

**SANTA CATARINA STATE UNIVERSITY – UDESC
CENTER OF TECHNOLOGICAL SCIENCES – CCT
GRADUATE PROGRAM IN APPLIED COMPUTING – PPGCA**

DOUGLAS MACEDO SGROTT

**A MULTI-OBJECTIVE DATA-DRIVEN EVOLUTIONARY ALGORITHM APPLIED
TO STEEL MODELING**

JOINVILLE

2022

DOUGLAS MACEDO SGROTT

**A MULTI-OBJECTIVE DATA-DRIVEN EVOLUTIONARY ALGORITHM APPLIED
TO STEEL MODELING**

Master thesis submitted to the Computer Science Department at the College of Technological Science of Santa Catarina State University in fulfillment of the partial requirement for the Master's degree in Applied Computing.

Supervisor: Rafael Stubs Parpinelli

JOINVILLE

2022

Para gerar a ficha catalográfica de teses e
dissertações acessar o link:
<https://www.udesc.br/bu/manuais/ficha>

Macedo Sgrott, Douglas

A multi-objective data-driven evolutionary algorithm
applied to steel modeling / Douglas Macedo Sgrott. -
Joinville, 2022.

104 p. : il. ; 30 cm.

Supervisor: Rafael Stubs Parpinelli.

.
Dissertação (Mestrado) - Universidade do Estado
de Santa Catarina, Centro de Ciências Tecnológicas,
Programa de Pós-Graduação em Computação Aplicada,
Joinville, 2022.

1. Palavra-chave. 2. Palavra-chave. 3. Palavra-chave.
4. Palavra-chave. 5. Palavra-chave. I. Stubs Parpinelli,
Rafael. II. , . III. Universidade do Estado de Santa
Catarina, Centro de Ciências Tecnológicas, Programa de
Pós-Graduação em Computação Aplicada. IV. Título.

DOUGLAS MACEDO SGROTT

**A MULTI-OBJECTIVE DATA-DRIVEN EVOLUTIONARY ALGORITHM APPLIED
TO STEEL MODELING**

Master thesis submitted to the Computer Science Department at the College of Technological Science of Santa Catarina State University in fulfillment of the partial requirement for the Master's degree in Applied Computing.

Supervisor: Rafael Stubs Parpinelli

EXAMINATION BOARD:

Dr. Rafael Stubs Parpinelli
UDESC (president/supervisor)

Members:

Dr. Yuri Kaszubowski Lopes
UDESC

Dr. Fabiano José Fabri Miranda
ArcelorMittal VEGA

Joinville, 22nd of February, 2022

I dedicate this work to my mother, who has been
throughout my life a constant source of light.

ACKNOWLEDGEMENTS

I would like to deeply thank Rafael Parpinelli for the immense patience he had with me during this journey and for showing me a door towards a new horizon. Not only is he an example of academic excellence, but also an incredible human being.

I would also like to thank my mother, for without her, nothing of this would be possible. She's an example that, bridging the gap between what is possible and impossible, is a matter of love.

To ArcelorMittal for enabling this work to be financially viable, and also Fabricio Cerqueira, Fabiano Miranda and José Francisco for aiding this work and believing in the partnership between ArcelorMittal and UDESC that made all of this feasible. To my friends, that brought warmth (but also beer and coffee) to keep me from falling. To my sister, that held me up when I fell. To all my teachers that gave me multiple new perspectives from which the world can be admired.

Last but not least, to everyone that passed through my life in the past two pandemic years. Despite these trying times, by looking back at past experiences, I can see hope in going forward.

"Simplicity is the ultimate sophistication"

Leonardo da Vinci

ABSTRACT

Steel is one of the world's most important engineering and construction material, being used in every aspect of our lives. Consequently, the research and development of novel steels with higher performance can yield economic benefits across a wide spectrum of industries. However, not all steels are made equally, and these differences can be made apparent under a microscope or evaluating their mechanical behavior. Hence, in this work is proposed a data-driven system using Machine Learning (ML) capable of predicting steel's mechanical properties, as well as optimize chemical composition and processing parameters in order to design new steels with specific mechanical properties. For the prediction task, an Artificial Neural Network (ANN) is trained on a metallurgical dataset using Automated Machine Learning (AutoML). Then, the ANN is coupled with a Multi-Objective Genetic Algorithm (MOGA) in order to evolve new steel designs in a surrogate-based optimization algorithm. Case studies were conducted in order to provide experimental validation of the prediction model, and computational validation of the optimization scheme.

Keywords: Artificial Neural Network. Optimization. Material Design. Steel. Machine Learning.

RESUMO

O aço é um dos materiais de engenharia e construção mais importantes do mundo, sendo usado em todos os aspectos de nossas vidas. Consequentemente, a pesquisa e o desenvolvimento de novos aços com desempenho superior podem gerar benefícios econômicos em um vasto espectro de indústrias. No entanto, nem todos os aços são criados exatamente da mesma forma, e essas diferenças podem ser aparentes sob um microscópio ou quando avaliando sua caracterização mecânica. Assim, neste trabalho é proposto um sistema baseado em dados usando Aprendizagem de Máquina capaz de prever as propriedades mecânicas do aço, assim como otimizar a composição química e os parâmetros de processamento, a fim de projetar novos aços com propriedades mecânicas específicas. Para a tarefa de predição, uma Rede Neural Artificial é treinada em um conjunto de dados metalúrgicos usando Aprendizado de Máquina Automatizada. Em seguida, a ANN é acoplada a um Algoritmo Genético Multi-Objetivo para desenvolver novos designs de aços em um algoritmo de otimização baseado em modelo substituto. Estudos de caso foram conduzidos a fim de fornecer validação experimental do modelo de predição e validação computacional do esquema de otimização.

Palavras-chave: Rede Neural Artificial. Otimização. Design de Materiais. Aços. Machine Learning.

LIST OF FIGURES

Figure 1 – Schematic illustration of a Strain-Stress curve and yielding point. Source: Own authorship.	25
Figure 2 – Graphic overview of different steel grades in relation to strength and elongation. Source: Hall e Fekete (2017)	26
Figure 3 – Schematic illustration of TRIP effect during deformation. Source: Inspired by Soleimani, Kalhor e Mirzadeh (2020)	29
Figure 4 – Illustration of a general supervised learning scheme for training and inference stage. Source: Own authorship.	30
Figure 5 – Representation of (a) neuron and (b) neural network. Source: Inspired by Haykin (1994)	31
Figure 6 – Example of a Pareto front with two objectives for minimization: Cost and number of faults. Source: Own authorship.	35
Figure 7 – Pareto front with two objectives, being f_1 and f_2 . Non-dominated points in blue, dominated points in orange. A and B are Pareto optimal solutions that dominate C . Source: Own authorship.	36
Figure 8 – Example of a Pareto ranking from a non-dominated sorting. Source: Own authorship.	41
Figure 9 – The different possible layers of a surrogate-based optimization process: real-world process, computational model and surrogate. Source: Own Inspired by Stork, Eiben e Bartz-Beielstein (2020)	42
Figure 10 – Example of a graphical HV calculation. Source: Inspired by Fonseca, Paquete e Lopez-Ibanez (2006).	44
Figure 11 – Example of a graphical IGD calculation. Source: Own authorship.	45
Figure 12 – Flowchart of the entire computational system. Source: Own authorship.	56
Figure 13 – Proposed surrogate modeling flowchart. Source: Own authorship.	58
Figure 14 – Proposed surrogate-based optimization flowchart. Source: Own authorship.	60
Figure 15 – Piece-wise linear function with unitary slope, where the global optima are located at p_1 and p_2	61
Figure 16 – Online parameter control using deterministic rules for Recombination and Mutation. Source: Own authorship.	63
Figure 17 – Stages of the steel manufacturing process and involved parameters. Source: Own authorship.	64
Figure 18 – SHAP feature importance ranking for ANNs trained on different steel grades.	67
Figure 19 – SHAP feature importance ranking for ANNs trained on (a) completely imbalanced and (b) less imbalanced data.	69
Figure 20 – Box-plot of the IGD indicator for different recombination (CX) and mutation (MUT) parameters. Source: Own authorship.	71

Figure 21 – Box-plot of the DCI indicator for different recombination (CX) and mutation (MUT) parameters. Source: Own authorship.	72
Figure 22 – Best ANN found for different activation functions, such as (a) ReLu, (b) Sigmoid and (c) Tanh. Source: Sgrott et al. (2021).	73
Figure 23 – Training and validation loss of training epochs using the NAS architecture for the ReLU activation function. Source: Sgrott et al. (2021).	74
Figure 24 – Sensitivity Analysis for the three different ANNs. Source: Sgrott et al. (2021).	75
Figure 25 – SHAP feature importance ranking for ANNs with different activation functions.	75
Figure 26 – Sensitivity analysis derived from the DOE results. Source: Own authorship.	77
Figure 27 – Average and standard deviation for each steel grade optimization metrics, namely HV, D_N^{PW} and IGD, for random steels. Source: Own authorship.	78
Figure 28 – Example of an optimized population’s genotypic and phenotypic distribution. Target symbol, star symbol and red dots represents the reference steel, its most similar optimized steel and the remaining population, respectively. Source: Own authorship.	78
Figure 29 – Pareto front from a random IF steel optimization. Source: Own authorship.	79
Figure 30 – Average and standard deviation for each steel grade optimization metrics, namely HV, D_N^{PW} and IGD, for steels with the largest associated prediction error. Source: Own authorship.	80
Figure 31 – Examples of an optimized population’s genotypic and phenotypic distribution for different steel. Target symbol, star symbol and red dots represents the reference steel, its most similar optimized steel and the remaining population, respectively. Source: Own authorship. Source: Own authorship.	81
Figure 32 – Pareto front from (a) a single optimization and (b) combined populations generated from all optimizations regarding a DP and TRIP steel with the largest associated prediction error. Source: Own authorship.	82
Figure 33 – Average and standard deviation for each steel grade optimization metrics, namely HV, D_N^{PW} and IGD, for a hypothetical and utopic steel. Source: Own authorship.	84
Figure 34 – Pareto front from (a) a single optimization and (b) combined populations generated from all optimizations regarding an IF steel with utopic mechanical property objectives. Source: Own authorship.	84
Figure 35 – Flowchart for a bottoms-up approach of steel development integrating a prediction model, optimization algorithm and a specialist’s knowledge. Source: Own authorship.	85
Figure 36 – Visual representation of the search scope’s upper and lower bounds in relation to (a) the dataset histogram and (b) the optimized populations genotypic and phenotypic histogram. Small red lines are discrete values from the histogram. Source: Own authorship.	87

Figure 37 – Mechanical properties box plots for different batches of steel experiments.	
Source: Own authorship.	89
Figure 38 – Mechanical properties scatter plots for different batches of steel experiments.	
Source: Own authorship.	90

LIST OF TABLES

Table 1	– Table regarding the number of found, filtered and selected studies from the SLR.	49
Table 2	– Concept Matrix from the SLR. <i>Reference</i> are the articles selected for study. <i>IN1</i> , <i>IN2</i> , <i>IN3</i> , <i>IN4</i> is the presence of microstructure, chemical composition, processing parameters and mechanical properties respectively as independent variables and <i>Output</i> the variable being predicted or optimized. <i>Algorithm</i> refers to the set of algorithms and techniques used in the research, while <i>Quant. Analysis</i> and <i>Quali. Analysis</i> is the presence of a quantitative and qualitative analysis, respectively.	50
Table 3	– Search space for the NAS parameters.	58
Table 4	– Description from the models variables, where CE, PP, DIM and MP stands for Chemical Element, Process Parameter, Dimension and Mechanical Property, respectively.	65
Table 5	– Proportions of each steel grade in each dataset.	68
Table 6	– Mean Relative Error from each steel grade obtained by different models trained on different subsets of data.	70
Table 7	– R^2 metric with mean and standard deviation using 10-fold cross-validation. .	73
Table 8	– DOE parameters setup.	76
Table 9	– Relative error from experimental validation.	77
Table 10	– Comparison between the real mechanical properties of the steel with the largest associated prediction error, its prediction values and the averaged predictions from the optimized populations.	80
Table 11	– Lower and upper bounds for the optimization's search space. Source: Own authorship.	86
Table 12	– Description of the different optimization targets. Source: Own authorship. . .	87
Table 13	– DOE parameters setup for advanced steel development case study.	88
Table 14	– List of all steel experiments proposed by specialist. All parameters in the table are measured as <i>ppm</i> . Source: Own authorship.	92
Table 15	– Overview from the results obtained from all steel experiments regarding their requirements satisfiability. Source: Own authorship.	93

LIST OF ABBREVIATIONS AND ACRONYMS

AI	Artificial Intelligence
AHSS	Advanced High-Strength Steel
ANN	Artificial Neural Network
ASE	Academic Search Engine
AutoML	Automated Machine Learning
BH	Bake Hardening
CART	Classification And Regression Tree
CD	Crowding Distance
CFD	Computational Fluids Dynamics
CRM	Cold Rolling Mill
DCI	Diversity Comparison Indicator
DOE	Design Of Experiments
DP	Dual Phase
EA	Evolutionary Algorithm
EL	Elongation
ES	Evolutionary Strategy
EP	Evolutionary Programming
FEM	Finite Element Method
GA	Genetic Algorithm
GDM	Genotypic Diversity Measure
GEP	Gene Expression Programming
GP	Genetic Programming
IF	Interstitial Free
IGD	Inverse Generational Distance
HDG	Hot Dip Galvanizing line
HSLA	High-Strength Low-Alloy
HSM	Hot Strip Mill
HSS	High-Strength Steel
HV	Hypervolume
LAB	Laboratory

MB-SGD	Mini Batch Stochastic Gradient Descent
ML	Machine Learning
MLP	Multi-Layer Perceptron
MOEA	Multi-Objective Evolutionary Algorithm
NAG	Nesterov Accelerated Gradient
NAS	Neural Architecture Search
NSGA	Non-dominated Sorting Genetic Algorithm
PSO	Particle Swarm Optimization
Q&P	Quench and Partitioning
RF	Random Forest
ReLU	Rectified Linear Unit
RQ	Research Question
SGD	Stochastic Gradient Descent
SHAP	Shapley Values
SLR	Systematic Literature Review
SM	Steel Mill
SPEA	Strength Pareto Evolutionary Algorithm
SVM	Support Vector Machine
Tanh	Hyperbolic Tangent
TRIP	Transformation Induced Plasticity
TWIP	Twinning-Induced Plasticity
TS or UTS	Tensile Strength
xAI	Explainable Artificial Intelligence
YS	Yield Strength

LIST OF SYMBOLS

Al	Aluminium
B	Boron
C	Carbon
Cr	Chrome
d_i^p	Distance between point i from the Pareto front and p points from the approximation front
D_{PW}^N	Pairwise diversity measurement
e_f	Elongation of a steel after it's fracture point
F	Pareto front
f_N	fitness objective regarding N parameter
$f(X)$	fitness function
HRC	Hardness Rockwell C
L1	Regularization using the absolute magnitude of the neural weights in the cost function
L2	Regularization using the squared magnitude of the neural weights in the cost function
L_a	Original gauge length of a steel
L_f	Fracture length of a steel
MAE	Mean Absolute Error
Mn	Manganese
MPa	Mega Pascal
MRE	Mean Relative Error
Mo	Molybdenum
N	Nitrogen
Nb	Niobium
Ni	Nickel
NMDF	Normalization with Maximum Diversity so Far
P	Phosphorus
PWLIF	Piece-Wise Linear Interpolation Function
PWLF	Piece-Wise Linear Function

POP_{SIZE}	Population size
r	Reference point
R^2	Coefficient of Determination
S	Sulfur
Si	Silicon
Ti	Titanium
V	Vanadium
x	Generic optimization's decision variable vector
w	Synaptic weights from the artificial neural network
X_{01-29}	Independent parameters used in modeling and as decision variables in optimization
Y_{01-03}	Dependent parameters used in modeling
λ_m	m-dimensional Lebesgue measure

CONTENTS

1	INTRODUCTION	20
1.1	MOTIVATION	22
1.2	OBJECTIVES	22
1.3	DOCUMENT STRUCTURE	22
2	BACKGROUND	24
2.1	STEEL TYPES AND PROPERTIES	24
2.1.1	STEEL PROPERTIES	24
2.1.1.1	<i>YIELD STRENGTH</i>	24
2.1.1.2	<i>TENSILE RESISTANCE</i>	24
2.1.1.3	<i>ELONGATION</i>	25
2.1.2	STEEL TYPES	26
2.1.2.1	<i>INTERSTITIAL FREE STEELS</i>	27
2.1.2.2	<i>BAKE HARDENING STEELS</i>	27
2.1.2.3	<i>HIGH-STRENGTH LOW-ALLOY STEELS</i>	27
2.1.2.4	<i>DUAL-PHASE STEELS</i>	28
2.1.2.5	<i>TRANSFORMATION INDUCED PLASTICITY STEELS</i>	28
2.2	DATA-DRIVEN APPROACH AND MACHINE LEARNING	29
2.2.1	SUPERVISED LEARNING	29
2.2.1.1	<i>ENSEMBLE MODELS</i>	30
2.2.2	ARTIFICIAL NEURAL NETWORKS	31
2.2.2.1	<i>NEURAL NETWORK PARAMETERS</i>	32
2.2.2.1.1	NUMBER OF HIDDEN LAYERS AND NEURONS	32
2.2.2.1.2	ACTIVATION FUNCTION	32
2.2.2.1.3	REGULARIZATION PARAMETERS	32
2.2.2.1.4	OPTIMIZER PARAMETERS	33
2.2.2.2	<i>NEURAL ARCHITECTURE SEARCH</i>	34
2.3	MULTI-OBJECTIVE OPTIMIZATION	34
2.3.1	PARETO OPTIMALITY	36
2.3.2	MULTI-OBJECTIVE EVOLUTIONARY ALGORITHMS	37
2.3.2.1	<i>GENETIC ALGORITHMS</i>	38
2.3.2.2	<i>NONDOMINATED SORTING GENETIC ALGORITHM II (NSGA-II)</i>	40
2.3.3	SURROGATE BASED EA OPTIMIZATION	41
2.3.4	OPTIMIZATION METRICS	43
2.3.4.1	<i>HYPERVOLUME INDICATOR</i>	43
2.3.4.2	<i>INVERTED GENERATIONAL DISTANCE</i>	44

2.3.4.3	<i>DIVERSITY COMPARISON INDICATOR</i>	45
2.3.4.4	<i>GENOTYPIC DIVERSITY MEASURE</i>	45
3	SYSTEMATIC LITERATURE REVIEW	47
3.1	RESEARCH METHOD	47
3.2	PLANNING THE REVIEW	47
3.3	CONDUCTING THE REVIEW	48
3.4	SUMMARY AND DISCUSSIONS	49
3.4.1	DIVERSITY OF MODELED STEEL CATEGORIES	51
3.4.2	INDEPENDENT VARIABLES	51
3.4.3	MODELING ALGORITHMS	52
3.4.4	OPTIMIZATION ALGORITHMS	53
3.4.5	QUANTITATIVE ANALYSIS	53
3.4.6	QUALITATIVE ANALYSIS	54
3.4.7	SLR CONSIDERATIONS	55
4	PROPOSED APPROACH	56
4.1	NEURAL NETWORK MODELING	57
4.1.1	NEURAL ARCHITECTURE SEARCH	57
4.2	SURROGATE-BASED MOEA FOR STEEL DESIGN OPTIMIZATION . .	59
4.2.1	OBJECTIVE FUNCTIONS	60
4.2.2	EA PARAMETER CONTROL	62
5	RESULTS AND ANALYSIS	64
5.1	NEURAL NETWORK MODELING	64
5.1.1	DATABASE	64
5.1.2	IMBALANCED DATASET AND ITS CONSEQUENCES	66
5.1.2.1	<i>STRATIFIED MODELS</i>	66
5.1.2.2	<i>UNIFIED MODEL</i>	68
5.1.3	QUANTITATIVE RESULTS	69
5.2	OPTIMIZATION ANALYSIS	70
5.2.1	ONLINE PARAMETER CONTROL ANALYSIS	71
5.3	CASE STUDIES	72
5.3.1	CASE STUDY 1: MODELING AND INFERENCE OF NOVEL BH INDUSTRIAL STEELS	72
5.3.1.1	<i>IF AND BH MODELING</i>	72
5.3.1.2	<i>INFERENCE OF NOVEL EXPERIMENTAL BH STEEL</i>	76
5.3.2	CASE STUDY 2: STEEL OPTIMIZATION EXPERIMENTS	77
5.3.3	CASE STUDY 3: STEEL DISCOVERY EXPERIMENTS	79
5.3.4	CASE STUDY 4: STRESSING THE OPTIMIZATION ALGORITHM .	83

5.3.5	CASE STUDY 5: ACCELERATED STEEL DEVELOPMENT	85
<i>5.3.5.1</i>	<i>STEEL OPTIMIZATION</i>	86
<i>5.3.5.2</i>	<i>SPECIALIST KNOWLEDGE AND DESIGN OF EXPERIMENTS</i>	88
<i>5.3.5.3</i>	<i>RESULTS</i>	88
6	CONCLUSIONS AND FUTURE RESEARCH DIRECTION	94
	BIBLIOGRAPHY	96

1 INTRODUCTION

Through several decades of continuous research and development, steel has proven itself a versatile and efficient material for many applications across different industries and the world in general. During this time, the evolving requirements for safety, durability and economy have driven steel producers and engineers to work together to develop new grades of steel aimed at the unique requirements of various areas of application. By carefully controlling processing and chemistry parameters, steel can be tailored to provide optimal performance for specific tasks. Hence, there are more than 200 individual grades of steel available today based on the different needs of a vast spectrum of industries (HALL; FEKETE, 2017).

Aside from the traditional manufacturing parameters that influence steel's mechanical behavior, it has become clear that the microstructure has a significant role in steel's characterization, particularly for modern and complex steel grades. Through the incorporation of alloying elements that go beyond the base iron and carbon, combined with modifications of mechanical and thermal processing conditions, the different crystal phases of steel, namely austenite, ferrite, cementite and martensite, are transformed into a wide range of microstructures and substructures (KRAUSS, 2017). Although invisible to the naked eye, these microstructures, with their different crystallinity and metallic bonding, governs the steel's mechanical behavior in the microscopic levels and can, thereafter, directly influence its macroscopic properties.

As steel research and development make progress, new steel grades and categories are discovered and manufactured. However, between the 1st generation Advanced High Strength Steel (AHSS) widely used in industries today and the novel steels from the 2nd generation AHSS, there is a gap. This gap represents the possibility to expand the current metallurgy body of knowledge, offering new novel steels with high performance in different applications and excellent mechanical behavior under reasonable costs (HALL; FEKETE, 2017). Thus, the search for technical, theoretical, mathematical and computational tools to aid in the development of steel from any grade can open new doors for industrial applications and marketing opportunities, advancing the limits of the current steel manufacturing landscape.

Such tools are widely used today. In Chen et al. (2020), mechanical properties were predicted for steels manufactured using Quench & Partitioning (Q&P) hot stamping process, using an integrated analytical model of microstructure evolution coupled with Finite Element Method (FEM). The use of data-driven models is also becoming increasingly common, modeling upon hard data instead of domain-specific rules. The use of such models can be seen in many recent works, such as in Wu et al. (2018) where an Artificial Neural Network (ANN) was modeled in order to predict the mechanical properties of the S360 steel alloy and coupled with a Strength Pareto Evolutionary Algorithm II (SPEA II) algorithm, in order to design new chemical compositions that achieves a defined set of mechanical properties. Still, there is an ongoing trend to compensate analytical models with data-driven techniques and vice-versa, such that physical metallurgy principles are included in Machine Learning (ML) models. According

to Shen et al. (2019), where a Support Vector Machine was modeled and used by a Non-dominating Sorting Genetic Algorithm II (NSGA-II) for steel designs, with and without the use of microstructural information. As reported by the authors, the computational system that did not model microstructure data obtained from software based on physical metallurgy principles exhibited less accuracy and optimized steels that were not feasible to manufacture in realistic conditions, as demonstrated by the authors' experimental validation.

In this work, a computational system intended for steel design is proposed using bio-inspired surrogate-based optimization algorithms. This system is divided into two parts.

In the first part, an ANN is created to be used as a surrogate model based on Interstitial Free (IF), Bake Hardening (BH), High-Strength Low-Alloy (HSLA), Dual Phase (DP) and Transformation Induced Plasticity (TRIP) steels. Using this ML modeling technique, many variables obtained in different manufacturing stages, each with its characteristics and complexities, can be modeled in relation to the steel's mechanical behavior. Since more than one property is being predicted, the ANN can be characterized as a multi-variate regression model (CHOI; PARK, 2001), mapping the functional relationship between inputs and multiple outputs using supervised learning (ZHOU; WU, 2011). The ANN architecture and hyperparameters are optimized using a Neural Architecture Search (NAS) in order to automate and optimize the modeling stage (SGROTT et al., 2021). The ANN inputs are chemical composition, processing parameters and dimensional information, while its outputs are mechanical properties, namely Tensile Strength (TS or UTS), Yield Strength (YS) and Elongation (EL).

In the second part, the ANN is coupled in a Multi-objective Evolutionary Algorithm (MOEA) as a surrogate model, where its predictions are used in the fitness function, evaluating the different optimized individuals evolved over generations. The MOEA's goal is to obtain a population of evolved individuals which presents mechanical properties sufficiently close to a desired set of targeted mechanical properties. Furthermore, each individual corresponds to a steel's chemical composition, processing parameters and dimensional information. In other words, the ANN inputs define the optimization search space, while its outputs, i.e., mechanical properties, corresponding to the decision space. Since there are conflicting relationships between the steel's mechanical behavior, NSGA-II was employed in order to handle the optimization's candidate solutions Pareto optimality (DEB et al., 2002).

Five case studies were performed in order to provide computational or experimental validation. In the first study, a NAS optimized model trained on IF and BH steels was used to predict the mechanical properties of a novel BH steel with high Niobium content under different galvanization temperatures.

The second, third and fourth case studies were performed to investigate the surrogate-based optimization's performance under different contexts, in order to understand its capabilities to compromise between realistic and novel steel designs.

Finally, the fifth study case exemplifies an end-to-end industrial application, connecting optimization and prediction stages in a methodology that promotes accelerated steel development.

1.1 MOTIVATION

There are many factors involved in steel manufacturing, such as manufacturing limitations and footprint, cost, sustainability, mechanical behavior and the feasibility to use specific steel in different applications, such as stamping, welding and painting (HALL; FEKETE, 2017). Consequently, the search for new and cheaper applicable steel designs that presents the intended mechanical properties under different applications can augment the steel manufacturing landscape and offer new products with new layers of sophistication. Thus, this defines space for improvement in steel manufacturing and material design through theoretical and computational models.

1.2 OBJECTIVES

The main goal of this work is to develop a system that is capable of proposing new chemical compositions and processing parameters in order to manufacture steels with intended mechanical properties. To accomplish that, the search mechanism used in the design optimization procedure must be **(a)** efficient in converging towards solutions that present small discrepancies between desired and optimized properties and **(b)** able to generate a wide pool of genetic diversity, such that multiples steels with different designs can be compared and analyzed. To achieve convergence while maintaining a diversified set of solutions, the search mechanism must handle multi-objective problems.

Following, some secondary objectives are listed, which will build the steps to accomplish the main objective:

- Accurately model different steel grades using ANNs, such that the surrogate model is sufficiently reliable.
- Gain a deeper understanding from the surrogate model by applying explainable artificial intelligence techniques.
- Develop the system employing bio-inspired optimization algorithms capable of dealing with the aforementioned problems, such as MOEA.
- Build an analysis procedure in order to appraise the hypothetical steels.

1.3 DOCUMENT STRUCTURE

The remaining chapters of this master's thesis are organized as follows: Chapter 2 presents the theoretical background required to comprehend this work, providing some metallurgical and computational information; Chapter 3 presents the Systematic Literature Review (SLR), concerning computational methods use in steels prediction and design; Chapter 4 provides a top-down and detailed description of the proposed system; Chapter 5 provides the information

about how the experiments were conducted, their results and analysis; finally, Chapter 6 presents the conclusions and some future directions.

2 BACKGROUND

This chapter provides the necessary metallurgical and computational background to ground the present study, especially regarding the use of data-driven methods for steel design and mechanical properties prediction.

2.1 STEEL TYPES AND PROPERTIES

According to Paupler (1988), industrial application of steel usually involves forces or loads, in which the steel may be a member or part of a machine or structure. Thus, it is necessary to understand the limiting values that the steel can withstand without physical failures.

Since there are many different applications of steels, many steels are created with different characteristics and behaviors. These characteristics arise from the difference in chemical composition and the relation of multiple complex metallurgical processes and parameters in steel production (HALL; FEKETE, 2017). When mechanical behavior is understood in terms of physical metallurgical principles, it can be possible to improve the mechanical properties or at least to control them (PAUPLER, 1988).

2.1.1 STEEL PROPERTIES

Although steels have many types of properties, such as electrical, magnetic, chemical and thermal, this work focuses on mechanical properties, specifically yield strength, tensile resistance and elongation.

2.1.1.1 *YIELD STRENGTH*

Yield strength can be defined as the stress which will begin to inflict permanent deformation on the steel, known as plastic deformation (CALLISTER; RETHWISCH, 2010). In engineering terms, this can be described as the limit of usable elastic behavior. As this limit is exceeded, plastic deformation begins to occur, which is permanent and non-reversible.

The material response is linear until the upper yield point is reached. As the plastic deformation of the material increases, the metal becomes stronger, a phenomenon named strain hardening. The yield point phenomenon is illustrated in Figure 1, where a typical stress-strain is presented. The y-axis denotes stress, which expresses the external forces acting over the cross-sectional area of the steel sample and the x-axis represents strain, a measure of the sample's deformation.

2.1.1.2 *TENSILE RESISTANCE*

The tensile strength (TS or UTS) is the engineering stress at the maximum of the stress-strain curve as indicated in Figure 1, indicating the maximum load divided by the original cross-sectional area of the steel specimen being measured. As strain increases beyond TS, stress

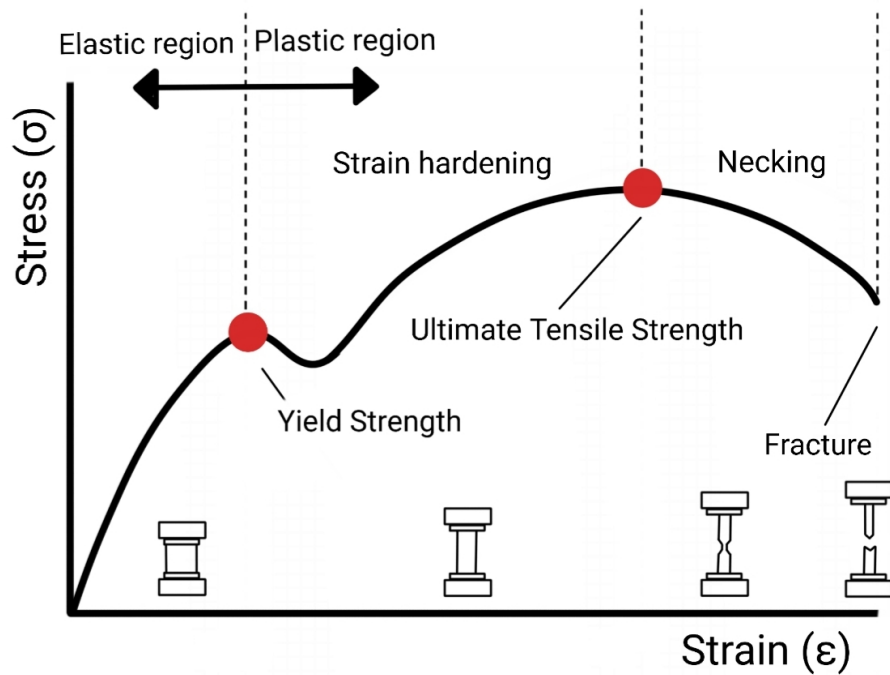


Figure 1 – Schematic illustration of a Strain-Stress curve and yielding point. Source: Own authorship.

begins to decrease, forming a neck, where the local cross-sectional area becomes smaller until fracture begins (CALLISTER; RETHWISCH, 2010).

In practice, this property can be regarded as a measure of the maximum load which the specimen can withstand under uniaxial loading. As stated in Paupler (1988), although this value is of little fundamental importance in regards to the strength of a metal under more realistic conditions, because of the long practice of using the tensile strength to determine the steel's strength, it is still a useful property of a material. Besides, this property is easy to determine and easily reproducible, corroborating its usefulness for specifications and quality control.

