

Data-driven Mechanical Property Prediction and Optimization of Hot Rolled Microalloyed Steels

Sushant Sinha



McGill

Department of Mining and Materials Engineering
McGill University, Montreal, Quebec, Canada
December 15, 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Master of Science.

©2024 Sushant Sinha

*Dedicated to my mother,
whose unwavering love and belief
have shaped the person I am today.*

Abstract

The primary objective of this research was to employ data-driven techniques to predict and optimize the mechanical properties of microalloyed steels during the thin slab direct rolling process. The data for this study were sourced from Algoma Steel Inc., located in Ontario, Canada, and the work was divided into two main parts.

In the first part, Deep Neural Network (DNN) models were developed to predict the mechanical properties, specifically Ultimate Tensile Strength (UTS) and Lower Yield Strength (LYS), of Nb-based microalloyed hot-rolled strips. To introduce explainability into the DNN models, Game theory-based SHapely Additive exPlanations (SHAP) were utilized. The SHAP values provided insights into the combined effects of chemical composition and thermomechanical processing parameters. The influence of chemical composition was corroborated by physical metallurgy theory, and correlations were established with the empirical relationship of the No-recrystallization temperature (T_{nr}) from existing literature. Additionally, the data were analyzed using SIMS Mean Flow Stress (MFS) against the inverse temperature, with comparisons across different gauges and compositions to support the model explanations and suggest underlying metallurgical mechanisms. This segment of the study highlighted significant opportunities for optimization of alloy composition, which led to the second part of the research.

The second part aimed to develop a data-driven framework for alloy design, considering

the processing schedules of the rolling mill. Initially, seven different supervised machine learning (ML) algorithms were employed to model UTS and % Elongation for V-based microalloyed steel. Global feature importance were derived from SHAP values for these models. Model-agnostic conformal predictions were implemented to quantify uncertainty, enhancing the reliability of predictions. Given the challenges of inverse design in industrial contexts—such as multiple objectives, non-unique solutions, and large search spaces—the problem was approached as a multi-objective optimization (MOO) task focusing on the trade-off between strength and ductility, i.e. generating the best combination of strength and ductility. The best performing ML models for UTS and % Elongation were utilized as objective functions in MOO, with the Non-dominated Sorting Genetic Algorithm II (NSGA-II) employed to derive optimized Pareto front solutions. The thermomechanical processing parameters were integrated as strict constraints in the decision variable space of the NSGA-II. To visualize the solutions, t-distributed Stochastic Neighbor Embedding (t-SNE) was used to map them along with original rolling data into a two-dimensional space, which was then clustered using the K-means algorithm. Select representative solutions from each cluster were chosen to identify unique alloys. This research provides key applications in developing online property prediction tools, enhancing process understanding and aiding in both process control, and alloy design.

Résumé

L'objectif principal de cette recherche était d'utiliser des techniques axées sur les données pour prédire et optimiser les propriétés mécaniques des aciers microalliés lors du processus de laminage direct de minces dalles. Les données pour cette étude provenaient d'Algoma Steel Inc., situé en Ontario, Canada, et le travail était divisé en deux parties principales.

Dans la première partie, des modèles de réseaux de neurones profonds (DNN) ont été développés pour prédire les propriétés mécaniques, spécifiquement la résistance à la traction ultime (UTS) et la limite d'élasticité inférieure (LYS), des bandes laminées à chaud microalliées à base de niobium. Pour introduire l'explicabilité dans les modèles DNN, des explications additives SHapely basées sur la théorie des jeux (SHAP) ont été utilisées. Les valeurs SHAP ont fourni des insights sur les effets combinés de la composition chimique et des paramètres de traitement thermomécanique. L'influence des composition chimique a été corroborée par la théorie de la métallurgie physique, et des corrélations ont été établies avec la relation empirique de la température de non-recristallisation (T_{nr}) issue de la littérature existante. De plus, les données ont été analysées en utilisant le Stress Moyen d'Écoulement (MFS) par rapport à la température inverse, avec des comparaisons entre différentes épaisseurs et compositions pour soutenir les explications des modèles et suggérer des mécanismes métallurgiques sous-jacents. Ce segment de l'étude a mis en évidence d'importantes opportunités pour l'optimisation de la composition des alliages, ce

qui a mené à la deuxième partie de la recherche.

La deuxième partie visait à développer un pipeline axé sur les données pour la conception d'alliages, en tenant compte des calendriers de traitement du laminoir. Initialement, sept algorithmes d'apprentissage automatique supervisé différents ont été utilisés pour modéliser l'UTS et le % d'Allongement pour l'acier microallié à base de vanadium. L'importance globale des caractéristiques a été dérivée des valeurs SHAP pour ces modèles. Des prédictions conformes indépendantes du modèle ont été mises en œuvre pour quantifier l'incertitude, améliorant ainsi la fiabilité des prédictions. Étant donné les défis de la conception inverse dans les contextes industriels—tels que les objectifs multiples, les solutions non uniques et les grands espaces de recherche—le problème a été abordé comme une tâche d'optimisation multi-objectifs (MOO) se concentrant sur le compromis entre la résistance et la ductilité. Les meilleurs modèles d'apprentissage automatique pour l'UTS et le % d'Allongement ont été utilisés comme fonctions objectifs dans le MOO, avec l'Algorithme Génétique de Tri Non Dominé II (NSGA-II) employé pour dériver des solutions optimales du front de Pareto. Les paramètres de traitement thermomécanique ont été intégrés comme contraintes strictes dans l'espace des variables de décision du NSGA-II. Pour visualiser les solutions, l'incorporation stochastique de voisins par t-distribution (t-SNE) a été utilisée pour mapper ces solutions ainsi que les données de laminage originales dans un espace bidimensionnel, qui a ensuite été regroupé en utilisant l'algorithme de clustering K-means. Des solutions représentatives de chaque groupe ont été

choisies pour identifier des alliages uniques. Cette recherche offre des applications clés dans le développement d'outils de prévision des propriétés en ligne, améliorant la compréhension des processus et aidant à la fois dans le contrôle des processus et la conception des alliages.

Acknowledgements

I am immensely grateful for the unwavering support and guidance I received from many as a graduate student and throughout my journey. My deepest appreciation goes to my family—my parents, and my lovely sister, whose constant encouragement and support were indispensable to my success.

I extend heartfelt thanks to my supervisor, Prof. Stephen Yue, whose constant support and guidance have been key in shaping me into an independent researcher. Prof. Yue's enthusiasm for exploring new ideas and his support during challenging times have been crucial to my research progress. I am equally grateful to my co-supervisor, Prof. Narges Armanfard, whose expertise and engaging discussions significantly enriched my research experience. Her guidance has been a continuous source of inspiration.

My gratitude extends to Dr. Xiaoping Ma and Dr. Kashif Rehman from Algoma Steel Inc., who not only hosted my visit to their industrial complex but also provided invaluable data and support essential for the success of my research.

I have been fortunate to work alongside many brilliant colleagues in the Hot Deformation Lab at McGill University. I thank all lab members, especially Evelin, Manoj and John, for their camaraderie. Special thanks to Barbara, June and Leslie for their invaluable assistance with administrative tasks.

I am also thankful for the support and encouragement from my friends—Aakashdeep

Prasad, Amit Kumar, Azba Shaikh, Farhan Ahmed, Rohit Tiwari, and Sourav Sinha—during this important phase of my life.

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and our industrial partner, Algoma Steel Inc., to whom I am deeply thankful for their financial support. Lastly, I am grateful for the support from MITACS through the Graduate Globalink Fellowship and McGill University for the Graduate Excellence Awards, which have greatly contributed to my studies.

Contribution of Authors

This thesis is manuscript-based and comprises a collection of two papers. One has already been published in a peer reviewed journal, while the other is currently in the publication process. The manuscripts were a joint effort between researchers from McGill University and Algoma Steel Inc. Below, the description of each author's contribution is provided.

