SANTA CATARINA STATE UNIVERSITY - UDESC
CENTER OF TECHNOLOGICAL SCIENCES - CCT
GRADUATE PROGRAM IN APPLIED COMPUTING - PPGCAP

BERNARD DA SILVA

AUTOMOTIVE STAMPING PROCESS OPTIMIZATION USING MACHINE LEARNING AND MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM

JOINVILLE

2023

BERNARD DA SILVA

AUTOMOTIVE STAMPING PROCESS OPTIMIZATION USING MACHINE LEARNING AND MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM

Dissertation submitted to the Postgraduate Program in Applied Computing at the Center for Technological Sciences of the State University of Santa Catarina, to obtain the degree of Master in Applied Computing.

Advisor: Dr. Rafael Stubs Parpinelli

JOINVILLE

2023

Ficha catalográfica elaborada pelo programa de geração automática da Biblioteca Universitária Udesc, com os dados fornecidos pelo(a) autor(a)

Silva, Bernard de
Automotive stamping process optimization using machine learning and multi-objective evolutionary algorithm / Bernard de Silva. -- 2024.
77 p.

Orientador: Rafael Stubs Parpinelli
Dissertação (mestrado) -- Universidade do Estado de Santa Catarina, Centro de Ciências Tecnológicas, Programa de Pós-Graduação em Computação Aplicada, Joinville, 2024.

1. Multi-objective optimization. 2. Surrogate model. 3. Evolutionary computing. 4. Automotive industry. 5. Metal stamping. I. Parpinelli, Rafael Stubs. II. Universidade do Estado de Santa Catarina, Centro de Ciências Tecnológicas, Programa de Pós-Graduação em Computação Aplicada. III. Título.

BERNARD DA SILVA

AUTOMOTIVE STAMPING PROCESS OPTIMIZATION USING MACHINE LEARNING AND MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM

Dissertation submitted to the Postgraduate Program in Applied Computing at the Center for Technological Sciences of the State University of Santa Catarina, to obtain the degree of Master in Applied Computing.

Adivisor: Dr. Rafael Stubs Parpinelli

EXAMINATION BOARD:

Rafael Stubs Parpinelli
President/Advisor

Members:

Rafael Stubs Parpinelli
UDESC

Chidambaram Chidambaram
UDESC

Carmelo Jose Albanez Bastos Filho
UPE

Joinville, December 15, 2023

I dedicate this work to my wife and daughter, who have always been by my side, serving as sources of inspiration and support.

ACKNOWLEDGEMENTS

I want to express my deep gratitude to Rafael Parpinelli for the valuable opportunity to participate in this project. Furthermore, I would like to extend my sincere thanks to collaborators Ana Athayde Carneiro, Lucas Gonçalves Erdmann, and Peter Brendel. Your support was fundamental to the success of this project.

I also express my deep gratitude to my wife and daughter. Without their support, none of this would have been possible. In challenging moments, their love proved to be a solid pillars that supported me at the top.

I want to express my sincere gratitude to ArcelorMittal for making this work financially viable. I want to recognize the essential collaboration of experts José Osvaldo Tepedino and Roan Sampaio Souza, who played a crucial role in the realization of this project and believed in the partnership between ArcelorMittal and UDESC, which made it all possible.

To my friends, who patiently listened to my daily rants and witnessed my exhausting journey, I want to express my deep gratitude for always supporting me and keeping me from giving up.

My sincere gratitude to all my teachers, who provided me with a wide range of new perspectives through which I could see the world in a new way.

Last, I would like to thank all the family members who supported me for the previous two years. The encouragement from each of you served as fuel for my determination.

“In the midst of difficulty lies opportunity.”

Albert Einstein

ABSTRACT

Stamping of automotive parts is a process of forming metal parts that are used in the manufacture of automobiles, such as the Inner Rear Door. The problem is characterized as a multi-objective optimization problem, as it involves the optimization of multiple antagonistic objectives simultaneously. In this study, Random Forest and Extra Tree multivariate regression algorithms were used to characterize the problem and use it as a replacement model for the Non-Dominated Genetic Sorting Algorithm II. Our goal is to explore the possibilities of simultaneously minimizing fracture, insufficient stretching, and wrinkling. In this work, two case studies were analyzed: one of a laboratory nature, called "Cup", and another related to the Internal Tailgate Door of a vehicle. The results obtained through the proposed approach highlight the effectiveness of the model, contributing to the identification of a more efficient solution. Notably, the results revealed an approximately 15% reduction in wrinkling and an 8% reduction in under stretch in the Internal Tailgate Door study when compared to the previously used methodology.

Key-words: Multi-objective Optimization, Surrogate Model, Evolutionary Computing, Automotive Industry, Metal Stamping.

RESUMO

A estampagem de peças automotivas é um processo de conformação de peças metálicas que são utilizadas na fabricação de automóveis, como a Porta Traseira Interna. O problema é caracterizado como um problema de otimização multiobjetivo, pois envolve a otimização de múltiplos objetivos antagônicos simultaneamente. Neste estudo, utilizou-se algoritmo de regressão multivariada Random Forest e Extra Tree para caracterizar o problema e utilizá-lo como modelo substituto para o Algoritmo Genético de Ordenação Não Dominada II. Nosso objetivo é explorar as possibilidades de minimizar simultaneamente Fratura, Alongamento Insuficiente e Enrugamento. Neste trabalho, foram analisados dois estudos de caso: um de natureza laboratorial, denominado "Copinho", e outro relacionado à Porta Traseira Interna de um veículo. Os resultados obtidos por meio da abordagem proposta destacam a eficácia do modelo, contribuindo para a identificação de uma solução mais eficiente. Notavelmente, os resultados revelaram uma redução de aproximadamente 15% no enrugamento e 8% no estiramento insuficiente no estudo da Porta Traseira Interna, ao compararmos com a metodologia previamente utilizada.

Palavras-chave: Otimização Multiobjetivo, Modelo Substituto, Computação Evolutiva, Indústria Automotiva, Estampagem de Metais.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ANN	Artificial Neural Network
BHF	Blank Holder Force
CLC	Conformation Limit Curve
DOE	Design of Experiments
EA	Evolutionary Algorithm
ET	Extra Tree Regressor
FEM	Finite Element Method
FRA	Fracture
FRC	Friction Coefficient
IS	Insufficient Stretch
KRG	Kriging
MOEA	Multi-Objective Evolutionary Algorithm
MOGA	Multi-Objective Genetic Algorithm
ML	Machine Learning
NBI	Normal Boundary Intersection
NNCM	Normalized Normal Constraints Methods
NSGA	Non-dominated Sorting Genetic Algorithm
RBF	Radial basis function
RES	Restraining Force Factor
RF	Random Forest
RSM	Response Surface Methodology
SA	Simulated Annealing
SLR	Systematic Literature Review
SPSA	Simultaneous Perturbation Stochastic Approximation
TS	Tensile Strength
SVM	Support Vector Machine

SVR	Support Vector Regression
WRI	Wrinkle
YS	Yield Strength

LIST OF SYMBOLS

ε_1	Major Strain
ε_2	Minor Strain
Rav	R average
$Eq.Line$	Equation of the curve line conformation limit test
MPa	Megapascal
RMSE	Root Mean Square Error
X_{01-05}	Independent parameters used in modeling and as decision variables in optimization

CONTENTS

1 INTRODUCTION

The Industry 4.0 concept emerged in 2011 at the Hanover fair in Germany, presenting several cutting-edge industrial futuristic concepts. Since then, in a natural process, companies have been rapidly transforming with the support of intelligent tools delivered to the market. Among the tools currently used, many are focused on computational intelligence, which has been generating several opportunities for research and development in different areas (COSTA, 2019).

The automotive industry tends to be a pioneer when it comes to investment in research and development. Because of this, the evolution of increasingly elaborate and innovative solutions has been presented to the market. During the last decades, a growing improvement in requirements such as safety, efficiency, ergonomics, and economy can be observed. This is due to significant improvements on different fronts, such as raw materials, with the creation of new, more efficient steel alloys; improvements in manufacturing processes, such as process automation and machinery evolution; and the development of intelligent tools to support decision-making (ATZEMA, 2017).

One of the manufacturing processes that has evolved over the last few years is the sheet metal stamping process, a process widely used in the automobile industry, where controlled deformation of metal sheets occurs, with high demand levels, high productivity, low scrapping, low cost, high finish, and accuracy (MA; HUANG et al., 2014). The process aims to generate a form from the raw materials used, which is very useful in serial industrial systems. The set of parts necessary to carry out a stamping is composed of a matrix with the negative shape of the part's profile, a blank to restrict the movement of the plate, and a punch that allows the penetration of the plate into the matrix, making the final part as the desired shape.

The stamping process has a significant influence on the final quality of molded parts (HOSFORD; CADDELL, 2011). Finite element analysis has often been used as a valuable testing method to evaluate stamping formability and to help define better parameters for the machines responsible for the stamping process in the factory. However, when a more in-depth analysis is required involving a high number of simulations, challenges arise in terms of industrial scalability. This is due to some factors, such as excessive time required to run the simulations, the need for qualified labor for analysis, and high-performance hardware to support the simulations (KAZAN; FIRAT; TIRYAKI, 2009).

Currently, professionals in the field use their knowledge to manually configure a wide range of parameters in order to find the best scenarios with the help of finite element analysis tools. The big problem with this method is making the process biased towards the developer's vision and experience since the ideal combination to minimize failures is primarily based on experience and trial and error steps (TEKKAYA, 2000). For this reason, remaining in the empirical state can delay development, as well as increase the risk of problems occurring throughout the production process. This gap represents the possibility of expanding knowledge of current metallurgy, opening new doors for new applications and opportunities, thus offering new mathematical and computational tools to assist in decision-making on the best parameters for automotive stamping.

The problem addressed involves minimizing failures in the stamping process of automotive parts to reduce considerably the financial and environmental losses. To deal with this challenge, it is essential to find an ideal set of parameters that when applied, minimize the occurrence of failures. However, it is essential to highlight that these objectives are often antagonistic, which means that minimizing one of them can directly affect the others.

Given the complexity of the stamping process, several approaches have been applied to address these challenges (PARK; KIM, 1995). One of them is the use of evolutionary multi-objective optimization techniques, which allow the advanced search for optimal solutions that would otherwise be unfeasible or very difficult to find manually. These evolutionary algorithms can explore the solution space more efficiently, finding trade-offs between objectives and providing a set of Pareto-optimal solutions (LINDEN, 2008).

Furthermore, the application of Machine Learning (ML) algorithms as surrogate models has demonstrated great potential in dealing with complex problems such as those presented in the works (SGROTT, 2022) (BAO et al., 2022) (BRENDEL et al., 2021). The substitute models aim to help represent these problems, allowing a quick and efficient approach. They can later be used as input models in other operations, such as optimization, through evolutionary algorithms (JIANG et al., 2020).

Faced with the numerous challenges in the stamping process, several studies have presented different approaches to improve this stage, such as the works (CUI et al., 2020) (BRENDEL et al., 2021) (XIE et al., 2022). Based on solutions previously described in the literature, this work proposes the development of a computational system for optimizing the automotive stamping process, suggesting stamping parameters, using ML algorithms as substitute models coupled to multi-objective optimization algorithms, with This aims to reduce the inside of regions with three main faults, namely, wrinkling (WRI), insufficient stretching (IS), and fracture (FRA).

This work uses ML, in particular, the ExtraTree Regressor (ET) algorithm, as a predictive model to update the stamping process of any automotive part that uses the "Deep draw" process, employing a multivariate regression approach (CHANKONG; HAIMES, 2008). The choice of ExtraTree Regressor as an algorithm is based on its intrinsic characteristics, such as speed in training and greater interpretability of the results obtained. The ET algorithm inputs include steel stamping parameters and domains, machine parameters, and dimensional information. At the same time, its outputs are the result of the part's conformation, the percentage of area with the possibility of FRA, IS, and WRI. Furthermore, ET is integrated with a Multi-Objective Evolutionary Algorithm (MOEA) for optimization, using ET predictions as an evaluation function to guide the evolution of different solutions towards the desired shape of the parts. Some studies, such as (RAHIMI et al., 2023) and (RIBEIRO, 2016), carried out comparative analyses of genetic algorithms to evaluate their performance in different scenarios. These investigations highlight the promising results obtained by the Non-Dominated Classification Genetic Algorithm (NSGA-II) in multi-objective problems. Therefore, in this work, NSGA-II is used to manage candidate solutions for Pareto optimal optimization (DEB et al., 2002).

Case studies were conducted to validate the proposed model, both through computer simulations and practical experiments. Through the using some algorithms, it was possible to evaluate the capacity of the model in different contexts. The analysis of different case studies allowed the identification of the limitations of the model and the understanding of the impacts of its application.

These case studies provided a comprehensive assessment of the model, exploring its effectiveness and performance under different scenarios and conditions. By performing a variety of experiments and simulations, it was possible to gain valuable insights into the capabilities and strengths of the model, as well as identify any constraints and challenges to be faced.

1.1 MOTIVATION

There are many factors involved in stamping automotive parts, such as material properties, manufacturing operating conditions, cost, sustainability, safety, final quality, and product scalability.

Consequently, the search for instruments that facilitate decision-making in choosing ideal stamping configurations, directing the granting of components with superior operational performance in various applications, becomes a growing trend. In this context, opportunities arise to improve stamping procedures through the application of intelligent computational models.

1.2 OBJECTIVES

The main objective of this work is to develop a prediction system based on computational intelligence capable of proposing the best parameters for stamping automotive parts. To achieve this, the prediction mechanism used in the project must contain a model capable of generating genetic diversity by exploring different solutions for users in a multi-objective scenario.

The following are some secondary objectives, which will build the steps to achieve the main objective:

- Research for implemented tools implemented to optimize the automotive stamping process.
- Develop software to help extract data from the AutoForm simulator, enabling an increase in the amount of data and accelerating the document extraction step.
- Develop algorithms capable of replicating numerical calculations performed by the AutoForm software since it does not provide the desired objective information, namely WRI, IS, and FRA.
- Develop exploratory data analysis, which enables a deeper understanding of the selection of parameters and domains used to generate the database.
- Create a model for each automotive part using ET so that the model can be trusted and accurately represent the modeled element.
- Deepen the understanding of the surrogate model through the application of sensitivity analysis techniques.
- Develop a multi-objective solution exploration system using evolutionary optimization algorithms.
- Apply methods and tools to analyze optimization results, allowing assessment of resulting diversity and convergence.

1.3 SCIENTIFIC CONTRIBUTIONS

During this work, some intermediate scientific contributions were made, which are briefly described below:

The paper titled "Automotive Stamping Process Optimization Using ML and Multi-objective Evolutionary Algorithm" (SILVA et al., 2023a), was presented at the International Conference on Intelligent Systems Design and Applications (ISDA) in 2022, aims to optimize the process of selecting stamping parameters for the manufacture of a fundamental part, the roof of a vehicle, a flat surface with few details.

In this study, an approach is employed that combines the ML Random Forest (RF) algorithm as a surrogate model coupled with the multi-objective genetic algorithm NSGA-II. The case study involved the analysis of three input variables and three output variables for a restricted database with a number of 200 experiments.

The results obtained in the study were promising, with a significant reduction in wrinkling and insufficient stretch objectives.

The second work is titled "Surrogate Model and Multi-objective Evolutionary Algorithm Applied to Automotive Stamping" (SILVA et al., 2023b). It was presented at the XVI Brazilian Conference on Computational Intelligence (CBIC) in 2023. In it, complementary analysis tools are presented to help understanding the intermediate steps of the hybrid model proposed in this thesis. The primary purpose of these tools is to assist users in making early decisions, mainly in anticipating possible errors in the inadequate selection of simulation parameters and domains.

These tools were applied in two case studies, one on a laboratory part and the other on a real automotive part. Both demonstrated good results, both in terms of optimization and the interpretability of the proposed model.

The results obtained in the study were promising, with a significant reduction in wrinkling and insufficient stretch in the Internal Tailgate case study. Furthermore, in the Copinho case study, the tools provided interesting triggers that guided the redefinition of the experiment plan, demonstrating their effectiveness in decision-making.

1.4 DOCUMENT STRUCTURE

The remaining chapters of this master's thesis are organized as follows: Chapter 2 details the concepts involved in the development of the proposed approach. In chapter 3, it presents the Systematic Literature Review (SLR) regarding the computational methods used in the prediction and definition of stamping parameters. The proposed model for solving the problem is detailed in Chapter 4. Chapter 5 presents how the case studies were conducted, their results, and the analyses performed. Finally, in Chapter 6, the conclusions obtained and the possibilities for future work are presented.

2 BACKGROUND

This chapter provides the necessary metallurgical and computational background to ground the present study, especially concerning the use of data-based methods for parameter suggestion design and automotive stamping domains.

2.1 AUTOMOTIVE MECHANICAL FORMING PROCESS AND PROPERTIES

It is believed that the first stamping processes were carried out in the manufacture of coins many years ago, a process based on the use of the weight of a tool (in this case, a hammer) on a mold. This method of hammering to generate shape has been used for many years.

With the global evolution, a growing demand for more elaborate solutions emerged, and with that, the stamping process needed to undergo improvements, mainly in terms of quality and productivity.

The automotive mechanical forming process is a fundamental process for the production of metallic parts. This process is used to convert flat sheet metal into specific shapes. Automotive mechanical forming can be accomplished in several ways, including stamping, forging, and rolling (HOSFORD; CADDELL, 2011).

In addition to the specific properties of each automobile mechanical forming method, several general properties are important to guarantee the quality and durability of the produced parts, as well as evaluating the tools to describe these characteristics. Among these properties are yield stress, tensile strength, and elongation (BOLJANOVIC, 2004).

2.1.1 AUTOMOTIVE MECHANICAL FORMING PROCESS

The manufacturing process of automotive parts is composed of several stages, including design, cutting, stamping, modeling, and others. This work focuses on the stamping process using the deep drawing technique and the intrinsic conformations of the process.

2.1.1.1 MATERIAL PROPERTIES AND OPERATING CONDITIONS

Metal stamping is considered one of the essential processes in the production process of automotive parts, as its quality directly impacts the final product. The success of this process strongly depends on the material properties and operating conditions.

Knowledge of the mechanical properties of materials, such as yield strength, tensile strength, and elongation, is essential to determine the material's ability to deform without breaking during the stamping process (KIM; PARK, 2002) (MA; HUANG et al., 2014).

In addition to material properties, operational conditions influence metal stamping, some of which stand out: Temperature, tooling configuration, stamping speed, lubrication, and dimensional tolerances. Among several conditions, these are some of the main ones that must be controlled (PARK; KIM, 1995).

Finally, it is essential to highlight that the optimization of the metal stamping process is a complex process with several challenges, as it involves the selection of several parameters and seeks to find a balance between antagonistic features, such as lightness and resistance of the part.

2.1.1.2 STAMPING PROCESS AND TOOLING

There are three main types of metal stamping techniques: progressive, four-blade, and deep draw.

The method is called "deep draw" when the depth of the stamped part exceeds its diameter. This type of forming is ideal for creating components that require different series of diameters and generating metal parts with complex and detailed shapes. Figure 1 shows a schematic of this conventional process. This type of forming is ideal for creating components that require significant variations in diameter and for producing metal parts with complex and detailed shapes. In short, the process begins with the creation of a die and the tools used in the press, including the punch and a set of cavities that define the final shape of the part. Next, the raw material is adequately lubricated and inserted between the punch and the die in the press, where it is subjected to a compression force that shapes it, pressing the material against the molds and resulting in a part with the desired shape. Stamping can be carried out cold or hot, and the choice of method depends on the composition of the metal and the complexity of the part (BOLJANOVIC, 2004).