2.1.1.3 ELONGATION

Elongation is a measure of ductility obtained from the tension test and as such, describes how much plastic deformation a material can sustain before a fracture occurs. Elongation can also serve as an indicator of changes in impurity level or processing conditions.

Elongation can be obtained after fracture from a destructive tension test by putting the specimen back together and measuring L_f and L_a and calculating the percent of elongation according to Equation 1.

$$e_f = \frac{L_f - L_a}{L_a} \quad (1)$$

where L_f is the fracture length and L_a is the original gauge length.

2.1.2 STEEL TYPES

There are many different types of steel due to different chemical compositions and processing parameters used in its manufacturing. This results in steels with unique microstructure and different phases at the microscopic level, resulting in a wide range of mechanical properties and mechanical behaviors (KRAUSS, 2017). Generally, we can group some of these steels in grades, so that most steels pertaining to the same grade have similarities between them, whether in alloying elements, manufacturing stages and technologies, processing parameters, microstructure and, primarily, by its mechanical property range. Figure 2 illustrates the characterization and superposition of different steel grades concerning its strength and ductility, in the x-axis and y-axis, respectively.

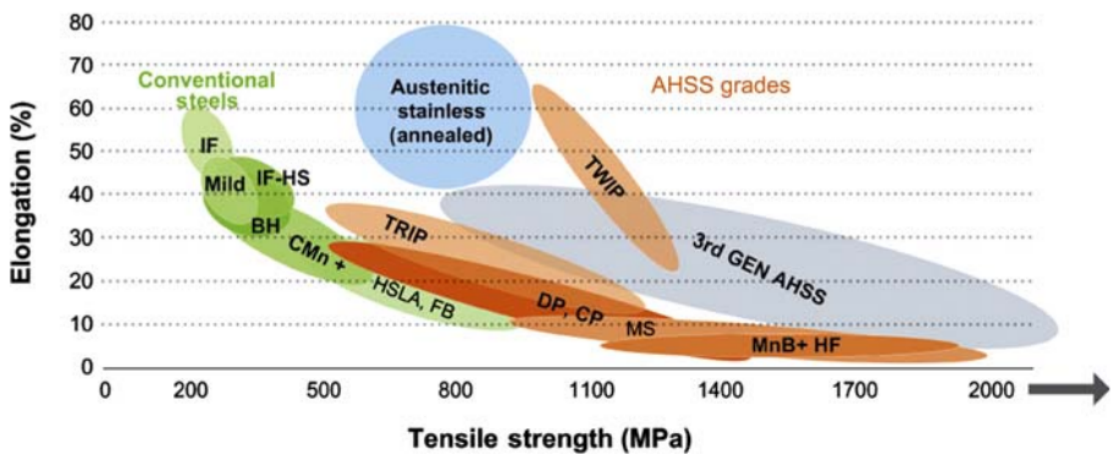


Figure 2 – Graphic overview of different steel grades in relation to strength and elongation.
Source: Hall e Fekete (2017)

In order to develop new steel grades with novel characteristics, research was done in many areas to enhance the understanding of steel manufacturing and its microscopic peculiarities. Coupled with a historical and technological context, this enables the categorization of different steel grades in specific clusters, such as Mild Steels (also called low-carbon steels), Conventional High-Strength steel (HSS), 1st generation Advanced High-Strength steel (AHSS), 2nd generation AHSS and 3rd generation AHSS.

Mild steels, Conventional HSS and 1st generation AHSS are widely used steels across several industries. Some of the grades in this category were developed and manufactured over many decades, but this doesn't mean that there are not new challenges, especially regarding 1st generation AHSS. There is still a lack of a complete understanding of the many factors that influence complex metallographic phases and microstructures. In some cases, even when this understanding is achieved, there can still be a lack of sufficient technological progress to manufacture novel steels. All the steels studied in this work belong to these categories and will be described briefly.

Second generation AHSS was developed to simultaneously provide high strength and formability and incorporate grades such as austenitic stainless steel and Twinning-Induced

Plasticity (TWIP) steels. High formability can be achieved with the presence of austenite. However, since austenite is unstable in ambient temperature, high amounts of alloying elements, including nickel and manganese, are needed to stabilize austenite in ambient temperature. High strength levels can be achieved by twinning or transforming austenite into martensite under strain. Nonetheless, despite the higher mechanical properties, these steels are costly to manufacture and can contain other limitations in their applicability because of problems associated with the high alloying content, such as cracking issues with traditional spot welding (HALL; FEKETE, 2017).

Still, by examining Figure 2, a gap can be found between 1st and 2nd generation AHSS. This gap indicates the possibility of new steel grades that could provide the best combination of strength, formability and weldability, at a reasonable cost. This gap, referred to as the 3rd generation AHSS, represents a very active field of research due to the many opportunities involved with these hypothetical grades with strength levels above 1000 Mpa and elongation at or above 20% (HALL; FEKETE, 2017).

2.1.2.1 INTERSTITIAL FREE STEELS

In Interstitial Free (IF) steels, there is no significant quantity of interstitial solute atoms to strain the solid iron lattice, resulting in a steel that contains interstitial free body-centered cubic (BCC) ferrite matrix. This is achieved by removing carbon, nitrogen and other gases during steelmaking by carefully controlling the chemical composition during fusion and through a vacuum degassing process (RAY; JONAS; HOOK, 1994), accompanied by the addition of Titanium (Ti) and/or Niobium (Nb) used for the stabilization of the residual interstitials and the formation of carbide and nitride precipitates.

2.1.2.2 BAKE HARDENING STEELS

Bake hardening (BH) steels are a class of forming steels submitted to an effect called bake hardening, corresponding to the ability of the metal to harden during a paint baking process. Thus, BH steels can provide low yield stress leading to high formability before baking and great strength resulting in significant dent resistance after baking (BERBENNI et al., 2004).

During the bake hardening, the interaction between the solute carbon and the dislocations generated during press forming, results in the strain aging process, where yield strength and tensile strength are increased and elongation decreases due to Cottrell atmosphere formation.

2.1.2.3 HIGH-STRENGTH LOW-ALLOY STEELS

High-Strength Low-Alloy steels are developed similarly to low carbon steels, with the addition of small amounts of alloying elements used to refine the grain microstructure or facilitate precipitation hardening. Common alloying elements are Nb, V and/or Ti (JOHNSON; SANDERS, 2012) but other elements such as Zr, Al and B can also be added. This addition, coupled with controlled rolling, a process developed between 1963 and 1967 (ADAMCZYK,

2006), enabled a finer austenite grain microstructure to be produced, which resulted in finer ferrite grain sizes after transformation during cooling (JOHNSON; SANDERS, 2012), improving mechanical properties such as strength and toughness.

2.1.2.4 *DUAL-PHASE STEELS*

Dual-phase (DP) steels are historically the first family of Advanced-High Strength Steels (AHSS) whose manufacturing involves specific control of phase transformations.

In metallurgy, the term phase refers to a physically homogeneous state of matter, with a specific chemical composition, a distinct type of atomic bonding and arrangement of elements. Each phase has distinct properties and depends on the chemical composition and thermal treatment to which the steel has been exposed. There are three phases involved in any steel: ferrite, carbide (cementite) and austenite (BOYER, 1984).

By explicitly controlling the thermomechanical process and the addition of alloying elements, ferrite-martensite microstructure can be developed in a relatively straightforward manner, providing a wide range of mechanical properties. However, other phases; austenite and carbide and different microstructures; pearlite, bainite and acicular ferrite can be developed depending on the processing route. These and many other parameters determine the mechanical behavior of DP steels, making them a very complex steel with many open questions (TASAN et al., 2015).

2.1.2.5 *TRANSFORMATION INDUCED PLASTICITY STEELS*

Transformation Induced Plasticity (TRIP) steels are AHSS alloys where the mechanical behavior is enhanced through the transformation of retained austenite to martensite during plastic deformation (SOLEIMANI; KALHOR; MIRZADEH, 2020). Although TRIP steels can have many different phases and different microstructures, the retained austenite is the most relevant phase constituent that postpones the necking phenomenon (a mode of tensile deformation), which results in a higher strength-ductility balance (FISCHER et al., 2000). Figure 3 illustrates the inhibition of the necking in a steel sample due to the TRIP effect. This transformation makes it possible to achieve higher ductility without compromising strength, as indicated in Figure 2.

Although the term TRIP is commonly associated with a type of steel from the 1st generation AHSS, it is essentially a metallurgical phenomenon and can be present in different grades of steel, including 2nd and 3rd generation AHSS (SOLEIMANI; KALHOR; MIRZADEH, 2020).

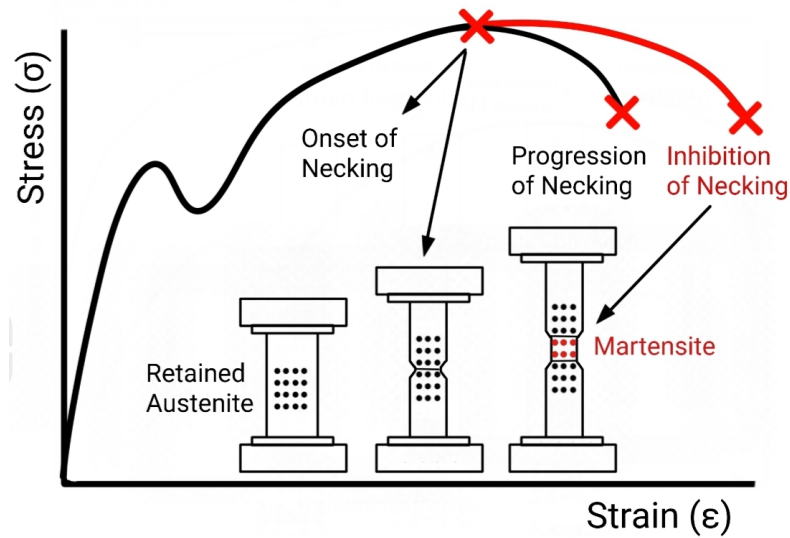


Figure 3 – Schematic illustration of TRIP effect during deformation. Source: Inspired by Soleimani, Kalhor e Mirzadeh (2020)

2.2 DATA-DRIVEN APPROACH AND MACHINE LEARNING

In the contemporary world, an unprecedented amount of data have been continually created. According to Citroen (2011), with new data acquisition and analysis methods, one of the greatest questions across many industries is how to benefit from it.

Regardless of the availability of the massive amounts of data today, data-driven approaches for many different tasks, including modeling, have been used in many fields for several decades. As stated by Küchler (1997), a data-driven approach is when decisions are based on the analysis and interpretation of hard data instead of observations, ensuring that sets of factual information support solutions and plans.

However, during the past decade, ML systems have been increasingly adopted in a broad spectrum of massive and complex data-intensive fields, such as astronomy, biology, climatology, medicine, finance and economy (AL-JARRAH et al., 2015). This is in part because of their ability to discover statistical patterns without expert knowledge. Hence, since data-driven approaches can include ML in a broader sense, these two terms will be used interchangeably in this work.

2.2.1 SUPERVISED LEARNING

Supervised learning is a modeling methodology which consists of providing the desired outputs with the corresponding inputs during what is called a training stage (WANG; LIN; DANG, 2020). Consequently, the data must have clear input-output pairs and properly labeled training data, which will build the set of training examples. If the training stage is adequate, the result is a model adjusted to the patterns contained within the mass of training data. Hence, in the inference stage, the model is capable of predicting the label for an unlabeled data due to how the model's adjusted parameters and computations. This is illustrated in Figure 4.

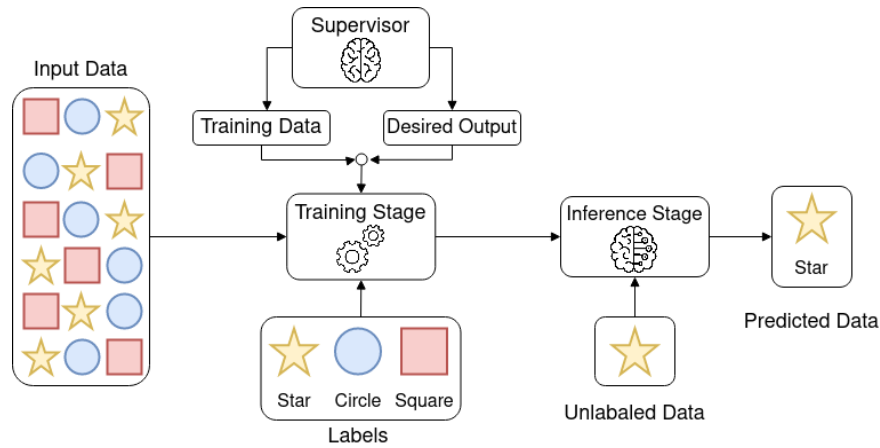


Figure 4 – Illustration of a general supervised learning scheme for training and inference stage.
Source: Own authorship.

Generally, a supervised learning pipeline involves several steps. First, there is data collection, cleaning and pre-processing. Then, a set of adequate features which will be used as input parameters is constructed using statistical methods, feature engineering or ad-hoc knowledge. After the data is ready, a modeling technique is selected and its parameters adjusted. Finally, the results obtained are evaluated to infer the model's performance.

With clean and representative data and a well-built modeling pipeline, state-of-the-art models can be obtained that can, in some cases, outperform humans experts or automate complex work that would be unrealistic for a human to complete (GEBRU et al., 2017).

Although there are numerous modeling techniques, the use of Artificial Neural Networks (ANN) and its many different structures (LEIJNEN; VEEN, 2020) are increasingly prevalent in recent years and will be described in more detail as follows.

2.2.1.1 ENSEMBLE MODELS

Ensemble modeling is an approach in which multiple singular models, namely base estimators, are combined in the inference stage. Since each base estimator is modeled differently, each with its own parameters or features, each model can present bias towards noise or undesirable trends from the dataset. By combining different possible predictions, an ensemble model is able to reduce the overall variance from the final prediction, possibly increasing the accuracy.

According to Moreira et al. (2012), ensemble models are generally concerned with three stages; ensemble generation, pruning and ensemble integration. In the first, the concerns are focused on how to generate different but still efficient base estimators. In the second, a subset of models are selected from a pool of models in order to reduce computational complexity and, if possible, to increase accuracy. Finally, in the latter, the focus is centered on how to combined the different predictions in order to obtain the final ensemble prediction.

For regression problems, a weighting function is commonly used in the ensemble integration. This weighting can be constant, being the same set of coefficients regardless of the

prediction's inputs, or can be non-constant. In the latter case, the coefficients change according to the prediction's inputs, either in a static manner defined in the learning stage or dynamic, defined in the inference stage (MOREIRA et al., 2012).

2.2.2 ARTIFICIAL NEURAL NETWORKS

ANN is a data-driven adaptive system that is inspired by the human brain and its functioning processes, capable of modifying their internal structure according to a cost function (GROSSI; BUSCEMA, 2007). In a supervised learning scenario, the ANN is capable of mapping the input-output relation of a dataset after having seen a number of training examples (ZHOU; WU, 2011). The structure of an ANN is generally a set of processing elements called neurons, connected in different manners by synapses. The weight of these connections is adjusted during a training phase, where a learning algorithm fits the neural network configuration according to a cost function, modeling the linear and nonlinear relationship of a problem or data. Each neuron combines the weighted sum of all its inputs and through an activation function, an output is calculated. The described neuron computational schematic is illustrated in Figure 5 (a).

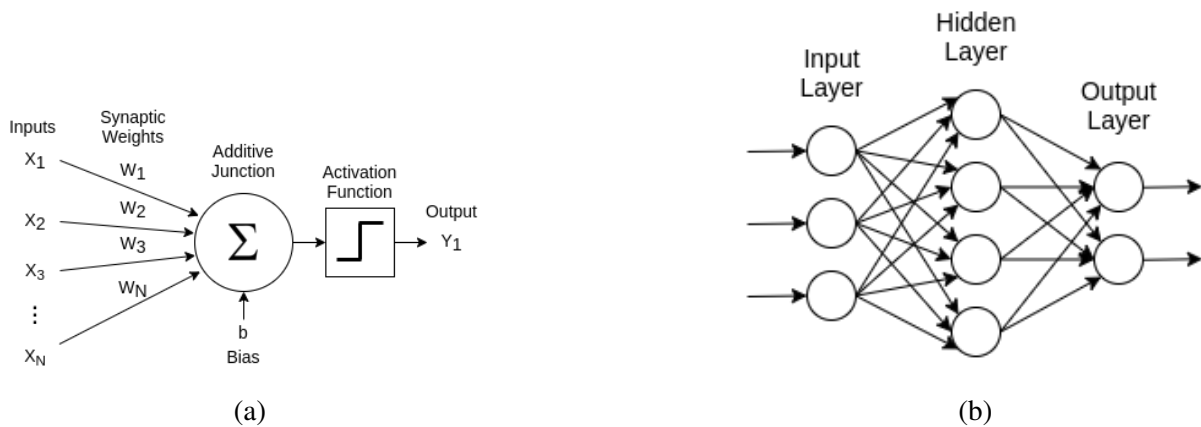


Figure 5 – Representation of (a) neuron and (b) neural network. Source: Inspired by Haykin (1994)

Neurons can be organized in different topological structures with different processing functions, depending on the type, quality and amount of input data and modeling complexity (HUNTER et al., 2012; KALCHBRENNER; GREFENSTETTE; BLUNSOM, 2014). The most common ANN topology is the so-called feed-forward ANN, also called Multilayer Perceptron (MLP). In this work, the term ANN will refer to neural networks with the feed-forward topology unless otherwise stated.

This structure is composed of three types of layers: input, hidden and output layers. These layers are fully connected; each neuron in one layer is connected to every neuron in the following layer. Each connection has an associated weight, which modifies the input or output of each neuron across the ANN. Input and output layers have a number of neurons corresponding to the number of input and output parameters, respectively (GROSSI; BUSCEMA, 2007). Hidden

layers can have an arbitrary number of neurons that perform nonlinear transformations on the network inputs.

Thus, as the information is forwarded across the network during a training stage, the model can generalize the input-output relationship of a collection of data, learning its patterns from examples in a supervised manner.

2.2.2.1 NEURAL NETWORK PARAMETERS

Although the model can generalize patterns without the need for explicit rules or knowledge about a dataset, an ANN's performance depends on numerous parameters, as will be discussed in the following sections.

2.2.2.1.1 NUMBER OF HIDDEN LAYERS AND NEURONS

As each neuron from the hidden layer computes the net input according to the connection weights and activation functions, non-linearity can be modeled by the ANN. By adding more neurons and hidden layers, an ANN can model more complex patterns and higher accuracy can be reached by the cost of computational power.

However, as the number of hidden layers grows, the model complexity also expands, making it prone to overfitting, modeling noise along with the data. If the number of neurons and hidden layers is too low, the model may be under-fitted, incapable of modeling the fundamental patterns from the dataset (KARSOLIYA, 2012).

Although there are many rules of thumb for estimating the number of neurons and hidden layers, there is no consensus of the ideal configuration for optimal modeling since this is dependable on the type, quality and complexity of the dataset samples for which the network is designed (HUNTER et al., 2012).

2.2.2.1.2 ACTIVATION FUNCTION

All non-input neurons in an ANN transform their net input using a typically nonlinear function called an activation function, resulting in a neuron's output. Since each neuron's output is forwarded to the following layers, if each neuron outputs a linear function, the ANN output would be a composition of linear functions, which is itself linear. One of the primary purposes of activation functions is to introduce non-linearity into the network. Although many activation functions exist, differentiable and bounded functions are preferred since they facilitate the model's convergence.

2.2.2.1.3 REGULARIZATION PARAMETERS

To avoid overly complex models and overfitting, regularization techniques can be applied during the training stage to reduce the effective number of free or redundant parameters of the

ANN. This can assist the model in generalizing better to samples on which the model has not been trained.

This can be done by introducing weight penalties, such as L1 and L2 regularization and soft weight sharing (NOWLAN, 2018). Mathematically, L1 and L2 can respectively be expressed according to Equations 2-3,

$$\text{Cost function} = \text{Loss} + \lambda * \sum |w| \quad (2)$$

$$\text{Cost function} = \text{Loss} + \lambda * \sum |w|^2 \quad (3)$$

where *Loss* denotes the modeling error, λ the regularization rate and w the synaptic weights from the ANN. By summing the values of the synaptic weights to the cost function, the training stage will penalize the modeling increase in complexity.

The Dropout technique also prevents overfitting by temporarily removing random neurons and their connections from an ANN during training. This results in a pruned network composed from the units that survived the dropout (SRIVASTAVA et al., 2014).

However, if the regularization rate or dropout probability is too big, this can lead to unstable training and under-fitted models. For L1 and L2, choosing a reasonable regularization rate can be tricky. In the case of dropout, training time is increased due to the noise in the parameter updates generated from dropping out neurons (SRIVASTAVA et al., 2014).

2.2.2.1.4 OPTIMIZER PARAMETERS

In the context of ANN, optimizers are algorithms used to update the network attributes such as weights and learning rate to reduce modeling error. During the training stage, as the training examples are provided to the ANN, optimizers are used to solve optimization problems by minimizing or maximizing an objective function. Typically, this objective function corresponds to the error to be minimized during training and as such, the objective function is often referred to as a loss function.

However, the objective function is often stochastic since it is common for different subsamples of data to be evaluated randomly during training. There can be other sources of noise, such as dropout regularization (SRIVASTAVA et al., 2014). Hence, the algorithm robustness and the rate of change of the parameter updates can significantly affect the model's convergence.

Nevertheless, despite the existence of many optimizers, such as Gradient Descent, Stochastic Gradient Descent (SGD), Mini Batch SGD (MB-SGD), Nesterov Accelerated Gradient (NAG), AdaDelta, Adam, among others, empirical comparisons of these optimizers depend heavily on the hyperparameter tuning protocol (CHOI et al., 2019).

2.2.2.2 NEURAL ARCHITECTURE SEARCH

With the recent explosion in the use of ML and the successes of deep learning applications, the need for efficient pipelines for ML development is becoming increasingly relevant (CHAUHAN et al., 2020). This trend motivates the automation of different stages involved in data-driven modeling, ranging from data acquisition to model evaluation, through the use of Automated Machine Learning (AutoML).

One of such applications is related to the search and optimization of neural architectures used in networks for different purposes. This process, named Neural Architecture Search (NAS), has been widely researched in recent years (ELSKEN; METZEN; HUTTER, 2019b). Since the search space for neural architectures can become huge and training a single network can be computationally expensive, one of the main challenges involved in NAS is the efficient exploration of the search space, in order to optimize a network under a feasible computational load while achieving better modeling performance (JIN; SONG; HU, 2019).

In Zoph e Le (2017), a Recurrent Neural Network (RNN) generated models descriptions of other networks and was trained using reinforcement learning in order to maximize accuracy. Although vast computational resources were needed, the optimized models obtained by Zoph e Le (2017) achieved a competitive performance on different benchmarks, such as CIFAR-10 and Penn Treebank, launching NAS into mainstream research. In Liu et al. (2018), a sequential model-based optimization strategy was used, where the structures of Convolutional Neural Networks (CNNs) are searched in order of increasing complexity while a surrogate model guides the search space. CNNs were also optimized through NAS in Tan et al. (2019) for mobile devices, explicitly incorporating model latency as an objective since the trade-off between model complexity and accuracy is greater on mobile devices. EAs were also used as a search mechanism in NAS. Although gradient-based methods were still used to optimize the neuron's weights, in Elsken, Metzen e Hutter (2019a) an EA approximates the entire Pareto-front of architectures under multiple optimization objectives, including predictive performance and number of parameters.

Computational efficiency is crucial for the mainstream application of NAS and several different methods can be applied in combination to accomplish this. Auto-Keras (JIN; SONG; HU, 2019) is an open-source NAS framework that enables Bayesian optimization to guide the search space while maintaining network morphism, that is, the functionality of a network while changing its neural architecture. A tree-structured acquisition function optimization algorithm, coupled with the development of a neural network kernel, efficiently explores the search space, enabling the use of NAS by mainstream practitioners.

2.3 MULTI-OBJECTIVE OPTIMIZATION

According to Deb e Kalyanmoy (2001), optimization consists of obtaining one or more feasible solutions that correspond to extreme values of one or more objectives. This is done by finding and comparing different available feasible solutions until no better solution can be found.

In a single-objective optimization problem, generally, the task is to find a solution that minimizes or maximizes a single objective function.

When more than one conflicting objective is being optimized, our problem can be characterized as a multi-objective optimization problem (MARLER; ARORA, 2004). In this case, there are multiple criteria involved in the decision-making and multiple objective functions being optimized simultaneously. Mathematically, a multi-objective optimization problem can be formulated according to Equation 4, where the minimization of the objective is intended.

$$\min_{x \in \mathbf{X}} \mathbf{F}(x) = [f_1(x) f_2(x) \cdots f_m(x)]^T \quad (4)$$

where $\mathbf{X} = (x_1, x_2, \dots, x_n) \in \mathcal{X}$ is the feasible set or solution vector of n decision variables and \mathcal{X} is the feasible search/design space. Correspondingly, $\mathbf{F}(x) = [f_1(x) f_2(x) \cdots f_m(x)]^T$ is the feasible set solution, determined by calculating all the objective functions f from the solution $x \in \mathbf{X}$ such that $f : \mathbf{X} \rightarrow \mathbb{R}^m$, with $m \geq 2$. The image of the feasible set $\mathcal{F} = \mathbf{F}(x) \in \mathbb{R}^m : x \in \mathbf{X}$ is called the objective space (AUDET et al., 2020).

For example, consider Figure 6, which illustrates a commonly found trade-off in many engineering and business scenarios. The x-axis denotes *Cost* and can be relative to a business, architectural, design or engineering project, a piece of machinery, among other things. The y-axis is the number of faults that could be present in this project or machine.

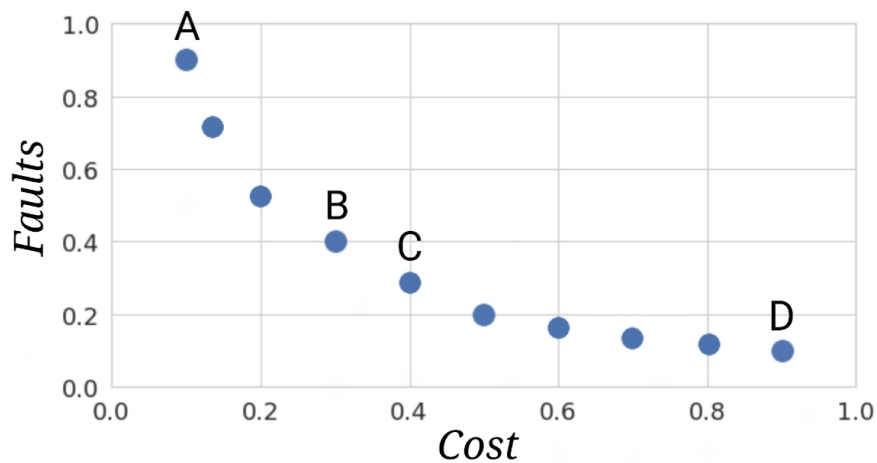


Figure 6 – Example of a Pareto front with two objectives for minimization: Cost and number of faults. Source: Own authorship.

According to Figure 6, point *D* denotes a very costly project with few, if any, number of faults. Point *A* it's the opposite, with a high number of faults but at a very cheap price tag. Points *B* and *C* offer relatively low cost and number of faults, but still higher than *A* and *D*, respectively. Without additional context, neither point can be taken as the best solution. On a tighter budget, point *D* can be impractical. On a critical system where faults cannot be tolerated, point *A* can be compromising. In either case, points *B* and *C* are only viable under a decision maker's scrutiny.

2.3.1 PARETO OPTIMALITY

When all objective functions for minimization cannot be optimized simultaneously, there does not typically exist a single solution that simultaneously optimizes every objective due to conflicting objectives trade-offs. In this case, the solutions can be regarded as Pareto optimal if none of the objective functions can be improved without deteriorating another objective value (DEB; KALYANMOY, 2001). Without an additional context or subjective preference, all Pareto optimal solutions are considered equally good.

Simply, if there exists a candidate solution A with better performance across all optimized objectives than a candidate solution B , it is said that A dominates B . If A is better in some objectives than candidate solution C , but worse in others, then both solutions can be considered optimal solutions in a Pareto set. This is illustrated in Figure 7.

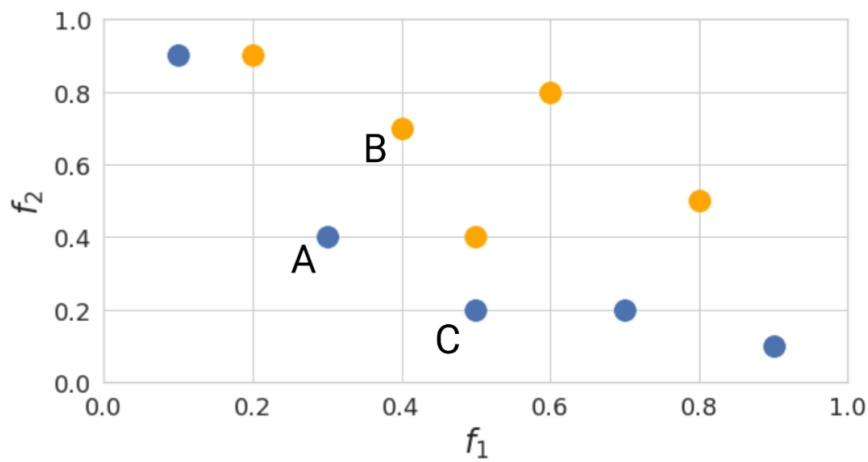


Figure 7 – Pareto front with two objectives, being f_1 and f_2 . Non-dominated points in blue, dominated points in orange. A and B are Pareto optimal solutions that dominate C .
Source: Own authorship.

Since Pareto optimality is intrinsically related to multi-objective optimization, some definitions will be formalized according to Marler e Arora (2004), Audet et al. (2020) in order to structure some concepts present in this study.

Definition 1 (Dominance relations). Given two decision vectors x and x' in \mathbf{X} ,

- $x \preceq x'$ (x weakly dominates x') if and only if $\mathbf{F}(x) \leq \mathbf{F}(x')$
- $x \prec x'$ (x dominates x') if and only if $x \preceq x'$ and at least one component of $\mathbf{F}(x)$ is strictly less than the corresponding one of $\mathbf{F}(x')$
- $x \prec\prec x'$ (x strictly dominates x') if and only if $\mathbf{F}(x) < \mathbf{F}(x')$.
- $x || x'$ (x and x' are incomparable) if neither x weakly dominates x' nor x' weakly dominates x .

Definition 2 (Pareto optimality and Pareto solutions). The vector $x \in \mathbf{X}$ is a Pareto optimal solution if there is no other vector in \mathbf{X} that dominates it. The set of Pareto optimal solutions is called the Pareto set, denoted \mathcal{X}_P and the image of the Pareto set is called the Pareto front, denoted \mathcal{F}_P .

Definition 3 (Pareto set approximation). A set of vectors A in the decision space is called a Pareto set approximation if no element of this set is dominated by any other. The image of such a set in the objective space is called a Pareto front approximation.

Definition 4 (Ideal point). The ideal point $F^I = (f_1^I, f_2^I, \dots, f_M^I)$ is the vector composed with the best objective values over the search space.

Definition 5 (Nadir point). The nadir point $F^N = (f_1^N, f_2^N, \dots, f_M^N)$ is the vector composed with the worst objective values over the Pareto set.

2.3.2 MULTI-OBJECTIVE EVOLUTIONARY ALGORITHMS

Evolutionary Algorithms (EAs) are inspired by the biological model of evolution and natural selection. Through a simplification of natural processes, EAs can simulate an environment and the biological pressures in which potential solutions can evolve (VIKHAR, 2016). By adapting and surviving, solution candidates can be optimized across multiple generations towards an approximate solution shaped according to the parameters and constraints of the problem at hand (BINITHA; SATHYA, 2012).

