UFSM

CIÊNCIA e NATURA

ISSN 2179-460X
OPEN ACCESS

Ci. e Nat., Santa Maria, v. 47, spe. 1, e89819, 2025• https://doi.org/10.5902/2179460X89819
Submitted: 11/18/24 • Approved: 11/21/24 • Published: 01/15/25

IV Jornada de Matematica e Matematica aplicada UFSM

ANN-MoC method for the inverse problem of source characterization

Método ANN-MoC para problema inverso de caracterização da fonte

Nelson Garcia Roman [I], **Pedro Costas dos Santos** [I], **Pedro Henrique de Almeida Konzen** [I]

[I]Universidade Federal do Rio Grande do Sul, RS, Brazil

ABSTRACT

Inverse problems of neutral particle transport have significant applications in engineering and medicine. In this study, we present a new application of the ANN-MoC method to solve inverse problems of source characterization. It involves estimating the source parameters based on measurements of particle density at the boundaries of a one-dimensional computational domain. In summary, the method employs an artificial neural network (ANN) as a regression model. The neural network is trained using data generated from solutions of the method of characteristics (MoC) for the associated direct transport problem. Results of three test cases are presented. In the first, we highlight the advantage of preprocessing the input data. For all cases, sensibility tests are provided to study the advantages and limitations of the proposed approach in solving inverses problems with noisy data.

Keywords: Artificial neural network; Method of characteristics; Particle neutral transport; Inverse problem

RESUMO

Problemas inversos de transporte de partículas neutras têm aplicações significativas em engenharia e medicina. Neste estudo, apresentamos uma nova aplicação do método ANN-MoC para resolver problemas inversos de caracterização de fonte. Isso envolve a estimativa dos parâmetros da fonte com base em medidas da densidade de partículas nas fronteiras de um domínio computacional unidimensional. Em resumo, o método emprega uma rede neural artificial (ANN) como um modelo de regressão. A rede neural é treinada usando dados gerados a partir de soluções do método das características (MoC) para o problema direto de transporte associado. Resultados de três casos de teste são apresentados. No primeiro, destacamos a vantagem do pré-processamento dos dados de entrada. Para todos os casos, testes de sensibilidade são fornecidos para estudar as vantagens e limitações da abordagem proposta na resolução de problemas inversos com dados ruidosos.

Palavras-chave: Rede neural artificial; Método das características; Transporte de partículas neutras; Problemas inversos

1 INTRODUCTION

The inverse problem of neutral particle transport has significant applications in engineering and medicine. For instance, in the context of radiative transport (Modest, 2013), applications include the development of quality control and safety protocols in high temperatures manufacturing processes, such as glass and ceramic manufactures (Larsen, Thömmesand, Klar, Seaıd, e Götz, 2002). In the realm of nuclear transport (Lewis e Miller, 1984; Stacey, 2007), safety protocols hold evident importance. In optical medicine (Hielscher, Alcouffe, e Barbour, 1998; Tarvainen, Vauhkonen, e Arridge, 2008; Wang e Wu, 2012), applications involving radiative transport (e.g., computed tomography) and neutron transport are found.

We assume that particle transport is modeled by the linear Boltzmann equation in a medium with isotropic scattering (Lewis e Miller, 1984; Modest, 2013)

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

where $I = I(x,\mu)$ represents the particle intensity at point $x \in D = (0,1)$ and in the direction $\mu \in (-1,1) \setminus 0$. The medium properties are given by the total absorption coefficient $\sigma_t = \kappa + \sigma_s > 0$, where $\kappa > 0$ is the absorption coefficient and $\sigma_s > 0$ is the scattering coefficient. Furthermore, we have the particle density

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

and given boundary conditions

$$\mu > 0: \ I(0,\mu) = I_0, \tag{3a}$$

$$\mu < 0: \ I(1,\mu) = I_1, \tag{3b}$$

where I_0, I_1 denote the given particle intensities entering the domain through the boundary. The source of the transport problem is characterized by the parameters $\alpha, \beta > 0$, which is defined according to the problem at hand.

Inverse problems are commonly formulated as optimization problems (Nocedal e Wright, 1999), where the aim is to minimize an objective function measuring the error between observed and model-predicted data. To solve these problems, optimization methods such as the conjugate gradient method (Li, 1997) and the Levenberg-Marquardt method (Mengüç e Manickavasagam, 1993) are employed.

In addition to traditional optimization methods, heuristic methods have been studied to solve global optimization problems in the context of inverse problems. Among these methods are the genetic algorithm (Kim, Baek, Kim, e Ryou, 2004) and the particle swarm optimization algorithm (Qi, Ruan, Zhang, Wang, e Tan, 2007).

In other cases, as an alternative, the Monte Carlo method (Kaipio e Somersalo, 2006) is used to address the uncertainty and non-uniqueness in estimating the model parameters from observed data. This method is based on random number sampling to generate approximate solutions to the inverse problem, thereby allowing a robust estimation of the probability distribution of the model parameters.

In this work, we apply the ANN-MoC method (Roman, Santos, e Konzen, 2023) to solve the inverse problem of source characterization based on measurements of the particle density at the domain's boundaries. The method involves training an artificial neural network (ANN) (Haykin, 2009) to estimate the parameters α and β of the source. The training set is generated from solutions of the method of characteristics (MoC) (Evans, 2010) of the associated direct problem. In previous the work of Santos, Melo, e Konzen (2022), the methodology was employed for distinct problems of source determination and localization. Here, the source characterization is pursued with a single ANN.

Next, we present a description of the ANN-MoC method. Subsequently, we present results from its application, followed by some final considerations.

2 METHOD ANN-MOC

The ANN-MoC method consists of training an artificial neural network (ANN) to solve the inverse transport problem based on solutions of the method of characteristics (MoC) of the associated direct problem.

2.1 MoC - Direct Problem Solutions

The training and validation sets for the ANN are obtained from solutions of the direct problem for selected values of α and β, considering the properties of the medium to be known.