- Chapter 3 : Sinha, S., Guye, D., Ma, X., Rehman, K., Yue, S., & Armanfard, N. (2024). **Neural network prediction of the effect of thermomechanical controlled processing on mechanical properties.** *Machine Learning with Applications*, 100531.

Authors : Sushant Sinha, Denzel Guye, Xiaoping Ma, Kashif Rehman, Stephen Yue, Narges Armanfard.

Sushant Sinha developed the machine learning models, carried out the computational analysis, and prepared the manuscript. Denzel Guye contributed to the methodology conceptualization, and writing the manuscript draft. Xiaoping Ma and Kashif Rehman provided the required data and reviewed the manuscript. Stephen Yue and Narges Armanfard supervised the project and edited the manuscript.

- Chapter 4 : Sinha, S., Ma, X., Rehman, K., Armanfard, N., & Yue, S., (2024). **An Interpretable and Reliable Framework for Alloy Discovery in**

Thermomechanical Processing (to be submitted).

Authors : Sushant Sinha, Xiaoping Ma, Kashif Rehman, Narges Armanfard, Stephen Yue.

Sushant Sinha conceptualized the project, developed the machine learning framework, carried out the computational analysis, and prepared the manuscript. Xiaoping Ma and Kashif Rehman provided the required data and reviewed the manuscript. Stephen Yue and Narges Armanfard supervised the project and edited the manuscript.

Table of Contents

Abstract	ii
Résumé	iv
Acknowledgements	viii
Contribution of Authors	x
List of Figures	xvii
List of Tables	xviii
1 Introduction	1
1.1 Thesis Objectives	4
1.2 Thesis Structure	4
2 Background	8
2.1 Thin Slab Direct Rolling	8
2.1.1 TSDR Process Overview	8
2.2 Basic Theory of Rolling	11
2.3 Metallurgical Events During Hot Rolling	16
2.3.1 Plastic Deformation	16
2.3.2 Recovery	17
2.3.3 Static Recrystallization	18

2.3.4	Dynamic Recrystallization	20
2.3.5	Metadynamic Recrystallization	22
2.3.6	Grain Coarsening	23
2.3.7	Precipitation	23
2.3.8	Phase Transformation	25
2.4	TMCP in TSDR	26
2.5	Microalloying	27
2.6	Machine Learning	29
2.6.1	Supervised Learning	29
2.6.2	Unsupervised Learning	31
3	Neural Network Prediction of the Effect of Thermomechanical Controlled Processing on Mechanical Properties	38
3.1	Introduction	39
3.2	Neural Network Approaches	42
3.3	Modeling Methods	44
3.3.1	Preprocessing	45
3.3.2	Prediction	47
3.3.3	Model Interpretation	50
3.4	Results	52
3.5	Discussion	61

3.6	Conclusion	67
4	An Interpretable and Reliable Framework for Alloy Discovery in Thermomechanical Processing	73
4.1	Introduction	74
4.2	Methodology	79
4.2.1	Dataset Background	79
4.2.2	Pre-processing	80
4.2.3	Predictive Modelling	81
4.2.4	Uncertainty Quantification	87
4.2.5	TreeSHAP : Model Interpretability	90
4.2.6	Multiobjective Optimization	92
4.2.7	Post-Pareto Analysis	94
4.2.8	Implementation Details	98
4.3	Results and Discussion	99
4.3.1	Predictive Modelling	99
4.3.2	Uncertainty Quantification	100
4.3.3	Model Interpretability	102
4.3.4	Multiobjective Optimization	102
4.3.5	Post-Pareto Analysis	106
4.4	Conclusion	109

5 Conclusion	125
5.1 Global Discussion and Conclusions	125
5.2 Proposed Applications	127
5.3 Future Work	128

List of Figures

1.1	General workflow proposed for alloy design in Chapter 4. The ML framework captures the crux of objectives outlined in the thesis.	5
2.1	Schematic of the TSDR mill at Algoma Steel Inc., Ontario, Canada	9
2.2	Schematic of Roll and Strip Geometry during Hot Deformation. (5)	12
2.3	Schematic showing various stages in the recovery of a plastically deformed material (11)	18
2.4	Schematic showing various stages of grain evolution in the dynamic recrystallization. (11)	22
2.5	Metallurgical mechanisms and reactions at different stages of TSDR process. (28)	26
3.1	Schematic of Direct Strip Processing Complex	42
3.2	Hyperparameter grid search for the neural network model with varying model complexity and regularization strength for predicting UTS. (a) Training Error and Validation Error, (b) PAM 5% and 2.5%. The model selected is the neural network trained with (3000, 100) nodes and no regularization.	53

3.3	Hyperparameter grid search for the neural network model with varying model complexity and regularization strength for predicting LYS. (a) Training Error and Validation Error, (b) PAM 5% and 2.5%. The model selected is the neural network trained with (3000, 2000) nodes and no regularization.	54
3.4	Prediction by neural network models with selected hyperparameters of (a) UTS and (b) LYS	57
3.5	Residual plot for the prediction of selected Neural Networks. (a) Residual plot for UTS, (b) Residual plot for LYS.	58
3.6	Effect of TMCP parameters on UTS. (a) Shows the combined effect of TMCP on UTS against the relative gauge, and (b) shows the combined effect of chemistry on UTS against the relative gauge	59
3.7	Effect of TMCP parameters on LYS. (a) Shows the combined effect of TMCP on LYS against the relative gauge, and (b) shows the combined effect of chemistry on LYS against the relative gauge	60
3.8	SIMS MFS Plot for low Nb composition in different thickness	63
3.9	(a) Flow stress behaviour of two different Nb compositions with mean gauge 4.5 mm. (b) Variation of estimated no recrystallization temperature (T_{nr}) with mean gauge based on changes in chemical composition.	65

3.10	Predicted effect of variation in strength due to chemical compositions observed in the dataset using a single deformation schedule on (a) UTS and (b) LYS. Note that the axes in (a) and (b) are different.	66
4.1	Performance of predictive models with uncertainty as colour mapping. (a) True vs Predicted for UTS , (b) True vs Predicted for Elongation.	101
4.2	Feature Importance plot based on the selected XGBoost Models and mean shap values (a) For Ultimate Tensile Strength (b) For Elongation.	103
4.3	NSGA-II results (a)Convergence of Objectives over generations , (b) Optimized Pareto Front with 95% Confidence Interval. The actual distribution is also presented in the background for comparison.	105
4.4	Overview of clustering results and analysis : (a) Clustered Pareto fronts using t-SNE and K-means, background is t-SNE compressed train data, (b) Selection of optimal number of clusters using elbow method for inertia, (c) Cluster statistics to understand the distinct characteristics of each cluster.	107

List of Tables

1.1	A summary of common alloying elements used in microalloyed steels and their influence (15).	3
4.1	Descriptive statistics of the dataset	84
4.2	Summary of XGBoost hyperparameters tuned in this work.	85
4.3	Performance Metrics for various ML models for UTS and Elongation (Test Set).	99

Chapter 1

Introduction

The commercial development of microalloyed steel began in 1958 when the Great Lakes Steel Corporation in the USA demonstrated significant improvements in strength by adding small amounts of Niobium (0.005-0.03%) to carbon-manganese steels. For instance, GLX-60-W, a carbon-manganese grade steel with an initial yield strength of 300 MPa, saw its strength increase to an impressive 415 MPa with these "micro" Niobium additions (1). Since this development, researchers have made substantial advancements in both the composition and rolling strategies of microalloyed steels, also known as high-strength low-alloy (HSLA) steels (2). These steels now constitute about 12% of global steel production (3) and are extensively used in applications such as bridges (4), structural components (5), transportation (6), and the oil and gas sectors (7).

The microalloying elements Nb, Ti, and V primarily contribute to steel strengthening through three mechanisms : austenite conditioning, precipitation hardening, and solid solution strengthening (8; 9). Among these, austenite conditioning is perhaps the most crucial, involving strain-induced precipitation within the austenite phase that hinders further recrystallization. This mechanism retains work hardening from deformation in the low-temperature austenite region, leading to the formation of pancaked grains and

deformation bands. Consequently, this creates more nucleation sites for phase transformation, promoting the development of a fine-grained ferrite microstructure (10).