Since many biological processes involve evolution, several types of EAs were proposed over the years with different inspirations, genetic representations, implementation details and how they are applied to a particular problem. According to Vikhar (2016), the main algorithms are:

- Genetic Algorithm (GA), Developed during the 1960s and 1970s by John Henry Holland and his students (HOLLAND, 1973), is the most popular EA used in optimization problems. GAs encode the solution of a problem in order to recombine and mutate different solutions each generation. The standard encoding is in the form of strings of Boolean or Real-coded numbers, but the encoding may change according to each problem domain (RONALD, 1997). The population can be either preserved during different generations or eliminated in each generation, passing along its genes through a recombination process.
- Genetic Programming (GP) represents programs or instruction sets as possible solutions to a problem. These computer programs, traditionally represented as tree structures, are evolved by applying operations analogous to natural genetic processes to the population of programs (KOZA, 2003). Every tree node has an operator function and every terminal node has an operand, which in association, can express mathematical functions, sets of computer instructions, ensembles of data processing functions, among other possibilities, according to the problem domain. Proposed in 1992 by John Koza, a Ph.D. student of

Holland, GP can be considered an extension of GA, where the solutions could be computer programs with variable-length representation instead of the fixed-length encoding used in GA. Thus, population diversity is obtained in the individual's gene values but also in its structure (BINITHA; SATHYA, 2012).

- Evolution Strategy (ES), like GAs, is defined by the application of selection, recombination and mutation in the possible solutions each generation. However, unlike GAs that is inspired in the micro-level (gene, chromosomes, alleles), ES is modeled after the macro-level (phenotype, hereditary, variation) (BINITHA; SATHYA, 2012). In ES, solutions are encoded as vectors of real numbers and its most prominent design principle is the use of a self-adaptive control of parameters. This technique was first proposed by Rechenberg in the 1960s and further developed by Schwefel (BäCK; HOFFMEISTER; SCHWEFEL, 1994).
- Evolutionary Programming (EP) (FOGEL, 1998) is very similar to ES; however, it has no restriction regarding the use of data types in its encoding. In EP, similarly to GP, the program's structure to be optimized is fixed, but its numerical parameters are allowed to evolve. Another main difference is that, in EP, the main genetic operator is mutation, while recombination is not traditionally present.

Many real-world problems involve concurrent optimization of multiple and sometimes conflicting objectives. When the problem's dimensionality increases drastically, or when the problem is too hard to formalize in prescriptive terms, the search for an exact solution can become computationally too demanding or outright infeasible. In such scenarios where a near-optimal solution is sufficient, Multi-objective EAs (MOEAs) can be effective at providing reasonable approximate solutions (ZHOU et al., 2011).

Although many different EAs exist, in the current work, a GA will be used due to its efficiency in problems where the search space is large and poorly known with no analytical equation provided to solve the problem (BINITHA; SATHYA, 2012). Thus, a more detailed description of GA's basic procedure and operators are provided, along with a NSGA-II explanation of its abilities to handle multi-objective problems.

2.3.2.1 GENETIC ALGORITHMS

GAs are stochastic population-based optimization algorithms that reflect the process of natural selection where the fittest individuals are selected for reproduction and other genetic operations in order to produce offspring of the next generation (BINITHA; SATHYA, 2012).

Several elements of this biological inspiration need to be translated into computational terms. A population is a set of individuals, where each individual is a candidate solution to the optimization problem. Each individual is characterized by their chromosome (or genotype),

corresponding to the set of parameters that defines a possible solution (DEB; KALYANMOY, 2001). An individual parameter is considered to be a gene in the chromosome.

Two requirements must be satisfied for the proper application of GAs in a problem domain. The first requirement is a suitable encoding of the candidate solutions, such that the design of the chromosome represents the solution parameters of the problem to be solved (CHAHAR; KATOCH; CHAUHAN, 2021). A good representation can facilitate optimization by limiting the search space.

The second requirement is the definition of a fitness function used to evaluate the solution candidates to guide the optimization across the search space towards optimal results. The fitness function can be represented in many ways, as long as it characterizes the performance of each solution closely correlated with the optimization goal. Thus, it can be represented by analytical equations, data-driven models and even simulations, although it is preferable to be computationally efficient (STORK; EIBEN; BARTZ-BEIELSTEIN, 2020), since GAs must be iterated many times to generate an adequate solution for non-trivial problems. The evaluation can also be outsourced to external agents, such as human specialists, through interactive GAs.

The basic outline of a GA follows an iterative pattern as follows:

1. Initialization: Creates an initial set (population) of candidate solutions (individuals). The population can be generated randomly across the search space or seeded in predetermined areas, where the likelihood of optimal solutions is higher or that satisfies more problem constraints, resulting in a faster convergence (TODOROVSKI; RAJICIC, 2006).
2. Selection: During each generation, some individuals from the population are selected for breeding. The selection process is usually carried through a fitness-based procedure, where solutions with higher fitness are more likely to be selected. Since premature convergence usually occurs due to lack of diversity in the population, the selection process plays an important role in maintaining good individuals without losing diversity (SHUKLA; PANDEY; MEHROTRA, 2015). Selection can also impact the GA ability to deal with Pareto optimality (BAO et al., 2017).
3. Crossover: After selection, the population is recombined through a crossover operator, which pairs the genes of different individuals, creating a new generation of solutions sharing characteristics from their parents. The best individuals selected from the prior generation typically ensure that the new generation's average fitness is higher. In contrast, the inclusion of less fitted individuals helps to maintain genetic diversity. Over the years, many application-dependent and independent crossover operators were proposed since their effects vary across different problem domains and encoding, each with its advantages and disadvantages (KORA; YADLAPALLI, 2017).
4. Mutation: With the offspring obtained using the crossover operator, a mutation operator is applied randomly to each chromosome, possibly altering the population with new genetic

structures by randomly modifying some of its building blocks (OTMAN; ABOUCH-ABAKA; TAJANI, 2012). This ensures diversity of solutions through a better exploration of the search space. Usually, the probability of mutation is set low to avoid a random search.

5. Evaluation and Iteration: Each candidate solution from the new generation is evaluated according to a fitness function. If a termination condition is met (number of generations, quality of solutions, or elapsed time), the GA is stopped. Otherwise, steps 2-5 are repeated.

Building upon this basic outline, it is possible to elaborate specialized GAs with new operators that offer better performance concerning population diversity, time complexity, or higher dimensionality.

2.3.2.2 *NONDOMINATED SORTING GENETIC ALGORITHM II (NSGA-II)*

In practical problems, the NSGA-II is arguably one of the most popular MOEAs used for optimization tasks and has been successfully applied in a vast range of problems across many industries (ZHOU et al., 2011). Its application is also widespread in the scientific community and is frequently used as a benchmark for novel algorithms.

The main features of NSGA-II are the assignment of fitness to candidate solutions based on non-dominated sorting and diversity preservation among solutions of the same non-dominated front (DEB et al., 2002).

The algorithm uses elitist selection, where the best chromosomes from the current generation are made available to be carried to the next generation. By doing so, the quality of the best solution found in each generation monotonically increases over time. If these solutions were discarded, it would be possible to lose the best solution due to stochasticity (BALUJA; CARUANA, 1995).

Non-dominated sorting classifies each solution according to their Pareto dominance. In this sorting, all the non-dominated solutions from population P are selected and assigned to F1 (the rank 1 front). Then, the same steps are applied to the remaining solutions and assigned to F2 (rank 2 front). This procedure is repeated until all solutions have been assigned to a front, as illustrated in Figure 8. Non-dominated sorting plays an essential role in determining the relative quality of the solutions in the selection operation (BAO et al., 2017).

To maintain diversity, a Crowding Distance (CD) operator is used as a criterion to select the non-dominated solutions, guiding the process toward a uniformly spread-out Pareto front. The CD measure of a solution can be interpreted as an estimation of the density of solutions surrounding a particular solution from the population (DEB et al., 2002). A higher CD is preferable, indicating that the solution is, in some sense, less crowded by other solutions. One of the main advantages of this approach is the lack of user-defined parameters.

Hence, with a fast non-dominated sorting procedure, an elitist strategy and a parameterless approach for diversity preservation (DEB et al., 2002), NSGA-II found increased attention and

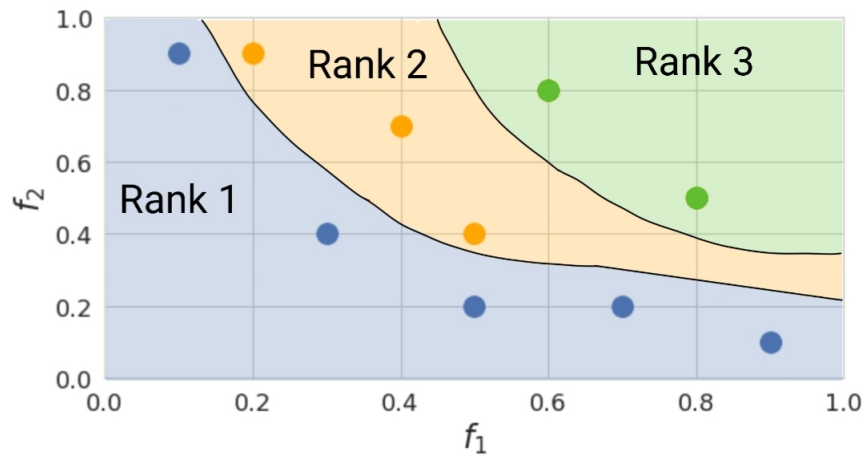


Figure 8 – Example of a Pareto ranking from a non-dominated sorting. Source: Own authorship.

was effectively applied in many different scenarios (ZHOU et al., 2011). Some of these scenarios involve problems with no analytical function to be evaluated or where the fitness evaluation is too computationally expensive. Such cases are common in real-world engineering problems and are described in more detail as follows.

2.3.3 SURROGATE BASED EA OPTIMIZATION

Many engineering problems require simulations or experiments to evaluate the performance and constraints as a function of independent variables. This includes computer simulations (HAN; ZHANG, 2012), where a single simulation can take minutes, hours, or even days to simulate. Since many metaheuristics require thousands or millions of evaluations, incorporating such complex models would take a prohibitive amount of time and resources.

Nonetheless, numerous studies propose an optimization for many of such real-world engineering problems. For instance, in Tao et al. (2017) a Kriging surrogate model coupled with a modified Particle Swarm Optimization optimized the design of automobile fenders made of 3D woven composite, intending to minimize its weight, achieving a 20.65% weight reduction. In Sreekanth e Datta (2010), modular neural networks and genetic programming in tandem with MOEA, was employed to derive the optimal pumping strategies for coastal aquifer management and the results from the model's performance were compared. Also, in Mengistu e Ghaly (2008), an ANN model coupled with MOEA. High fidelity Computational Fluids Dynamics (CFD) simulations were used to generate a dataset used in training the ANN, with the objective of enhancing the aerodynamic performance of a transonic turbine stator and subsonic compressor rotor. The use of the ANN instead of CFD simulations reduced the computing time by ten folds.

What all these studies have in common is the use of a surrogate model, which approximates and substitutes the original expensive objective function, physical simulation, real-world or computational process during the optimization. A surrogate model can be considered as an educated guess about how an original function behaves, such as an engineering-related function, based on a few points of data where it was affordable to measure the function values (FOR-

RESTER; SOBESTER; KEANE, 2008). This data could be from a computer simulation or directly from physical experimentation or production.

According to S ndergaard (2003), surrogate models are either simplified physical or numerical models based on knowledge about the physical system, or data-driven models based on knowledge acquired from empirical evaluations and sparse sampling of the problem's parameter space. Although the original function's bigger picture cannot be told from the sampled data, under the assumptions that the function and its derivatives are continuous, a surrogate model can become a very effective low-cost replacement of an original model (FORRESTER; SOBESTER; KEANE, 2008).

In the study presented in Stork, Eiben e Bartz-Beielstein (2020), the process of surrogate-based optimization is defined by three possible layers: L_1 - real-world process; L_2 - computational model and L_3 - surrogate model, constructed either for the real-world application (L_1) or computational model (L_2). The arrows represent different possible data streams. Some arrows are in the background, connecting different layers. Solid arrows represent the implementations used in the current work, while dotted arrows are not implemented.

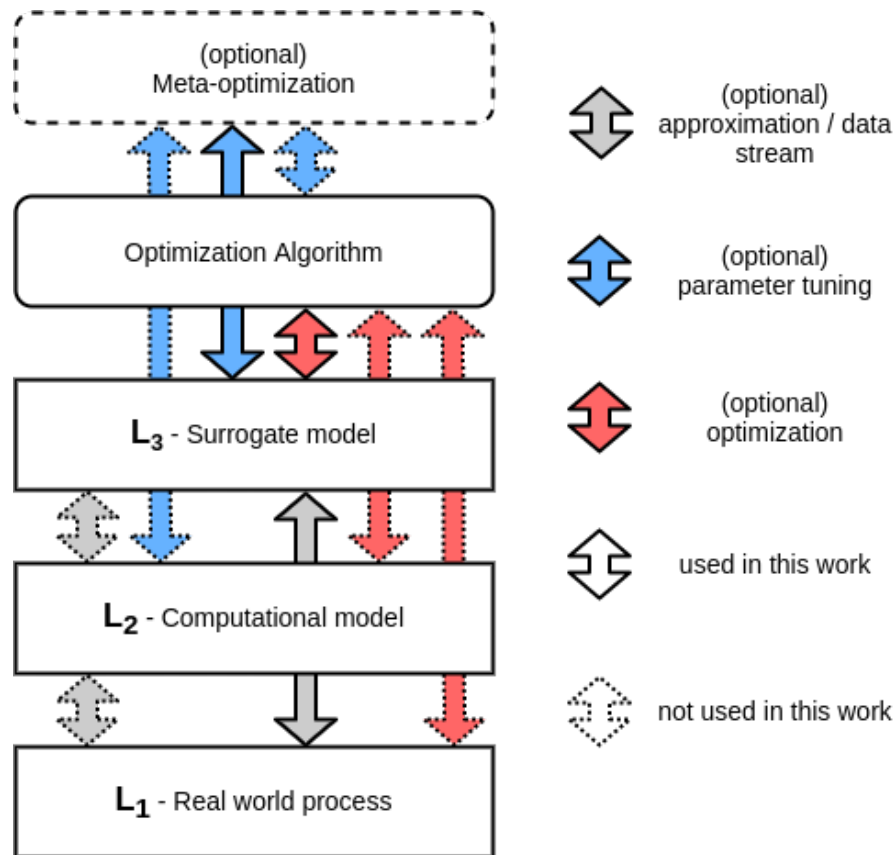


Figure 9 – The different possible layers of a surrogate-based optimization process: real-world process, computational model and surrogate. Source: Own Inspired by Stork, Eiben e Bartz-Beielstein (2020)

In reference to Figure 9, each layer can be the target of optimization or used to sample information to guide the optimization process. Besides the real, computational and surrogate

objective function layers, defined as L_1 , L_2 and L_3 respectively, the surrogate-based optimization cycle also includes the optimization process itself and an optional meta-optimization, controlling the parameters of different stages of the cycle. Although the direct optimization of L_1 and L_2 could be possible, only the optimization of L_3 constitutes as a surrogate-based optimization, where L_1 or L_2 is used only for verification of promising solution candidates (STORK; EIBEN; BARTZ-BEIELSTEIN, 2020).

In the current work, there is no available L_2 layer for solving or simulating the entire metallurgical process including different steel grades. Thus, the real-world process (L_1) is approximated by a surrogate (L_2), optimized through NAS (meta-optimization) and then coupled in the optimization process itself (NSGA-II). This implementation is defined by the solid arrows in Figure 9.

Next, some optimization metrics are discussed in order to provide some quantitative method of comparing and analyzing an optimization's performance.

2.3.4 OPTIMIZATION METRICS

Generally, in a multi-objective optimization context, it is not possible to find or enumerate all elements of the Pareto front (AUDET et al., 2020). Thus, the optimization results consist of a discrete approximation of a Pareto front. In order to solve the many-objective optimization problem, it would be necessary to compare the obtained approximation with the best discrete representation of the Pareto front.

However, according to (AUDET et al., 2020) evaluating the quality of a Pareto front approximation is not a trivial task since it involves several factors such as the closeness to the Pareto front and the coverage in the objective space. Due to these complexities, many metrics, or performance indicators, have been introduced to measure the quality of a Pareto front approximation by assigning a score according to different measurement objectives (LI; YAO, 2019).

2.3.4.1 HYPERVOLUME INDICATOR

The hypervolume indicator is defined as the volume of the space in the objective space dominated by the Pareto front approximation S and delimited from above by a reference point $r \in \mathbb{R}^m$ such that for all $z \in S$, $z \prec r$. Mathematically, the hypervolume is given by Equation 5. The hypervolume indicator is given by

$$HV(S, r) = \lambda_m\left(\bigcup_{z \in S} [z; r]\right) \quad (5)$$

where λ_m is the m -dimensional Lebesgue measure.

The hypervolume indicator is widely used to quantify the quality of the approximate Pareto front, due to its practicality and its properties, particularly the strict monotonicity, i.e., if a

Pareto front approximation A dominates another approximation B , then $HV(A, r) > HV(B, r)$. The hypervolume indicator is still the only known metric that exhibits this property.

An hypervolume indicator illustration is shown for the bi-objective example where $m = 2$ in Figure 10.

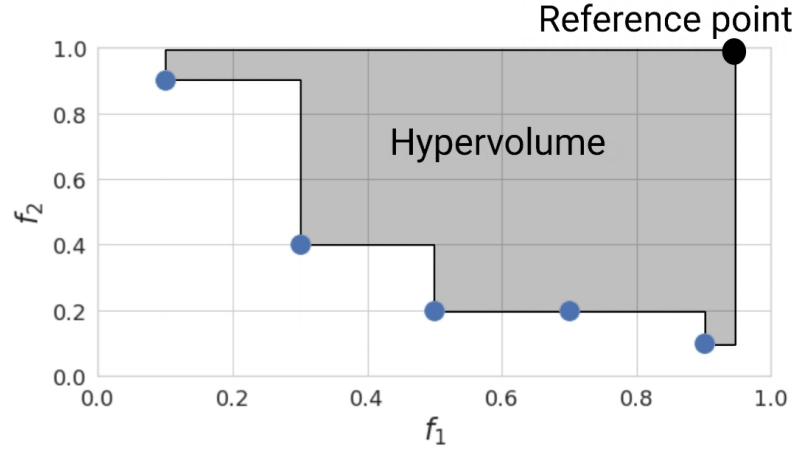


Figure 10 – Example of a graphical HV calculation. Source: Inspired by Fonseca, Paquete e Lopez-Ibanez (2006).

The HV computation used in this work is based on variant three of the algorithm proposed by Fonseca, Paquete e Lopez-Ibanez (2006).

2.3.4.2 INVERTED GENERATIONAL DISTANCE

The Inverted Generational Distance (IGD) measures the distance of each point from the Pareto front in relation to p points from the approximation set, as defined in Equation 6 and illustrated in Figure 11.

$$IGD(S, F) = \frac{1}{|F|} \left(\sum_{i=1}^{|F|} d_i^p \right)^{\frac{1}{p}} \quad (6)$$

where d is the distance between a point i from the Pareto front and the closest point pertaining to the Pareto approximation. Generally $p = 2$, as shown in Figure 11.

The IGD metric has a straightforward computation, but it is very sensitive to the number of points found by a given approximation front. For example, if the algorithm identifies a single point in the Pareto front, the indicator's value will equal zero. This metric has a bias where it favors Pareto approximations with few non-dominated points close to the Pareto front against more distributed Pareto approximations (AUDET et al., 2020). Thus, it is used as a convergence measure and it's not representative of the solutions distribution and spread.

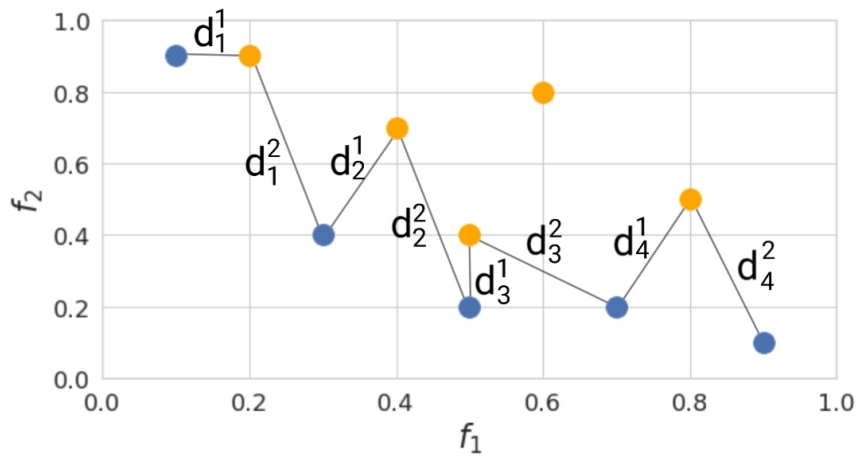


Figure 11 – Example of a graphical IGD calculation. Source: Own authorship.

2.3.4.3 DIVERSITY COMPARISON INDICATOR

The Diversity Comparison Indicator (DCI) (LI; YANG; LIU, 2014) is a k -nary spread indicator; that is, it uses k Pareto approximations as arguments to calculate the extent of the approximation.

To calculate this metric, the approximations are delimited by lower and upper bounds and divided into a number of hyperboxes. Then, a contribution coefficient is calculated for each Pareto front approximation in relation to the hyperboxes where non-dominated points are found. For each approximation, DCI indicates the average of contribution coefficients relative to all hyperboxes of interest.

However, a drawback of the DCI metric is the need to manually choose the number of hyperboxes, which can wrongly favor some Pareto front approximations over others.

DCI cannot be used when only one Pareto front approximation is known and is mainly used to compare the spread of different Pareto front approximations.

2.3.4.4 GENOTYPIC DIVERSITY MEASURE

Genotypic Diversity Measurement (GDM) is a helpful metric for monitoring and managing the exploration of an optimization process and can also be used to avoid premature convergence towards a suboptimal solution due to lack of diversity within the population.

GDMs can be classified into two categories. In the distance-based (D) category, the distance between different individuals is measured. There are many ways to evaluate distance, ranging from the average spatial position of the population like the distance-to-average-point measure (D_{DTAP}), the moment of inertia measure (D_{MI}), or from the position of the fittest individual (D_{ED}). The position of each individual can also be used in the calculation, ranging from the pairwise distance (D_{PW}) to the maximum distance (D_{RP}) between individuals. With gene frequency category (GF), each gene representation is counted in order to measure how representative it is in the overall genetic pool. Since the encoding used in the present work

involves real-coded numbers, all genes are continuous and with a possibly infinite number of values, which can hamper the frequency scanning. Consequently, the genes histogram should be partitioned and each interval considered as a possible allele. The interval number involved in this discretization can severely limit the GDM efficacy, especially for small populations, high dimensionality problems, or data distributions with high skewness or multiple peaks (CORRIVEAU et al., 2015).

Despite the numerous formulations proposed in the literature and for different encodings, every GDM formulation has its limitations, such that there is no measurement capable of reflecting the true diversity of a population and its peculiarities. These include the ability to capture single or multi-site convergence, stability, insensitivity with respect to: dimensions, population size, interval and finally, robustness to outliers.

All things considered, the development of GDM can aid in online feedback mechanisms used in adaptive methods to circumvent premature convergence, compare the performance of multi-site convergence and adjust the exploration-exploitation balance driving the optimization process.

In this work, the normalized pairwise diversity measurement (D_{PW}^N), described by Equation 7, will be adopted due to its excellent capabilities regarding single-site convergence, stability, insensitivity with respect to dimension and outliers (CORRIVEAU et al., 2012), being one of the best GDM for standard EAs according to Corriveau et al. (2015).

$$D_{PW}^N = \frac{2}{POP_{SIZE}(POP_{SIZE}-1)} \sum_{i=2}^{POP_{SIZE}} \sum_{j=1}^{i-1} \sqrt{\sum_{k=1}^{POP_{SIZE}} (x_{i,k} - x_{j,k})^2} \quad (7)$$

where POP_{SIZE} is the population size, i and j are individuals, k is the gene locus and $NMDF$ is a normalization factor.

3 SYSTEMATIC LITERATURE REVIEW

According to Okoli e Schabram (2010), a literature review can be conducted with different standards of rigor, ranging from a collection of similar works without any explicitly defined criteria to a scientifically rigorous synthesis of a body of primary research using a predefined method.

Although all literature review should aim to be open-minded and transparent about its focus and objectives, for a literature review to be considered systematic, there must also be an explicit and reproducible approach to identify, evaluate and synthesize the body of work incorporating elements of analytical criticism (OKOLI; SCHABRAM, 2010)

The use of data-driven models for the prediction and design of steel's properties is not new and has recently been the topic of numerous works. However, there is no comprehensive systematic literature review (SLR) about the topic. This may be due to numerous possible reasons, such as the lack of a perceived necessity to systematically organize the body of work or the difficulty that arises from the multidisciplinary nature of these works and their vast application range.

Therefore, a systematic literature review was conducted during July and August of 2020 to find answers related to the prediction and design of steel's mechanical properties using data-driven techniques.

3.1 RESEARCH METHOD

With the main goal to provide an overview of the topic and identify the characteristics and trends within it, the following research objective was defined:

- Investigate the data-driven techniques used to model the mechanical properties of steels.

with the following research questions (RQs) to assist in the understanding and synthesis of the research objective:

1. What are the most frequently used techniques?
2. How is data-driven modeling relevant compared to analytical models derived from physical metallurgy?
3. What are the peculiarities of modeling different steel categories?

3.2 PLANNING THE REVIEW

With the proposed research objective and questions, the planning stage of the SLR was defined first by creating a search phrase selecting the Academic Search Engines (ASEs) where the search would take place

Through different iterations, it was decided to use nine different keywords such that only eight logical operators would be needed to compose the search phrase, making it compatible with all the selected ASEs. The search phrase was inserted in the Abstract and Title search fields when such options were available. Finally, the search phase was defined as:

- steel AND (predict* OR "neural net") AND (optimization OR design OR property) AND (elongation OR "tensile strength" OR yield)

Due to the multidisciplinary nature of the topic, relevant ASEs in both engineering and computer science were selected. The chosen ASEs are Association for Computing Machinery Digital Library (ACM DL), IEEExplore, Science Direct, Scopus and Web of Science.

3.3 CONDUCTING THE REVIEW

After defining the ASEs and the search phrase, the next step in conducting an SLR is to outline the objective, inclusion and exclusion criteria used as a practical screen to decide which studies should be considered for the review. This procedure is carried out so early in the SLR due to the vast quantities of studies usually retrieved from the search engines that are irrelevant to the research objective and questions of the review (OKOLI; SCHABRAM, 2010). These criteria are presented as follows:

- Objective Criteria
 - Year range: 2010 - 2020
 - Document type: Articles, dissertation and thesis
 - Language: English
 - Availability: Open Access
 - Completeness: Full or short
- Inclusion Criteria
 - Include studies if it uses data-driven methods to model the mechanical properties of steels
 - Include studies if it applies metaheuristics to optimize steel parameters towards new or desired mechanical properties
- Exclusion Criteria
 - Remove studies that predicts or optimizes only non-mechanical properties of steels
 - Remove studies that predicts properties or optimizes nonmetallic materials

- Remove studies that predicts properties or optimizes steels under practical engineering context, such as welding, load bearing structures, etc.

Given the inclusion parameters, it was established that if a work fits into any excluding criteria or does not fit in any objective criteria, the study should be disregarded from the further steps of the SLR. The criteria were applied in the following order: objective, exclusion and inclusion.

Once all the articles were selected through the practical screen and quality appraisal, the number of articles chosen from each ASEs is illustrated in Table 1. The column Found refers to the number of articles returned by using the search phrase alone. Filtered is the number of articles after the objective, inclusion and exclusion criteria were applied. Finally, Selected is the number of articles that were selected for further study.

Table 1 – Table regarding the number of found, filtered and selected studies from the SLR.

ASE	Found	Filtered	Selected
ACM DL	23	18	2
IEEEExplore	44	1	1
Science Direct	934	59	4
Scopus	3004	184	10
Web of Science	3547	372	4

The information collected from the body of work to be tabulated is:

- Input variables from the model, classified into categories IN1, IN2, IN3 and IN4 defined as follows:
 - IN1: Microstructure
 - IN2: Chemical composition
 - IN3: Processing parameters
 - IN4: Mechanical properties (as independent variables)
- Output variables from the model, that is, the properties being predicted or optimized.
- The algorithm used in modeling the data.
- Presence of a quantitative analysis of the results.
- Presence of a qualitative analysis of the results.

3.4 SUMMARY AND DISCUSSIONS

From the selected studies, a matrix concept was generated and is illustrated as Table 2. The table contains information such as the modeling or optimization technique employed, input and output variables and quantitative or qualitative analysis from the results achieved.

Table 2 – Concept Matrix from the SLR. *Reference* are the articles selected for study. *IN1*, *IN2*, *IN3*, *IN4* is the presence of microstructure, chemical composition, processing parameters and mechanical properties respectively as independent variables and *Output* the variable being predicted or optimized. *Algorithm* refers to the set of algorithms and techniques used in the research, while *Quant. Analysis* and *Quali. Analysis* is the presence of a quantitative and qualitative analysis, respectively.

Reference	IN1	IN2	IN3	IN4	Output	Algorithm	Quant. Analysis.	Quali. Analysis.
(DENG et al., 2020)		X			TS, YS, H	ANN	X	X
(FAIZABADI et al., 2014)		X		X	H	ANN		
(GHAISARI; JANNESARI; VATANI, 2012)		X	X		YS, TS, EL	ANN	X	X
(ERES-CASTELLANOS et al., 2020)		X	X		YS	Analytical Eq.	X	X
(BRIGHT et al., 2011)		X	X		YS, TS, EL	CART	X	X
(SARAVANAKUMAR et al., 2012)		X	X		YS, TS, EL	ANN		
(PANDYA; SHAH, 2014)			X	X	EL	ANN		
(DI SCHINO; RICHETTA, 2017)	X	X			YS, TS	Regression		X
(WU et al., 2018)		X	X		YS, TS, EL	ANN e EA	X	X
(CHEN et al., 2020)	X	X	X		YS, TS, EL, H	Analytical Eq. e FEM		
(XIONG; ZHANG; SHI, 2020)		X	X		FS, TS, H	RF, LLS, KNN, ANN		
(BHATT; PARAPPAGAUDAR, 2015)		X			YS, TS, EL	DOE e ANOVA		X
(WANG et al., 2020)		X	X		YS, TS	ANN	X	X
(GARCÍA-MARTINO et al., 2020)		X	X		YS, TS	Regression, ANN e MARS		
(KANCA; ÇAVDAR; ERSEN, 2016)		X	X	X	YS, TS, EL	GEP	X	X
(JIA et al., 2011)		X	X		YS, TS, EL	ANN, PSO		
(GALINDO-NAVA; RAINFORTH; CASTILLO, 2016)	X	X	X		YS	Analytical Eq.	X	X
(WANG et al., 2020)		X	X		YS IT	RF, MOEA	X	X
(CHOU; TSAI; CHOU, 2016)		X			YS, TS	ANN, PSO		
(WU et al., 2020)		X	X		YS, TS, EL	SAFER-BODNN	X	
(COSTA et al., 2019)	X		X		YS, TS, EL	Regression, MOEA	X	X
Present work		X	X		YS, TS, EL	ANN, MOEA	X	

3.4.1 DIVERSITY OF MODELED STEEL CATEGORIES

One of the main difficulties in compiling and comparing the studies is the many types of steel and metallic alloys used across different industries. Each of these steels can have its own characteristic behaviors and specific industrial production stages, which can vary according to each manufacturing line. Furthermore, different properties can be measured in different ways according to different technical standards and testing equipment.

This, in turn, creates a more complex task in discerning the limitations of each study, as well as understanding different models in the scope of a centralized objective since each study can present different sets of input and output variables.

3.4.2 INDEPENDENT VARIABLES

According to Table 2, it is prevalent for chemical composition and processing parameters to be present as independent variables in modeling, which is expected since they are the main factors in steel production. Even when the chemical composition is not treated as variables, they are still considered in the study in the form of constants, as in Costa et al. (2019), to evaluate in isolation other variables concerning the mechanical properties, such as processing parameters and microstructure.

Despite chemical composition and processing parameters being frequently included, these sets of variables are not homogeneous across different studies, given the peculiarities of each manufacturing line or study objective. Thus, some alloying elements may be present in some of the selected works while not present in others.

When both the dependent and independent variables include mechanical properties, the prediction model usually is used to forecast the effects of a specific processing stage of steel production, as in the work of Kanca, Çavdar e Ersen (2016) where the goal was to predict the mechanical properties after the cold rolling process, taking into account the value of these mechanical properties before the process.