The application of the MoC is carried out using the discrete ordinates method (DOM) approximation and the source iteration (SI) scheme (Modest, 2013). Assuming the Gaussian quadrature $\{(\mu_i, \omega_i)\}_{i=1}^{N}$, the DOM & SI formulation of (1)-(3) is given as follows:

$$1 \leq i \leq N: \ \mu_i \cdot \frac{\partial}{\partial x} I_i^{(j)}(x) + \sigma_t I_i^{(j)} = \sigma_s \Psi^{(j-1)}(x) + q(x, \mu_i), \ \forall x \in \mathcal{D}, \tag{4a}$$

$$\forall \mu_i > 0: \ I_i^{(j)}(0) = I_0, \tag{4b}$$

$$\forall \mu_i < 0: \ I_i^{(j)}(1) = I_1, \tag{4c}$$

where $I_i^{(j)} \approx I^{(j)}(x, \mu_i)$, $j = 1, 2, \ldots, L$, with a given initial approximation $\Psi^{(0)}(x)$. The j-th approximation of the particle density is given by

$$\Psi^{(j)}(x) = \frac{1}{2} \sum_{i=1}^{N} \omega_i I_i^{(j)}(x). \tag{5}$$

The problem (4) consists of a system of first-order partial differential equations with boundary conditions. The method of characteristics (MoC) provides the following solution along the characteristics $x(s) = x_0 + s\mu_i$, $s \in \mathbb{R}$,

$$I_i^{(j)}(s) = I_i^{(j)}(0) e^{-\int_0^s \sigma_t \, ds'} + \int_0^s \left[\Psi^{(j)}(s') + q(s, \mu_i)\right] e^{-\int_{s'}^s \sigma_t \, ds''} \, ds', \tag{6}$$

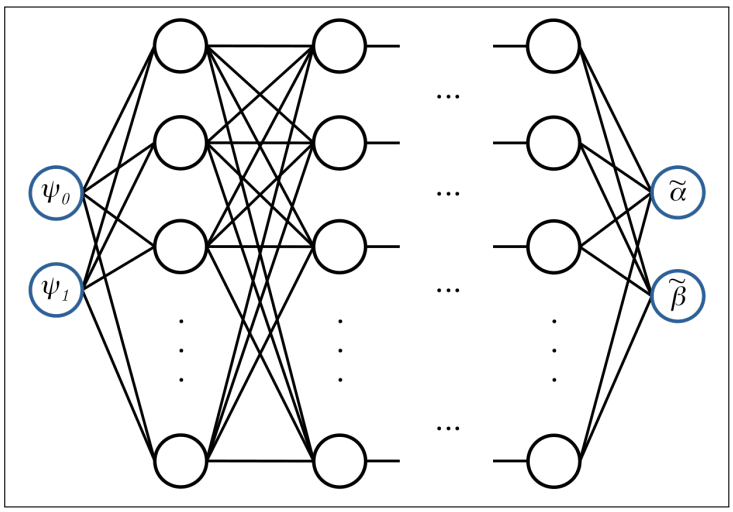
where $I_i^{(j)}(s) = I_i^{(j)}(x(s))$. Refer to the works of Santos et al. (2022) and Roman et al. (2023) for further details on our implementation of the MoC.

2.2 ANN - Solving the Inverse Problem

To solve the inverse problem, we apply a multilayer perceptron (MLP) neural network (Haykin, 2009). The MLP is designed as a nonlinear regression model of the given particle density values at the domain boundaries $(\Psi(0), \Psi(1))$ to the expected

values of α and β for source characterization $q_{\alpha,\beta}$ (see Figure 1). Through a supervised training approach, a training set (calibration) $\{\boldsymbol{\psi}^{(s)}, \boldsymbol{\gamma}^{(s)}\}_{s=1}^{n_{\text{train}}}$, where $\boldsymbol{\psi}^{(s)} = \left(\Psi^{(s)}(0), \Psi^{(s)}(1)\right)$, with n_{train} samples, is computed for selected parameters $\boldsymbol{\gamma}^{(s)} = \left(\alpha^{(s)}, \beta^{(s)}\right)$ using the MoC solution of the direct problem. Here, with some notation overlap, the superscript denotes the sample indexing.

Figure 1 – The MLP neural network architecture $2 - n_n \times n_h - 2$



Source: the authors (2024)

In the context of this work, the ANN is denoted by

$$\tilde{\boldsymbol{\gamma}} = \mathcal{N}\left(\boldsymbol{\psi}; \left\{\left(W^{(l)}, \boldsymbol{b}^{(l)}, \boldsymbol{f}^{(l)}\right)\right\}_{l=1}^{n_h+1}\right), \tag{7}$$

where $\left(W^{(l)}, \boldsymbol{b}^{(l)}, \boldsymbol{f}^{(l)}\right)$ denotes the triple of weights $W^{(l)}$, biases $\boldsymbol{b}^{(l)}$, and activation function $\boldsymbol{f}^{(l)}$ in the l-th layer of the MLP, $l = 1, 2, \ldots, n_h + 1$.

Given a network architecture $2 - n_n \times n_h - 2$ (2 inputs, n_h hidden layers with n_n neurons each, 2 outputs), and the activation functions, training the network involves solving the following minimization problem

$$\min_{\left\{\left(W^{(l)}, \boldsymbol{b}^{(l)}\right)\right\}_{\forall l}} \underbrace{\frac{1}{n_s} \sum_{s=1}^{n_s} \left\| \tilde{\boldsymbol{\gamma}}^{(s)} - \boldsymbol{\gamma}^{(s)} \right\|^2}_{=:\varepsilon}, \tag{8}$$

where ε denotes the loss function to be minimized.

More concisely, training the network involves calibrating its weights and biases to minimize the mean squared error (MSE) function. To achieve this, we apply the

backpropagation method (Haykin, 2009) with the Adam optimizer (Kingma e Ba, 2017). The implementations were carried out in the Python language with the help of the PyTorch machine learning package. The code parameters are problem-dependent and will be given later along with the following discussed test cases. Model validation is performed with a new data set $\{\boldsymbol{\psi}^{(s)}, \boldsymbol{\gamma}^{(s)}\}_{s=1}^{n_{\text{valid}}}$ and different from the training one, also obtained from MoC solutions of the direct problem presented earlier.

3 RESULTS

In this section, we present the results of applying the ANN-MoC method to solve three inverse problems of source characterization.

The numerical solution of the direct problem by the MoC depends on the parameters n_x (number of cells in the computational mesh) and N (number of pairs in the Gaussian quadrature). Numerical experiments indicated that the choices $n_x = N = 100$ are sufficient to obtain solutions with an accuracy of at least 3 significant digits for the particle density. In all the following cases, the domain is assumed $[a, b] = [0, 1]$ with boundary conditions $I_0 = I_1 = 0$.

For the training of the ANN models the loss tolerance of $\varepsilon \leq 10^{-5}$ has been set as a stop criterion. All the tests assume a model architecture of $2 - n_n \times n_h - 2$ (2 inputs, n_h hidden layers each with n_n neurons, and 2 outputs), and the tangent hyperbolic and the identity as the activation functions in hidden and output layers, respectively.

