

Método ANN-MoC para Problema Inverso de Caracterização da Fonte

Nelson Garcia Roman*¹, Pedro Costa dos Santos², Pedro Henrique de Almeida
Konzen³

^{1,3}PPGMAP, IME, UFRGS, Porto Alegre, RS

²IME, UFRGS, Porto Alegre, RS

Resumo

Problemas inversos de transporte de partículas neutras têm importantes aplicações na engenharia e na medicina. Neste trabalho, apresentamos uma nova aplicação do método ANN-MoC na resolução de um problema inverso de caracterização de fonte. O problema consiste em estimar parâmetros da fonte de partículas com base em medições da densidade de partículas nos extremos de um domínio computacional 1D. Em resumo, o método emprega uma ANN (do inglês, *artificial neural network*) como modelo de regressão para a estimativa dos parâmetros. A rede é treinada com base em dados gerados de soluções MoC (do inglês, *method of characteristics*) do problema direto de transporte associado. Para um estudo de caso, resultados são apresentados e indicam a robustez do método em fornecer estimativas precisas da caracterização da fonte.

1 Introdução

Problema inverso de transporte de partículas neutras tem importantes aplicações na engenharia e na medicina. Por exemplo, no contexto de transporte radiativo (Modest, 2013) encontra aplicações de desenvolvimento de protocolos de controle de qualidade e de segurança em processos industriais de manufaturas em altas temperaturas. No contexto de transporte nuclear (Lewis & Miller, 1984; Stacey, 2007), protocolos de segurança tem evidente importância. Na medicina ótica (Wang & Wu, 2007), encontra-se aplicações envolvendo o transporte de radiativo (por exemplo, tomografia computadorizada) como o de nêutrons.

Assumimos que o transporte de partículas é modelado pela equação linear de Boltzmann em um meio com espalhamento isotrópico (Lewis & Miller, 1984; Modest, 2013)

$$\forall \mu : \mu \frac{\partial I}{\partial x} + \sigma_t I(x, \mu) = \sigma_s \Psi(x) + q_{\alpha, \beta}(x), \quad x \in D, \quad (1)$$

*ngroman1992@gmail.com

onde $I = I(x, \mu)$ é a intensidade de partículas no ponto $x \in D = (0, 1)$ e na direção $\mu \in (-1, 1) \setminus \{0\}$. As propriedades do meio são dadas pelo coeficiente de absorção total $\sigma_t = \kappa + \sigma_s > 0$, onde $\kappa > 0$ é o coeficiente de absorção e $\sigma_s > 0$ o coeficiente de espalhamento. Ainda, tem-se a densidade de partículas

$$\Psi(x) := \frac{1}{2} \int_{-1}^1 I(x, \mu) d\mu \quad (2)$$

e dadas condições de contorno

$$\mu > 0 : I(0, \mu) = I_a, \quad (3a)$$

$$\mu < 0 : I(1, \mu) = I_b, \quad (3b)$$

onde I_a, I_b denotam as dadas intensidades de partículas que entram no domínio por suas extremidades. A fonte do problema de transporte é caracterizada pelos parâmetros $\alpha, \beta > 0$ e definida por

$$q_{\alpha, \beta}(x) := \begin{cases} \alpha & , |x - \beta| \leq \delta, \\ 0 & , |x - \beta| > \delta \end{cases} \quad (4)$$

Neste trabalho, aplicamos o método ANN-MoC (Roman et. al, 2023) para resolver o problema inverso de caracterização da fonte com base em medições da densidade de partículas nos extremos do domínio. O método consiste em treinar uma ANN (do inglês, *artificial neural network*) (Haykin, 2009) para estimar os parâmetros α e β da fonte. O conjunto de treinamento é gerado de soluções MoC (do inglês, *method of characteristics*) (Evans, 2010) do problema direto associado. Em trabalho anterior de Santos et. al (2022), a metodologia foi empregada para problemas distintos de determinação e localização da fonte. Aqui, a caracterização da fonte é buscada com uma única ANN.

Na sequência, apresentamos uma descrição método ANN-MoC. Em seguida, apresentamos resultados da aplicação e, então, fazemos algumas considerações finais.

