Enhancing Neural Network Performance in Skin-Pass Mill Modeling through Generalized Extremal Optimization

Italo Pinto Rodrigues [*,**] **Gabriel Alberto Rodrigues** [**]
Bruno Lima dos Santos [**]

[*] *Engenharia e Tecnologia Espaciais,*
Coordenação de Ensino, Pesquisa e Extensão,
Instituto Nacional de Pesquisas Espaciais (INPE), SP
(e-mail: italoprodrigues@gmail.com).

[**] *Faculdade de Engenharia Elétrica,*
Centro Universitário de Volta Redonda (UniFOA), RJ
(e-mail: {gabriel.alberto, bruno.santos1 }@unifoa.edu.br).

Abstract: This study is justified by the need to precisely adjust controllers in skin-pass mills to ensure the quality of steel sheets, given the complexities of the rolling process. Traditional models fail to incorporate the historical wear of the mills, thus data-driven modeling is proposed as a viable approach to accurately represent these conditions and optimize production quality. The aim is to optimize the architectures of Artificial Neural Networks (ANNs) using the Generalized Extremal Optimization (GEO) algorithm, which stands out for its ability to adjust network architectures with a single free variable, simplifying configuration and enhancing computational efficiency. The methodology involved using the GEO algorithm to balance error minimization and computational cost, suitable for implementations in Programmable Logic Controllers (PLCs) with limited resources. As a significant result, the optimized ANN architecture achieved a Mean Relative Squared Error (MRSE) of 0.8418%, demonstrating an efficient solution given the application's constraints.

Resumo: Este estudo justifica-se pela necessidade de ajustar precisamente controladores em laminadores de encruamento para garantir a qualidade das chapas de aço, enfrentando as complexidades do processo de laminação. Dado que modelos tradicionais falham em incorporar o desgaste histórico dos moinhos, a modelagem orientada a dados surge como uma abordagem viável para representar essas condições com precisão e otimizar a qualidade da produção. O objetivo deste trabalho é resolver a otimização de arquiteturas de Redes Neurais Artificiais (ANNs) utilizando o algoritmo de Otimização Extrema Generalizada (GEO), destacando-se pela capacidade de ajustar arquiteturas de rede com uma única variável livre, simplificando a configuração e aumentando a eficiência computacional. A metodologia aplicada envolveu o uso do algoritmo GEO para equilibrar a minimização de erros e o custo computacional, adequado para implementações em Controladores Lógicos Programáveis (PLCs) com recursos limitados. Como resultado expressivo, a arquitetura de ANN otimizada alcançou um Erro Quadrático Médio Relativo (MRSE) de 0.8418%, demonstrando uma solução eficiente em termos das restrições impostas pela aplicação..

Keywords: Optimization; Artificial Neural Network; Generalized Extremal Optimization (GEO); Skin-pass mill; Industry 4.0

Palavras-chaves: Otimização; Redes Neurais Artificiais; Otimização Extrema Generalizada (GEO); Laminador de Encruamento, Indústria 4.0

1. INTRODUCTION

Accurate controller adjustment is essential for ensuring the quality finish of steel sheets in the skin-pass mill. The effectiveness of these controllers depends heavily on representing real-world conditions accurately, as mismatches can compromise the steel's quality. Seo et al. (2020) highlight the critical nature of proper tuning, without which the product's integrity is at risk.

The complex nature of the steel milling process introduces significant modeling challenges, involving various process variables, non-linearities, and time delays, as noted by Alves et al. (2012). These complexities are compounded by operational conditions and mill speed, which substantially affect the system's behavior. Over time, if a controller operates on a system that has strayed from its initial settings, the quality of the steel sheet may degrade. Shen et al. (2022) further discuss how traditional physical models fail to incorporate the wear history of mills, which is crucial after multiple material passes.

To counter these modeling challenges and to accurately depict the mill, data-driven modeling—or empirical modeling—has been identified as a viable approach. This method, essential for maintaining steel sheet quality and preventing financial losses, requires operational system data for the identification process. The significance of this approach is supported by the work of Shi et al. (2022) and Shen et al. (2022), ensuring that models accurately reflect the mill's current state and support optimal production quality.

Artificial Neural Networks (ANNs) provide a powerful tool for this kind of identification due to their ability to effectively map relationships between various inputs and outputs. The application of ANNs, as demonstrated in studies by Colla (2022) and Shi et al. (2022), is crucial for reducing losses and enhancing the predictive accuracy required for high-quality outputs. These advanced computational techniques assist in modeling the intricate interactions within the milling process.

The primary objective of this study is to develop an optimized Artificial Neural Network (ANN) architecture using the Generalized Extremal Optimization (GEO) algorithm, and to contrast the results with the previous empirical efforts by Rodrigues et al. (2023) to determine the ANN architecture. This approach seeks to demonstrate the advantages of GEO over traditional optimization methods. A key benefit of the Generalized Extremal Optimization algorithm is that it has only one free parameter, or hyperparameter, which simplifies its configuration. This comparison aims to showcase the effectiveness and precision of the GEO algorithm in refining ANN configurations, providing a more streamlined and accurate method for establishing the optimal network structure for specific applications.