2.1.2 STEEL PROPERTIES

There are many types of steel properties, such as electrical (ability to conduct), magnetic (ability of a material to be attracted or become a magnet), chemical (molecular structure of the material), and thermal (ability to conduct or dissipate heat). In this context, the following subsections will focus on a brief discussion of the three fundamental mechanical properties: yield strength, tensile strength, and elongation.

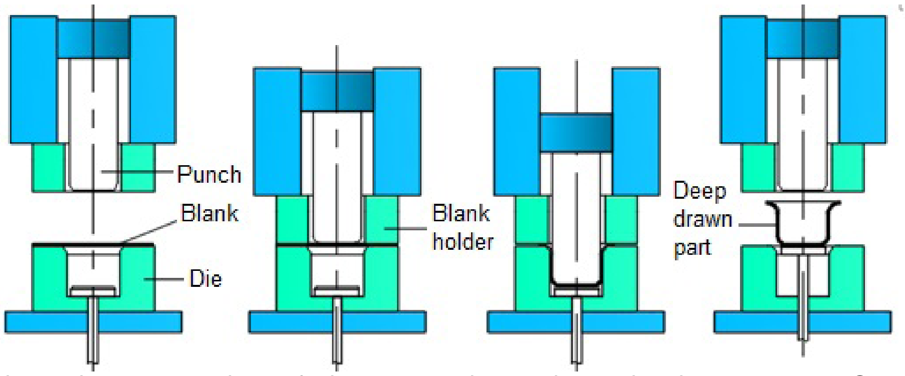

Figure 1 – Schematic presentation of the conventional deep drawing process. Source:(GÜRÜN; KARAAĞAÇ, 2015).

2.1.2.1 YIELD STRENGTH

Yield strength is a mechanical property of metals that is directly related to their ability to resist plastic deformation.

Yield strength is defined as the amount of stress a material can withstand before it begins to permanently deform. In other words, it is the amount of force acting on the cross-sectional area of the material to a point where it can no longer return to its original shape, as illustrated in Figure 2 (SGROTT, 2022).

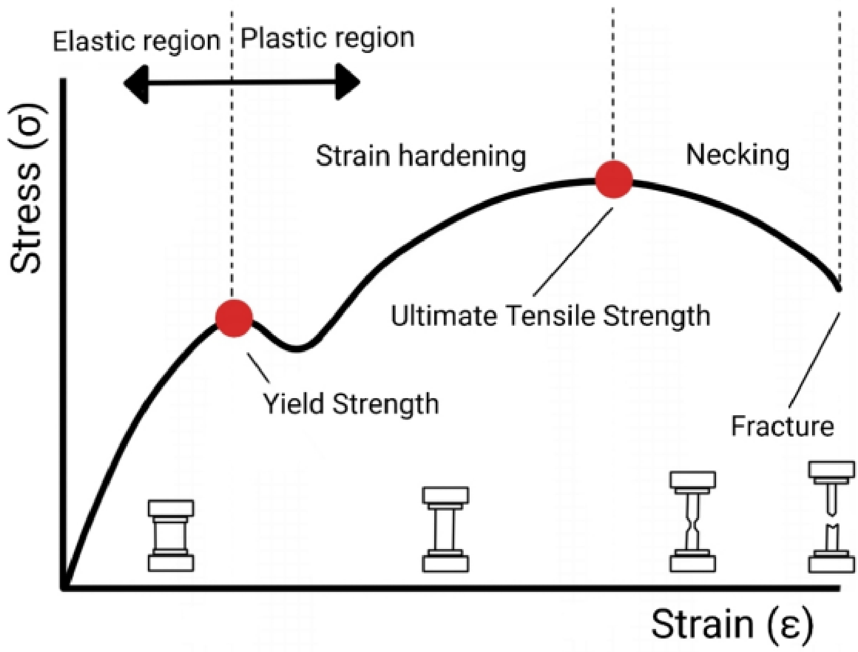

Figure 2 – Schematic presentation of the conventional deep drawing process. Source: (SGROTT, 2022).

Materials with a high yield strength are more difficult to deform. Consequently, more robust equipment is needed to enable the stamping process to be carried out. On the other hand, materials with low yield strength are easier to deform, inherent to the formation of wrinkles in the stamped part.

Figure 3 presents a comparison of different materials with their respective deformations when a known stress is applied to the material. In the illustration, it is possible to observe the behavior of materials with high (DP980) and low yield strength (DQSK). The DP980 material can withstand very high tension (approximately 1000 MPa). However, its tensile capacity is minimal, and fracture occurs quickly after reaching maximum deformation stress (next to 10%). On the other hand, the DQSK material can be highly tensile (next to 45%) and has a low resistance to applied tension (approximately 200 MPa).

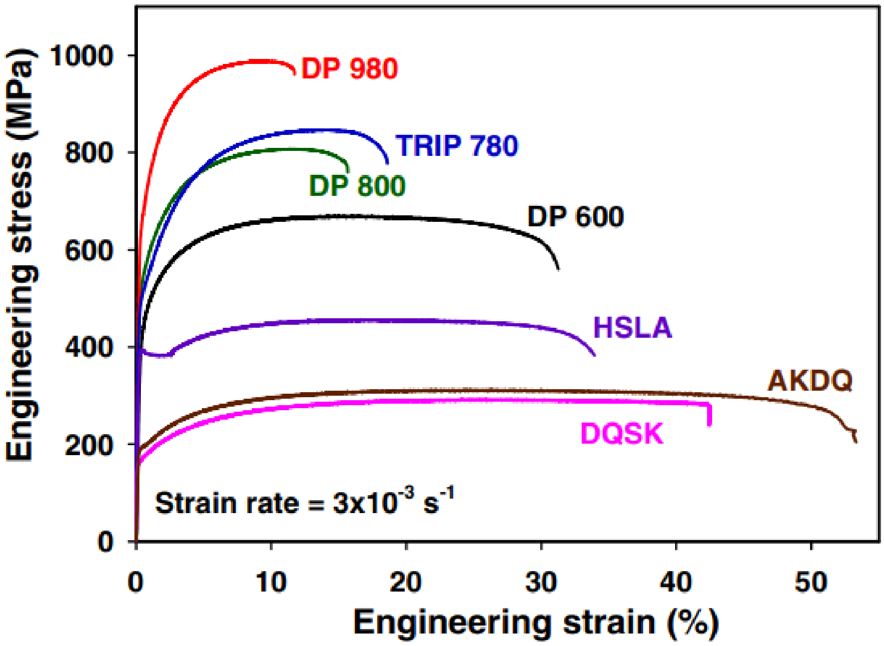

Figure 3 – Typical stress x strain curve for different steel's. Source: (LIM et al., 2012).

2.1.2.2 TENSILE RESISTANCE

Tensile strength is defined as the maximum tensile load that a material can withstand, that is, the maximum of the deformation stress as shown in Figure 2 indicated by curve (ultimate tensile strength).

Generally, high-strength steels have a higher tensile strength, as illustrated by the example of DP980 steel (Figure 3). On the other hand, milder or lower-strength steels, such as DQSK and AKDQ (Figure 3), have lower tensile strength. High-strength steels are commonly used in applications that require high tensile strength, such as structural parts that need to withstand heavy loads. Low-strength steels, on the other hand, are often used in applications where structural strength is not the main criterion, such as aesthetic components, where other factors are decisive for material selection (BOLJANOVIC, 2004).

2.1.2.3 ELONGATION

Elongation is a measure of the ability of a material to plastically deform before breaking when subjected to tensile stress. Such a measure is obtained after destructive tests carried out in the laboratory, where after the part fractures, it is possible to calculate the percentage of elongation obtained according to Equation 2.1. Therefore, elongation is an important property to evaluate the ductility and formability of the material (BOLJANOVIC, 2004).

$$e_f = \frac{L_f - L_a}{L_a} \tag{2.1}$$

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

Tensile strength was the subject of study by (LIM et al., 2012), in which he observed that high-strength steels demonstrate a greater capacity for tensile strength. This is exemplified by DP980 steel, as illustrated in Figure 3. In contrast, softer or low-strength steels, such as DQSK and AKDQ (Figure 3), exhibit lower tensile strength.

2.1.2.4 CONFORMATION LIMIT CURVE (CLC)

The Forming Limit Curve (CLC) is a tool used to describe the relationship between stress and strain during the steel forming process, whose typical appearance is shown in Figure 4. Provides essential information about the level and type of deformation that a material can withstand when subjected to forming stresses. CLC has been widely used in the comparative evaluation of materials and the prediction of failures in automotive parts, as applied by Tepedino in (TEPEDINO, 2014), which aims to predict rupture at the edges of stamped parts in the automotive industry.

The Forming Limit Curve (CLC) can be used to identify areas of WRI, IS, and FRA during the steel forming process. This information helps in the early assessment of the quality and integrity of the parts produced. Furthermore, they can be used in a complementary way in applications aimed at optimizing the stamping process, as demonstrated in the methodology used in this study.

2.2 DATA MODELING

With technological advances in recent years, access to an ever-increasing amount of data has become possible, driven by the development of more powerful hardware and advanced signal processing tools. As pointed out by a study (SCHONFELD, 2010), the amount of data generated will continue to increase in the coming years, opening up countless opportunities. These opportunities stand out, particularly

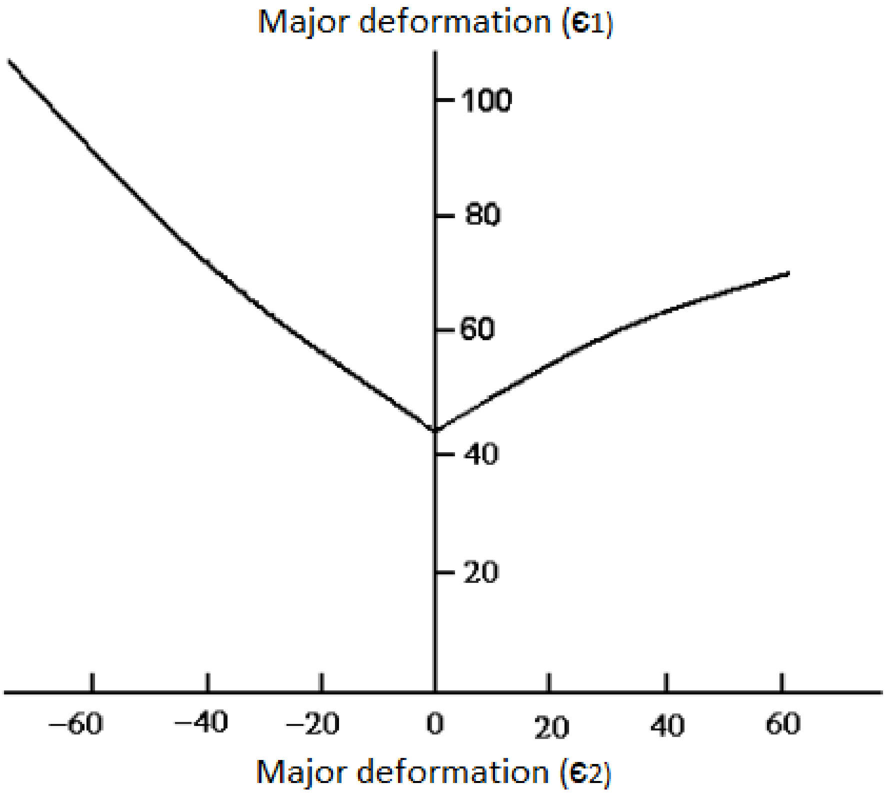

Figure 4 – Conformation Limit Curve - CLC. Source: (TEPEDINO, 2014)

in applications that make use of sophisticated signal processing tools, which can efficiently extract information that would be practically inaccessible by manual means.

In particular, data modeling tools have gained prominence and are widely adopted in areas that deal with massive and complex data sets. This is due to its ability to discover statistical patterns in data and use these patterns to make predictions or make informed decisions (AL-JARRAH et al., 2015).

These data modeling tools have proven to be indispensable for dealing with the vast amount of information currently available. They allow you to extract valuable insights, identify correlations and trends, and assist in making informed decisions across a wide variety of industries (BONACCORSO, 2017).

As technologies continue to evolve, data modeling tools are expected to become even more sophisticated and efficient, allowing us to explore the potential of data further and drive innovation in many areas.

In the context of data modeling through ML, two main types stand out: supervised and unsupervised. The fundamental distinction between them lies in the presence or absence of a reference during training, that is, labels that associate the inputs with the desired outputs. While in supervised learning, there is this explicit guidance, in unsupervised learning, the training aims to identify patterns autonomously without relying on predefined labels.

2.2.1 SUPERVISED LEARNING

Supervised learning is a modeling methodology that consists of training a model from labeled examples, where each example consists of a set of attributes (input data) and a known response (labels). The objective is to learn a function capable of mapping the attributes to the correct answers, thus allowing to make predictions on new data not previously seen. With an adequate training stage, the model is adjusted to the patterns present in the training data. They allow the model to make accurate predictions for unlabeled data in the inference phase due to the parameters and calculations adjusted during training. The model learns to generalize the relationships between attributes and labels, allowing it to make accurate estimates even on new data (MAHESH, 2020).

Figure 5 demonstrates the steps in a summarized form of a supervised learning algorithm, where, initially, the collection, cleaning, and pre-processing of the data is carried out. Once the data are prepared, the selection of an appropriate modeling technique can be defined, considering the type of problem and the specific objectives. At this stage, model parameters are tuned to optimize its performance. It is essential to strike a balance between model complexity and generalizability. Finally, the evaluation of the obtained results is carried out to infer the performance of the model. By using clean and representative data and building a robust model, good performance can be achieved (BONACCORSO, 2017).

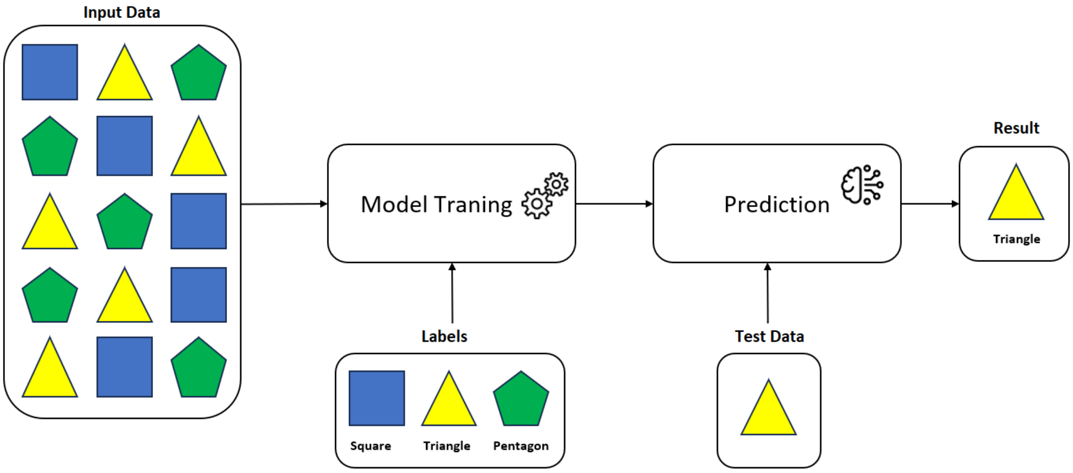

Figure 5 – Illustration of the supervised learning process for training and inference stage. Source: Own authorship

The modeling of multi-objective problems involves several approaches. Among them, some intrinsic characteristics of decision tree algorithms stand out in this context, such as speed in training and greater interpretability of the results obtained. In the following topics, some of these algorithms will be discussed in detail.

2.2.2 DECISION TREE

Decision tree algorithms are ML techniques used to solve classification and regression problems. These algorithms create a tree-like structure to represent a set of decisions. The decision tree is divided into root nodes and child nodes, where, at each step, the algorithm selects the attribute that best divides the data based on commonly used criteria, information gain, Gini index, or entropy. Based on the selected split attribute, child nodes are created, where, the data is split into smaller subsets. Each subset corresponds to a specific value of the split attribute. This process is repeated until a stopping criterion is met. This criterion can be a tree depth limit, minimum number of samples in a node, or any other user-defined criteria (BONACCORSO, 2017).

There are also variants of decision tree algorithms, such as Random Forest and Extra Tree, which use multiple trees to obtain more robust results and improve performance, in addition to randomness factors that help to define a more robust model.

2.2.2.1 RANDOM FOREST AND EXTRATREE REGRESSOR

Random Forest (RF) is a ML algorithm based on the Decision Tree technique. It can be used for both classification and regression problems, being able to deal with categorical and continuous predictor variables. This algorithm has several attractive features that make it a popular implementation choice (CUTLER; CUTLER; STEVENS, 2012).

One of the main advantages of RF is its ability to perform multiclass classification efficiently. Furthermore, it stands out for its speed in both the training and the prediction phase, making it a practical choice in several applications mainly involving industrial applications (BREIMAN, 2001).

Additionally, the RF demonstrates competence in dealing with high-dimensional problems in which there are a large number of predictor variables. It is capable of automatically selecting the most relevant variables, making it an attractive option for addressing complex challenges that involve an extensive set of attributes.

Figure 6 presents a general representation of the RF, an algorithm that combines several decision trees to make a final decision. At the beginning, the root node of the tree encompasses all predictor data. Nodes that are not split are called terminal nodes, forming the final partitions of the predictor space. Each non-terminal node splits into two child nodes, one on the left and one on the right, based on the value of one of the predictor variables. In the case of a continuous predictor variable, the division is determined by a cutoff point. Data whose predictor value is less than the cutoff point are directed to the left node, while the rest are directed to the correct node.

In addition to creating several trees, the algorithm uses random selection to perform node divisions. This process makes the Random Forest more robust against noise when compared to traditional decisions.

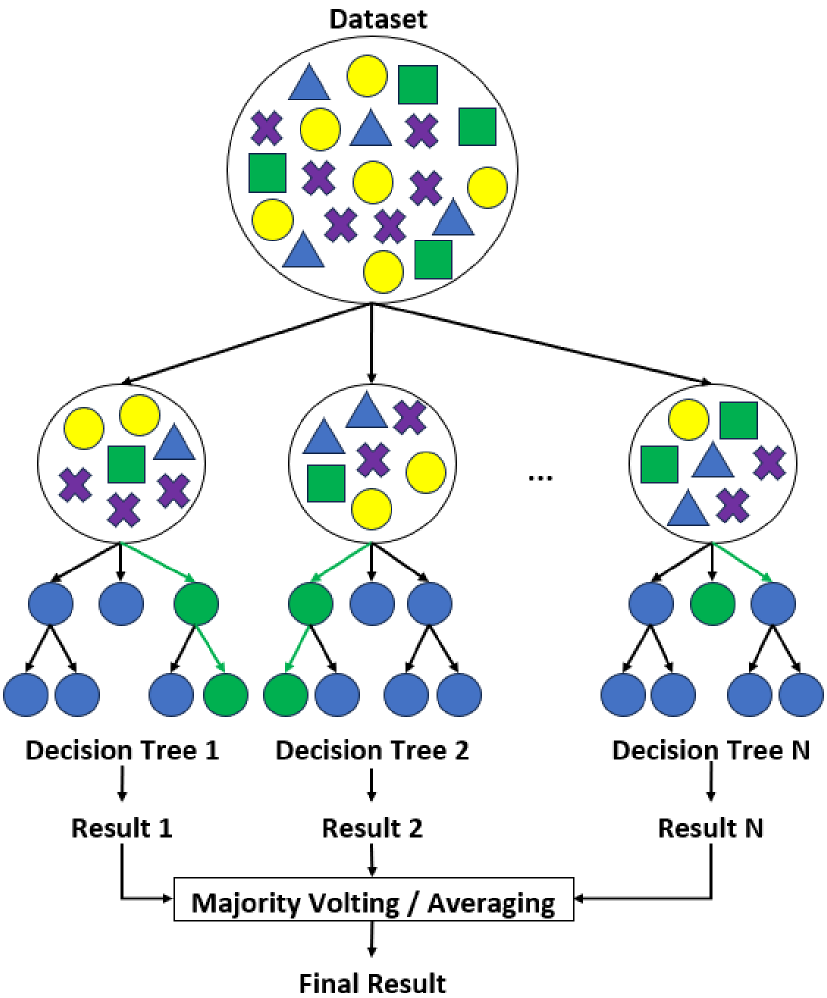

Figure 6 – Illustration of random forest trees. Source: Own authorship.