2 Método ANN-MoC

O método ANN-MoC consiste em treinar uma ANN para resolver o problema inverso de transporte com base soluções MoC do problema direto associado.

2.1 MoC - Resolução do Problema Direto

Os conjuntos de treinamento e validação para a ANN são obtidos de soluções do problema direto para selecionados valores de α e β , considerando-se conhecidas as propriedades do meio.

A aplicação do MoC é feita a partir da aproximação DOM (do inglês, *discrete ordinates method* e do esquema SI (do inglês, *source iteration*) (Modest, 2013). Assumindo-se a quadratura gaussiana $\{(\mu_i, \omega_i)\}_{i=1}^N$, a formulação DOM & SI de (1) é dada como segue

$$1 \leq i \leq N : \mu_i \cdot \frac{\partial}{\partial x} I_i^{(j)}(x) + \sigma_t I_i^{(j)} = \sigma_s \Psi^{(j-1)}(x) + q(x, \mu_i), \quad \forall x \in \mathcal{D}, \quad (5a)$$

$$I_i^{(j)}(0) = I_a, \quad \forall \mu_i > 0; \quad I_i^{(j)}(1) = I_b, \quad \forall \mu_i < 0, \quad (5b)$$

onde $I_i^{(j)} \approx I^{(j)}(x, \mu_i)$, $j = 1, 2, \dots, L$, com dada aproximação inicial $\Psi^{(0)}(x)$. A j -ésima aproximação da densidade de partículas é dada por

$$\Psi^{(j)}(x) = \frac{1}{2} \sum_{i=1}^N \omega_i I_i^{(j)}(x). \quad (6)$$

O problema (5) consiste em um sistema de equações diferenciais parciais de primeira ordem com condições de contorno. O método MoC fornece a seguinte solução nas características $x(s) = x_0 + s\mu_i$, $s \in \mathbb{R}$,

$$I_i^{(j)}(s) = I_i^{(j)}(0) e^{-\int_0^s \sigma_t ds'} + \int_0^s \left[\Psi^{(j)}(s') + q(s, \mu_i) \right] e^{-\int_{s'}^s \sigma_t ds''} ds', \quad (7)$$

onde $I_i^{(j)}(s) = I_i^{(j)}(x(s))$. Consulte o trabalho de Santos et. al (2022) para maiores detalhes sobre nossa implementação do MoC.

2.2 ANN - Resolução do Problema Inverso

Para a resolução do problema inverso aplicamos uma ANN do tipo MLP (do inglês, *multi-layer perceptron*) (Haykin, 2009). A MLP é tratada como um modelo de regressão não linear do conjunto dos valores dados da densidade de partículas nos extremos do domínio ($\Psi(0)$, $\Psi(1)$) nos valores esperados de α e β para a caracterização da fonte. Com base em uma abordagem de treinamento supervisionado, um conjunto de treinamento (calibração) $\{\boldsymbol{\psi}^{(s)}, \boldsymbol{\gamma}^{(s)}\}_{s=1}^{n_{\text{train}}}$, $\boldsymbol{\psi}^{(s)} = (\Psi^{(s)}(0), \Psi^{(s)}(1))$, com n_{train} amostras é computado para selecionados parâmetros $\boldsymbol{\gamma}^{(s)} = (\alpha^{(s)}, \beta^{(s)})$, usando-se a solução MoC do problema direto. Aqui, com alguma sobreposição de notação, o sobre-índice denota a indexação da amostra.

No contexto deste trabalho, a ANN é denotada por

$$\tilde{\boldsymbol{\gamma}} = \mathcal{N} \left(\boldsymbol{\psi}; \left\{ \left(W^{(l)}, \mathbf{b}^{(l)}, \mathbf{f}^{(l)} \right) \right\}_{l=1}^{n_h+1} \right), \quad (8)$$

onde $(W^{(l)}, \mathbf{b}^{(l)}, \mathbf{f}^{(l)})$ denota a tripla de pesos $W^{(l)}$, bias $\mathbf{b}^{(l)}$ e função de ativação $\mathbf{f}^{(l)}$ na l -ésima camada da MLP, $l = 1, 2, \dots, n_h + 1$.