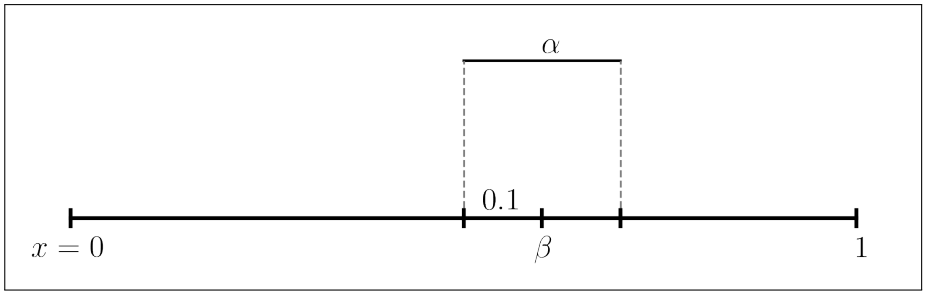
3.1 Inverse Problem 1

The inverse problem 1 consists of estimating the parameters of the particle source characterized as

$$q_{\alpha,\beta}(x) := \begin{cases} \alpha & ,|x-\beta| \leq 0.1, \\ 0 & ,|x-\beta| > 0.1, \end{cases} \tag{9}$$

where, here, the source intensity is $0.1 \leq \alpha \leq 1.0$ and its location $0.1 \leq \beta \leq 0.9$ (see Figure 2). The medium properties are assumed $\kappa = 0.75$ and $\sigma_s = 0.25$.

Figure 2 – Inverse Problem 1: source illustration



Source: the authors (2024)

The ANN model is trained from a set of $n_{\text{train}} = 170$ samples, which were generated from solving the direct problem for each of the combinations of $\alpha = 0.1 + (i-1)0.1$, $i = 1, 2, \ldots, 10$, and $\beta = 0.1 + (j-1)0.05$, $j = 1, 2, \ldots, 17$. For the model validation, a set of 200 samples was generated with random values (uniform distribution) $0.1 \leq \alpha \leq 1.0$ and $0.1 \leq \beta \leq 0.9$.

The ANN model architecture has been chosen after several tests. Due to the stochasticity of the training method, each test has been repeated three times. Table 1 presents the results with no data preprocessing. The tabulated values correspond to $\bar{n}_e/\bar{\varepsilon}$, where $\bar{n}_e$ is the average number of epochs required (maximum fixed at 5000) and $\bar{\varepsilon}$ is the average final value of the error function (stopping criterion $\varepsilon < 10^{-5}$). It is observed that an MLP with architecture $2 - 20 \times 4 - 2$ is sufficient to learn the training data in fewer than 3000 epochs. To enhance the training, we have then performed trials with data preprocessing. Inputs of the training samples have been scaled with the Standard Scaler, several MLP architectures have been tested, and the results can be found in Table 2. The enhancement with the preprocessing is notable, especially for the $2 - 40 \times 4 - 2$ architecture, which has been chosen in the following presented results.

Table 1 – Inverse Problem 1: choice of network architecture without data preprocessing - tabulated values $\bar{n}_e/\bar{\varepsilon}$

$n_n \backslash n_h$	1	2	3	4
5	$5000/1.5 \times 10^{-3}$	$5000/2.1 \times 10^{-4}$	$5000/6.2 \times 10^{-5}$	$5000/8.7 \times 10^{-5}$
10	$5000/1.1 \times 10^{-3}$	$5000/7.8 \times 10^{-5}$	$5000/2.4 \times 10^{-5}$	$3632/ < 1 \times 10^{-5}$
20	$5000/1.0 \times 10^{-3}$	$5000/2.5 \times 10^{-5}$	$3689/ < 1 \times 10^{-5}$	$2723/ < 1 \times 10^{-5}$
30	$5000/8.6 \times 10^{-4}$	$5000/2.5 \times 10^{-5}$	$5000/1.6 \times 10^{-5}$	$3954/ < 1 \times 10^{-5}$
40	$5000/1.1 \times 10^{-3}$	$5000/7.6 \times 10^{-5}$	$3840/ < 1 \times 10^{-5}$	$3652/ < 1 \times 10^{-5}$

Source: the authors (2024)

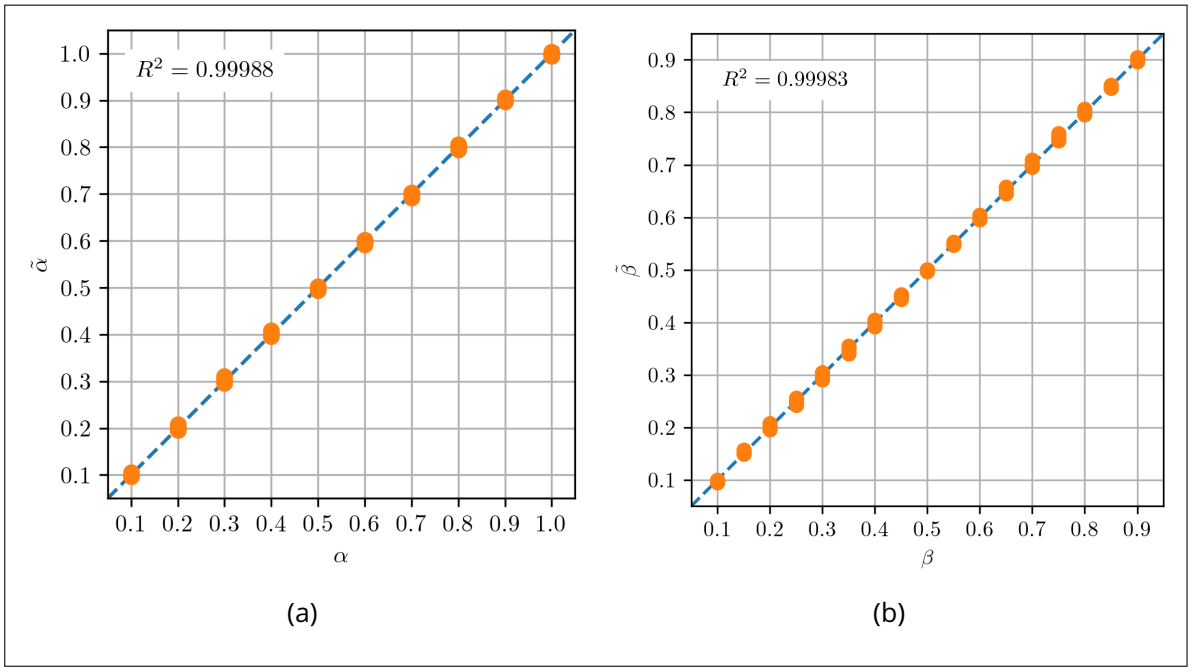
Table 2 – Inverse Problem 1: choice of network architecture with data preprocessing - tabulated values $\bar{n}_e/\bar{\varepsilon}$

