MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ENGENHARIA DE ENERGIA

EFFICIENCY AND NOX EMISSION OPTIMIZATION BY GENETIC ALGORITHM OF A COAL-FIRED STEAM GENERATOR MODELED WITH ARTIFICIAL NEURAL NETWORKS

por

Bárbara Pacheco da Rocha

Monografia apresentada à Comissão de Graduação do Curso de Engenharia de Energia da Escola de Engenharia da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do diploma de Bacharel em Engenharia de Energia.

Porto Alegre, Dezembro de 2019

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

ESCOLA DE ENGENHARIA

ENGENHARIA DE ENERGIA

EFFICIENCY AND NOX EMISSION OPTIMIZATION BY GENETIC ALGORITHM OF A COAL-FIRED STEAM GENERATOR MODELED WITH ARTIFICIAL NEURAL NETWORKS

por

Bárbara Pacheco da Rocha

ESTA MONOGRAFIA FOI JULGADA ADEQUADA COMO PARTE DOS REQUISITOS PARA A OBTENÇÃO DO TÍTULO DE
BACHAREL EM ENGENHARIA DE ENERGIA.
APROVADA EM SUA FORMA FINAL PELA BANCA EXAMINADORA

Prof. Letícia Jenisch Rodrigues
Coordenador do Curso de Engenharia de Energia

Orientador: Prof. Dr. Paulo Smith Schneider

Coorientadora: Natália de Assis Brasil Weber

Banca examinadora:

Eng. Bruno Venanzio Trasatti – SENAI

Prof. Dr. Sergio Luis Haffner – DELAE/UFRGS

Prof. Dr. Paulo Smith Schneider – DEMEC / UFRGS

Porto Alegre, 11 de dezembro 2019.

ACKNOWLEDGMENTS

To my family and, particularly, to my parents, for their unconditional love and support throughout all stages of my life.

To my thesis advisor, Paulo Smith Schneider, for the years of guidance throughout my entire academic career. Thank you for having believed in my potential. I am also thankful to my co-advisor, Natália de Assis Brasil Weber, for her contributions.

To my friends and colleagues that somehow shared with me these intense university years. You made this experience better.

I'd like to thank the university and the Conselho Nacional de Desenvolvimento Científico e Tecnológio (CNPq) for the financial support during my first year of scientific research.

Finally, I would like to acknowledge EDP - Energia de Portugal for the financial support that enabled the development of the SMART-PECÉM R&D project. On behalf of Guilherme de Oliveira, I would like to thank the entire EDP technical team, who contributed greatly to the development of this work.

ROCHA, B. P. **Efficiency and NOx Emissions Optimization by Genetic Algorithm of a Coal-Fired Steam Generator Modeled with Artificial Neural Networks.** 2019. 37 páginas. Monografia (Trabalho de Conclusão do Curso em Engenharia de Energia) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

RESUMO

Este trabalho faz parte do desenvolvimento de um modelo de apoio à decisão para a operação de um gerador de vapor real. O estudo propõe uma otimização combinada que visa encontrar pontos de operação que atinjam a maior eficiência do gerador de vapor associada à menor emissão de NO_x, aplicando algoritmo genético na saída de modelos de redes neurais artificiais (RNA). A base de dados é formada por 10 parâmetros de operação coletados durante um ano e meio com passo de meia hora e tratados estatisticamente. O comportamento do gerador de vapor é modelado por redes neurais artificiais Perceptron de várias camadas, com saídas separadas para eficiência e emissão de NOx. As métricas de avaliação empregadas nas RNAs foram o erro médio absoluto (MAE), erro quadrático médio (MSE), erro médio percentual (MAPE) e coeficiente de determinação (R^2). A RNA para predizer o comportamento da eficiência apresenta MSE e MAE do seu teste de 0,7572 e 0,6206, respectivamente e a RNA para NO_x apresenta MSE e MAE do seu teste de 312,43 e 12,36. A otimização tem como alvo atingir 98% de eficiência do gerador de vapor e 220,00 mg/m_N^3 de emissões de NOx, e se aproxima dessas metas com 97,95% de eficiência e 222,28 mg/m_N^3 de emissões de NOx.

PALAVRAS-CHAVE: Projeto de Experimentos, Gerador de Vapor a Carvão Pulverizado, Metamodelo, Otimização Combinada, Termelétrica a Carvão.

ROCHA, B. P. **Efficiency and NOx Emissions Optimization by Genetic Algorithm of a Coal-Fired Steam Generator Modeled with Artificial Neural Networks**. 2019. 37 pages. Monografia (Trabalho de Conclusão do Curso em Engenharia de Energia) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

ABSTRACT

This work is part of the development of a decision support model for the operation of a real steam generator. The study proposes a combined optimization that aims to find operating points that achieve the highest efficiency of the steam generator associated with lower NO_x emissions, applying genetic algorithm to the output of artificial neural network (ANN) models. The database consists of 10 operating parameters collected over a year and a half with a half-hour step and treated statistically. The behavior of the steam generator is modeled by multilayer Perceptron artificial neural networks with separate outputs for efficiency and NOx emission. The evaluation metrics applied to the ANNs were mean absolute error (MAE), mean square error (MSE), mean percentage error (MAPE) and coefficient of determination (R^2). The ANN for predicting efficiency behavior presents test MSE and MAE of 0.7572 and 0.6206, respectively, and RNA for NOx has test MSE and MAE of 312.43 and 12.36. The optimization targets 98% efficiency of the steam generator and 220.00 mg/m_N^3 of NOx emissions, and approaches these goals with 97.95% efficiency and 222.28 mg/m_N^3 of NOx emissions.

KEYWORDS: Coal Power Plant, Combined Optimization, Design of Experiments, Metamodel, Pulverized Coal Steam Generator.

INDEX

LIST OF INITIALS AND ABBREVIATIONS

ANOVA	Analysis of Variance
ANN	Artificial Neural Network
CO_2	Carbon Dioxide
DEAP	Distributed Evolutionary Algorithms in Python
DoE	Design of Experiments
GA	Genetic Algorithm
HHV	Higher Heating Value
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
MLP	Multi-Layer Perceptron
NO_x	Nitric Oxide
OFA	Over Fired Air
OFaT	One-Factor-at-a-Time
ONS	National System Operator
ReLU	Rectifier Linear Unit
RSM	Response Surface Methodology
SCADA	Supervisory Control and Data Acquisition
SSE	Sum of Squares Errors
SHSG	Superheated Steam Generator
Tanh	Hyperbolic Tangent
TSS	Total Sum of Squares
UG2	Generating Unit 2

LIST OF SYMBOLS

η	Efficiency of the superheated steam generator
$\dot{m}_{MS}$	Main steam flow rate (ton/h)
$\dot{m}_{RS}$	Reheat steam flow rate (ton/h)
$\dot{m}_{coal}$	Fuel mass flow (ton/h)
h_{MS}	Main steam enthalpy (kJ/kg)
h_{RS}	Reheated steam enthalpy (kJ/kg)
h_{SR}	Enthalpy of the steam to be reheated (kJ/kg)
h_f	Feed water enthalpy (kJ/kg)
R^2	Coefficient of determination
f	Fitness function
a	Ponderation to weight efficiency
b	Ponderation to weight NO_x emissions
η_{opt}	Targeted efficiency of the SHSG
η_{pred}	SHSG efficiency predicted by the ANN
$[NO_x]_{opt}$	Targeted NO_x emissions
$[NO_x]_{pred}$	NO_x emissions predicted by the ANN

1 INTRODUCTION

Annual world energy consumption grew by 2.3% in 2018 driven by a strong global economy paired with rising demand for heating and cooling The increase in electricity demand accounted for about half of this growth (IEA, 2018a), which represents almost twice the average growth rate seen since 2010.

Global coal demand grew for the second year in a row in 2018, but its share on the energy mix continued to fall. While coal's share of primary energy demand and electricity generation continues to slowly decline, it remains the world's largest source of electricity and the second largest source of primary energy (IEA, 2018b). As a result of higher energy consumption, 2018 CO_2 emissions increased by 1.7% over the previous year and set a new record. Coal-fired power generation remains the largest single emitter, accounting for 30% of all energy-related carbon dioxide emissions (IEA, 2018a).

Although renewable energy is constantly growing in the global energy matrix, fossil fuel still remains predominant at the base. The Brazilian electric system can be considered mostly hydrothermal, where thermoelectric generation represents 24.52% of the installed power (ANEEL, 2019). Coal has a 12.9% stake in thermoelectric generation (EMPRESA DE PESQUISA ENERGÉTICA - EPE, 2018), concentrated mainly in southern Brazil, where the largest deposits are located.

Coal offers the benefit of the lower fuel cost among fossil fuels, but in addition to having higher initial construction costs, it is more difficult to operate compared to oil or gas plants (GP STRATEGIES, 2013). In China, coal-fired power plants are the main suppliers of electricity, as well as the largest consumer of coal and water resources and the largest emitter of SO_x, NO_x and greenhouse gases (GHGs) (XU et al., 2011). Therefore, it is important to establish a comprehensive, scientific, reasonable and feasible evaluation system for coal-fired thermal power plants to guide them in the multiple optimization of their thermal, environmental and economic performance.

Fossil fuel boilers have been in their present form since the early 1900s. While designs have evolved into larger sizes, better materials and better efficiency, the basic concept of heat transfer generated from the combustion reaction to water cooled pipes remains the same. The boiler's main objectives are to mix combustion air and fuel, burn the air-fuel mixture, transfer the maximum amount of heat from the combustion process to the working fluid and exhaust combustion through the products. Conventional boilers typically operate at about 85 to 90% efficiency (GP STRATEGIES, 2013).

Steam generators are complex and highly relevant heat exchangers within the simulation of thermoelectric plants. Traditional mathematical methods that make use of mass and energy balances can become complicated due to the large number of parameters and the nonlinearity of the phenomena involved. This difficulty of implementation drove the use of artificial intelligence to do so.

Technological advances in data acquisition and computational power over the last decades have enabled the implementation of artificial intelligence algorithms to support the solution of real engineering problems. Machine learning models, such as artificial neural networks (ANN), have the ability to recognize patterns and infer relationships from a dataset. Artificial neural networks enable easy-to-implement modeling with quick and appropriate responses to problems in many areas, including complex physical problems such as steam generators.

ANNs have already been successfully applied to reproduce and simulate the behavior of heat transfer problems involving gas modeling, energy efficiency optimization and NOx emissions, energy resource prediction, among others (GHUGARE et al., 2014). ANN models can be developed and applied to existing systems using actual plant data stored, and this dataset can be further updated with new plant data. ANN modeling of real plant data has been previously investigated by DE et al., 2007; MESROGHLI; JORJANI; CHEHREH CHELGANI, 2009; SMREKAR et al., 2009; STRUŠNIK; GOLOB; AVSEC, 2015. It is worth noting that there is no generic model and it is necessary to develop a specific model to reproduce the actual equipment or system.

Hybrid models that combine experimental data with artificial neural network modeling and optimization algorithms, have already been implemented to assist obtaining adjustment points for variables of interest with opposite behaviors. Liu et al. (2016) and Chang (2014) have proposed the application of a genetic algorithm on an ANN output to improve coal-fired boiler efficiency while reducing pollutant emissions caused by NO_x.

The objective of this paper is to present a methodology for optimizing the combined effect of the efficiency and NOx emissions of a coal-fired steam generator modeled by artificial neural network. The

proposed work is part of a set of tools aiming to guide the steam generator operation and support the operator's decisions.

The specific objectives are as follows:

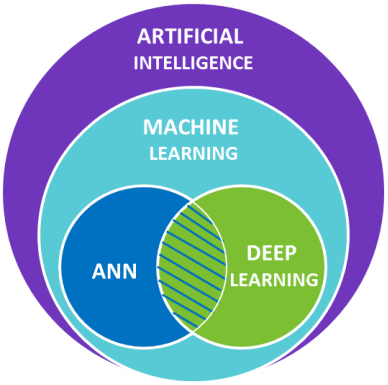
- Assemble a data set to build ANNs based on actual measured data from the Pecém power plant through statistical analysis.
- Analyze and validate different artificial neural network topologies to define the most suited to represent the problems at hands.
- Propose different importance weights of efficiency and NOx emissions on the genetic algorithm fitness function.

2 THEORETICAL BACKGROUND

2.1 History of Artificial Neural Networks and their relationship with Artificial Intelligence

Artificial Intelligence is the large area defined by John McCarthy as "The science and engineering of producing intelligent machines." Within, there are still other subareas such as Machine Learning and Deep Learning that can best be seen in Figure 2.1 along with the location of Artificial Neural Networks.