In addition to microalloying elements, several other alloying elements present in steel, as detailed in Table 1.1, interact with the Thermomechanical Controlled Processing (TMCP) parameters as well. This huge number of influencing parameters and their inter-relationships complicates the study of their influence on mechanical properties through experiments or isolated trials, which can be costly as well as time-consuming. Therefore, this complexity underscores the necessity for a "model" that can accurately represent and optimize the coupled effects of composition and TMCP. The conventional method for modeling material properties has largely relied on mathematical modeling rooted in empirical relationships (11; 12). Consider the dynamics of static recrystallization, which is well-understood to involve nucleation and growth mechanisms (13); researchers typically model these processes using the Avrami equation (14). However, the coefficients derived from these equations can vary widely between studies. Furthermore, parameters calibrated under controlled laboratory conditions do not always translate well to industrial conditions. Similarly, the structure-property relationship for yield stress is frequently modeled using an extended Hall-Petch equation, which accounts for various strengthening mechanisms. Yet, extending such an approach to predict other mechanical properties such as Ultimate Tensile Strength (UTS) or Percent Elongation, is far from the obvious. The use of multiple sub-models in this framework can amplify errors, leading to unreliable final

predictions, as each layer of the model introduces potential inaccuracies.

Element	wt-% in steel	Influence
C	<0.25	Strengthener
Mn	0.5–2.0	Delays austenite decomposition during accelerated cooling ; Decreases ductile to brittle transition temperature ; Strong sulphide former
Si	0.1–0.5	Deoxidiser in molten steel ; Solid solution strengthener
Al	<0.02	Deoxidiser ; Limits grain growth as AlN
Nb	0.02–0.06	Very strong ferrite strengthener as Nb(C,N) ; Delays $\gamma \rightarrow \alpha$ transformation
Ti	0–0.06	γ grain size control by TiN ; Strong ferrite strengthener
V	0–0.10	Strong ferrite strengthener by V(C,N)
Zr	0.002–0.05	γ grain size control [Zr(C,N)] ; Strong sulphide former
N	<0.012	Forms nitrides and carbonitrides
Mo	0–0.3	Promotes bainite formation ; Ferrite strengthener
Ni	0.5	Increases fracture toughness
Cu	0–0.55	Improves corrosion resistance ; Ferrite strengthener
Cr	0.1–1.25	With Cu, increases atmospheric corrosion resistance
B	0.0005	Promotes bainite formation

Table 1.1 – A summary of common alloying elements used in microalloyed steels and their influence (15).

With these challenges in mind, there is a strong incentive to investigate alternative methods, especially in industries like hot rolling, where there is ample stored data to support more flexible and dynamic approaches. Machine learning offers an obvious solution with its ability to learn from data and provide iterative improvement.

1.1 Thesis Objectives

This thesis is driven by two primary objectives :

- (I) **Development of Predictive Models** : To develop machine learning models capable of predicting the mechanical properties of hot rolled microalloyed steel. A crucial aspect of this objective is to use these models to derive and enhance metallurgical understanding of the processes involved.
- (II) **Alloy Design** : To integrate machine learning models with evolutionary genetic algorithms to establish a comprehensive framework for alloy design, tailored to specific processing schedules. The focus is on proposing alloys that exhibit superior strength-ductility combinations. The study also aims to incorporate tools that enhance the explainability and reliability of the developed frameworks, ensuring their practical applicability and acceptance in the field.

1.2 Thesis Structure

This manuscript-based thesis is organized into five chapters, each designed to systematically address the research objectives outlined above. The first and current chapter introduces the broader research context, motivation and establishes the rationale for the study. Chapter 2 discusses the key background concepts of Thin Slab Direct Rolling

An Interpretable & Reliable Data-driven Framework for Alloy Discovery

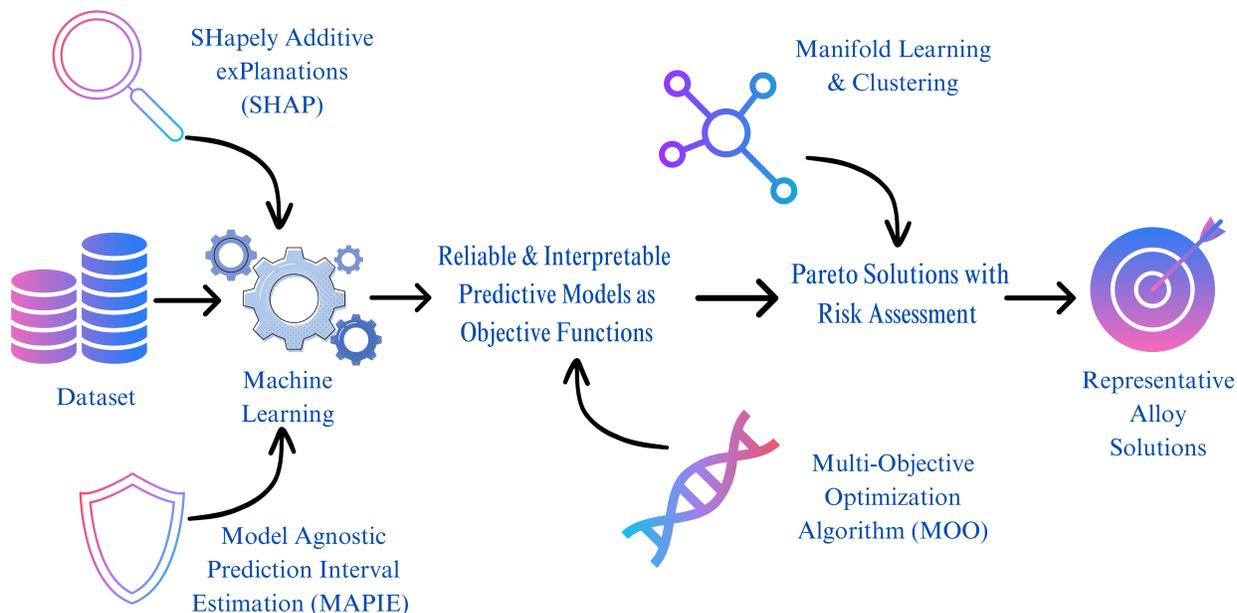


Figure 1.1 – General workflow proposed for alloy design in Chapter 4. The ML framework captures the crux of objectives outlined in the thesis.

(TSDR), theory of hot deformation, key metallurgical events, and basics of machine learning. Chapters 3 and 4 are the manuscripts which consist of the results and discussion of the work. In Chapter 3, we address the first objective by developing machine learning models to predict mechanical properties and assess the effect of thermomechanical controlled processing and the composition. Chapter 4 expands the scope to alloy design, detailing the integration of machine learning with NSGA-II, an evolutionary genetic algorithm. Graphic abstract of chapter 4 is presented in figure 1.1. The final and concluding chapter synthesizes the key findings, suggests industrial applications, and proposes directions for future research.

In order to avoid redundancy, the background of specific ML methods, optimization algorithms, explainability and uncertainty quantification methods is not discussed in the 2nd chapter i.e., Background. In the manuscripts (chapters 3rd & 4th), these methods are discussed and their literature survey related to the field of this work is summarized as well.

References

- [1] F. W. Starratt, Columbium-treated steels low-cost, high-strength, JOM 10 (12) (1958) 799–799.
- [2] W. Morrison, Microalloy steels—the beginning, Materials Science and Technology 25 (9) (2009) 1066–1073.
- [3] [Www.totalmateria.com/en-us/articles/application-of-microalloyed-hsla-steel-1/](http://www.totalmateria.com/en-us/articles/application-of-microalloyed-hsla-steel-1/), Accessed online : July 10, 2024.
- [4] D. Jarreta, Niobium micro-alloyed high carbon steel wires : innovative solution for bridges and infrastructure, SEASI Quarterly J 6 (2) (2017) 29–36.
- [5] D. Bhattacharya, Microalloyed steels for the automotive industry, Tecnologia em Metalurgia, Materiais e Mineração 11 (4) (2014) 371.
- [6] J. Adamczyk, Development of the microalloyed constructional steels, Journal of Achievements in Materials and Manufacturing Engineering 14 (1-2) (2006) 9–20.