This paper introduces the pioneering application of the (GEO) algorithm in the steel industry. This marks the first utilization of GEO within this sector, expanding its application beyond the aerospace field where it has previously shown substantial benefits (Chagas et al., 2014).

2. PROBLEM STATEMENT

In the field of skin-pass mills, the need for precise process identification has led researchers to explore various methods. Rodrigues et al. (2013) highlighted the N4SID (Numerical Algorithm for Subspace State Space System Identification) method, though they noted its limitations in capturing the aging process of mills due to insufficient data sampling. To overcome these limitations, the application of Artificial Intelligence (AI), particularly Artificial Neural Networks (ANNs), has been advanced for their capability to model complex input-output relationships, as emphasized by Colla (2022) and Shi et al. (2022). These methods are instrumental in reducing process losses.

Further investigations into AI's potential were conducted by Escribano et al. (2012), who employed ANNs and regression trees to model the behavior of cold rolling mills, achieving promising results. Remarkably, Shen et al. (2022) showed that ANN-based models could improve the prediction accuracy of the rolling force by up to 50% compared to traditional models, integrating variables such as steel chemistry and rolling parameters.

However, empirical tuning of ANN hyperparameters remains a pervasive challenge, as evidenced in the works of Pican et al. (1996), Wiklund and Sandberg (2002), He and Liu (2005), and Ren et al. (2021). These studies often lack a comprehensive comparison of different network architectures, which would elucidate the effects of various topologies on performance. Addressing these gaps, this project employs GEO algorithm, to optimize ANN architectures, facilitating a structured evaluation against previously used empirical methods.

Additionally, Santos and Barcelos (2020) explored the use of neural networks within the control systems of a cold strip rolling mill, focusing on architecture that minimizes output discrepancies. However, this approach did not consider whether the architecture achieved an optimal balance between error reduction and computational efficiency—a critical factor when implemented on a Programmable Logic Controller (PLC), where resources are limited.

Rodrigues et al. (2023) successfully demonstrated the feasibility of using ANNs in skin-pass mills with a limited dataset. They identified an optimal structure with approximately 30 neurons across three layers, achieving an error rate of 0.7288%. While this error rate is impressive, the study did not delve into optimizing the architecture for computational efficiency, crucial for PLC environments.

This article proposes the use of Generalized Extremal Optimization (GEO) to optimize ANNs, aiming to balance error minimization with computational cost, addressing the shortcomings identified in prior studies like that of Rodrigues et al. (2023). Employing GEO can refine ANN architectures to ensure they are both precise and computationally efficient, an essential balance for applications in PLCs and similar environments. One significant advantage of using GEO is its simplicity, due to having only one free parameter, which facilitates easier configuration compared to other optimization methods. This streamlined approach, supported by a larger dataset, is expected to enhance model training, leading to more accurate and

robust systems—key for real-world industrial applications where accuracy and efficiency are paramount.

3. BACKGROUND

3.1 GEO Algorithm

The GEO algorithm, proposed by Sousa et al. (2003), has been increasingly applied to complex optimization problems within the aerospace sector (Sousa et al., 2007; Mainenti-Lopes et al., 2012, 2016; Freitas et al., 2018).

The GEO algorithm is a global search meta-heuristic developed by Sousa. It is inspired by natural evolutionary processes and based on the theory of *Self-Organized Criticality* (SOC), used to explain power-law signatures in complex systems. Unlike methods such as Simulated Annealing or Genetic Algorithms, GEO does not rely on derivatives and can handle non-convex or disjoint problems with variables that are continuous, discrete, or mixed. A unique feature of GEO is its single free parameter, τ, which controls the search determinism from a random walk ($\tau = 0$) to a deterministic local search ($\tau \to \infty$). The typical range for τ in applications is between 1 and 5. (Sousa et al., 2003; Galski, 2006)

In the GEO framework, a set of L solution components, represented as bits in a binary sequence (similar to a chromosome in genetic algorithms), encode the design variables. Each bit's fitness is determined by the impact of its toggling on the objective function value. The algorithm ranks bits by their fitness and applies mutations probabilistically, governed by τ. (Sousa et al., 2003; Galski, 2006)

The operational steps of the GEO algorithm are as follows (Sousa et al., 2003; Galski, 2006):

(1) Initialize a binary string C of length L encoding N design variables. Convert C to X, calculate the objective function $F(X)$, and set $X_{\text{best}} = X$ and $F(X_{\text{best}}) = F(X)$.
(2) For each bit i in C:
   - Toggle i to get a new configuration C_i, convert C_i to X_i, and compute $F(X_i)$.
   - Assign an adaptation index $\Delta F(X_i) = F(X_i) - F(X_{\text{best}})$ and revert i to its original value.
(3) Sort all bits by $\Delta F(X_i)$ from the least to the most adapted and pair each bit's position with its rank.
(4) Select a bit to mutate based on a probability $P(i) = k^{-\tau}$ where k is the rank, and mutate if a generated random number is less or equal to $P(i)$.
(5) Update C to C_{esc}, convert to X, and compute $F(X)$.
(6) If $F(X) < F(X_{\text{best}})$, update $X_{\text{best}} = X$ and $F(X_{\text{best}}) = F(X)$.
(7) Repeat the above steps until a stop criterion is met and return X_{best} and $F(X_{\text{best}})$.