Figure 2.1 - Artificial intelligence and its subareas.



Source: The author.

While artificial intelligence can be defined as science capable of mimicking human skills, Machine Learning is a specific strand that trains machines to learn from data. ANNs are one of the existing methods within Machine Learning and the one chosen to be applied in this work.

The history of artificial intelligence begins in conjunction with Artificial Neural Networks, tracing back the work done by MCCULLOCH and PITTS (1943) whose proposed the first artificial neuron model. They presented an analogy between living cells and electronic processes, simulating the behavior of a biological neuron. The proposed neuron model presented binary activity and had no weighting. From this work, several studies on RNA's have been proposed over the decades.

HEBB (1949) proposed a learning law, demonstrating that the network's learning potential depends on the activation of pre- and postsynaptic cells, which once activated simultaneously lead to a change in synaptic weight. In 1958, ROSENBLATT suggested the neural model called the perceptron with the goal of training an RNA to achieve greater synaptic efficiency by inserting weights at each input. RUMELHART, G. E. HINTON and R.J.WILLIAMS (1986) developed the training method called backpropagation algorithm for the training of neural networks using multilayer perceptrons. This algorithm made it possible to solve more complex and nonlinearly separable problems.

2.2 Artificial Neural Networks

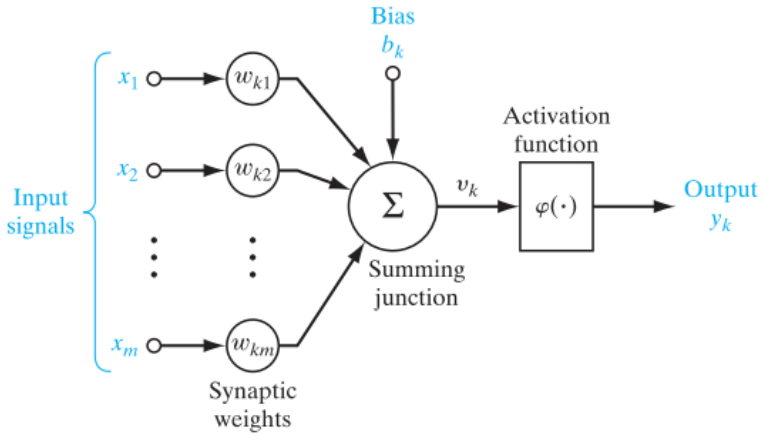
An artificial neural network consists of an information processing system created based on the functioning of biological neurons. Resembling the human brain, ANNs are composed of a large number of simple processing elements called neurons, which will gather information from the environment through a learning process, which will be further described in the following section. Each neuron is connected to other

neurons via targeted communication links reproducing a synapse, each with an associated weight (HAYKIN, 2014).

Haykin (2014) identifies three basic elements of the neural model that can be seen in Figure 2.2:

1. A set of synapses, or connection links characterized by a weight. Specifically, a signal x_j at the synapse input j connected to neuron k is multiplied by the synaptic weight w_{kj}, where k refers to the neuron in question, and j to the input parameter to which the weight refers. Weight may be in a range that includes both negative and positive values.
2. An adder to sum the input signals, weighted by the respective synaptic forces of the neuron.
3. An activation function φ to limit the output amplitude of a neuron.

Figure 2.2 - Model of a neuron



Source: HAYKIN, 2014

The neural model of also includes an externally applied bias denoted by b_k. The bias is the adjustable value whose effect is to increase or decrease the net input of the activation function in order to transfer it to the axis. In mathematical terms, we can describe the neuron k by writing the pair of equations:

$$v_k = \sum_{i=1}^{m} x_i w_{ki} + b_k \tag{2.1}$$

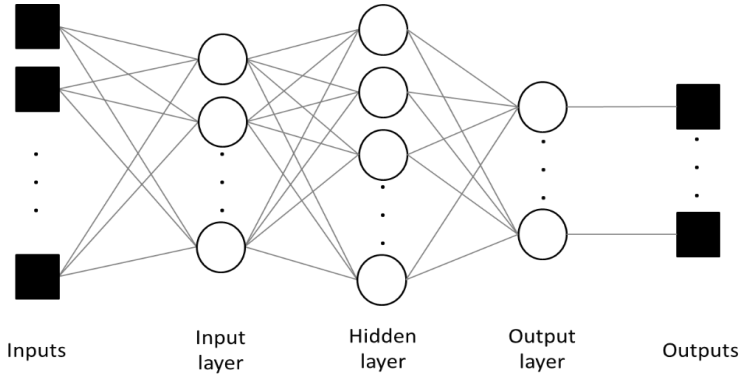
$$y_k = \varphi(v_k) \tag{2.2}$$

Eq. (2.1) describes the internal activity level v_k of the neuron, composed by the sum of the weighted inputs $x_i w_{ki}$ plus the bias b_k of the neuron k. It is observed in Eq. (2.2) that the output of neuron k will be v_k applied to the activation function $\varphi(.)$.

2.2.1 Multi-Layer Perceptron

The perceptron network is a model that is bounded by one layer of input neurons and another by output neurons. Whenever intermediate layers are added, the model is called Multi-Layer Perceptron (MLP), which is an extension of the perceptron model proposed by Rosenblatt. It is composed of several intermediate or hidden layers of artificial neurons. Figure 2.3 shows a schematic representation of the MLP architecture:

Figure 2.3 - Perceptron Multi-Layer Scheme.



Source: HAYKIN, 2014.

The MLP architecture houses an input layer, an output layer, and intermediate layers called hidden layers. The inputs are associated with neurons in the left layer of the input, where external information feeds the network. As a next step, the information passes to the hidden layer to be processed. The processed information is then transferred to the output layer.

The MLP model stands out for three characteristics: nonlinear activation function, hidden neurons and high degree of connectivity. The enable function should exhibit smooth nonlinearity for gradient variation and error to be reduced. Hidden neurons are responsible for the absorption of progressive knowledge, allowing the execution of more complex tasks. Finally, it is important to emphasize that the network has high connectivity of its synapses and that any modification to the network requires that it be restructured (HAYKIN, 2014).

It must be determined whether the expected output meets the stipulated precision requirements. If the expected output and actual output error do not meet the accuracy requirement and do not reach the maximum training time, it will enter the error propagation phase. This occurs when the error is transferred layer by layer from through the hidden layers to the input layer. The error signal of each neuron will then change the value of each neuron. This weighting process is the learning network training process, responsible for a continuous loop until the network output error is reduced to the required accuracy or to a predefined maximum number of times.

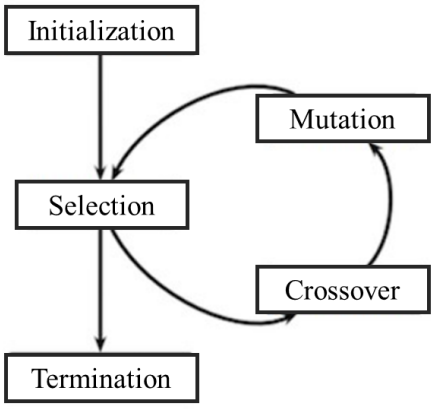
2.3 Genetic algorithm

Evolutionary optimization methods are a family of heuristic-based algorithms typically inspired by some phenomena from nature, wildly used to solve challenging optimization problems. As ROMANYCIA and PELLETIER (1985) defined, heuristic techniques are practical methods that cannot guarantee finding a global optimal, but are able to reach short-term, satisfactory solutions for impractical problems. Evolutionary algorithms are extensively used in the analysis, design, and operation of systems that are highly nonlinear, high dimensional and noisy or for solving problems that are not easily dealt by classical deterministic methods of optimization (VENKATESWARLU et al., 2020).

Genetic algorithms (GA) combine the concepts of genetics and evolution into an optimization algorithm that involves iterative search procedures inspired by the natural selection process (Darwinism) (MEYER-BAESE et al., 2014). GAs, as they are known today, were first introduced by John Holland in the 1960s. HOLLAND (1975) defined in his book “Adaptation in Natural and Artificial Systems” the method for moving from one population of “chromosomes” (binary strings representing candidate solutions for a problem) to a new population, using selection together with the genetic operators of crossover, mutation and inversion (MITCHELL; STEPHANIE FORREST, 1994).

Figure 2.4 describes in a scheme the general operations made in a genetic algorithm.

Figure 2.4 - Schematic to present genetic algorithm functioning.



Source: SONI, 2018.

The algorithm is initialized with a number of randomly selected individuals, which correspond to a set of potential solutions to the problem in question. A percentage of these individuals are subjected to reproduce and form offspring through crossover. Another percentage can suffer mutations in their genes and become new individuals. The resulting population of candidates composed of the individuals that were untouched and the new ones, is evaluated according to the fitness function. Only the winning candidates are allowed to pass on to a new generation and restart the process.

Liu, Li, Gao, 2016 and Chang, 2014 proposed a combined objective function f (Eq. (2.3)) based on the minimization of deviations to find operational inputs able to reach targeted values of efficiency η and NO_x emissions.

$$\min f = a\left(\eta_{opt} - \eta_{pred}\right) + b([NO_x]_{pred} - [NO_x]_{opt}) \tag{2.3}$$

with η_{pred} and $[NO_x]_{pred}$ as the predicted values and η_{opt} and $[NO_x]_{opt}$ the targeted ones; a and b are two weighting variables.

2.4 Design of Experiments

Design of Experiments (DoE) is a statistic approach employed to acquire and assess data by exploiting the coupled sensitivity of multi-input factors on the responses of a experiment (MONTGOMERY, 2013). The methodology is based on ANOVA (analysis of variance) and the parameters significance is determined through hypothesis testing. Hypothesis testing is a statistical assistance tool used to state some conjecture about the problem situation, reflecting whether a proposal hypothesis is true or false within a confidence interval. The p-value is used to report whether the null hypothesis was or was not rejected. For the case of parameters significance, the null hypothesis is that there is no significant correlation between the parameters and the alternative hypothesis is its opposite (MONTGOMERY, 2013).

Among many DoE techniques, the Response Surface Methodology (RSM) Box-Behnken with central composition has proven to be more efficient, generating a smaller number of combinations when compared to 3^k factorial designs and avoiding experiments performed under extreme conditions (FERREIRA et al., 2007).

2.5 Pearson correlation

Pearson Correlation ρ is used to verify the correlation between the parameters under analysis. It enables the measurement of the intensity and direction of the linear association between two variables. This correlation is given by Eq. (2.4).

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \tag{2.4}$$

with $\bar{x}$ the sample mean for the first variable, s_x the standard deviation for the first variable, $\bar{y}$ the sample mean for the second variable, s_y the standard deviation for the second variable and n the number of sample elements or events. Pearson correlation levels are presented in Table 2.1.

Table 2.1 - Pearson correlation levels

Correlation Size	*Coefficient Interpretation*
0.8 to 1.0	Very strong correlation
0.6 to 0.8	Strong correlation
0.4 to 0.6	Moderate correlation
0.2 to 0.4	Weak correlation
0.0 to 0.2	Very weak or nonexistent correlation

Source: SALKIND, 2013.

P-value is an important indicator to be looked along with the Pearson correlation. The p-value assess whether a correlation coefficient is significantly different from 0 compared to a significance level α. You can only conclude that the correlation is different from 0 with p-value ≤ α. Otherwise, you cannot conclude anything.

2.6 Metrics

Different errors can be used as evaluation metrics to many analyses. Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2) are here defined. In all the equations below, Y_i stands for the observed values, $\hat{Y}_i$ the predicted values, $\overline{Y_i}$ the average of the values being predicted and n the number of data points.

- **MAE**

The mean absolute error is calculated as in Eq. (2.5).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right| \tag{2.5}$$

MAE is an average of the absolute error between a variable and its prediction, measuring the magnitude of the residuals for which all individuals have equal weight. This error uses the same scale as the measured data and therefore cannot be directly compared to MAE of other variables with different scales .

- **MSE**

The mean squared error is calculated as in Eq. (2.6).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|^2 \tag{2.6}$$

MSE is a measure of the quality of an estimator and incorporated both the variance of the estimator (how widely spread are the predictions from the observed data) and its bias (distance from the average). The squaring of the errors gives more weight to larger differences. MSE is expressed by the same unit as the square of the evaluated unit.