During the prediction step, individual trees vote on their predictions, and the most frequent class or value is selected as the final prediction. This is called a prediction combination. These randomization and prediction combination strategies make RF a powerful and effective algorithm for dealing with ML problems, offering greater accuracy and addressing the adverse effects of overfitting.

The Extra Tree (ET) algorithm is also based on decision trees and is considered a variation of the RF algorithm. Although both algorithms follow a similar approach, there are essential differences between them. The main difference between ET and RF algorithms is in the construction of the individual trees. In RF, each tree is built from a random sample of the training data (with replacement) and takes into account a random selection of features for each node split. In ET, trees are built using the complete training set and a random selection of attributes for each node split (GEURTS; ERNST; WEHENKEL, 2006).

This difference implies that ET tends to be even more random than RF. For example, in RF, each tree is built from a random sample, but the algorithm still evaluates several options to select the best split at each node. In ET, as all trees are built using the entire training set, there is no such additional evaluation of splitting options. This makes ET faster in training but may result in a slight loss in accuracy compared to RF. However, when properly tuned, ET can provide a better characterization of the problem at hand (GEURTS; ERNST; WEHENKEL, 2006).

2.3 MULTI-OBJECTIVE OPTIMIZATION

Optimization is a process that aims to find one or more feasible solutions that correspond to the extreme values of one or more objectives. It is performed by comparing different available solutions until no better solution can be found.

In single-objective optimization problems, the task is to find a solution that minimizes or maximizes a single function. However, when more than one conflicting objective is being optimized, we deal with a multi-objective optimization problem. In this case, there are negotiated criteria in decision-making and several objective functions being optimized simultaneously, and such objectives tend to be conflicting.

Multi-objective optimization seeks to find a set of non-dominated solutions, that is, solutions that cannot be improved on all objectives simultaneously. Instead of looking for a single optimal solution, the objective is to provide a set of solutions that enable the decision maker to choose the solution that best meets their needs (DEB, 2011).

In recent years, genetic algorithms have been widely used in solving multi-objective optimization problems due to their high capacity to deal with complex problems. These problems are often encountered in our everyday lives, making genetic algorithms a valuable tool for finding efficient and effective solutions (AMOUZGAR, 2012).

What makes genetic algorithms so attractive is their ability to search for high-quality solutions in a multidimensional search space. They use selection, recombination, and mutation techniques to generate and improve a population of decisive solutions over several generations. This approach mimics the process of natural evolution, where only the most adapted solutions survive and reproduce, producing correspondingly better solutions over time (RIBEIRO, 2016).

2.3.1 PARETO OPTIMALITY

In the context of multi-objective optimization, it is important to understand the concept of domination to identify optimal solutions. Domination is a partial ordering relation between two solutions (DEB, 2011).

A solution is considered dominant over another solution if two conditions are met:

- The solution x^1 is not worse than x^2 in all considered objectives.
- Solution x^1 is strictly better than x^2 in at least one objective.

Domination allows you to identify solutions that offer different tradeoffs between objectives. The non-dominated solutions form the so-called Pareto front, which represents the best solutions in terms of options among the considered criteria. Each point on the Pareto front corresponds to a solution that cannot be improved on one criterion without getting worse on another criterion.

By exploring the Pareto front, decision-makers can analyze the different options available and choose the solution that best suits their preferences and constraints. This allows you to make informed decisions, considering trade-offs between objectives and selecting the solution that best fits the specific needs of the problem at hand.

Figure 7a illustrates the pairwise comparison between a set of solutions, using the aforementioned definition to determine whether one point dominates another. In this way, it is possible to establish which points are dominated and which are not. In the example shown in the figure, points 3, 5, and 6 are identified as non-dominated points.

Figure 7b illustrates the resulting Pareto front for the set of six solutions. This Pareto front represents the best non-dominated solutions, where each point represents a compromise between the different objectives considered. By exploring this front, decision-makers can analyze the available options and choose the solution that best fits their specific preferences and constraints.

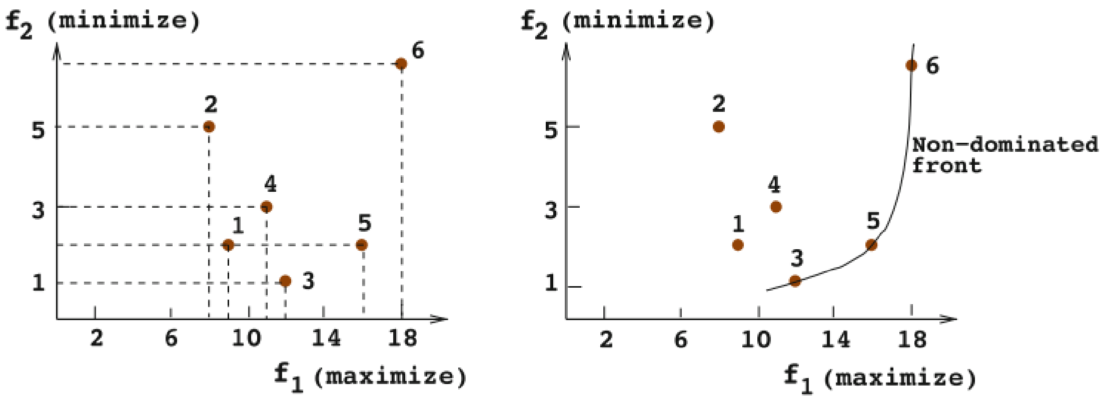

Figure 7 – Illustration of a set of points with their Pareto front. Source: (DEB, 2011).

2.3.2 MULTI-OBJECTIVE EVOLUTIONARY ALGORITHMS

Evolutionary Algorithms (EAs) are inspired by the biological model of evolution and natural selection, allowing to simulate an environment and the biological pressures under which potential solutions can evolve. Through adaptation and survival over several generations, solution candidates can be optimized towards an approximate solution, shaped according to the specific parameters and constraints of the problem at hand (BINITHA; SATHYA et al., 2012).

In many real-world problems, it is necessary to simultaneously optimize multiple objectives, which can sometimes be conflicting. When the complexity of the problem increases considerably or when it is difficult to formalize it precisely, the search for an exact solution can become computationally expensive or even unfeasible (VIKHAR, 2016). In those scenarios, where an approximate solution is sufficient, multi-objective Evolutionary Algorithms (MOEAs) can be effective in obtaining reasonable solutions.

There are several variations of EAs, but Genetic Algorithms (GA) stand out due to their efficiency in solving problems with an ample search space and little prior knowledge, especially when analytical equations are not readily available to address the problem. Therefore, a more detailed description of the basic GA procedure and operators will be provided, along with an explanation of NSGA-II, which demonstrates skills in handling multi-objective problems.

2.3.2.1 NONDOMINATED SORTING GENETIC ALGORITHM II (NSGA-II)

In applications of multi-objective Evolutionary Algorithms (MOEAs), NSGAII stands out for its solid performance when compared to other algorithms in the same category. Studies carried out by authors such as (RAHIMI et al., 2023) and (ISHIBUCHI et al., 2016) provide comprehensive analyses of the relative capabilities of these algorithms.

One of the main features of NSGA-II is that it follows the general scheme of a genetic algorithm but with modifications for mating selection and survival. The first step of the algorithm is to perform the selection concerning the non-dominated fronts. During this process, there may be a need to divide a front, as not all individuals can survive. In this division of fronts, solutions are selected based on the cluster distance (DEB et al., 2002).

The crowding distance is calculated using the Manhattan Distance in objective space. However, it is desirable to maintain the extreme points in each generation. The algorithm implements an elitist selection strategy, in which the best chromosomes from the current generation are preserved and passed on to the next generation. This approach guarantees that the quality of the best solution found increases progressively over time (DEB et al., 2002).

By keeping the best chromosomes, the algorithm benefits from the valuable information contained in these high-performance solutions. These chromosomes are considered the fittest and represent high-quality solutions in the search space. By preserving them, the algorithm prevents the loss of these promising solutions, ensuring that they have the opportunity to influence and improve the quality of subsequent generations (DEB et al., 2002).

These changes to NSGA-II aim to improve the convergence and diversity of solutions across generations. By using a selection based on non-dominated fronts and cluster distances, the algorithm promotes the maintenance of solutions and the preservation of diversity in populations. Tournament selection also intensifies selective pressure, favoring more promising solutions in each generation.

2.3.3 SURROGATE MODEL OPTIMIZATION

A surrogate model, also known as an emulator model or approximation model, is a technique widely used in data analysis and process optimization. These models aim to represent complex systems or processes, allowing the prediction of the system's behavior under different conditions or the optimization of the process without the need to resort to the natural system repeatedly (QUEIPO et al., 2005).

Building a surrogate model involves using statistical or ML techniques to create a behavioral approximation of the system. This is especially useful when the system in question is complex to measure or calculate directly or your process is complicated to reproduce due to time constraints.

Surrogate models are also applicable to multi-objective problems, in which optimization seeks to find a set of solutions that uniformly meet the requirements of all objectives. There are several approaches to building surrogate models for multi-objective problems, such as the one used in the work (DASARI; CHEDDAD; ANDERSSON, 2019), which used models based on regression trees to approximate the actual model. These types of models offer several benefits, such as reducing the time and costs required to perform experiments, ease of modeling, the ability to perform experiments under different conditions, and the ability to analyze and optimize without affecting the natural system.

The selection of evaluation methods for a substitute model must be guided by the particularities of the problem under analysis, taking into account its specific nature. These methods can be applied independently or combined to provide a comprehensive assessment of the model from several aspects. The effectiveness of a surrogate model is intrinsically linked to its ability to predict the behavior of the original system accurately.

2.3.4 OPTIMIZATION METRICS

Evaluating the quality of a Pareto front approximation is a complex challenge due to several factors, such as proximity to the actual Pareto front and adequate coverage in the objective space. Given these complexities, several metrics, known as performance indicators, have been developed to measure the quality of a Pareto front approximation. These metrics assign a score based on different evaluation criteria (LI; YAO,).

The Pareto front approximation aims to find a set of non-dominated solutions that represent a balance between the objectives of the problem. Assessing the quality of this approximation involves determining how close it is to the true Pareto frontier and how adequately it covers the objective space.

The evaluation metrics mainly consider two aspects: convergence and diversity. Convergence refers to how close the approximation is to the true Pareto frontier, evaluating the distance between the found solutions and the known optimal solutions. Diversity evaluates the distribution of solutions along the approximate boundary, ensuring comprehensive coverage of the objective space.

The proper evaluation of these metrics allows a clear understanding of the performance of the algorithm in relation to the optimization objectives, helping in the selection and adjustment of parameters of the genetic algorithm.

In the following topics, details of the metrics used in this work will be presented to help understanding the effectiveness of GA in finding optimal or approximate solutions for a given problem.

2.3.4.1 HYPERVOLUME INDICATOR

The hypervolume (HV) metric is a widely used metric to evaluate the quality of Pareto front approximations in multi-objective problems. It provides a quantitative measure of the extent of coverage of the objective space by a given approximate solution, as illustrated in Figure 8.

The idea behind the HV calculation is that a good approximation of the Pareto front should encompass a region of objective space with a significant volume. The larger the HV, the better the approximation quality.

The HV indicator is a metric used to measure the quality of a Pareto front approximation in multi-objective problems. It represents the volume of the objective space that is dominated by the approximate solution of the Pareto front S, and is delimited su-

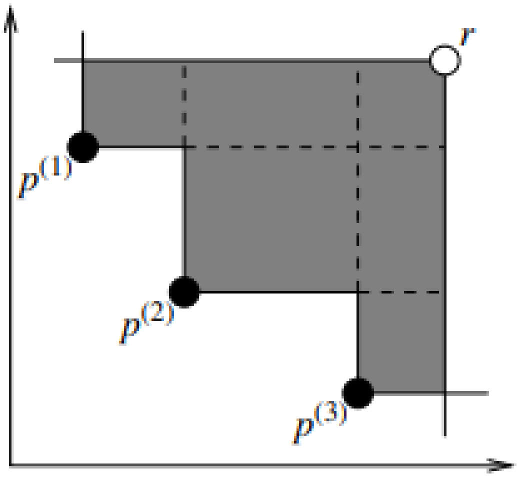

Figure 8 – Example of a graphical HV calculation. Source: (FONSECA; PAQUETE; LÓPEZ-IBÁNEZ, 2006).

periorly by a reference point r $\in \mathbb{R}^m$, where for all z $\in$ S, z $\prec$ r. The calculation of the HV is performed through Equation 2.2.

$$HV(S, r) = \lambda m(\bigcup_{z \in S} [z; r]) \tag{2.2}$$

The HV calculation used in this work is based on variant three of the algorithm proposed by (FONSECA; PAQUETE; LÓPEZ-IBÁNEZ, 2006).

2.3.4.2 GENOTYPIC DIVERSITY MEASURE

The Genotypic Diversity Measure (GDM) is a metric used to assess the genetic diversity of a GA population. This metric is applied specifically at the level of genotypes of individuals in the population.

Genotypic diversity is an important measure to assess the exploration and search capacity of the GA. The more genotypic diversity is in the population, the greater the variety of solutions represented, which can lead to a broader exploration of the search space.

This metric is often used to monitor genotypic diversity across GA generations. A decline in genotypic diversity over time may indicate premature convergence of the algorithm, where the population is restricted to a specific region of the search space, resulting in suboptimal solutions. On the other hand, a high genotypic diversity indicates a broader exploration of the search space and a greater chance of finding optimal or close to optimal solutions.

In this work, the normalized pairwise diversity measurement (DNPW), which is widely discussed in the literature and different codifications, will be adopted due to its remarkable capabilities regarding convergence, stability, and insensitivity to dimension and outliers. This measure, described by Equation 2.3, is effective in a variety of contexts, making it a solid choice for this study (CORRIVEAU et al., 2012).

$$D_{PW}^{N} = \frac{\frac{2}{POP_{size}(POP_{size}-1)} \sum_{i=2}^{POP_{size}} \sum_{j=1}^{i-1} \sqrt{\sum_{k=1}^{POP_{size}} (x_{i,k} - x_{j,k})^2}}{NMDF} \tag{2.3}$$

where POPSIZE is the population size, i and j are individuals, k is the gene locus and NMDF is a normalization factor.

3 SYSTEMATIC LITERATURE REVIEW

The use of data-driven models to predict the best process parameters and stamping of metal materials is not new. However, there is a paucity of documentation available on the topic. Surprisingly, there is no comprehensive systematic literature review (SLR) that addresses this issue. This gap can be attributed to several possible reasons, such as the lack of a general perception of the need to systematically organize the existing body of knowledge or the difficulties encountered in solving this challenging multi-objective problem.

Thus, a systematic review of the literature was carried out until November 2023 to address the project of predicting and optimizing the automotive stamping process using machine learning techniques and multi-objective evolutionary algorithms.

3.1 RESEARCH METHODOLOGY

With the main objective of offering a comprehensive view of the subject and identifying its characteristics and trends, the defined research objective is the following:

- Investigate data-driven techniques used to optimize the automotive stamping process or suggest the best process parameters and materials to employ.

The following questions to assist in the understanding and synthesis of the research objective:

1. What are the most used techniques?
2. How is data-driven modeling relevant compared to empirical analytical models?
3. Difficulties related to the stamping process?

3.2 PLANNING THE REVIEW

With the proposed research objective and the elaborated questions, the SLR planning stage began using the Prisma methology (LIBERATI et al., 2009), in which the search phrase was established. The search phrase was used in the abstract and title search fields, whenever these options were available.

After several iterations, it was decided to use five different keywords, to require only seven logical operators to compose the search phrase, making it compatible with all selected Academic Search Engines (ASE). The search phrase was entered into the abstract and title search fields when these options were available. Finally, the search phase was defined as follows:

With the proposed research objective and the questions elaborated, the RSL planning stage began using the Prisma methodology, in which the search phrase was established. The search phrase was used in the abstract and title search fields whenever these options were available.

- (“Automotive Stamping” OR "Metal Stamping“ OR "Steel Stamping") AND (“Optimization” OR "Improvement*")

Considering the multidisciplinary nature of the topic, it was decided to select the areas of Engineering, Materials Science, and Computer Science as the main focus of this research. The chosen search engines are IEEXplore, Science Direct, Web of Science, and Scopus.

3.3 THE REVIEW PROCESS

After defining the search engines and search phrase, the next step in conducting a systematic research review is to outline the objective, inclusion, and exclusion criteria used as a practical screen to decide which studies should be considered for the review. The defined criteria are presented below:

- Objective criteria:
  - Year range: 2010 - 2023
  - Type of document: Articles, dissertation, and thesis
  - English language
- Inclusion criteria:
  - Include studies that apply metaheuristics to optimize the steel stamping process
  - Include studies that apply methods or tools to suggest process parameters or materials used for stamping steel
- Exclusion criteria:
  - Remove studies that predict or optimize only factory processes
  - Remove studies that predict properties or optimize non-metallic materials

- Remove studies that aim to present only the influence of parameters in the stamping process

Based on the established criteria, it was defined that, if a study meets any exclusion criteria or does not meet any objective criteria, it should be excluded from the subsequent steps of the systematic review. The criteria were applied in the following order: objective, exclusion and inclusion. After practical screening and quality assessment, the selected articles were counted and presented in Table 1. The "Found" column represents the number of articles returned by the search phrase alone. "Filtered" indicates the number of articles after applying the objective, inclusion and exclusion criteria. Finally, "Selected" shows the number of articles that were chosen for further study.

Table 1 – Table of references or articles, filtered and selected from the search SLR.

ASE	Found	Filtered	Selected
IEEEXplore	13	4	2
Science Direct	32	7	2
Web of Science	48	15	8
Scopus	72	15	2
Others	6	6	6

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

- Model input variables, classified into categories IN1, IN2 and IN3 defined as follows:
  - IN1: Type of stamping
  - IN2: Stamped part
  - IN3: Process parameters or materials used
- Model output variables, that is, the properties being predicted or optimized:
  - The tool used to optimize the process (statistics, machine learning, numerical simulation, among others)
  - Presence of quantitative analysis of results
  - Presence of qualitative analysis of results

3.4 SUMMARY AND DISCUSSIONS

Based on the selected studies, a matrix concept was elaborated, which is presented in Table 2. This table includes relevant information, such as the modeling and optimization techniques used, the input and output variables considered, and the analysis of the results obtained, whether quantitative or qualitative.

Table 2 – Table with a fraction of studies selected from the SLR.

