REVISTA Mundi
Engenharia, Tecnologia e Gestão
INSTITUTO FEDERAL Paraná

REDUCING THE SENSITIVITY OF ANN REGRESSION MODEL FOR PARAMETER SOURCE ESTIMATES

REDUCCIÓN DE LA SENSIBILIDAD DEL MODELO DE REGRESIÓN DE RNA PARA LA ESTIMACIÓN DE PARÁMETROS DE LA FUENTE

Nelson Garcia Roman[1]

Pedro Henrique de Almeida Konzen[2]

Abstract: Inverse problems in neutral particle transport have many important applications in engineering and in optical medicine. This work presents a study on data selection for reducing the sensitivity of regression models based on artificial neural networks (ANNs) for the estimation of source parameters. It addresses two selected sensitive problems consisting of estimating source parameters based on detector measurements of the particle density at the domain boundaries over time. Multilayer perceptron (MLP) type ANNs are employed as non-linear regression models from detector data to parameter estimations. The model sensitivity on noisy input data is analyzed by means of the mean absolute percentage error (MAPE) and the factor of noise propagation from input to the output estimates. The results indicate that the sensitivity of the ANN models decreases significantly as proper detector data over time are provided. The importance of ensuring adequate data selection for mitigating model sensitivity is highlighted. The study confirms the potential of ANN-based regression models in solving inverse problems on source characterization of neutral particle transport.

Keywords: Artificial neural network. Method of characteristics. Neutral particle transport. Inverse problem. Model sensitivity.

Resumen: Los problemas inversos en el transporte de partículas neutras tienen importantes aplicaciones en ingeniería y medicina óptica. Este trabajo presenta un estudio sobre la selección de datos para reducir la sensibilidad de modelos de regresión basados en redes neuronales artificiales (RNA) en la estimación de parámetros de la fuente. Se analizan dos problemas sensibles que consisten en estimar dichos parámetros a partir de mediciones de la densidad de partículas en los límites del dominio a lo largo del tiempo. Se emplean redes neuronales de tipo perceptrón multicapa (PMC) como modelos de regresión no lineales. La sensibilidad frente a datos de entrada con ruido se evalúa mediante el error porcentual absoluto medio (EPAM) y el factor de propagación del ruido desde la entrada hasta las estimaciones. Los resultados muestran que la sensibilidad de los modelos disminuye significativamente cuando se seleccionan adecuadamente los datos de los detectores. El estudio destaca la importancia de una correcta selección de datos y confirma el potencial de los modelos basados en RNA para resolver problemas inversos de caracterización de fuentes en el transporte de partículas neutras.

Palabras clave: Redes neuronales artificiales. Método de las características. Transporte de partículas neutras. Problemas inversos. Modelo de sensibilidad.

[1] PPGMAp, IME, UFRGS, ngroman1992@gmail.com
[2] PPGMAp, IME, UFRGS, pedro.konzen@ufrgs.br

1 INTRODUCTION

Inverse problems of neutral particle transport have important applications in engineering and in optical medicine. In radiative transport (MODEST, 2023), they are essential for developing quality control protocols in high-temperature industrial processes. In nuclear transport (LEWIS, 1984; STACEY, 2007), they are fundamental to ensure safety. In optical medicine (WANG, 2012), they are found as applications of radiative transport (e.g., computed tomography) and neutron transport.

Since the 90s, deep learning has established itself as a fundamental tool in advanced research and modern technology. This approach, which involves multiple layers of compositions of nonlinear transformations, has demonstrated a remarkable ability to approximate highly nonlinear functions, making it particularly useful for solving complex problems, such as inverse problems in neutral particle transport. Deep neural networks are recognized for their universal approximation capability, a property supported by many works (BARRON, 1993; CYBENKO,1989; HORNIK,1989; YAROTSKY, 2017).

Artificial neural networks (ANNs) are machine learning approaches that can establish a mapping from input to output data, being effective as non-linear regression models. Examples of applications in the context of particle transport can be found in the specialized literature. In the work of Verma *et al.* (2023), an ANN model is used to predict coolant activity and defect states, while in the work of Oliveira *et al.* (2010), a committee of ANNs is employed to estimate radiative transfer parameters, such as diffuse reflectivity and scattering albedo.

First published in the work of Raissi *et al.* (2019), the physics-informed neural networks (PINNs) are a related deep learning approach that has been successfully applied in the simulation of direct and inverse problems, including applications in neutral particle transport. Mishra *et al.* (2021) developed an algorithm based on PINNs to simulate this phenomenon, establishing generalized error bounds. Other studies, such as those by Huhn *et al.* (2023) and Riganti *et al*. (2023), extended the use of PINNs to various radiative transport configurations, demonstrating their versatility and effectiveness in this field.

In our previous works (ROMAN, 2024; ROMAN, 2025), we have already explored ANNs as non-linear regression models for source characterization and absorption coefficient estimates in the stationary transport equation. However, the models studied were very sensitive to noisy data in some applications. By assuming a transient transport model, the present work is a study on data selection to mitigate the sensitivity in ANN-based regression models for source characterization in neutral particle transport problems. Two particularly sensitive test cases are studied, in which the source parameters are estimated from particle density measurements (detector data) at the domain boundaries over time. These problems were previously studied in Roman (2025) and are revisited here to improve the generalization of the ANN regression models for noise levels greater than 3%, a challenge identified in that work.