- **MAPE**

The mean absolute percentage error is calculated as shown in Eq (2.7).

$$MPE = \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \tag{2.7}$$

MAPE measures the size of the error in percentage terms. This error is scale sensitive and easy to interpret. Usually used for assessing forecast accuracy in statistics methods and in loss functions for machine learning regression problems.

- **R^2**

The coefficient of determination can be calculated through the equations presented in Eq.(2.8),(2.9) and (2.10).

$$R^2 = 1 - \frac{SSE}{TSS} \tag{2.8}$$

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{2.9}$$

$$TSS = \sum_{i=1}^{n} (y_i - \bar{y}_i)^2 \tag{2.10}$$

with SSE the sum of the squared errors and TSS the total sum of the squares. R^2 measures how well the predicted values are replicated by a model, based on the proportion of total variation of the values being predicted.

2.7 Efficiency

The efficiency of a steam generator can be evaluated by two methods: the direct and the indirect. Both provide different results. The direct method accounts the energy gained by the working fluid compared to energy contained in the fuel, while the indirect method includes all heat losses of the system and compares it with the energy input.

The indirect method accounts takes into account several process parameters and imposes difficulties to measure all losses. Therefore, the use of the direct method to calculate the steam generator efficiency parameter presents advantages due to its ease of implementation with instant process data. CHETAN, VIJAY and BHAVESH (2013) and M.RAUT, KUMBHARE and THAKUR (2014) present the efficiency direct method as the ratio of heat output per heat input as in Eq. (2.11):

$$\eta = \frac{Steam\ flow\ rate\ (steam\ enthalpy - feed\ water\ enthalpy)}{(Fuel\ firing\ rate)\ (Higher\ Heat\ Value)}\ 100 \tag{2.11}$$

Considering the fraction of regenerated steam, the direct efficiency can be calculated as in Eq (2.12).

$$\eta = \frac{\dot{m}_{MS}(h_{MS} - h_f) + \dot{m}_{RS}(h_{RS} - h_{SR})}{\dot{m}_{coal}\ HHV}\ 100 \tag{2.12}$$

With $\dot{m}_{MS}$ the main steam flow rate in t/h, $\dot{m}_{RS}$ the reheat steam flow rate in t/h, $\dot{m}_{coal}$ the fuel mass flow in t/h, h_{MS} the main steam enthalpy in kJ/kg, h_f the feed water enthalpy in kJ/kg, h_{RS} the reheated steam enthalpy in kJ/kg, h_{SR} the enthalpy of the steam to be reheated in kJ/kg and HHV the higher heating value of the fuel in kJ/kg.

3 PROBLEM DESCRIPTION

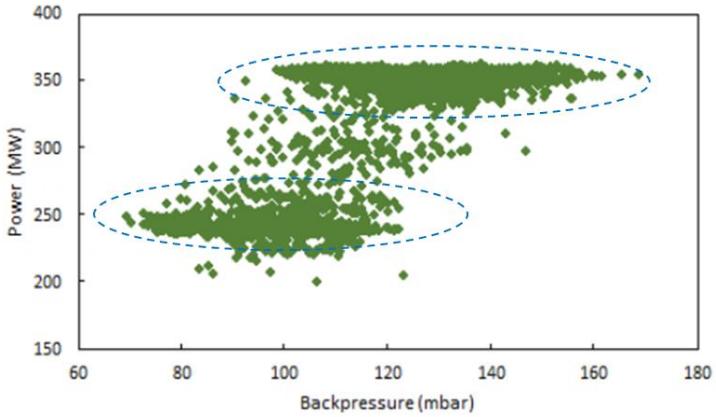
The Pecém power plant is a complex composed of 3 identical and independent generation groups each equipped with a pulverized coal superheated steam generator (SHSG) designed to meet 360MW of electrical generation. The first generator set went into commercial operation in 2012 while the second and third went into operation in 2013. The plant operates under a subcritical Rankine cycle originally designed for high rank coal burning[1]. Figure 3.1 presents an overview of the power plant located in São Gonçalo do Amarante, Ceará.

Figure 3.1 - Pecém Thermal Power Plant and its location.

Source: The author.

Pecém mostly operates in two basic generation ranges, around 240 MW and 360 MW, the maximum rated power. Power demand is dispatched as requested by the National System Operator (ONS) that can be viewed in the data clouds circled in Figure 3.2. Intermediate measurements represent transient operation between the two power levels.

Figure 3.2 - Pecém electrical power as a function of condensation vapor pressure.



Source: The author.

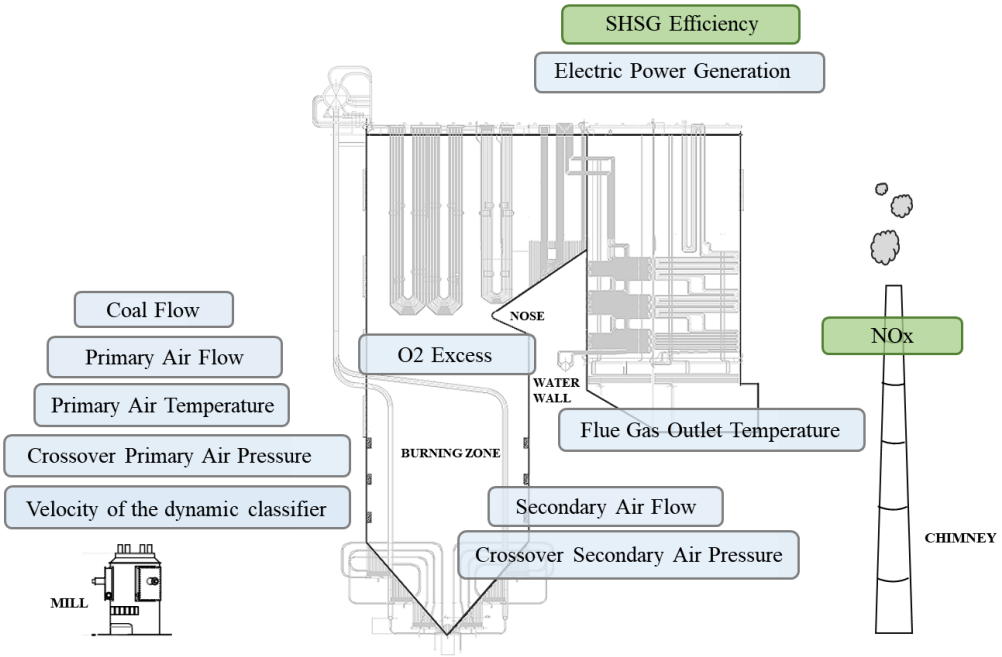
The main object of this study is the Superheated Steam Generator (SHSG) of PECEM, capable of producing 1200 t/h of superheated steam flow at 540 ° C and 18 MPa. Pulverized fuel is introduced to the furnace via twenty-four Low NOx Axial Swirl Burners, completed by twelve after-air ports (Over Fired Air – OFA) to reach complete combustion. The burners are arranged in two rows of six each on the furnace

[1] *Rank* is the term for the degree of evolution/carbonification of the process of transformation of plant matter.

front and rear walls. The OFA ports are arranged in two single rows of six each above the top rows of pulverized fuel burners. Each of the pulverized fuel burners is equipped with co-axial light fuel oil burners which provide for the boiler light up and flame stabilization. The oil burners are able to fire the boiler up to a load of 30% load boiler maximum continuous rating.

Figure 3.3 shows a scheme of the steam generator, mills and chimney, followed by the selected parameters considered in this work.

Figure 3.3 - Scheme of steam generator, mills and chimney along with the studied parameters.



Source: The author.

Descriptions of each input parameter of the system is presented in Table 3.1

Table 3.1 – Input parameters ranges and units

Input parameters	*Minimum*	*Average*	*Maximum*	*Unit*
Primary Air Flow	63.43	78.40	115.71	kg/s
Secondary Air Flow	192.47	246.66	285.45	kg/s
Coal Flow	124.71	139.10	151.41	t/h
Velocity of the Dynamic Classifier	49.71	70.16	102.42	rpm
O2 Excess	0.98	2.732	4.49	%
Primary Air Temperature	316.08	337.16	371.67	°C
Flue Gas Outlet Temperature	330.27	354.77	390.61	ºC
Crossover Primary Air Pressure	74.84	83.17	94.40	mbar
Crossover Secondary Air Pressure	14.60	18.22	23.97	mbar
Electric Power Generation	345.00	355.38	364.85	MW

Source: The author.

The primary air flow carries the pulverized coal from the mills into the furnace and is linked to the flame stability. Primary air flow and temperature also influence in the coal drying. Secondary air flow participates in the combustion process by locally adjusting the stoichiometric relation and, also, enhancing the combustion reaction by promoting the air fuel mixture. The amount of fuel fed in the furnace can be seen through the coal flow measurement. The dynamic classifier is located at the mill and its velocity is responsible for regulating the coal particle size. The O_2 excess indicates the global stoichiometry of the combustion process of the entire SHSG. The crossover duct gets split up in two: one part goes to the mills

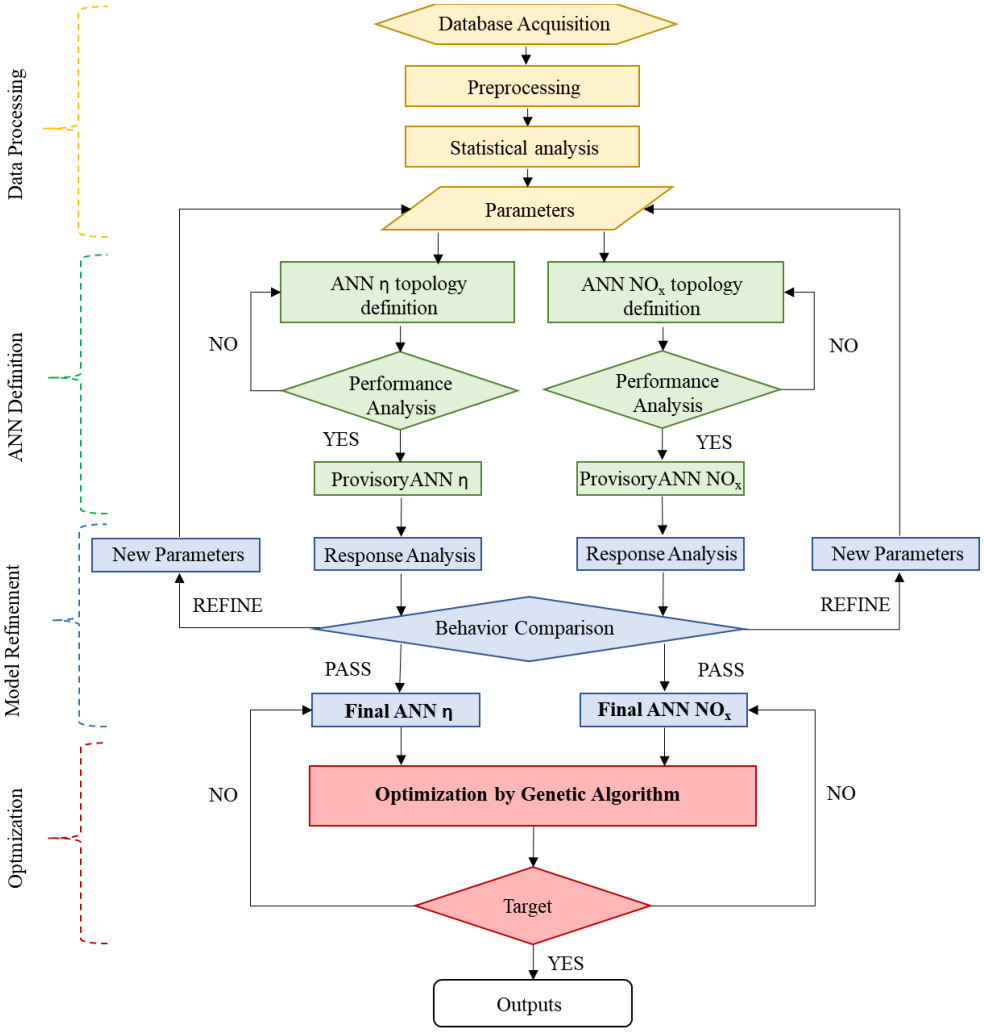
controlling the primary air through its pressure; and the other one controls the pressure of the secondary air flow. This stream splits itself again in the two wind boxes that will lead to the burners. The flue gas flow rate and temperature are measured at the exit of the SHSG. NO_x emissions are measured at the chimney and corrected for the current O_2 concentration. At Pecém powerplant, almost all the NO_x is derived from the fuel formation mechanism, with very little coming from thermal formation. UG2 usually operates NO_x emission ranges of 100 to 800 mg/m_N^3 under normal conditions. The Brazilian National Counsel of the Environment (CONAMA) states that the maximum limit of NO_x emissions is 1000 mg/m_N^3 on dry basis and 3% excess oxygen (CONAMA, 2006).