Reference	IN1	IN2	IN3	Output	Simulator	Algorithms	Optimizer	Quant.	Quali.
(YAN et al., 2020)	X	X	X	WRI, FRA	DYNAFORM	ANN	AG		X
(OUJEBBOUR; HABBAL; ELLAIA, 2013)	X	X	X	WRI, FRA	LS-DYNA	SA, SPSA, NBI, NNCM	NSGAII		
(WANG et al., 2018)	X	X	X	WRI, IS, FRA	LS-DYNA	RBF, KRG, SVM	NSGAII	X	X
(INGARAO; Di Lorenzo, 2010)	X	X	X	WRI, IS, FRA	DOE + LS-DYNA	RSM			X
(INGARAO; LORENZO, 2010)	X	X	X	WRI, IS, FRA	DOE + LS-DYNA	RSM			X
(KIM; KIM; LEE, 2023)	X	X	X	WRI, IS, FRA	DOE + Altair Hyperform	RSM			X
(HAMDAOUI et al., 2015)	X		X	WRI, IS, FRA	LS-DYNA	KRG	NSGAII	X	
(DANG; LAFON; LABERGERE, 2017)	X	X	X	WRI, IS	DOE + LS-DYNA	RBF			
(XIE, 2011)	X		X	WRI, FRA	DOE + LS-DYNA	KRG		X	X
(SU et al., 2017)	X	X	X	WRI, FRA	LS-DYNA				X
(XIE, 2011)	X	X	X	FRA	DOE + LS-DYNA	KRG		X	X
(BHUYAN et al., 2015)	X	X	X	FRA	DOE + LS-DYNA				X
(ESENER; ERCAN; FIRAT, 2014)	X	X	X	WRI, IS, FRA	DOE + DYNAFORM + CATIA				X
(HU et al., 2018)	X	X	X	WRI, FRA	AutoForm				X
(MA; HUANG et al., 2014)	X	X	X	WRI, FRA	DOE + Dynaform				X
(RAFIZADEH et al., 2017)	X	X	X	WRI	DOE + ABAQUS	RSM, RNA		X	X
(SINGH; GUPTA, 2010)	X	X	X	WRI	LS-DYNA	SVR, RNA			X
(CUI et al., 2020)	X	X	X	FRA	UG Software	RSM	NSGAII	X	X
(XIE et al., 2022)	X	X	X	WRI	Autoform	RNA	AG	X	X
Present work	X	X	X	WRI, IS, FRA	Autoform	ET	NSGAII	X	X

3.4.1 DIVERSITY VARIABLES

As shown in Table 2, there is a predominance of research that employs stamping parameters and includes additional information about the part or process. Although most of the studies focus mainly on process and material-related variations, other information such as type of stamping, type of automotive part, and study materials are provided to enrich the understanding of the elements involved in the research, as in the works (YAN et al., 2020), (ESENER; ERCAN; FIRAT, 2014) and (BHUYAN et al., 2015).

Concerning input variables, which encompass both process factors and those related to materials, their configuration varies according to the specific focus of the study and the type of part to be stamped. This is due to the unique characteristics and varied purposes of each automotive component, which are adjusted according to their particular application. Therefore, it is crucial to consider that there are more secure configurations for each application context.

Furthermore, the objective of the research can vary significantly. In some cases, such as the study (OUJEBBOUR; HABBAL; ELLAIA, 2013), the emphasis is on identifying the ideal domains for each process variable. These domains will serve as a basis for the adequate definition of manufacturing parameters and the construction of the components involved in the stamping process, thus allowing the analysis of the impact of manufacturing parameters on the roof of a vehicle.

On the other hand, in research such as that carried out by Wang et al. in (WANG et al., 2018), which aims to define material-related parameters, the main focus lies on choosing the appropriate domains to guide the selection of the ideal material for the specific part in question. In this context, the study focuses on determining the ideal ranges for both process variables and materials. These ranges serve as a basis not only for the optimization of manufacturing processes but also for the selection of the material that will contribute to the creation of a failure-free part with high rigidity, especially in collision situations.

It is important to highlight the significant diversity in the approaches and restrictions adopted in each study, since they may have different objectives, consequently resulting in different definitions of inputs and sets of objectives to be achieved.

3.4.2 OPTIMIZATION MODELS

In the literature analysis, several models were identified that vary in complexity and approach, three of which stand out. The first follows a relatively simple approach, as evidenced in the studies by (SU et al., 2017) and (HU et al., 2018), where Design of Experiments (DOE) methods are used combined with numerical simulation software.

However, these methods often result in limited optimizations in the selection of the best stamping parameters due to the restrictions imposed by the need to evaluate a large number of possibilities empirically.

The second approach involves integrating the Response Surface Methodology (RSM) or applying the Kriging algorithm (KRG) to the Designed Experiments (DOE) model run in numerical simulation software. This methodology is widely adopted in studies related to the research topic, such as (INGARAO; Di Lorenzo, 2010), (KIM; KIM; LEE, 2023), and (XIE, 2011).

In the case of RSM, it focuses on analyzing the interactions between input variables derived from DOE in simulation software, aiming to find the ideal combination of parameters to optimize a specific objective function.

In the context of the KRG algorithm, the emphasis is on the interpolation of sample data (DOE) provided by the simulator, taking advantage of a technique that models the spatial correlation between sample points, aiming to estimate values in non-simulated locations based only on the information collected.

However, it is essential to highlight that adopting this approach (whether using RSM or KRG) results in more substantial margins of error in predicting ideal stamping parameters due to limited data availability and the restrictions inherent to the tool used. Despite of theses challenges, this methodology continues to be widely used in several studies, many of which present notable and significant results.

Finally, we identified a more advanced third approach, inspired by the work of author Maomao Cui, with the study (CUI et al., 2020). This methodology served as the basis for the approach developed in this study. The strategy adopted includes the combination of finite element simulation with the RSM algorithms and the NSGA-II multi-objective genetic algorithm to explore the optimal stamping parameters.

In summary, three distinct approaches were identified, and the most complex approach stood out for its capacity and efficiency in optimizing ideal stamping parameters. However, it is essential to highlight that each project has its limitations that deserve consideration, whether due to the availability of tools, such as the use of numerical simulators, or the intrinsic nature of the objectives of the proposed study.

3.4.3 QUANTITATIVE ANALYSIS

In applications involving black box models, it is common practice to use metrics such as Mean Squared Error (MSE), Coefficient of Determination (R^2), Mean Absolute Error (MAE), and other performance measures. Although these metrics are primarily valuable from a practical point of view during the development process, they alone offer little contribution when it comes to explain or evaluate the performance of models in different contexts. However, many studies still choose to use them solely to validate the performance of their models.

Some works, such as (HAMDAOUI et al., 2015), go further and explore the relationship between input and output variables, deepening the understanding of the experimental plan. This makes it possible to anticipate decisions even before creating a representative model.

In applications involving genetic algorithms, it is expected to direct metrics to evaluate two main objectives: convergence and diversity. The work (BHUYAN et al., 2015) is one of the few that presents an approach to evaluating optimization convergence across generations.

Notably, most studies focus on understanding the phenomena resulting from metal-mechanical processes, with little emphasis on the quantitative analysis of the algorithms used. This raises doubts about the interpretability on the part of the authors concerning the tools used.

3.4.4 QUALITATIVE ANALYSIS

The evaluation of the results achieved in the optimization proposal, in comparison with theoretical and practical bases, often involves qualitative analyses. In studies focused on the topic of optimizing the stamping process, validation of the best parameters is usually conducted using a numerical simulator, in which the results are compared and the prediction error is calculated.

In the work carried out by Giuseppe Ingarao (INGARAO; Di Lorenzo, 2010), after the optimization process, the author examines the resulting Pareto front and identifies the individuals that meet the desired validation criteria. It then proceeds to re-execute the simulation process, allowing an assessment of the effectiveness and accuracy of the proposed model.

In contrast, in the study (KIM; KIM; LEE, 2023), a different approach was used. In addition to the numerical simulation steps, the author stamped a laboratory part for a practical understanding of the results proposed by the models, comparing them with results obtained in laboratory experiments.

Most of the selected works adopt a qualitative assessment, often using simulation software and practical experimental principles to compare and evaluate results in the area of metal-mechanical phenomena.

3.4.5 CONSIDERATIONS OF THE SYSTEMATIC LITERATURE REVIEW

After an in-depth review of the literature, it becomes evident that modeling and predicting parameters in the stamping process of materials for automotive parts is not only possible but also highly promising. However, as the number of objectives to be optimized increases, such as WRI, IS, and FRA, the challenge naturally becomes more complex, demanding increasingly sophisticated modeling approaches. Given that the parameters and configurations of simulation and stamping machines are often empirically determined and susceptible to errors and anomalies, a deep understanding of metallurgical principles becomes essential to establish the limits of these models.

In this context, the development of tools that help optimize and accelerate the process of defining stamping parameters has become a growing trend. This is mainly due to the positive results that have been presented, demonstrating its effectiveness.

Among the studies researched, a common point in all of them is the wide adoption of numerical simulation tools to generate synthetic data. This focus has been driven mainly by the costs associated with acquiring accurate data. Although the time required to carry out the simulations is considerable, it remains substantially less than the time spent executing the natural process in a factory. This not only saves time but also opens up ample opportunities for experimentation and refining the techniques used.

Among the techniques frequently used, Artificial Neural Networks (ANNs) and Response Surface Methodologies (RSM) for modeling stand out, as well as Genetic Algorithms for optimization. However, the particularities of different datasets can affect the modeling process in different ways. Therefore, in the absence of a consolidated data set as a reference, it becomes unfeasible to determine which algorithm represents the state of the art based solely on the performance of its metrics. Instead, the state of the art should be defined based on robustness, meticulous evaluation, results achieved, and compliance with relevant theoretical principles in the field of metallurgy.

Concerning quantitative assessments, there is a need for more approaches of this type, which results in algorithmic models that often need to be improved in representativeness. This, in turn, generates growing concern regarding the methodologies used.

4 PROPOSED APPROACH

In this chapter, we will present the proposed model to improve the optimization of the automotive stamping process. To promote well understanding, we illustrate the steps in the flowchart as shown in Figure 9. In addition, each step will be described in detail throughout this chapter.

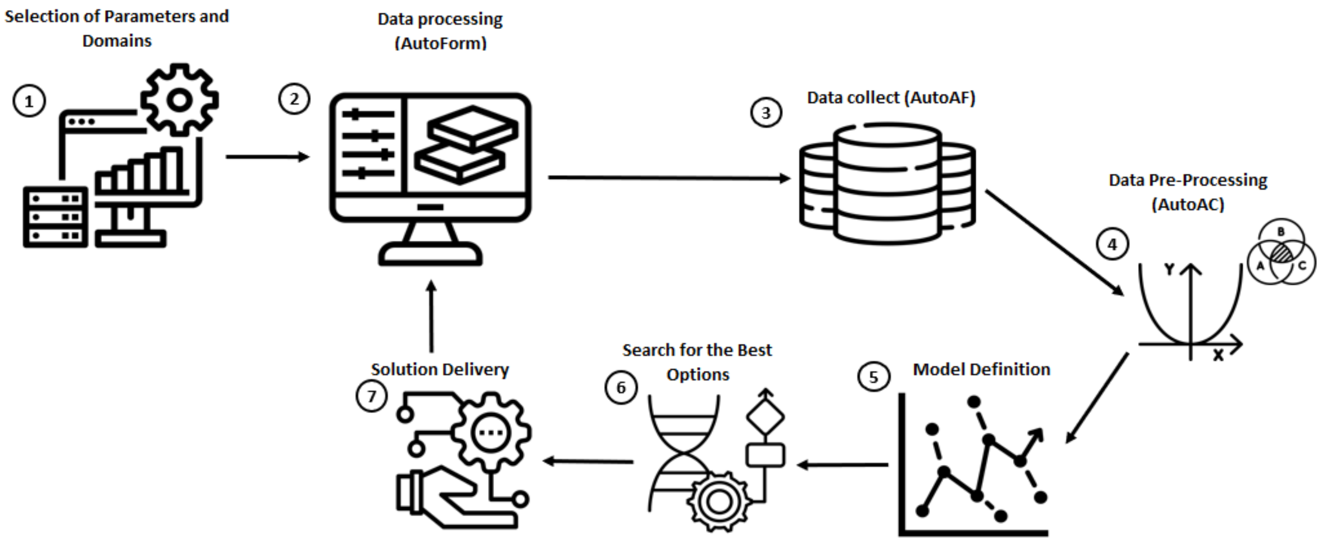

Figure 9 – Flowchart of the proposed model. Source: Own authorship.

4.1 SELECTION OF PARAMETERS AND DOMAINS

The first step consists of defining the parameters and domain selection process for designing the experiments in the AutoForm[1] simulation software. This software is responsible for modeling and simulating the behavior of automotive parts, generating the necessary database for the entire process. This stage is carried out by a specialist in metallurgy, who uses his skills to identify the critical variables in the stamping process of the desired part. The precise definition of these variables has a direct impact on the evaluation of the regression model. The expert ideally creates a plan with a series of experiments that uniformly span the search space of process and product variables.

To facilitate this comprehensive configuration, the AutoForm-Sigma tool is used. This tool aims to simplify the process of defining parameters and domains, creating a set of simulation possibilities based on the specified limits and the number of simulations requested. In addition to this, the tool allows the parallel execution of multiple simulations, thus accelerating data generation.

[1] AutoForm: <www.autoform.com>

4.2 DATA PROCESSING AND DATA COLLECT

After defining the appropriate parameters and domains, the process of the simulations is carried out. This step is critical for the project, considering the time required to carry out all the simulations. In steps 1 and 2 in the flowchart, as shown in Figure 9, it is essential to find a balance between the execution time and the desired performance of the model (Model Definition Step 5). Carrying out a large number of simulations generates significant operational costs for the company, especially when the simulated part is large, generating excessive execution time. In contrast, for the proposed model, which is data-based, the amount of available data directly impacts all subsequent steps (creation of the surrogate model and optimization using multi-objective evolutionary algorithms).

It is essential to highlight that the balance in the number of planned simulations will directly reflect on the subsequent stages of this project. Therefore, finding this balance is considered one of the project's challenges.

After completing the simulations planned in steps 1 and 2, it is necessary to extract the generated data, due to the limitations of the AutoForm software in relation to integration with other systems. In order to obtain data more efficiently, the AutoAF software was developed. AutoAF is a program developed in Python language, designed to carry out the automated export of simulation data.

AutoAF is software developed in Python language to meet the demands of large volumes of data and numerous simulations, which require an excessively long time for a specialist to perform the extraction manually. However, AutoAF significantly reduces the time required for this task, providing greater efficiency and freeing the specialist to focus on other analysis and interpretation activities of the data obtained.

By automating data extraction, AutoAF makes the process more agile, accurate, and reliable, ensuring that all simulation results are correctly collected and usable in subsequent stages of the project.

4.3 DATA PRE-PROCESSING

The data pre-processing process consists of two steps: area calculation and exploratory data analysis. This step is necessary due to the nature of the data exported from the AutoForm simulator, which consists of files containing the mesh resulting from the simulation, that is, values of nodes and simulation elements. The following topics describe in more detail the steps required to process this data and the expected result.

4.3.1 AREA CALCULATION

The AutoForm software provides a simulation database, consisting of two files for each simulation. These files contain information about the elements and their respective nodes, represented by three-dimensional spatial coordinates, as illustrated in Figure 10.

The nodes play the role of reference points in which the physical properties or variables of the problem are calculated and stored, in this case, their three-dimensional location demonstrates their spatial position referring to the base element. Depending on the nature of the problem, the elements can have different shapes and sizes. In the context of this simulation, they have associated attributes to allow the determination of their location in the CLC, such as the minimum and maximum strain applied to the element.

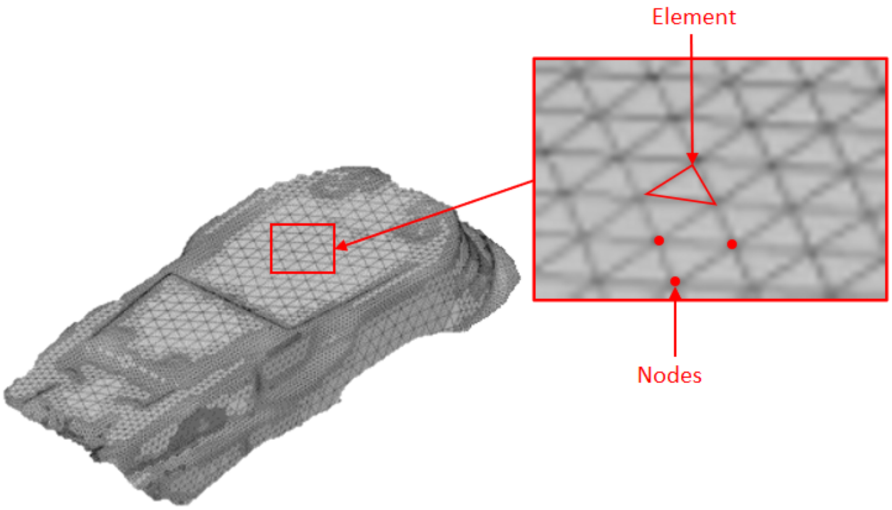

Figure 10 – Nodes and elements in a mesh. Source: Own authorship.

The flowchart shown in Figure 11 demonstrates the necessary steps to perform the area calculation, in which the final objective is to determine the area percentage of each objective (WRI, IS, and FRA).

The area calculation process starts with the organization of the elements, where the nodes and strain limits of each element are found, linking the spatial location to the base element.

The next step consists of calculating the area of each array element employing a vector function, in which the output will be the total area of the element. This output will be used in the following steps to determine the total area and the specific area of each objective.

After determining the area of the element, it is verified if it meets the criteria of one of the objectives, that is, if it meets the criteria of Equations 4.1, 4.2 and 4.3. If

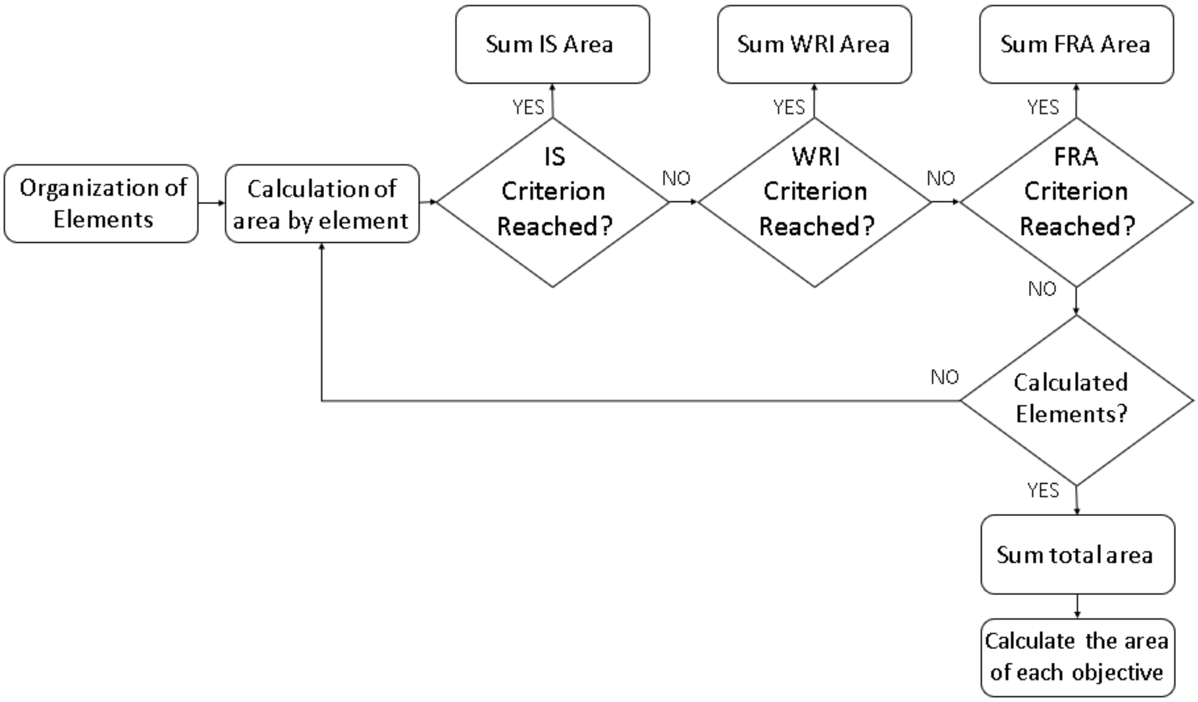

Figure 11 – Area calculation flowchart. Source: Own authorship.

one of the conditions is true, the area of the element is added to the total of the chosen objective.

$$F_{wri} = (\varepsilon_2 < -0.00995) \wedge (\varepsilon_1 > -\tan(\frac{45 * \Pi}{180}) * \varepsilon_2) \tag{4.1}$$

$$\begin{gathered} F_{is} = (\varepsilon_2 \geq -\ln(1.01) \wedge (\varepsilon_1 < (-\varepsilon_2 - \ln *(0.98)))\vee \\ (\varepsilon_1 < (-\varepsilon_2 * -\ln(0.98) \wedge -\ln *\varepsilon_1 > ((-\varepsilon_2) * (\frac{1 + Rav}{Rav}))) \end{gathered} \tag{4.2}$$

$$F_{fra} = (\varepsilon_1 > Eq.Line) \tag{4.3}$$

where, ε_1 and ε_2 are the major and minor strain, Rav is the average R material parameter, and $Eq.Line$ is the equation of the curve line conformation limit test (CLC).