Still, another challenge is the use and influence of microstructure as an independent variable, due to the need for sophisticated laboratory equipment such as high-precision microscopes for accurate measurement. A possible alternative is to infer an initial microstructure through the use of analytical equations or surrogate models, such as computer simulations using physical metallurgical principles, as was done in Chen et al. (2020). Although less time consuming and possibly more cheap, these methods are prone to errors and there is no guarantee that it considers every alloying element or processing parameter in its calculations or simulations.

The use of microstructure not only can make the numerical prediction more accurate but can also avoid the design of infeasible steels. This was demonstrated by Chou, Tsai e Chou (2016), where two models were trained: one with microstructure information and one without. The microstructureless model guided a metaheuristic optimization towards steels that couldn't exist in a realistic metallurgical scenario. Thus, while not foolproof, the use of microstructure as

a dependent variable can significantly enhance the robustness of steel modeling.

As the complexity of steel production and understanding increases in order to obtain novel steels, including 3rd generation AHSS, the microstructure plays a bigger role in the comprehension of available possibilities as well as technical limitations. In this context, a state-of-the-art model considering the steel microstructure can aid in new development and answer open questions, possibly being regarded as a future trend.

3.4.3 MODELING ALGORITHMS

In the literature study, there is wide use of ANN, with different training algorithms including SGD, Quasi-Newton method, Levenberg-Marquardt algorithm, Bayesian Framework, among others. What all these ANNs had in common was their relative simplicity, containing one or two hidden layers and few dozens of neurons.

From the selected articles, only one study proposed a more complex architecture containing up to four hidden layers and a hundred neurons per layer. This was done in Wu et al. (2018) in two steps. First, data augmentation was accomplished with semi-supervised learning, imputing missing values from the dataset. Thus, the size of the dataset was increased to 49505 entries, the most numerous from the literature review. Then, through hyperparameter optimization, a complex ANN architecture was found that could provide high performance.

Besides ANNs, many other computational algorithms are used for modeling data, either by black-box models or generating analytical equations through regression and evolutionary algorithms. In Kanca, Çavdar e Ersen (2016), Gene Expression Programming was used to generate analytical models, relating the independent to the independent variables. This technique optimizes and evolves tree structures that learn and adapt by changing shapes, sizes and other attributes. Then, each tree can be decoded into an analytical equation that aims to model a predictor.

Although there are studies that implement classical regression techniques, most studies implement ANN as uni or multi-variate regression model. For example, in Jia et al. (2011) multiple ANNs were modeled, each model intended for a specific mechanical property. In Saravanakumar et al. (2012), different mechanical properties (YS, TS and EL) were modeled by a single network. This shows the ANN's ability to be used as a universal function approximator. However, from the SLR, there was no study comparing multiple ANNs with a single prediction output against a single ANN with multiple predictions.

Although ANNs with simple architectures and other simple data-driven algorithms can adequately model the data, given the recent progress in machine learning, one can wonder what can be attained by using more recent and sophisticated modeling techniques and if this complexity is necessary at all. In either case, the validation of the model is as important as the modeling itself.

3.4.4 OPTIMIZATION ALGORITHMS

Optimization algorithms can be used in different contexts. They can be applied in hyperparameter optimization, as was done in Wu et al. (2018) to enhance the ANN performance. But they can also be used in the optimization of hypothetical steel designs, by proposing new possible chemical solutions and/or processing parameters.

For instance, in Costa et al. (2019), an experimental central composite design and a multi-variate regression model was applied to generate a surrogate model that, coupled with NSGA-II, was used to optimize the continuous galvanizing heat treatment. This optimization was carried in order to generate a DP steel with a TS greater than 1100 Mpa, YS between 550 - 750 Mpa and a minimum elongation of 10%. The objective was to minimize YS and maximize both TS and EL. The chemical composition was kept constant throughout the experiment and only the galvanizing heat treatments were able to be optimized.

In Shen et al. (2019) Support Vector Machines were used as surrogate models and used in tandem with NSGA-II to improve the chemical composition and ageing conditions of stainless steel strengthened by the R-phase precipitates. The sole objective was to maximize the steel's hardness (HRC). Physical metallurgical principles, Support Vector Classifiers and experimental trials were used to validate the system. The steels used are from the category of martensitic stainless steels and were manufactured by complete homogenisation-quenching-tempering routes.

3.4.5 QUANTITATIVE ANALYSIS

In many studies, the only value representative of the model's performance is a regression metric, mainly Mean Relative Error (MRE), Coefficient of Determination (R^2), Mean Absolute Error (MAE), among others. Although useful in a practical sense, these metrics do little in terms of explaining or measuring the models performance in different contexts.

Some works present quantitative analysis that better explore the relationship between the input and output variables in more detail, making it possible to analyze possible inconsistencies or gain insights into the models behavior and explainability, which is particularly useful for black box models.

To better understand the global impact of the input variable on the models output, it is possible to calculate and visualize the weight of the influence of each parameter in relation to the output variable, as was done in Wang et al. (2020), Wu et al. (2018) for ANNs models. There are many different techniques which can be used to calculate this weight, including an adaptation of Shapley values (LUNDBERG; LEE, 2017), a solution used in cooperative game theory, as demonstrated in Sgrott et al. (2021). For analytical regression models, this comes much more naturally as the weights associated with each variable correspond directly to the coefficient values in the mathematical equation, providing much more interpretable results.

If a local analysis is desired, that is, an analysis that only considers a single or a few samples from the dataset, a sensitivity analysis can be realized, as was done in Costa et al. (2019),

Ghaisari, Jannesari e Vatani (2012). By doing this, a single input variable is varied while the others remain fixed, making it feasible to analyze the direction and magnitude of the changes in the model's output in relation to the input variable. If two variables are varied simultaneously, a gradient plot can be obtained that represents as a 3D surface or contours the effects caused by the selected input variables interaction (WANG et al., 2020). Although this procedure is limited by the number of free input variables and by considering one sample at a time, the results can be useful for investigating specific steel samples and comparing the data-driven models output behavior to physical metallurgical principles, in search of correlations or possible contradictions.

Depending on the computational model created, it is possible to obtain many different or specific analyses, such as the Classification and Regression Tree (CART) obtained in Bright et al. (2011) by using tree based classification models. Although easy to graphically understand through a decision tree, this method is limited to very simple models, since its interpretability is quickly lost as the complexity of the tree based model increases.

Albeit accurate and powerful, a black box model can suffer from bias, technical and ethical limitations. As machine learning interpretability and explainability is further investigated by the scientific community, possible solutions to these problems are proposed. Consequently, there is a trend for the development of more human-centered, interpretable and explainable models trained on tabular data that still manages to provide high performance.

3.4.6 QUALITATIVE ANALYSIS

Qualitative analyses can be used to ground the outcomes achieved from modeling in comparison to theoretical and practical backgrounds, basically asserting a quality associated with the modeling or the steel optimization along with its outcome.

Such analysis can be fundamental for a practical understanding of the algorithmic limitations, as they allow a cross-reference between different models, empirical observations, physical laws, chemical and other phenomenological models.

In the work of Galindo-Nava, Rainforth e Castillo (2016), the authors created a modeling framework based on physical metallurgical principles relating the microstructure and mechanical properties of maraging steels, as well as a verification of which chemical elements have the greatest effect on the hardness and elongation of steel.

In order to generate a hypothetical steel with specific mechanical properties, in Wang et al. (2020) generated steels using genetic algorithms and compared their results with steels already used in the market. The differences between the chemical compositions of these real and hypothetical steels were analyzed considering metallurgical principles to verify the feasibility of the steels found by the optimization process.

From the selected works in the current literature review, a minority of studies present a qualitative analysis comparing the computational findings with experimental or theoretical backgrounds. With increasingly complex steels being modeled or optimized

3.4.7 SLR CONSIDERATIONS

Through the literature study, it can be seen that the modeling and prediction of mechanical properties of different types of steel is possible. However, as steels become intrinsically more complex, more complete and robust data becomes increasingly necessary for an adequate modeling. Since this data is usually empirically collected and thus susceptible to errors and anomalies, it becomes increasingly important to have a broad and deep knowledge of any metallurgical principles that can be helpful towards a better understanding of the limits of these models.

From all the used techniques, ANNs for modeling and MOEAs for optimization are the most common. However, peculiarities of different datasets can affect the modeling stage in different ways. Thus, without a consolidated dataset to serve as a benchmark, it is infeasible to determine which algorithm serves as the state of art from its metrics performance alone. Instead, the state of the art can be defined by the rigor and scrutiny present in the systems research and validation approach, its achieved results and its ability to correspond with relevant theoretical background from the metallurgy field.

When compared to analytical and theoretical models based on physical metallurgy principles, some studies (COSTA et al., 2019; WANG et al., 2020) asserts that data-driven models present better performance. Nonetheless, despite many black box data-driven techniques being capable of modeling an arbitrary number of inputs to output variables, performance can still be disregarded if explainability is a modeling objective. The fusion of data-driven human-centered models and expert knowledge is still an open question, for steel modeling and optimization purposes.

In conclusion, due to the complexity involved in steelmaking and its products, there are pitfalls that can hamper the use of prediction models, both phenomenological and data-driven. With the ongoing research and development in machine learning and the different branches of metallurgy, there are many possibilities to explore novel models that present better modeling capabilities, interpretability or multimodalism.

4 PROPOSED APPROACH

In this chapter, the flowchart used for steel mechanical properties prediction using the surrogate model and the surrogate-based steel optimization is described.

Figure 12 illustrates a simplified flowchart coupling the surrogate modeling and its application in steel optimization.

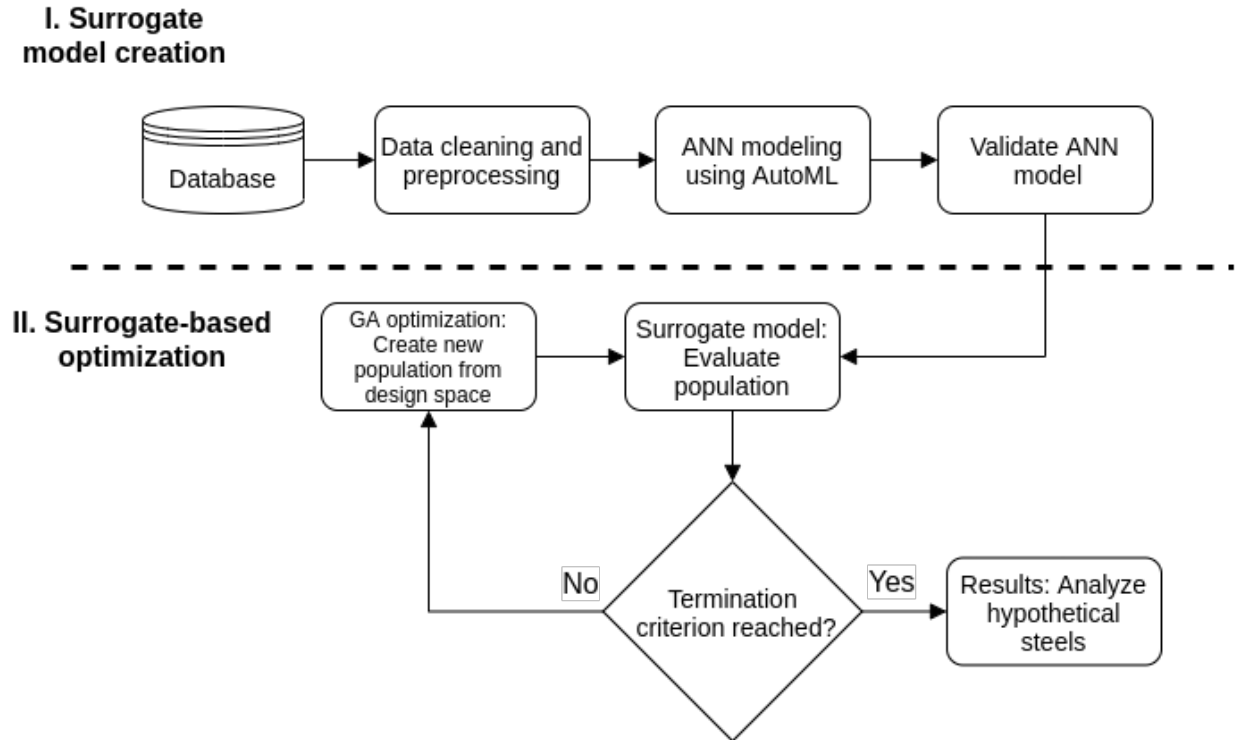


Figure 12 – Flowchart of the entire computational system. Source: Own authorship.

The surrogate-based optimizer is a MOEA coupled with an ANN trained on the problem domain dataset. Thus, the surrogate model inputs are chemical composition, processing parameters and dimensional information. Its output is mechanical properties evaluations, which are fed in the fitness calculation of the MOEA, which in turn creates a new generation of steels. Then, the surrogate model evaluates these hypothetical steels and the cycle repeats. Therefore, new steel designs can be achieved by the optimization procedure through the surrogate model guidance.

In Figure 12, the surrogate model creation starts with the data collection. After the relevant data for modeling is gathered from the steel's manufacturing historical database, data cleaning and pre-processing are done to make the data suitable for ANN training. The model training is automated with NAS, where the best models are obtained and validated. For the surrogate-based optimization, the best-validated model obtained from the creating procedure is coupled inside an NSGA-II algorithm, where it evaluates the population of steels generated by stochastic evolutionary procedures. If the optimization's termination criterion is reached, the results are then analyzed.

The decision to use NSGA-II as the optimization algorithm is due to several reasons such as: it is a state-of-the-art algorithm commonly used as benchmark; it is a suitable algorithm for optimizing multiple and conflicting variables; it is suitable when dealing with numeric variables; it has been successfully used to solve common problems in the literature as seen in Chapter 3 and there are multiple computational frameworks written in Python that already employs the algorithm's standard mechanism and operators.

4.1 NEURAL NETWORK MODELING

From the SLR, many studies indicate that it is commonplace to model steel mechanical properties using data-driven methods. This is partly due to the practicality of employing these methods compared to the arduous and time-consuming modeling of physical metallurgy principles into analytical equations, which may not be entirely applicable to different manufacturing lines.

Figure 13 represents the proposed surrogate modeling flowchart using Automated Machine Learning (AutoML). After the training and validation dataset is cleaned and pre-processed, AutoML optimizes the ANN hyperparameters. If the best models don't perform well on a validation dataset, then the automation was incapable of delivering sufficiently accurate models, prompting the data pre-processing stage to be reanalyzed. The accuracy thresholds are defined by specialists in the area of metallurgy. Then, the optimized ANNs are cross-validated in order to verify their reproducibility. If the best model's performance presents a low standard deviation, it is selected as the surrogate model. Otherwise, the second-best model is cross-validated and analyzed. This repeats until an accurate and reproducible model is found. This procedure is used to select a single ANN, but can be repeated multiple times in order to generate multiple ANNs to create an ensemble model.

In the present work, the mechanical properties of steels were modeled using ANNs due to their ability to model complex non-linear patterns. The ANN inputs consist of twenty-nine variables in total, being that fourteen of these are chemical elements and thirteen processing parameters from different manufacturing stages and two-dimensional parameters. The ANN outputs are three variables that correspond to the predicted mechanical properties, namely YS, TS and EL. By predicting the different outputs in a single network, we can define the ANN as a multi-variate regression model. All input and output variables can naturally be represented as integers or floating-point numbers.

4.1.1 NEURAL ARCHITECTURE SEARCH

Due to the many different hyperparameters involved in defining and training a neural architecture, albeit powerful, ANN's performance can be challenging to optimize. A common way to perform hyperparameter optimization is through an exhaustive search of a manually specified subset of the hyperparameter space, that is, the possible combinations of each hyper-

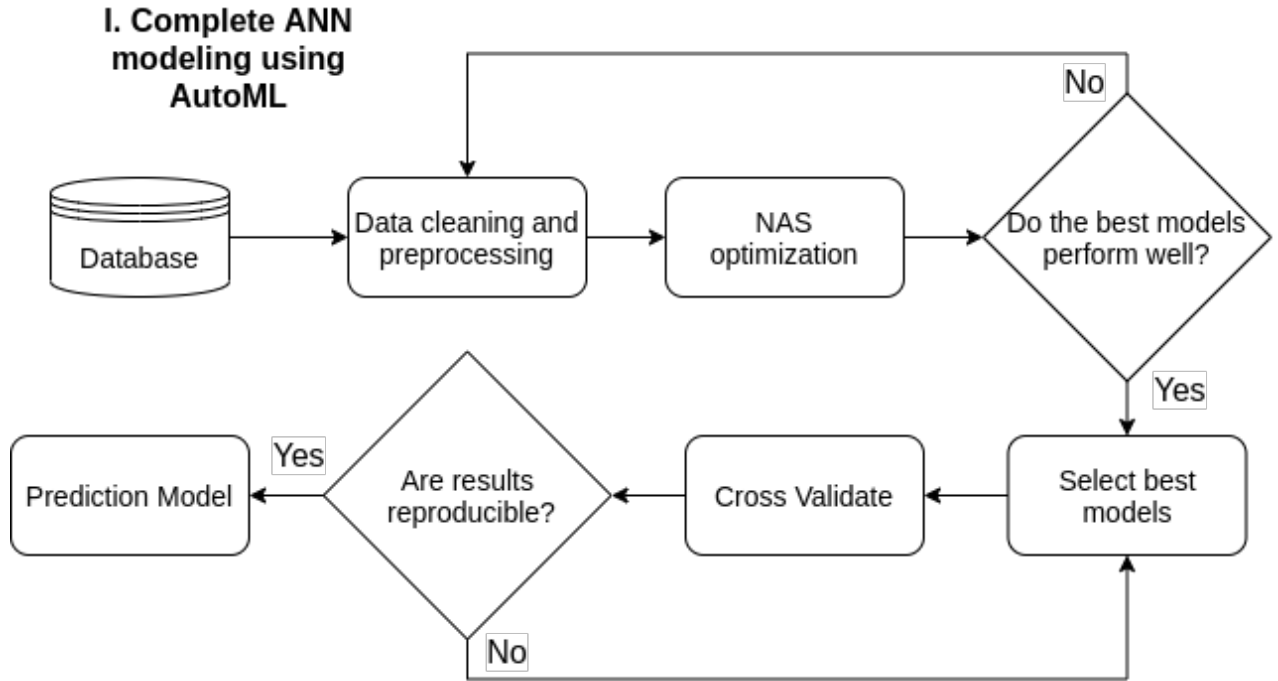


Figure 13 – Proposed surrogate modeling flowchart. Source: Own authorship.

Table 3 – Search space for the NAS parameters.

Parameter	Values
Number of layers	[1, 2, 3]
Number of neurons	[16, 32, 64, 128, 256, 512, 1024]
Use of drop-out in hidden layers	[Yes, No]
Drop-out value for hidden layers	[0, 0.25, 0.5]
Use of drop-out in output layer	[Yes, No]
Drop-out value for output layer	[0, 0.25, 0.5]
Use of batch-regularization	[Yes, No]
Optimizer	[stochastic gradient descent, adam, adam weighted decay (adam WD)]
Learning rate	[0.1, 0.01, 0.001, 0.0001, 0.00001, 0.00002]

parameter's values. Although simple, this trial and error can be computationally expensive and time-consuming since the search space increases geometrically in relation to the linear increase of each hyperparameter possible value.

AutoML was used to automate the task of hyperparameter optimization through the use of Auto-Keras, a Neural Architecture Search (NAS) (JIN; SONG; HU, 2019) framework. Auto-Keras uses Bayesian optimization to guide the network morphism while developing a neural network kernel and a tree-structured acquisition function optimization algorithm to explore the ANN architecture search space efficiently.

The hyperparameters that define the search space for the NAS are specific to the number of layers, neurons, regularization and optimization of the ANN and are defined according to Table 3.

The number of neurons per hidden layer can be different for each individual layer, which increases the search space significantly.

The metrics selected to measure the performance of the models during the architecture search were the Mean Squared Error (MSE), the Mean Absolute Error (MAE) and the Coefficient of Determination (R^2). The loss function was the MSE, which the model should seek to minimize during training. The hold-out method is used to build the model, using 80% of the steel samples as training data and 20% as validation data.

The hold-out method is used to build the model. From the dataset, 15426 entries have been used for ANN training and 3864 as the validation set.

Based on an initial random architecture within the scope of the parameters defined in Table 3, the NAS algorithm searches through different architectures by testing distinct combinations of parameters.

A different NAS was conducted for three types of activation functions, using the search space defined within the scope of Table 3. The activation functions used are the Rectified Linear Unit (ReLU), the Sigmoid and the Hyperbolic Tangent (Tahn). By defining 200 model evaluations as the stopping criterion, the model with the highest performance from each NAS is selected to be compared. Although the activation function itself could be selected through NAS, since there are many other parameters, there would be no guarantee that the number of ANNs with different activation functions would be balanced, which could add bias to the search process.

To obtain the mean and standard deviation of the performance from the selected ANNs, a K-Fold Cross-Validation technique is applied, partitioning the dataset into ten folds and retraining the models for each fold. Then, sensitivity analysis and the impact of the input variables are analyzed for each ANN.

4.2 SURROGATE-BASED MOEA FOR STEEL DESIGN OPTIMIZATION

Since modeling steels mechanical properties by using data-driven methods is established in the literature, these models are suitable for surrogate-based optimization. However, as seen in the SLR, since there are many different steels with distinct sets of input parameters, each design optimization can consequently carry different objectives and limitations. For instance, in Costa et al. (2019) DP steels had heat treatment parameters optimized to minimize YS and maximize both TS and EL. In contrast, in Shen et al. (2019) and Dutta et al. (2019) studies, both chemical composition and processing parameters were optimized for different input parameters, surrogate models, steel categories, mechanical properties and objectives.

Hence, there are many possibilities, improvements and open questions related to the capabilities and limitations of steel design optimization, particularly for complex steels. In the present work, the IF, BH, HSLA, DP and TRIP steel grades are modeled by a single ANN coupled with an NSGA-II to investigate such possibilities and limitations. The surrogate-based optimization flowchart is illustrated in Figure 14, where the optimized ANN found using NAS is

used to evaluate the population optimized using NSGA-II. First, an initial population is created from random values. The individuals from the population are ranked and evaluated by the surrogate model. From this evaluation, the parents are selected for the next genetic operators, where mutation and recombination are applied following deterministic rules. To the new pool of parents and offspring, elitism is applied in order to maintain the best individuals from the past generation, thus, creating the updated generation with new characteristics. This population would be evaluated by the surrogate, repeating the cycle until the termination criteria is reached. If so, the results are analyzed.

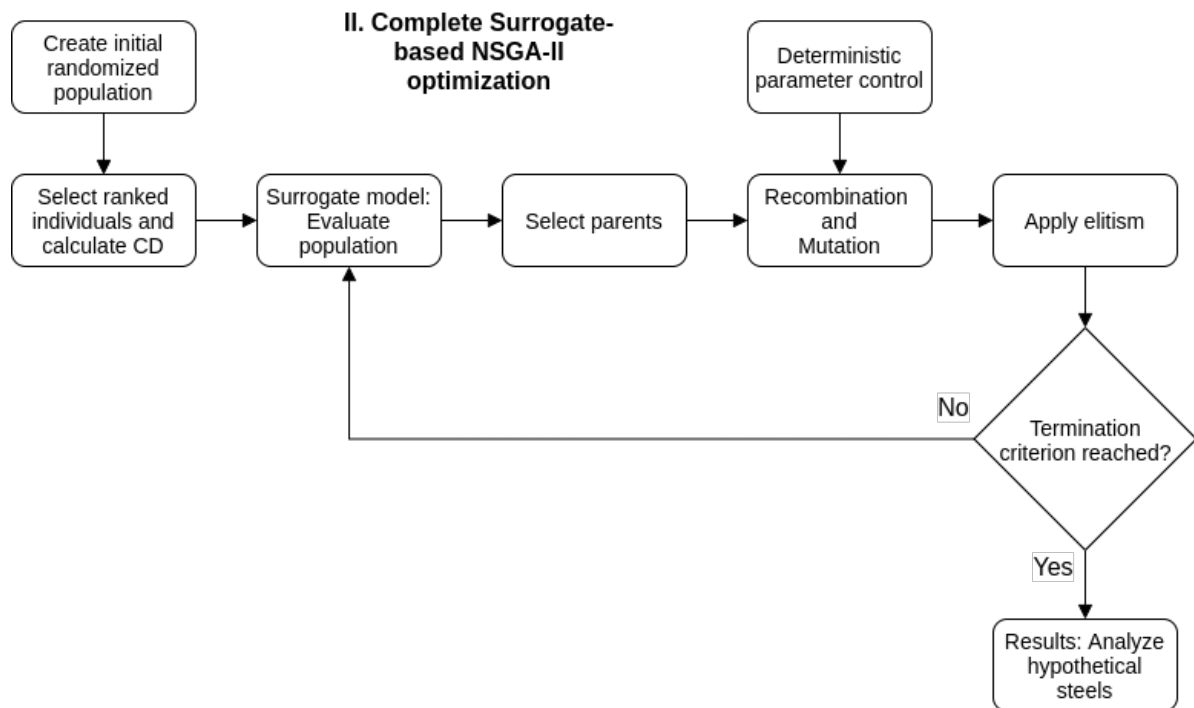


Figure 14 – Proposed surrogate-based optimization flowchart. Source: Own authorship.

4.2.1 OBJECTIVE FUNCTIONS

In this work, instead of maximizing or minimizing certain mechanical properties, an alternative objective is proposed in order to evolve steel designs whose properties are sufficiently close to a predefined set of desired mechanical properties.

By doing this, different steels can be evolved that satisfies specific mechanical property criteria. This is useful in a manufacturing context where:

1. the steels mechanical properties need to be within certain technical specifications and meet specific standards for a client or application
2. different chemical compositions are compared in order to cut costs by manufacturing steels with cheaper alloying elements
3. alternative processing parameters are required due to manufacturing limitations

4. multiple values for the same type of mechanical property is desired

This is done by creating an interpolation function for each objective that closely mimics a piece-wise linear function as described by Equation 8, where in this context p_1, p_2, \dots, p_n are user defined global optima, m_1, m_2, \dots, m_{n-1} are the mid-point between each consecutive pair of global optima and x the mechanical property.

$$PWLF(x) = y(x) = \begin{cases} -x + p_1 & x < p_1 \\ x - p_1 & p_1 \leq x < m_1 \\ -x + p_2 & m_1 \leq x < p_2 \\ x - p_2 & p_2 \leq x < m_2 \\ \vdots & \\ x - p_n & x > m_{n-1} \end{cases} \quad (8)$$

By defining piece-wise linear functions in this manner, it is possible to systematically create user defined objective functions with an arbitrary number of global optima. Besides, by defining a single global point p , Equation 8 can be written as $y = |x - p|$, which denotes the absolute error in relation to the global optimum, providing a pragmatic interpretation of the fitness function. Consequently, Equation 8 can be interpreted as the mechanical property prediction error relative to the nearest global optima. Figure 15 illustrates an example of Equation 8 for hypothetical global optima p_1 and p_2 .

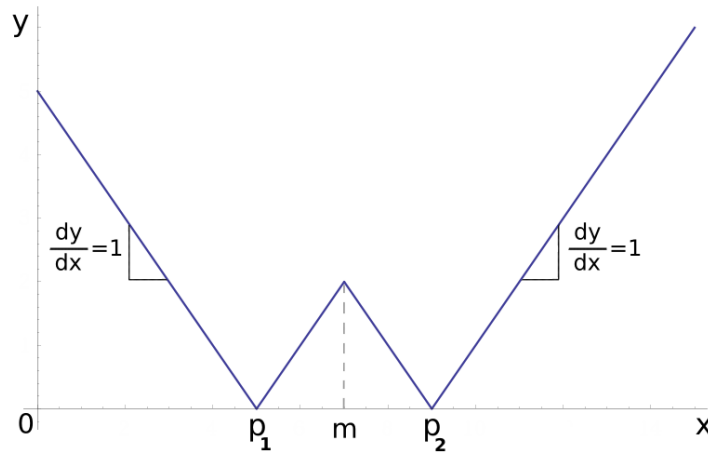


Figure 15 – Piece-wise linear function with unitary slope, where the global optima are located at p_1 and p_2 .

However, since there are three mechanical properties being optimized, the overall fitness function can be written according to Equation 9,

$$f(X) = \begin{cases} f_{YS} = PWLIF_{YS}(YS_{OPT}) \\ f_{TS} = PWLIF_{TS}(TS_{OPT}) \\ f_{EL} = PWLIF_{EL}(EL_{OPT}) \end{cases} \quad (9)$$

where *PWLIF* stands for Piece-Wise Linear Interpolation Function, representing the objective function for each mechanical property as denoted by Equation 8, while YS_{OPT} , TS_{OPT} and EL_{OPT} are the mechanical properties predicted by the surrogate model from the NSGA-II evolved population.

4.2.2 EA PARAMETER CONTROL

In this study, the genetic operators involved in the optimization, namely mutation and recombination, are not fixed throughout the NSGA-II run time. These parameter values are controlled according to deterministic rules, a strategy where no feedback information from the optimization is used. Instead, the rules are specified by the user before starting the optimization (ALETI; MOSER, 2011). Since the parameters are changed while the algorithm is running, it belongs to a category of online parameter control.

Deterministic parameter controls are generally very simple in nature and computationally cheaper than more robust strategies. However, since this control technique does not take into consideration the NSGA-II behavior, the actual progress of the optimization itself, or information from the fitness function (PARPINELLI et al., 2019), it can lead to suboptimal results for some problems. This is especially the case when optimizing problems of different scales; since complexity can alter the speed of the search progress and it is not obvious to predict the number of generations the NSGA-II will take to converge, possibly undermining the determined rules efficacy set a priori (ALETI; MOSER, 2011).

In this work, a deterministic rule is used such that the mutation and recombination rate is updated in every generation, according to Equations 10 - 11, such that the recombination rate starts at 0% and rapidly grows with a decreasing slope until it reaches 100% at g_F , while the mutation rate starts with 20% and is decreased to 0% at g_F , where g_F refers to the last generation. The recombination and mutation rate curves are illustrated in Figure 16.

$$\text{Recombination rate} = \frac{\log_{10}g}{\log_{10}g_F} \quad (10)$$

$$\text{Mutation rate} = 0.2 \cdot \left(1 - \frac{\log_{10}g}{\log_{10}g_F}\right) \quad (11)$$

Even with its simplicity and absence of feedback information, the optimization results achieved by using deterministic controls are usually better than using static values throughout the search process (PARPINELLI et al., 2019), due to a better exploration-exploitation balance by the online parameter control.

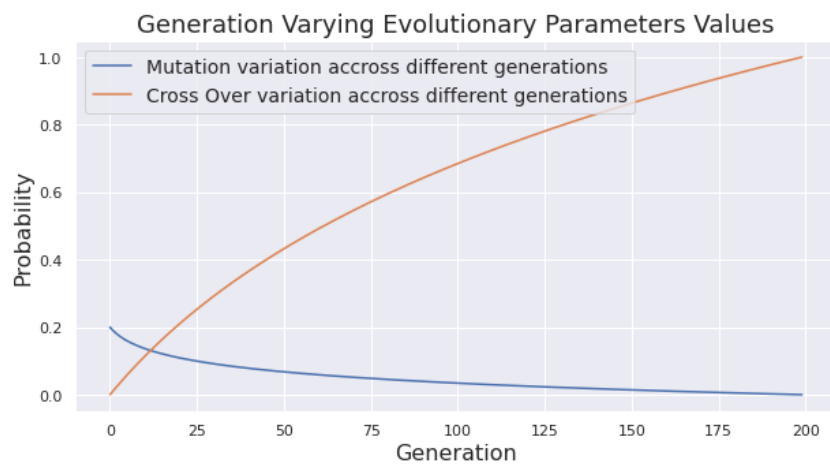


Figure 16 – Online parameter control using deterministic rules for Recombination and Mutation.
Source: Own authorship.

5 RESULTS AND ANALYSIS

This chapter presents different results with corresponding analyses of different applications from the models and algorithms achieved so far.

5.1 NEURAL NETWORK MODELING

In this section, the dataset used in the ANN modeling is described. However, the dataset is composed of many steel grades on different and imbalanced quantities. The effects of modeling the data under different magnitudes of imbalance is studied in order to verify its consequences on the regressor using xAI (explainable Artificial Intelligence) techniques.