The present work is part of a large project that seeks to develop a tool able to assist the Pecém thermoelectric power plant operation by predicting the steam generator operating and combustion process behaviors. The work carried out in the next sessions refer to Pecém generating unit 2 (UG2). Analyzes are made for the nominal operating range of the plant from 340 to 365 MW at steady state, since Pecém operates most of the time in this stable condition.

4 METHODOLOGY

The steps followed in this work in order to increase the SHSG efficiency and reduce NOx emissions are represented in the flowchart of Figure 4.1.

Figure 4.1 - Flow chart presenting the steps followed in order to optimize the efficiency of the SHSG and NOx emissions.



Source: The author.

4.1 Data Processing

Plant operating data is constantly collected over time through its SCADA system (Supervisory Control and Data Acquisition), that enables real-time visualization of the plant as well as the download of its data history in spreadsheets.

It is possible to preselect parameters for the database by knowing the problem and delimiting a control volume in the interest object. The input parameters are chosen from a large data set due to two main arguments: its controllability by the operators and their ability to reflect the steam generator operating conditions. Controllable parameters are those that can be directly manipulated by manual command and present an independent behavior among each other. These data are then downloaded as measurement averages within the chosen time interval.

Data preprocessing is an important step for getting accurate results from the model. All measurements are subjected to imperfections that reflect into inadequate data. Data must be queried, summarized and visualized before and after training the models in search of any special pattern, as well as the presence of outliers. The evaluation is made according to three characteristics, which are location (central tendency), variation (dispersion) and shape. Negative and null observations were also eliminated from the dataset.

Statistical analyzes can assess the identification of the best parameters to represent the problem in question. Pearson correlation is used to investigate the relationship between the pairs of continuous variables.

4.2 ANN Definition

Two different ANN models are built in this study, one for modeling efficiency and a different one for NO_x emissions. This approach was chosen to value the freedom of working with different topologies for each response. Therefore, enabling more precision to model the two distinct behaviors. This choice also facilitates building the optimization algorithm that will sweep each of the responses fields.

The number of hidden layers, hidden neurons per each hidden layer and activation functions of the ANN, also known as hyperparameters, were defined through an interactive approach. Hyperparameters configurations were tested by a trial and error method guided by doubling the number of hidden neurons on each try.

In all tested ANN topologies, the input layer has the number of neurons equal to the number of ANN input parameters. The first layer topology tested started with two hidden layers and 8 hidden neurons at each. As the result of the evaluative network metrics and their outputs, presented in subsection 4.4, the network topology is modified in an iterative process until the most appropriate configuration is found within the performance sought.

The evaluation of the developed ANNs is performed using as metrics: MSE, MAE, MPE and R^2. Training and validation plots must be analyzed to prevent overfitting the model.

4.3 Model Refinement

In this step, the already defined ANNs are used as models of the SHSG where statistical analysis through DoE and sensitivity will be applied to study the inputs behaviors.

4.3.1 DoE

DoE is applied to study the correlations between the inputs on one specific output parameter calculated by the already established ANN models. An input parameter can only be removed from the dataset only when it is found to be statically not significant to both models, i.e., for the efficiency and the NO_x ANNs. Whenever that double behavior is identified, new and simpler ANNs can be developed for each output.

The DoE method chosen was the Box-Behnken. Parameters are selected according to their statistical significance, where the high order terms and the interactions between different input parameters are eliminated first, within the significance interval of 95%. Terms with p-value greater than 0.05 must be eliminated according to the hypothesis testing. Residuals are assumed to be normally and independently distributed. Residual plots are checked to assure a precise and unbiased model.

4.3.2 Sensitivity analysis

An ANN sensitivity analysis is performed following a One-Factor-at-a-Time (OFaT) approach. The method consists of varying each factor (input parameter) over its range while the other ones are held constant at a baseline set level. The major disadvantage of the OFAT strategy is that it fails to consider any possible interaction between the factors, justifying the previous use of DoE (MONTGOMERY, 2013). The sensitivity analysis assesses whether a heuristic-based approach to optimize the problem in necessary by scoping for inflexions in the curves. Monotonic curves point that the problem could be solved by classical gradient approach optimization methods.

4.4 Optimization

A combined optimization of the ANN outputs is applied by means of a genetic algorithm, looking for raising the SHSG efficiency while decreasing NOx emissions. The objective function is applied based on the Eq. (2.3) presented in the theoretical background.

5 RESULTS AND DISCUSSION

Results were obtained by applying the steps provided in Figure 4.1. The numbering of the result sections is consistent with that suggested in the methodology.

5.1 Database Processing

It was collected a 9194-sample dataset of the 10 parameters presented earlier on Table 3.1, stored every half hour within the period from January 2018 to May 2019. The steam generator efficiency was added to the database, calculated with the aid of Eq. (2.12), and NO_x emissions were added directly from the supervisory.

According to the Pearson correlation levels presented in Table 2.1, correlations below 0.2 are considered to be very weak or nonexistent. Correlation can only be assumed different than 0 for p-values lower or equal to the significance interval, in this case 0.05. Parameters that presented a correlation lower than 0.2 in respect the efficiency of the steam generator and p-value lower or equal to 0.05 were: O2 excess, crossover secondary air pressure and electric power generation. Parameters that presented a correlation lower than 0.2 with the NOx emissions and p-value lower or equal to 0.05 were: crossover secondary air pressure, crossover primary air pressure and electric power generation.

A table containing all the parameters Pearson correlations and the p-values can be found at APPENDIX A.

5.2 ANN Definition

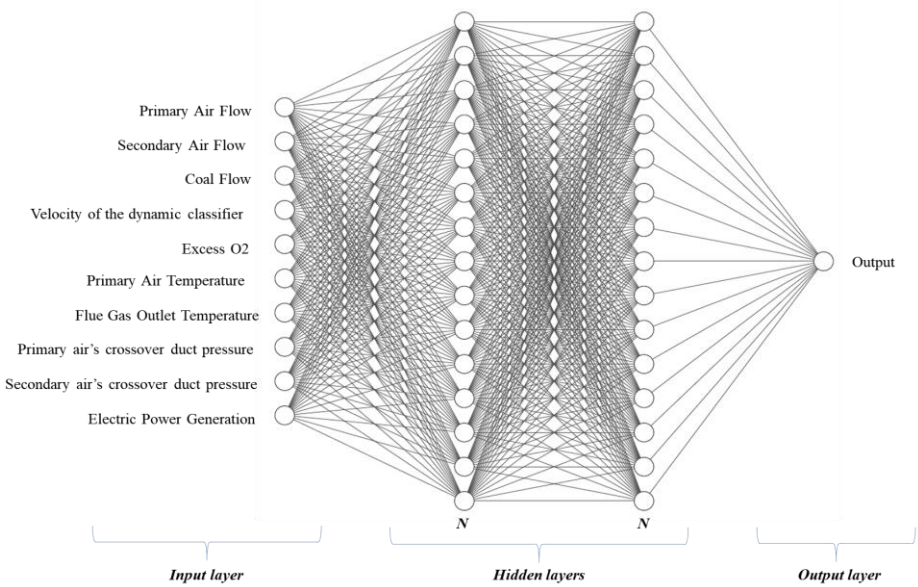
Dataset was randomized and divided into 70% training, 15% test, 10% validation and 5% sample to be used to create the two neural networks, for efficiency and NOx emissions Input parameters were standardized in respect to each of their standard deviations.

Dataset was randomized and divided into 70% training, 15% test, 10% validation and 5% sample. Two different neural networks were developed separately to predict each of the sought outputs. Input parameters were standardized in respect to each of their standard deviations.

The topology of the ANN hyperparameters followed the approach presented in the methodology. Combinations of 8, 16, 32 and 64 hidden neurons with 2 to 3 hidden layers were evaluated. The number of neurons in each hidden layer was either kept constant or changed in the second layer by assuming half the neurons of the first hidden layer. Tested activation functions were the hyperbolic tangent (*tanh*) and the rectified linear unit (ReLU).

Figure 5.1 presents an illustrative scheme of the ANN topologies, composed by 10 nodes for the input layer, 2 hidden layers with N number of neurons and the output layer, with one single neuron for the response (efficiency or NOx).

Figure 5.1 - Illustrative topology of the ANNs developed.



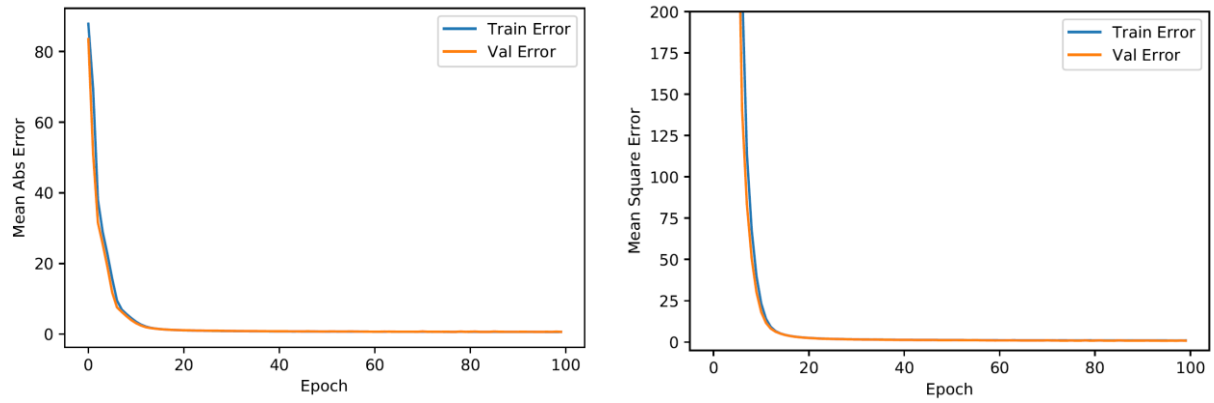
Source: The author.

The ANNs were developed using the Python programming language through the Jupyter Notebook[2] compiler. The Keras[3] programming interface, made available by the Tensorflow[4] machine learning library, was used for its construction.

5.2.1 Efficiency

Different topologies were tested in order to define the best suited to describe the efficiency behavior. The chosen topology was built with one input layer containing the 10 input parameters, two hidden layers of 32 nodes each and one output layer. In the first hidden layer the hyperbolic tangent activation function was used and in the second, ReLU. Training was performed with 100 epochs using as loss function MSE. The MSE and MAE of the training were respectively 0.8647 and 0.6033. The MSE and MAE of the test were respectively 0.7572 and 0.6206 (Figure 5.2).

Figure 5.2 - Training and validation MAE and MSE for efficiency.



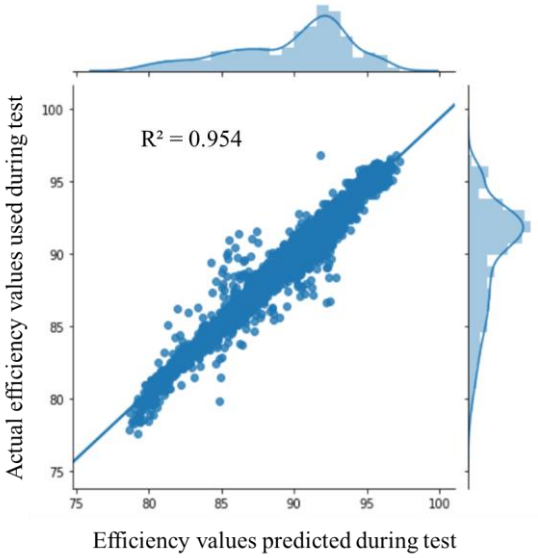
Source: The author.

[2] https://jupyter.org/try
[3] https://keras.io/
[4] https://www.tensorflow.org/

Both MAE and MSE errors decreased exponentially before completing 40 epochs. Both curves were already stable with 100 epochs so training could stop. The proximity of the train and validation errors was an important point to observe in order to avoid model overfitting.

Another model validating procedure was to plot actual values of efficiency from our test database against ANN predicted efficiency values, as presented in Figure 5.3.

Figure 5.3 - Actual efficiency values plotted against ANN predictions.