3.2 Artificial Neural Network (ANN)

The artificial neural network used in this study is a multilayer perceptron (MLP) due to its simplicity and ability to map inputs to outputs, as described Haykin (2009).

The GEO algorithm initiates its process by randomly generating a population. Each individual in this population encapsulates a unique set of architectural parameters for an Artificial Neural Network (ANN), which includes the number of hidden layers, the neurons in each layer, and the chosen activation function. These parameters define each ANN's structure, triggering a separate training cycle for every configuration in the population. Throughout these cycles, the neural network receives inputs and computes outputs, represented as $\hat{y}$, according to the following equations (Haykin, 2009):

$$\hat{y}_{(j)} = a_j^{[l+1]} = z_j^{[l+1]}, \tag{1}$$

$$a_j^{[l]} = \sigma\left(z_j^{[l]}\right), \tag{2}$$

$$z_j^{[l]} = \left(\sum_{k=1}^{m} w_{jk}^{[l]} x_j^{[i]} + b_j^{[l]}\right). \tag{3}$$

The output of a neuron, denoted by a, is calculated using the activation function σ. The weights w and inputs x from the previous layer contribute to this output, alongside the bias b. In this formula, m represents the total number of input connections to the neuron. Layer and neuron specificity are indicated by superscripts and subscripts: $[l]$ marks the l^{th} layer, (i) identifies the i^{th} training example, j refers to each input unit, and k specifies the particular neuron within the layer.

The activation function establishes a non-linear relationship between the input and output of a neuron, critical for enabling complex pattern recognition in neural networks. Prominent among these transfer functions are the hyperbolic tangent and the sigmoid function (Haykin, 2009):

$$\sigma_{tanh}(z) = \frac{2}{e^{-2z}+1} - 1; \tag{4}$$

log-sigmoid

$$\sigma_{sig}(z) = \frac{1}{1+e^{-z}}; \tag{5}$$

and gaussian or radial basis

$$\sigma_{gauss}(z) = e^{-z^2}. \tag{6}$$

In this research, the Levenberg-Marquardt (LM) method was selected for training the network. This method combines strategies initially developed by Levenberg and Marquardt into a unified algorithm, as detailed in the cited literature (Haykin, 2009). The performance of the Artificial Neural Network (ANN) is evaluated using the mean-squared error (MSE) metric

$$E_{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left[(y_{(i)} - \hat{y}_{(i)})^2\right]. \tag{7}$$

4. METHOD

This research enhances the framework originally developed by Rodrigues et al. (2021), which was focused on simulator applications where timing is crucial. Here, we adapt the approach for use in environments with significant computational limitations, such as Programmable Logic Controllers (PLCs). This adaptation is vital for applying the methodology in settings with stringent hardware constraints, thereby improving the utility of Artificial Neural Networks (ANNs) not just in simulated environments but also in real-world industrial applications. By extending the original methodology to PLCs, this work addresses the need for systems that operate effectively within the limited computational capacities often present in these environments, increasing the practicality and range of neural network applications.

The customized approach for data from the skin-pass mill is demonstrated in Figure 1. This figure visually outlines the adapted process flow, elucidating how the methodology optimizes ANNs within the specific constraints of a skin-pass mill.

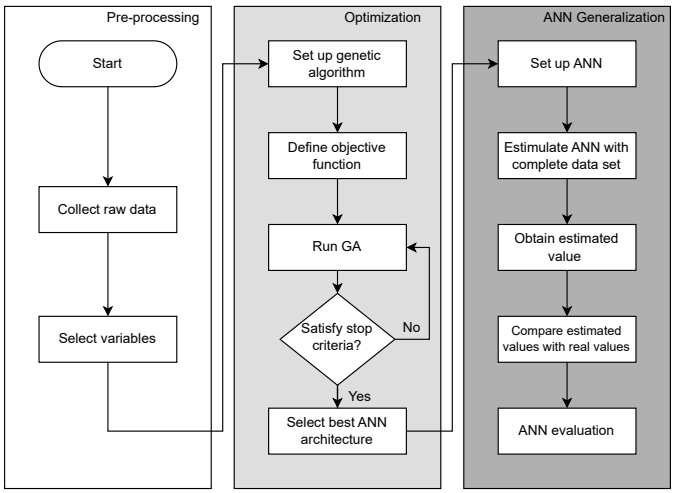

Figure 1. Steps of the methodology. Source: Adapted from Rodrigues et al. (2021).

During the pre-processing phase, data are gathered and readied for subsequent use. Specifically, the data are extracted from the database of the skin-pass mill and subsequently stored in spreadsheet format. Following this, the dataset intended for training the ANN is curated and selected.