$n_n \backslash n_h$	1	2	3	4
5	$5000/5.6 \times 10^{-4}$	$5000/2.1 \times 10^{-5}$	$5000/2.6 \times 10^{-5}$	$5000/4.4 \times 10^{-5}$
10	$5000/1.7 \times 10^{-4}$	$5000/3.7 \times 10^{-5}$	$3700/ < 1 \times 10^{-5}$	$3506/ < 1 \times 10^{-5}$
20	$5000/7.3 \times 10^{-5}$	$2974/ < 1 \times 10^{-5}$	$1847/ < 1 \times 10^{-5}$	$1863/ < 1 \times 10^{-5}$
30	$5000/6.8 \times 10^{-5}$	$2785/ < 1 \times 10^{-5}$	$1736/ < 1 \times 10^{-5}$	$1603/ < 1 \times 10^{-5}$
40	$5000/6.6 \times 10^{-5}$	$2465/ < 1 \times 10^{-5}$	$1594/ < 1 \times 10^{-5}$	$1305/ < 1 \times 10^{-5}$

Source: the authors (2024)

With the chosen architecture and the network trained, we proceed to analyze the results in solving the inverse transport problem. Figure 3 shows the training curves for the parameters α (a) and β (b) that characterize the source. The points represent the results for each training sample, and the dashed line is line fitted by least squares. It can be observed that the network provides estimates $\tilde{\alpha}$, $\tilde{\beta}$ with high precision compared to the expected values. For both parameters, the coefficient of determination $R^2 = 0.9998$ has been reached.

Figure 3 – Inverse Problem 1: training curves



Source: the authors (2024)

Training an ANN as a regression model can lead to overfitting, where the network fits the training data very well but fails to accurately estimate new data. To validate the robustness and stability of the trained MLP in this problem, we tested it on the validation

set $\{\boldsymbol{\psi}^{(s)}, \boldsymbol{\gamma}^{(s)}\}_{s=1}^{n_{\text{valid}}}$ described above. Additionally, a sensitivity test was applied, involving adding uniformly distributed noise to the input data.

The results indicate that the MLP model is relatively robust to moderate and slightly high levels of noise in the input data. Table 3 shows the results of the mean squared error R2 and the mean absolute squared error (MAPE) for different levels of noise. The noise is propagated to the output by a factor of 1.6 times for α (Table 3, left) and 1.7 times for β (Table 3, right). An $R^2 > 0.90$ is achieved even with a noise level of up to 10%. Figure 4 displays the expected versus estimated α (a) and β (b) of the test dataset with noise levels of 0%, 4%, 6%, and 8%. In the figures, the identity line is plotted as a dashed line as a guide. We observe the absence of outliers, which also indicates good generalization of the ANN-MoC method.

Table 3 – Inverse Problem 1: sensitivity tests for α (left) and β (right)

$Noise(\%)$	R^2	$MAPE(\%)$
0	0.9999	0.53
1	0.9998	0.54
2	0.9997	0.80
3	0.9992	1.48
4	0.9973	3.01
5	0.9939	4.15
6	0.9874	6.48
7	0.9779	8.11
8	0.9604	11.36
9	0.9417	13.50
10	0.9920	15.49

$Noise(\%)$	R^2	$MAPE(\%)$
0	0.9999	0.75
1	0.9997	0.77
2	0.9995	1.03
3	0.9987	1.73
4	0.9967	3.00
5	0.9904	4.46
6	0.9845	6.52
7	0.9616	8.72
8	0.9420	11.62
9	0.9181	13.89
10	0.8961	17.56

Source: the authors (2024)

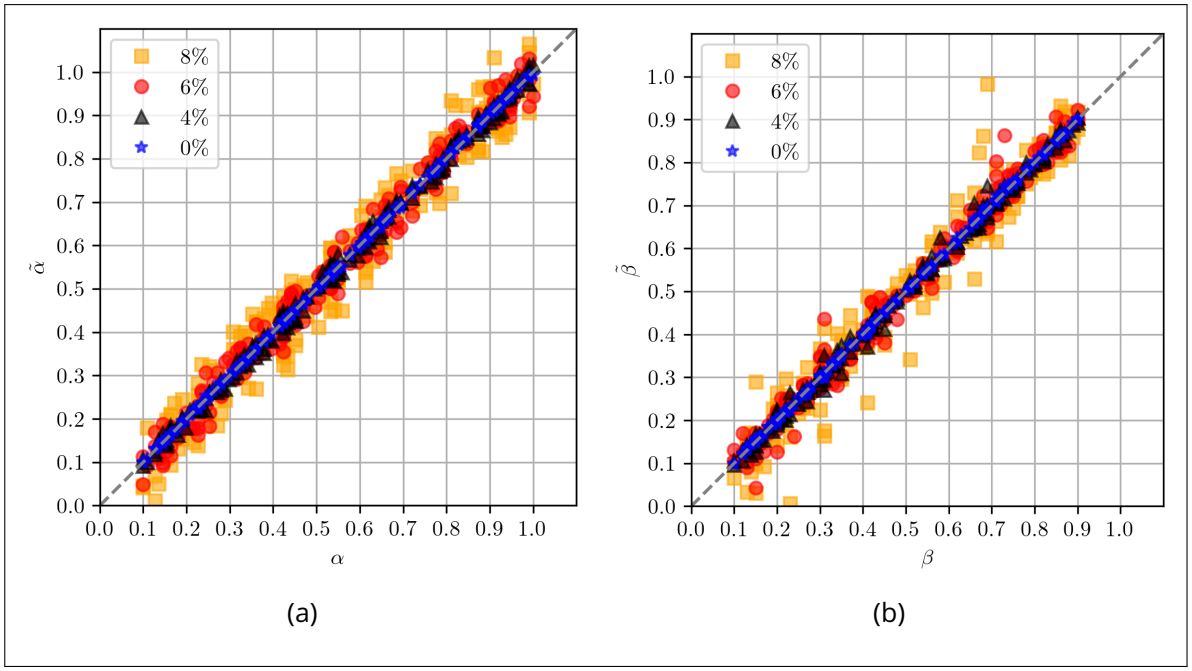
3.2 Inverse Problem 2

The inverse problem 2 consists of estimating the parameters of the particle source characterized as

$$q_{\alpha,\beta}(x) := \begin{cases} \alpha & , 0.4 \leq x \leq 0.5 \\ \beta & , 0.5 < x \leq 0.6, \\ 0 & , \text{otherwise}, \end{cases} \tag{10}$$

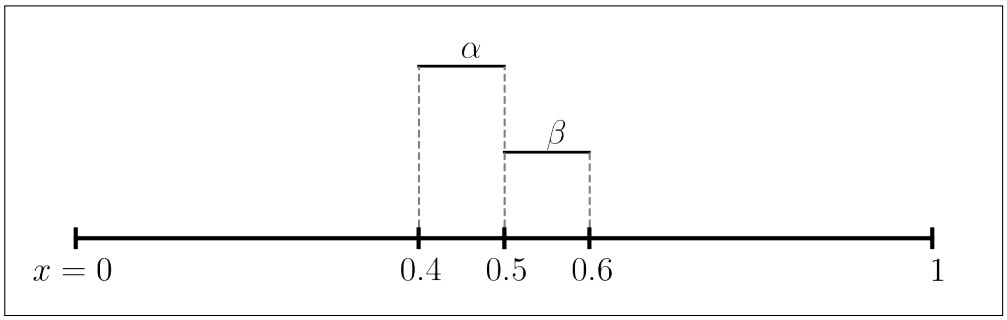
where, here, the source intensities are $0.1 \leq \alpha, \beta \leq 1.0$ (see Figure 5). The medium properties are assumed $\kappa = 0.5$ and $\sigma_s = 0.5$.