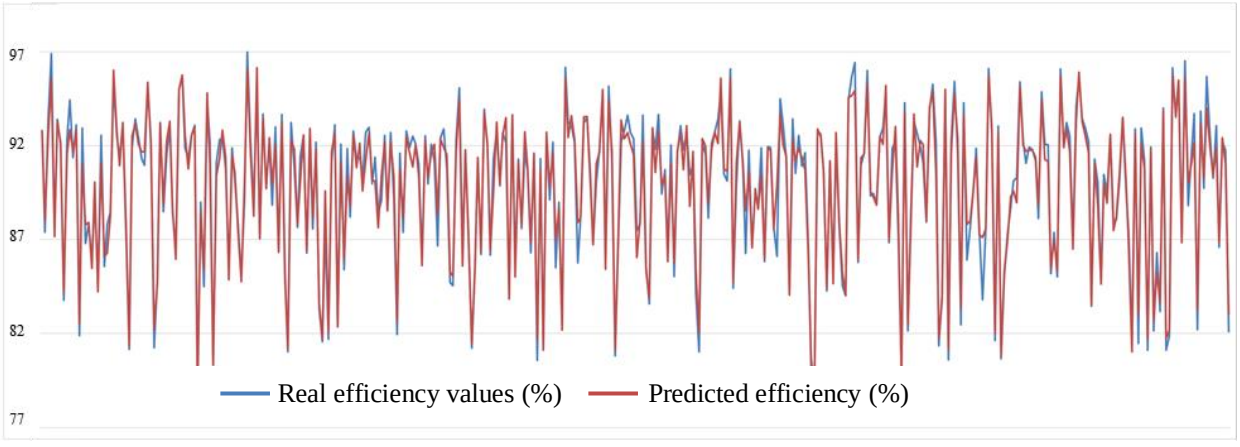


Source: The author.

The model was able to predict the efficiency behavior with a R^2 of 0.954, with little data dispersion and only few points more distanced from the regression line, that are likely to be outliers.

Sample set with unseen data was inputted to analyze whether the model was generalizing the efficiency behavior. MPE and MSE of the predicted values against the historical ones were respectively 0.63% and 0.6449. Figure 5.4 shows the data plotted in a line chart to compare the behavior learned by the ANN with the one observed in the supervisory.

Figure 5.4 - Predicted values of efficiency along with data from the supervisory.



Source: The author.

The trend of behavior was very similar, with its inflections coinciding while only at extreme points of observations, of very low or very high values, the ANN prediction curve was smoother, indicating an adequate generalization of the model.

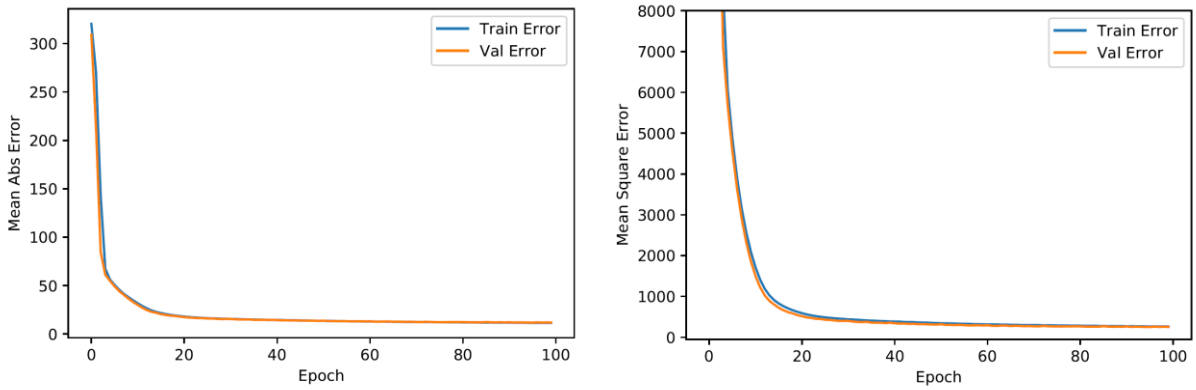
A table containing the tested topologies and its performances can be found at the APPENDIX B.

5.2.2 NO_x

Different topologies were also studied to define the best suited to describe NO_x behavior. The chosen topology was composed with one input layer, 2 hidden layers of 64 nodes each and one output layer. Hyperbolic tangent activation function was used in the first hidden layer and ReLU in the second. Training was performed with 100 epochs using as loss function MSE. The MSE and MAE of the training were respectively 247.04 and 11.25. The MSE and MAE of the test were respectively 312.43 and 12.36. NO_x model errors are higher when compared to efficiency. That can be explained not only by its more complex behavior, but especially due to the errors character that are relative to the range of values the response can assume.

MAE and MSE evolution throughout training and validation can be seen in the graphs below as function of the epochs in Figure 5.5.

Figure 5.5 - Training and validation MAE and MSE for NOx emissions

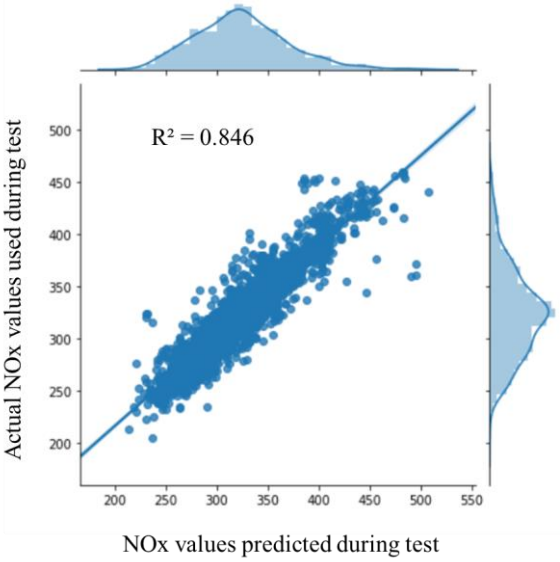


Source: The author.

Both MAE and MSE errors decreased exponentially before completing 40 epochs. Training stopped at 100 epochs. The proximity of the train and validation errors indicate that the model is not overfitting.

Actual NO_x emission values from the test are plotted against NO_x values predicted by the ANN in Figure 5.6.

Figure 5.6 - Actual NOx emission values plotted against ANN predictions

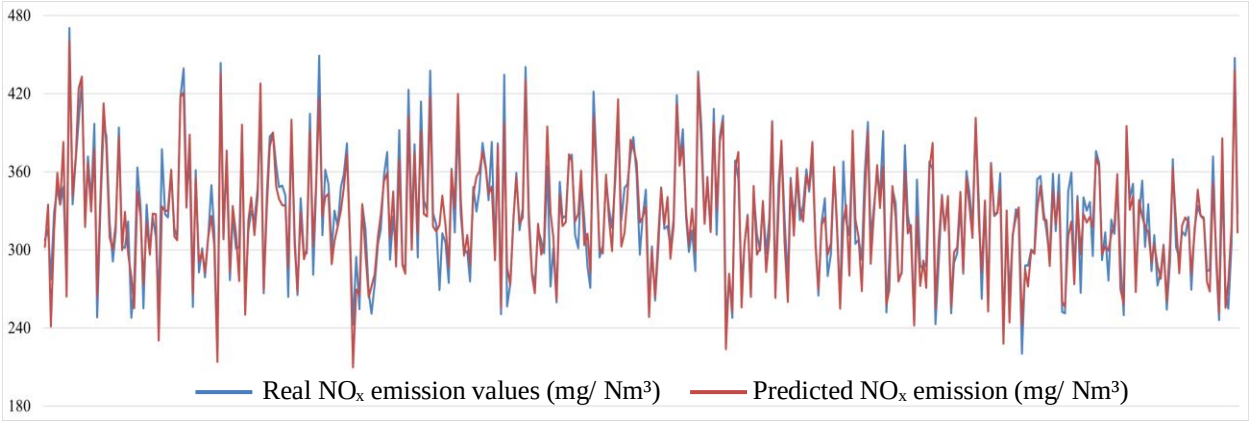


Source: The author.

The model was able to predict the NO_x behavior with a R^2 of 0.846, with more data dispersion than in the efficiency model. More distanced points are likely to be outliers or points outside normal operation.

The unseen sample set was inputted in the model to analyze its NO_x behavior generalization capacity. MPE and MSE of the predicted values compared to the historical ones were respectively 3.49% and 207.80. Figure 5.7 shows the data plotted in a line chart to compare the behavior learned by the ANN with the one observed.

Figure 5.7 - Predicted values of NOx emissions along with data from the supervisory.



Source: The author.

The model has more associated error and deviation due to NO_x behavior complexity, but is still able to learn and generalize it. The trend of both lines is similar, with almost all its inflections coinciding. Extreme points with values farther from the average, have smoother curves and more deviation, still indicating the generalization of a complex response model.

The table containing the tested topologies and its performances can be found at the APPENDIX C

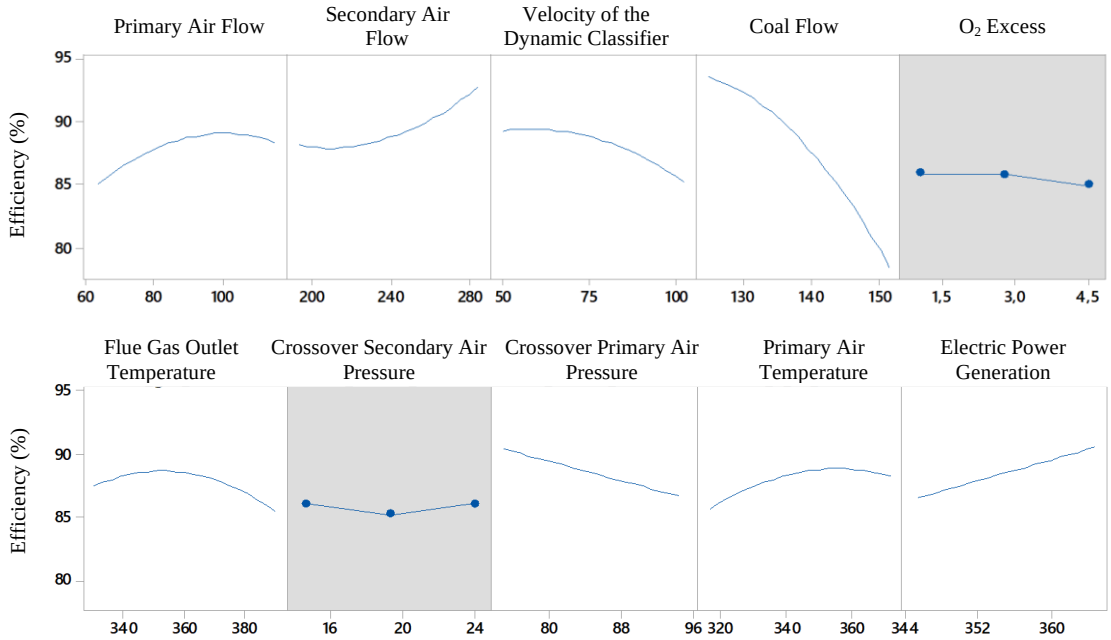
5.3 Model Refinement

DoE methodology and sensitivity analyzes were performed after the models were established.

5.3.1 DoE

The main effects of the parameters on efficiency and NO_x responses were evaluated separately. Main effects for the SHSG efficiency can be seen in Figure 5.8.

Figure 5.8 - Main Effects Graphs for Efficiency.