4.1 Optimization

In the optimization phase, multiple configurations of the neural network are evaluated. The hyperparameters that are modified encompass: i) the number of neurons within each hidden layer, ii) the number of hidden layers , and iii) the type of activation function employed. The hidden layers and the neurons (depicted as circles) are illustrated in Figure 2.

Objective Function The objective of this phase is to identify an architecture that optimizes performance through

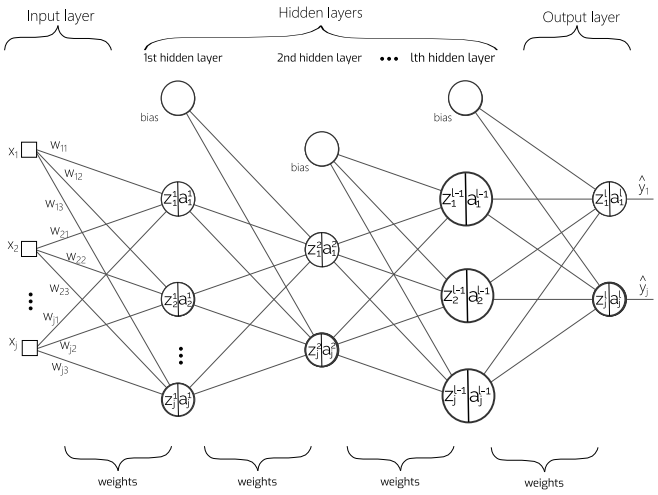

Figure 2. Hidden layers of neural networks. Source: Adapted from Haykin (2009); Rodrigues et al. (2021).

the minimization of the objective function, while concurrently diminishing the computational demands related to the number of neurons and the length of the learning process. The objective function employed for this evaluation is presented in (8) (Carvalho et al., 2011)

$$f_{obj} = penalty \times \left(\frac{\rho_1 \times E_{train} + \rho_2 \times E_{gen}}{\rho_1 + \rho_2} \right), \quad (8)$$

The parameters ρ_1 and ρ_2 ($\rho_1,\ \rho_2 > 0;\ \rho_1,\ \rho_2 \in \Re$) are used to weigh the importance of errors, with E_{train} representing the training error and E_{gen} indicating the generalization error. Both errors are quantified according to (7). The training error originates from the learning process of the ANN with the training data set, while the generalization error is evaluated using the test data set. It is essential to recognize that the data set employed during the optimization stage is a condensed version of the full data set, making it imperative for the ANN to be capable of generalizing to previously unseen data points. When $\rho_2 > \rho_1$, the focus shifts towards enhancing the network's performance on generalized data (Carvalho et al., 2011). The penalty term

$$penalty = \left(e^{5\times10^{-8}\times(nw)^2} \right) + \left(5 \times 10^{-5} \times epochs + 1 \right), \quad (9)$$

where nw signifies the total number of weights in the ANN and $epochs$ indicates the number of iterations necessary to train the ANN. This metric evaluates the complexity of the ANN architectures and applies penalties to configurations that involve a large number of neurons or prolonged training periods. Architectures with such complexity are generally ill-suited for real-time applications, like those required in an operational satellite simulator (Carvalho et al., 2011).

Running the GEO Table 1 specifies the values for each design variable used. During the execution of the GEO Algorithm, the configuration of neurons in the first hidden layer and the output layer remains unchanged. The number of neurons in the first hidden layer mirrors the count

of input nodes, while the number of neurons in the output layer matches the quantity of target variables.

In the implementation of the GEO algorithm, the parameter τ was set to 3. Two termination conditions were defined: i) the number of generations, indicating the frequency at which the population is evaluated by the objective function; ii) the average relative change in the value of the objective function over 50 assessments, which must be 1×10^{-6}. Furthermore, the parameters ρ_2 and ρ_1, as mentioned in (8), were established at 1.0 and 0.1, respectively.

Table 1. Parameters used to define ANN architecture.

Design Variable	Values
Neurons in each hidden layer	1...10
Hidden layers	$1, 2, 3$
Activation function	hyperbolic tangent, log-sigmoid, gaussian

During the search for the optimal ANN architecture, the GEO algorithm strives to minimize the objective function value (8); however, each iteration requires training an ANN. Throughout this training process, the dataset is partitioned into distinct subsets as follows (Haykin, 2009): i) 70% is allocated to the training set, utilized to calculate errors and adjust weights and biases across the training iterations; ii) 15% is designated for the validation set, used to prevent overfitting; iii) 15% forms the testing set, employed to assess the model's ability to generalize errors. It is important to note that the complete dataset consists of $267,870$ samples per variable, with $53,574$ samples being employed in the phase dedicated to identifying the most effective architecture.

Upon completion of the GEO algorithm, the optimal network is identified as the one that achieves the lowest value of the objective function while also minimizing computational demand, evidenced by having the smallest number of neurons.

4.2 ANN Generalization

Once optimization is concluded, the determined parameters, such as the number of neurons, the configuration of hidden layers, and the choice of activation function, are utilized to configure the ANN for subsequent performance evaluations.

