeias de Markov, e amostragem via grade fixa. Ambas conduziram a resultados satisfatórios para os experimentos realizados e são por natureza mais informativas que os métodos tradicionais de regressão.

Palavras-chave. Calibração, Modelagem Matemática, Crescimento Tumoral.

1 Introdução

Muitos trabalhos vêm sendo desenvolvidos atualmente para explicar como ocorre o surgimento e a evolução tumoral, um fenômeno complexo, não linear, que envolve mecanismos regulatórios em diversas escalas espaço-temporais. Dentre os estudos realizados, alguns têm como enfoque o desafio da construção de modelos matemáticos preditivos, como em [3–5]. Para que um modelo seja preditivo, isto é, capaz de prever o fenômeno que se propõe a representar, é necessário considerar três etapas:

- 1. Verificação:** avaliar a diferença entre os resultados produzidos pelos modelos numérico e matemático. Pressupõe a verificação da solução, a qual envolve principalmente a estimativa do erro *a posteriori*. Nesta etapa, o erro devido a discretização é quantificado, estimado ou limitado de algum modo e eventualmente julgado aceitável;

¹annacmr@lncc.br

²heberlr@lncc.br

³rcca@lncc.br

⁴rssr@lncc.br

⁵lima@ices.utexas.edu

2. **Calibração:** ajustar as informações sobre os parâmetros (ou a falta delas) obtidas *a priori* de modo a tornar o modelo definido no cenário específico consistente com as observações disponíveis (dados), os quais usualmente vêm acompanhados de incertezas;
3. **Validação:** a meta do processo de validação é assegurar a validade da teoria (isto é, do modelo matemático). Os métodos de validação incluem, quando possível, comparações das previsões computacionais com observações, testes físicos e experimentos e devem levar em consideração a questão fundamental do entendimento dos impactos das incertezas no modelo matemático.

A calibração pode ser realizada através de técnicas como mínimos quadrados, abordagem Bayesiana, máxima verossimilhança, etc [8]. O processo de calibração através da abordagem Bayesiana é foco deste trabalho. Na inferência Bayesiana, as incertezas e/ou falta de confiança na teoria/modelo também são incluídas na análise através da probabilidade condicional de que o modelo irá prever as observações, dados os parâmetros (função de verossimilhança). Assim, o teorema de Bayes combina estes elementos de uma forma bastante natural e sistemática, permitindo inferir tanto a distribuição *a posteriori* dos parâmetros que melhor adequam o modelo teórico aos dados assim como seus valores mais prováveis [4].

2 Processo de Calibração

O crescimento tumoral está atrelado a eventos que ocorrem em diferentes escalas biológicas, espaciais e temporais. Sob essa perspectiva, para a obtenção de um modelo matemático preditivo, é crucial definir o foco de investigação e integrar dados experimentais para que o modelo possa ser calibrado e posteriormente validado. A confiabilidade da predição do modelo depende de como as incertezas são quantificadas e propagadas, sendo cada fase deste processo acompanhada de incertezas: percepção do modelo, aquisição de dados experimentais, calibração, etc. A abordagem Bayesiana é utilizada aqui para calibrar os parâmetros de um modelo de crescimento tumoral avascular sem restrição nutricional.

2.1 Modelo Matemático

Para ilustrar o processo de calibração utilizamos um modelo simples de crescimento exponencial para representar a evolução do volume tumoral (N). Este modelo descreve a taxa de variação do volume tumoral como contribuição do balanço entre proliferação e morte, descrita por uma taxa de crescimento constante. Esta dinâmica corresponde somente à fase de crescimento avascular, na qual o tumor invade tecidos adjacentes até que o fornecimento de nutrientes deixa de ser suficiente para manter seu ritmo desgovernado de crescimento. Sob as hipóteses descritas, o modelo proposto, definido em $(0, \tau_{\max}]$, onde τ_{\max} representa o tempo máximo de evolução, é dado por:

$$\begin{cases} \frac{dN}{dt} = rN; \\ N(t_0 = 0) = N_0; \end{cases} \implies \boxed{N = N_0 \exp(rt)}, \quad (1)$$

Solução Exata

em que $r > 0$ é a taxa de crescimento (dia^{-1}) e N_0 é o volume tumoral inicial (cm^3). Os dados publicados em [1] (Seção 2.2) são utilizados para calibrar os parâmetros r e N_0 .