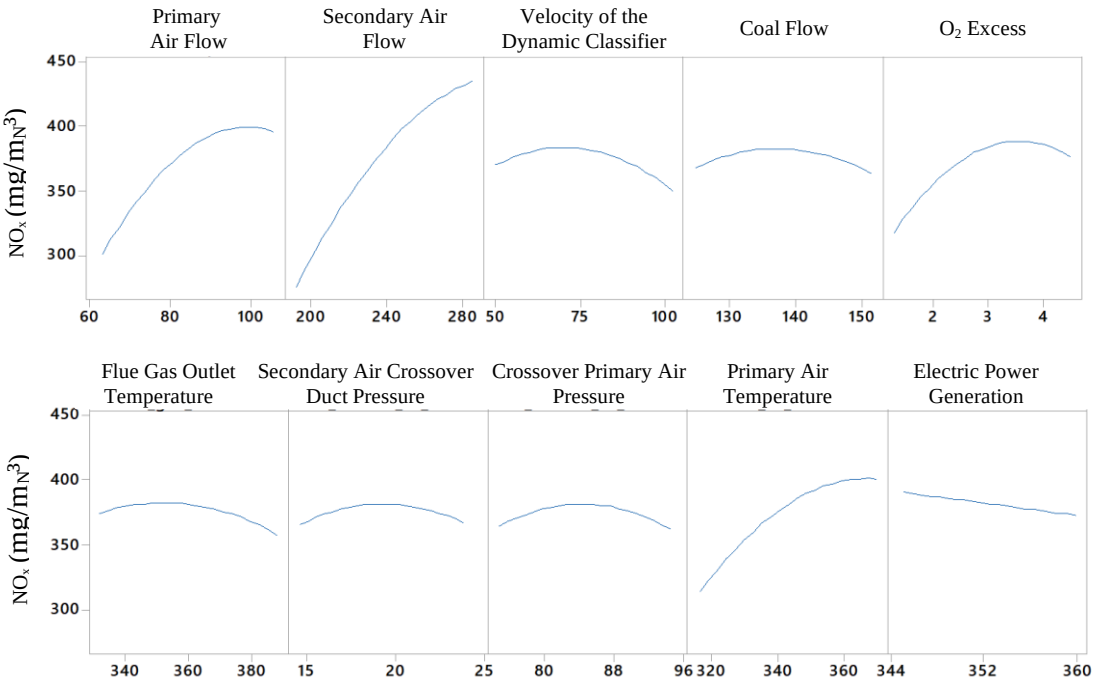


Source: The author.

Most parameters showed to be important to the model, with coal flow the one that impacts the most on the result with a negative correlation. Gray boxes highlight the parameters that were found as not statistically significant to the model, showing that the efficiency was not affected by O2 excess and crossover secondary air pressure.

The same statistical method was applied to verify direct influence of the parameters on the NO_x emissions behavior. The graphs with the main effects are presented in Figure 5.9.

Figure 5.9 - Main Effects Graphs for NOx.



Source: The author.

DoE has shown that for the NO_x model all the parameters were statistically significant, unlike in the efficiency analysis. The most influential input was the secondary air flow with a positive correlation.

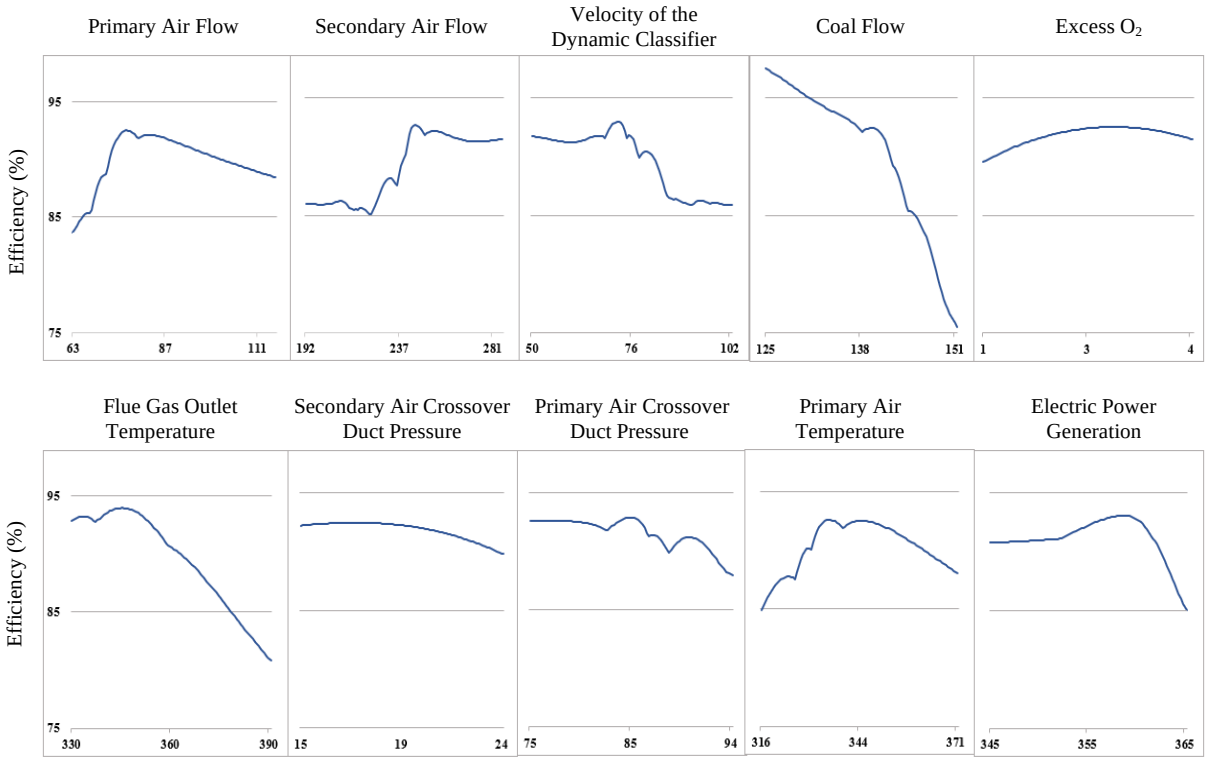
Efficiency and NO_x presented different results for the importance of the parameters on their behaviors. Therefore, no input parameter could be deleted from the models and the ANNs remained the same.

The analysis was performed using the Minitab[5] software.

5.3.2 Sensitivity analysis

The OFaT sensitivity analysis was performed varying each of the input parameters for 100 different steps along its defined range. The efficiency model sensitivity to each parameter is presented in Figure 5.10

Figure 5.10 - Model sensitivity analysis for efficiency response.



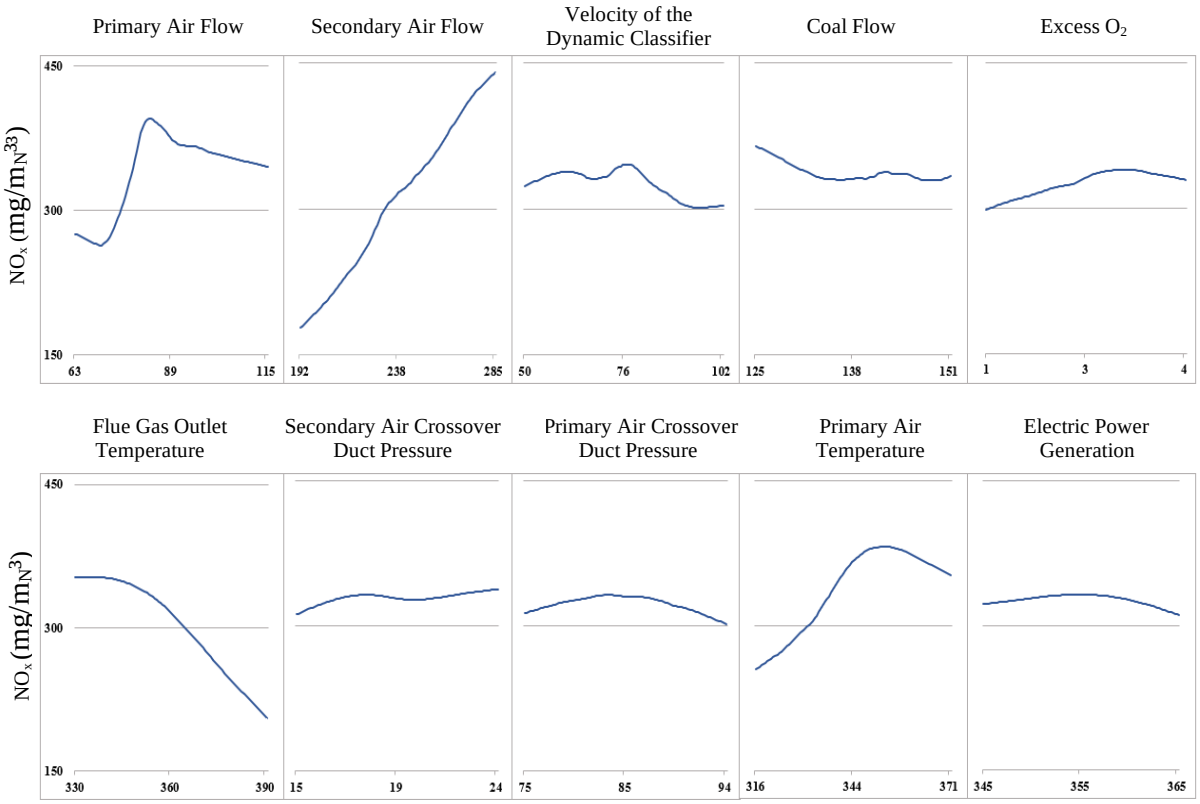
Source: The author.

Several local minimum and maximum points can be observed in the sensitivity graphs. Parameters such as the Secondary Air Flow and Velocity of the Dynamic Classifier presented complex curves that are ideal for the application of heuristic-based optimization algorithms. O2 excess and crossover secondary air pressure presented themselves less sensible, as expected according to their non-significance to the response as the DoE has shown.

Sensitivity analysis of the NO_x emissions model is presented in Figure 5.11.

[5] https://www.minitab.com/

Figure 5.11 - Model sensitivity analysis for NOx response.



Source: The author.

The analysis showed a less sensitive behavior of NOx emissions to disturbances caused by the parameters, when compared with the efficiency curves. However, the parameters also presented inflexions that justified the use of a genetic algorithm. Primary Air Flow and the Velocity of the Dynamic Classifier presented the most complex behaviors among them.

5.4 Optimization

The objective function was based on the one proposed in Eq. (2.3) and aimed to find the input parameters that minimize the deviations in respect to the target values of efficiency (98%) and NOx emissions (220 mg/m_N^3), as shown at Eq. (5.1).

$$\min f = a\left(98 - \eta_{pred}\right) + b([NO_x]_{pred} - 220) \tag{5.1}$$

with η_{pred} the predicted efficiency (%), $[NO_x]_{pred}$ the predicted emissions (mg/m_N^3) and a and b the weights to ponder SHSG efficiency and NOx emissions as deemed relevant.

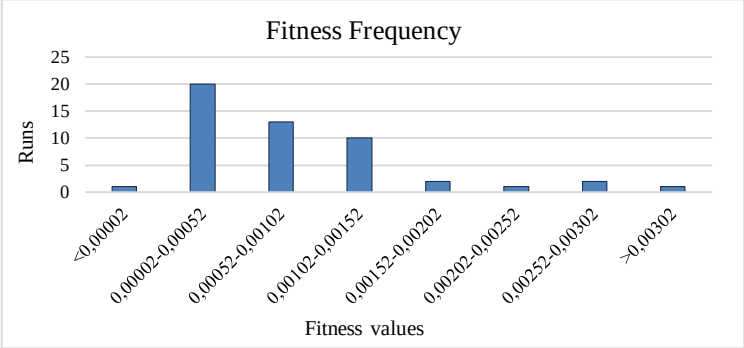
NO_x was normalized from 0 to 100 to assume the same range of the efficiency and facilitate weighting their contributions. The boundaries of the input parameters respected their standardized operating ranges presented in Table 3.1. Population was initialized according to those boundaries, with random integer individuals from -3 to 3. Individuals previously analyzed and known to return low fitness values were inserted in the initial population to guarantee good offspring of the next generations. The mutation used was the flip bit type with a 10% probability for each individual. A two-point crossover of the input sequence was applied with a 50% probability, where the two individuals were modified in place and both kept their original length. The selection of the individuals that pass to the next generation was made through tournament, where 3 randomly chosen individuals compete against each other comparing their fitness values.

Different combinations of population and generation numbers were tested to compare the fitness efficiency and NO_x values returned by the genetic algorithm. The fitness function was pondered to weight both responses equally, with 50% importance each. The three proposed configurations were run 50 times in order to generate the next histograms. Different combinations of *a* and *b* in Eq. (5.1) were tested increasing the efficiency importance.

- Population of 300 individuals and 1000 generations

The fitness values of each result returned by the genetic algorithm were computed and displayed in the form of a histogram presented below in Figure 5.12.