After setting up the ANN, it undergoes testing with the entire unmodified dataset. This evaluation is critical to assess the ANN's capacity to generalize beyond the data used in its training phase. The responses of the ANN to these inputs are then examined to evaluate its effectiveness.

The evaluation of the ANN's performance primarily utilizes the mean squared error (MSE), detailed in 7, and the mean relative squared error (MRSE),

$$MRSE(\%) = \frac{1}{n}\sum_{t=1}^{no}\left[\sqrt{\frac{\sum_{i=1}^{N}(y_{(i)} - \hat{y}_{(i)})^2}{\sum_{i=1}^{N}(\hat{y}_{(i)})^2}}\right] \times 100, \quad (10)$$

as key metrics.

where n indicates the aggregate number of output variables, and N corresponds to the total data samples used for evaluating the performance of the ANN. Given that there are two outputs, the overall Mean Relative Squared Error (MRSE) is calculated as the average of the MRSE for each individual output.

The Mean Relative Squared Error (MRSE) and Mean Squared Error (MSE) are statistical measures that quantify the variations between the actual outputs of the system and the outputs predicted by the ANN. Metrics nearing zero reflect that the ANN models the system's dynamics with substantial precision (Borjas and Garcia, 2011).

5. CASE STUDY: SKIN-PASS MILL

The dataset employed in this study was obtained from a skin-pass mill. The particular section under investigation is referred to as the "mill region," which is illustrated in Figure 3. This area contains the equipment used for data collection as noted by Rodrigues et al. (2013):

- Tensiometers: Comprising three interconnected cylinders attached to a load cell. These instruments are utilized to measure the force necessary for assessing mechanical stress.
- Backup roll: These rollers make direct contact with the material undergoing processing.
- Central tensioner: Featuring three tensioning rollers driven by a unified motor, this apparatus is crucial for sustaining the mechanical stress at the output of the mill.

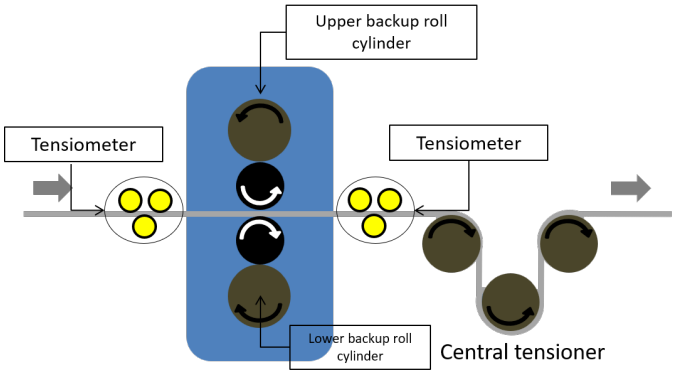

Figure 3. Skin-Pass mill diagram. Source: Rodrigues et al. (2013)

The hydraulic actuation facilitates the operation of the backup roll, with the lower backup roll cylinder approaching the upper cylinder. This action applies force ($\boldsymbol{F}$), resulting in the elongation of the steel sheet. To ensure proper elongation, it is essential to tension the steel strip at both the entry and exit of the mill to maintain full extension of the sheet (Rodrigues et al., 2013).

The variables for identification are outlined in Table 2, indicating their association with particular equipment, measurement techniques, units of measurement, and their classification within the system as either input or output variables.

The mechanical stress level necessary for processing the steel sheet effectively, without defects, loosening, or sheet

breakage, is contingent upon its dimensions, including thickness and width. Input mechanical stress data informs the motor control of the lower backup roll cylinder, adjusting its speed accordingly. Similarly, output mechanical stress data are conveyed to the motor control of the central tensioner, regulating the exit stress of the rolling mill (Rodrigues et al., 2013).

6. RESULTS

Over a period of about 61 hours and 27 minutes [1], a total of 3520 neural network architectures were evaluated. Towards the end of the GEO's execution, no further architectural tests were conducted as the objective function's cost had not been minimized for over 50 iterations. The cost distribution of the objective function is illustrated in Figure 4.

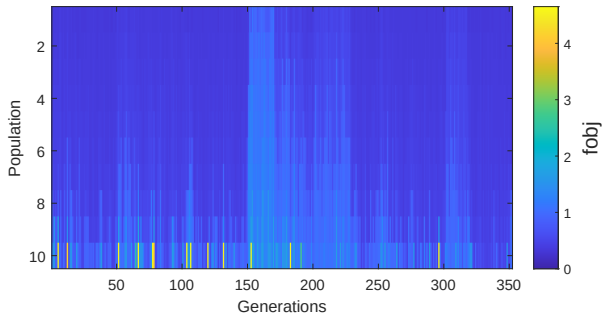

Figure 4. The objective function map.

Out of the architectures assessed, the one demonstrating superior performance consisted of three hidden layers, comprising 9, 9, and 5 neurons respectively, and employed a log-sigmoid activation function. This architecture yielded an objective function value of $fobj = 0.3167$.