2.2 Dados Experimentais

O trabalho apresentado em [1] tem como objetivo desenvolver e caracterizar um novo modelo animal para estudo do carcinoma de células escamosas (SCC), o mais comum câncer de cabeça e pescoço (HNSCC) [6]. Através da administração subcutânea de uma origem humana da linhagem celular de HNSCC (SCC-4) em ratos atímicos (que não possuem o timo, órgão de defesa do organismo), os autores avaliam, dentre outras propriedades, o crescimento do tumor. Os resultados obtidos em [1] são reproduzidos na Fig. 1, com um tumor mensurável a partir de 7 dias da inoculação de SCC-4. Os dados foram obtidos parcialmente do próprio artigo (Tabela 1, com média \pm desvio padrão) e complementados com extração digital utilizando a ferramenta online `WebPlotDigitizer 3.11` [7]. Na Fig. 1 indicamos também o ajuste via mínimos quadrados não-linear, denotado por `MQ-gnu`, realizado usando o programa `Gnuplot` através de uma implementação do algoritmo de Levenberg-Marquardt (via comando `fit`). Este ajuste `MQ-gnu` conduziu a $N_0 = 0,085758$ e $r = 0,172668$.

Tabela 1: Dados experimentais [1].

Dias após Inoculação	Volume Tumoral (cm^3)
10	0.40 ± 0.10
13	0.72 ± 0.16
15	1.01 ± 0.20
20	2.73 ± 0.33
24	5.19 ± 1.52

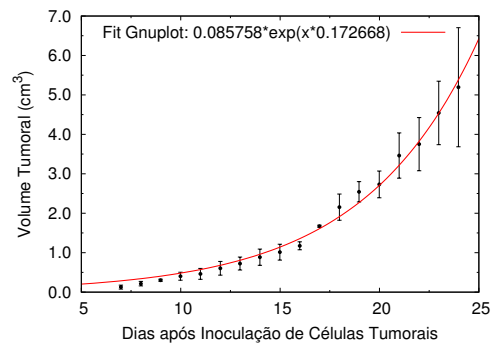


Figura 1: Evolução tumoral após administração subcutânea de SCC-4.

2.3 Abordagem Bayesiana para Calibração

Nosso objetivo com a calibração é ajustar os parâmetros r e N_0 de (1) para que o volume tumoral resultante do modelo proposto represente os resultados obtidos experimentalmente por [1]. Usando a nomenclatura introduzida em [8], definimos $\mathbf{m} = \{r, N_0\}$ como o conjunto dos parâmetros e \mathbf{d} como o conjunto dos dados. O Teorema de Bayes

estabelece que:

$$\pi_{\text{post}}(\mathbf{m}|\mathbf{d}) = \frac{\pi_{\text{ver}}(\mathbf{d}|\mathbf{m})\pi_{\text{prior}}(\mathbf{m})}{\pi_{\text{evid}}(\mathbf{d})}, \quad (2)$$

em que π_{prior} representa o conhecimento inicial (distribuição conjunta de probabilidade *a priori*) acerca dos parâmetros \mathbf{m} , π_{ver} é a função verossimilhança que permite atualizar o conhecimento de \mathbf{m} a partir do conjunto de dados \mathbf{d} e π_{evid} é a evidência (constante de normalização), que representa a informação contida nos dados \mathbf{d} [8]. O paradigma da calibração Bayesiana está descrito em detalhes em algumas referências como [4, 8] e consiste na atualização do conhecimento *a priori* π_{prior} a partir dos dados obtidos. A Eq. (2), portanto, é utilizada como um mecanismo para aprimorar o conhecimento sobre \mathbf{m} sempre que novas informações estiverem disponíveis.

Assumimos que os dados são independentes entre si e que as incertezas obtidas nos experimentos realizados por [1] seguem distribuições Gaussianas com desvios-padrão σ_i . Deste modo, a função de verossimilhança π_{ver} é dada por:

$$\pi_{\text{ver}}(\mathbf{d}|\mathbf{m}) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{d_c^i - d^i}{\sigma_i}\right)^2\right), \quad (3)$$

em que \mathbf{d}_c representa os dados para o procedimento de calibração [1] e n o tamanho do conjunto \mathbf{d} (resultado do problema direto). Para todos os testes consideramos um valor fixo $\sigma_i = \sigma = 0.319$, que equivale à média dos valores experimentais obtidos em [1]. As distribuições de probabilidade *a priori* foram definidas com base no ajuste MQ-gnu. Especificamente, definimos $\pi_{\text{prior}}(\mathbf{m}) = \pi_{\text{prior}}^r \pi_{\text{prior}}^{N_0}$ com $\pi_{\text{prior}}^r \sim \mathcal{U}(0.15, 0.19)$ e $\pi_{\text{prior}}^{N_0} \sim \mathcal{U}(0.07, 0.13)$. Assim, π_{post} é dada por:

$$\pi_{\text{post}}(\mathbf{m}|\mathbf{d}) \propto \prod_1^n \left(\frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{d_c^i - d^i}{\sigma}\right)^2\right] \right) \pi_{\text{prior}}(\mathbf{m}). \quad (4)$$