- [7] D. Stalheim, Metallurgical optimization of microalloyed steels for oil and gas transmission pipelines, in : Proceedings of 6th International Conference on High Strength Low Alloy Steels (HSLA Steels 2011). Beijing China, 2011.
- [8] A. DeArdo, M. Hua, K. G. Cho, C. Garcia, On strength of microalloyed steels : an interpretive review, *Materials Science and Technology* 25 (9) (2009) 1074–1082.
- [9] T. Baker, Microalloyed steels, *Ironmaking & Steelmaking* 43 (4) (2016) 264–307.
- [10] S. Yue, Thermomechanical processing of ferrous alloys 14A (2005) 286–296.
- [11] H. PD, A mathematical model to predict the mechanical properties of hot rolled c-mn and microalloyed steels, *ISIJ international* 32 (12) (1992) 1329–1338.
- [12] F. Siciliano, Mathematical modeling of the hot strip rolling of nb microalloyed steels, Ph.D. thesis, McGill University (1999).
- [13] J. Kliber, I. Schindler, Recrystallization/precipitation behaviour in microalloyed steels, *Journal of materials processing technology* 60 (1-4) (1996) 597–602.
- [14] M. Avrami, Kinetics of phase change. i general theory, *The Journal of chemical physics* 7 (12) (1939) 1103–1112.
- [15] S. Vervynckt, K. Verbeken, B. Lopez, J. Jonas, Modern hsla steels and role of non-recrystallisation temperature, *International Materials Reviews* 57 (4) (2012) 187–207.

Chapter 2

Background

2.1 Thin Slab Direct Rolling

The use of Thin Slab Direct Rolling (TSDR) technology in steel production has brought remarkable improvements in efficiency, cost, and environmental impact. TSDR was first introduced at Nucor, Crawfordsville, USA in 1989 based on Compact Strip Casting (CSP) technology developed by the SMS company. The key idea was to cast steel directly into thin 50 mm slabs, which are directly fed to the hot rolling mill, preserving a significant amount of heat from the casting phase (1). There have been several developments and hence variants of this technology since then. Here we only discuss the general idea behind it with focus on the TSDR mill installed at Algoma Steel Inc., Ontario, Canada (DSPC). The figure 2.1 depicts the schematic of the mill.

2.1.1 TSDR Process Overview

This process encompasses several stages, the key sections of which are discussed below (2; 3) :

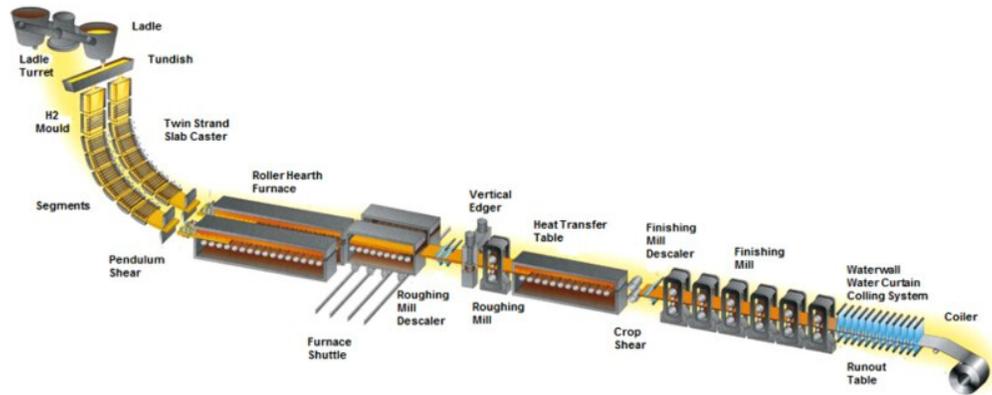


Figure 2.1 – Schematic of the TSDR mill at Algoma Steel Inc., Ontario, Canada (2)

2.1.1.1 Continuous Caster

The continuous caster is essential for solidifying molten steel into thin slabs about 70-80 mm in thickness and 800-1600 mm in width. Steel flows from a ladle into a tundish, which then feeds it into a water-cooled mold. This setup facilitates rapid solidification, crucial for achieving the desired upstream slab macrostructure for further rolling. Enhanced cooling rates and dynamic soft reduction techniques are employed to minimize internal defects and achieve uniformity in the cast slab. Macrostructure control in TSDR becomes more important given the lack of long reheating times in the process, hence plays a key role in the final property of the strip.

The as-cast slab is cut using the pendulum shear and further passed through a rotary descaler, which produces a scale-free slab surface.

2.1.1.2 Tunnel Furnace

The tunnel furnace serves to reheat and homogenize the temperature of the slab to obtain the desired flow behaviour for subsequent rolling. Typically, the average residence times are approximately 20 minutes, varying based on the chemistry. By regulating the oxygen % in the furnace, the roller hearth furnace can control the slab temperature. Oxygen % is also important for controlled growth of secondary scales. The shuttle section of the furnace is utilized to deliver the homogenized slabs to the roughing mill as quickly possible.

2.1.1.3 Roughing Mill

Before the roughing mill, the slab is passed through secondary descaling unit and vertical edger. The edgers provide width control of the slab before rolling and improves the yield of tapered slabs. The roughing stand here is a single pass and non-reversing mill. The main purpose is to drastically reduce the slab to a transfer bar about 30-45 mm thick, ideal to feed the finishing mill.

2.1.1.4 Finishing Mill

The transfer bar from the roughing mill is passed through a heated transfer table, a third descaling unit to ensure surface cleanliness, and a rotary drum crop shear. This prepares the bar for rolling in a six-stand finishing mill. The key focus of the finishing mill is to ensure tight dimensional tolerance of the product. This is unarguably the most interesting section

where the majority of metallurgical action occurs. The strain and strain rate applied at the subsequent stands drives metallurgical reactions like recrystallization or precipitation, which, in-turn, play a key role in the final mechanical property of the rolled strip.

2.1.1.5 Runout Table and Coiler

The steel strip is rapidly cooled to a coiling temperature with tight temperature tolerance along the strip on the runout table following the finishing mill. The importance of accelerated cooling and the cooling stop temperature lies in its ability to obtain a uniform and refined microstructure, giving the mechanical properties needed for specific applications.

Following this the strip is coiled in the downcoiler. This coil can either be dispatched in as-hot-rolled condition or further processed downstream. The mechanical properties discussed in this work are experimentally measured by sampling from these as-hot-rolled coils.

2.2 Basic Theory of Rolling

The theory of rolling developed by Orowan (4) is comprehensive as it considers both the inhomogeneous distribution of deformation across the roll gap and friction at the material-roll interface, aspects previously ignored by other theories. He highlighted the variability in frictional conditions within a single pass, shifting from the Coulomb condition, where the interfacial shear stress τ is equal to μs (μ being the friction coefficient and s the local pressure), to conditions where τ equals the shear yield stress of the material. He introduced

an inhomogeneity factor to account for this variability in deformation, moving away from the simplistic assumption of "plane sections remain plane."

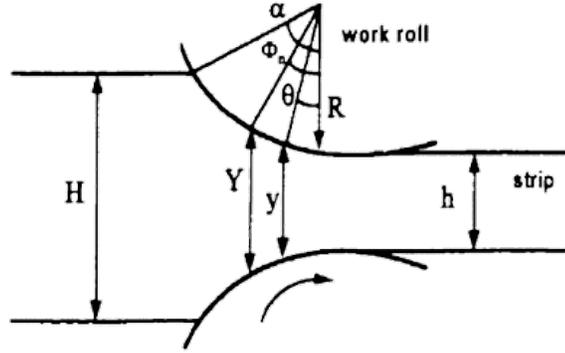


Figure 2.2 – Schematic of Roll and Strip Geometry during Hot Deformation. (5)

To incorporate the non-uniform deformation and the temperature and strain rate dependencies of the yield stress, Orowan used numerical integration, a method complex enough for practical applications that later researchers sought simpler solutions. Alexander (6) left out the inhomogeneity factor and used a fourth order Runge-Kutta technique to determine the normal pressure distribution and then the roll force. Sims (7) made still more simplifying assumptions to Orowan's theory to obtain an analytical solution, which is still widely accepted.