The mentioned checks are performed for all simulation elements. Upon completion of the scan, the total area of the simulated part is summed, and then the area percentage is calculated for each objective, based on the previously summed areas for WRI, IS, and FRA.

4.3.2 EXPLORATORY DATA ANALYSIS

After structuring the database, an exploratory analysis is carried out, using visualizations and correlation analyses of the variables using the Spearman (SPEARMAN, 1961) statistical method, as shown in Figure 12. The objective is to use a visual and statistical tool to evaluate whether any of the parameters defined in the first stage of the process presents a low correlation with the other variables (inputs and outputs). If any of the parameters present a low correlation, the selection of parameters and domains must be reevaluated, as this variable has little impact on the model results and may even significantly impair the quality of the creation of the surrogate model.

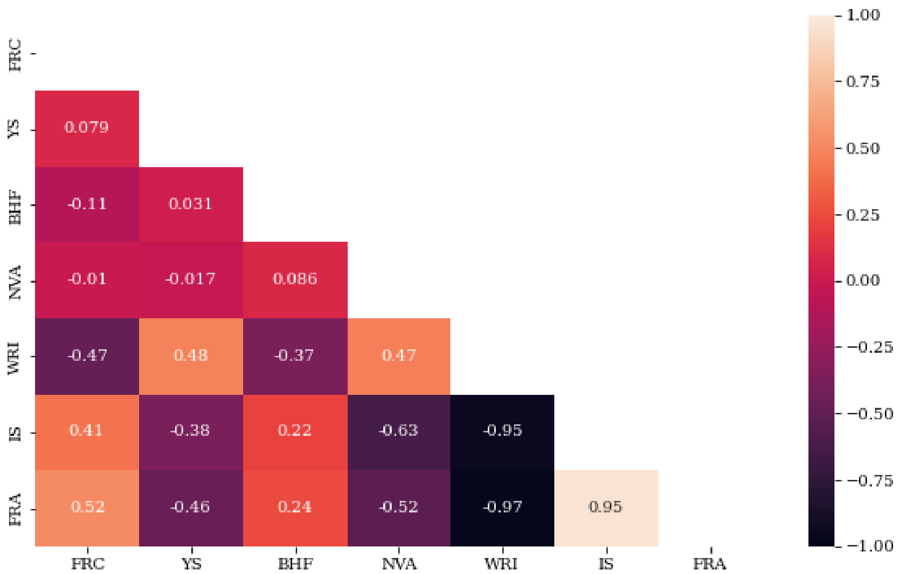

Figure 12 – Exploratory analysis using data correlation visualization. Source: Own authorship.

This visualization aims to support the user's decision-making about moving to the next step (creation of the replacement model) or reevaluating the parameters and domains, redoing the previous steps, including running new simulations.

To evaluate the visualization of the correlation analysis, which is represented by a heat map, it is necessary to observe the intensity of the colors or values in the cells corresponding to the combinations of the analyzed variables. The higher the value, the greater the correlation between the variables. Furthermore, the sign that accompanies the correlation value indicates the direction of that relationship. A direct or positive correlation is indicated by a positive value or a lighter color, while an inverse or negative correlation is indicated by a negative value or a darker color. It is important to emphasize that the direction of the correlation indicates whether the variables move in the same direction (positive correlation) or opposite directions (negative correlation) as their values change.

4.4 MODEL DEFINITION

This step consists of creating the surrogate model using ML tools, to represent the behavior of the addressed problem. Because it is a multi-objective problem, given the need to minimize different objectives in a balanced way, the use of two regressors in this study was defined, both based on regression trees, namely, RF and ET. The use of Neural Networks ended up being discarded at the beginning of the project, based on the results obtained in (BRENDEL, 2021) work.

Thus, one of the stages of this work is to evaluate the performance of the RF and ET regressors. For this, the Root Mean Squared Error (RMSE) metric was used to measure the performance of the models during the search for the best regressor.

The leave-one-out method is used to build the model, due to data quantity limitations. Then, as an auxiliary tool, sensitivity analysis and the impact of input variables are analyzed for the created model.

4.5 SEARCH FOR THE BEST OPTIONS AND SOLUTIONS DELIVERY

After creating a surrogate model that represents the problem in question, it is inserted as a fitness function in the evolutionary algorithm NSGA-II, which will have the challenge of finding the Pareto front, which will represent the set of optimal points or non-dominated solutions in a given multi-objective scenario. Therefore, the objective in this step is to find the input parameters that result in objectives with the lowest percentage of failure, that is, minimized.

In addition to looking for the ideal parameters to optimize the main objectives, experts often consider additional criteria, such as the availability of materials and tooling. Therefore, a wide variability in the provided domains allows for a more comprehensive assessment by experts. In this study, the genetic operators essential for optimization, i.e., mutation and recombination, are not fixed during the execution of the NSGA-II algorithm. Instead, these parameters are adjusted based on deterministic rules, following a real-time control strategy. The definition of the rules is the user's responsibility and is established before the optimization process begins (ALETI; MOSER, 2011).

Although deterministic parameter controls are generally more straightforward and more economically advantageous compared to more complex strategies, their limitation in considering the specific behavior of the NSGA-II, the actual optimization progress, and the information provided by the fitness function can lead to suboptimal results in specific scenarios (PARPINELLI et al., 2019). Furthermore, predicting the number of generations necessary for the convergence of an optimization problem is a challenging task, which can affect the effectiveness of previously established deterministic rules.

In this study, an approach with a deterministic rule was used that adjusts the mutation and recombination rate at each generation, following the Equations 4.4 and 4.5. Initially, the recombination rate starts at 0% and increases progressively, reaching 100% as generations advance. On the other hand, the mutation rate starts at 40% and gradually decreases until it reaches 0%, where g_f refers to the last generation.

Figure 13 graphically illustrates the evolution curves of recombination and mutation rate. This dynamic parameter tuning strategy is adopted to address the uncertainty related to NSGA-II convergence, allowing for more accurate adaptation throughout the optimization process.

$$Recombination\, rate = \frac{\log_{10} g}{\log_{10} g_f} \tag{4.4}$$

$$Mutation\, rate = 0.4. \left(1 - \frac{\log_{10} g}{\log_{10} g_f}\right) \tag{4.5}$$

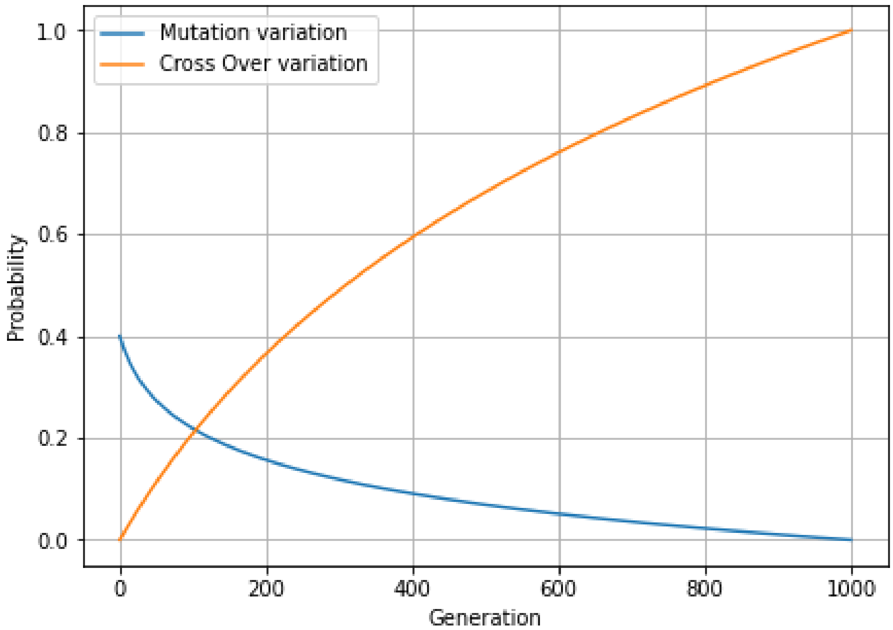

Figure 13 – Evolution of recombination and mutation rate. Source: Own authorship.

To enable an evaluation of the evolutionary algorithm, the HV metric and parallel coordinates are used, which allow the evaluation of the convergence of the algorithm and visual analysis of the diversity of non-dominated solutions.

Finally, the final step presents the most relevant selected (non-dominated) results, including the machine parameters and the percentage of the predicted area for each objective. To verify the accuracy of the proposed model, it is necessary to select one of the generated solution vectors as output and parameterize a new experiment in the AutoForm software with the characteristics of that solution vector. If the validation is not satisfactory, the NSGA-II parameterization is revised and adjusted.

5 RESULTS AND ANALYSIS

In this topic, the results of the proposed methodology are presented, which are divided into two case studies. The first case study is related to a laboratory part called "Cup", while the second covers a genuine car part called "Internal Tailgate".

The case study "Cup" concerns a laboratory solid whose shape is represented in Figure 14. Its unique geometry gives it a wide versatility of applications in the field of mechanical engineering, making it a frequently used element in studies related to materials. The "Cup" plays a crucial role in validating and testing new developments, being widely used in laboratory environments. This part is regularly used in material resistance tests, including mechanical tests to evaluate physical properties, such as tension, compression, and flexion, in samples composed of a variety of materials.

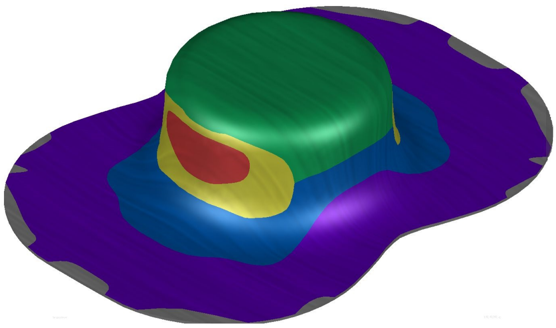

Figure 14 – Case Study 1 - Cup. Areas of wrinkling (WRI - highlighted in purple), under stretch (IS - highlighted in grey), and fracture index (FRA - highlighted in red). Source: Own authorship

The second case study aims to analyze the automotive part called the "Internal Tailgate", which is a fundamental component used in automobiles. This part is positioned at the rear of the vehicle, more specifically in the area corresponding to the trunk.

Figure 15 presents construction details of the internal tailgate, demonstrating information about its structure and characteristics.

The structure of this chapter follows the following organization: Initially, the experiment plan is presented, and then the results of each case study.

5.1 EXPERIMENTS PROTOCOL

The proposed approach was developed in two programming languages, LabVIEW for the area calculation step (AutoAC software) and Python with the Sklearn library[1] and pymoo (BLANK; DEB, 2020) for model definition steps and search for the

[1] Sklearn site: <https://scikit-learn.org/>

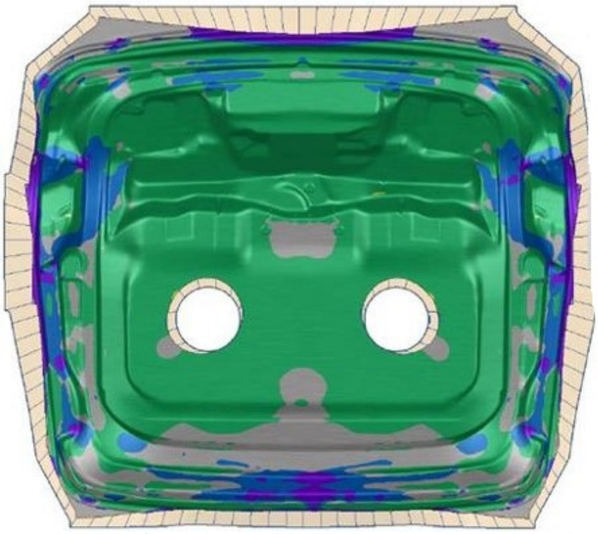

Figure 15 – Internal Tailgate - Areas of wrinkling (highlighted in purple) and under stretch (highlighted in grey). Source: Own authorship.

best solutions.

To create the replacement models, we use multivariate regressors, namely ET and RF. Additionally, we apply hyperparameter tuning tools to optimize model performance in each case study. The training was carried out by the leave-one-out (LOO) cross-validation (CV) method, using the root mean squared error (RMSE) metric in a sensitivity analysis set to evaluate the quality of the regression models.

In the "Cup" case study, the variables Coefficient of Friction (FRC), Swift Hockett Sherby Combination Factor (SWI), Yield Stress (YS), Blank Support Force (BHF), and NValue (NVA) were initially considered by the specialist. The table 3 presents the domains and parameters used to generate the database by the AutoForm software. Based on these parameters, it was possible to generate a database with 747 AutoForm simulations, with five inputs (FRC, SWI, YS, BHF, and NVA) and three outputs (WRI, IS, and FRA).

In the "Internal Tailgate" case study, the variables Blank Support Force (BHF), Friction Coefficient (FRC), Tensile Strength (TS), and Yield Resistance (YS) were initially considered by the specialist. Table 4 presents the domains and parameters used to generate the database by the AutoForm software, and below, the purpose of each parameter in the deep drawing process is briefly described. Based on these parameters, it was possible to generate a database with 200 AutoForm simulations, with four inputs (BHF, FRC, TS, and YS) and three outputs (WRI, IS, and FRA).

- FRC: the parameter is responsible for quantifying the existing sliding restriction between the surfaces of the plate and the stamping tool. Lower coefficients of friction decrease the risk of fractures but can increase the likelihood of wrinkling

Table 3 – Parameters and domains of case study 1 (Cup). Source: Own authorship

Name	Parameters	Minimum	Medium	Maximum
X_{01}	FRC	0.13	0.15	0.19
X_{02}	SWI	0.5	0.75	1
X_{03}	YS	145.867	171.608	240.251
X_{04}	BHF	38000	47500	57000
X_{05}	NVA	0.159	0.199	0.239

Table 4 – Parameters and domains of case study 2 (Internal Tailgate). Source: Own authorship.

Name	Parameters	Minimum	Medium	Maximum
X_{01}	BHF	4e^{5}	9e^{5}	1.4e^{6}
X_{02}	FRC	0.13	0.15	0.17
X_{03}	TS	300	356.435	400
X_{04}	YS	205.692	244.385	274.385

and even result in insufficient stretching.

- SWI: this variable relates to the way the steel hardens, having a significant effect on the material's response to deformation during the stamping process as the intensity of this parameter is adjusted.

- YS: the parameter represents the stress corresponding to the transition between the elastic and plastic regions of the steel. Tensions above the yield point promote definitive deformations in the material (necessary in stamping processes). The greater the yield strength of steel, the greater its hardness and mechanical strength. However, there is a tendency for steels with higher yield strengths to have lower formability.

- BHF: the parameter regulates the sheet feeding into the stamping die. The application of higher Blank Holder Forces is necessary to avoid the formation of wrinkles, especially in the peripheral regions of the stamped part, in addition to contributing to greater rigidity in the central areas. Excessive force, however, can lead to the formation of cracks during stamping.

- NVA: the value "n," also known as the hardening coefficient, is a measure that evaluates a material's ability to withstand deformations before reaching the necking stage, which precedes rupture. Furthermore, this value represents the ability of a steel to acquire additional resistance, known as hardening. Therefore, materials that have higher "n" values have a greater capacity for both deformation and increased mechanical resistance during the stamping process.

- TS: the parameter consists of defining the ability of a material to withstand the application of tensile forces without breaking, that is, the greater the tensile strength of a material, the greater its ability to withstand tensile forces without failing.

The parameters used to adjust the regression model of these studies are presented in Table 5. These parameters were defined using a factorial design.

Table 5 – Parameters used to define models. Source: Own authorship.

Case study	Algorithm	Estimator	Depth
Cup	RF	130	20
Cup	ET	120	27
Internal Tailgate	RF	80	20
Internal Tailgate	ET	60	21

The implementation of objective functions and parameterization were the fundamental steps to start the evolutionary process. At this stage, we choose the NSGA-II algorithm to perform process optimization. Thus, we perform 1000 generations, each one composed of a population of 100 individuals (candidate solutions) with actual coding, performing 30 executions. Each individual represents a solution vector containing several input variables, and the objective functions considered were WRI, IS, and FRA.

All AutoForm runs were performed on hardware equipped with a 2.20 GHz i7-8750H processor and 32 GB of RAM. To run the surrogate models and genetic algorithms, we used hardware with an i5-11300H processor, 8 GB of RAM, and a frequency of 3.1 GHz/3.11 GHz.

5.2 ANALYSIS

This section presents the experiments carried out and the results obtained in two case studies. Both studies aim to minimize the objectives of WRI, IS, and FRA and carry out the evaluation of two multivariate algorithms, namely ET and RF.

5.2.1 CASE STUDY - CUP

To evaluate the significance of the defined parameters and domains, after the database generation step using the AutoForm numerical simulation software, correlation analysis is used as shown in detail a) of Figure 16. This tool allows for a preliminary investigation of the data, aiming for a more in-depth understanding of the relationships between the variables before creating the replacement model.