5.1.1 DATABASE

The dataset used in the present work comprises twenty-nine input parameters taken and organized from different processing lines, being Steel Mill (SM), Hot Strip Mill (HSM), Cold Rolling Mill (CRM) and Hot Dip Galvanizing (HDG) line. In SM, the iron ore is melted and different alloying elements can be added. In HSM, the steel is shaped and has its transverse section reduced from the pressure exerted on one or multiple rolling stands. This is done in very high temperatures to achieve a greater reduction in the transverse section, hence the term hot in its name. This is opposite to the CRM, which is not done in high temperatures. The CRM stage, although similar to HSM, achieves different properties due to being processed close to ambient temperatures, such as higher resistance due to the stress involved in the rolling. Then, at the HDG, the steel is reheated for further thermal processing and dipped into a molten zinc bath in order to gain better performance against corrosion. Figure 17 shows the different stages of the process and the origin of the variables used in the model.

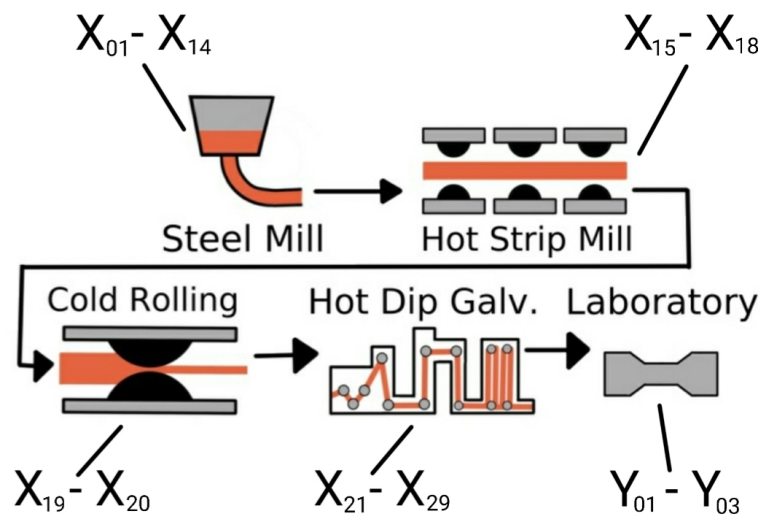


Figure 17 – Stages of the steel manufacturing process and involved parameters. Source: Own authorship.

From this set, fourteen parameters consist of chemical composition and thirteen are related to processing lines and two are dimensional variables. Three different mechanical properties are defined as output parameters, which are measured at the laboratory (LAB). Each variable is described in Table 4 and was selected by an expert from the field.

Table 4 – Description from the models variables, where CE, PP, DIM and MP stands for Chemical Element, Process Parameter, Dimension and Mechanical Property, respectively.

Name	Description	Type	Unit	Origin
X ₀₁	Carbon content	CE	ppm	SM
X ₀₂	Manganese content	CE	ppm	SM
X ₀₃	Phosphorus content	CE	ppm	SM
X ₀₄	Silicon content	CE	ppm	SM
X ₀₅	Sulfur content	CE	ppm	SM
X ₀₆	Nickel content	CE	ppm	SM
X ₀₇	Aluminium content	CE	ppm	SM
X ₀₈	Chrome content	CE	ppm	SM
X ₀₉	Niobium content	CE	ppm	SM
X ₁₀	Molybdenum content	CE	ppm	SM
X ₁₁	Titanium content	CE	ppm	SM
X ₁₂	Vanadium content	CE	ppm	SM
X ₁₃	Boron content	CE	ppm	SM
X ₁₄	Nitrogen content	CE	ppm	SM
X ₁₅	Coil width	DIM	mm	HDG
X ₁₆	Coiling temperature	PP	°C	HSM
X ₁₇	Heating temperature before hot rolling	PP	°C	HSM
X ₁₈	Heating temperature after hot rolling	PP	°C	HSM
X ₁₉	Cold rolling reduction	PP	%	CRM
X ₂₀	Thickness after hot rolling	DIM	mm	CRM
X ₂₁	Line speed	PP	m/s	HDG
X ₂₂	Mean Skin Pass Mill elongation	PP	%	HDG
X ₂₃	Temperature from Pyrometer 03	PP	°C	HDG
X ₂₄	Temperature from Pyrometer 04	PP	°C	HDG
X ₂₅	Temperature from Pyrometer 10	PP	°C	HDG
X ₂₆	Temperature from Pyrometer 12	PP	°C	HDG
X ₂₇	Temperature from Pyrometer 13	PP	°C	HDG
X ₂₈	Temperature from Pyrometer 15	PP	°C	HDG
X ₂₉	Temperature from Pyrometer 16	PP	°C	HDG
Y ₀₁	Yield Strength	MP	MPa	LAB
Y ₀₂	Ultimate Tensile Strength	MP	MPa	LAB
Y ₀₃	Elongation	MP	%	LAB

The collected data are from 2018-2020 and were filtered to select data from the IF, BH, HSLA, DP and TRIP steel grades. After eliminating data inconsistencies, incomplete data and extreme outliers, the dataset comprises a total of 52390 data entries.

The database is composed by a collection of industrial data collected by the company ArcelorMittal. Most of the data are derived from automated measurement systems such as sensors and are centralized in a computer database. Some variables, such as mechanical properties, are inputted directly from the user measuring such properties. The data related to the SM and HSM stages are collected in the industrial plant located in the state of Espírito Santo in Brazil, while the data from CRM, HDG and LAB stages are from the industrial plant in Santa Catarina, Brazil.

5.1.2 IMBALANCED DATASET AND ITS CONSEQUENCES

As seen in Section 2.1.2, there are many different types of steels, each with their own particularities. Although most of their parameters are common to each other, some steels can have their own grade depending on the quantity of such parameters and the metallurgical phenomena that arises from such combinations, which isn't apparent by just inspecting the tabular data format presented as the training data.

The problem of imbalanced datasets is very well known in classification tasks (KOTSIANTIS; KANELLOPOULOS; PINTELAS, 2005; HE; GARCIA, 2009; YANMINSUN; WONG; KAMEL, 2011). However, it isn't very discussed in regression tasks, even though a regression model trained on an imbalanced dataset can carry many of the same problems which occur in a classification task (BRANCO; TORGO; RIBEIRO, 2015).

The current sub-section analyses the results and particularities of modeling ANNs on datasets with different degrees of imbalance.

5.1.2.1 STRATIFIED MODELS

To qualitatively and quantitatively measure the consequences of unbalanced classes on the trained regressor, different models were trained exclusively on different parts of the dataset. The steels were grouped in four clusters and for each cluster of steel grades an ANN was trained on and the most important features were ranked using SHAP analysis, with results presented in Figure 18.

Carbon appears as a major feature in IF steels. Since one of the metallurgical principles of IF steels is the major reduction in carbon content atoms present in the steel's ferritic matrix and given that carbon is a known ferrite strengthener (SANTOS et al., 2018), it is a crucial element to be controlled during the manufacturing process. This corroborates with the results found by Mohanty, Bhattacharjee e Datta (2011), which also modeled IF using ANNs and analyzed its feature importance. Also, according to Santos et al. (2018), phosphorus, silicon and manganese can be added to the steel in order to increase its resistance by promoting solid solution hardening, elements which are present in the IF feature ranking. Titanium and/or niobium are microalloying elements that are added to the steel to promote the formation of precipitates (SANTOS et al., 2018) and are also on the top ranking features.

Similarly with IF steels, carbon can play a major role in BH steels, since, according

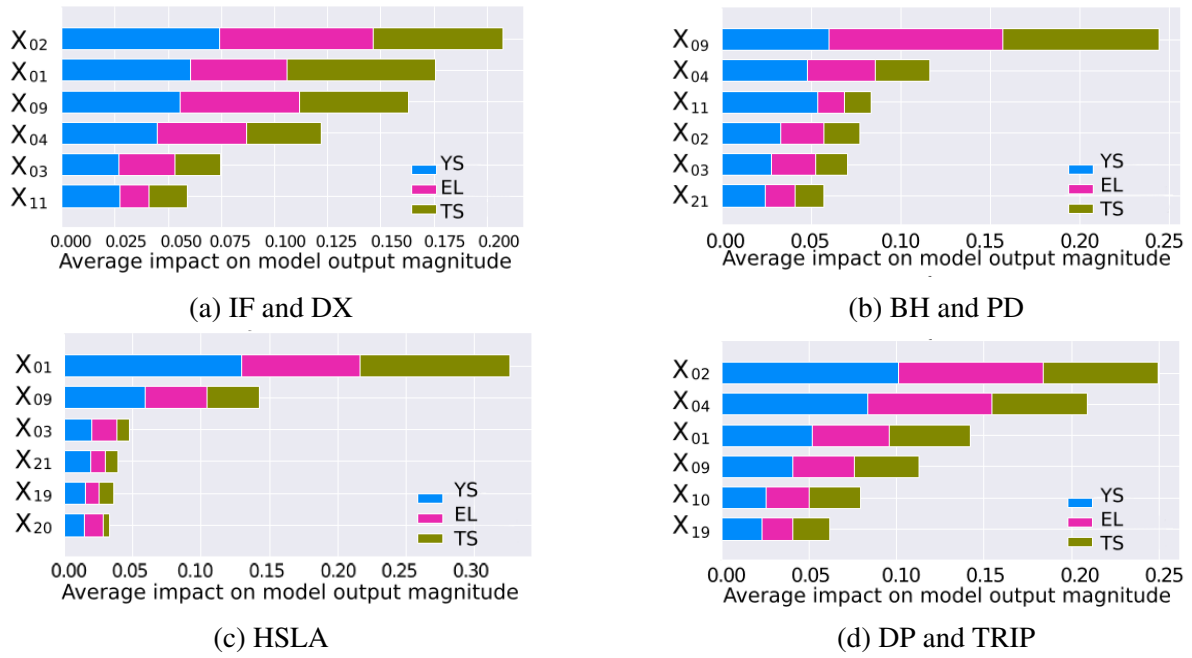


Figure 18 – SHAP feature importance ranking for ANNs trained on different steel grades.

to Baker, Daniel e Parker (2002), carbon is one of the solute atoms responsible for the bake hardening response in BH steels. According to the same authors, special attention during coiling, annealing and temper rolling are needed to ensure that a certain amount of carbon remains in solution, in order to promote the bake hardening results, which can possibly explain the inclusion of a processing feature and a dimensional feature, result of the processing stage.

The HSLA steel in our dataset are microalloyed steels, which means they have alloying elements added to their composition. There are many existing empirical equations, but according to Kestenbach, Campos e Morales (2006) it is common to use structure-property relationship developed by Pickering (1978), expressed as:

$$\sigma_y[MPa] = 15.4(3.5 + 2.1pctMn + 5.4pctSi + 23pctN_f + 1.13d^{-1/2}) \quad (12)$$

where pct Mn, pct Si and pct N_f are the percentages of manganese, silicon and free nitrogen, respectively, dissolved in ferrite and d is the ferrite grain size in mm. If there are strong carbide/carbonitride forming elements in sufficient amounts present in the steel, such as niobium, vanadium and/or titanium, N_f can be assumed to be negligible (SHOW et al., 2010). This can possibly explain the four highest ranking features in the HSLA-modeled ANN.

While molybdenum only appears in the top ranking features of DP and TRIP steels, its effects were studied by many authors for both DP and TRIP steels (MUKHERJEE; HAZRA; MILITZER, 2009; ABBASI; RAINFORTH, 2019). The conclusions from the authors state that the molybdenum is one of the elements with a strong connection between the microstructural evolution and precipitation coarsening, although this behavior changes greatly depending on the steel, processing parameters and alloying elements.

Both DP and TRIP steels can be considered to be complex steels on their own, making it more difficult to correlate the top ranking features with metallurgical literature findings, making an in-depth analysis out of the scope of this work.

Nonetheless, without any ad-hoc knowledge, the driving factors behind many of each steels important metallurgical phenomena seems to appear in the top ranking modeled features. Although the ANN can still be considered a black box and correlation does not imply causation, this could be considered a good sign regarding the model's training algorithm and interpretability.

5.1.2.2 UNIFIED MODEL

One of the objectives from this present work is to create an algorithm capable of theoretically discovering new and novel steels. Thus, instead of modeling each steel grade individually, integrating all steel grades into a unified training dataset can be beneficial, since each grade have it's own peculiarities and can provide an extended data interval when combined.

To analyze the consequences of class imbalances, two models were trained on different datasets. The first model was trained on the entire dataset, which due to the high class imbalance is named *completely imbalanced*. The second model was trained on a pre-processed dataset which aims to remove redundant data from classes that outnumber smaller classes and is named *less imbalanced*. The proportion of each steel grade in the datasets are presented in Table 5.

Table 5 – Proportions of each steel grade in each dataset.

Dataset	IF %	BH %	HSLA %	DP and TRIP %
Completely imbalanced	60.7	23.6	11.9	3.65
Less imbalanced	21.1	36.3	32.6	10

With a better notion about how the model is interpreting the features from the training data for different steel grades, as seen in 5.1.2.1, the consequences of class imbalances can be analyzed by comparing the top ranking features obtained from models trained on unified datasets in relation to the models trained on stratified datasets. The unified model's ranking features are illustrated on Figure 19.

For brevity's sake, only 12 from the 29 features are shown in Figure 19. For both models, carbon is the most influential feature. However, since the scales are different, the carbon's influence on the *completely imbalanced* model is twice as much from the *less imbalanced*, overshadowing the other features' contributions. The following 5 features, namely titanium, manganese, silicon, niobium and phosphorus, are common to both rankings, albeit in slightly different order. Moreover, the *less imbalanced* feature rank is the only one that presents a parameter from the HDG line - P3 Med and an important feature from DP and TRIP steels - molybdenum.

Thus, in order to satisfy the discoverability of novel steels, the *less imbalanced* dataset and model is preferred due to its ability to better aggregate the contributions from different

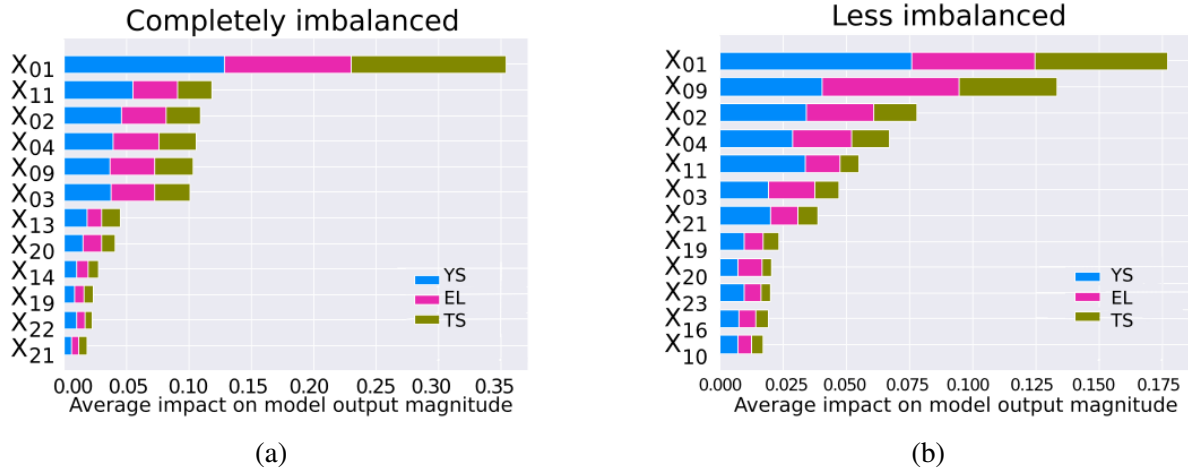


Figure 19 – SHAP feature importance ranking for ANNs trained on (a) completely imbalanced and (b) less imbalanced data.

features into its predictions.

5.1.3 QUANTITATIVE RESULTS

Although xAI is capable of providing insights about the inner functioning of black-box models, some of these insights rely on human interpretation and most do not include direct quantitative measurements of performance. For instance, it was concluded previously that despite the smaller quantities of data, a model trained on a less imbalanced dataset is preferred rather than a completely imbalanced dataset due to the stronger contributions that each parameter is able to provide.

However, quantitative results are still needed and the jury's still out regarding the advantages of combining different steel grades instead of using an ensemble of multiple ANNs, each trained on a different cluster of steel grades from the dataset.

Thus, an experiment was conducted to measure the general strengths and flaws of modeling an ANN on different steel grades. First, stratified models regarding specific steel grades and a unified model on less imbalanced dataset were trained. This resulted in 5 different models trained on different subsets of data. Then, the performance of each model on all the different steel grade clusters were measured by averaging the prediction's *MRE* of all different mechanical properties.

For instance, a model trained on IF dataset had its performance measured against IF, BH, HSLA, DP and TRIP steels. The same happened for models trained on other dataset, stratified or unified. Hence, since all models are trained on the same set of variables but in different data intervals, this experiment measures how well each model can extrapolate their predictions under different input intervals.

Table 6 presents the results from repeating this experiment 50 times in order to obtain mean and standard deviation data. Cells with a darker background represent that the row's model was trained on a subset of the columns data.

Table 6 – Mean Relative Error from each steel grade obtained by different models trained on different subsets of data.

Grades used in modeling	MRE [%] of steel grades during inference stage			
	IF	BH	HSLA	DP and TRIP
IF	3.15 ± 0.07	5.73 ± 0.91	8.62 ± 1.80	33.7 ± 4.31
BH	16.2 ± 1.50	2.75 ± 0.06	60.7 ± 7.22	81.5 ± 9.59
HSLA	35.7 ± 5.30	25.9 ± 4.01	2.83 ± 0.13	34.8 ± 6.65
DP and TRIP	72.2 ± 5.39	47.3 ± 4.01	32.4 ± 2.90	3.83 ± 0.14
All grades	3.84 ± 0.67	3.12 ± 0.19	2.92 ± 0.16	4.73 ± 0.50

From Table 6 it can be seen that all stratified models present smaller errors than the unified model regarding the steel grade in which they were trained. That is, an ANN modeled after IF steels are better at predicting IF steels than an ANN modeled on all combined steels and the same happens for all other stratified models.

However, a unified model can still demonstrate a high accuracy regarding all steel grades simultaneously. Furthermore, stratified models can show terrible results when predicting steels from grades they weren't trained on. Although this can be expected, it clearly shows that, albeit more accurate on a specific steel grade, stratified are more limited than a unified model.

Since one of the objectives from this present work is to use the ANN as a surrogate model for novel steel discovery, by using a unified model instead of an ensemble of different stratified models, the overall complexity involved can be reduced while still attaining reasonable results. Hence, from the qualitative results, quantitative performance and by applying Occam's razor, unified modeling is preferred. Furthermore, ensembles of unified models can still be used in order to improve performance and reduce the prediction's variance.

Finally, the quantitative results shown in Table 6 not only corroborates with the qualitative findings presented by the SHAP values, but also exemplifies the efficiency of using an ANN as a surrogate model, by being able to adequately model multiple output variables of many different types of steels simultaneously.

5.2 OPTIMIZATION ANALYSIS

This section is intended to present the performance from the modifications or additions made to the NSGA-II algorithm presented in Chapter 4, in order to justify such developments.

Since the objective functions defined for the optimization task are related to the minimization of the discrepancies between optimized and the desired mechanical properties, the optimization complexity depends on the desired properties and the steel's predefined process parameters. Thus, for all experiments in this section, random steels were selected in order to sample without bias from the dataset. The optimizations used a population of 500 individuals and 500 generations.

5.2.1 ONLINE PARAMETER CONTROL ANALYSIS

To demonstrate and compare the effectiveness of genetic operators' online parameter control using deterministic rules, twenty random steels from the database had their chemical composition optimized in order to achieve a set of desired mechanical properties. These properties were equal to the real steel's properties from the database.

For each randomized steel, the optimization was performed multiple times using different fixed recombination and mutation rates and online parameter control. The results can be seen in Figure 20 and 21.

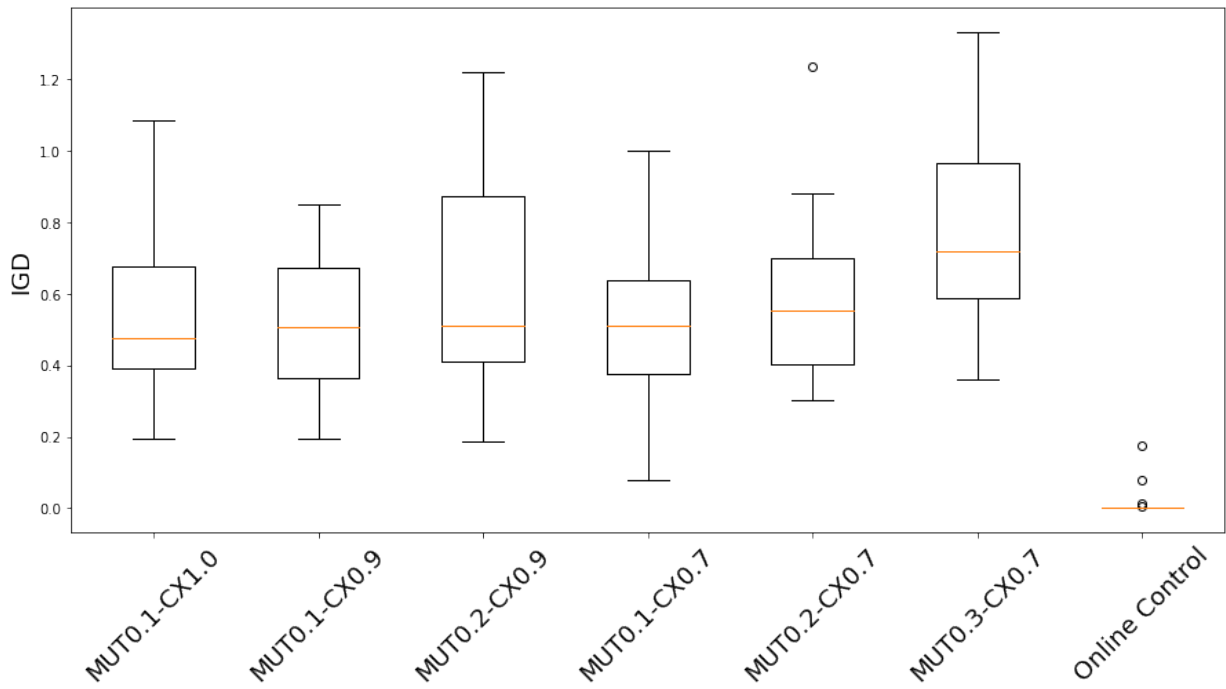


Figure 20 – Box-plot of the IGD indicator for different recombination (CX) and mutation (MUT) parameters. Source: Own authorship.

From Figure 20, it can be seen that the NSGA-II using online control achieved the smaller values of the IGD indicator, indicating better performance. Since IGD is a binary indicator that requires a discrete representation of the Pareto set (AUDET et al., 2020), the Pareto set used for this calculation is the collection of all non-dominated solution candidates from all Pareto approximations. Given the IGD measurement drawback presented in Chapter 2, it only indicates the convergence of the different Pareto approximations towards an idealized Pareto approximation.

In Figure 21, the NSGA-II using online control achieved the highest values of the DCI indicator, indicating better performance. Since the DCI is a k-nary indicator, recommended to compare the spread of different Pareto front approximations (LI; YANG; LIU, 2014), it can be concluded that the solution candidates found using online parameter control not only dominates the solutions found using fixed rates but are also more spread out in the objective space.

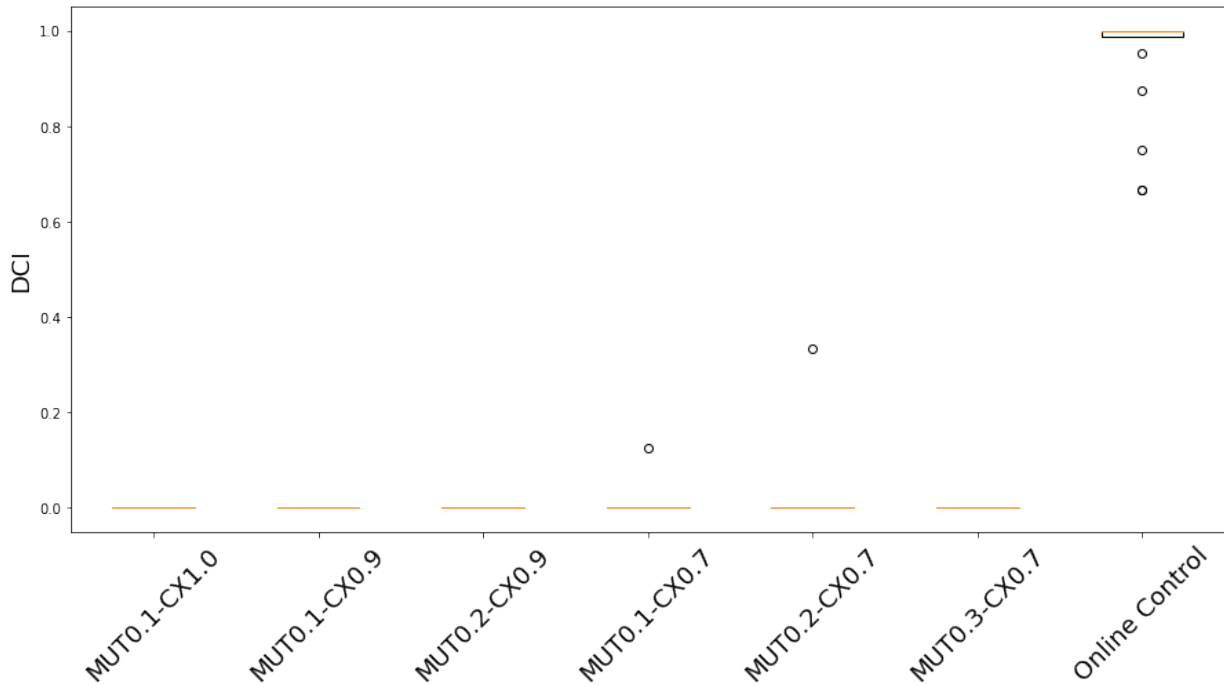


Figure 21 – Box-plot of the DCI indicator for different recombination (CX) and mutation (MUT) parameters. Source: Own authorship.

Finally, since online parameter control achieves better performance in relation to a convergence measure - IGD and a spread indicator - DCI, it can be demonstrated that the application of deterministic rules for recombination and mutation rates, albeit being a simple strategy, achieves better results than fixed rates.

5.3 CASE STUDIES

Different case studies have been made in order to study the application and practical performance of the present work's algorithm, including the modeling and inference of novel steels, the capabilities of the NSGA-II to optimize and discover new steels and finally, the integration of both optimization, prediction and specialist knowledge in order to propose novel steel designs.

5.3.1 CASE STUDY 1: MODELING AND INFERENCE OF NOVEL BH INDUSTRIAL STEELS

In this case study, an ANN modeled on IF and BH steel was used in order to predict the mechanical properties of novel and experimental BH steels. The case study is organized in two parts: modeling and inference.

5.3.1.1 IF AND BH MODELING

The architecture from the best model found from each NAS is illustrated in Figure 22.

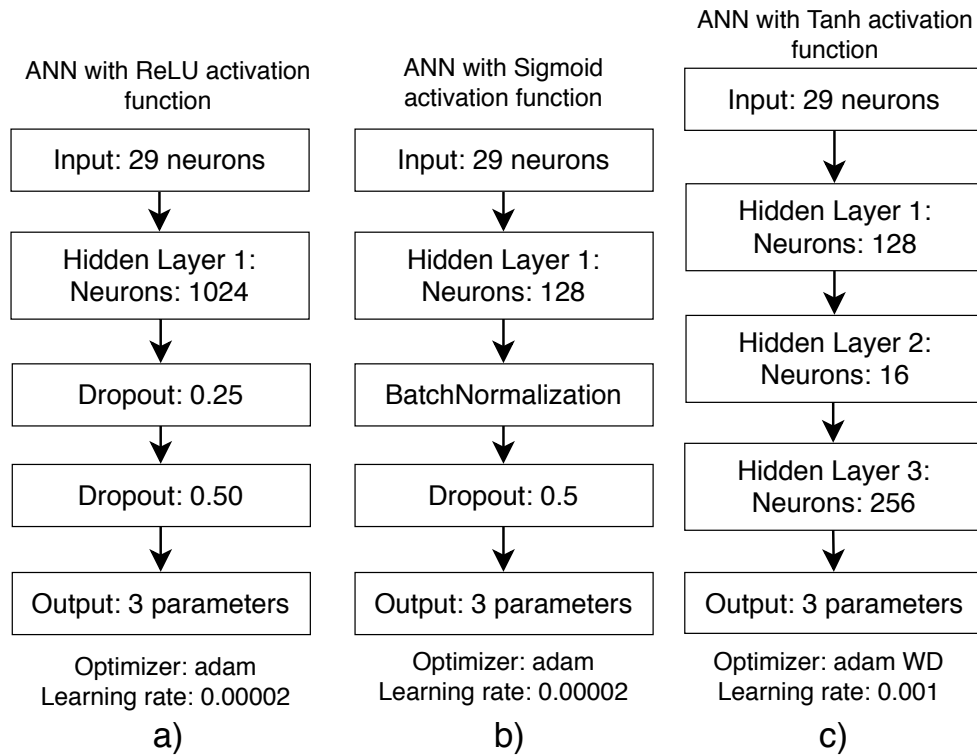


Figure 22 – Best ANN found for different activation functions, such as (a) ReLu, (b) Sigmoid and (c) Tanh. Source: Sgrott et al. (2021).

As shown in Figure 22, through the use of NAS, Auto-Keras can find and optimize distinct ANN architectures with diversified sets of hyperparameters. For instance, the ANN with ReLU and Sigmoid as activation function contain only a single fully connected hidden layer and the use of Drop-out. In contrast, the Tanh ANN contains three hidden layers of different sizes without Drop-out. The ANN evolved using the Sigmoid activation function is the only one that employs Batch Normalization.

By retraining these neural architectures using the cross-validation technique, it is feasible to evaluate their performance taking into consideration the modeling reproducibility. The results are shown in Table 7 for a 10-fold cross-validation.

Table 7 – R^2 metric with mean and standard deviation using 10-fold cross-validation.

Optimized ANNs	R^2 metric
ReLU-based ANN	0.908 ± 0.0021
Sigmoid-based ANN	0.887 ± 0.0033
Tanh-based ANN	0.899 ± 0.0029

According to Table 7, it can be seen that the ReLU activation function provided superior results, followed by the Tanh and Sigmoid functions. However, besides the diversified parameters such as different layers, training optimizer and activation functions, all ANN exhibits a low standard deviation associated with their metric, indicating that the performance obtained from

their training is reproducible.

Concerning the model loss of the ANN for the training and validation data, Figure 23 illustrates that the ANN does not suffer from overfitting or underfitting. Although Figure 23 illustrates the model loss from an ANN architecture using ReLU, similar results were obtained using different activation functions.

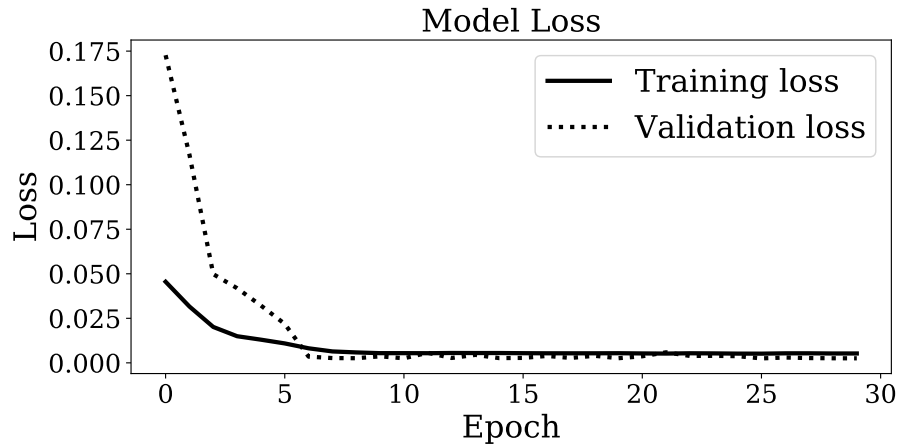


Figure 23 – Training and validation loss of training epochs using the NAS architecture for the ReLU activation function. Source: Sgrott et al. (2021).

Numerical experiments have been performed to investigate the effects of different input parameters on the mechanical properties using sensitivity analysis. As seen from the SLR, sensitivity analysis is a common quantitative analysis used in the context of modeling steel properties in order to investigate the design space approximation. Sensitivity analysis is defined as the study of how uncertainty in the output of a model can be attributed to different sources of uncertainty in the model input (SALTELLI et al., 2008).