Fixada uma arquitetura de rede $2 - n_n \times n_h - 2$ (2 entradas, n_h camadas escondidas com n_n neurônios em cada, 2 saídas) e as funções de ativação, o treinamento da rede consiste em resolver o seguinte problema de minimização

$$\min_{\{(W^{(l)}, \mathbf{b}^{(l)})\}_{l=1}^{n_h+1}} \underbrace{\frac{1}{n_s} \sum_{s=1}^{n_s} \|\tilde{\boldsymbol{\gamma}}^s - \boldsymbol{\gamma}^s\|^2}_{=:\varepsilon}, \quad (9)$$

onde ε denota a função erro a ser minimizada.

De forma mais sucinta, o treinamento da rede consiste em calibrar seus pesos e *biases* de forma a minimizar a função erro MSE (do inglês, *mean squared error*). Para tanto, aplicamos o método de retropropagação (Haykin, 2009). As implementações foram realizadas em linguagem Python com a ajuda do pacote de aprendizagem de máquina PyTorch. Os parâmetros do código são problema dependentes e serão apresentados mais adiante junto com os resultados alcançados.

Uma vez calibrada a rede, propõe-se utilizar um conjunto de validação $\{\boldsymbol{\psi}^{(s)}, \boldsymbol{\gamma}^{(s)}\}_{s=1}^{n_{\text{valid}}}$ como teste. Essas amostras são todas diferentes das de treinamento e também são obtidas da solução do problema direto apresentado anteriormente.

3 Resultados

Neste seção, apresentados os resultados da aplicação do método ANN-MoC para a resolução de problema inverso de caracterização da fonte.

3.1 Dados de treinamento e validação

O conjunto de treinamento foi gerado da resolução do problema direto fixadas as propriedades do meio $\kappa = 0.75$ e $\sigma_s = 0.25$. As condições de contorno foram assumidas $I_a = I_b = 0$ (vácuo). Foram geradas $n_{\text{train}} = 170$ amostras escolhendo-se todas as combinações de $\alpha = 0.1 + (i - 1)0.1$, $i = 1, 2, \dots, 10$, e $\beta = 0.1 + (j - 1)0.05$, $j = 1, 2, \dots, 17$.

A solução numérica do problema direto pelo MoC depende dos parâmetros n_x (número de células na malha computacional) e N (número de pares da quadratura gaussiana). Experimentos numéricos indicaram que escolhas de $n_x = N = 100$ foram suficientes para obtermos soluções com precisão de 4 dígitos significativos para a densidade de partículas.

Para a validação, gerou-se um conjunto de 200 amostras com valores randômicos (distribuição uniforme) $0.1 \leq \alpha \leq 1.0$ e $0.1 \leq \beta \leq 0.9$.

$n_n \backslash n_h$	1	2	3	4
5	5000/1.5E-3	5000/2.1E-4	5000/6.2E-5	5000/8.7E-5
10	5000/1.1E-3	5000/7.8E-5	5000/2.4E-5	3632/<1E-5
20	5000/1.0E-3	5000/2.5E-5	3689/<1E-5	2723/<1E-5

Tabela 1: Escolha da arquitetura da rede.

3.2 Resultados de Treinamento

De posse do conjunto de treinamento $\{\boldsymbol{\psi}^{(s)}, \boldsymbol{\gamma}^{(s)}\}_{s=1}^{n_{\text{train}}}$, buscamos uma ANN com função erro $\varepsilon < 10^{-5}$. A Tabela 1 apresenta resultados de testes com várias arquiteturas de rede $2 - n_n \times n_h - 2$. Em todos os testes, escolhemos tangente hiperbólica nas camadas escondidas e função identidade na camada de saída como funções de ativação. O otimizador utilizado foi o método Adam (Kingma & Ba, 2014) com taxa de aprendizado $l_r = 10^{-2}$. Dada a estocasticidade do algoritmo de treinamento, cada teste foi executado três vezes. Os valores tabelados correspondem a $\bar{n}_e / \bar{\varepsilon}$, onde \bar{n}_e é o número médio de épocas demandadas (máximo fixado em 5000) e $\bar{\varepsilon}$ é o valor final médio da função erro (critério de parada $\varepsilon < 10^{-5}$). Observa-se que uma MLP com arquitetura $2 - 40 \times 4 - 2$ é suficiente para aprender os dados de treinamento em menos de 3000 épocas. Tendo em vista

que arquiteturas maiores não geraram um ganho significativo, escolhemos esta arquitetura para a geração dos resultados a seguir.