We assume the particle transport is modeled using the linear Boltzmann equation in a participating medium with isotropic scattering (LEWIS, 1984; MODEST, 2023)

$$\forall\mu\colon \frac{1}{c}\frac{\partial}{\partial t}I(t,x,\mu) + \mu\frac{\partial}{\partial x}I(t,x,\mu) + \sigma_t I(t,x,\mu) = \sigma_s\Psi(t,x) + q_{\alpha,\beta}(t,x,\mu), (t,x) \in (0,t_f] \times D, \quad (1a)$$

$$\forall\mu\colon I(0,x,\mu) = I_0(x,\mu), x \in D, \quad (1b)$$

$$\forall\mu > 0\colon I(t,a,\mu) = I_{in,a}(t,\mu), t \in (0,t_f], \quad (1c)$$

$$\forall\mu > 0\colon I(t,a,\mu) = I_{in,a}(t,\mu), t \in (0,t_f]. \quad (1d)$$

Considering a radiative transport problem, $I(t,x,\mu)$ (W/sr) denotes the radiation intensity at time interval $0 \leq t \leq t_f$ (s) at point $x \in D = [a,b]$ (m), and in the direction $-1 < \mu < 1$ (sr), $\mu \neq 0$. The average speed of light in the medium is denoted by c (m/s). The total absorption coefficient is given by $\sigma_t = \kappa + \sigma_s$ $(2/m)$, where κ $(1/m)$ and σ_s $(1/m)$ are, respectively, the absorption and scattering coefficients. The source in the domain is denoted by $q_{\alpha,\beta}(t,x,\mu)$ $(W/(msr))$ and characterized by unknown parameters α and β. At the boundaries, the known source of particles are $I_{in,a}(t,\mu), I_{in,b}(t,\mu)$ (W/sr). At $t = 0$, the initial condition $I = I_0(x,\mu)$ (W/sr) is assumed. The particle density Ψ (W) is defined as

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

2 ANN REGRESSION MODEL

A regression model aims to capture the functional relationships between input and output variables, allowing the prediction or estimation of unknown values based on observable information. In this work, an artificial neural network of the type multilayer perceptron (MLP) (HAYKIN, 2009) is employed to establish the mapping between particle density measurements (detector data) and the parameters that characterize the particle sources. The MLP acts as a nonlinear regression model that associates the detector measurements of the particle density values at the domain boundaries over time with the source parameters, i.e., input data are $\boldsymbol{d} = (\Psi(\boldsymbol{t}_d, 0), \Psi(\boldsymbol{t}_d, 1))$ and the outputs are α and β, where $\boldsymbol{t}_d$ is a vector of selected discrete time instants of detector measurements.

Based on a supervised training approach, a calibration dataset $\{\boldsymbol{\Psi}^{(s)}, \boldsymbol{\gamma}^{(s)}\}_{s=1}^{n_{train}}$ must be provided, where $\boldsymbol{\Psi}^{(s)} = (\boldsymbol{d}_0, \boldsymbol{d}_1)$ represents the detector measurements and $\boldsymbol{\gamma}^{(s)} = \left(\alpha^{(s)}, \beta^{(s)}\right)$ corresponds to the source parameters. The measurements are expressed as

$$\boldsymbol{d}_0 = \left(\Psi(t_{d,1}, 0), \Psi(t_{d,2}, 0), \dots, \Psi(t_{d,n_d}, 0)\right) \tag{3a}$$

$$\boldsymbol{d}_1 = \left(\Psi(t_{d,1}, 1), \Psi(t_{d,2}, 1), \dots, \Psi(t_{d,n_d}\ 1)\right) \tag{3b}$$

where $\boldsymbol{t}_d = \left(t_{d,1},\ t_{d,2}, \dots, t_{d,n_d}\right)$, represents the n_d measurement time instants of the particle density $\Psi\left(t_{d,i}, 0\right)$ and $\Psi\left(t_{d,i}, 1\right)$ obtained at the domain boundaries for $i = 1, 2, \dots, n_d$.

Alternatively to experimental data, the training set was generated synthetically for several parameter values α and β by solving the related direct problem (1) by our implementation of a method of characteristics solver (MoC).

The implementation details of our MoC solver for the neutral particle transport problem (1) is available in our previous work (ROMAN, 2024) and in the master's dissertation of Roman (2025). For the completion of the text, the solver is summarized here. The MoC formulation is obtained by approximating the problem (1) using the discrete ordinates method (DOM), followed by temporal discretization through the implicit Euler method (IE). The resulting system of ordinary differential equations is iteratively decoupled using a source iteration scheme (SI). The final MoC formulation is obtained after a suitable change of variables, and it provides the solution $I(x,\mu) = I(x_0 + r\mu)$ parameterized by $r \in R$ in the direction μ. The resulting numerical scheme depends on the parameters n_x (the number of cells in the computational mesh), N (the number of discrete ordinates given by the Gauss quadrature), and h_t (the time step). The study on these numerical parameters convergence to produce solutions with precision of up to 3 significant digits is discussed in our above cited publications.

A MLP of architecture $2n_d - n_h \times n_n - 2$ ($2n_d$ inputs, n_h hidden layers each with n_n neurons, 2 outputs) is assumed and denoted as

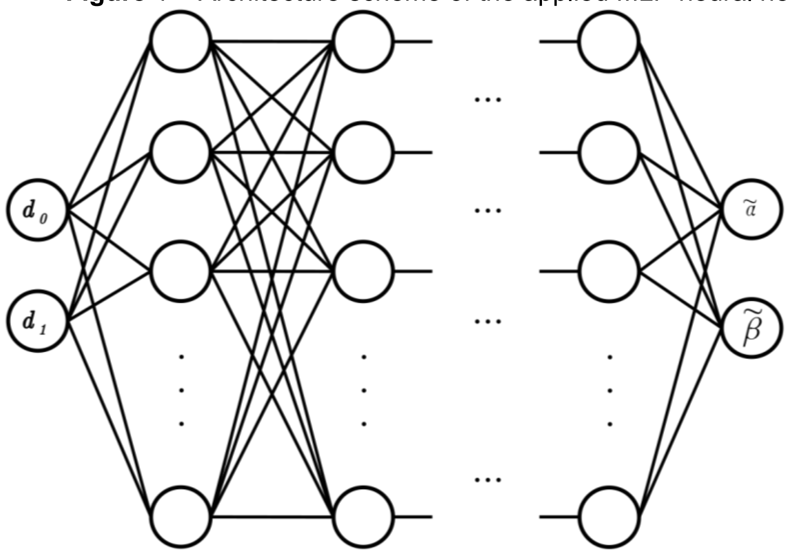
$$\widetilde{\boldsymbol{\gamma}} = N\left(\boldsymbol{\Psi}; \left\{\left(\boldsymbol{W}^{(l)}, \boldsymbol{b}^{(l)}, \boldsymbol{f}^{(l)}\right)\right\}_{l=1}^{n_h+1}\right), \quad (4)$$