In Figure 24, the manganese content was varied while the remaining parameters were kept fixed. For all three different ANNs, the addition of Manganese strengthens the steel, an expected result based on the available literature (TAKECHI, 1994). Although this behavior is generally known, the magnitude of this parameter's influence can vary depending on the other chemical composition and process parameters, that in this case, were fixed. The sensitivity analysis, illustrated in Figure 24 shows similar behavior for all the three ANNs, despite their different architectures. However, the differences in their rate of change, for YS and TS properties, the predictor is monotonically increasing while monotonically decreasing for EL property. Although this is a local analysis, it can carry good quantitative comparability to study different models' behavior on specific steel samples.

Although the localized sensitivity analysis is useful to understand the effects of individual input, it cannot consider how multiple input features interact with each other and with the model output simultaneously. This makes the sensitivity analysis not representative of the global effect of predictors upon the model's output.

To better understand the global impact of the input parameters relative to the mechanical

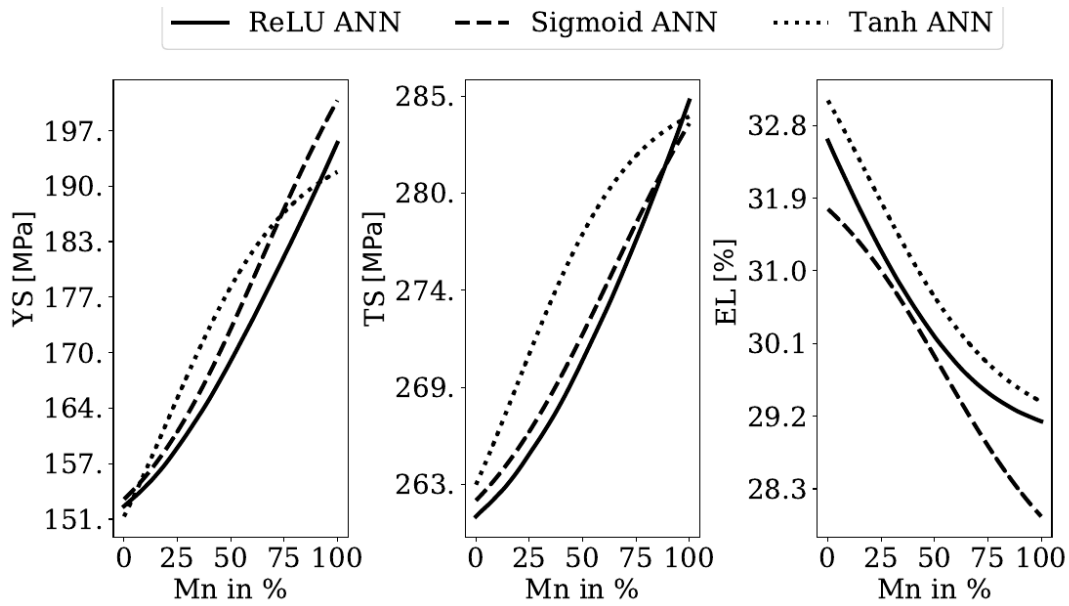


Figure 24 – Sensitivity Analysis for the three different ANNs. Source: Sgrott et al. (2021).

properties, SHAP (LUNDBERG; LEE, 2017) was used to calculate the distribution of each feature's impacts on the model output. When applied to regression models, this technique enables the estimation and graphical evaluation of every feature's relative importance for every data sample, as illustrated in Figure 25.

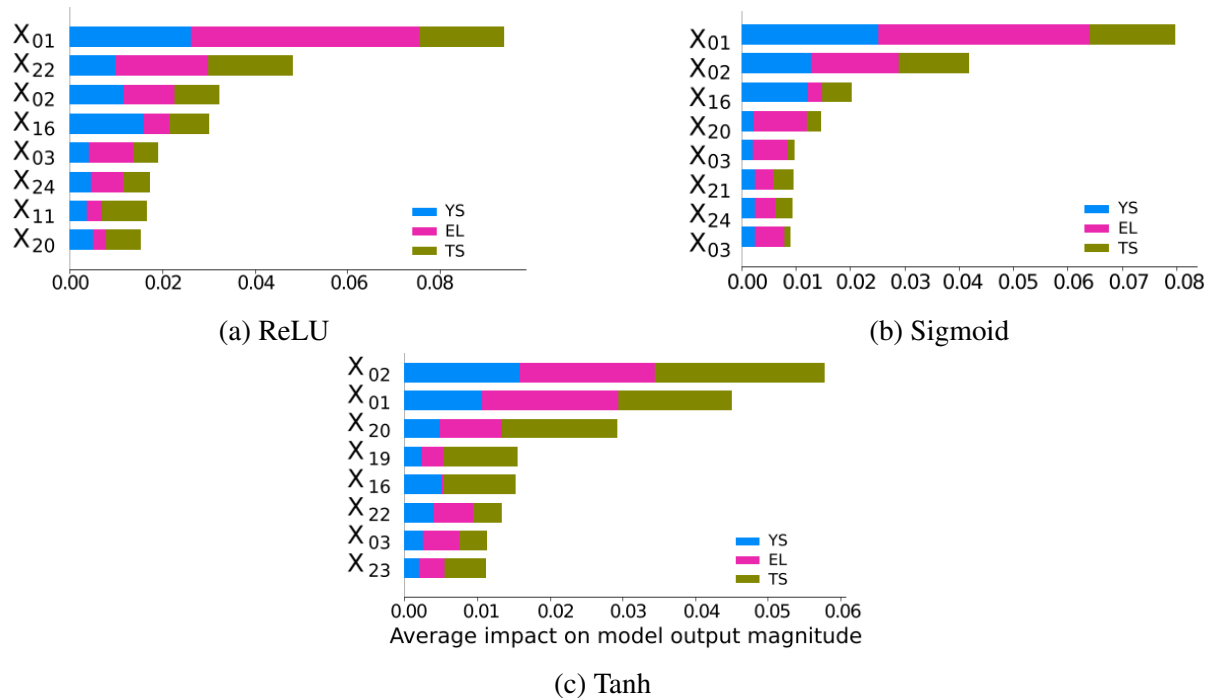


Figure 25 – SHAP feature importance ranking for ANNs with different activation functions.

It can be seen in Figure 25 that Carbon and Manganese are the parameters with the highest influence in magnitude over the output for all the selected ANNs, regardless of their architectures.

Since the ReLU based ANN showed superior performance in terms of R^2 , has comparable results in relation to the other two models and is computationally cheaper due to its piecewise linear activation function, it was decided to discard the Sigmoid and Tanh ANNs and use the ReLU ANN as the selected model.

5.3.1.2 INFERENCE OF NOVEL EXPERIMENTAL BH STEEL

In order to perform an experimental validation for the prediction model, a Design of Experiment (DOE) was conducted to predict the outcome obtained by changing some of the input parameters from three different steels. The aim was to investigate which steel and which parameters could be more appropriate to manufacture in order to achieve the desired mechanical properties. The selected variables and the values they could admit are illustrated in Table 8.

Table 8 – DOE parameters setup.

Input parameter	Possible values
X_{21} (Line speed)	60, 80, 100, 120 [m/s]
X_{22} (Mean Skin Pass Mill force)	0.4, 0.8, 1.0, 1.2, 1.6, 2.0 [%]
X_{23} and X_{24} (P3, P4 temperature)	750, 770, 790, 810, 830 [°C]

The variables X_{23} and X_{24} are changed simultaneously. Thus, since there are three combinations of variable input parameters, being X_{21} , X_{22} and X_{23} plus X_{24} with four, five and six possible values respectively, there are one hundred and twenty possible steel configurations. Given that there are three different steel samples, each with its specific attributes, there are three hundred and sixty possible configurations in total.

Although it could be time-consuming for a human to analyze each hypothetical steel from the DOE individually, for a computer model such as an ANN, this can be done much faster while providing custom data visualization. For instance, Figure 26 illustrates the effects that the temperature at pyrometer P3 plus P4 (X_{23} and X_{24}) and the line speed (X_{21}), from the HDG, have on the mechanical properties. It can be seen that the increase in temperature also increases the strength but decreases elongation, while a higher line speed (X_{21}) and mean skin pass mill force (X_{22}) can increase all properties. Under the assumption that the model is accurate, this can aid the specialist in making a decision on the parameter selection.

After particular steel was selected for manufacturing, the final product was obtained and its mechanical properties were measured. In Table 9 the results indicate that, for a certain steel manufactured in different P3 and P4 temperatures in the HDG, the ANN achieved very good results at predicting the steel YS and TS and reasonably good accuracy at predicting EL. Given that the model was not trained with the information from this steel, the results can be considered adequately accurate.

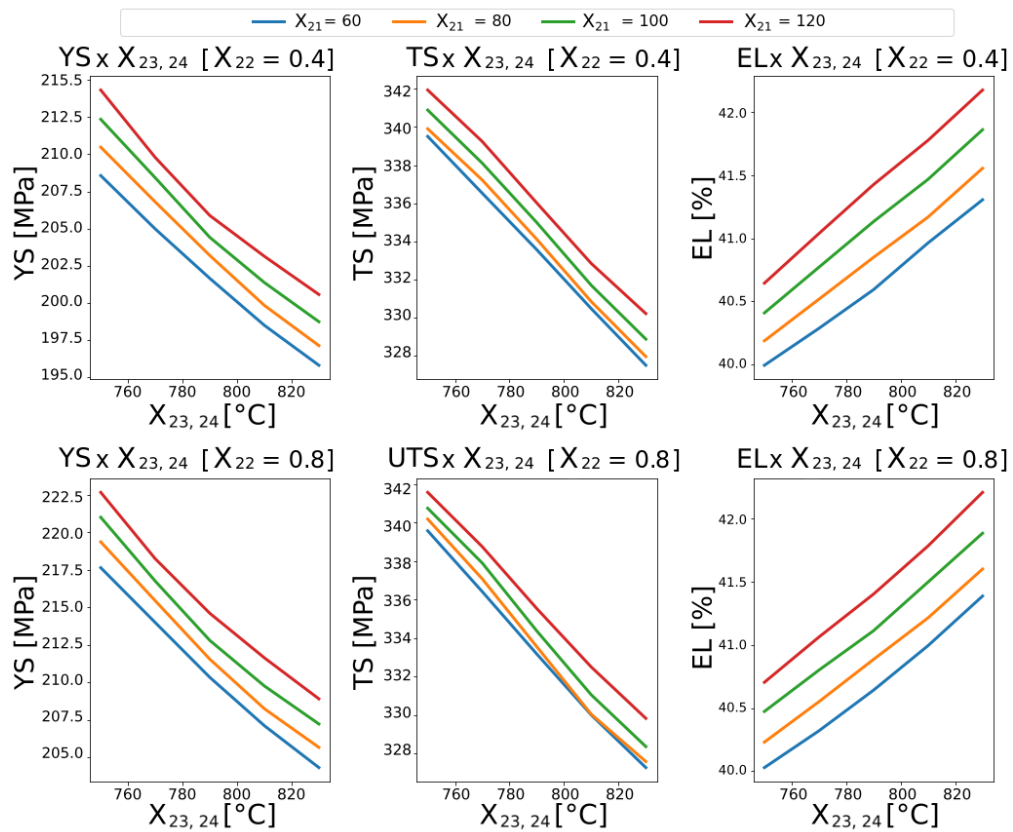


Figure 26 – Sensitivity analysis derived from the DOE results. Source: Own authorship.

Table 9 – Relative error from experimental validation.

	YS [%]	TS [%]	EL [%]	Temp. [°C]
	1.40	2	10.50	750
	0.80	1.60	10.30	790
	1.70	4.30	12.80	830
Accuracy rate	98.7	97.4	88.8	

5.3.2 CASE STUDY 2: STEEL OPTIMIZATION EXPERIMENTS

Regarding steel optimization, this case study is focused on optimizing random steels from the dataset in order to verify the algorithm's capabilities of creating steels with different mechanical properties

This is done by selecting a random steel from the dataset, hereby named reference steel and using its processing parameters as fixed parameters, mechanical properties as optimization targets and allowing the NSGA-II to generate new chemical compositions for the algorithm's population. Thus, the processing parameters are used as limitations in order for new chemistry to be created that satisfies a certain mechanical property requirement.

Since there are different steel grades, 100 random steels were sampled from each stratified cluster of grades, as was done in 5.1.2.1. The average and standard deviation of the optimization metrics - HV, $D_N^P W$ and IGD, are plotted in Figure 27. The optimization ran for 100 generations

using 100 individuals and a model trained on all steel grades.

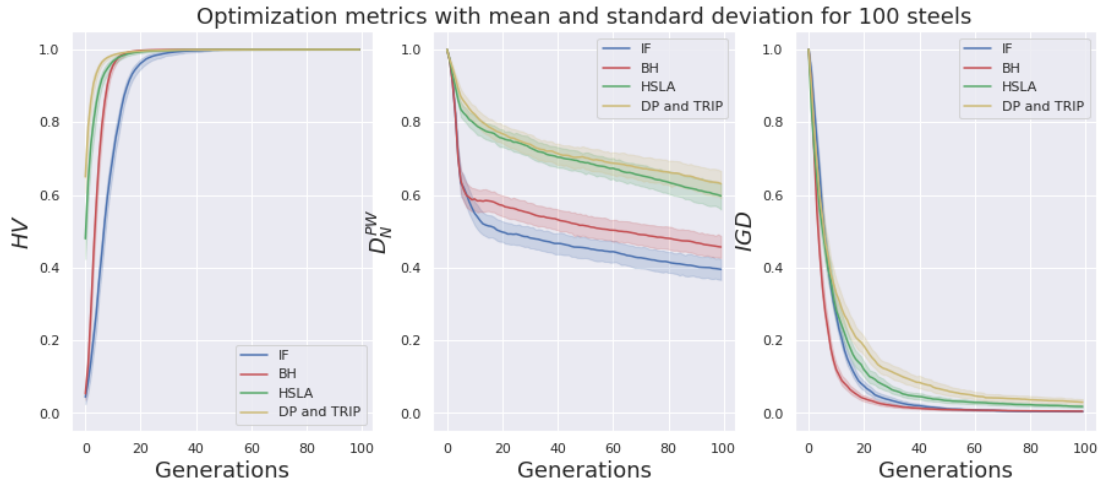


Figure 27 – Average and standard deviation for each steel grade optimization metrics, namely HV, D_N^{PW} and IGD, for random steels. Source: Own authorship.

Besides these optimization metrics, another aspect that can be valuable to analyze is the capability of the algorithm to create feasible steels, that is, steels that are possible to manufacture and which mechanical predictions are reliable. One way to possibly measure how realistic the optimized steel's results are, is to compare it with the dataset, which can be seen as a source of truth.

Thus, for each optimization experiment, the chemical composition's pairwise distance between each optimized steel and the reference steel was calculated. This calculation's output is a real number that indicates the distance between an optimized steel and the reference steel, or, how closely the newfound chemistry resembles a real, industry-validated one. The parameters are normalized in order to avoid bias due to different scales. An example is shown in Figure 28.

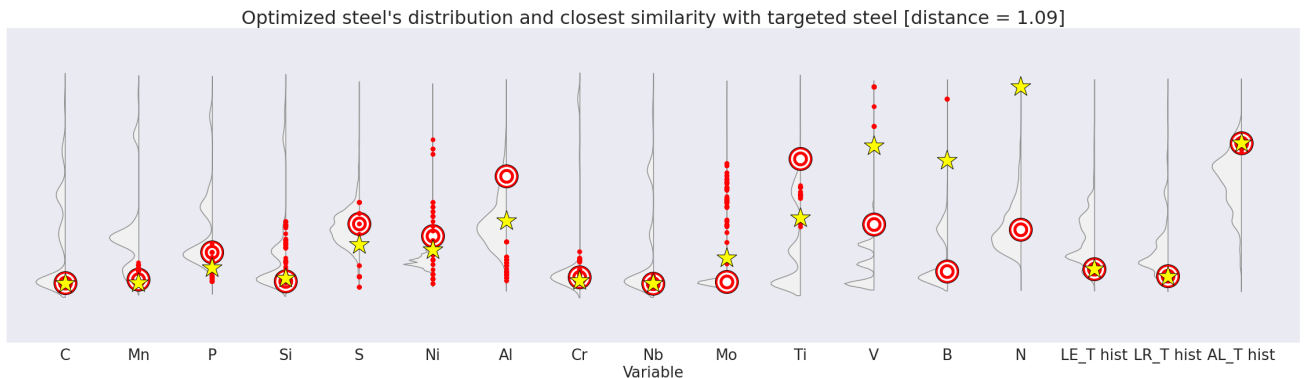


Figure 28 – Example of an optimized population's genotypic and phenotypic distribution. Target symbol, star symbol and red dots represents the reference steel, its most similar optimized steel and the remaining population, respectively. Source: Own authorship.

In Figure 28, a steel with fixed processing parameters and free chemical composition is optimized towards a YS, TS and EL of 183 MPa, 294 MPa and 45.6%, respectively. The

reference steel is illustrated by the target symbol, although its chemical composition does not influence the algorithm in this study case. The star symbol is the closest optimized steel in relation to the reference steel, while the red dots are the remaining individuals.

To create a feasible steel in this context, the goal would be to pinpoint all star symbols inside or near the target symbols. It can be seen that, while not completely on top of the reference chemistry, there is a reasonably high similarity between the reference and closest optimized steel. Besides, since some chemical elements have low influences on the output, such as boron e vanadium, they tend to have a higher variance compared to more influential elements, as shown in the SHAP ranking in 5.1.2.1.

Figure 29 illustrates the Pareto front from one of the random steel optimizations. In Figure 32, dominated solution candidates are represented as gray points and non-dominated as colored points. The color represents a third dimension, corresponding with the associated error from another mechanical property.

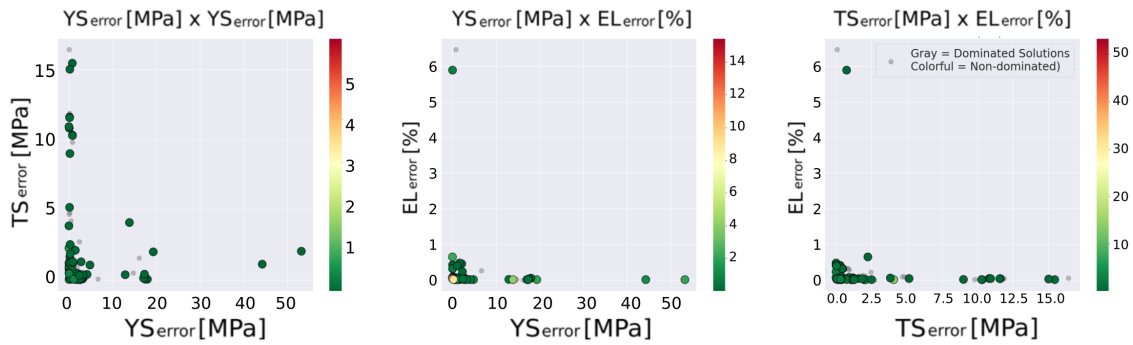


Figure 29 – Pareto front from a random IF steel optimization. Source: Own authorship.

Since the axis of the plots in Figure 29 corresponds with the population's prediction error in relation to the optimization objectives, the origin of the plots represents the point where the prediction error is null, or in this case, the ideal point. Since it can be seen from Figure 29 that the individuals converged towards or near the origin point, the convergence of this Pareto set corroborates with the results presented in the mechanical properties columns from Figure 28 and the optimization metrics in Figure 27.

Furthermore, the small distance values demonstrated in Figure 28, combined with the high convergence in Figure 27, shows that the algorithm is capable of delivering feasible steels under normal conditions, that is, when the optimized parameters are within or near well-known regions of the dataset.

5.3.3 CASE STUDY 3: STEEL DISCOVERY EXPERIMENTS

Moreover, it is important to consider that one of the main objectives of the present work is to generate novel steels, or at least, new insights about them. Hence, this case study is similar to the previous case study, but instead of sampling random steels, the reference steel is the one that presents the greatest prediction error from the database.

The objective is to verify if the algorithm can propose optimized steels with a smaller prediction error than the reference steel, but also to verify the diversity contained in the solution candidates, given that the newfound chemical composition may not necessarily be close to the reference steel's chemistry, due to the high prediction errors associated with it.

Similarly to the previous case study, this experiment was conducted for each stratified cluster of steel grades 100 times, using 100 individuals, 100 generations and a model trained on all steel grades. The results are shown in Figure 30 and Table 10.

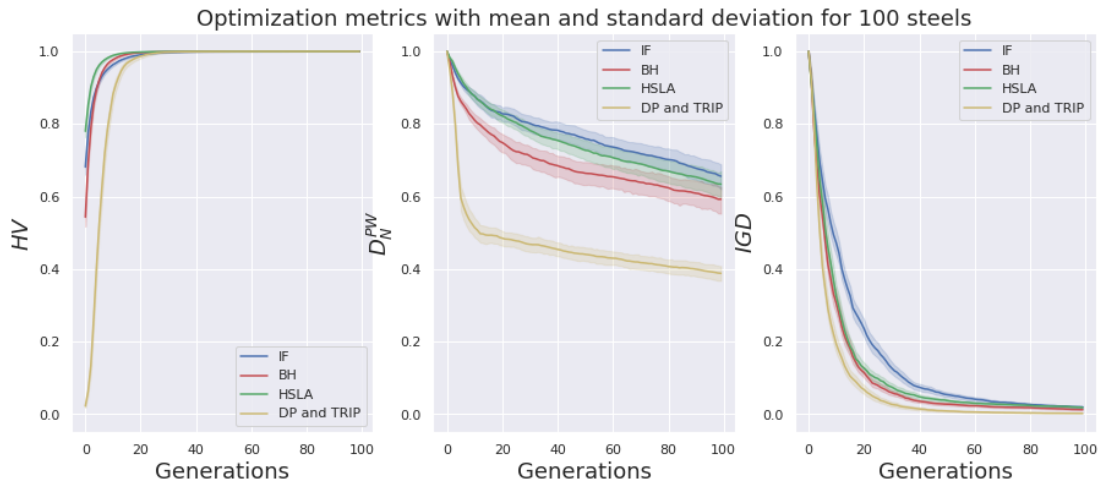


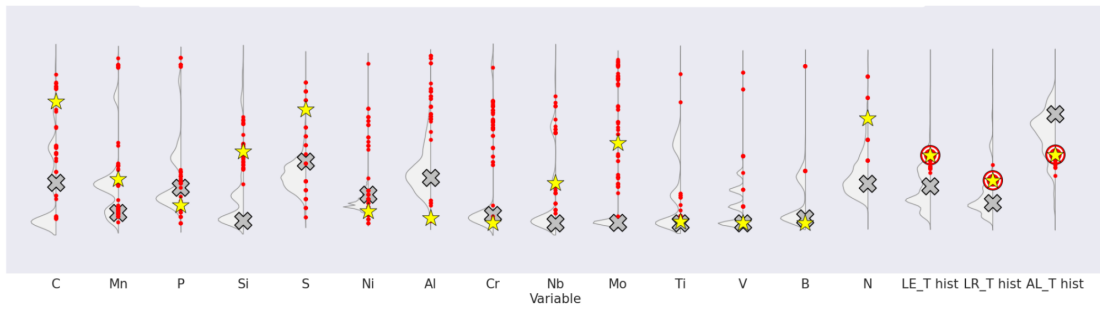
Figure 30 – Average and standard deviation for each steel grade optimization metrics, namely HV, D_N^{PW} and IGD, for steels with the largest associated prediction error. Source: Own authorship.

Table 10 – Comparison between the real mechanical properties of the steel with the largest associated prediction error, its prediction values and the averaged predictions from the optimized populations.

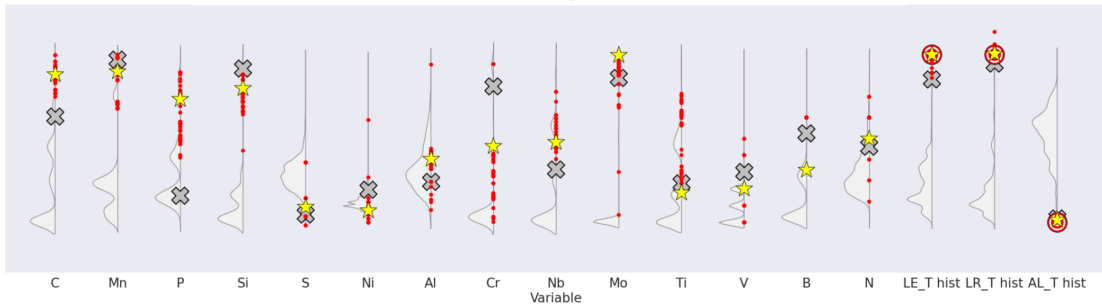
	Real mechanical properties			Reference steel's prediction			Optimized population's prediction		
Steel grade	$YS[MPa]$	$TS[MPa]$	$EL[\%]$	$YS[MPa]$	$TS[MPa]$	$EL[\%]$	$YS[MPa]$	$TS[MPa]$	$EL[\%]$
IF	446	490	29.3	300	368	42.2	436	492	28.6
BH	355	406	32.5	287	412	33.3	349	410	32.3
HSLA	535	557	22.6	480	546	22.7	528	565	22.4
DP and TRIP	914	1166	7.5	798	1116	8.9	912	1169	8.0

From the final metrics seen in Figure 30, the algorithm was able to optimize steels towards the desired mechanical properties. Table 10 presents the steel reference's prediction error, along with the average from the 100 optimizations for each cluster of steel grades. From Table 10, it can be clearly seen that the algorithm can indeed create more assertive steels in relation to the reference steel.

Some insights from the optimization successes can be found by analyzing the chemical distribution of the optimized steel. Two examples are shown, in Figure 31.



(a) Optimized population's genotypic and phenotypic distribution for IF steel.



(b) Optimized population's genotypic and phenotypic distribution for DP and TRIP steel.

Figure 31 – Examples of an optimized population's genotypic and phenotypic distribution for different steel. Target symbol, star symbol and red dots represents the reference steel, its most similar optimized steel and the remaining population, respectively.

Source: Own authorship. Source: Own authorship.

In Figure 31, the star symbol represents the closest steel to the reference steel. The optimization target, that is, the real mechanical properties found in the dataset is denoted by the target symbol. But since the reference steel is by design a bad reference in this case study, it is represented by a gray X letter.

Compared to Figure 28, in Figure 31 the optimized steels' genotypes are much farther away from the reference steel. This may be due to the high prediction errors that this reference steel genotype presumably generates. Besides, the optimized population's mechanical properties are closer to the targeted mechanical properties than the reference steel, corroborating to the metrics displayed in Figure 30.

Moreover, the results display a high diversity of solutions. For instance, according to Table 10, the IF steel with the highest prediction error is a very resistant steel. In Figure 31 (a), which is the result of one of the optimizations regarding the IF steel, a majority of the optimized individuals present a high amount of carbon. However, IF steels are known to have low quantities of carbon due to the metallurgical phenomena intrinsic to their definition. Nonetheless, some of the optimized individuals also present small quantities of carbon, presenting different alternatives in the final population. The same happens for many other chemical elements in the optimized genotype.

Parameter extrapolation can also occur regarding AHSS steel optimization, such as DP and TRIP. In Figure 31 (b), which represents a DP and TRIP optimization experiment, the

optimized steels have higher phosphorus and titanium contents than normal, more akin to BH and IF steels respectively. Analogous extrapolations occur in other parameters.

Finally, an overview of the mechanical properties distribution can be seen in Figure 32, where a Pareto front from a single optimization is presented in (a) and the combination of all optimizations regarding DP and TRIP steels is presented in (b) as a Pareto front. Similar results are obtained from other steel grades.

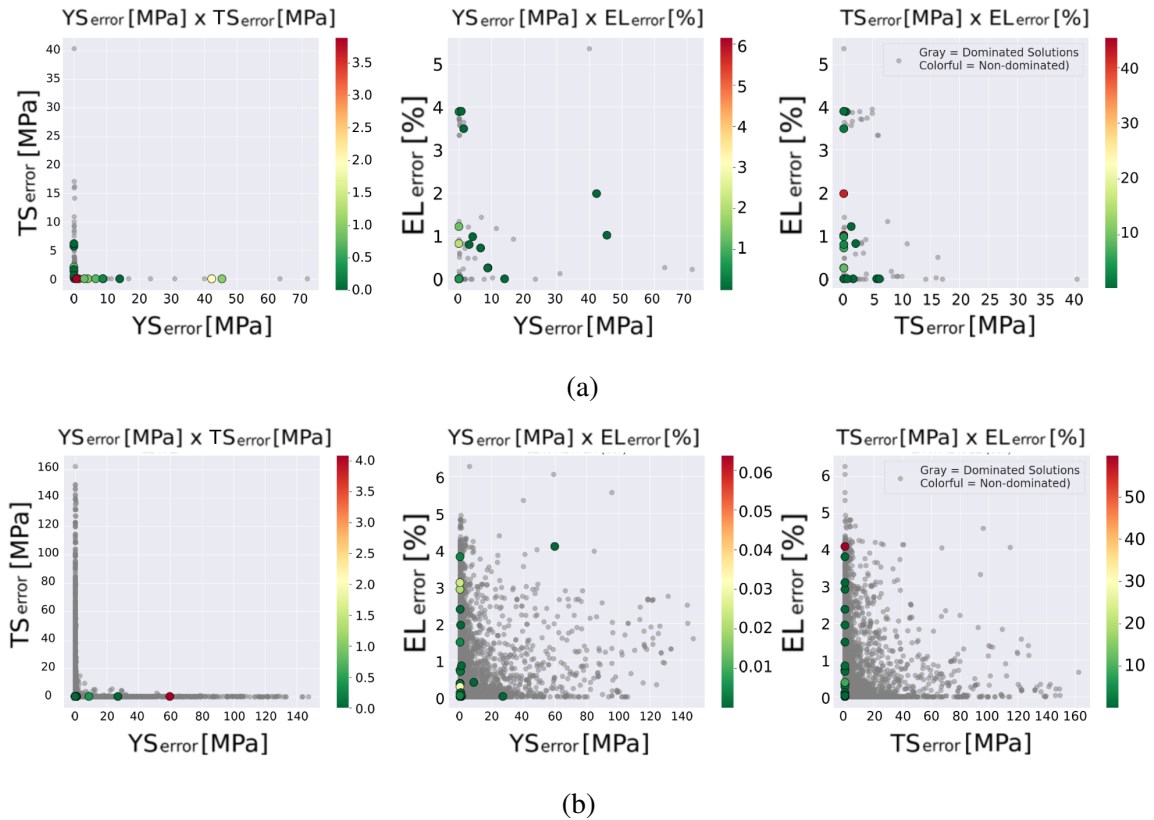


Figure 32 – Pareto front from (a) a single optimization and (b) combined populations generated from all optimizations regarding a DP and TRIP steel with the largest associated prediction error. Source: Own authorship.

Although the Pareto front distribution in Figure 32 is not explicitly indicative of a multi-objective algorithm's output with conflicting objectives, it is still a valid result due to the capability of the NSGA-II in satisfying the optimization requirements, namely the reduction of the associated prediction error towards a certain specified set of mechanical properties. In fact, the combined solution candidates can still be ordered and ranked by using a Non-dominated Sorting algorithm, thus providing the dominated and non-dominated classifications presented in Figure 32. Also, some of the optimizations were so successful in minimizing the prediction error that most of the solutions in Figure 32 (b) are dominated, that is, colored in gray.

Nonetheless, when the optimization is successful in creating the steel with the desired properties, the solutions converge towards a singular point while still taking into account the conflicting objectives. This does not jeopardize, however, the genotypic diversity, since different

chemical composition and process parameter combinations can still lead to equal mechanical properties.

In conclusion, this study case illustrates that by combining the data intervals of different steel grades and using a model trained upon this unified dataset, it is possible to extrapolate and combine different aspects of each steel grade in order to create novel steels. However, it can be argued that this diminishes the feasibility of the algorithm to generate realistic steels.

Thus, based on the previous and the present case studies, it is demonstrated that the algorithm has the capabilities to recreate what is known and to innovate into the unknown, depending on the manner in which the algorithm is used and how the limitations are imposed by the user.

5.3.4 CASE STUDY 4: STRESSING THE OPTIMIZATION ALGORITHM

In the previous two case studies, the focus was on generating steels that closely approaches a set of defined mechanical properties. In the first optimization study case, by randomly selecting random steels and analyzing how distant the population is from the reference steel, the algorithm's capability to generate feasible steels can be analyzed. In the second optimization study case, by choosing the steel with the largest associated prediction error, it is possible to investigate how the NSGA-II is able to take advantage of an extended data interval in order to extrapolate different parameters beyond what is commonly used in each steel grade individually.