Com a arquitetura escolhida e a rede treinada, passamos a analisar os resultados da rede na resolução do problema inverso de transporte. A Figura 1 mostra as curvas de treinamento para os parâmetros α e β de caracterização da fonte. Observa-se que a rede fornece estimativas $\tilde{\alpha}$, $\tilde{\beta}$ com alta precisão em relação aos valores esperados α e β . Para ambos os parâmetros, o coeficiente de determinação R^2 foi maior que 0.9999.

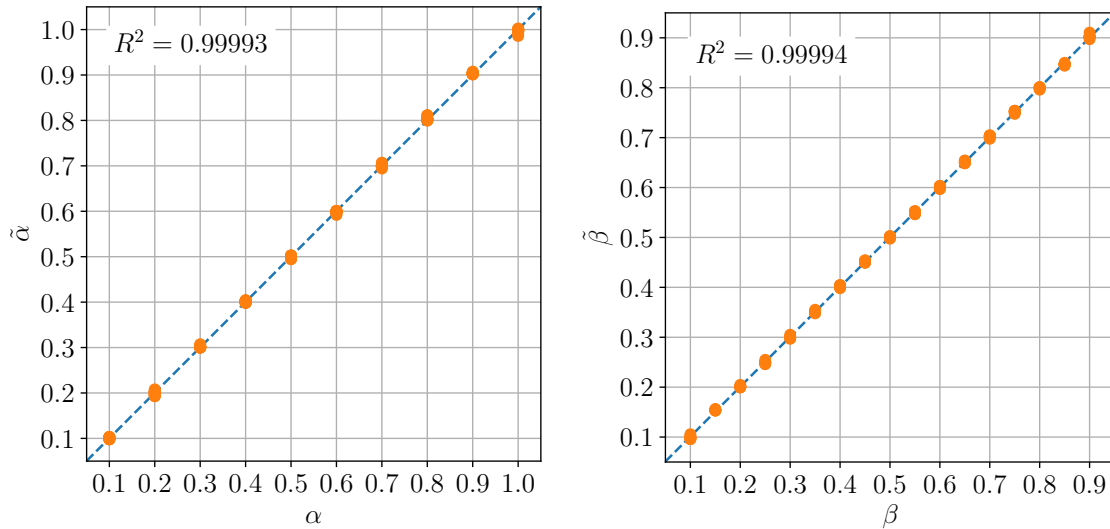


Figura 1: Curva de treinamento. Esquerda: α . Direita: β . Os pontos representam os resultados para cada amostra de treinamento e, em linha tracejada tem-se a reta ajustada por mínimos quadrados.