3 Calibração do Modelo

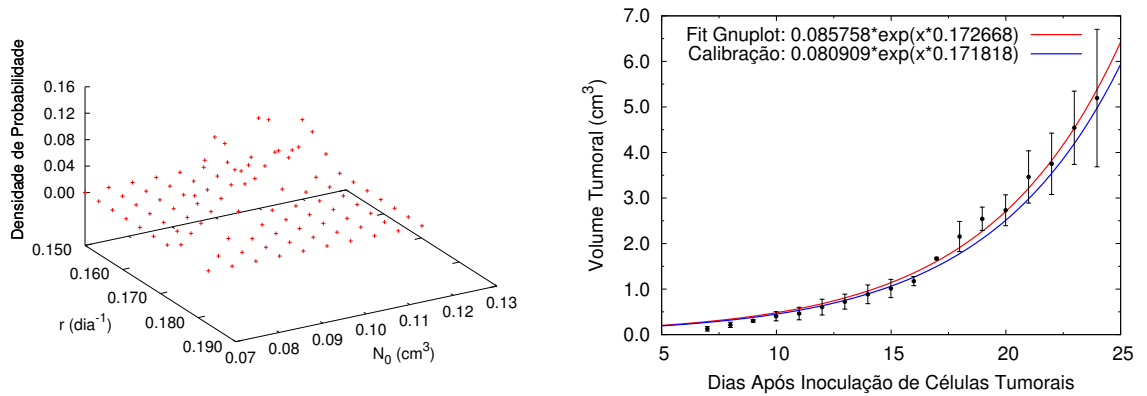
As distribuições de probabilidade *a posteriori* sobre os parâmetros são exploradas utilizando duas técnicas:

- i) **Amostragem via grade fixa:** o espaço paramétrico é particionado uniformemente em pequenos subintervalos;
- ii) **Amostragem via algoritmo de Metropolis-Hastings:** um método de Monte Carlo via cadeias de Markov (MCMC) [2].

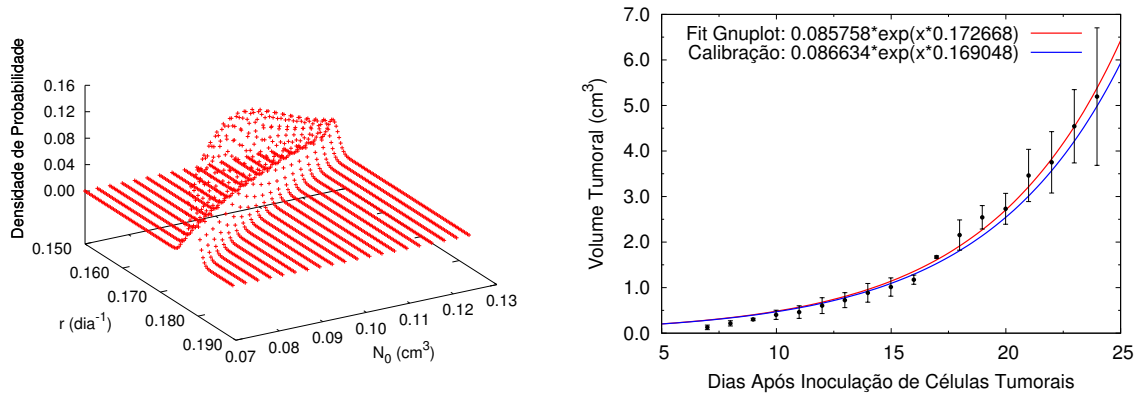
A amostragem via grade fixa é a mais simples pois só depende da partição do espaço paramétrico, a qual tem que ser definida *a priori*. O método de Metropolis-Hastings faz a varredura aleatória do espaço paramétrico a partir de um dado ponto inicial. O caminho percorrido é modificado com base em regras de aceitação e/ou rejeição, e tem como objetivo alcançar uma distribuição alvo [2]. Sua eficácia depende do número de

amostras utilizadas e sua natureza sequencial é uma de suas limitações, que pode tornar o algoritmo custoso pelo número de iterações necessárias para que este atinja convergência.