Figure 5.12 - Frequency histogram of fitness values for a GA with 300 individuals and 1000 generations

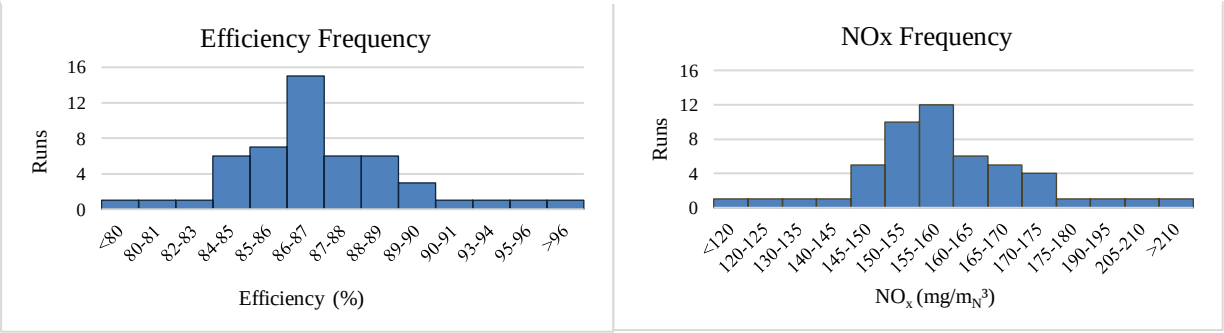


Source: The author.

20 over 50 runs returned fitness values that ranged from 0.00002 to 0.00052. Occurrences steadily decreased while fitness values increased.

Efficiency and NOx occurrence frequencies were also collected and are presented in Figure 5.13.

Figure 5.13 - Frequency histogram of efficiency and NOx values returned for a GA with 300 individuals and 1000 generations.



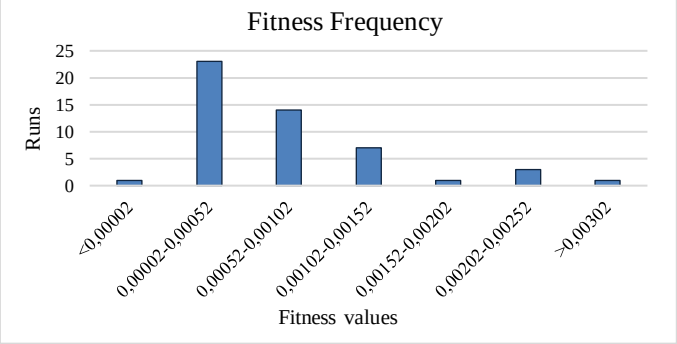
Source: The author.

Efficiency and NO_x values displayed normal distributions centered in the ranges of 86-87%and 155-160 mg/m_N^3, respectively.

- Population of 300 individuals and 500 generations

The fitness values computed throughout the 50 GA runs for the second individual-generation configuration is presented in Figure 5.14.

Figure 5.14 - Frequency histogram of fitness values for a GA with 300 individuals and 500 generations.

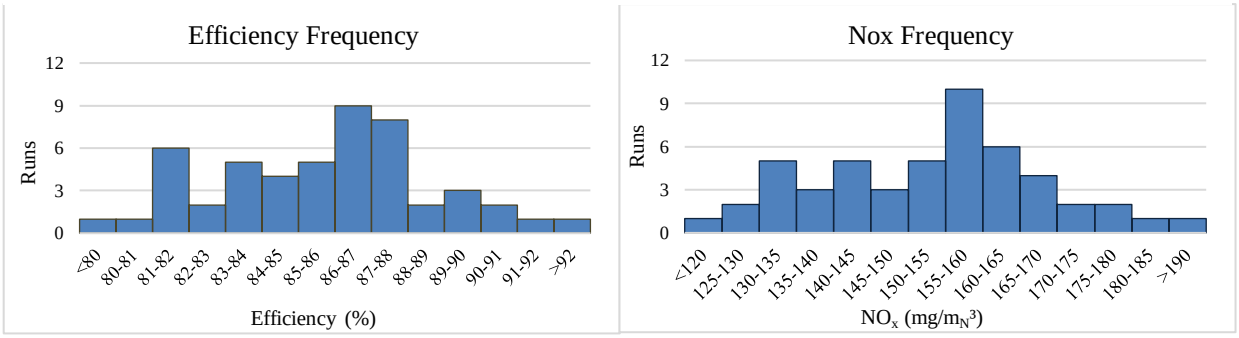


Source: The author.

The fitness values histogram presented a very similar behavior when compared to the first configuration (Figure 5.13), with most of the occurrences in the range of 0.00002 to 0.00052.

Efficiency and NOx occurrence frequency is presented in Figure 5.15.

Figure 5.15 - Frequency histogram of efficiency and NOx values returned for a GA with 300 individuals and 500 generations.



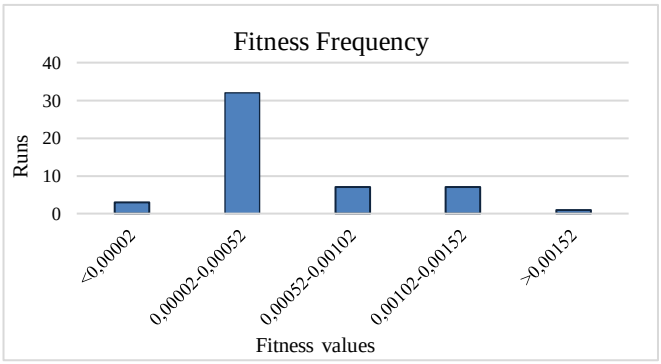
Source: The author.

Though the efficiency and NOx ranges that appeared the most are the same as in the first configuration, the histograms presented more scattered occurrences, less concentrated around the main value.

- Population of 500 individuals and 300 generations

Fitness values returned from the iterations of the GA containing populations of 500 individuals and 300 generations are shown in the graph on Figure 5.16.

Figure 5.16 - Frequency histogram of fitness values for a GA with 500 individuals and 300 generations

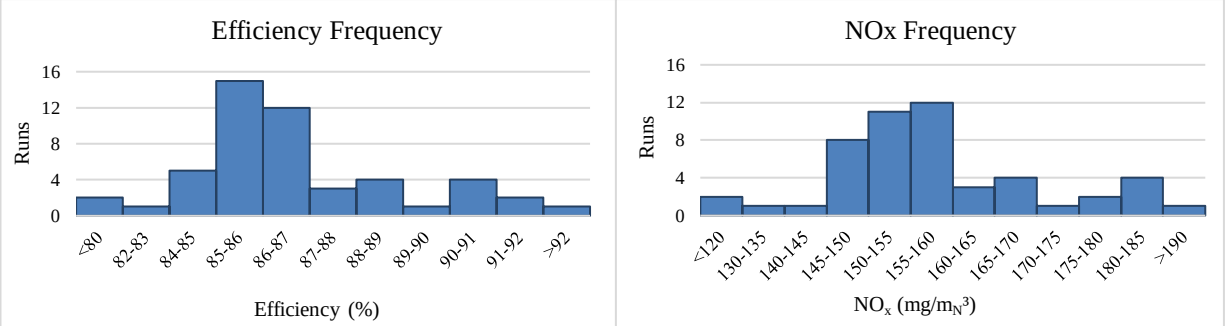


Source: The author.

32 over 50 iterations returned values within the range of 0.00002 – 0.00052 and there were no occurrences higher than 0.00302 as in the first two analysis with smaller populations.

Frequency histograms for efficiency and NOx occurrences are presented in Figure 5.17.

Figure 5.17 - Frequency histogram of efficiency and NOx values returned for a GA with 300 individuals and 500 generations.



Source: The author.

The histograms of efficiency and NOx values presented scattered data in more frequency bins than the first configuration of 1000 generations, but more concentrated when compared to the configuration of 500 generations. The two former graphs display the majority of occurrences in the ranges of 85-86 of efficiency and 155-160 for NOx, as in the other NO_x configurations. The running time is directly related to the number of generations, and therefore this configuration was the faster one.

- Different weights for the variables on the objective function

The first configuration, 300 individuals and 1000 generations, was used for this analysis due to its lesser data dispersion. The weights a and b were adjusted to increase the importance of the efficiency on the optimization algorithm. Three different ponderations were tested: the standard one that weights both efficiency and NOx equally, 75% ponderation of the efficiency versus 25% NOx and 90% efficiency versus 10% NOx. Each of these ponderations was run and recorded 10 times. The average of the values returned for the fitness function and the efficiency and NOx can be seen in Table 5.1.

Table 5.1 - Average of fitness, efficiency and NOx values for different weight combinations of *a* and *b* Efficiency and NOx emission target values of 98% and 220 mg/m_N^3

a	*b*	*Fitness values*	*η (%)*	*NOx (mg/m_N^3)*
0.50	0.50	$4e^{-4}$	86.15	155.35
0.75	0.25	$3e^{-4}$	95.11	171.09
0.90	0.10	$6e^{-4}$	97.95	222.28

Source: The author.

As expected, with the increase of the SHSG efficiency NOx increased as well. The algorithm was able to find different operating points according to the weights of each response. The last combination of *a* and *b*, with more importance to the efficiency, results very close to the targets established of 98% efficiency and 220 $mg/m_N^3$of NOx emissions were found.

Three samples of different operation points for 0.90 *a* and 0.10 *b* can be seen in Table 5.2.

Table 5.2 - Operation points for 0.90 *a* and 0.10 *b*

Primary Air Flow	*Secondary Air Flow*	*Velocity Dynamic Classifier*	*Coal Flow*	*O2 Excess*	*Flue Gas Outlet Temperature*	*Crossover Secondary Air Pressure*	*Crossover Primary Air Pressure*	*Primary Air Temperature*	*Electric Power Generation*	*n*	*Nox*
82.66	256.47	78.18	132.34	2.82	332.89	19.59	74.84	316.40	358.61	**98.12**	**225.83**
83.28	265.12	52.22	127.81	2.01	380.08	23.97	79.20	316.99	360.03	**97.43**	**191.80**
82.66	229.78	67.83	127.52	2.72	336.29	23.91	72.22	319.86	350.90	**97.88**	**213.87**

Source: The author.

Operation point samples for the other combinations of *a* and *b* can be found in APPENDIX D.

The genetic algorithm was developed using the DEAP[6] (Distributed Evolutionary Algorithms in Python) library available in Python.

[6] https://deap.readthedocs.io/en/master/

6 CONCLUSION

The main focus of this work was to model through artificial neural networks the superheated steam generator efficiency and the NOx emissions. Afterwards, a genetic algorithm was applied to perform a combined optimization that aimed to maximize the efficiency while minimizing NOx emissions. The relevance of this study is to work with two different variables of interest that present opposing behaviors. Therefore, operating both of them is not a straightforward task.

The analysis includes as inputs the parameters: primary air flow, secondary air flow, primary air temperature, coal flow, velocity of the dynamic classifier, O_2 excess, flue gas outlet temperature and crossover primary air pressure, crossover secondary air pressure, and electric power generation.

Two different ANNs were developed, one for each response. The best topology for efficiency behavior was found to be an ANN with 2 hidden layers of 32 neurons. MSE and MAE of the training were respectively 0.8647 and 0.6033. MSE and MAE of the test were respectively 0.7572 and 0.6206. The second ANN model developed to describe NO_x emissions behavior was composed of 2 hidden layers of 64 neurons each. MSE and MAE of this ANN's training were respectively 247.07 and 11.25. MSE and MAE of the test were respectively 312.43 and 12.36. It is worth to emphasize that the errors resulted from the NO_x model were higher than the efficiency model mostly due to the character of the errors working at the same higher range of values that NOx can assume.

A DoE approach was applied, through Box-Behnken, to evaluate the importance of the parameters chosen to describe each response. While for efficiency O_2 excess and crossover secondary air pressure were found to be statistically not significant, all input parameters were important to describe the NO_x behavior. Thus, no parameter could be retrieved from the models. Also, the application of a genetic algorithm to search for satisfactory values of efficiency and NOx was validated through a One-Factor-at-a-Time approach to analyze the sensitivity of the models.

Three different combinations of individual and generation numbers were tested for the genetic algorithm proposed. After that, different importance ponderations for efficiency and NO_x in the fitness function were implemented. With 50/50 ponderation for the responses, the genetic algorithm returned an efficiency of 86.15% and NO_x emissions of 155.35 mg/m_N^3. For a 75/25 ponderation favoring efficiency over NOx, the genetic algorithm returned 95.11% of efficiency and 171.09 mg/m_N^3 of NO_x emissions. Lastly, a ponderation of 90/10 of efficiency over NO_x was tested, returning 97.95% of efficiency and 222.28 mg/m_N^3 of NO_x emissions, achieving the proposed target.

The proposed methodology was able to successfully connect three distinct methods and have them work together, each one with its contribution. Through ANN, it was possible to model two completely different responses of the steam generator. Design of Experiments assessed the importance of each input parameter to the responses being analyzed. Finally, the genetic algorithm was able to work with different penalties to find the best configuration of our system within the Brazilian legislation.