Figures 5 and 6 present the ANN's performance in mapping the input and output mechanical tension of the skin-pass mill. The system's response, shown in blue, corresponds to the different setpoints applied according to the material properties. Although the ANN achieved an average MRSE of 6.82%, it is notable that the samples between $200,000$ and $250,000$ did not produce satisfactory results, revealing a significant disparity between the expected and ANN-estimated values.

To assess the performance of the architecture developed in this study, the dataset used by Rodrigues et al. (2023) was employed for comparative purposes. The obtained average MRSE of 0.8418% was slightly higher than the 0.7288% MRSE achieved by Rodrigues et al. (2023). However, it is noteworthy that the architecture proposed in this research requires fewer computational calculations, with approximately 185 calculations needed compared to 278 in the previous study. This indicates a subtle improvement in processing efficiency, underscoring the potential for a more resource-efficient solution tailored to the application's requirements.

[1] Hardware: Dell G7 with 32 GB of RAM, an i7 8750H processor, and 1 TB of SSD storage.

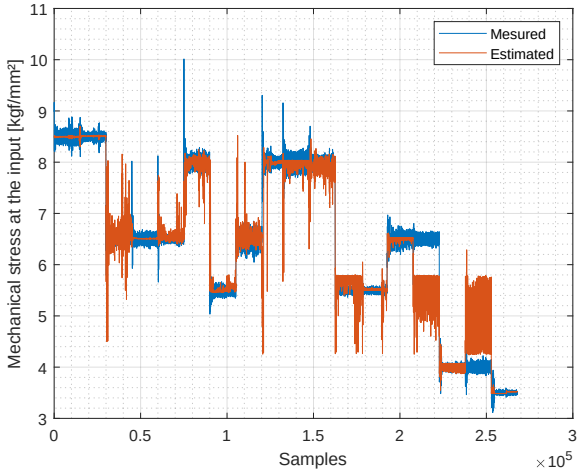

Figure 5. Mechanical stress at the input.

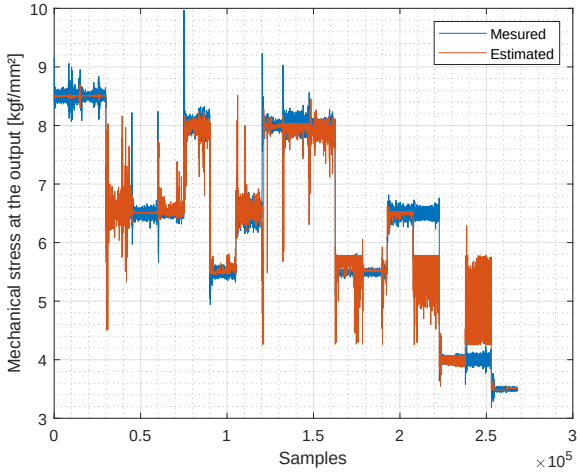

Figure 6. Mechanical stress at the output.

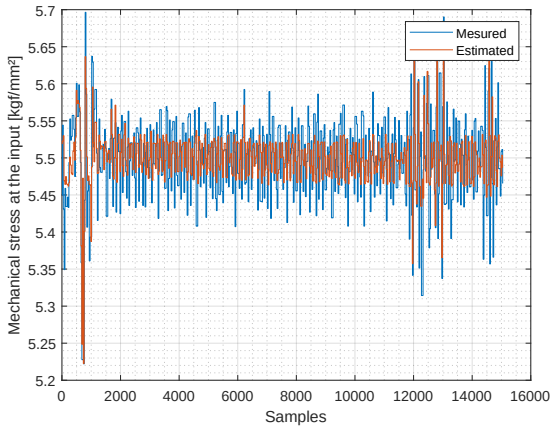

Figure 7. Mechanical stress at the input with dataset from Rodrigues et al. (2023).

7. CONCLUSIONS

The application of the Generalized Extremal Optimization (GEO) algorithm in optimizing Artificial Neural Network (ANN) architectures for skin-pass mills has yielded promis-

Table 2. Identification variables and their specifications. Source: Rodrigues et al. (2013)

Variable	Equipment	Sensor	Unit	Classification
Mechanical stress at the input	Tensiometer	Load cell	kgf/mm^2	Output
Backup roll cylinder speed	Lower backup roll cylinder	Encoder	mpm	Input
Mechanical stress at the output	Tensiometer	Load cell	kgf/mm^2	Output
Tensioner speed	Central tensioner	Encoder	mpm	Input

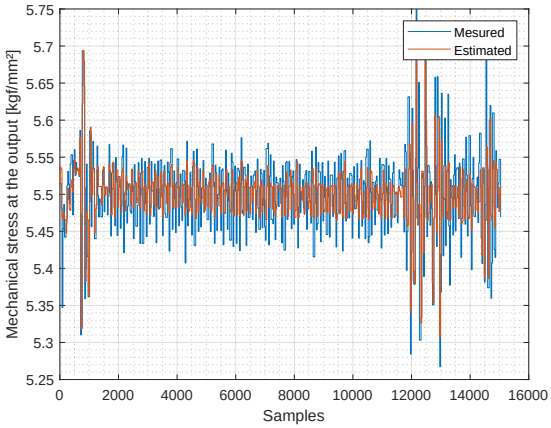

Figure 8. Mechanical stress at the output with dataset from Rodrigues et al. (2023).