Figure 4 – Inverse Problem 1: validation curves for different levels of noise



Source: the authors (2024)

Figure 5 – Inverse Problem 2: source illustration



Source: the authors (2024)

The training set has been generated with $n_{\text{train}} = 100$ samples by selecting all combinations of $\alpha, \beta = 0.1 + (i-1)0.1$, $i = 1, 2, \ldots, 10$. For the validation, a set of 200 samples was generated with random values (uniform distribution) $0.1 \leq \alpha, \beta \leq 1.0$.

Here, we directly apply the Standard Scaler to preprocessing the inputs of the model. Several network architectures have been tested, each trained three times, due to the stochastic nature of the training approach. Analogous for the inverse problem 1, Table 4 presents the values of $\bar{n}_e/\bar{\varepsilon}$, where $\bar{n}_e$ is the average number of epochs required (with a maximum of 5000) and $\bar{\varepsilon}$ is the average final value of the error function. It is observed that with an MLP architecture of $2 - 40 \times 2 - 2$, it is sufficient to learn the

training data in less than 1000 epochs. Increasing the network size does not guarantee better learning results.

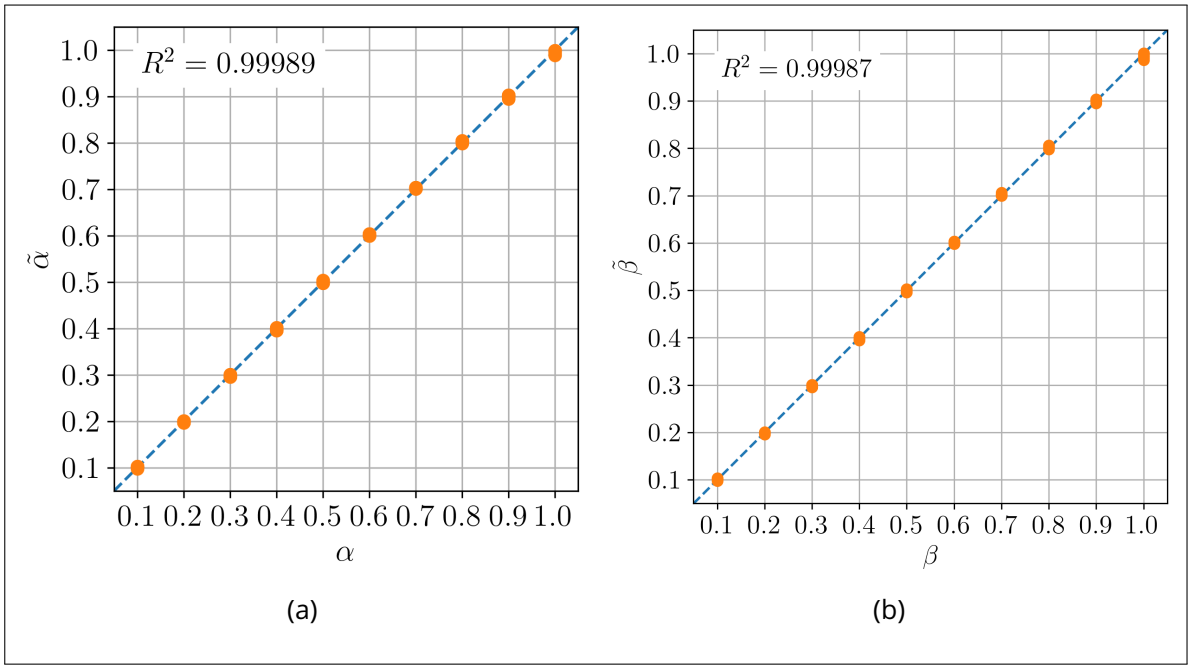
With the chosen architecture and the trained network, we proceed to analyze the results of the network in solving the inverse transport problem.Figure 6 shows the training curves for the parameters α (a) and β (b) that characterize the source. The points represent the results for each training sample, and the dashed line is line fitted by least squares. It can be observed that the network provides estimates $\tilde{\alpha}$, $\tilde{\beta}$ with high precision compared to the expected values α and β. For both parameters, the coefficient of determination R^2 was greater than 0.9998.

Table 4 – Inverse Problem 2: choice of network architecture with data preprocessing - tabulated values $\bar{n}_e/\bar{\varepsilon}$

$n_n \backslash n_h$	1	2	3	4
5	$3402/ < 1 \times 10^{-5}$	$3997/ < 1 \times 10^{-5}$	$5000/1.18 \times 10^{-5}$	$5000/1.6 \times 10^{-5}$
10	$1634/ < 1 \times 10^{-5}$	$1325/ < 1 \times 10^{-5}$	$1915/ < 1 \times 10^{-5}$	$1563/ < 1 \times 10^{-5}$
20	$1138/ < 1 \times 10^{-5}$	$717/ < 1 \times 10^{-5}$	$1201/ < 1 \times 10^{-5}$	$1467/ < 1 \times 10^{-5}$
30	$967/ < 1 \times 10^{-5}$	$563/ < 1 \times 10^{-5}$	$1040/ < 1 \times 10^{-5}$	$2367/ < 1 \times 10^{-5}$
40	$1119/ < 1 \times 10^{-5}$	$464/ < 1 \times 10^{-5}$	$1145/ < 1 \times 10^{-5}$	$1671/ < 1 \times 10^{-5}$

Source: the authors (2024)