Na Fig. 2 apresentamos a distribuição conjunta *a posteriori* utilizando o método de grade fixa com duas malhas diferentes. Discretizamos o espaço paramétrico em $10(r) \times 10(N_0)$ e $100(r) \times 20(N_0)$ divisões, resultando nos valores mais prováveis $r = 0,171818$; $N_0 = 0,080909$ e $r = 0,169048$; $N_0 = 0,086634$. As evoluções do Modelo (1) com estes parâmetros são apresentadas nas Figs. 2(a) e 2(b). As diferenças são sutis quando comparados os valores mais prováveis obtidos, apesar do aumento do número de simulações (de 100 para 2000) e consequente aumento do custo computacional.



(a) Método de grade fixa, discretizando o espaço paramétrico com $10(r) \times 10(N_0)$ divisões. A distribuição conjunta *a posteriori* é apresentada à esquerda, em que $r = 0,171818$ e $N_0 = 0,080909$ são os valores mais prováveis. À direita, comparamos a evolução do volume tumoral obtida com o Modelo (1) e o ajuste MQ-gnu.



(b) Método de grade fixa, discretizando o espaço paramétrico com $100(r) \times 20(N_0)$ divisões. A distribuição conjunta *a posteriori* é apresentada à esquerda, em que $r = 0,169048$ e $N_0 = 0,086634$ são os valores mais prováveis. À direita, comparamos a evolução do volume tumoral obtida com o Modelo (1) e o ajuste MQ-gnu.

Figura 2: Método de grade fixa. Adotando $\sigma = 0,319$, calibramos o parâmetro r e a condição inicial N_0 do Modelo (1) com duas partições do espaço paramétrico.

Para a calibração com o método de Metropolis-Hastings utilizamos 1000 pontos amostrais. A convergência do método MCMC com relação aos valores mais prováveis de r e N_0 são apresentadas na Fig. 3, resultando em $r = 0,172422$ e $N_0 = 0,085504$. Após o

transiente inicial (região de *burn-in*), o método rapidamente captura a região de maior probabilidade. A convergência é alcançada mais rapidamente em N_0 (perto da iteração 100), enquanto que r converge perto da iteração 300. Na Fig. 4 apresentamos a distribuição conjunta *a posteriori* (sem a região de *burn-in*) e a evolução do volume tumoral adotando os valores mais prováveis.

A evolução do volume tumoral dada pelo Modelo (1) com os parâmetros ajustados por ambas as técnicas foram qualitativamente semelhantes ao ajuste MQ-gnu. As técnicas conduziram a resultados satisfatórios para os experimentos realizados e são por natureza mais informativas que os métodos tradicionais de regressão pois permitem obter as distribuições de probabilidade de cada parâmetro (marginais de $\pi_{\text{post}}(\mathbf{m}|\mathbf{d})$).

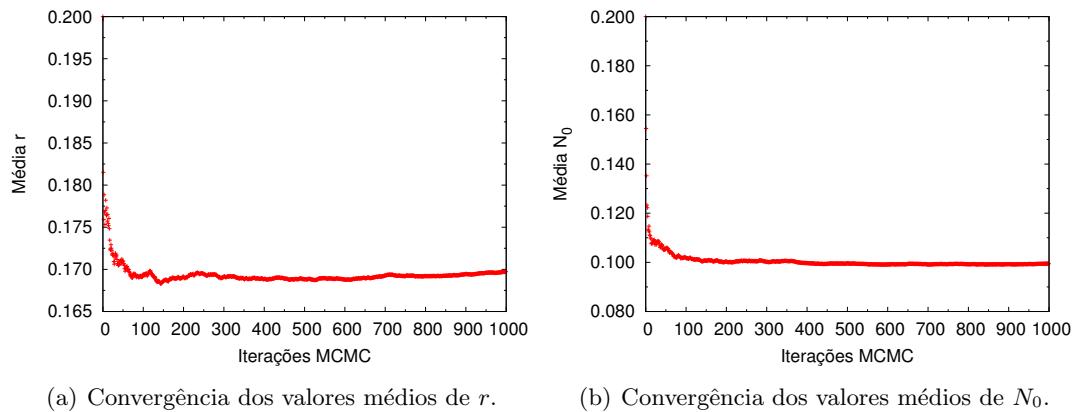


Figura 3: Método de Metropolis-Hastings, utilizando 1000 pontos amostrais e $\sigma = 0,319$. Convergência dos valores médios de r (a) e N_0 (b) com o método MCMC.