where $\boldsymbol{W}^{(l)}$, $\boldsymbol{b}^{(l)}$ and $\boldsymbol{f}^{(l)}$ denote the weights, the biases (network trainable parameters), and the activation function in the l-th layer, respectively. The activation functions and the parameters n_h and n_n are chosen *a priori*. Denoting a sample input as $\boldsymbol{y}^{(0)} = \boldsymbol{\Psi}^{(s)}$ (s denotes the sample index), the forward propagation through the network is the following iteration of compositions

$$\boldsymbol{y}^{(l)} = \boldsymbol{f}^{(l)}\left(\boldsymbol{W}^{(l)}\boldsymbol{y}^{(l-1)} + \boldsymbol{b}^{(l)}\right) \quad (5)$$

and the final output corresponds to the estimated source parameters, *i,e*.

Figure 1 – Architecture scheme of the applied MLP neural networks.



Source: From authors.

$\boldsymbol{y}^{(n_h+1)} = \left(\tilde{\alpha}, \tilde{\beta}\right)$.

The training of the ANN model consists of solving the following minimization problem

$$\min_{\{(\boldsymbol{W}^{(l)},\boldsymbol{b}^{(l)},\boldsymbol{f}^{(l)})\}_{l=1}^{n_h+1}} \left(\varepsilon := \frac{1}{n_s} \sum_{s=1}^{n_s} \left\| \tilde{\boldsymbol{\gamma}}^{(s)} - \boldsymbol{\gamma}^{(s)} \right\|^2 \right) \tag{6}$$

where ε presents the loss function, which is the mean squared error (MSE) of the expected values $\boldsymbol{\gamma}^{(s)}$ and the estimated values $\tilde{\boldsymbol{\gamma}}^{(s)}$of the source parameters of all the training samples $s = 1, 2, \dots, n_s$. Following the classical backpropagation algorithm (HAYKIN, 2009), the training procedure involves calibrating the network parameters (the weights and biases) to minimize the ε. At each training iteration (called an epoch), the Adam optimizer (KINGMA, 2017) has been applied. The implementations were carried out in Python using the PyTorch machine learning package (PASZKE *et at.*, 2016). The specific code parameters (including the MLP architecture) are problem-dependent and are presented later, along with the test cases discussed below.

3 RESULTS AND DISCUSSIONS

With the aim of studying the generalization of the proposed ANN models in the context of noisy input data, two particular selected problems are now addressed. The generalization is analyzed by measuring the sensitivity of the regression models in relation to the level of noise considered in the detectors' data.

In all cases, the transport model (1) is assumed with the homogeneous medium properties $\kappa = 0.5$ and $\sigma_s = 0.5$. The boundary conditions are taken as $I(t, 0, \mu) = 0$, for all $\mu > 0$, and $I(t, 1, \mu) = 0$, for all $\mu < 0$ (vacuum boundary conditions with no influxes), while the initial condition is $I(0, x, \mu) = 0$ for $x > 0$. The detector measurements correspond to the particle density at the boundaries of the domain $D = [0, 1]$ at n_d discrete times $t_{d,k}$.

For the ANN models, the training and test data has been manufactured by numerically solving the transport model (1) with our implementation of the MoC solver. Previous numerical experiments indicated that the choices of $n_x = N =$

100 and $h_t = 0.01$ were sufficient to obtain solutions with 4 digits of precision for the particle density (ROMAN, 2025).

For each test case, $n_{train} = 100$ training samples and $n_{test} = 200$ test samples have been generated for different values of α and β. The training data have been produced for the expected source parameters $\alpha^{(i)} = 0.1 + (i - 1)h$, $\beta^{(j)} = 0.1 + (j - 1)h$, with $h = 0.1$, $i, j = 1, 2, \ldots, 10$. The test set was generated with $n_{test} = 200$ samples for randomly uniformly distributed values of the source parameters in the range $0.1 \leq \alpha, \beta \leq 1.0$.

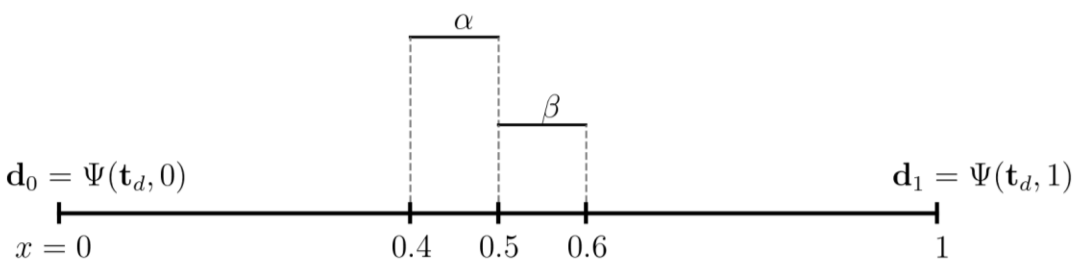
In order to improve the training procedure, the inputs of all samples were previously preconditioned using the standard scaler (SCIKIT-LEARN DEVELOPERS, 2025). The training stop criterion was assumed $\varepsilon < 10^{-5}$. After many empirical numerical tests, the MLP architecture has been fixed to $2n_d - 4 \times 40 - 2$ ($n_h = 4$ hidden layers, each with $n_n = 40$ neuron units) for all the following inverse problems to be addressed. See Roman (2025), for more details on the choice of the network architecture.