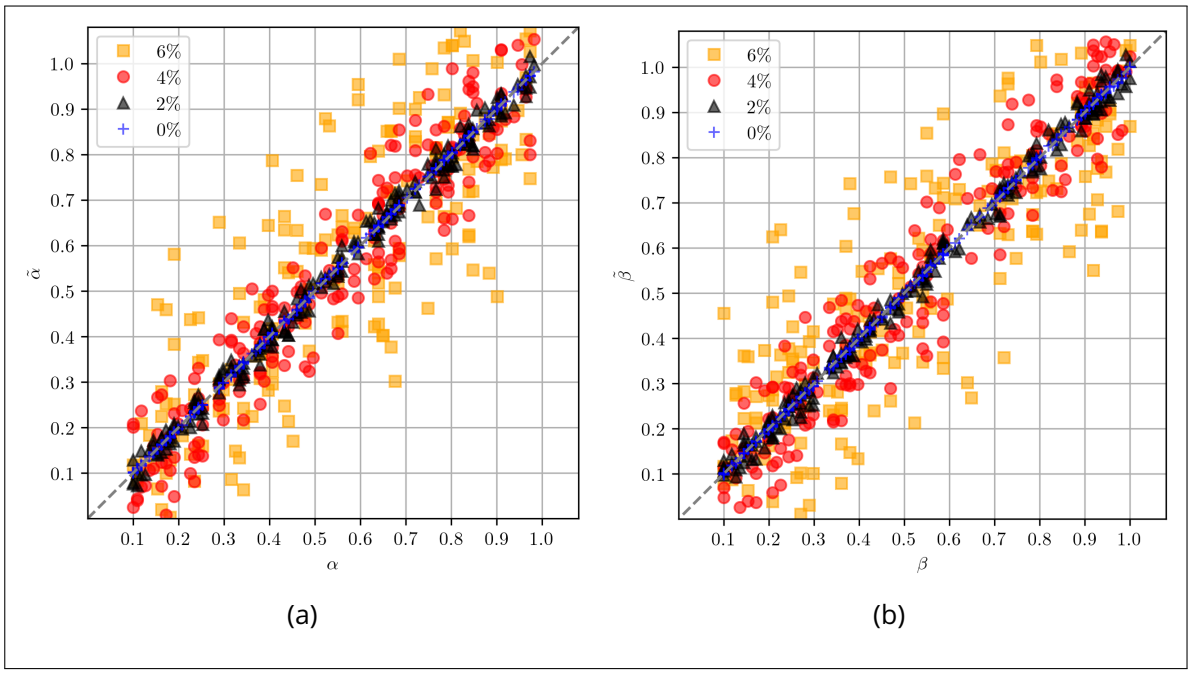
Figure 6 – Inverse Problem 2: training curves



Source: the authors (2024)

We now discuss the validation and test the sensibility of the model for this inverse problem. See Figure 7, for the validation and sensitivity curves, and Table 5 for the R^2 and $MAPE$ obtain with different levels of noise,for α (Table 5, left) and for β (Table 5, right).

Figure 7 – Inverse Problem 2: validation curves for different levels of noise



Source: the authors (2024)

Table 5 – Inverse Problem 2: sensitivity tests for α (left) and β (right)

$Noise(\%)$	R^2	$MAPE(\%)$
0	0.9999	0.49
1	0.9995	1.15
2	0.9933	4.48
3	0.9716	8.80
4	0.9185	17.48
5	0.7955	25.87
6	0.6602	36.75
7	0.5539	52.94
8	0.3579	64.04
9	0.3008	80.47
10	0.2181	105.77

$Noise(\%)$	R^2	$MAPE(\%)$
0	0.9999	0.51
1	0.9996	1.15
2	0.9943	4.56
3	0.9749	9.85
4	0.9155	19.00
5	0.8480	26.53
6	0.6808	42.72
7	0.5506	53.70
8	0.4385	67.96
9	0.2975	92.08
10	0.2044	104.15

Source: the authors (2024)

The results indicate that, without noise, the estimation error is just about 0.5%. The sensibility test, indicates that errors remain moderate (less than 10%) for a noise

level up to 3%. However, here, the errors grow exponentially with the noise. An $R^2 > 0.90$ is achieved with a noise level up to 4%.

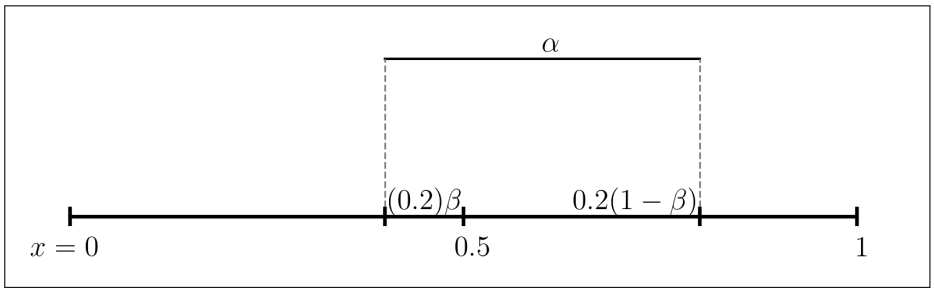
3.3 Inverse Problem 3

The inverse problem 3 consists of estimating the parameters of the particle source characterized as

$$q_{\alpha,\beta}(x) := \begin{cases} \alpha & , 0.5 - \beta \cdot 0.2 \leq x \leq 0.5 + (1-\beta) \cdot 0.2, \\ 0 & , \text{otherwise}, \end{cases} \tag{11}$$

where, here, the source intensity is $0.1 \leq \alpha \leq 1.0$ and $0.1 \leq \beta \leq 1.0$ is the factor of displacement from the center of the domain (see Figure 8). The medium properties are assumed $\kappa = 0.5$ and $\sigma_s = 0.5$.

Figure 8 – Inverse Problem 3: source illustration



Source: the authors (2024)

The training set has been generated with $n_{\text{train}} = 100$ samples by selecting all combinations of $\alpha, \beta = 0.1 + (i-1)0.1$, $i = 1, 2, \ldots, 10$. For validation, a set of 200 samples was generated with random values (uniform distribution) $0.1 \leq \alpha, \beta \leq 1.0$.