In both case studies however, the optimizations are successful, which means that the Pareto solutions converge towards an ideal point where the prediction error is minimized or even nullified. Although this can be thought of as a desirable result in practice, in order to justify the choice of using a multi-objective optimization algorithm recommended for dealing with conflicting objectives such as NSGA-II, the algorithm needs to be tested at its limits.

In order to do this, all three objectives will be maximized. This can be done by defining all three desired objectives, namely YS , TS and EL , to an arbitrarily large number. Consequently, this should drive the algorithm to search for steels with increasingly large strength and ductility simultaneously, until some limit is found.

The processing parameters chosen for this are the steels which present the largest associated prediction error, that is, the same steels from Case Study 3 and the chemical compositions are free to be optimized. The desired mechanical properties are set as the largest properties from the dataset. The optimization ran for 200 generations with a population of 100 individuals. The optimization metrics and Pareto front are presented in Figure 33 and Figure 34, respectively.

From Figure 33, it can be clearly seen that, from the HV and IGD metrics, the optimization is far from reaching an optimal convergence state. Although the optimization target directly affects the metrics values, the chosen mechanical properties used as targets are still obtained from the dataset, so their values are not unreasonable. However, since there is a negative correlation between the strength and the ductility of a steel, it is in fact unreasonable to expect a steel that satisfies all these optimization requirements simultaneously, justifying the low HV and IGD

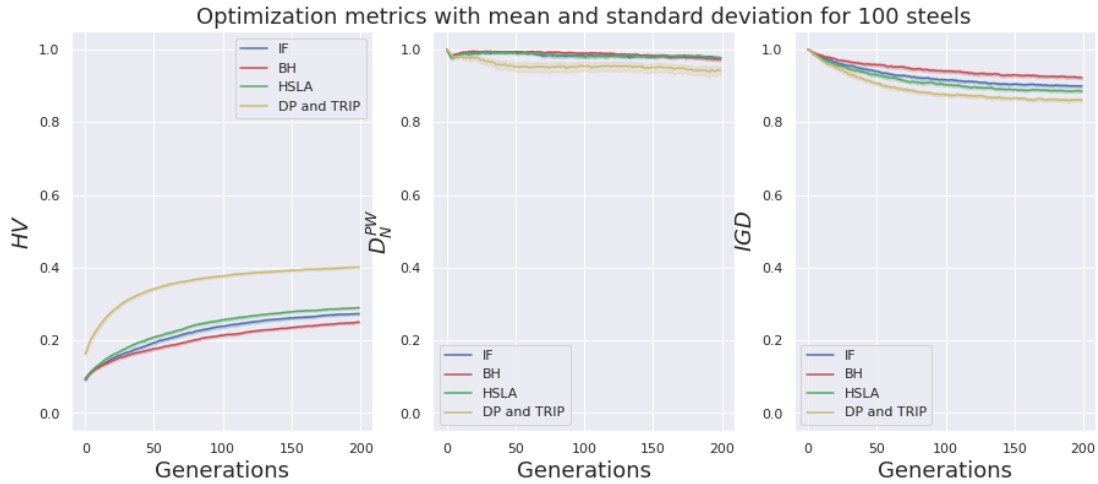


Figure 33 – Average and standard deviation for each steel grade optimization metrics, namely HV, D_N^{PW} and IGD, for a hypothetical and utopic steel. Source: Own authorship.

values. Nonetheless, the diversity metric shows a very high value, meaning that the euclidean distance between the steels are still large, indicating that convergence towards an ideal point hasn't occurred.

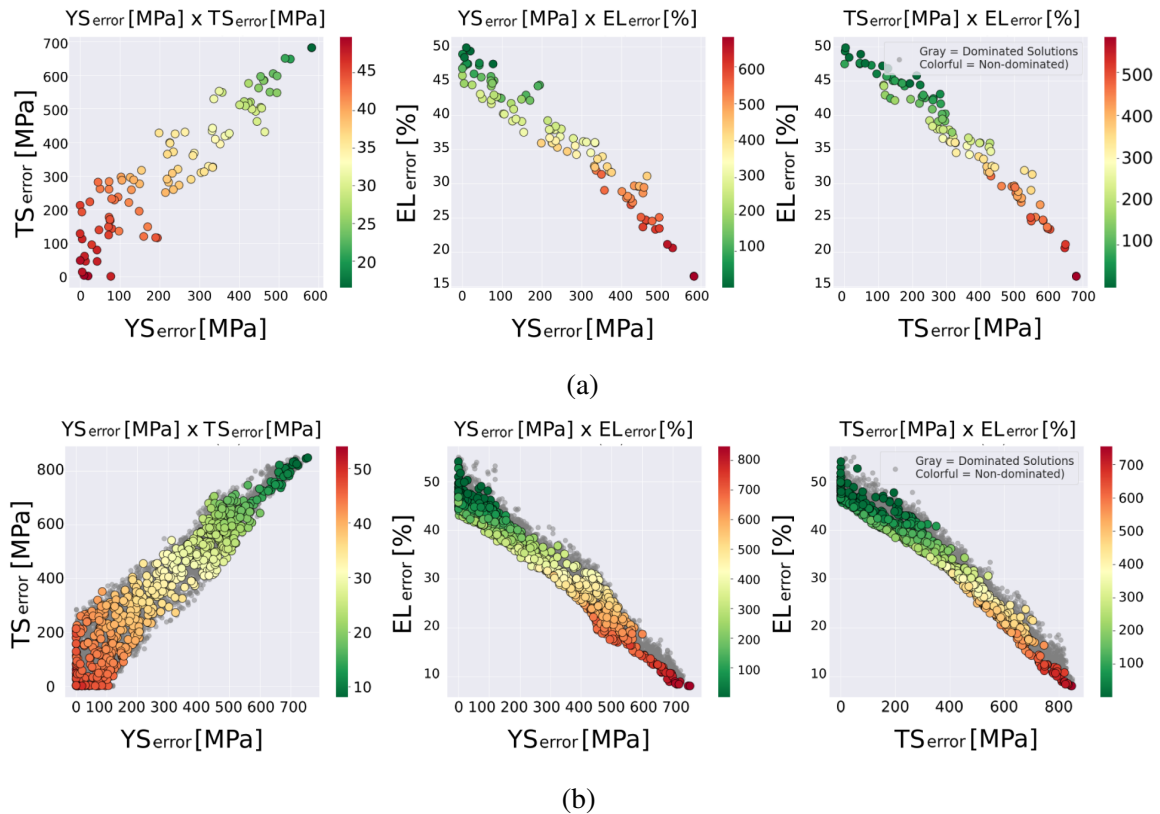


Figure 34 – Pareto front from (a) a single optimization and (b) combined populations generated from all optimizations regarding an IF steel with utopic mechanical property objectives. Source: Own authorship.

From Figure 34, the reasons behind the metrics values and the conflicting objectives are more clearly exemplified. Figure 34 (a) illustrates the Pareto front from a single optimization,

while (b) presents the Pareto front from all the populations combined by using the processing parameters for the IF steel specifically, although the results for the other steel grades are very similar. By comparing (a) and (b) it can be seen that the optimization results in terms of mechanical properties are similar between different experiments and that although the number of dominated points are smaller than in the previous case studies, they still exist.

Still, the non-dominated points clearly illustrate the relationship between different mechanical properties. From the first column of scatter plots - $YS \times TS$, the data points closely resemble an identity function, showing that as the YS error increases, so does the TS associated error. Moreover, as the strength increases, the data point's color goes from a deep green to a deep red, representing an increase in the error associated with the strength's ductility. Thus, as the strength increases in unison, the ductility decreases.

Similar conclusions can be obtained by analyzing the second and third column plots from Figure 34, the $YS \times EL$ and $TS \times EL$ plots. As the data points go from right to left in relation to the horizontal axis, the associated error in both strength decreases, represented from the axis values and the data points' color transition from red to green. Furthermore, by analyzing the interaction between both axes, it is clearly shown that, as one property's error decreases, the other increases, or rather, as the strength increases, the ductility decreases.

Although this is well-known in the metallurgy field, being even illustrated in Figure 2, from this case study it can be seen that this behavior also appears during optimization and, consequently, is a pattern learned from the data modeling stage.

5.3.5 CASE STUDY 5: ACCELERATED STEEL DEVELOPMENT

In the previous case study, although different processing parameters were experimented with the prediction model, the chemical composition was already defined beforehand. In this case study, an end to end application of the present models and algorithms is presented according to the steps illustrated in Figure 35.

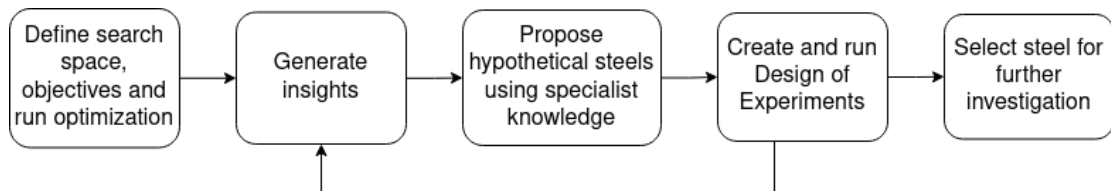


Figure 35 – Flowchart for a bottoms-up approach of steel development integrating a prediction model, optimization algorithm and a specialist's knowledge. Source: Own authorship.

Figure 35 describes a scenario that aims to create a steel with a desired set of mechanical properties which satisfies imposed engineering requirements following a bottoms-up approach. First, the search space and objectives are defined and fed to the optimization algorithm. Then, if the optimization is successful, the resultant data is analyzed by a specialist for new insights to

be generated on how to combine different design parameters. Using this as an initial stepping stone, along with the specialist's experience and knowledge, hypothetical steels can be created to test different ideas. Using these steels, Design of Experiments are performed in order to obtain more information regarding possible parameter variations and their effects. If necessary, the hypothetical steels are adjusted and new Design of Experiments are performed, until a final hypothetical steel is selected as a prototype or for further investigations.

5.3.5.1 STEEL OPTIMIZATION

In the first stage, the search space of the optimization algorithm is defined. This is done by selecting which parameters will be kept fixed and which will be optimized, along with the definition of a set of mechanical properties as an optimization target.

The search space is defined in Table 11, which contains the upper and lower bounds permissible for the optimized steels. Parameters not present in Table 11 were kept constant and not included in the optimization. For completeness, the optimization targets are also included in the table.

Table 11 – Lower and upper bounds for the optimization's search space. Source: Own authorship.

Parameter	Lower Bound	Upper Bound	Parameter	Lower Bound	Upper Bound	Parameter	Lower Bound	Upper Bound
X_{20} [mm]	0.6	1.2	X_{02} [ppm]	0.05	1	X_{09} [ppm]	0.01	0.1
X_{19} [%]	75	85	X_{03} [ppm]	0.005	0.04	X_{10} [ppm]	0.005	0.01
X_{15} [mm]	1200	1700	X_{04} [ppm]	0.003	0.1	X_{11} [ppm]	0	0.1
X_{18} [°C]	850	950	X_{05} [ppm]	0.0005	0.005	X_{12} [ppm]	0	0.01
X_{16} [°C]	550	700	X_{06} [ppm]	0.003	0.01	X_{13} [ppm]	0	0.002
X_{17} [°C]	1150	1250	X_{07} [ppm]	0.02	0.1	X_{21} [m/s]	100	170
X_{01} [ppm]	0.001	0.05	X_{08} [ppm]	0.005	0.05	X_{22} [%]	0.4	2
Y_{01} [MPa]	220	280	Y_{02} [MPa]	320	380	Y_{03} [%]	43	Boundless

Since the present optimization algorithm shown in this work is not properly suited for converging towards an interval of mechanical properties, five separate optimizations were executed using different mechanical properties as targets, presented in Table 12. Although EL's lower bound was set as 43% without any upper limit, the target EL for optimization was set at 50% in order to allow margin for error.

The algorithm ran for 200 generations using 2000 individuals. Thus, by combining the solutions from all different optimizations experiments, the final population contained 10000 solution candidates.

The resulting population of all experiments described in Table 12 were joined. This was done in order to optimize the mechanical properties and genotype towards a diversified range of values, as stated by Table 11 requirements. Results are illustrated graphically in Figure 31.

Table 12 – Description of the different optimization targets. Source: Own authorship.

	YS target [MPa]	TS target [MPa]	EL target [%]
Experiment 1	220	320	50
Experiment 2	220	380	50
Experiment 3	280	320	50
Experiment 4	280	380	50
Experiment 5	250	350	50

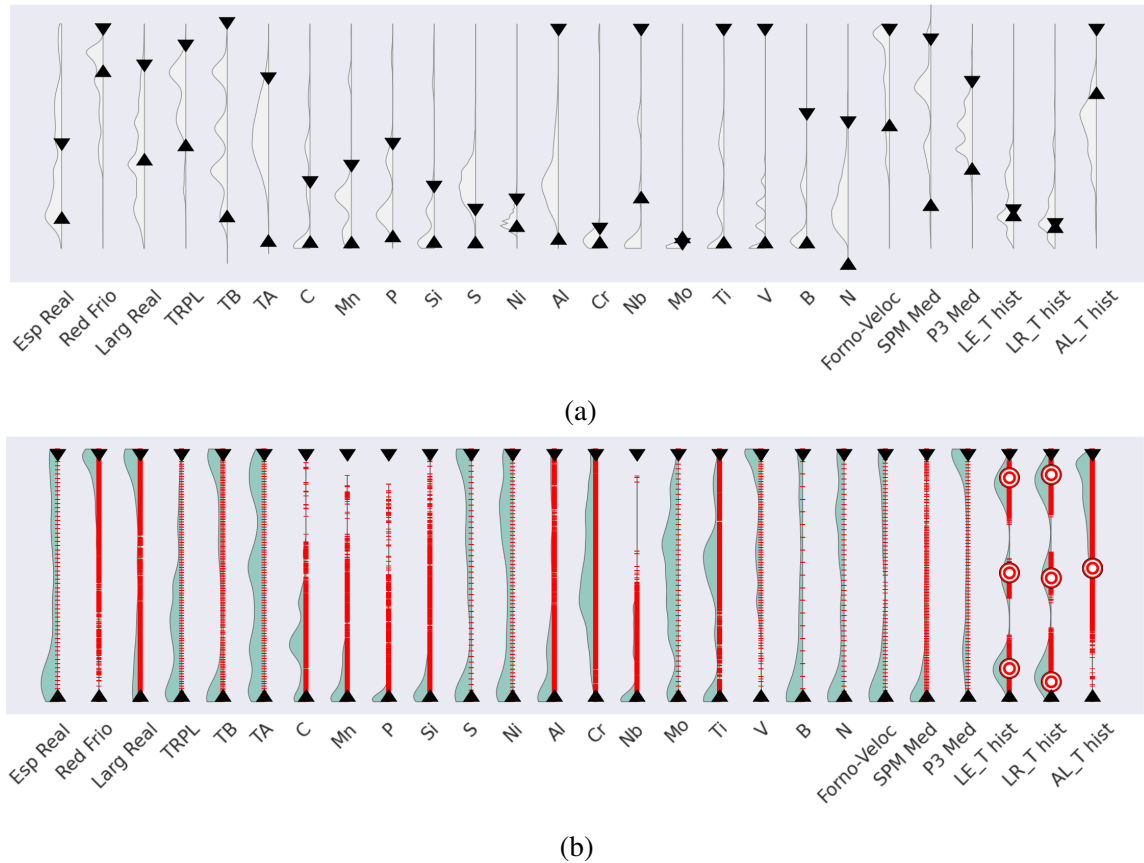


Figure 36 – Visual representation of the search scope's upper and lower bounds in relation to (a) the dataset histogram and (b) the optimized populations genotypic and phenotypic histogram. Small red lines are discrete values from the histogram. Source: Own authorship.

Figure 31 presents the search scope and obtained results in conjunction. In Figure 31 (a), the white histogram represents the dataset histogram. For Figure (b), the green histogram corresponds to the optimized population's histogram, that is, the density of the obtained solution candidates. To complement this visualization, the population is also represented as discrete red strokes. In both (a) and (b), the arrows pointing upward and downward represent the lower and upper bounds described in Table 11, respectively, albeit in different scales for better visualization. Finally, the target symbols correspond to the desired mechanical properties.

As can be seen in Figure 31 (a), some parameters have a greater allowance of variance, such as aluminum, niobium and titanium, among others. However, some other parameters are

kept within a very reduced search space, such as chrome and molybdenum. The delimitation of this search space was defined by a specialist in metallurgy in order to balance different project and engineering requirements.

Still, from Figure 31 (b) it can be seen that, the results from all five experiments described on Table 12, when aggregated, presents a comprehensive coverage of the search space. Besides, despite the algorithm's tendencies to converge towards a single mechanical property, as illustrated from the clear gaps in YS and TS Figure 31 (b), when combined, a majority of the interval is covered.

5.3.5.2 SPECIALIST KNOWLEDGE AND DESIGN OF EXPERIMENTS

Due to the sheer size of the obtained data, it was critical to employ the knowledge of a metallurgy specialist in order to filter, sort, organize and scrutinize the data, making sense of it and how it relates to the metallurgical perspective.

Thus, based on the 10000 candidate solutions, a set of hypothetical chemical compositions were created from the ground up by the specialist. Then, to analyze how different processing parameters influence the mechanical properties, a Design of Experiment was built, with the varying parameters described in Table 13.

Table 13 – DOE parameters setup for advanced steel development case study.

Input parameter	Possible values
X_{21} (Line speed)	100, 120, 140, 160 [m/s]
X_{16} (Coiling temperature)	700, 660, 620, 550 [°C]
X_{22} (Mean Skin Pass Mill force)	0.2, 0.4, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0 [%]
X_{23} and X_{24} (P3, P4 temperature)	750, 770, 790, 810, 830 [°C]

As illustrated in Figure 35, this process was done in iterations, such that after each iteration, gradual improvements were proposed for the hypothetical chemical compositions, new Design of Experiments were conducted and so on. Following this approach, 28 different chemical compositions in total, described in Table 14, were created and had their mechanical properties experimented on.

5.3.5.3 RESULTS

Although the number of hypothetical steels generated through the DoE can be high, data analysis can be somewhat simplified with the aid of appropriate data visualizations and graphical analysis tools. In this context, each steel composition, along with its processing parameter's variations derived from the DoE, will be denominated as a steel experiment.

As an example of graphical analysis, Figure 37 and Figure 38 illustrate the mechanical properties box plot and scatter plot, respectively. In both figures, (a) and (b) correspond to different sets of steel experiments, derived from different iterations of the steel development

presented in this case study. Hashed red lines represent lower or upper bounds, while the green background denotes the inner interval of acceptable mechanical properties, as defined by Table 11.

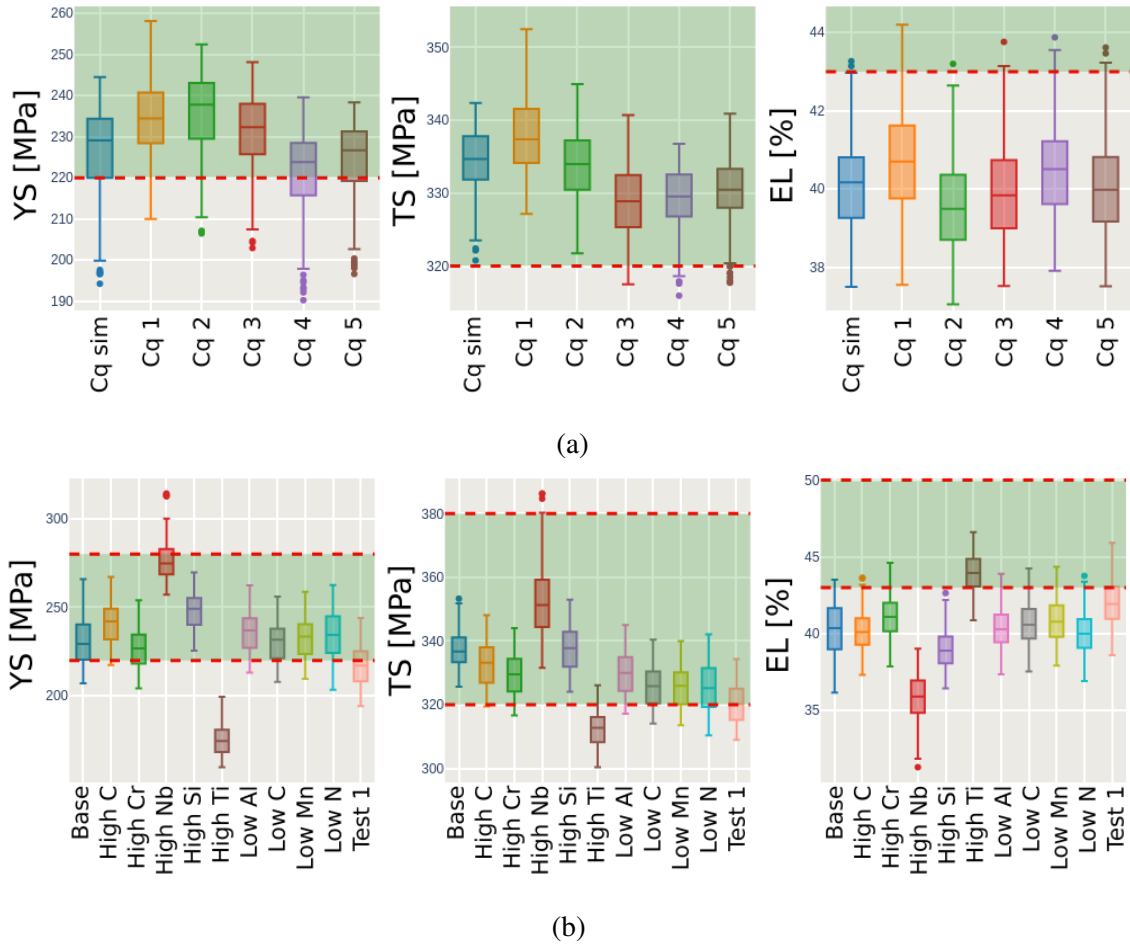


Figure 37 – Mechanical properties box plots for different batches of steel experiments. Source: Own authorship.

Regardless of the quantity of individual data points, in Figure 37 (a) it can be seen that most steel experiments have a majority of their hypothetical steels satisfying the YS and TS requirements, that is, most of the boxes are inside the green background. However, the opposite is true regarding EL , that is, only a minority of the steel experiments, including some outliers, presents an EL higher than 43%.

The same happens for Figure 37 (b), with some peculiarities. The EL box plot corresponding to *High Ti* experiment have the majority of its steels satisfying the ductility requirements, but the opposite happens regarding its strengths. In fact, all the *High Ti* have lower YS than the minimum required. Still, nearly a quarter of the *Test 1* experiment have more than 43% of EL , but since this graphical tool doesn't correlate the different mechanical properties, additional analysis is required.

Nonetheless, the box plot is powerful for its simplicity and its capabilities to accentuate the overall tendencies of each experiment. For example, Figure 37 (b) shows that, compared

to *Base*, the additional carbon presented on *High C* increases the *YS* but decreases *TS*. The addition of niobium on *High Nb* dramatically increases the overall strength of the steel, but also diminishes significantly its ductility. The opposite occurs regarding titanium, as shown in *High Ti* experiment. Finally, lowering aluminum, carbon, manganese and nitrogen leads to statistically similar results, as can be seen in *Low Al*, *Low C*, *Low Mn* and *Low N*.

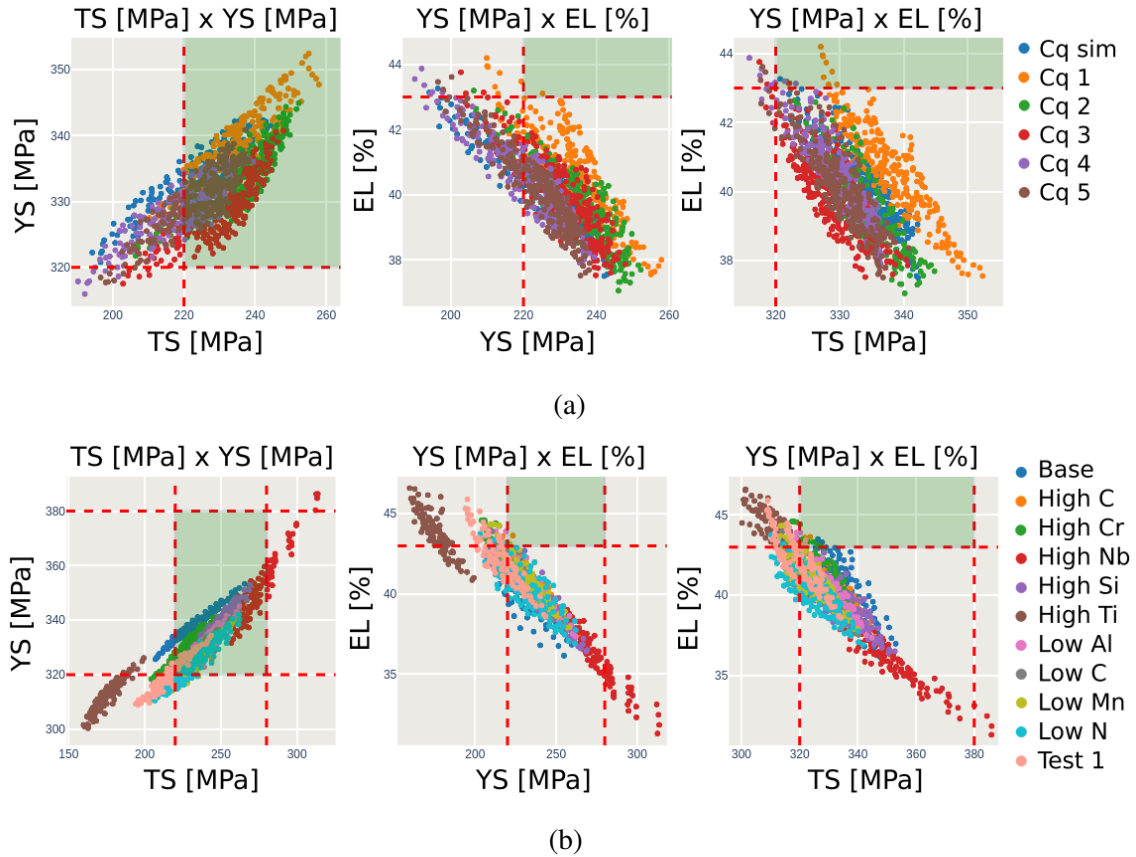


Figure 38 – Mechanical properties scatter plots for different batches of steel experiments.
Source: Own authorship.

By considering the relationship between the different mechanical properties, Figure 38 (a) and (b) illustrates that, while the relationship between *YS* and *TS* lends to a majority of the steel experiments to be within acceptable strength limits, with the exception of *High Ti*, this satisfiability does not occur regarding the strength and ductility relationship. Although this graphical analysis can become somewhat polluted due to the quantity of data, the plot is interactive and can, in practice, aid the steel development.

Finally, an overview of the achieved results from all steel experiments is summarized in Table 15, where the percentage of steels that satisfies the mechanical properties requirements is presented.

Although very few steels satisfy all the mechanical properties requirements, illustrating the challenge that this steel design proposes, some steels were deemed as potential solutions. Given that several different steel experiments were conducted and adjusted in a relatively short period of time, it is possible to apply the prediction model to accelerate steel development. Fur-

thermore, by using the optimization algorithm with a specialist's guidance, different hypothetical steels can be developed following a bottoms-up approach, making it possible for new venues of novel and hypothetical steel discovery to be explored.

Table 14 – List of all steel experiments proposed by specialist. All parameters in the table are measured as *ppm*. Source: Own authorship.

	C	Mn	P	Si	S	Ni	Al	Cr	Nb	Mo	Ti	V	B	N
UCB-Nb	0,0025	0,2000	0,0150	0,0150	0,0080	0,0100	0,1000	0,0300	0,0200	0,0050	0,0050	0,0050	0,0003	0,0004
UBC-NbTi	0,0025	0,3000	0,0100	0,0300	0,0080	0,0050	0,0300	0,0300	0,0300	0,0050	0,0200	0,0050	0,0005	0,0040
EBC-Nb1	0,0100	0,1000	0,0100	0,0500	0,0040	0,0090	0,0700	0,0450	0,0150	0,0100	0,0030	0,0100	0,0005	0,0030
EBC-Nb2	0,0100	0,4000	0,0100	0,0500	0,0040	0,0090	0,0700	0,0450	0,0150	0,0100	0,0030	0,0100	0,0005	0,0030
EBC-NbTi	0,0100	0,4000	0,0100	0,0300	0,0080	0,0900	0,0700	0,0400	0,0350	0,0100	0,0250	0,0100	0,0005	0,0030
Base	0,0025	0,2000	0,0150	0,0150	0,0080	0,0100	0,1000	0,0300	0,0200	0,0050	0,0050	0,0050	0,0003	0,0040
Low C	0,0010	0,2000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0300	0,0200	0,0050	0,0050	0,0050	0,0003	0,0040
High C	0,0040	0,2000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0300	0,0200	0,0050	0,0050	0,0050	0,0003	0,0040
Low Mn	0,0025	0,1000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0300	0,0200	0,0050	0,0050	0,0050	0,0003	0,0040
High Si	0,0025	0,2000	0,0150	0,0500	0,0080	0,0100	0,0300	0,0300	0,0200	0,0050	0,0050	0,0050	0,0003	0,0040
Low Al	0,0025	0,2000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0300	0,0200	0,0050	0,0050	0,0050	0,0003	0,0040
High Cr	0,0025	0,2000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0800	0,0200	0,0050	0,0050	0,0050	0,0003	0,0040
High Nb	0,0025	0,2000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0300	0,0400	0,0050	0,0050	0,0050	0,0003	0,0040
High Ti	0,0025	0,2000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0300	0,0200	0,0050	0,0200	0,0050	0,0003	0,0040
Low N	0,0025	0,2000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0300	0,0200	0,0050	0,0050	0,0050	0,0003	0,0025
Test 1	0,0010	0,1000	0,0150	0,0150	0,0080	0,0100	0,0300	0,0800	0,0200	0,0050	0,0050	0,0050	0,0003	0,0040
Test 1.1	0,0030	0,1000	0,0150	0,0150	0,0080	0,0010	0,0300	0,0800	0,0200	0,0050	0,0050	0,0050	0,0003	0,0025
TiNb Low N	0,0025	0,2000	0,0015	0,0150	0,0080	0,0010	0,0300	0,0800	0,0200	0,0050	0,0100	0,0050	0,0003	0,0025
TiNb2 Low N	0,0025	0,2000	0,0015	0,0150	0,0080	0,0010	0,0300	0,0800	0,0200	0,0050	0,0200	0,0050	0,0003	0,0025
High Cr 2	0,0025	0,2000	0,0015	0,0150	0,0080	0,0010	0,0300	0,0800	0,0200	0,0050	0,0050	0,0050	0,0003	0,0025
Cr + Ti	0,0025	0,2000	0,0015	0,0150	0,0080	0,0010	0,0300	0,1000	0,0200	0,0050	0,0100	0,0050	0,0003	0,0025
Medium Cr	0,0025	0,2000	0,0015	0,0150	0,0080	0,0010	0,0300	0,0500	0,0200	0,0050	0,0100	0,0050	0,0003	0,0025
Cq1	0,0025	0,2000	0,0180	0,0300	0,0080	0,0100	0,0500	0,0600	0,0200	0,0080	0,0050	0,0050	0,0003	0,0040
Cq2	0,0025	0,2000	0,0180	0,0300	0,0080	0,0100	0,0500	0,0800	0,0200	0,0080	0,0050	0,0050	0,0003	0,0025
Cq3	0,0025	0,1000	0,0180	0,0300	0,0080	0,0100	0,0500	0,0800	0,0200	0,0080	0,0050	0,0050	0,0003	0,0025
Cq4	0,0025	0,2000	0,0180	0,0300	0,0080	0,0100	0,0500	0,0800	0,0200	0,0080	0,0010	0,0050	0,0003	0,0025
Cq5	0,0025	0,1000	0,0180	0,0300	0,0080	0,0100	0,0500	0,0800	0,0250	0,0080	0,0010	0,0050	0,0003	0,0025
Cq sim	0,0025	0,3000	0,0180	0,0300	0,0080	0,0100	0,0500	0,0800	0,0200	0,0080	0,0010	0,0050	0,0003	0,0025

Table 15 – Overview from the results obtained from all steel experiments regarding their requirements satisfiability. Source: Own authorship.