3.1 Inverse Problem 1

In this first inverse problem, the source of the transport problem is assumed as $q_{\alpha,\beta}(x) = \alpha$, for $0.4 \leq x \leq 0.5$, $q_{\alpha,\beta}(x) = \beta$, for $0.5 < x \leq 0.6$, and $q_{\alpha,\beta} = 0$ otherwise. *I.e.*, the source is characterized by two fixed localized pulse functions with intensities α and β (see Figure 2).

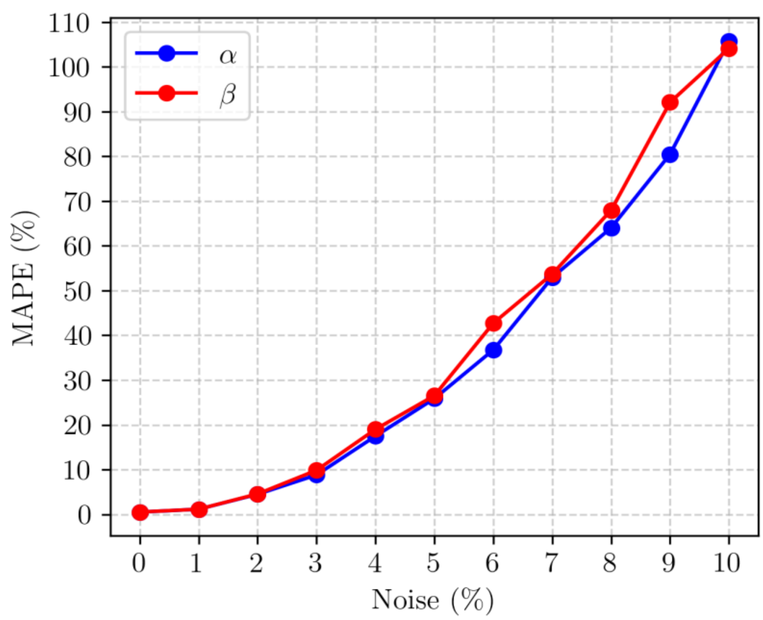
This problem has been previously addressed in the paper of Roman *et al.* (2025) by assuming a stationary transport model. With this strong assumption, it was possible to achieve good parameter estimations only for data with a noise level of up to 4%. Figure 3 shows the curves of the mean absolute percentage error (MAPE) values corresponding to the estimations of α and β for different noise levels on the input data. In that work, the noise propagation factor through the MLP network is defined as the slope of the least square line of noise level by the MAPE. A factor of 1 means the MAPE of the output estimations is at the same level as the noise in the input data. A factor of less than 1, indicates the ANN regression model could learn to filter the noise. In contrast, as greater than 1 is the factor, more sensitive in the ANN model to the level of noise. In this previous work, the sensitivity factor was 10.3 in the estimations of the α parameter, and 10.8 in the estimations of the β parameter.

Figure 2 – The source for the inverse problem 1, two pulse functions with intensities α and β.



Source: From the authors.

Figure 3 – MAPE *vs.* noise for the ANN model trained from stationary data.

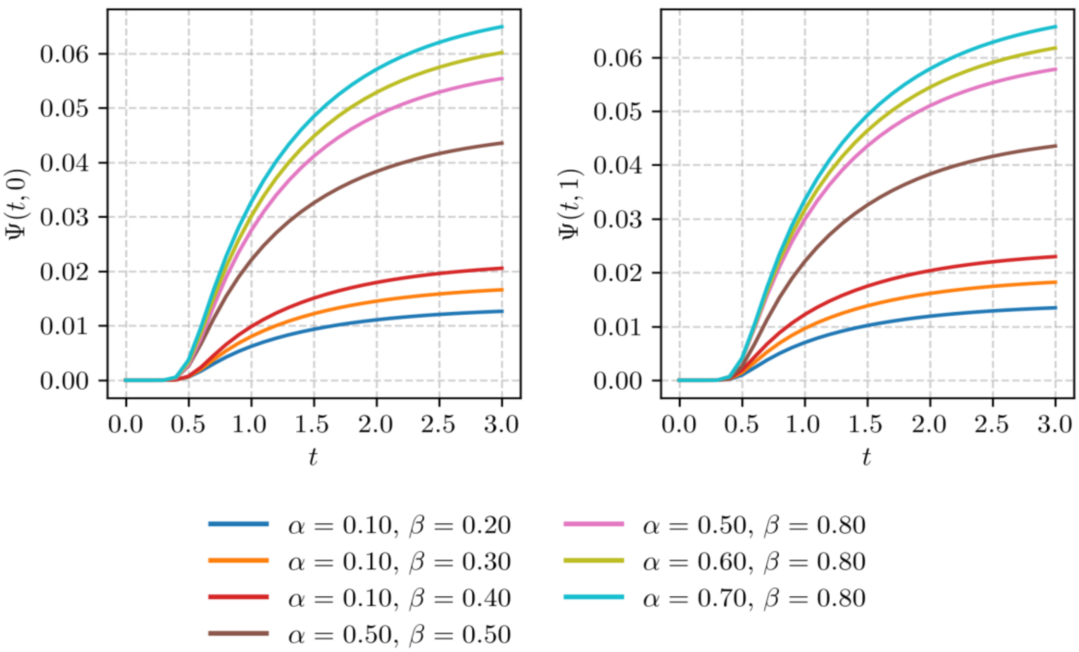


Source: From authors.

Since that regression model was particularly sensitive to noise in this problem setup, a reformulation of the inverse problem is studied here. The idea is that additional dynamic information about the evolution of the particle density, may mitigate the sensitivity of the model to the data noise. By considering measurements at the domain boundaries over time, the hypothesis is that an ANN regression model can filter the noise and enhance its generalization capability, as it is now studied in the sequence.