We first consider the differential equation 2.1 as the base which is basically derived from considering the equilibrium of a thin segment of unit width bounded by planes parallel to the roll axes planes.

$$s(\tan \theta \pm \mu)R' d\theta - \frac{1}{2}d(ty) = 0 \quad (2.1)$$

In this equation, s represents the normal roll pressure, which is the pressure exerted by the rolls perpendicular to the surface of the metal. The angular coordinate θ measures the position along the arc of contact in the roll gap, and μ is the coefficient of friction between the rolls and the metal. The term R' indicates the deformed radius of the work roll given by Hitchcock's formula (8), taking into account the flattening effect due to the rolling pressure. The differential element $d\theta$ represents a small change along the angle θ , and $d(ty)$ signifies the differential change in the product of the horizontal stress t of the rolled material and its thickness y .

Sims considered that sticking friction occurs between the work roll and the workpiece, simplifying the calculation process. This approach is commonly utilized due to its simplicity and ability to derive analytical results. Two key assumptions are made : firstly, sticking friction $\mu_s = \frac{k}{2}$ is assumed to prevail over the entire arc, and secondly, the angle θ in hot rolling is considered so small that we can approximate $\sin \theta = \tan \theta = \theta$, $\cos \theta = 1$, and $(1 - \cos \theta) = \frac{\theta^2}{2}$. Under these assumptions, the equation transforms to :

$$\frac{dT}{d\theta} = 2R's\theta + R'k, \quad (2.2)$$

where T is the horizontal force per unit width acting on the strip, and k is the yield stress of the material under plane compression.

Sims also adopted the assumption from Orowan that the rolling process resembles deformation between two inclined plates, an idea initially proposed by Nadai (9). This

leads to :

$$T = y\left(s - \frac{\pi}{4}k\right), \quad (2.3)$$

Substituting Equation 2.3 into the simplified model, we derive :

$$\frac{d\left[y\left(s - \frac{\pi}{4}k\right)\right]}{d\theta} = 2R's\theta \pm R'k, \quad (2.4)$$

Considering geometry and using the small angle approximation, the relation for y becomes :

$$y = h + R'\theta^2, \quad (2.5)$$

where h is the thickness of the rolled strip. Substituting Equation 2.5 into Equation 2.4, and assuming k remains constant over the arc of contact, we establish :

$$\frac{d\left(s/k - \frac{\pi}{4}\right)}{d\theta} = \frac{R'\pi\theta}{2(h + R'\theta^2)} \pm \frac{R'}{(h + R'\theta^2)}, \quad (2.6)$$

Finally, solving for the above equation 2.6 for exit side from the neutral plane, we obtain :

$$s^+/k = \frac{\pi}{4} \log_e \left(\frac{y}{h}\right) + \frac{\pi}{4} + \sqrt{\frac{R'}{h}} \tan^{-1} \left(\sqrt{\frac{R'}{h}}\right), \quad (2.7)$$

and for the entry side, the relationship becomes :

$$s^-/k = \frac{\pi}{4} \log_e \left(\frac{y}{H} \right) + \frac{\pi}{4} + \sqrt{\frac{R'}{h}} \tan^{-1} \left(\sqrt{\frac{R'}{h}} \right) \alpha - \sqrt{\frac{R'}{h}} \alpha - \sqrt{\frac{R'}{h}} \tan^{-1} \left(\sqrt{\frac{R'}{h}} \right) \theta. \quad (2.8)$$

To determine the plane of intersection in the rolling process, we can solve for the point where the solutions for the entry and exit sides converge. This results in the following equation :

$$\frac{\pi}{4} \log_e(1-r) = 2\sqrt{\frac{R'}{h}} \tan^{-1} \left(\sqrt{\frac{R'}{h}} \right) \phi - \sqrt{\frac{R'}{h}} \tan^{-1} \left(\sqrt{\frac{r}{1-r}} \right) \quad (2.9)$$

The difference between the normal roll pressure and the vertical pressure becomes minimal when the angular coordinate θ is small, assuming that plane deformation occurs. Under these conditions, the specific roll load can be expressed as :

$$P = R' \int_0^\alpha s d\theta, \quad (2.10)$$

Substituting from equations 2.7 and 2.8 into equation (2.10), the expression for P becomes

more complex, involving the integration of several terms :

$$P = R'k \left\{ \int_0^\phi \left(\frac{\pi}{4} \log_e \left(\frac{y}{H} \right) + \frac{\pi}{4} + \sqrt{\frac{R'}{h}} \tan^{-1} \left(\sqrt{\frac{R'}{h}} \right) \alpha - \sqrt{\frac{R'}{h}} \alpha - \sqrt{\frac{R'}{h}} \tan^{-1} \left(\sqrt{\frac{R'}{h}} \theta \right) \right) d\theta \right. \\ \left. + \int_\phi^\alpha \left(\frac{\pi}{4} \log_e \left(\frac{y}{h} \right) + \frac{\pi}{4} + \sqrt{\frac{R'}{h}} \tan^{-1} \left(\sqrt{\frac{R'}{h}} \right) \right) d\theta \right\}. \quad (2.11)$$

After simplification and integration, the specific roll force is obtained as :

$$P = R'k \left\{ \frac{\pi}{2} \sqrt{\frac{h}{R'}} \tan^{-1} \left(\sqrt{\frac{r}{1-r}} \right) - \frac{\pi\alpha}{4} - \log_e \left(\frac{y}{H} \right) + \frac{1}{2} \log_e \left(\frac{H}{h} \right) \right\}, \quad (2.12)$$

where $r = \frac{\delta}{H}$ represents the reduction ratio, δ is the draft (difference between the entry and exit thickness), H is the entry thickness, and h is the exit thickness. This equation can be rearranged to obtain SIMS mean flow stress (MFS) formulae after accounting for a factor of $\frac{2}{\sqrt{3}}$ to allow for plane strain.

2.3 Metallurgical Events During Hot Rolling

2.3.1 Plastic Deformation

The initial and visible effect of applying strain during rolling is plastic deformation of the material. This results in the generation and subsequent movement of dislocations leading to work hardening, or strain hardening. As the deformation continues, an increase

in dislocation density leads to increased interactions among dislocations themselves and with existing grain boundaries. These interactions impede the movement of dislocations, requiring higher stresses to sustain further deformation (10). Hence, the work hardened material experiences an increase in strength.

2.3.2 Recovery

Recovery is a softening mechanism which serves to release the strain energy stored within the material due to dislocation multiplication. This can be either static i.e. post-deformation or dynamic i.e. during-deformation. Recovery does not involve grain boundary motion but rather the local annihilation and rearrangement of dislocations into lower energy configurations. Both annihilation and rearrangement are achieved by glide, climb, and cross-slip of dislocations. The process may consist of a series of steps as detailed in figure 2.3

Apart from the strain and temperature, the most important parameter affecting recovery is the stacking fault energy (SFE) of the material which determines the rate of dislocation climb and cross slip. Hence, for high SFE materials like Al, BCC-Fe recovery is favourable unlike for low SFE materials like FCC-Fe (11).

Note that, recovery and recrystallization are competing processes. For medium and low SFE materials the slow kinetics of recovery promotes dislocation density increase, facilitating recrystallization.