ing results in this study. By exploring over 3500 neural network configurations, we identified an architecture with three hidden layers and a total of 23 neurons, utilizing a log-sigmoid activation function, which achieved an objective function value of 0.3167. Despite a slight increase in the mean relative squared error (MRSE) compared to previous work, our architecture demonstrated improved computational efficiency, requiring fewer calculations while maintaining a low error rate.

The success of this study underscores the effectiveness of the GEO algorithm in streamlining the process of ANN architecture optimization, providing a simpler yet accurate model for skin-pass mills. Future research avenues could explore the adaptation of the GEO algorithm with an adaptive approach to its single free parameter, potentially enhancing its performance further. Additionally, increasing the diversity of the dataset and addressing outliers could improve the generalization capabilities of the ANN architecture, paving the way for more robust industrial applications in the future.

In conclusion, while this study has demonstrated the efficacy of the Generalized Extremal Optimization (GEO) algorithm in optimizing Artificial Neural Network (ANN) architectures for skin-pass mills, there are opportunities for further refinement and improvement. One notable aspect to consider is the limited portion of the dataset utilized for the optimization process, representing approximately 20% of the available samples. Expanding the dataset could enhance data diversity, potentially leading to the discovery of an ANN architecture that more effectively generalizes the data, especially considering the variation in samples attributable to different materials processed in the skin-pass mill.

ACKNOWLEDGEMENTS

The authors would like to express heartfelt appreciation to the Centro Universitário de Volta Redonda for their invaluable institutional support during the course of this research. Additionally, sincere thanks are extended to CAPES for the financial assistance provided to I.P. Rodrigues (Process No. 88882.444522/2019-01) during their doctoral studies, which greatly contributed to the successful completion of this work.

REFERENCES

Alves, P.G., Castro, J.A.d., Moreira, L.P., and Hemerly, E.M. (2012). Modeling, simulation and identification for control of tandem cold metal rolling. *Materials Research*, 15(6), 928–936. doi:10.1590/s1516-14392012005000137. URL `http://dx.doi.org/10.1590/S1516-14392012005000137`.

Borjas, S.D.M. and Garcia, C. (2011). Subspace identification for industrial processes. *Trends in Applied and Computational Mathematics*, 12(3), 183–194. doi:10.5540/tema.2011.012.03.0183. URL `https://tema.sbmac.org.br/tema/article/view/547`.

Carvalho, A.R., Ramos, F.M., and Chaves, A.A. (2011). Metaheuristics for the feedforward artificial neural network (ann) architecture optimization problem. *Neural Computing and Applications*, 20(8), 1273 – 1284. URL `http://dx.doi.org/10.1007/s00521-010-0504-3`.

Chagas, R.A.J., Galski, R.L., and Sousa, F.L.d. (2014). Orbgeo - an orbit selection tool for satellite constellations using the multiobjective generalized extremal optimization (mgeo) algorithm. In *Proceedings...* International Systems and Concurrent Engineering for Space Applications Conference, 6. URL `http://congrexprojects.com/2014-events/14c08/`.

Colla, V. (2022). A big step ahead in metal science and technology through the application of artificial intelligence. *IFAC-PapersOnLine*, 55(21), 1–6. doi:https://doi.org/10.1016/j.ifacol.2022.09.234. URL `https://www.sciencedirect.com/science/article/pii/S2405896322014641`. 19th IFAC Symposium on Control, Optimization and Automation in Mining, Mineral and Metal Processing MMM 2022.

Escribano, R., Lostado, R., de Pisón, F.M., Pernía, A., and Vergara, E. (2012). Modelling a skin-pass rolling process by means of data mining techniques and finite element method. *Journal of Iron and Steel Research, International*, 19(5), 43–49. doi:https://doi.org/10.1016/S1006-706X(12)60098-3. URL `https://www.sciencedirect.com/science/article/pii/S1006706X12600983`.

Freitas, V.L.d.S., Sousa, F.L.d., and Macau, E.E.N. (2018). Reactive model for autonomous vehicles formation following a mobile reference. *Applied Mathematical Modelling*, 61, 167–180. doi:10.1016/j.apm.2018.04.

011. URL http://dx.doi.org/10.1016/j.apm.2018.04.011.

Galski, R.L. (2006). *Desenvolvimento de versões aprimoradas, híbridas, paralela e multiobjetivo do método da otimização extrema generalizada e sua aplicação no projeto de sistemas espaciais*. Ph.D. thesis, Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos. URL http://urlib.net/ibi/6qtX3pFwXQZGivnK2Y/NjtT4.

Haykin, S. (2009). *Neural Networks and Learning Machines*. Bookman, 3nd edition.