The inverse problem is well-posted when detector data is taken over time. Figure 4 presents the MoC direct solution curves of $\Psi(t,0)$ (right) and $\Psi(t,1)$ (left) for various combinations of the parameters α and β. We observe that the curves clearly differ for different values of the parameters. For example, in the case of $t = 3$, we observe that the solution for equal values of α varying β, and for equal

Figure 4 – Detectors curves for several combinations of α and β. Left: Solution curves for $\Psi(t,0)$. Right: Solution curves for $\Psi(t,1)$. Centered below: Legend for solution curves.



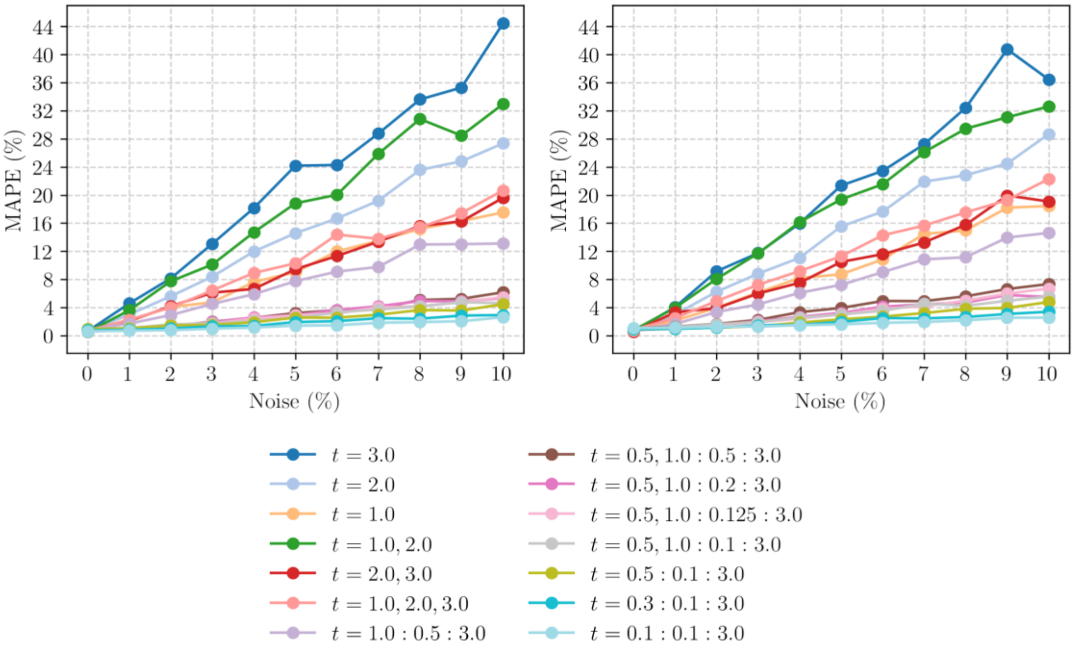
Source: From authors.

values of β varying α, show significant differences in the estimates.

The hypothesis that the ANN models can deal with noisy data if detector data over time is taken, can be confirmed by analyzing Figure 5. The figure presents the MAPE of the parameters estimations ($\tilde{\alpha}$ and $\tilde{\beta}$ for the test set) in relation to the level of noise in the input data (detectors measurements) for several configurations setups. It can be observed that as more detector measurements are considered over time, the less is the noise propagation in the trained ANN regression models. It is notable, that even for the level noise of *10%*, estimations with a MAPE of less than 4% could be reached by taking detectors' data at times $t = 0.1, 0.2, \ldots, 3.0$. In the legend of the Figure 5, we assume the two-dot notation, where, for instance, $t = 0.1 : 0.1 : 3.0$ denotes the sequence of discrete times $t = 0.1, 0.2, \ldots, 3.0$.

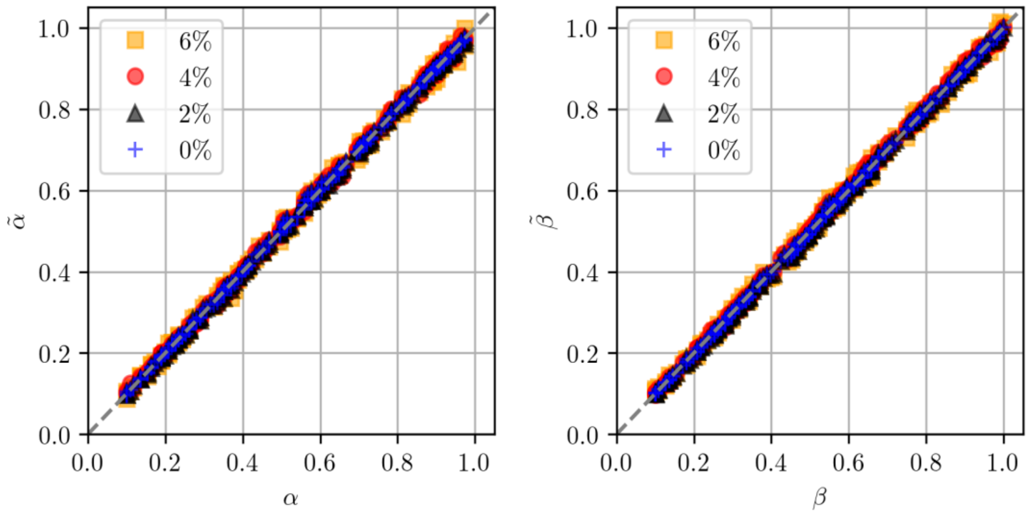
Figure 6 shows the expected *versus* estimated values of α (right) and β (left) from the test set with detector data taken in discrete times $t = 0.1 : 0.1 : 3.0$ and for noise levels of 0%, 2%, 4% and 6%. In the figures, the identity line is represented as a dashed line for reference. With this data setup, the sensitivity factor of the trained ANN model was of 1.3 for α and 1.4 for β. We observe the absence of outliers, which also indicates good generalization of the ANN model.