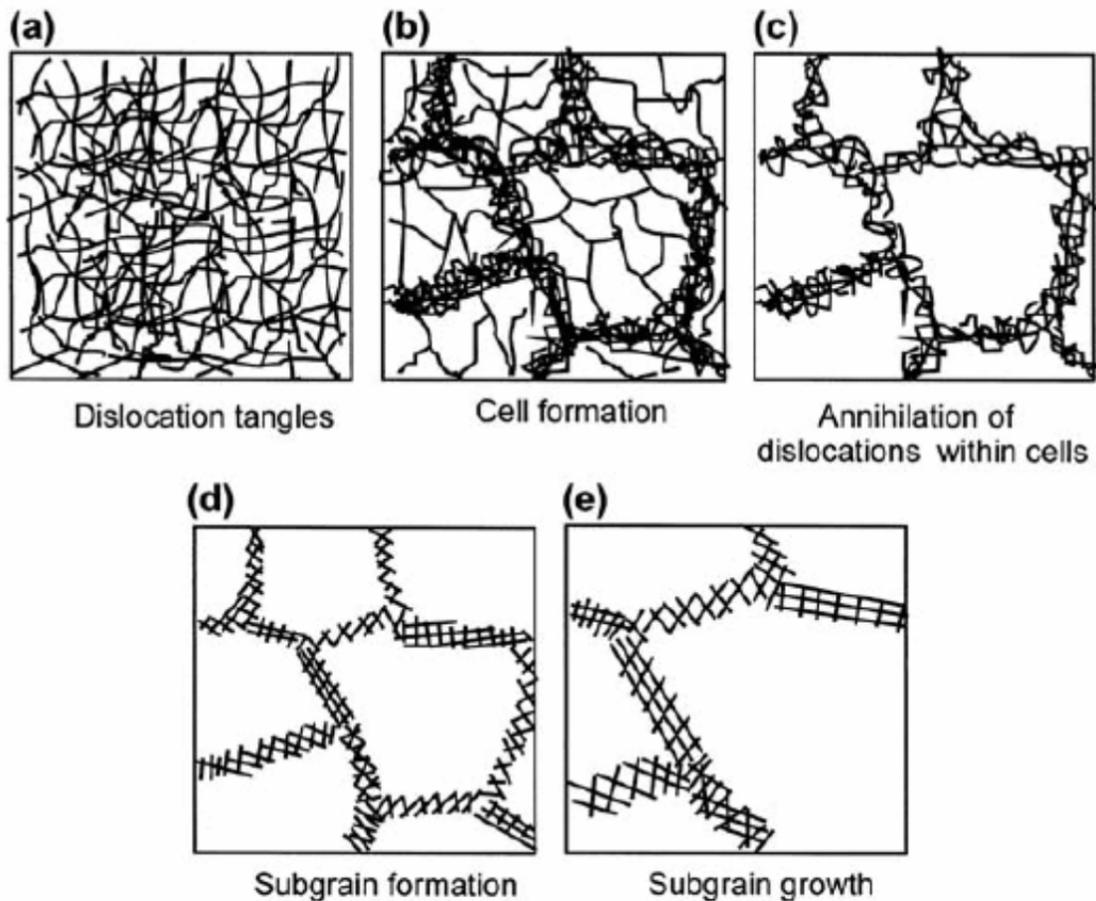


Figure 2.3 – Schematic showing various stages in the recovery of a plastically deformed material (11)

2.3.3 Static Recrystallization

Static recrystallization (SRX) involves the migration of high-angle boundaries, which effectively annihilates dislocations and thereby releases the stored energy within the material. Typically occurring between passes, particularly after deformation and during the interpass time in hot rolling, the driving force behind SRX is the strain energy previously stored in

the material in the form of dislocations (12).

SRX follows a nucleation and growth transformation. Nucleation of new grains takes place at sites with highest local deformation. These sites can be grain boundaries, deformation bands, and inclusions. Nucleation is a thermally activated process. The mechanism of nucleation may involve subgrain growth, or subgrain coalescence, or strain induced grain boundary migration. In subgrain growth, dislocations rearrange themselves as low angle grain boundaries in the area of high deformation. The subsequent interaction of dislocations with subgrain boundaries leads to increased misorientation and hence results in formation high angle grain boundaries. In subgrain coalescence, rotation of subgrain to reduce misfit strain with its neighbour results in nucleation. The common grain boundary disappears and high angle grain boundaries are created. In strain induced grain boundary migration, the grain boundaries between low-strain and high-strain subgrains bulges out to form a coarser grain. As a result of the difference in strain of neighbouring grains, the dislocations are swept away during bulging and leads to a strain free area (11; 13). Once a high angle boundary is formed, the growth process follows. The growth of new grains will depend on the boundary mobility and and net driving force on the boundary (11).

SRX being a nucleation and growth process, its kinetics can be described by the Avrami (14) equation. The transformation kinetics are represented by the equation :

$$X = 1 - \exp(-bt^n) \tag{2.13}$$

where X is the fraction transformed, b is a constant dependent on the nucleation and growth rates, t is the time, and n is the time exponent. Equation 2.13 will describe a sigmoidal-type curve typical of nucleation and growth transformations under isothermal conditions (11). For practical purposes, the recrystallization kinetics are defined in terms of 50% recrystallization time and hence the recrystallized fraction can be obtained from the Avrami equation.

2.3.4 Dynamic Recrystallization

Dynamic Recrystallization or DRX is a rapid softening mechanism occurring during deformation when the applied strain exceeds a critical strain, ϵ_c . This critical strain corresponds to a critical dislocation density which is more favourable to reach in low SFE materials such as austenite ($\gamma - Fe$). There exists a characteristic peak strain (ϵ_p) in stress-strain plot which denotes the maximum flow stress reached before eventually reaching a steady state (15).

The ϵ_p parameter is represented as :

$$\epsilon_p = BD_0^m Z^p \quad (2.14)$$

where, B, m and p are material dependent coefficient, D_0 is the initial grain size, and Z is the Zener-Hollomon parameter given by the equation :

$$Z = \dot{\epsilon} \exp\left(\frac{Q_{\text{def}}}{RT}\right) \quad (2.15)$$

where, $\dot{\epsilon}$ is the strain rate, Q_{def} is the apparent activation energy for deformation, R is the gas constant (8.31 J/mol · K), and T is the absolute temperature. The critical strain can then be represented as (16) :

$$\epsilon_c = k \cdot \epsilon_p \quad (2.16)$$

Figure 2.4 depicts the typical stages involved in DRX. New grains typically nucleate along the high-angle grain boundaries; which may be the original grain boundaries, boundaries of dynamically recrystallized grains, or high-angle boundaries associated with deformation bands or deformation twins. However, with continued material deformation, these new grains accumulate dislocations. This accumulation diminishes the driving force necessary for further grain growth, eventually limiting the growth of recrystallizing grains. The growth of these new grains may be further constrained by the nucleation of additional grains along the migrating boundaries. If there is a large difference between the initial grain size and the recrystallized grain size, the "necklace" structure can be observed (11).