He, H.T. and Liu, H.M. (2005). The research on integrated neural networks in rolling load prediction system for temper mill. In *2005 International Conference on Machine Learning and Cybernetics*, volume 7, 4089–4093 Vol. 7. doi:10.1109/ICMLC.2005.1527653.

Mainenti-Lopes, I., Souza, L.C.G.d., and Sousa, F.L.d. (2016). A comparative study between control strategies for a solar sailcraft in an earth–mars transfer. *Mechanical Systems and Signal Processing*, 79, 289–296. doi:10.1016/j.ymssp.2016.02.026. URL http://dx.doi.org/10.1016/j.ymssp.2016.02.026.

Mainenti-Lopes, I., Souza, L.C.G.d., and Sousa, F.L.d. (2012). Design of a nonlinear controller for a rigid-flexible satellite using multi-objective generalized extremal optimization with real codification. *Shock and Vibration*, 19(5), 1–10. doi:10.3233/SAV-2012-0702. URL http://dx.doi.org/10.3233/SAV-2012-0702. Setores de Atividade: Atividades profissionais, científicas e técnicas.

Pican, N., Alexandre, F., and Bresson, P. (1996). Artificial neural networks for the presetting of a steel temper mill. *IEEE Expert*, 11(1), 22–27. doi:10.1109/64.482953.

Ren, X.Y., Gao, H.M., Xu, H.W., Huang, H.G., and Sun, J.N. (2021). Identification and control of elongation system of skin passing mill based on intelligent algorithm. *Journal of Physics: Conference Series*, 1820(1), 012154. doi:10.1088/1742-6596/1820/1/012154. URL http://dx.doi.org/10.1088/1742-6596/1820/1/012154.

Rodrigues, I.P., Jorge, J.M., and Oliveira, K.F.d. (2013). Identificação e controle de um laminador de encruamento em malha fechada através de métodos de subespaços. Trabalho de Conclusão de Curso.

Rodrigues, I.P., Oliveira, P.A., Ambrosio, A.M., and Chagas, R.A. (2021). Modeling satellite battery aging for an operational satellite simulator. *Advances in Space Research*, 67(6), 1981–1999. doi:https://doi.org/10.1016/j.asr.2020.12.031. URL https://www.sciencedirect.com/science/article/pii/S0273117720309042.

Rodrigues, I.P., Rodrigues, G.A., and Santos, B.L.d. (2023). Artificial neural networks in skin-pass mill modeling: an empirical analysis of architecture and hyperparameters. *Cadernos UniFOA*, 18(53), 1–12. doi:10.47385/cadunifoa.v18.n53.4672. URL http://dx.doi.org/10.47385/cadunifoa.v18.n53.4672.

Santos, B.C.d. and Barcelos, A.F. (2020). Aplicação da rede neural artificial como ferramenta de diagnóstico e controle do sistema de tensão de um laminador de tiras a frio. In *XVII SEGeT - Simpósio de Excelência em Gestão e Tecnologia*. Accessed: 31/03/2024.

Seo, M., Ban, J., Koo, B.Y., and Kim, S.W. (2020). Static model identification for sendzimir rolling mill using noise corrupted operation data. *IEEE Access*, 8, 150685–150695. doi:10.1109/ACCESS.2020.3017025.

Shen, S., Guye, D., Ma, X., Yue, S., and Armanfard, N. (2022). Multistep networks for roll force prediction in hot strip rolling mill. *Machine Learning with Applications*, 7, 100245. doi:https://doi.org/10.1016/j.mlwa.2021.100245. URL https://www.sciencedirect.com/science/article/pii/S2666827021001237.

Shi, P., Gao, H., Yu, Y., Xu, X., and Han, D. (2022). Intelligent fault diagnosis of rolling mills based on dual attention- guided deep learning method under imbalanced data conditions. *Measurement*, 204, 111993. doi:https://doi.org/10.1016/j.measurement.2022.111993. URL https://www.sciencedirect.com/science/article/pii/S0263224122011897.

Sousa, F.L.d., Soeiro, F.J.C.P., Neto, A.J.S., and Ramos, F.M. (2007). Application of the generalized extremal optimization algorithm to an inverse radiative transfer problem. *Inverse Problems in Science and Engineering*, 15(7), 699–714. doi:10.1080/17415970701198415. URL https://doi.org/10.1080/17415970701198415.

Sousa, F.L.d., Ramos, F.M., Paglione, P., and Girardi, R.M. (2003). New stochastic algorithm for design optimization. *AIAA Journal*, 41(9), 1808–1818. doi:10.2514/2.7299. URL https://doi.org/10.2514/2.7299.

Wiklund, O. and Sandberg, F. (2002). Chapter 15 - modelling and control of temper rolling and skin pass rolling. In J. Lenard (ed.), *Metal Forming Science and Practice*, 313–343. Elsevier Science Ltd, Oxford. doi:https://doi.org/10.1016/B978-008044024-8/50015-1. URL https://www.sciencedirect.com/science/article/pii/B9780080440248500151.