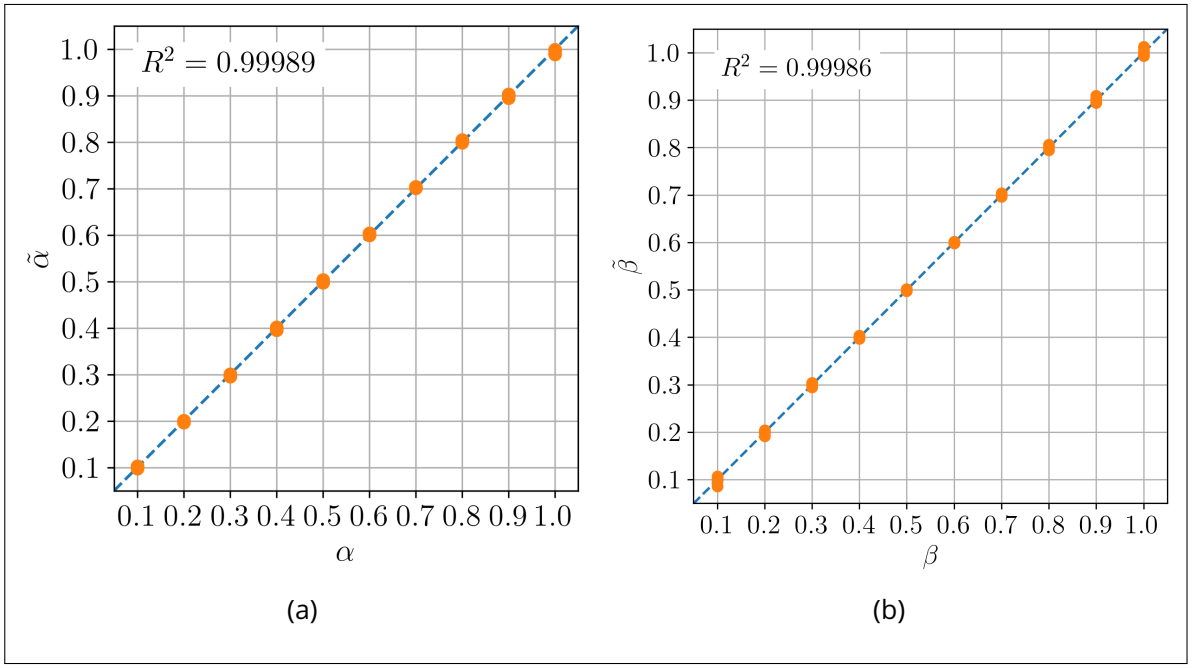
Following and analogous study of the previous inverse problems, network architecture is chosen after the tests presented in Table 6 presents the values of $\bar{n}_e/\bar{\varepsilon}$, where $\bar{n}_e$ is the average number of epochs required (with a maximum of 5000) and $\bar{\varepsilon}$ is the average final value of the loss function in training. It is noteworthy that an MLP architecture of $2 - 40 \times 2 - 2$, is sufficient to learn the training data in less than 1000 epochs. However, increasing the size of the network does not guarantee an improvement in learning outcomes.

Table 6 – Inverse Problem 3: choice of network architecture with data preprocessing - tabulated values $\bar{n}_e/\bar{\varepsilon}$

$n_n \backslash n_h$	1	2	3	4
5	$5000/1.6 \times 10^{-5}$	$5000/2.6 \times 10^{-5}$	$5000/1.1 \times 10^{-5}$	$5000/3.0 \times 10^{-5}$
10	$5000/4.9 \times 10^{-5}$	$2288/ < 1 \times 10^{-5}$	$2093/ < 1 \times 10^{-5}$	$2214/ < 1 \times 10^{-5}$
20	$5000/8.7 \times 10^{-5}$	$1911/ < 1 \times 10^{-5}$	$1125/ < 1 \times 10^{-5}$	$1260/ < 1 \times 10^{-5}$
30	$5000/4.4 \times 10^{-5}$	$1643/ < 1 \times 10^{-5}$	$1233/ < 1 \times 10^{-5}$	$1067/ < 1 \times 10^{-5}$
40	$5000/5.6 \times 10^{-5}$	$1291/ < 1 \times 10^{-5}$	$1069/ < 1 \times 10^{-5}$	$844/ < 1 \times 10^{-5}$

Source: the authors (2024)

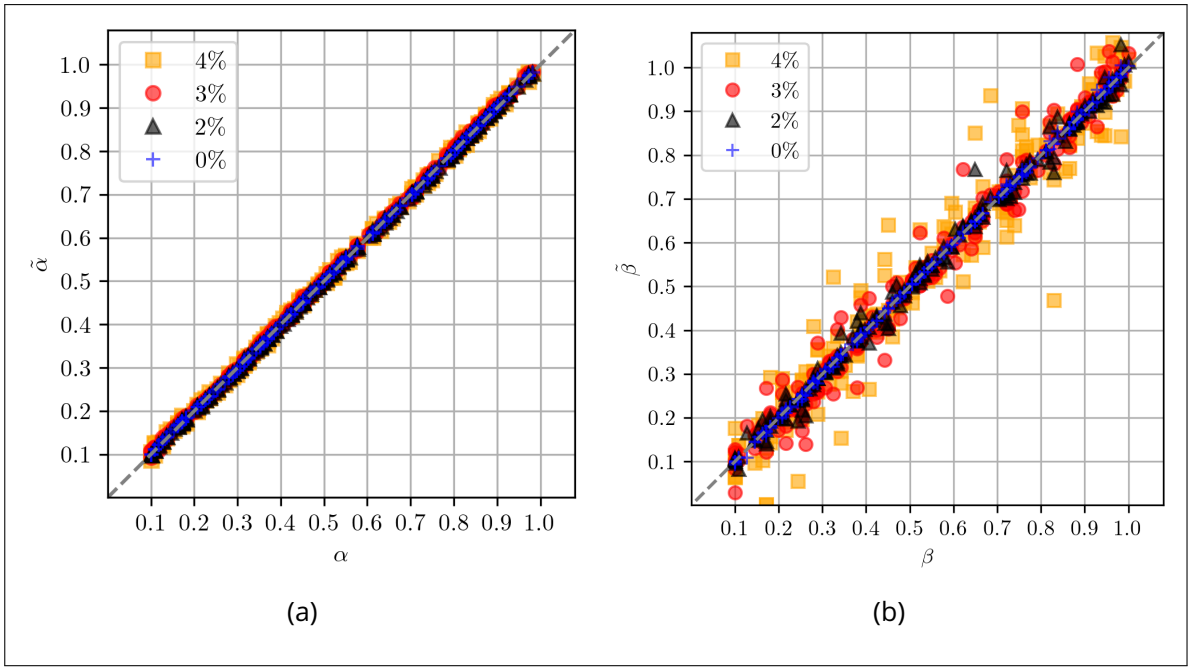
Figure 9 – Inverse Problem 3: training curves



Source: the authors (2024)

With the chosen architecture and the trained network, we proceed to analyze the results of the network in solving the inverse transport problem. Figure 9 shows the training curves for the parameters α (a) and β (b) that characterize the source. The points represent the results for each training sample, and the dashed line is line fitted by least squares. For no noisy data, it can be observed that the network provides estimates $\tilde{\alpha}$, $\tilde{\beta}$ with good precision compared to the expected values α and β. For both parameters, the coefficient of determination R^2 was greater than 0.9998.