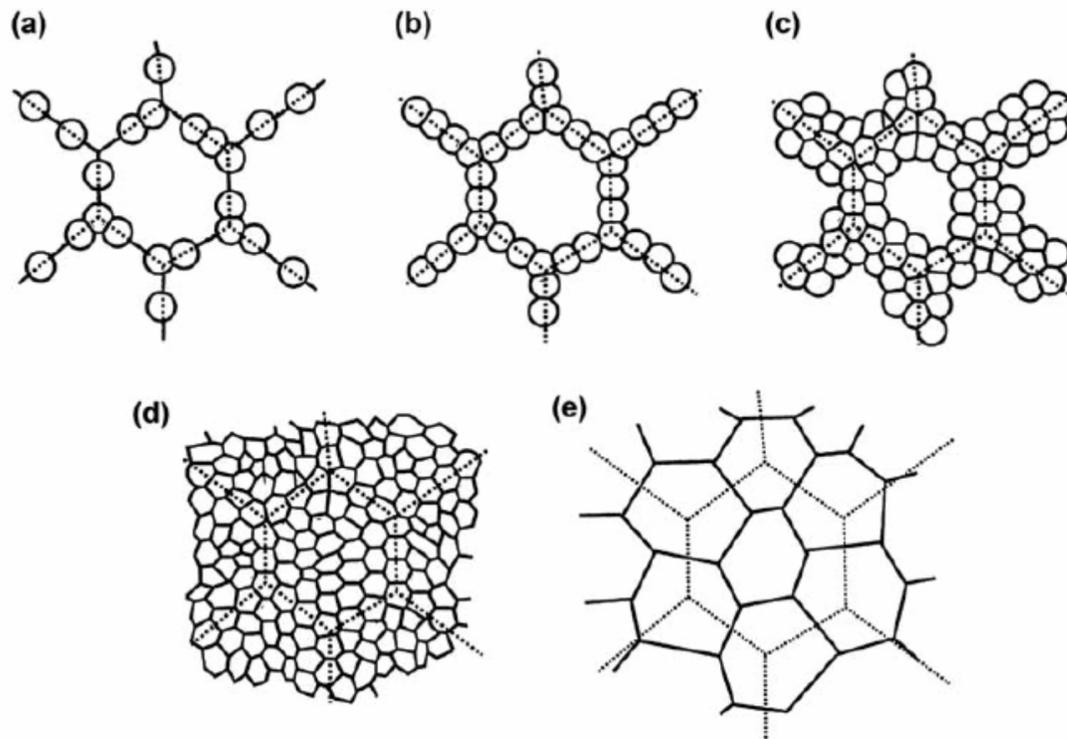


Figure 2.4 – Schematic showing various stages of grain evolution in the dynamic recrystallization. (11)

2.3.5 Metadynamic Recrystallization

Metadynamic recrystallization (MDRX) is actually post-dynamic recrystallization. After deformation stops, the nuclei initiated by dynamic recrystallization continues to grow which is called as MDRX. Hence, MDRX does not require any incubation time as the nuclei are already formed from DRX. This also explains the rapid kinetics in MDRX (11). As a result, DRX microstructures rapidly change post-deformation which results in coarser grain size (17). This can be favoured in hot rolling process schedules where the strain applied exceeds

the peak strain (ϵ_p) but never reach the steady state region.

In MDRX, the strain rate, rather than the strain itself, determines the kinetics. Higher strain rates would mean higher steady state stress which may result in a higher dislocation density and, therefore, a stronger driving force (18).

2.3.6 Grain Coarsening

Grain coarsening follows the completion of recrystallization. Note that it is not the growth stage of recrystallization, hence the use of the word 'coarsening'. In grain growth, the grain boundary area is reduction is the driving force, which therefore results in coarser grains. The high temperature conditions of hot rolling favours the same (5). The general empirical equation can be written as (11) :

$$D^n = D_R^n + ct \exp\left(-\frac{Q_g}{kT}\right) \quad (2.17)$$

where c , n , and Q_g are constants.

Since grain coarsening takes place in fully recrystallized regions, some authors consider 95% recrystallization time as the time when grain coarsening is initiated (5; 19).

2.3.7 Precipitation

Microalloying additions such as Nb, V, or Ti in the solution precipitate as carbides, nitrides or complex carbonitrides. Precipitation can be accelerated, either strain induced or

transformation induced.

Strain induced precipitates are generally incoherent do not contribute to strength but can retard recrystallization (20). The retention of work hardening occurs due to deformation in the low temperature austenite region, results in the formation of pancaked austenite grains and deformation bands. Thus, more sites for the γ to α transformation are created, encouraging the formation of a fine-grained ferrite microstructure. Since, strain induced precipitation is also a nucleation and growth mechanism, its kinetics can be studied empirically using the well-known Avrami equation (14). In mathematical modelling, it is commonly assumed that if precipitation begins, represented by 5% precipitation time, before recrystallization is complete during the interpass time, it becomes the controlling mechanism and hinders further recrystallization (21).

Strengthening of the matrix occurs when coherent precipitates are formed during or after transformation. Strengthening occurs as the motion of dislocations is impeded by precipitates. When a dislocation encounters a precipitate, it can react in two possible ways. If the precipitate is impenetrable, the dislocation may form a loop around the particle, which is called the Orowan mechanism (22). On the other hand, the dislocation might pass through the precipitate and shear it (23). The equations below describe the increase in strength resulting from particle looping and particle cutting (24) :

$$\sigma_{\text{prec}} \propto \frac{f^{1/2}}{r} \quad (\text{particle looping}) \quad (2.18)$$

$$\sigma_{\text{prec}} \propto (fr)^{1/2} \quad (\text{particle cutting}) \quad (2.19)$$

where f is the volume fraction and r is the radius of the precipitates.

2.3.8 Phase Transformation

The primary goal in processing microalloyed steels is to maximize the number of ferrite grains derived from a single austenite grain, effectively increasing the ferrite nucleation rate (25). Nucleation begins at corner sites, the junctions between four grains, then at edge sites where three grains meet, and finally at grain surfaces between two adjacent grains (26).

In fully recrystallized austenite structure, ferrite nucleation primarily occurs at recrystallized grain boundaries. The higher grain boundary area per unit volume accelerates the transformation and leads to a significant refinement of the ferrite structure. Alternatively, if precipitates impede recrystallization, austenite undergoes significant deformation, resulting in ‘pancaking’ and the creation of many intragranular nucleation sites, thereby increasing the ferrite nucleation rate in comparison to fully recrystallized austenite (27).

2.4 TMCP in TSDR

The final mechanical property of the rolled material is reflected through its microstructure. Hence, it is essential to understand the microstructure evolution during the process and the mechanisms/reactions driving them. Figure 2.5 shows the different metallurgical events that happen during the TSDR process.

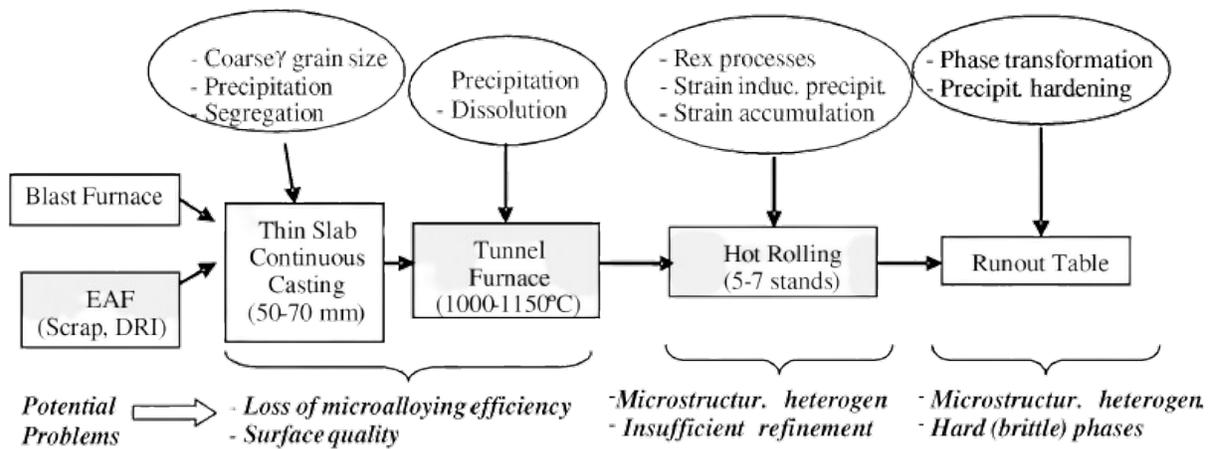


Figure 2.5 – Metallurgical mechanisms and reactions at different stages of TSDR process. (28)

The as-cast macrostructure of the slabs in TSDR has very coarse austenite grain size when compared with the traditional hot rolling process. In direct charging there is no reheating times; hence no austenite to ferrite transformation during slab cooling post-casting and hence, no ferrite to austenite transformations during reheating which can refine the as-cast macrostructure (28).

Each rolling pass can be characterized by its strain, strain rate, temperature and the

interpass time between stands. The inter-relationship between them decide which mechanism dominates. The mechanisms involved have been discussed in the previous section. For a more detailed quantitative modelling in TSDR, an interested reader can follow Uranga et. al (21).