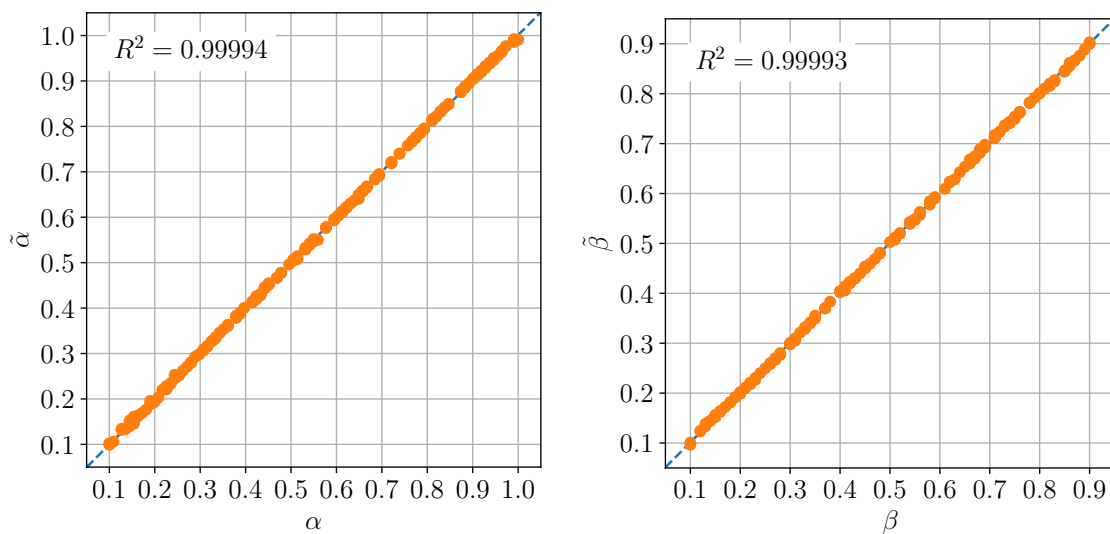


Figura 2: Curva de validação. Esquerda: α . Direita: β . Os pontos representam os resultados para cada amostra de validação e, em linha tracejada tem-se a reta ajustada por mínimos quadrados.

3.3 Resultados de Validação

O treinamento de uma ANN como um modelo de regressão pode resultar em sobre-ajuste, quando a rede se ajusta muito bem aos dados de treinamento, mas falha em estimar dados novos. Para verificar a robustez da MLP treinada, testamo-la ao conjunto de validação $\{\psi^{(s)}, \gamma^{(s)}\}_{s=1}^{n_{\text{valid}}}$ descrito acima. Os resultados obtidos são apresentados na Figura 2. Observa-se que a rede forneceu estimativas com boa precisão, com coeficientes de determinação R^2 novamente acima de 0.9999. Os resultados confirmam a robustez da rede em fornecer a caracterização da fonte com boa precisão.

4 Considerações Finais

Apresentamos a aplicação do método ANN-MoC na resolução de um problema inverso de transporte para a caracterização da fonte de partículas neutras. O problema consiste em determinar parâmetros da fonte com base em medições da densidade de partículas nos extremos do domínio. O método resume-se em treinar uma ANN com dados gerados de soluções MoC do problema de transporte direto associado. Uma ANN do tipo MLP foi empregada para fornecer estimativas dos parâmetros de caracterização da fonte. Os resultados de treinamento e de validação indicam a robustez da rede em fornecer estimativas precisas com base nas medições das densidades. A aplicação do método ANN-MoC em problemas mais realísticos requer o emprego de um modelo de transporte adequado e esforço deve ser focado na geração de dados de qualidade (de treinamento e de validação) para o treinamento da rede.

Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e a Universidade Federal do Rio Grande do Sul (UFRGS) pelo apoio a este trabalho através da concessão de bolsas de estudos.

Referências Bibliográficas

- Evans, L.C. (2010). *Partial differential equations*, 2ed., AMS.
- Haykin, S. (2009). *Neural networks and learning machines*, 3ed., Pearson.
- Kingma, D.P., Ba, J. (2014). Adam: a method for stochastic optimization. *arXiv*, 1412.6980.
- Lewis, E.E. and Miller, W.F. (1984). *Computational methods of neutron transport*, 1ed., Wiley.
- Modest, M.F. (2013). *Radiative heat transfer*, 3ed., Elsevier.
- Roman, N.G., Santos, P.C., Konzen, P.H.A. (2023) “ANN-MoC method for inverse transient transport problems in one-dimensional domain”. Submetido em: *XXVI ENMC/XIV ECTM*, Nova Friburgo, Brasil.
- Santos, P.C., Melo, G.R., Konzen, P.H.A. (2022) “RNAs aplicadas a determinação e localização

de fonte de partículas em problemas de transporte unidimensional”. Em: *Anais do XXV ENMC/XIII ECTM/9ª MCSul/IX SEMENGO*, evento online, Brasil.

Stacey, W.M. (2007). *Nuclear reactor physics*, 2ed., Wiley.

Wang, L.V., Wu, H. (2007). *Biomedical optics: principles and imaging*, 1ed., Wiley.