REFERENCES

ANEEL. **ANEEL. BIG - Banco de Informações de Geração**. 2019. Disponível em: <http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/capacidadebrasil.cfm>.

CHETAN, Patel T.; VIJAY, Patel K.; BHAVESH, Patel K. Efficiency with different GCV of coal and efficiency improvement. **International Journal of Innovative Research in Science, Engineering and Technology**, [s. l.], v. 2, n. 5, p. 1518–1527, 2013.

CONAMA. Resolucão normativa nº 382/2006 de 26 de dezembro. **Conama**, [s. l.], p. 37, 2006. Disponível em: <http://www.mma.gov.br/port/conama/res/res06/res38206.pdf>

DE, S. et al. **Development of an artificial neural network model for the steam process of a coal biomass cofired combined heat and power (CHP) plant in Sweden**. [s.l: s.n.]. v. 32

EMPRESA DE PESQUISA ENERGÉTICA - EPE. Balanço Energético Nacional 2018- BEN 2018 - Síntese (ano base 2017). [s. l.], p. 62, 2018. Disponível em: <http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-303/topico-397/Relatório Síntese 2018-ab 2017vff.pdf#search=co2>

FERREIRA, S. L. C. et al. Box-Behnken design: An alternative for the optimization of analytical methods. **Analytica Chimica Acta**, [s. l.], v. 597, n. 2, p. 179–186, 2007. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0003267007011671>. Acesso em: 1 dez. 2019.

GHUGARE, Suhas et al. **Prediction of Higher Heating Value of Solid Biomass Fuels Using Artificial Intelligence Formalisms**. [s.l: s.n.]. v. 7

GP STRATEGIES. **Curso de Heat Rate Awareness**, 2013.

HAYKIN, Simon. **Neural Networks and Learning machines**. [s.l: s.n.]. Disponível em: <https://arxiv.org/pdf/1312.6199v4.pdf>

HEBB, D. O. The Organization of Behavior:A Neuropsychological Theory. **New York:Wiley**, [s. l.], 1949.

HOLLAND, J. H. **Adaptation in Natural and Artificial Systems**. [s.l.] : University of Michigan Press, 1975.

IEA. **Global Energy & CO2 Status Report**. 2018a. Disponível em: <https://www.iea.org/geco/>.

IEA. **Global Energy & CO2 Status Report, Coal**. 2018b. Disponível em: <https://www.iea.org/geco/coal/>.

M.RAUT, Sachin; KUMBHARE, Sanjay B.; THAKUR, Krishna C. Energy Performance Assessment of Boiler at P.S.S.K. Ltd, Basmathnagar, Maharashtra State. **International Journal of Emerging Technology and Advanced Engineering**, [s. l.], v. 4, n. 12, p. 1–12, 2014.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, [s. l.], v. 5, n. 4, p. 115–133, 1943. Disponível em: <https://doi.org/10.1007/BF02478259>

MESROGHLI, Sh.; JORJANI, E.; CHEHREH CHELGANI, S. Estimation of gross calorific value based on coal analysis using regression and artificial neural networks. **International Journal of Coal Geology**, [s. l.], v. 79, n. 1–2, p. 49–54, 2009. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0166516209000573>. Acesso em: 6 maio. 2019.

MEYER-BAESE, Anke et al. Genetic Algorithms. **Pattern Recognition and Signal Analysis in Medical Imaging**, [s. l.], p. 135–149, 2014. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780124095458000054?via%3Dihub>. Acesso em: 25 nov. 2019.

MITCHELL, Melanie; STEPHANIE FORREST. Genetic Algorithms and Artificial Life. **Artificial Life**, [s. l.], 1994.

MONTGOMERY, Douglas C. **Design and Analysis of Experiments**. [s.l: s.n.].

ROMANYCIA, Marc H. J.; PELLETIER, Francis Jeffry. What is a heuristic? **Computational Intelligence**, [s. l.], v. 1, n. 1, p. 47–58, 1985.

ROSENBLATT, F. The Perceptron:A probabilistic model for information storage and organization in the brain. **Psychological Review**, [s. l.], v. 65, p. 386–408, 1958.

RUMELHART, D. E.; G. E. HINTON; R.J.WILLIAMS. **Learning internal representations by error propagation**. [s.l: s.n.]. v. 1

SALKIND, Neil J. **Statistics for people who (think they) hate statistics**. [s.l: s.n.]. v. 38

SMREKAR, J. et al. Development of artificial neural network model for a coal-fired boiler using real plant data. **Energy**, [s. l.], v. 34, n. 2, p. 144–152, 2009. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0360544208002880>. Acesso em: 27 fev. 2019.

SONI, Devin. **Introduction to Evolutionary Algorithms**. 2018. Disponível em: <https://towardsdatascience.com/>. Acesso em: 25 nov. 2019.

STRUŠNIK, Dušan; GOLOB, Marjan; AVSEC, Jurij. **Artificial neural networking model for the prediction of high efficiency boiler steam generation and distribution**. [s.l: s.n.]. v. 57

VENKATESWARLU, Ch. et al. Stochastic and evolutionary optimization algorithms. **Stochastic Global Optimization Methods and Applications to Chemical, Biochemical, Pharmaceutical and Environmental Processes**, [s. l.], p. 87–123, 2020. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780128173923000041?via%3Dihub>. Acesso em: 25 nov. 2019.

XU, Gang et al. Comprehensive evaluation of coal-fired power plants based on grey relational analysis and analytic hierarchy process. **Energy Policy**, [s. l.], v. 39, n. 5, p. 2343–2351, 2011. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0301421511000656?via%3Dihub>. Acesso em: 2 maio. 2019.

APPENDIX

APPENDIX A – Pearson Correlations between parameters

	Primary Air Flow	Secondary Air Flow	Velocity of the dynamic classifier	Coal Flow	Excess O2	Flue Gas Outlet Temperature	Secondary air's crossover duct pressure	Primary air's crossover duct pressure	Primary Air Temperature	Electric Power Generation	Efficiency
Secondary Air Flow	-0,487										
	0										
Velocity of the dynamic classifier	0,025	-0,324									
	0,013	0									
Coal Flow	0,445	-0,214	0,009								
	0	0	0,368								
Excess O2	-0,432	0,361	-0,003	-0,254							
	0	0	0,778	0							
Flue Gas Outlet Temperature	0,236	0,097	-0,257	0,243	-0,105						
	0	0	0	0	0						
Secondary air's crossover duct	-0,052	0,179	-0,09	0	0,305	0,167					
	0	0	0	0,979	0	0					
Primary air's crossover duct pressure	0	-0,166	0,089	0,467	0,026	-0,196	-0,015				
	0,973	0	0	0	0,012	0	0,153				
Primary Air Temperature	0,249	0,132	-0,267	0,07	-0,119	0,925	0,119	-0,343			
	0	0	0	0	0	0	0	0			
Electric Power Generation	-0,073	0,04	0,145	-0,132	-0,115	-0,033	-0,078	0,049	-0,004		
	0	0	0	0	0	0,001	0	0	0,732		
Efficiency	-0,288	0,424	-0,212	-0,801	0,15	0,011	0,027	-0,686	0,21	0,141	
	0	0	0	0	0	0,262	0,009	0	0	0	
Nox	0,067	0,589	-0,343	0,006	0,27	0,379	0,108	-0,15	0,502	-0,093	0,235
	0	0	0	0,604	0	0	0	0	0	0	0

APPENDIX B – Different ANNs topologies tested to model the efficiency behavior.

	MPE	MSE	Test	Train
8t_8r	1,03%	1,3934	Mse: 1.6745	Mse: 1.6849
			Mae: 0.9667	Mae: 0.9516
16t_16r	0,81%	1,0377	Mse: 1.0318	Mse: 1.1797
			Mae: 0.7464	Mae: 0.7287
32t_32r	0,63%	0,6449	Mse: 0.7572	Mse: 0.8647
			Mae: 0.6206	Mae: 0.6033
64t_64r	0,64%	0,6035	Mse: 1.1963	Mse: 1.3643
			Mae: 0.7255	Mae: 0.6685
8t_8t_8r	0,70%	0,8295	Mse: 0.9096	Mse: 1.0179
			Mae: 0.6558	Mae: 0.6333
16t_16t_16r	2,52%	8,7532	Mse: 11.1484	Mse: 10.1192
			Mae: 2.3096	Mae: 2.3163
32t_32t_32r	0,70%	0,7656	Mse: 1.1963	Mse: 1.3643
			Mae: 0.7255	Mae: 0.6685
32t_16r	0,72%	0,7854	Mse: 0.9533	Mse: 1.0194
			Mae: 0.7125	Mae: 0.6806
16t_16t	0,73%	0,8980	Mse: 0.9585	Mse: 1.0655
			Mae: 0.6962	Mae: 0.6727
32t_32t	0,58%	0,5245	Mse: 0.6417	Mse: 0.7134
			Mae: 0.5690	Mae: 0.5367
64t_64t	0,49%	0,3653	Mse: 0.6181	Mse: 0.4566
			Mae: 0.4831	Mae: 0.4406
32t_16t	0,68%	0,7482	Mse: 0.8464	Mse: 0.9370
			Mae: 0.6540	Mae: 0.6375
16r_16r	0,71%	0,9508	Mse: 0.9585	Mse: 1.0655
			Mae: 0.6962	Mae: 0.6727
32r_32r	0,65%	0,6450	Mse: 0.7639	Mse: 0.8341
			Mae: 0.6086	Mae: 0.5859
64r_64r	0,51%	0,3584	Mse: 0.5287	Mse: 0.5743
			Mae: 0.5165	Mae: 0.4723
32r_16r	0,63%	0,6950	Mse: 0.8075	Mse: 0.9192
			Mae: 0.6108	Mae: 0.5951

t stands for tanh and *r* for ReLU

APPENDIX C – Different ANNs topologies tested to model the NO_x emissions behavior.

	MEP	MSE	Test	Train
8t_8r	5,32%	467,137	Mse: 541.9567	Mse: 501.0870
			Mae: 17.3118	Mae: 16.7990
16t_16r	4,35%	323,837	Mse: 426.8316	Mse: 378.1133
			Mae: 14.9274	Mae: 14.1961
32t_32r	4,89%	390,231	Mse: 515.6654	Mse: 454.1765
			Mae: 16.5319	Mae: 15.5698
64t_64r	3,49%	207,798	Mse: 312.4339	Mse: 247.0415
			Mae: 12.3622	Mae: 11.2536
8t_8t_8r	4,93%	383,862	Mse: 472.9458	Mse: 432.9626
			Mae: 16.2378	Mae: 15.5535
16t_16t_16r	3,44%	205,953	Mse: 332.6059	Mse: 256.2370
			Mae: 12.5953	Mae: 11.4979
32t_32t_32r	4,38%	384,632	Mse: 331.9112	Mse: 229.0484
			Mae: 11.7421	Mae: 11.0400

t stands for tanh and *r* for ReLU

APPENDIX D – Operation points of the input parameters that reached desired efficiency and NO_x emissions.

- $a = 0.5; b = 0.5$

Primary Air Flow	*Secondary Air Flow*	*Velocity Dynamic Classifier*	*Coal Flow*	*O2 Excess*	*Flue Gas Outlet Temperature*	*Crossover Secondary Air Pressure*	*Crossover Primary Air Pressure*	*Primary Air Temperature*	*Electric Power Generation*	*n*	*Nox*
66.56	223.02	77.00	139.06	2.71	344.55	23.97	83.47	316.65	355.05	**86.52**	**157.39**
67.22	216.90	62.70	142.95	2.82	354.19	15.12	83.47	316.51	355.05	**85.06**	**149.41**
81.23	214.94	70.46	134.13	2.82	354.20	18.57	82.58	316.64	355.05	**86.74**	**158.56**

- $a = 0.75; b = 0.25$

Primary Air Flow	*Secondary Air Flow*	*Velocity Dynamic Classifier*	*Coal Flow*	*O2 Excess*	*Flue Gas Outlet Temperature*	*Crossover Secondary Air Pressure*	*Crossover Primary Air Pressure*	*Primary Air Temperature*	*Electric Power Generation*	*n*	*Nox*
78.11	246.60	52.79	128.39	1.58	372.33	20.35	71.54	319.61	358.92	**96.24**	**191.22**
88.33	246.60	54.06	135.94	1.52	368.53	16.52	78.35	316.95	355.05	**96.62**	**197.40**
78.04	242.88	50.34	129.94	2.47	354.20	14.65	83.47	316.07	357.96	**94.63**	**164.86**