Figure 5 – MAPE *vs*. noise levels for several data configurations. Left: α. Right: β. Centered below: Legend.



Source: From authors.

Figure 6 – Expected *versus* estimated parameters. Left: α. Right: β.

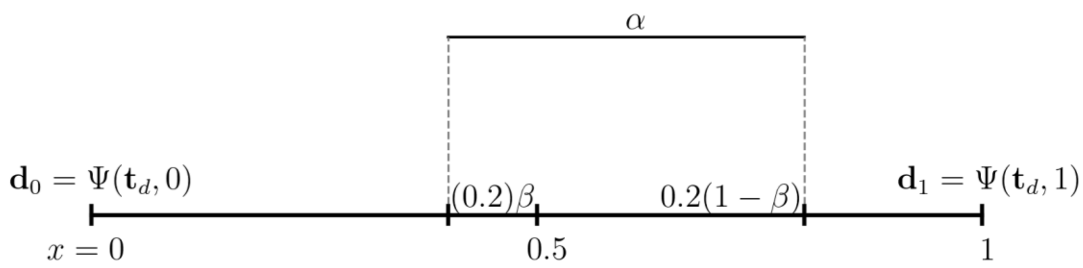


Source: From authors.

3.2 Inverse Problem 2

In this second inverse problem, the source of the transport problem is assumed as $q_{\alpha,\beta}(x) = \alpha$, for $0.5 - \beta \cdot 0.2 \leq x \leq 0.5 + (1 - \beta) \cdot 0.2$ and $q_{\alpha,\beta}(x) = 0$ otherwise. *I.e.*, the source is characterized by a step-like function with intensity α and displacement β (see Figure 7).

Figure 7 – The source for the inverse problem, a step source with intensity α and 0.2β displaced from the center.

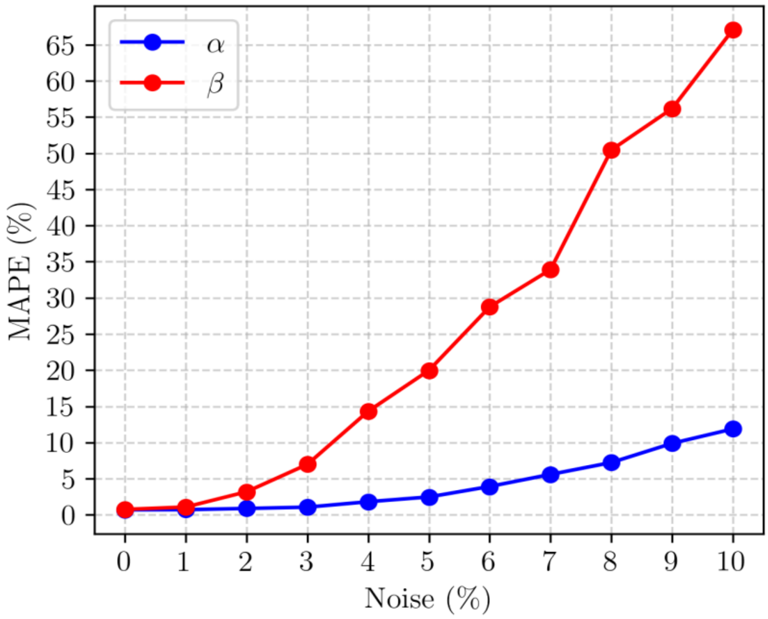


Source: From the authors.

This problem was also previously addressed in the work of Roman *et al.* (2025), assuming a stationary transport model. With this assumption, good estimates of the parameter α were obtained for data with noise levels up to 10%, while for the parameter β good estimates were only achieved for noise levels up to 4%. Figure 8 shows the MAPE values curves corresponding to the estimates of α and β for different noise levels. The trained ANN regression model had an sensitivity factor of 1.1 in the estimates of α, while a factor of 6.9 in the estimates of β, *i.e.* the model was much more sensible to the noise for the estimates of the source displacement parameter.

In order to achieve a ANN regression model that can better generalize, the approach based on the previous inverse problem (Subsection 3.1) can be applied. The model now takes detector data at discrete times $t = 0.1 : 0.1 : 3.0$and it learns from data generated by the transient transport model.

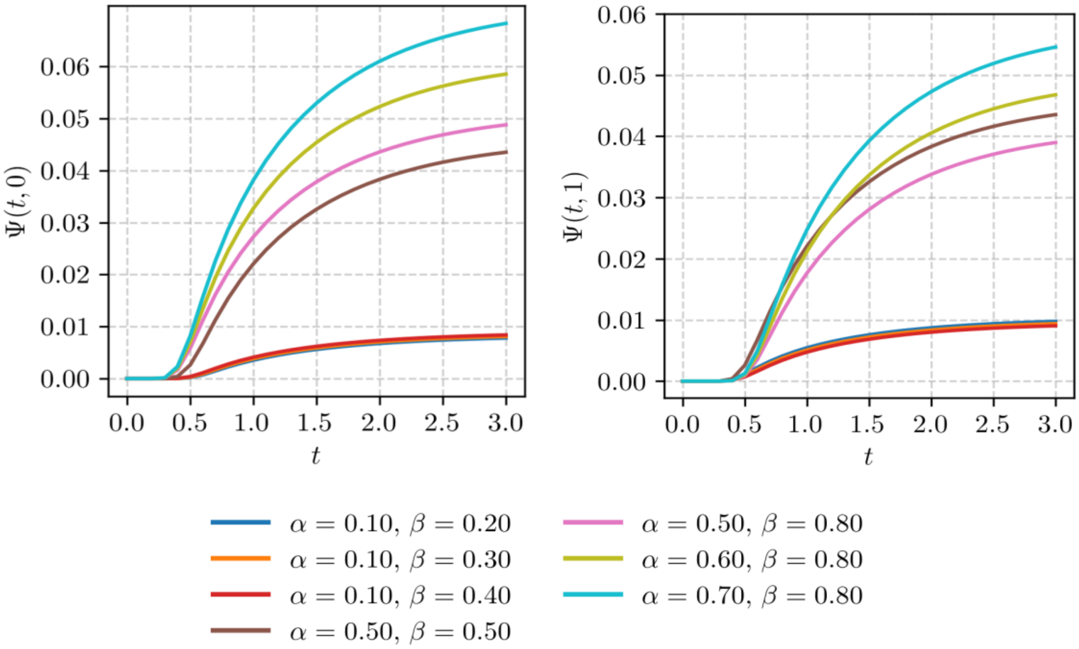
Figure 8 – MAPE *vs.* noise for the ANN model trained from stationary data.