Based on the analysis of the correlations (detailed in item a) of Figure 16) between the parameters and the objectives, it is possible to infer that the SWI variable does not present a significant correlation with any of the objectives. This indicates that the inclusion of the SWI variable in the analysis model would be inappropriate, as it is

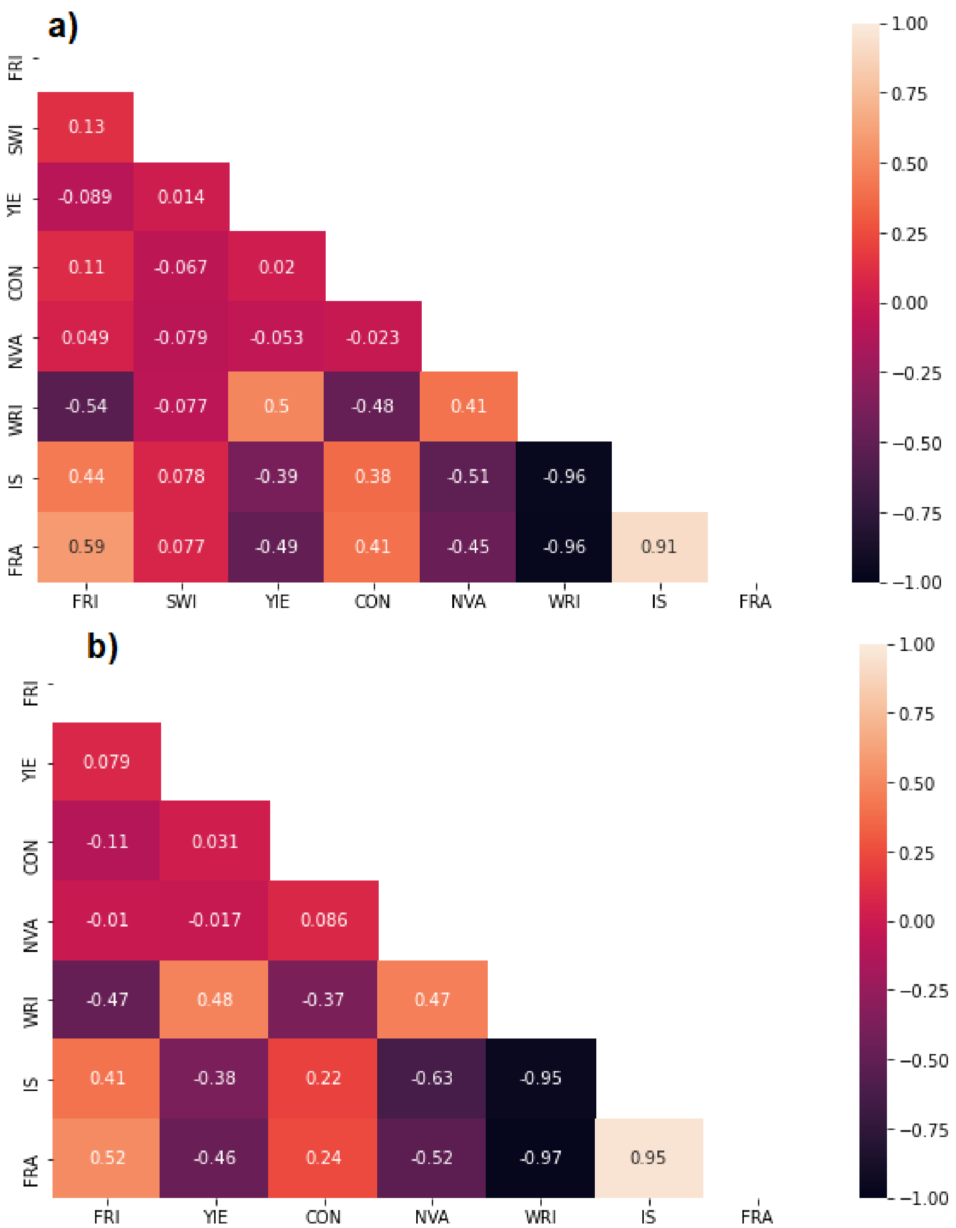

Figure 16 – Correlation between variables - Cup case study (a) 5 variables, (b) 4 variables

not associated with the objectives in question and could introduce unnecessary noise into the analysis. Therefore, the optimization process goes back to the first step, "Selection of parameters and domains", excluding the SWI variable and carrying out new numerical simulations to generate the database. Maintaining the integrity of the project flow, the subsequent steps are performed to evaluate the impacts on the creation of the replacement model, aiming to achieve more accurate and relevant results for the research.

On the other hand, it is noted that the FRI and CON parameters show a similar correlation with the analyzed objectives. They present a moderate positive correlation with the IS and FRA objectives, while they demonstrate an inverse correlation with the WRI objective. In turn, the YIE and NVA variables present an opposite correlation to those of FRI and CON. That is, they maintain a negative relationship with the IS and FRA objectives, while they have a positive correlation with WRI.

5.2.1.1 SURROGATE MODEL

One of the objectives of the project is to develop a specific model for each automotive part using ML techniques, which can reliably represent the intrinsic characteristics of the analyzed part. To achieve this objective, two algorithms belonging to the family of tree-based algorithms, ET and RF, were subjected to evaluation.

The training results of both algorithms, using the LOO cross-validation method, are presented in Figure 17. Notably, both algorithms present low RMSE, lower than the median value of 0.11. Furthermore, most models demonstrate low errors, as evidenced between the first and third quartiles.

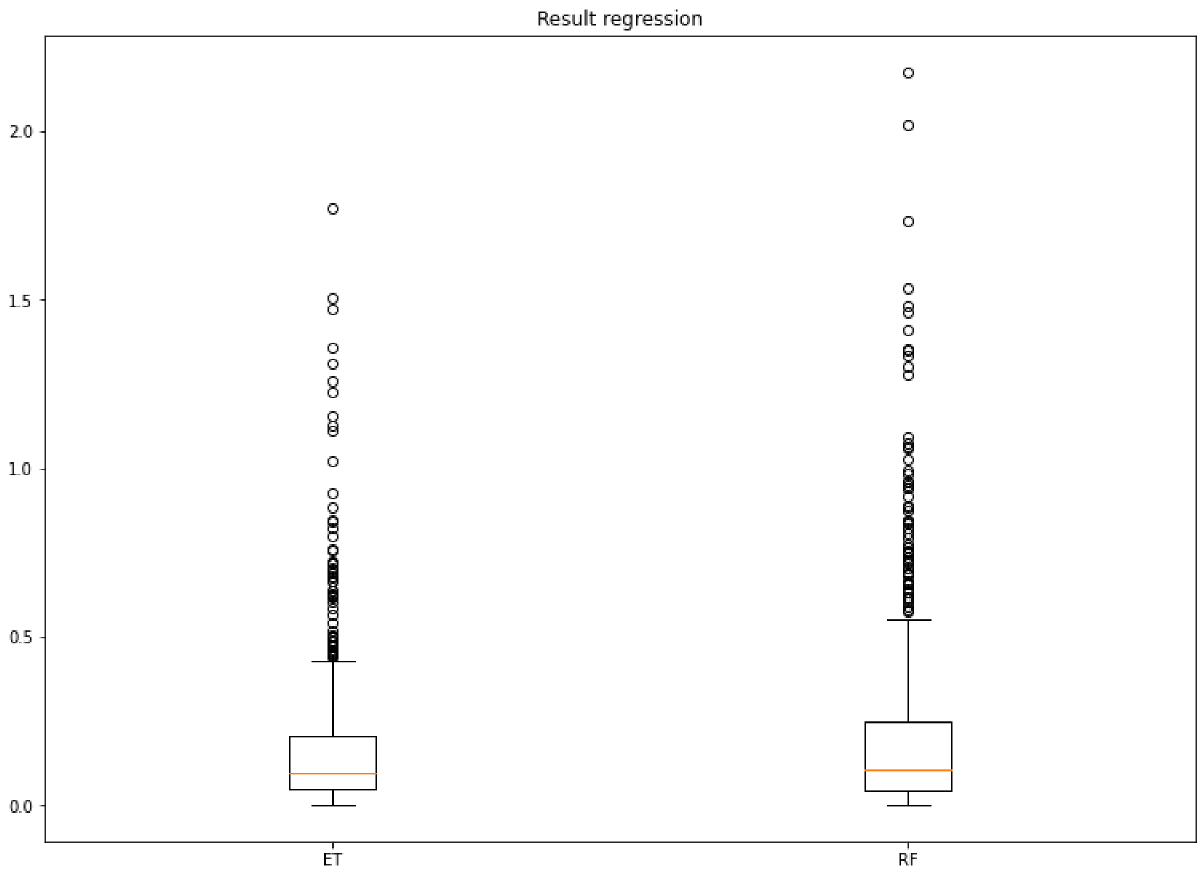

Figure 17 – Model training results using LOO and RMSE evaluation metric. Source: Own authorship

To improve the evaluation of the most outstanding model among the 747 created, a sensitivity analysis is conducted, the results of which are detailed in Figure 18. In this process, 100 fractional values were predicted for each input variable within the ranges determined during parameter selection (represented on the x-axis). The predicted values are then plotted on the graph (y-axis).

The SWI variable, presented in Figure 18, demonstrates difficulties during prediction. This behavior can be noted by the restriction of the range of results on the y-axis, indicating a limited correlation between the variable and the objective, resulting in a reduced influence on the construction of the model. Furthermore, abrupt results are observed, suggesting that the model faces difficulties in learning the most crucial relationships or may be suffering from underfitting.

The other variables exhibit appropriate behavior, demonstrating the sensitivity of the trained model across the entire range of values entered. This is noticeable by the sensitivity demonstrated during the prediction of the introduced values (y-axis). It is important to highlight that the model that uses the RF algorithm manifests more pronounced variations compared to the ET algorithm. This discrepancy may be related to the intrinsic characteristics of the algorithms.

Based on the results presented, the initial hypothesis, formulated during the exploratory data analysis, that the SWI variable exerts a negative influence which is reinforced by the sensitivity analysis of the adopted model. Given this scenario, a new experiment plan was conducted, covering only four inputs (FRC, YS, BHF, and NVA). Then, a new model was developed and submitted to the optimization phase, playing the role of a surrogate model from the part under analysis.

5.2.1.2 SEARCH FOR THE BEST PARAMETER OPTION

Table 6 presents the mean value and standard deviation of the WRI, IS, and FRA objectives obtained. Both surrogate models (ET and RF) were used in this experiment since the regression algorithms demonstrated similar results in the previous steps. Therefore, it became essential to evaluate the efficiency of both models when applied in the optimization process and subsequent validation with the AutoForm simulation software.

Table 6 – Mean and standard deviation (std) of WRI, IS, and FRA for 30 runs - Cup case study. Source: Own authorship

Algorithm	WRI	IS	FRA
ET (mean $\pm$ std)	32.239 $\pm$ 6.118e^{-5}	2.754 $\pm$ 5.640e^{-3}	0.025 $\pm$ 0.009
RF (mean $\pm$ std)	32.254 $\pm$ 1.420e^{-14}	2.737 $\pm$ 0.027	9.202e^{-3} $\pm$ 2.939e^{-3}

Through the implementation of surrogate models using machine learning algorithms, it was feasible to employ the NSGA-II algorithm to optimize predefined objectives, aiming to find the most effective solutions in a multi-objective search space. To evaluate the efficiency of the optimization process, the same convergence and diversity metrics applied to the Cup case study, such as hypervolume, Dnpw, and Pareto

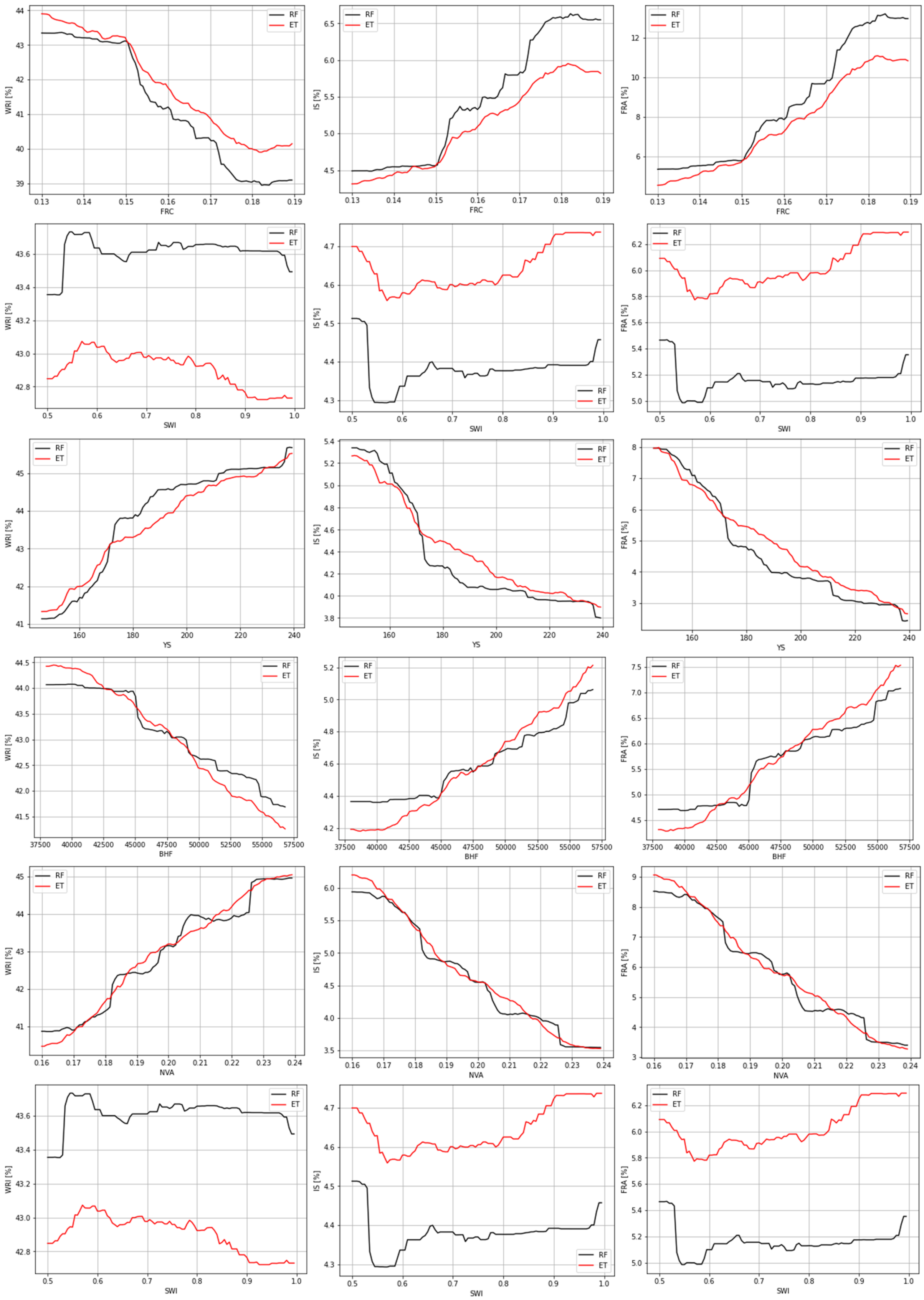

Figure 18 – Sensitivity analysis - Cup case study 5 variables. Source: Own authorship

frontal visualization, were used. The average values of the optimization metrics and the Pareto front are presented in Figures 19, 20 and 21, respectively.

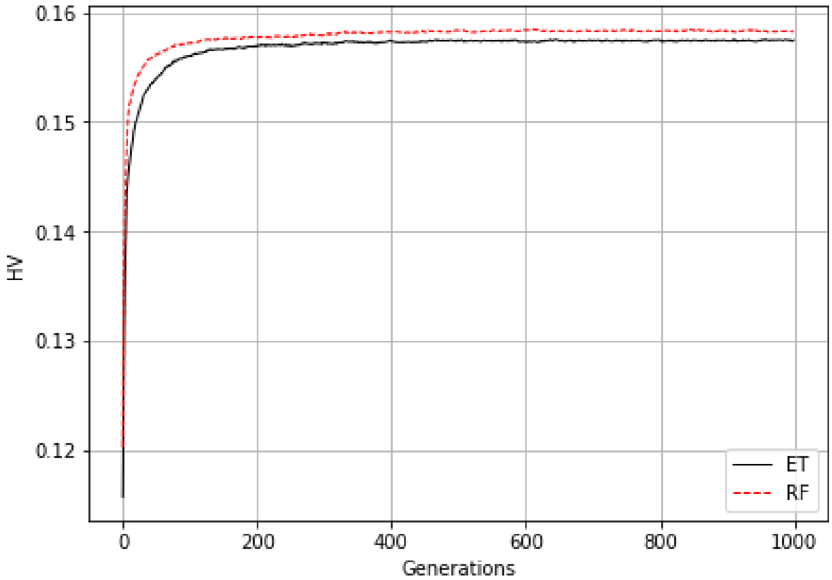

Figure 19 – Hypervolume - Cup case study. Source: Own authorship

As for convergence, in Figure 19, it is possible to notice that both algorithms demonstrated a similar convergent trajectory. The RF algorithm showed a slight advantage, reaching a maximum value of approximately 0.158, while the ET algorithm reached 0.157. This difference is not significant enough to declare the unfeasibility or assert the superiority of any of the models.

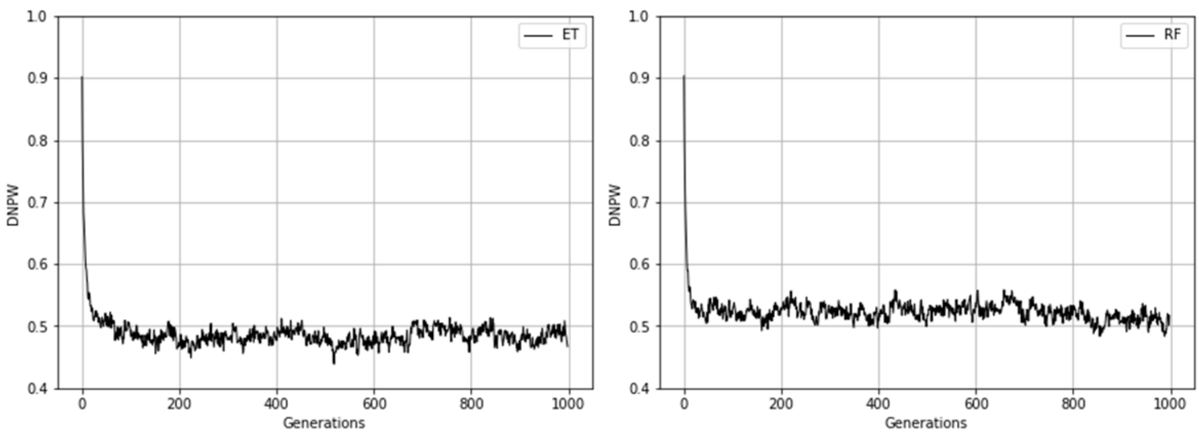

Figure 20 – DNPW (left) ET and (right) RF - Internal Tailgate case study. Source: Own authorship

In Figure 20, it is possible to observe a slight advantage in the diversity provided by the RF algorithm compared to ET throughout the evolutionary generations. The average values recorded were 0.523 for RF and 0.487 for ET, indicating that the RF algorithm maintained a more expressive diversity throughout the evolutionary process.

The Pareto Frontier analysis reveals how effective the algorithm was in providing the best solutions. Figure 21 graphically demonstrates that both algorithms present

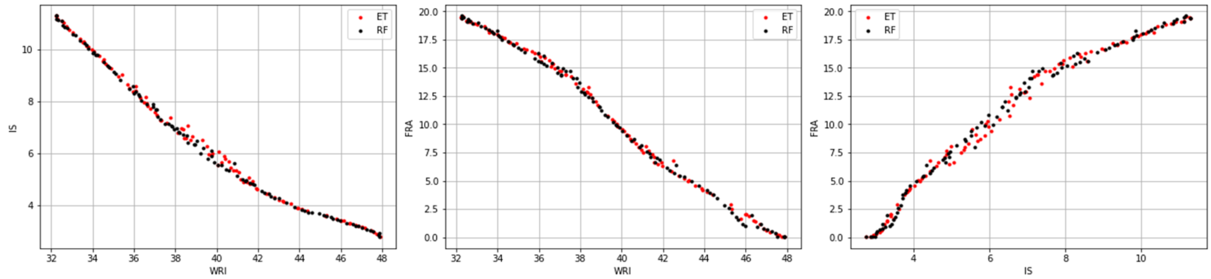

Figure 21 – Pareto front - Cup case study. Source: Own authorship

notable similarity in results, although the RF algorithm exhibits more significant variance in the data delivered.

A notable feature of NSGA-II is its consistent ability to generate Pareto fronts, enabling the selection of individuals from a single run to analyze the results. Validation of 20 non-dominated individuals from each set of solutions occurred in the AutoForm software. In other words, 20 individuals were evaluated using the RF surrogate model, while another 20 were evaluated using the ET surrogate model, seeking to evaluate the accuracy of the proposed design model. The results obtained are presented in Figure 22 and Table 7.