Steel Experiment	% of steels that satisfies YS requirements	% of steels that satisfies TS requirements	% of steels that satisfies EL requirements	% of steels that satisfies all requirements
EBC-Nb1	37	99	0	0
EBC-Nb2	20	89	0	0
EBC-NbTi	0	100	0	0
UBC-Nb	95	100	0.5	0
UBC-NbTi	41	99	4.5	0
Base	75	100	2.8	0
Low C	82	77	6.5	0
High C	99	98	3.7	1.9
Low Mn	84	81	9.2	1.9
High Si	100	100	0	0
Low Al	91	91	3.7	0
High Cr	69	88	13	0.9
High Nb	62	96	0	0
High Ti	0	8.2	79	0
Low N	84	71	2.8	0
Test 1	38	50	25	0
Cr + Ti	0	42	21	0
High Cr 2	37	67	6.5	0
Medium Cr	8.3	44	15	0
Test 1.1	40	55	12	0
TiNb Low N	2.8	44	18	0
TiNb2 Low N	0	34	31	0
Cq1	91	100	3.0	0.3
Cq2	92	100	0.3	0
Cq3	87	98	1.1	0
Cq4	63	97	2.6	0
Cq5	74	97	1.5	0
Cq sim	75	100	0.7	0

6 CONCLUSIONS AND FUTURE RESEARCH DIRECTION

The identification of the relationships between structure and mechanical property is fundamental to discovering new materials based on required attributes, including steels. Nonetheless, the ability to comprehensively understand and design structure-property relationships of steels is very challenging due to the diversity of steel grades and the complexity of its microstructure. As a result, data-driven discovery and design of novel advanced steels, through surrogate-based optimization, requires the application of advanced techniques such as big data, data mining and machine learning to accelerate research and development. It is through the foundation provided by steel data and machine learning that this study is made possible, enabling the ongoing flourishing of a data-driven material discovery paradigm that integrates metallurgy domain knowledge and machine learning, expanding the research field of materials informatics.

In this work, a surrogate model was built using data-driven algorithms in order to predict the mechanical properties of steels. The capabilities of such models were presented in Case Study 1, where a model trained on IF and BH steels predicted with high accuracy the properties of a new, unseen BH steel. This model was generated using an AutoML training pipeline, consequently removing the need of manual hyperparameter tuning through modern machine learning techniques and frameworks.

Following this validation, more complex steel grades were added to the training stage. Although the model is a multi-variate regressor without nominal variables, that is, explicit classes of data, special care was taken into consideration with imbalanced steel representations. The consequences of combining different steels grades in the surrogate's modeling were analyzed through xAI and by reducing the dataset imbalances, it was possible to increase the importance of a diversified set of parameters towards the model's output. The quantitative performance of different models were measured and it was decided that a model capable of multi-variate regression trained on different steels grades simultaneously was preferred. Thus, this conclusion came not only from quantitative measures, but also from a further understanding of the black-box's inner working, creating ultimately an efficient but still relatively simple model.

Then, this model was coupled with an optimization algorithm, namely NSGA-II in order to search for novel steel compositions. Since one of the main objectives of this work is to be able to find different hypothetical steels, different case studies were built to consider how the optimization algorithm behaves. The case studies illustrates that a compromise between novelty and feasibility can be attained by properly defining the search scope. Case Study 1 illustrates that the algorithm is capable of optimizing already known steel compositions, while Case Study 2 presents how the optimization can extrapolate and combine the domains of chemical element from different steel grades, creating novel steels in the process. Finally, Case Study 3 illustrates the algorithm limitations and by doing so, the fundamental compromise between a steel strength and ductility.

Finally, in Case Study 5, a novel steel was proposed by following a methodology of

accelerated steel development. By combining the optimization's search for an initial steel composition with a specialist's knowledge, different new insights can be generated relating how different parameters can be combined in order to achieve the desired mechanical properties, while still respecting the engineering requirements. Then, by iterating on several design of experiments, different specific steels can be analyzed and adjusted by a specialist, ultimately accelerating the brainstorming and investigations stage and reducing the time-to-market of hypothetical and novel steels.

Still, there are many challenges and open questions to be answered in the development and understanding of steel manufacturing with its many microstructures and complex phases, how to appropriately model such complexities in a surrogate model, the technical and operational limitations involved in such process, but also the checks and bounds of how far current modeling and optimization algorithms can lead us. From the ongoing research trends in many material discovery projects, the unification of different bodies of knowledge across multiple fields seems to point towards new scientific and industrial possibilities.

The main objective of this work in progress is demanding, but just as much as exciting. The objective of developing novel steel designs in a shorter time frame while still incorporating many flavors of the complexities associated with metallurgy is already bearing fruits. Nonetheless, much more lies on the horizon. Future works can involve the use of better data processing techniques, including data augmentation, modeling new data such as microstructure, new state-of-the-art modeling techniques, a deeper investigation, explanation and interpretation of the prediction model by using new xAI tools and finally, revamping the optimization algorithm such that it maintains a more diversified population across its evolution, by using novel algorithms, incorporating niching methods, implementing adaptive controls or redefining the fitness functions.

BIBLIOGRAPHY

ABBASI, Erfan; RAINFORTH, W. Effect of nbmo addition on the precipitation behaviour of v microalloyed steel during intercritical annealing. v. 15, p. 1–11, 01 2019. Cited on page 67.

ADAMCZYK, J. Development of the microalloyed constructional steels. **Journal of Achievements in Materials and Manufacturing Engineering**, v. 14, n. 1, p. 9–20, 2006. Cited on page 28.

AL-JARRAH, Omar Y. et al. Efficient machine learning for big data: A review. **Big Data Research**, v. 2, n. 3, p. 87–93, 2015. ISSN 2214-5796. Big Data, Analytics, and High-Performance Computing. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2214579615000271>>. Cited on page 29.

ALETI, Aldeida; MOSER, I. Predictive parameter control. In: . [S.l.: s.n.], 2011. p. 561–568. Cited on page 62.

AUDET, Charles et al. Performance indicators in multiobjective optimization. **European Journal of Operational Research**, v. 292, 11 2020. Cited 5 times on pages 35, 36, 43, 44, and 71.

BAKER, Laura; DANIEL, S.; PARKER, Jonathan. Metallurgy and processing of ultralow carbon bake hardening steels. **Materials Science and Technology**, v. 18, p. 355–368, 04 2002. Cited on page 67.

BALUJA, Shumeet; CARUANA, Rich. **Removing the Genetics from the Standard Genetic Algorithm**. USA, 1995. Cited on page 40.

BAO, Chunteng et al. A novel non-dominated sorting algorithm for evolutionary multi-objective optimization. **Journal of Computational Science**, Elsevier B.V., v. 23, p. 31–43, 2017. ISSN 18777503. Disponível em: <<http://dx.doi.org/10.1016/j.jocs.2017.09.015>>. Cited 2 times on pages 39 and 40.

BERBENNI, S. et al. A micromechanical approach to model the bake hardening effect for low carbon steels. **Scripta Materialia**, v. 51, n. 4, p. 303–308, 2004. ISSN 13596462. Cited on page 27.

BHATT, A.; PARAPPAGOUDAR, M. B. Modeling and Analysis of Mechanical Properties in Structural Steel-DOE Approach. **Archives of Foundry Engineering**, v. 15, n. 4, p. 5–12, 2015. ISSN 22992944. Cited on page 50.

BINITHA, S.; SATHYA, S.S. A survey of bio inspired optimization algorithms. **International Journal of Soft Computing and Engineering**, v. 2, p. 137–151, 01 2012. Cited 2 times on pages 37 and 38.

BOYER, Howard E. **Practical Heat Treating**. [S.l.: s.n.], 1984. ISBN 0871701782. Cited on page 28.

BRANCO, Paula; TORGO, Luís; RIBEIRO, Rita. A survey of predictive modelling under imbalanced distributions. 05 2015. Cited on page 66.

BRIGHT, G. W. et al. Variability in the mechanical properties and processing conditions of a High Strength Low Alloy steel. **Procedia Engineering**, v. 10, n. 01639, p. 106–111, 2011. ISSN 18777058. Cited 2 times on pages 50 and 54.

BÄCK, Thomas; HOFFMEISTER, Frank; SCHWEFEL, Hans-Paul. A survey of evolution strategies. **Proceedings of the Fourth International Conference on Genetic Algorithms**, 10 1994. Cited on page 38.

CALLISTER, William D.; RETHWISCH, David G. **Materials science and engineering : an introduction / William D. Callister, Jr., David G. Rethwisch**. 8th ed.. ed. Hoboken, NJ: John Wiley & Sons, 2010. ISBN 9780470448649. Cited 2 times on pages 24 and 25.

CHAHAR, Vijay; KATOCH, Sourabh; CHAUHAN, Sumit. A review on genetic algorithm: Past, present, and future. **Multimedia Tools and Applications**, v. 80, 02 2021. Cited on page 39.

CHAUHAN, Karansingh et al. Automated machine learning: The new wave of machine learning. In: **2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)**. [S.l.: s.n.], 2020. p. 205–212. Cited on page 34.

CHEN, Yang et al. Integrated Modelling of Microstructure Evolution and Mechanical Properties Prediction for Q&P Hot Stamping Process of Ultra-High Strength Steel. **Chinese Journal of Mechanical Engineering (English Edition)**, Springer Singapore, v. 33, n. 1, 2020. ISSN 21928258. Disponível em: <<https://doi.org/10.1186/s10033-020-00461-3>>. Cited 3 times on pages 20, 50, and 51.

CHOI, Dami et al. On Empirical Comparisons of Optimizers for Deep Learning. n. 1, 2019. Disponível em: <<http://arxiv.org/abs/1910.05446>>. Cited on page 33.

CHOI, Dong-Jin; PARK, Heekyung. A hybrid artificial neural network as a software sensor for optimal control of a wastewater treatment process. **Water Research**, v. 35, n. 16, p. 3959–3967, 2001. ISSN 0043-1354. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0043135401001348>>. Cited on page 21.

CHOU, Ping Yi; TSAI, Jinn Tsong; CHOU, Jyh Horng. Modeling and Optimizing Tensile Strength and Yield Point on a Steel Bar Using an Artificial Neural Network with Taguchi Particle Swarm Optimizer. **IEEE Access**, IEEE, v. 4, p. 585–593, 2016. ISSN 21693536. Cited 2 times on pages 50 and 51.

CITROEN, Charles L. The role of information in strategic decision-making. **International Journal of Information Management**, v. 31, n. 6, p. 493–501, 2011. ISSN 0268-4012. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0268401211000235>>. Cited on page 29.

CORRIVEAU, Guillaume et al. Review and study of genotypic diversity measures for real-coded representations. **Evolutionary Computation, IEEE Transactions on**, v. 16, p. 695–710, 10 2012. Cited on page 46.

CORRIVEAU, Guillaume et al. Evaluation of genotypic diversity measurements exploited in real-coded representation. 06 2015. Cited on page 46.

COSTA, Patricia et al. Optimization of the continuous galvanizing heat treatment process in ultra-high strength dual phase steels using a multivariate model. **Metals**, v. 9, n. 6, p. 1–19, 2019. ISSN 20754701. Cited 6 times on pages 50, 51, 53, 54, 55, and 59.

DEB, Kalyanmoy; KALYANMOY, Deb. **Multi-Objective Optimization Using Evolutionary Algorithms**. USA: John Wiley & Sons, Inc., 2001. ISBN 047187339X. Cited 3 times on pages 34, 36, and 39.

DEB, Kalyanmoy et al. A fast and elitist multiobjective genetic algorithm: NSGA-II. **IEEE Transactions on Evolutionary Computation**, v. 6, n. 2, p. 182–197, 2002. ISSN 1089778X. Cited 2 times on pages 21 and 40.

DENG, Yong et al. Mechanical Performance and Microstructure Prediction of Hypereutectoid Rail Steels Based on BP Neural Networks. **IEEE Access**, IEEE, v. 8, p. 41905–41912, 2020. ISSN 21693536. Cited on page 50.

DI SCHINO, Andrea; RICHETTA, Maria. Evaluation of the metallurgical parameters effect on tensile properties in austenitic stainless steels. **Acta Metallurgica Slovaca**, v. 23, n. 2, p. 111–121, 2017. ISSN 13381156. Cited on page 50.

DUTTA, Tanusree et al. Designing dual-phase steels with improved performance using ANN and GA in tandem. **Computational Materials Science**, Elsevier, v. 157, n. October 2018, p. 6–16, 2019. ISSN 09270256. Disponível em: <<https://doi.org/10.1016/j.commatsci.2018.10.020>>. Cited on page 59.

ELSKEN, Thomas; METZEN, Jan Hendrik; HUTTER, Frank. **Efficient Multi-objective Neural Architecture Search via Lamarckian Evolution**. 2019. Cited on page 34.

ELSKEN, Thomas; METZEN, Jan Hendrik; HUTTER, Frank. **Neural Architecture Search: A Survey**. 2019. Cited on page 34.

ERES-CASTELLANOS, Adriana et al. An integrated-model for austenite yield strength considering the influence of temperature and strain rate in lean steels. **Materials and Design**, The Authors, v. 188, p. 108435, 2020. ISSN 18734197. Disponível em: <<https://doi.org/10.1016/j.matdes.2019.108435>>. Cited on page 50.

FAIZABADI, Mohammad Javad et al. Predictions of toughness and hardness by using chemical composition and tensile properties in microalloyed line pipe steels. **Neural Computing and Applications**, v. 25, n. 7-8, p. 1993–1999, 2014. ISSN 09410643. Cited on page 50.

FISCHER, F. D. et al. New view on transformation induced plasticity (TRIP). **International journal of plasticity**, v. 16, n. 7, p. 723–748, 2000. ISSN 07496419. Cited on page 28.

FOGEL, David B. Artificial intelligence through simulated evolution. In: _____. **Evolutionary Computation: The Fossil Record**. [S.l.: s.n.], 1998. p. 227–296. Cited on page 38.

FONSECA, C.M.; PAQUETE, L.; LOPEZ-IBANEZ, M. An improved dimension-sweep algorithm for the hypervolume indicator. In: **2006 IEEE International Conference on Evolutionary Computation**. [S.l.: s.n.], 2006. p. 1157–1163. Cited 2 times on pages 9 and 44.

FORRESTER, Alexander; SOBESTER, Andras; KEANE, Andy. **Engineering Design Via Surrogate Modelling: A Practical Guide**. [S.l.: s.n.], 2008. ISBN 978-0-470-06068-1. Cited on page 42.

GALINDO-NAVA, E. I.; RAINFORTH, W. M.; CASTILLO, P. E.J. Rivera-Díaz-del. Predicting microstructure and strength of maraging steels: Elemental optimisation. **Acta Materialia**, Elsevier Ltd, v. 117, p. 270–285, 2016. ISSN 13596454. Disponível em: <<http://dx.doi.org/10.1016/j.actamat.2016.07.020>>. Cited 2 times on pages 50 and 54.

GARCÍA-MARTINO, Ángel et al. Prediction of mechanical properties for high strength low alloyed steels in a commercial hot dip galvanizing line without soaking section. **Metals**, v. 10, n. 5, p. 1–17, 2020. ISSN 20754701. Cited on page 50.

GEBRU, Timnit et al. Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the united states. **Proceedings of the National Academy of Sciences**, National Academy of Sciences, v. 114, n. 50, p. 13108–13113, 2017. ISSN 0027-8424. Disponível em: <<https://www.pnas.org/content/114/50/13108>>. Cited on page 30.

GHAISARI, J.; JANNESARI, H.; VATANI, M. Artificial neural network predictors for mechanical properties of cold rolling products. **Advances in Engineering Software**, Elsevier Ltd, v. 45, n. 1, p. 91–99, 2012. ISSN 09659978. Disponível em: <<http://dx.doi.org/10.1016/j.advengsoft.2011.09.016>>. Cited 3 times on pages 50, 53, and 54.

GROSSI, Enzo; BUSCEMA, Massimo. Introduction to artificial neural networks. **European Journal of Gastroenterology and Hepatology**, v. 19, n. 12, p. 1046–1054, 2007. ISSN 0954691X. Cited on page 31.

HALL, J.N.; FEKETE, J.R. 2 - steels for auto bodies: A general overview. In: RANA, Radhakanta; SINGH, Shiv Brat (Ed.). **Automotive Steels**. Woodhead Publishing, 2017. p. 19–45. ISBN 978-0-08-100638-2. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B978008100638200002X>>. Cited 6 times on pages 9, 20, 22, 24, 26, and 27.

HAN, Zhong-Hua; ZHANG, ke-shi. Surrogate-based optimization. In: _____. [S.l.: s.n.], 2012. ISBN 978-953-51-0146-8. Cited on page 41.

HAYKIN, Simon. **Neural Networks: A Comprehensive Foundation**. 1st. ed. USA: Prentice Hall PTR, 1994. ISBN 0023527617. Cited 2 times on pages 9 and 31.

HE, Haibo; GARCIA, Eduardo A. Learning from imbalanced data. **IEEE Transactions on Knowledge and Data Engineering**, v. 21, n. 9, p. 1263–1284, 2009. Cited on page 66.

HOLLAND, John H. Genetic algorithms and the optimal allocation of trials. **SIAM Journal on Computing**, v. 2, n. 2, p. 88–105, 1973. Disponível em: <<https://doi.org/10.1137/0202009>>. Cited on page 37.

HUNTER, David et al. Selection of proper neural network sizes and architectures—a comparative study. **IEEE Transactions on Industrial Informatics**, v. 8, n. 2, p. 228–240, 2012. Cited 2 times on pages 31 and 32.

JIA, Tao et al. The optimal design for the production of hot rolled strip with "tight oxide scale" by using multi-objective optimization. **ISIJ International**, v. 51, n. 9, p. 1468–1473, 2011. ISSN 09151559. Cited 2 times on pages 50 and 52.

JIN, Haifeng; SONG, Qingquan; HU, Xia. Auto-keras: An efficient neural architecture search system. In: **Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. [S.l.: s.n.], 2019. p. 1946–1956. ISBN 9781450362016. Cited 2 times on pages 34 and 58.

JOHNSON, N.; SANDERS, P. High strength low alloy (HSLA) aluminum. **International Journal of Metalcasting**, v. 6, n. 1, p. 61–62, 2012. ISSN 19395981. Cited 2 times on pages 27 and 28.

KALCHBRENNER, Nal; GREFFENSTETTE, Edward; BLUNSON, Phil. **A Convolutional Neural Network for Modelling Sentences**. 2014. Cited on page 31.

KANCA, E.; ÇAVDAR, F.; ERSEN, M. M. Prediction of mechanical properties of cold rolled steel using genetic expression programming. **Acta Physica Polonica A**, v. 130, n. 1, p. 365–369, 2016. ISSN 1898794X. Cited 3 times on pages 50, 51, and 52.

KARSOLIYA, Saurabh. Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture. **International Journal of Engineering Trends and Technology**, v. 3, n. 6, p. 714–717, 2012. ISSN 2231-5381. Cited on page 32.

KESTENBACH, H.-J; CAMPOS, Sheila; MORALES, Eduardo Valencia. Role of interphase precipitation in microalloyed hot strip steels. **Materials Science and Technology**, v. 22, p. 615–626, 06 2006. Cited on page 67.

KORA, Padmavathi; YADLAPALLI, Priyanka. Crossover operators in genetic algorithms: A review. **International Journal of Computer Applications**, v. 162, p. 34–36, 03 2017. Cited on page 39.

KOTSIANTIS, Sotiris; KANELLOPOULOS, D; PINTELAS, P. Handling imbalanced datasets: A review. **GESTS International Transactions on Computer Science and Engineering**, v. 30, p. 25–36, 11 2005. Cited on page 66.

KOZA, John R. **Genetic Programming : On the Programming of Computers By Means of Natural Selection Complex Adaptive Systems**. [s.n.], 2003. 1–609 p. ISBN 0262111705. Disponível em: <papers2://publication/uuid/5DADD85F-EE2F-42E1-8BF8-2CC6959C4FA0>. Cited on page 37.

KRAUSS, G. Physical metallurgy of steels: An overview. In: RANA, Radhakanta; SINGH, Shiv Brat (Ed.). **Automotive Steels**. Woodhead Publishing, 2017. p. 95–111. ISBN 978-0-08-100638-2. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780081006382000043>>. Cited 2 times on pages 20 and 26.

KüCHLER, Ingeborg. **Data Driven Statistical Methods**. [s.n.], 1997. v. 18. 3153-3154 p. ISBN 9781315140780. Disponível em: <<https://www.taylorfrancis.com/books/mono/10.1201/9781315140780/data-driven-statistical-methods-jim-zidek-peter-sprent>>. Cited on page 29.

LEIJNEN, Stefan; VEEN, Fjodor van. The neural network zoo. **Proceedings**, v. 47, n. 1, 2020. ISSN 2504-3900. Disponível em: <<https://www.mdpi.com/2504-3900/47/1/9>>. Cited on page 30.

LI, Miqing; YANG, Shengxiang; LIU, Xiaohui. Diversity comparison of pareto front approximations in many-objective optimization. **IEEE Transactions on Cybernetics**, v. 44, n. 12, p. 2568–2584, 2014. Cited 2 times on pages 45 and 71.

LI, Miqing; YAO, Xin. Quality evaluation of solution sets in multiobjective optimisation: A survey. **ACM Computing Surveys**, v. 52, p. 1–38, 03 2019. Cited on page 43.

LIU, Chenxi et al. Progressive neural architecture search. In: **Proceedings of the European Conference on Computer Vision (ECCV)**. [S.l.: s.n.], 2018. Cited on page 34.

LUNDBERG, Scott M; LEE, Su-In. A unified approach to interpreting model predictions. In: GUYON, I. et al. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2017. v. 30. Disponível em: <<https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf>>. Cited 2 times on pages 53 and 75.

MARLER, R.; ARORA, Jasbir. Survey of multi-objective optimization methods for engineering. **Structural and Multidisciplinary Optimization**, v. 26, p. 369–395, 04 2004. Cited 2 times on pages 35 and 36.

MENGISTU, Temesgen; GHALY, Wahid. Aerodynamic optimization of turbomachinery blades using evolutionary methods and ANN-based surrogate models. **Optimization and Engineering**, v. 9, n. 3, p. 239–255, 2008. ISSN 13894420. Cited on page 41.

MOHANTY, Itishree; BHATTACHARJEE, Debashish; DATTA, Shubhabrata. Designing cold rolled if steel sheets with optimized tensile properties using ann and ga. **Computational Materials Science**, v. 50, n. 8, p. 2331–2337, 2011. ISSN 0927-0256. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0927025611001492>>. Cited on page 66.

MOREIRA, João et al. Ensemble approaches for regression: A survey. **ACM Computing Surveys**, v. 45, p. 10:1–10:40, 11 2012. Cited 2 times on pages 30 and 31.

MUKHERJEE, Krishnendu; HAZRA, Sujoy; MILITZER, M. Grain refinement in dual-phase steels. **Metallurgical and Materials Transactions A**, v. 40, p. 2145–2159, 09 2009. Cited on page 67.

NOWLAN, Steven J. Simplifying Neural Networks by Soft Weight Sharing. **The Mathematics of Generalization**, v. 493, p. 373–394, 2018. Cited on page 33.

OKOLI, Chitu; SCHABRAM, Kira. Working Papers on Information Systems A Guide to Conducting a Systematic Literature Review of Information Systems Research. **Working Papers on Information Systems**, v. 10, n. 2010, 2010. ISSN 1535-6078. Cited 2 times on pages 47 and 48.

OTMAN, Abdoun; ABOUCHABAKA, Jaafar; TAJANI, Chakir. Analyzing the performance of mutation operators to solve the travelling salesman problem. **Int. J. Emerg. Sci.**, v. 2, 03 2012. Cited on page 40.

PANDYA, Dipali; SHAH, Dhaval. Experimentation and its Prediction of Process Parameters Effects on Elongation in Tensile Test of AISI 1008 Steel Using ANN Model. **Procedia Technology**, Elsevier B.V., v. 14, p. 282–289, 2014. ISSN 22120173. Disponível em: <<http://dx.doi.org/10.1016/j.protcy.2014.08.037>>. Cited on page 50.

PARPINELLI, Rafael et al. A review of techniques for online control of parameters in swarm intelligence and evolutionary computation algorithms. **International Journal of Bio-Inspired Computation**, v. 13, p. 1, 01 2019. Cited on page 62.

PAUPLER, P. G. e. dieter. mechanical metallurgy. 3rd ed., mc graw-hill book co., new york 1986. xxiii + 751 p., dm 138.50, isbn 0–07–016893–8. **Crystal Research and Technology**, v. 23, n. 2, p. 194–194, 1988. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/crat.2170230211>>. Cited 2 times on pages 24 and 25.

PICKERING, F. B. Physical metallurgy and the design of steels. In: . [S.l.: s.n.], 1978. Cited on page 67.

RAY, R. K.; JONAS, J. J.; HOOK, R. E. **Cold rolling and annealing textures in low carbon and extra low carbon steels**. [S.l.: s.n.], 1994. v. 39. 129–172 p. ISSN 17432804. ISBN 0000000000. Cited on page 27.

RONALD, S. Robust encodings in genetic algorithms: a survey of encoding issues. In: **Proceedings of 1997 IEEE International Conference on Evolutionary Computation (ICEC '97)**. [S.l.: s.n.], 1997. p. 43–48. Cited on page 37.

SALTELLI, Andrea et al. **Global Sensitivity Analysis: The Primer**. [S.l.: s.n.], 2008. v. 76. 452–452 p. ISSN 03067734. ISBN 9780470059975. Cited on page 74.

SANTOS, Andrea Pedroza da Rocha et al. Texture, microstructure and anisotropic properties of if-steels with different additions of titanium, niobium and phosphorus. **Journal of Materials Research and Technology**, v. 7, n. 3, p. 331–336, 2018. ISSN 2238-7854. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S223878541730827X>>. Cited on page 66.

SARAVANAKUMAR, P. et al. Rolling process using artificial neural network model. **Procedia Engineering**, v. 38, p. 3418–3425, 2012. ISSN 18777058. Cited 2 times on pages 50 and 52.

SGROTT, Douglas et al. Modelling if steels using artificial neural networks and automated machine learning. In: _____. [S.l.: s.n.], 2021. p. 659–668. ISBN 978-3-030-73049-9. Cited 6 times on pages 10, 21, 53, 73, 74, and 75.

SHEN, Chunguang et al. Physical metallurgy-guided machine learning and artificial intelligent design of ultrahigh-strength stainless steel. **Acta Materialia**, Elsevier Ltd, v. 179, p. 201–214, 2019. ISSN 13596454. Disponível em: <<https://doi.org/10.1016/j.actamat.2019.08.033>>. Cited 3 times on pages 21, 53, and 59.

SHOW, B.K. et al. Effect of vanadium and titanium modification on the microstructure and mechanical properties of a microalloyed hsla steel. **Materials Science and Engineering: A**, v. 527, n. 6, p. 1595–1604, 2010. ISSN 0921-5093. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0921509309011782>>. Cited on page 67.

SHUKLA, Anupriya; PANDEY, Hari Mohan; MEHROTRA, Deepti. Comparative review of selection techniques in genetic algorithm. In: **2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)**. [S.l.: s.n.], 2015. p. 515–519. Cited on page 39.

SOLEIMANI, Maryam; KALHOR, Alireza; MIRZADEH, Hamed. Transformation-induced plasticity (TRIP) in advanced steels: A review. **Materials Science and Engineering A**, Elsevier B.V., v. 795, n. July, 2020. ISSN 09215093. Cited 3 times on pages 9, 28, and 29.

SØNDERGAARD, Jacob. **Optimization using surrogate models - by the space mapping technique**. 203 p. Tese (Doutorado) — Informatics and Mathematical Modelling, Technical University of Denmark, DTU, 01 2003. Cited on page 42.

SREEKANTH, J.; DATTA, Bithin. Multi-objective management of saltwater intrusion in coastal aquifers using genetic programming and modular neural network based surrogate models. **Journal of Hydrology**, v. 393, n. 3, p. 245–256, 2010. ISSN 0022-1694. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0022169410005408>>. Cited on page 41.

SRIVASTAVA, Nitish et al. Dropout: A simple way to prevent neural networks from overfitting. **Journal of Machine Learning Research**, v. 15, n. 56, p. 1929–1958, 2014. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>. Cited on page 33.

STORK, Jørg; EIBEN, A.; BARTZ-BEIELSTEIN, Thomas. A new taxonomy of global optimization algorithms. **Natural Computing**, p. 1–24, 11 2020. Cited 4 times on pages 9, 39, 42, and 43.

TAKECHI, Hiroshi. Metallurgical Aspects on Interstitial Free Sheet Steel from Industrial Viewpoints. **ISIJ International**, v. 34, n. 1, p. 1–8, 1994. ISSN 09151559. Cited on page 74.

TAN, Mingxing et al. Mnasnet: Platform-aware neural architecture search for mobile. In: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2019. Cited on page 34.

TAO, Wei et al. Multi-scale design of three dimensional woven composite automobile fender using modified particle swarm optimization algorithm. **Composite Structures**, v. 181, p. 73–83, 2017. ISSN 0263-8223. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0263822317316239>>. Cited on page 41.

TASAN, C. C. et al. An Overview of Dual-Phase Steels: Advances in Microstructure-Oriented Processing and Micromechanically Guided Design. **Annual Review of Materials Research**, v. 45, n. April, p. 391–431, 2015. ISSN 15317331. Cited on page 28.

TODOROVSKI, M.; RAJICIC, D. An initialization procedure in solving optimal power flow by genetic algorithm. **IEEE Transactions on Power Systems**, v. 21, n. 2, p. 480–487, 2006. Cited on page 39.

VIKHAR, Pradnya A. Evolutionary algorithms: A critical review and its future prospects. In: **2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)**. [S.l.: s.n.], 2016. p. 261–265. Cited on page 37.

WANG, Chenchong et al. Design of comprehensive mechanical properties by machine learning and high-throughput optimization algorithm in RAFM steels. **Nuclear Engineering and Technology**, Elsevier Ltd, v. 52, n. 5, p. 1008–1012, 2020. ISSN 2234358X. Disponível em: <<https://doi.org/10.1016/j.net.2019.10.014>>. Cited 4 times on pages 50, 53, 54, and 55.

WANG, Xiangwen; LIN, Xianghong; DANG, Xiaochao. Supervised learning in spiking neural networks: A review of algorithms and evaluations. **Neural Networks**, v. 125, p. 258–280, 2020. ISSN 0893-6080. Cited on page 29.

WANG, Yuxuan et al. Prediction and analysis of tensile properties of austenitic stainless steel using artificial neural network. **Metals**, v. 10, n. 2, 2020. ISSN 20754701. Cited on page 50.

WU, Siwei et al. Elevating Prediction Performance for Mechanical Properties of Hot-Rolled Strips by Using Semi-Supervised Regression and Deep Learning. **IEEE Access**, v. 8, p. 134124–134136, 2020. ISSN 21693536. Cited on page 50.

WU, Siwei et al. High-Dimensional Data-Driven Optimal Design for Hot Strip Rolling of Microalloyed Steel. **Steel Research International**, v. 89, n. 7, p. 1–11, 2018. ISSN 16113683. Cited 4 times on pages 20, 50, 52, and 53.

XIONG, Jie; ZHANG, Tong Yi; SHI, San Qiang. Machine learning of mechanical properties of steels. **Science China Technological Sciences**, v. 63, n. 7, p. 1247–1255, 2020. ISSN 1862281X. Cited on page 50.

YANMINSUN; WONG, Andrew; KAMEL, Mohamed S. Classification of imbalanced data: a review. **International Journal of Pattern Recognition and Artificial Intelligence**, v. 23, 11 2011. Cited on page 66.

ZHOU, Aimin et al. Multiobjective evolutionary algorithms: A survey of the state of the art. **Swarm and Evolutionary Computation**, Elsevier B.V., v. 1, n. 1, p. 32–49, 2011. ISSN 22106502. Disponível em: <<http://dx.doi.org/10.1016/j.swevo.2011.03.001>>. Cited 3 times on pages 38, 40, and 41.

ZHOU, Yu; WU, Yali. Analyses on influence of training data set to neural network supervised learning performance. **Advances in Intelligent and Soft Computing**, v. 106, n. 1, p. 19–25, 2011. ISSN 18675662. Cited 2 times on pages 21 and 31.

ZOPH, Barret; LE, Quoc V. **Neural Architecture Search with Reinforcement Learning**. 2017. Cited on page 34.