Source: From authors.

As in the previous test problem, this second inverse problem is well-posed when detector data are taken over time. Figure 9 presents the MoC direct solution curves of $\Psi(t,0)$ (right) and $\Psi(t,1)$ (left) for various combinations of the parameters α and β. It can be observed that the curves differ for different parameter values. When α is kept constant and β is changed, the curves are near each other, which explain the large sensitivity of the regression models on the parameter β.

Figure 9 – Detectors curves for several combinations of α and β. Left: Solution curves for $\Psi(t,0)$. Right: Solution curves for $\Psi(t,1)$. Centered below: Legend for solution curves.

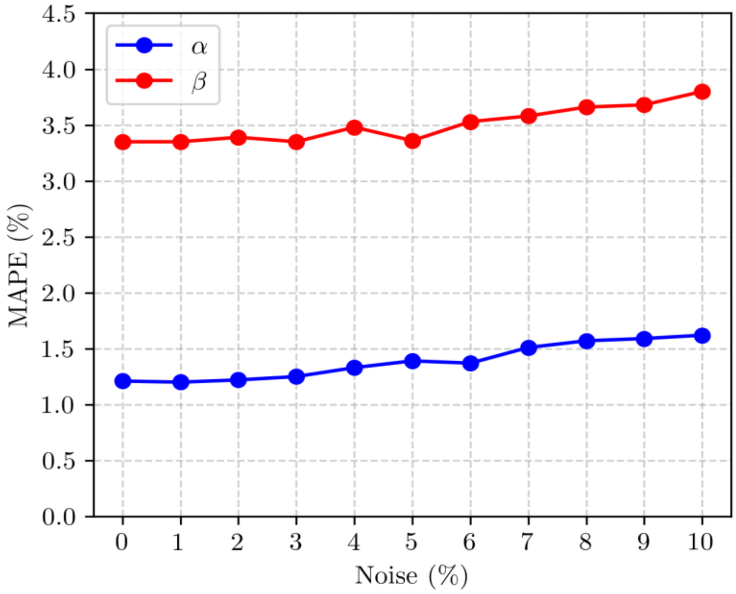


Source: From authors.

Figure 10 presents the MAPE of the estimates of the parameters α and β from the test set as a function of the noise level in the input data (detector measurements) for the configuration $t = 0.1{:}\,0.1{:}\,3.0$. It is noteworthy that with a noise level of 10%, the ANN estimates achieved a MAPE lower than 2% for α and 4% for β. With this data configuration, the sensitivity factor of the trained ANN model was 0.05 for both α and β.

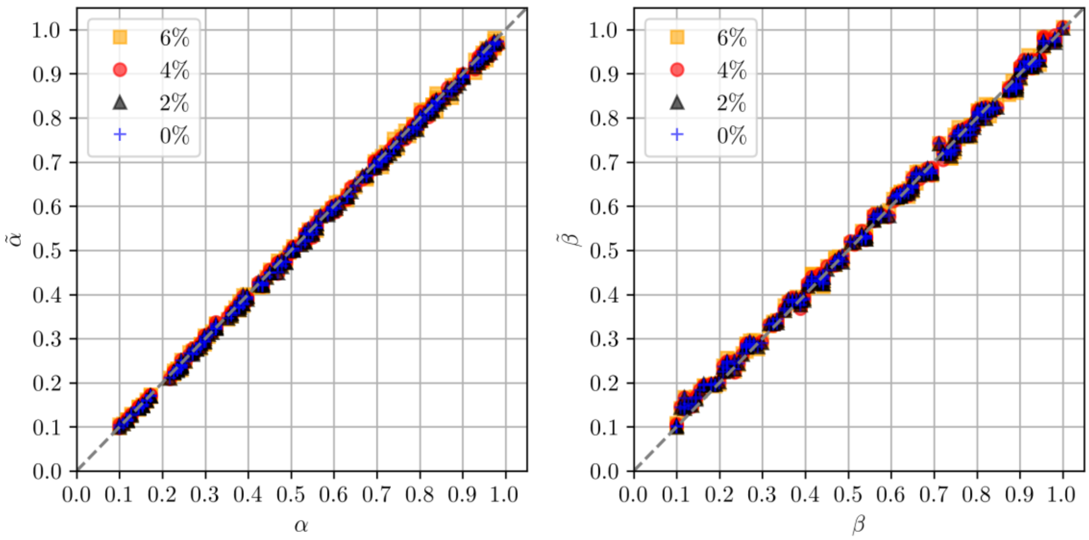
Figure 11 shows the expected *versus* estimated values of α (right) and β (left) from the test set, using detector data taken at discrete times $t = 0.1 : 0.1 : 3.0$ and for noise levels of 0%, 2%, 4%, and 6%. In the figures, the identity line is represented by a dashed line as a reference. This indicates that the ANN model could filter out the noise in the input data. The absence of outliers is observed, which also indicates the good generalization of the ANN model.

Figure 10 – MAPE *vs.* Noise for the ANN model trained with transient data.



Source: From authors.

Figure 11 – Expected *versus* estimated parameters for several noise levels. Left: α. Right: β.



Source: From authors.