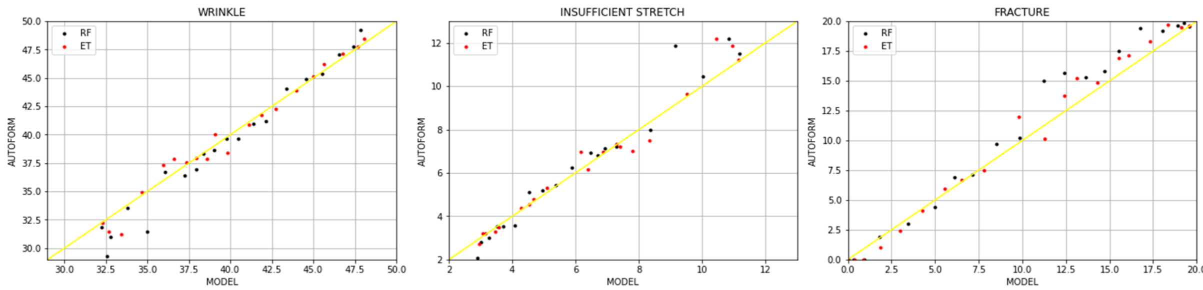

Figure 22 – Validation of results - Model Proposed x AutoForm. Source: Own authorship

Table 7 – Validation of results - Model x AutoForm. Source: Own authorship

Algorithm	Metric	WRI	IS	FRA
ET	R^2	0.970	0.956	0.976
	RMSE	0.837	0.552	1.010
RF	R^2	0.930	0.881	0.947
	RMSE	1.301	0.942	1.514

By analyzing Figure 22 and Table 7, it is possible to verify that both algorithms exhibited positive performance, confirming evidence previously observed throughout the analyses. However, it is noteworthy that the ET algorithm demonstrated better performance for all objectives. Table 7 presents metrics such as R^2 and RMSE, used for a qualitative assessment of the predictive capacity of the proposed model. In this case study, such metrics revealed remarkable efficiency in prediction, especially for the WRI objective.

Finally, the model provided positive results during the selection phase of the best parameters for automotive stamping. From the exploratory data analysis, the model enriched the specialist's perceptions, providing relevant information in advance and allowing more informed decision-making.

In this case study, the hypothesis that the SWI variable would not contribute significantly to the forming process was validated. Furthermore, the slight superiority of the ET algorithm in terms of assertiveness was highlighted, which is one of the main decisive factors in this specific scenario.

5.2.2 CASE STUDY - INTERNAL TAILGATE

The methodology used in this research involved the validation of two case studies, highlighting the second case, in this section, in which the proposed methodology was applied to a real automobile part. At this stage, in addition to the phases previously conducted in the evaluation of the laboratory case study (Cup), the last phase was included, which includes the comparison between methodologies. This analysis covers the empirical approach currently conducted by experts in the field and contrasts it with the model proposed in this study.

The assessment of the significance of the parameters and domains established in this case study is carried out through the application of correlation analysis, as illustrated in Figure 23.

After analyzing the correlations between the parameters and the objectives, we observe results that allow us to conclude that all the input variables have a substantial correlation with the established objectives. This finding reinforces the validity and effectiveness of the experimental design proposed by specialists in the field.

It is worth highlighting the presence of a significantly strong reverse and direct correlation between the BHF input variable and the WRI and FRA objectives, respectively. This indicates that the BHF variable has a relevant influence on the objectives mentioned above. An unusual factor to highlight is the notable similarity of the TS and YS variables in relation to all objectives. This results in a perfect correlation of one between these variables.

5.2.3 SUROGATE MODEL

After validating the parameters and domains, surrogate models, namely RF and ET, were developed. Subsequently, using the leave-one-out (LOO) cross-validation methodology in conjunction with the RMSE metric, the model with the best performance for each algorithm was listed. The figure 24 shows the distribution of results obtained in both algorithms during training using the LOO approach. Notably, both

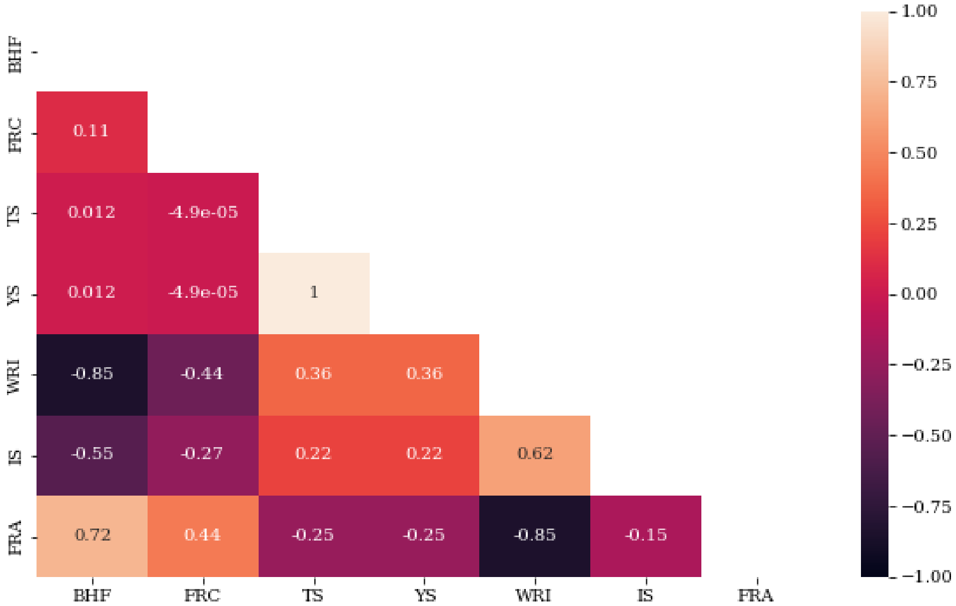

Figure 23 – Correlation between variables - Internal Tailgate case study. Source: Own authorship.

algorithms demonstrate low RMSE, recording values below the median of 0.2. The minimum values achieved were 0.0091 with the ET algorithm and 0.0176 for the RF algorithm, suggesting the excellent performance of the trained models.

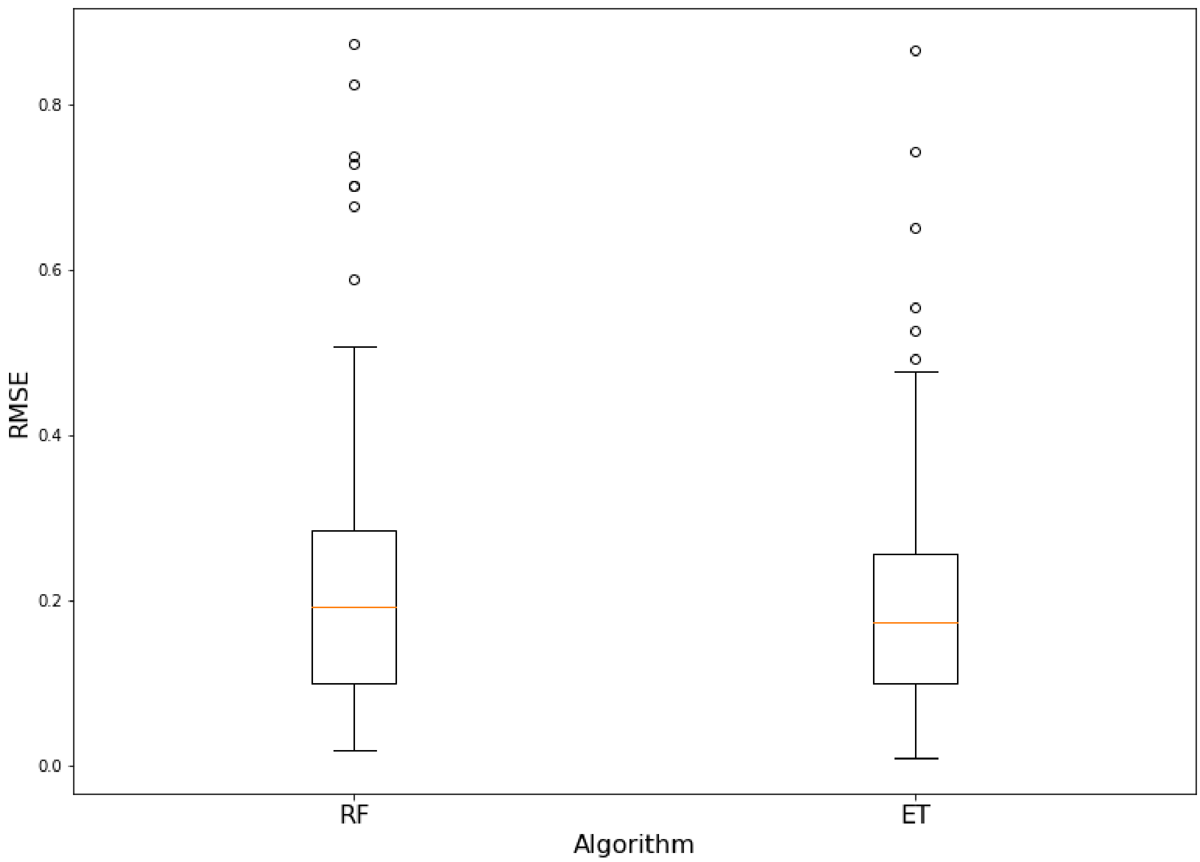

Figure 24 – Model creation using leave-one-out cross-validation. Source: Own authorship.

To obtain a deeper understanding of the developed model, a sensitivity analysis is performed to assess its effectiveness. This analysis aims to examine the influence of input variables on the model output to determine whether the model can adequately represent the problem at hand. Figure 25 shows the analysis performed, where some points can be observed, namely:

- The two regression algorithms produced similar results, showing a close correspondence between variations in the domains (x-axis) and corresponding objectives (y-axis). There is only a slight discrepancy between the models in the analyses carried out by the TS and YS variables.
- The BHF variable is the most relevant for all objectives, which can be verified by the variation expressed on the y-axis.
- The TS and YS variables exhibit analogous behavior in relation to the distribution along the y-axis. Furthermore, they are less relevant for all objectives, corroborating evidence previously identified during the correlation analysis.
- Both algorithms reveal more incredible difficulty in characterizing the TS and YS variables, highlighting a discrepancy in the results. In this context, it is noted that the ET algorithm displays higher values in relation to the RF on the y-axis.

Based on the obtained results, it was observed that both algorithms presented similar RMSE metric and sensitivity analysis in relation to the input parameters. Therefore, both algorithms will be taken to the next step of the design, in which the objective is to determine which replacement model will be more suitable to accurately delineate the problem. In this analysis, it is important to highlight that the TS and YS parameters have less influence on the defined objectives, and that BHF stood out positively for all objectives.

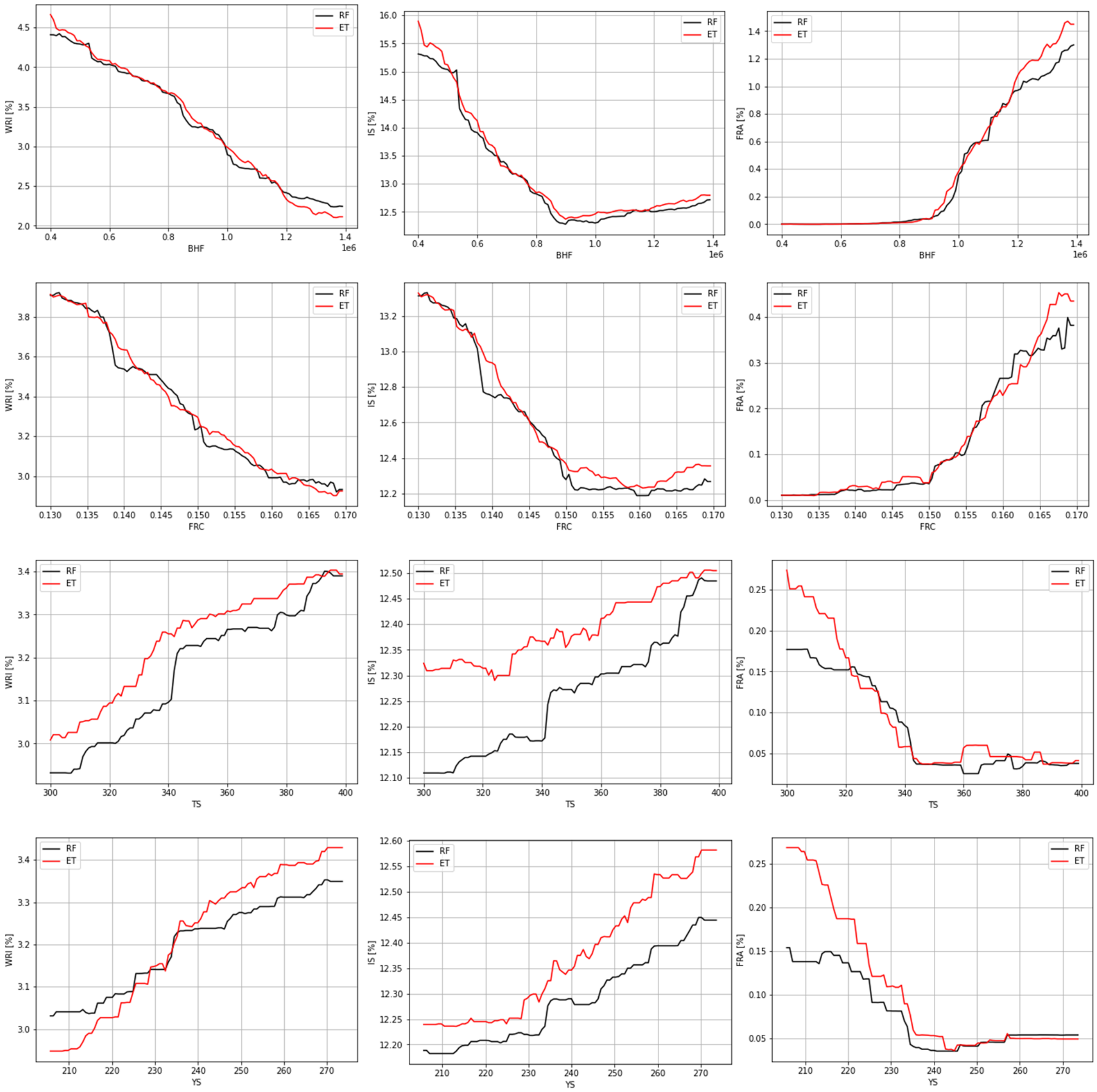

Figure 25 – Sensitivity analysis - Internal Tailgate case study. Source: Own authorship.

5.2.4 SEARCH FOR THE BEST PARAMETER OPTION

In a similar way to the Cup case study, table 8 exposes the average values and standard deviations achieved for the WRI, IS, and FRA objectives. Both algorithms presented comparable results with regard to the minimum value obtained, highlighting the RF algorithm, which managed to achieve minimization levels higher than ET in all objectives, in addition to exhibiting lower variance in the results obtained.

To further evaluate the effectiveness of the optimization process, convergence and diversity metrics were used, as shown in Figures 26, 27, and 28. These figures present the average values of the hypervolume, Dnpw, and Pareto front metrics, respectively.

As for convergence, in Figure 26, it is evident that both algorithms reached

Table 8 – Mean and standard deviation (std) of WRI, IS, and FRA for 30 runs - Internal Tailgate case study. Source: Own authorship

Algorithm	WRI	IS	FRA
ET (mean ± std)	$1.300 \pm 1.168e^{-3}$	11.921 ± 0.010	$1.240e^{-5} \pm 1.200e^{-5}$
RF (mean ± std)	$1.299 \pm 2.220e^{-16}$	$11.869 \pm 1.662e^{-3}$	$7.960e^{-6} \pm 2.640e^{-6}$

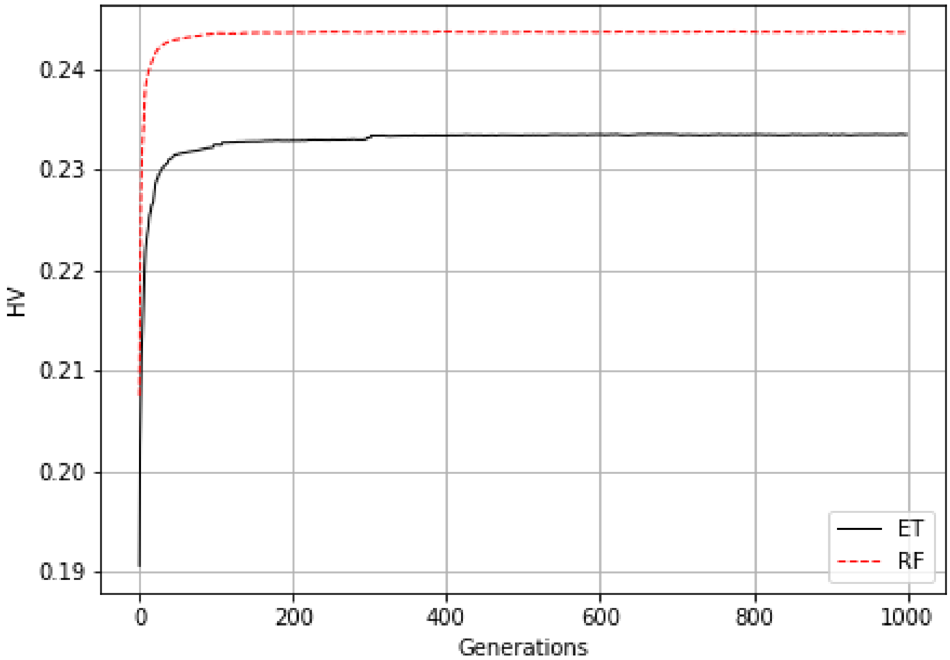

Figure 26 – Hypervolume - Internal Tailgate case study. Source: Own authorship.

maximum optimization convergence quickly, close to generation 100. Furthermore, the RF algorithm achieved greater convergence than ET. It is possible to observe the low convergence of the genetic algorithm, which can be attributed to the minimization challenges inherent to this multi-objective problem.

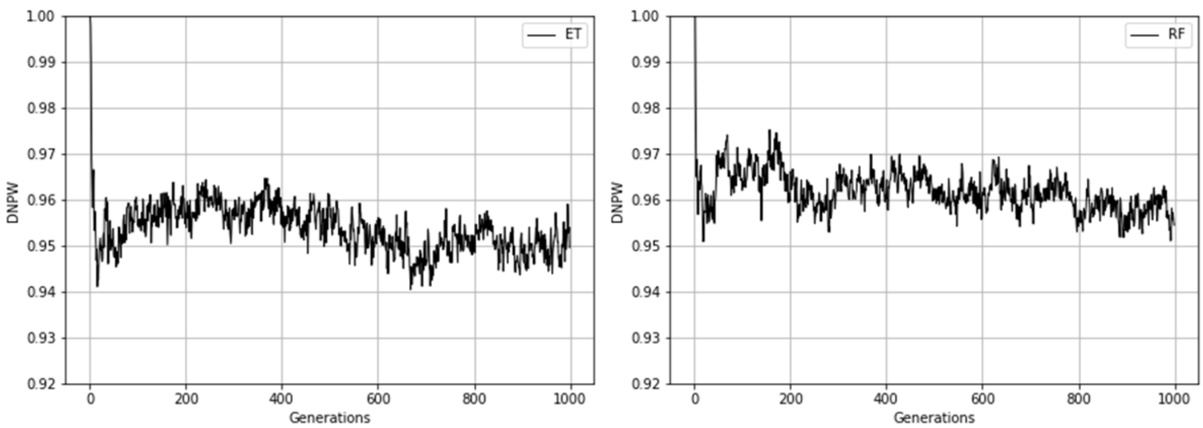

Figure 27 – DNPW (left) ET and (right) RF - Internal Tailgate case study. Source: Own authorship

Figure 27 highlights a significant diversity in the solutions generated by the two replacement models, with a notable increase in the first 100 generations due to the continuous convergence of the algorithm. Additionally, a higher average value for the RF algorithm stands out. After generation 100, it is noticeable that, together with the convergence metrics, the local optimum optimization point was reached.