Figure 10 – Inverse Problem 3: validation curves for different levels of noise



Source: the authors (2024)

Table 7 – Inverse Problem 3: sensitivity tests for α (left) and β (right)

$Noise(\%)$	R^2	$MAPE(\%)$
0	0.9999	0.69
1	0.9999	0.71
2	0.9998	0.88
3	0.9997	1.06
4	0.9992	1.81
5	0.9983	2.47
6	0.9961	3.92
7	0.9923	5.59
8	0.9881	7.23
9	0.9767	9.88
10	0.9693	11.88

$Noise(\%)$	R^2	$MAPE(\%)$
0	0.9997	0.75
1	0.9995	1.08
2	0.9949	3.19
3	0.9813	6.98
4	0.9180	14.32
5	0.8418	19.98
6	0.7506	28.76
7	0.6814	33.91
8	0.5257	50.45
9	0.3853	56.17
10	0.3057	67.08

Source: the authors (2024)

We now discuss the validation and sensitivity of the model for this inverse problem. See Figure 10 for the validation and sensitivity curves, and Table 7 for the R^2 and $MAPE$ obtained with different levels of noise for α (Table 7, left) and for β (Table 7, right).

One can observe how the noise propagates in the output of Figure 10 (b) compared to Figure 10 (a). This indicates that our MLP architecture is more sensitive in estimating β, where the noise propagates by a factor of 6.9 times. This suggests that it is necessary to improve the precision of the data generated by the direct solver. To

achieve this, one could consider further refining the spatial mesh and increasing the number of quadrature points.

4 FINAL CONSIDERATIONS

In this work, we have presented the application of the ANN-MoC method to solve three inverse transport problems for the characterization of a source of neutral particles. The problems consist of estimating the source parameters based on measurements of particle density at the domain boundaries. The method involves training an ANN with data generated from MoC solutions of the associated direct transport problem. An MLP neural network type was employed as a regression model to provide estimates of the source parameters.

After several numerical tests, we found that small MLPs could provide good estimates. Better results were obtained by preprocessing the input data with the Standard Scaler. A sensitivity test was also reported for each problem to study noise propagation. The training and validation results indicate the robustness of the network in providing accurate estimates based on density measurements.

Applying the ANN-MoC method to more realistic problems requires the use of an appropriate transport model and the effort should be focused on generating high-quality data (both for training and validation) for network training. Finally, the use of the proposed methodology for realistic problems depends on how good the direct transport model is for the intended application.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Universidade Federal do Rio Grande do Sul (UFRGS) for their support of this work through the provision of scholarships.

REFERENCES

Evans, L. (2010). *Partial Differential Equations*. (19th ed). American Mathematical Society.

Haykin, S. (2009). *Neural Networks and Learning Machines*. (3th ed). Pearson.

Hielscher, A. H., Alcouffe, R. E., Barbour, R. L. (1998). Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissues. *Physics in Medicine & Biology, 43*(5), 1285-1302. Recovered from: https://iopscience.iop.org/article/10.1088/0031-9155/43/5/017.

Kaipio, J., Somersalo, E. (2006). *Statistical and computational inverse problems*. (160 vol). Springer Science & Business Media.

Kim, K. W., Baek, S. W., Kim, M. Y., Ryou, H. S. (2004). Estimation of emissivities in a two-dimensional irregular geometry by inverse radiation analysis using hybrid genetic algorithm. *Journal of Quantitative Spectroscopy and Radiative Transfer, 87*(1), 1-14. Recovered from: https://doi.org/10.1016/j.jqsrt.2003.08.012.

Kingma, D. P., Ba, J. (2017). Adam: A Method for Stochastic Optimization. *arXiv, 4*, 1-15. Recovered from: https://doi.org/10.48550/arXiv.1412.6980.

Larsen, E. W., Thömmesand, G., Klar, A., Seaıd, M., Götz, T. (2002). Simplified P_N approximations to the equations of radiative heat transfer and applications. *Journal of Computational Physics, 183*(2), 652–675. Recovered from: https://doi.org/10.1006/jcph.2002.7210.

Lewis, E., Miller, W. (1984). *Computational methods of neutron transport*. (1th ed). Wiley.

Li, H. (1997). Inverse radiation problem in two-dimensional rectangular media. *Journal of thermophysics and heat transfer, 11*(4), 556-561. Recovered from: https://doi.org/10.2514/2.6279.

Mengüç, M., Manickavasagam, S. (1993). Inverse radiation problem in axisymmetric cylindrical scattering media. *Journal of Thermophysics and Heat Transfer, 7*(3), 479–486. Recovered from: https://doi.org/10.2514/3.443.

Modest, M. F. (2013). *Radiative heat transfer*. (3th ed). Elsevier Science.

Nocedal, J., Wright, S. J. (1999). *Numerical optimization*. Springer.

Qi, H., Ruan, L., Zhang, H., Wang, Y., Tan, H. (2007). Inverse radiation analysis of a one-dimensional participating slab by stochastic particle swarm optimizer algorithm. *International journal of thermal sciences, 46*(7), 649–661. Recovered from: https://doi.org/10.1016/j.ijthermalsci.2006.10.002.

Roman, N., Santos, P., Konzen, P. (2023). ANN-MoC method for inverse transient transport problems in one-dimensional geometry. *Latin-American Journal of Computing, 11*(2), 41-50.

Santos, P., Melo, G., Konzen, P. (2022). Rnas aplicadas a determinação e localização de fonte de partículas em problemas de transporte unidimensional. Em: , In *Anais do Encontro Nacional de Modelagem Computacional, Encontro de Ciência e Tecnologia de Materiais, Conferência Sul em Modelagem Computacional e Seminário e Workshop em Engenharia Oceânica*, Pelotas, RS, Brasil.

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

Tarvainen, T., Vauhkonen, M., Arridge, S. (2008). Gauss–Newton reconstruction method for optical tomography using the finite element solution of the radiative transfer equation. *Journal of Quantitative Spectroscopy and Radiative Transfer, 109*(17-18), 2767–2778. Recovered from: https://doi.org/10.1016/j.jqsrt.2008.08.006.

Wang, L., Wu, H. (2012). *Biomedical optics: principles and imaging*. Wiley.

Author contributions

1 – Nelson Garcia Roman (Corresponding Author)
Student, Master in Applied Mathematics
https://orcid.org/0009-0006-8794-9500 • ngroman1992@gmail.com
Contribution: Methodology; Production of results, Writing – Original Draft Preparation

2 – Pedro Costas dos Santos
Student, Undergraduate in Applied Mathematics
https://orcid.org/0009-0001-9927-2860 • pedro.costa4137@gmail.com
Contribution: Data preprocessing; Numerical tests

3 – Pedro Henrique de Almeida Konzen
Professor, Doctor in Applied Mathematics
https://orcid.org/0000-0002-0411-1563 • pedro.konzen@ufrgs.br
Contribution: Conceptualization, Methodology, Writing – Review

How to cite this article

Román, N. G., do Santos, P. C., & Konzen, P. H. de A. (2024). ANN-MoC method for the inverse problem of source characterization. *Ciência e Natura*, Santa Maria, v. 47, spe. 1, e89819. DOI: https://doi.org/10.5902/2179460X89819.