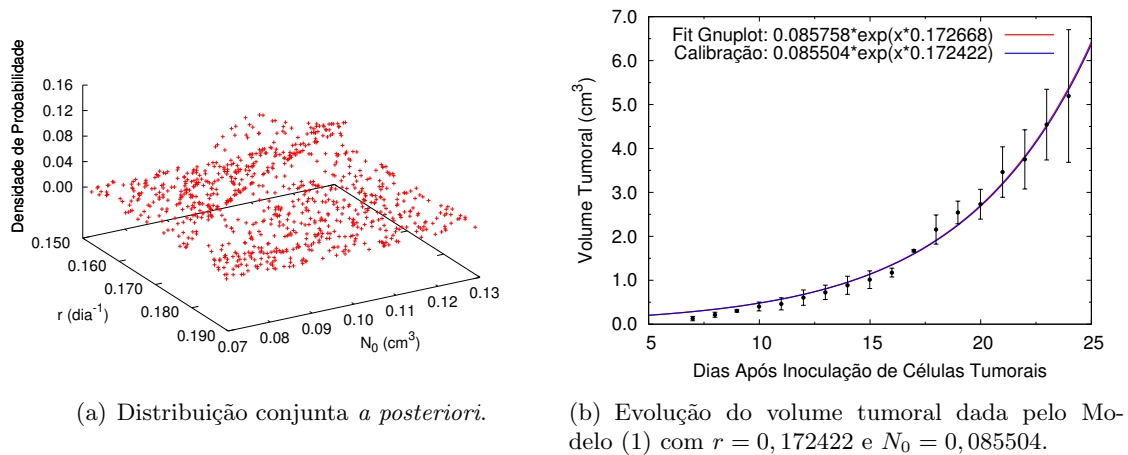


Figura 4: Método de Metropolis-Hastings, utilizando 1000 pontos amostrais e $\sigma = 0,319$. Distribuição conjunta *a posteriori* (a) em que $r = 0,172422$ e $N_0 = 0,085504$ são os valores mais prováveis e (b) evolução do volume tumoral obtida com o Modelo (1) com os parâmetros calibrados e o ajuste MQ-gnu.

4 Conclusão

O foco deste trabalho é a calibração de um modelo simples de crescimento tumoral avascular. A calibração foi feita com base na abordagem Bayesiana, utilizando duas técnicas para a obtenção das distribuições *a posteriori* sobre os parâmetros a serem calibrados: amostragem via algoritmo de Metropolis-Hastings e amostragem via grade fixa. Ambas as técnicas conduziram a resultados satisfatórios para os experimentos realizados e são por natureza mais informativas que os métodos tradicionais de regressão, pois além dos valores mais prováveis fornecem a descrição da incerteza associada.

Agradecimentos

ACMR e HLR agradecem o apoio concedido pelo CNPq e CAPES através dos projetos de número 140271/2016-3 e 1576953/2016-3.

Referências

- [1] A. Bao et al. Setup and characterization of a human head and neck squamous cell carcinoma xenograft model in nude rats, *Otolaryngology – Head and Neck Surgery*, 135(6):853–857, 2006. DOI: 10.1016/j.otohns.2006.06.1257.
- [2] W. R. Gilks, S. Richardson and D. Spiegelhalter. Markov Chain Monte Carlo in Practice. *Chapman & Hall/CRC Interdisciplinary Statistics Series*. Chapman & Hall, London, 1995. ISBN: 0412055511,9780412055515.
- [3] E. A. B. F. Lima et al. Selection, calibration, and validation of models of tumor growth. *Mathematical Models and Methods in Applied Sciences*, 26(12):2341–2368, 2016. DOI: 10.1142/S021820251650055X.
- [4] J. T. Oden, A. Hawkins and S. Prudhomme. General diffuse-interface theories and an approach to predictive tumor growth modeling. *Mathematical Models and Methods in Applied Sciences*, 20(3):477–517, 2010. DOI: 10.1142/S0218202510004313.
- [5] J. T. Oden, E. E. Prudencio and A. Hawkins-Daarud. *Selection and assessment of phenomenological models of tumor growth*, 23(7):1309–1338, 2013.
- [6] M. L. Prasad and A. G. Huvos. Pathology of head and neck tumors. *American Cancer Society atlas of clinical oncology: cancer of the head and neck*, volume 1, chapter 2, pages 19–55, 2001.
- [7] A. Rohatgi. WebPlotDigitalizer 3.11: HTML5 based online tool to extract numerical data from plot images. URL: <<http://arohatgi.info/WebPlotDigitizer/>> (accessed on January 2017).
- [8] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, Philadelphia, 2005. ISBN: 0-89871-572-5.