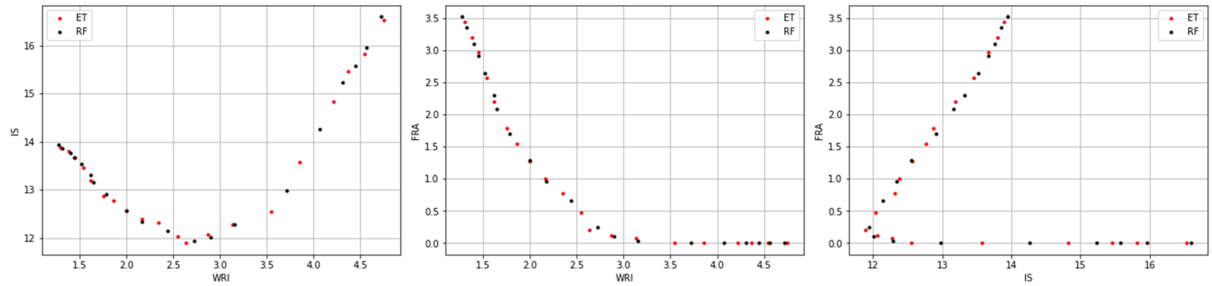

Figure 28 – Pareto front - Internal Tailgate case study. Source: Own authorship.

Figure 28 shows the 100 non-dominated solutions for the problem objectives using the ET and RF surrogate models. This approach provides the specialist with diverse access to solutions, simplifying decision-making in selecting the most appropriate parameters and domains to obtain the best results. The graphical representation also helps in visualizing the diversity in the suggestions provided by the model.

In this context, it is observed that both algorithms present similar results in both convergence and diversity, suggesting that they converge to similar solutions.

In the final stage of the project, some individuals were validated in the AutoForm software, the results of which are presented in Figure 29 and Table 9.

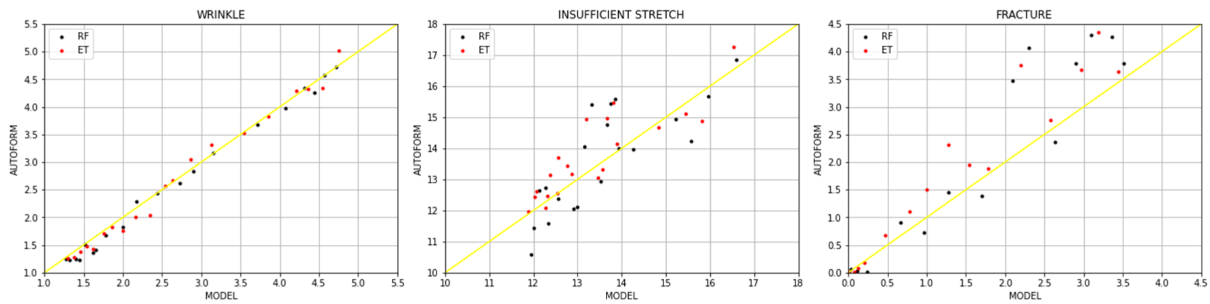

Figure 29 – Validation of the parameters suggested by the model. Source: Own authorship.

Table 9 – Validation of results - Model x AutoForm. Source: Own authorship

Algorithm	Metric	WRI	IS	FRA
ET	R^2	0.983	0.636	0.784
	RMSE	0.146	0.786	0.548
RF	R^2	0.989	0.441	0.741
	RMSE	0.125	0.985	0.655

Figure 29 shows the values obtained after validation in the AutoForm simulation software. The goal is to achieve highly accurate model suggestions, that is, minimizing the discrepancy between the model suggestions and the actual AutoForm outputs, characterizing it as a well-adjusted and practical model.

Despite the complexities inherent to the optimization of multi-objective problems, it was possible to identify positive results through the suggestions generated by the proposed model when compared with the empirical analysis previously conducted by experts in the field. Table 10 details the configuration of the variables and the results of the objectives obtained by the original stamping of the part, contrasting with the configuration suggested by the proposed approach.

Table 10 – Difference in proportion of affected area, in percentage, by objective. Source: Own authorship

	BHF	FRC	TS	YS	WRI	IS	FRA
Empirical	$3.90e^5$	0.15	350	245	4.63	14.76	0
Proposed Model	$4.73e^5$	0.15	324.92	218.59	3.93	13.6	0

The suggested model resulted in the production of a fracture-free part while reducing the initial wrinkling by approximately 15% and decreasing the stretch region by approximately 8%. In more precise terms, the model results represent a substantial improvement in both the shape of the part and the costs related to its development. Furthermore, it offers essential insights into defining stamping parameters and assists in the innovation of new materials.

Comprehensively, the proposed methodology obtained positive results both in the laboratory case study and in practical application on an actual automobile part. Regarding the comparison of efficiency between the algorithms, a slight superiority of the RF algorithm is noted in terms of diversity and convergence. However, the ET algorithm demonstrated superiority in the assertiveness of the proposed solutions, that is, it presented more realistic suggestions, directing experts in the area more appropriately.

6 CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

Steel forming by stamping is a process widely used in the automotive industry, where controlled deformation of metal sheets occurs. Properly defining the process parameters and materials for each part is essential to maintain high production criteria with high quality. However, this task is challenging due to the intrinsic characteristics of the problem since it is characterized as a multi-objective problem, where its objectives are usually antagonistic, making it challenging to find the desired local minimum of the objectives together.

To reduce the incidence of defects such as WRI, IS, and FRA as much as possible, advanced optimization techniques and machine learning algorithms have been used in recent years as carried out in (CUI et al., 2020), where the strategy adopted includes the combination of simulation of finite elements with the RSM algorithms and the NSGA-II multi-objective genetic algorithm to explore the optimal stamping parameters. These approaches have significantly accelerated research and development by combining empirical analyses carried out by experts in the field.

Therefore, in this work, we develop an alternative model using data-driven algorithms to predict the process parameters and materials of steel intended for stamping. The capabilities of these models were highlighted in two case studies: one in a laboratory environment called “Cup” and another involving a real automobile part called “Internal Tailgate”. By employing a hybrid solution that combines a machine learning model as a surrogate model coupled with a multi-objective optimizer, we achieve a low margin of error in parameter prediction. This not only speeds up decision-making by experts in the field but also significantly raises the technological and intellectual level in the specific domain.

With the aim of supporting experts in the field and streamlining the decision-making process, intermediate tools were incorporated that provide targeted information on the progress of the stages of the proposed model. In the “Cup” laboratory case study, we were able to observe these tools in operation. After creating the replacement model, it was possible to identify and validate the hypothesis that the SWI parameter was not suitable for application during the tests.

When compared to the systematic literature review carried out, this work presented notable scientific contributions, such as the following aspects:

The first highlight lies in the implementation of a robust and efficient methodology. The work adopted a hybrid solution, integrating machine learning and multi-objective optimization techniques.

Another relevant point is the development of auxiliary tools designed to facilitate the integration between the simulation system and the developed algorithms. This proved necessary due to the intrinsic limitations of the software used.

Furthermore, the work innovated by applying sensitivity analysis techniques during the process of selecting the best parameters. This approach deepened the understanding of the database used, providing valuable insights for the precise definition of stamping parameters and domains.

Despite the positive results obtained, this work identified several challenges and opportunities that can be explored in future research. Among them, the possibility of expanding the number of objectives addressed stands out. Although this study explored the three main objectives related to automotive stamping, there is an opportunity to evaluate other objectives of lesser significance, which would imply an increase in the number of objectives. Although the proposed model is able to deal with such challenges, it would be beneficial to carry out a comparative study between different algorithms or optimization methodologies. Additionally, an excellent research opportunity could arise by exploring the coupling with the project carried out by Douglas Sgrott (SGROTT, 2022).

BIBLIOGRAPHY

AL-JARRAH, O. Y. et al. Efficient machine learning for big data: A review. **Big Data Research**, Elsevier, v. 2, n. 3, p. 87–93, 2015.

ALETI, A.; MOSER, I. Predictive parameter control. In: **Proceedings of the 13th annual conference on Genetic and evolutionary computation**. .: ., 2011. p. 561–568.

AMOUZGAR, K. **Multi-objective optimization using Genetic Algorithms**. 2012.

ATZEMA, E. Formability of auto components. In: METAL, W. P. S. in; ENGINEERING, S. (Ed.). **Automotive Steels**. United Kingdom: Elsevier, 2017. p. 47–93.

BAO, L. et al. Multi-objective optimization of partition temperature of steel sheet by nsga-ii using response surface methodology. **Case Studies in Thermal Engineering**, v. 31, p. 101818, 2022. ISSN 2214-157X. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2214157X22000648>.

BHUYAN, D. R. et al. **Design For Six Sigma (DFSS) for Optimization of Stamping Simulation Parameters to Improve Springback Prediction**. ., 2015.

BINITHA, S.; SATHYA, S. S. et al. A survey of bio inspired optimization algorithms. **International journal of soft computing and engineering**, Citeseer, v. 2, n. 2, p. 137–151, 2012.

BLANK, J.; DEB, K. Pymoo: Multi-objective optimization in python. **Ieee access**, IEEE, v. 8, p. 89497–89509, 2020.

BOLJANOVIC, V. Mechanical behavior of materials. In: **Sheet metal forming processes and die design**. New York: Industrial Press Inc., 2004. p. 13–19.

BONACCORSO, G. Elements of multiobjective decision problem. In: **Machine learning algorithms**. New York: Packt Publishing Ltd, 2017. p. 3–23.

BREIMAN, L. Random forests. **Machine learning**, Springer, v. 45, p. 5–32, 2001.

BRENDEL, P. L. **Otimização Multiobjetivo Aplicada ao Processo de Estampagem de Peça Automotiva Utilizando Algoritmos Evolutivos**. 55 f. Monografia (Bacharelado) — Bacharelado em Ciência da Computação, Universidade do estado de Santa Catarina, Joinville, 2021.

BRENDEL, P. L. et al. Otimização do processo de estampagem de peça automotiva utilizando rede neural artificial e algoritmo evolutivo multiobjetivo. XV Congresso Brasileiro de Inteligência Computacional, 2021.

CHANKONG, V.; HAIMES, Y. Y. Elements of multiobjective decision problem. In: **Multiobjective decision making: theory and methodology**. New York: Courier Dover Publications, 2008. p. 3–23.

CORRIVEAU, G. et al. Review and study of genotypic diversity measures for real-coded representations. **IEEE Transactions on Evolutionary Computation**, v. 16, n. 5, p. 695–710, 2012.

COSTA, P. R. M. Princípios e cenários da indústria 4.0: Uma revisão de literatura. In: **CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO**. .: ., 2019. v. 9.

CUI, M. et al. Numerical simulation and multi-objective optimization of partition cooling in hot stamping of the automotive b-pillar based on rsm and nsga-ii. **Metals**, MDPI, v. 10, n. 9, p. 1264, 2020.

CUTLER, A.; CUTLER, D. R.; STEVENS, J. R. Random forests. **Ensemble machine learning: Methods and applications**, Springer, p. 157–175, 2012.

DANG, V. T.; LAFON, P.; LABERGERE, C. Surrogate models for sheet metal stamping problem based on the combination of proper orthogonal decomposition and radial basis function. In: AIP PUBLISHING. **AIP Conference Proceedings**. ., 2017. v. 1896, n. 1.

DASARI, S. K.; CHEDDAD, A.; ANDERSSON, P. Random forest surrogate models to support design space exploration in aerospace use-case. In: SPRINGER. **Artificial Intelligence Applications and Innovations: 15th IFIP WG 12.5 International Conference, AIAI 2019, Hersonissos, Crete, Greece, May 24–26, 2019, Proceedings 15**. Greece, 2019. p. 532–544.

DEB, K. Multi-objective optimisation using evolutionary algorithms: an introduction. In: **Multi-objective evolutionary optimisation for product design and manufacturing**. London: Springer, 2011. p. 3–34.

DEB, K. et al. A fast and elitist multiobjective genetic algorithm: Nsga-ii. **IEEE transactions on evolutionary computation**, IEEE, v. 6, n. 2, p. 182–197, 2002.

ESENER, E.; ERCAN, S.; FIRAT, M. A sensitivity analysis by using design of experiment and its application in stamping. In: **An International Conference on Engineering and Applied Sciences Optimization**. .: ., 2014. p. 1388–1394.

FONSECA, C. M.; PAQUETE, L.; LÓPEZ-IBÁNEZ, M. An improved dimension-sweep algorithm for the hypervolume indicator. In: IEEE. **2006 IEEE international conference on evolutionary computation**. Vancouver, BC, Canada, 2006. p. 1157–1163.

GEURTS, P.; ERNST, D.; WEHENKEL, L. Extremely randomized trees. **Machine learning**, Springer, v. 63, p. 3–42, 2006.

GÜRÜN, H.; KARAAĞAÇ, İ. The experimental investigation of effects of multiple parameters on the formability of the dc01 sheet metal. **Strojniski Vestnik-Journal Of Mechanical Engineering**, v. 61, n. 11, 2015.

HAMDAOUI, M. et al. Kriging surrogates for evolutionary multi-objective optimization of cpu intensive sheet metal forming applications. **International Journal of Material Forming**, Springer, v. 8, p. 469–480, 2015.

HOSFORD, W. F.; CADDELL, R. M. Stamping. In: **Metal forming: mechanics and metallurgy**. United States: Cambridge university press, 2011. p. 263–275.

HU, M. et al. Study on simulation of stamping forming and die surface optimization of aluminum alloy plate. **Key Engineering Materials**, Trans Tech Publ, v. 764, p. 303–311, 2018.

INGARAO, G.; Di Lorenzo, R. A new progressive design methodology for complex sheet metal stamping operations: Coupling spatially differentiated restraining forces approach and multi-objective optimization. **Computers Structures**, v. 88, n. 9, p. 625–638, 2010. ISSN 0045-7949. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0045794910000349>.

INGARAO, G.; LORENZO, R. D. Optimization methods for complex sheet metal stamping computer aided engineering. **Structural and multidisciplinary optimization**, Springer, v. 42, p. 459–480, 2010.

ISHIBUCHI, H. et al. Performance comparison of nsga-ii and nsga-iii on various many-objective test problems. In: IEEE. **2016 IEEE Congress on Evolutionary Computation (CEC)**. ., 2016. p. 3045–3052.

JIANG, P. et al. Surrogate-model-based design and optimization. In: SPRINGER. **Surrogate-model-based design and optimization**. Singapore, 2020. p. 135–216.

KAZAN, R.; FIRAT, M.; TIRYAKI, A. E. Prediction of springback in wipe-bending process of sheet metal using neural network. **Materials & design**, Elsevier, ., v. 30, n. 2, p. 418–423, 2009.

KIM, S.-Y.; KIM, B.-G.; LEE, T.-G. A study on the parametric optimization of drawing metal stamping process for aluminum alloy tailgate parts using response surface methodology. **Materialwissenschaft und Werkstofftechnik**, v. 54, n. 4, p. 502–511, 2023. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mawe.202200277>.

KIM, Y.; PARK, J. Effect of process parameters on formability in incremental forming of sheet metal. **Journal of materials processing technology**, Elsevier, v. 130, p. 42–46, 2002.

LI, M.; YAO, X. Quality evaluation of solution sets in multiobjective optimisation.

LIBERATI, A. et al. The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. **Annals of internal medicine**, American College of Physicians, v. 151, n. 4, p. W–65, 2009.

LIM, H. et al. Time-dependent springback of advanced high strength steels. **International Journal of Plasticity**, Elsevier, v. 29, p. 42–59, 2012.

LINDEN, R. Ga's conceitos básicos. In: BRASPORT. **Algoritmos genéticos (2a ediçao)**. Rio de Janeiro, 2008. p. 40–65.

MA, G.; HUANG, B. et al. Optimization of process parameters of stamping forming of the automotive lower floor board. **Journal of Applied Mathematics**, Hindawi, v. 2014, 2014.

MAHESH, B. Machine learning algorithms-a review. **International Journal of Science and Research (IJSR).[Internet]**, v. 9, n. 1, p. 381–386, 2020.

OUJEBBOUR, F.-Z.; HABBAL, A.; ELLAIA, R. Optimization of concurrent criteria in the stamping process. In: **Proceedings of 2013 International Conference on Industrial Engineering and Systems Management (IESM)**. .: ., 2013. p. 1–10.

PARK, K.; KIM, Y. The effect of material and process variables on the stamping formability of sheet materials. **Journal of materials processing technology**, Elsevier, v. 51, n. 1-4, p. 64–78, 1995.

PARPINELLI, R. S. et al. A review of techniques for online control of parameters in swarm intelligence and evolutionary computation algorithms. **International Journal of Bio-Inspired Computation**, Inderscience Publishers (IEL), v. 13, n. 1, p. 1–20, 2019.

QUEIPO, N. V. et al. Surrogate-based analysis and optimization. **Progress in aerospace sciences**, Elsevier, v. 41, n. 1, p. 1–28, 2005.

RAFIZADEH, H. et al. Wrinkling prediction in deep drawing by using response surface methodology and artificial neural network. **Transactions of FAMENA**, Fakultet strojarstva i brodogradnje, v. 41, n. 2, p. 17–28, 2017.

RAHIMI, I. et al. A comparative study on evolutionary multi-objective algorithms for next release problem. **Applied Soft Computing**, Elsevier, p. 110472, 2023.

RIBEIRO, R. de S. **Comparação De Algoritmos Evolucionários Para Problemas Com Muitos Objetivos**. Tese (Doutorado) — Universidade Federal de Minas Gerais, 2016.

SCHONFELD, D. The evolution of signal processing [from the editor]. **IEEE Signal Processing Magazine**, v. 27, n. 5, p. 2–6, 2010.

SGROTT, D. M. **A multi-objective data-driven evolutionary algorithm applied to steel modeling**. 104 p. Dissertação (Master's in Applied Computing) — State University of Santa Catarina, Joinville, 2022.

SILVA, B. da et al. Automotive stamping process optimization using machine learning and multi-objective evolutionary algorithm. In: ABRAHAM, A. et al. (Ed.). **Intelligent Systems Design and Applications**. Cham: Springer Nature Switzerland, 2023. p. 351–360. ISBN 978-3-031-27440-4.

SILVA, B. da et al. Surrogate model and multi-objective evolutionary algorithm applied to automotive stamping. 2023.

SINGH, S. K.; GUPTA, A. K. Application of support vector regression in predicting thickness strains in hydro-mechanical deep drawing and comparison with ann and fem. **CIRP Journal of manufacturing Science and Technology**, Elsevier, v. 3, n. 1, p. 66–72, 2010.

SPEARMAN, C. The proof and measurement of association between two things. Appleton-Century-Crofts, 1961.

SU, Y. et al. Numerical simulation and parameter optimization of automobile reinforced inner plate forming process. In: AIP PUBLISHING. **AIP Conference Proceedings**. ., 2017. v. 1864, n. 1.

TEKKAYA, A. E. State-of-the-art of simulation of sheet metal forming. **Journal of Materials Processing Technology**, Elsevier, v. 103, n. 1, p. 14–22, 2000.

TEPEDINO, J. O. A. Aplicação de curvas limite de conformação na previsão de rupturas em bordas de peças estampadas. Universidade Federal de Minas Gerais, 2014.

VIKHAR, P. 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)**. Jalgaon, India: ., 2016. p. 261–265.

WANG, H. et al. Multi-objective optimisation on crashworthiness of front longitudinal beam (flb) coupled with sheet metal stamping process. **Thin-Walled Structures**, v. 132, p. 36–47, 2018. ISSN 0263-8231. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0263823118303732>.

XIE, Y. Robust design of sheet metal forming process based on kriging metamodel. In: AMERICAN INSTITUTE OF PHYSICS. **AIP Conference Proceedings**. ., 2011. v. 1383, n. 1, p. 927–934.

XIE, Y. et al. Optimization of stamping process parameters based on an improved particle swarm optimization–genetic algorithm and sparse auto-encoder–back-propagation neural network model. **Engineering Optimization**, Taylor & Francis, p. 1–22, 2022.

XIE, Y. M. Multi-objective optimal approach based on kriging model in a deep drawing process. **Key Engineering Materials**, Trans Tech Publ, v. 474, p. 205–210, 2011.

YAN, G. et al. Optimization of stamping process parameters based on orthogonal test and intelligent algorithm. In: **2020 3rd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM)**. .: ., 2020. p. 393–397.