4 FINAL CONSIDERATIONS

This work addressed the reduction of sensitivity in regression models based on ANNs for the source characterization of neutral particle transport problems. It highlights the importance of proper data selection in the regression model setup. Model sensitivity on noisy detector data could be mitigated by availability of transient data over time.

The methodology was tested on two particular sensitive problems. In the first, the particle source is given as a two localized pulse functions with intensities to be estimated. Proper data selection could provide an ANN regression model with a sensitivity factor of about 1.4, which indicates that the source parameters can be estimated with a MAPE proportional to the level of noise assumed in the detectors.

The second problem considered the source as a pulse function with intensity and displacement to be estimated. With proper data selection, the ANN model could very well generalize its estimations and have almost filter-out the noise in the detectors data.

The presented methodology indicates that ANN regression models can be very effective in inverse source characterization on transport problems, with potential applications in engineering and medicine. Future research could focus on optimizing the use of advanced neural networks, refining training and validation datasets, and investigating the model sensitivity on normal distributed noisy data. Expanding the method to multidimensional geometry problems is also feasible, further improving its accuracy and applicability in more complex scenarios.

4 ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

REFERENCES

BARRON, A. R. Universal approximation bounds for superpositions of a sigmoidal function. **IEEE Transactions on Information Theory**, v. 39, n. 3, p. 930–945, 1993. DOI: https://doi.org/10.1109/18.256500.

CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of Control, Signals and Systems**, v. 2, n. 4, p. 303–314, 1989. DOI: https://doi.org/10.1007/BF02551274

HAYKIN, S. **Neural networks and learning machines**. 3. ed. Upper Saddle River: Pearson, 2009. ISBN 978-0-13-147139-9
HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural Networks**, v. 2, n. 5, p. 359–366, 1989. DOI: https://doi.org/10.1016/0893-6080(89)90020-8.
HUHN, Q. A.; TANO, M. E.; RAGUSA, J. C. Physics-informed neural network with Fourier features for radiation transport in heterogeneous media. **Nuclear Science and Engineering**, v. 197, n. 9, p. 2484–2497, 2023. DOI: https://doi.org/10.1080/00295639.2023.2184194.
KINGMA, D. P.; BA, J. **Adam: A method for stochastic optimization**. 2017. Disponível em: https://arxiv.org/abs/1412.6980
LEWIS, E.; MILLER, W. **Computational methods of neutron transport**. 1. ed. New York: Wiley, 1984. ISBN 978-0-471-09245-2.
MODEST, M. F. **Radiative heat transfer**. 3. ed. Oxford: Elsevier Science, 2013. ISBN 978-0-12-386944-9.
MISHRA, S.; MOLINARO, R. Physics-informed neural networks for simulating radiative transfer. **Journal of Quantitative Spectroscopy and Radiative Transfer**, v. 270, p. 107705, 2021. DOI: https://doi.org/10.1016/j.jqsrt.2021.107705.
OLIVEIRA, R.; ACEVEDO, N.; SILVA NETO, A.; BIONDI NETO, L. Aplicação de um comitê de redes neurais artificiais para a solução de problemas inversos em transferência radiativa. **Trends in Computational and Applied Mathematics**, v. 11, n. 2, p. 171–182, 2010. DOI: https://doi.org/10.5540/tema.2010.011.02.0171.
PASZKE, A.; GROSS, S.; CHINTALA, S.; CHANAN, G. **PyTorch**. Disponível em: <https://pytorch.org/>. Acesso em: 15 out. 2025.
RAISSI, M.; PERDIKARIS, P.; KARNIADAKIS, G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. **Journal of Computational Physics**, v. 378, p. 686–707, 2019. DOI: https://doi.org/10.1016/j.jcp.2018.10.045.
RIGANTI, R.; NEGRO, L. D. Auxiliary physics-informed neural networks for forward, inverse, and coupled radiative transfer problems. **Applied Physics Letters**, v. 123, n. 17, 2023. DOI: https://doi.org/10.1063/5.0167155.
ROMAN, N. G. **Modelos de Regressão de Aprendizagem Profunda em Problemas Inversos de Transporte de Partículas Neutras**. 2025. 92 pag. Master's Dissertation (Master in Applied Mathematics) — Universidade Federal do Rio Grande do Sul, Porto Alegre, 2025. Available at: https://lume.ufrgs.br/bitstream/handle/10183/295565/001280963.pdf?sequence=1.
ROMAN, N. G.; SANTOS, P. C.; KONZEN, P. H. A. ANN-MoC method for inverse transient transport problems in one-dimensional geometry. **Latin American Journal of Computing**, p. 41–50, 2024. DOI: https://doi.org/10.5281/zenodo.12191947.
ROMAN, N. G.; SANTOS, P. C.; KONZEN, P. H. A. ANN-MoC method for the inverse problem of source characterization. **Ciência e Natura**, v. 47, n. 1, e89819, 2025. DOI: https://doi.org/10.5902/2179460X89819.

STACEY, W. **Nuclear reactor physics**. 2. ed. Weinheim: Wiley-VCH, 2007. ISBN 978-3-527-61105-8.

SCIKIT-LEARN DEVELOPERS. **StandardScaler**. Available at: https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html. Acesso em: 11 set. 2025.

VERMA, L.; KREMER, F.; CHEVALIER-JABET, K. Defective PWR fuel rods detection and characterization using an artificial neural network. **Progress in Nuclear Energy**, v. 160, p. 104686, 2023. DOI: https://doi.org/10.1016/j.pnucene.2023.104686.

WANG, L.; WU, H. **Biomedical optics: principles and imaging**. 1. ed. Hoboken: Wiley, 2012. ISBN 978-0-470-17700-6.

YAROTSKY, D. Error bounds for approximations with deep ReLU networks. **Neural Networks**, v. 94, p. 103–114, 2017. DOI: https://doi.org/10.1016/j.neunet.2017.07.002.