2.5 Microalloying

Niobium

Nb has three fold impact : grain size refinement during TMCP, lowering the phase transformation temperature and precipitation hardening. The primary reason niobium aids in grain refinement is because Nb in form of carbonitride precipitates can impede or delay recrystallization before finish rolling by pinning austenite grain boundaries. By inhibiting recrystallization, austenite develops flattened grains and a high dislocation density, promoting the formation of ferrite. In addition, Nb reduces the γ to α transformation temperature (A_{r3}), leading to increased ferrite nucleation rate and decreased grain growth. A combination of these mechanisms creates a fine-grained transformation structure (29; 30).

Titanium

The role of Ti is of austenite grain refinement as well as precipitation hardening by forming nitrides, carbides and carbonitrides. In TSDR process, fine precipitates of TiN and

TiC—often existing as Ti(C, N) with varying C/N ratios—have been observed to precipitate in austenite along strain-induced precipitates during rolling. These precipitates retain their fine size due to the rapid cooling rates, avoiding coarsening. Ti-precipitates, mainly TiN, refine grains by pinning austenite grain boundaries and delay the recrystallization kinetics of austenite (31). Ti is often used in combination with Nb and V (32; 33).

Vanadium

Vanadium exhibits higher solubility in austenite compared to titanium or niobium, enabling it to stay dissolved until the austenite-to-ferrite transformation occurs. This enhances its role in strengthening through the formation of fine carbides, nitrides, or carbonitrides, which contribute via precipitation hardening. VC is highly effective due to its lower temperature precipitation in ferrite and finer granularity compared to VN. VC and VN have markedly different solubility products, with VN being about two orders of magnitude lower than VC. Hence, Nitrogen plays a crucial role in vanadium-microalloyed steels by controlling driving force for precipitation. The interaction of V with other microalloying elements often results in the formation of mixed carbonitrides, which further reduces the amount of vanadium dissolved in austenite (34; 35).

2.6 Machine Learning

Machine learning (ML) focuses on two key interrelated questions : How can computer systems automatically improve through experience, and what are the fundamental laws governing all learning systems, including computers, humans, and organizations? The area of ML answers these fundamental questions and aids in the development of highly practical applications across different fields (36). The call for such automated methods for data analysis becomes important as we enter the era of big data. Formally, ML can be defined as a set of methods that can automatically detect patterns in data and then use them for future predictions to aid in decision-making under uncertainty. ML methods can be broadly classified into three categories : supervised learning, unsupervised learning, and reinforcement learning (37). We will only touch upon the first two categories briefly, as they have been utilized in this work. To prevent repetition, we refrain from discussing the specific methods used in the thesis since they have already been explained in their respective manuscripts.

2.6.1 Supervised Learning

Supervised learning involves learning a function that maps an input to an output based on example input-output pairs. Let \mathcal{X} be the input space and \mathcal{Y} be the output space. Given a training dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^n$, where $x_i \in \mathcal{X}$ and $y_i \in \mathcal{Y}$, the goal is to learn a function

$f : \mathcal{X} \rightarrow \mathcal{Y}$ that can predict the output y for a given input x . Here, we assume that the output y can be predicted from the input x using a function f such that (38) :

$$y = f(x) + \epsilon \quad (2.20)$$

where ϵ is the error term representing the noise or randomness in the relationship between x and y . The function f is typically parameterized by a set of parameters θ , so we write $f(x; \theta)$. To learn the parameters θ , we define a loss function $L(y, \hat{y})$, which measures the difference between the true output y and the predicted output $\hat{y} = f(x; \theta)$. Common loss functions include mean squared error (MSE) and cross-entropy loss for regression and classification tasks respectively. The objective is to find the parameters θ that minimize the expected loss over the training dataset which is achieved using optimization algorithms. After training, it becomes essential to evaluate its performance over a separate test-set to ensure generalization to unseen-data. Common evaluation metrics include :

- For regression : Mean Absolute Error (MAE), Root Mean Squared Error (RMSE)
- For classification : Accuracy, Precision, Recall, F1-score, ROC-AUC

In this work, two supervised ML algorithms, namely Neural Networks and XGBoost, were utilized to model mechanical properties. Detailed explanations of these algorithms can be found in section 3.3.2 and 4.2.3.1 respectively.

2.6.2 Unsupervised Learning

Unsupervised learning involves finding hidden patterns or intrinsic structures in input data without labeled outputs. Let $\mathcal{X} = \{x_i\}_{i=1}^n$ be the input data. The goal is to model the underlying distribution or structure of the data (39). Unsupervised learning can be broadly classified into two categories : dimensionality reduction and cluster analysis (40).

Dimensionality reduction involves representing data in a more simplified form. For instance, the most basic form of such a method is Principal Component Analysis (PCA). PCA transforms the data into a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (principal component), the second greatest variance on the second coordinate, and so on (41). Given a dataset \mathcal{X} , the method finds the principal components by solving the eigenvalue problem :

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v} \tag{2.21}$$

where \mathbf{C} is the covariance matrix of \mathcal{X} , λ are the eigenvalues, and \mathbf{v} are the eigenvectors. The eigenvectors corresponding to the largest eigenvalues form the principal components. A more advanced version of dimensionality reduction technique is t-SNE, which is used in chapter 4, and hence described therein (section 4.2.7.1).

Clustering algorithms aim to partition the dataset into groups (clusters) such that data points within the same cluster are more similar to each other than to those in other clusters.

One of the most popular clustering methods is K-means clustering. Given K , the number of clusters, the K-means algorithm aims to minimize the within-cluster sum of squares (WCSS)

(42) :

$$\min_{C, \mu} \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2 \quad (2.22)$$

where C_k is the set of points in cluster k and μ_k is the centroid of cluster k . A more involved discussion of the K-means algorithm can be found in section 4.2.7.2.

References

- [1] J. Rodriguez-Ibabe, Thin slab direct rolling of microalloyed steel (2007).
- [2] T. Zhou, P. Zhang, K. Kuuskman, E. Cerilli, S.-H. Cho, D. Burella, H. S. Zurob, Development of medium-to-high carbon hot-rolled steel strip on a thin slab casting direct strip production complex, *Ironmaking & Steelmaking* 45 (7) (2018) 603–610.
- [3] R. Borsi, M. Rotti, Algoma’s direct production plant for quality steel, *Steel Times International* 22 (2) (1998) 22.
- [4] E. Orowan, The calculation of roll pressure in hot and cold flat rolling, *Proceedings of the Institution of Mechanical Engineers* 150 (1) (1943) 140–167.
- [5] F. Siciliano, Mathematical modeling of the hot strip rolling of nb microalloyed steels, Ph.D. thesis, McGill University (1999).

- [6] J. Alexander, H. Ford, Simplified hot-rolling calculations, *J Inst Met* 92 (1964) 397–404.
- [7] R. Sims, The calculation of roll force and torque in hot rolling mills, *Proceedings of the Institution of Mechanical Engineers* 168 (1) (1954) 191–200.
- [8] H. Hitchcock, Roll neck bearings, appendix i, *Elastic Deformation of Rolls during Cold Rolling* (1935) 33–41.
- [9] A. Nadai, A. M. Wahl, E. C. Bingham, *Plasticity, a mechanics of the plastic state of matter* (1931).
- [10] R. W. K. Honeycombe, *The plastic deformation of metals* (1968).
- [11] F. J. Humphreys, M. Hatherly, *Recrystallization and related annealing phenomena*, elsevier, 2012.
- [12] P. Cotterill, P. R. Mould, *Recrystallization and grain growth in metals* (1976).
- [13] R. W. Cahn, P. Haasen, *Physical metallurgy*, Vol. 1, Elsevier, 1996.
- [14] M. Avrami, Kinetics of phase change. i general theory, *The Journal of chemical physics* 7 (12) (1939) 1103–1112.
- [15] C. Sellars, Modelling microstructural development during hot rolling, *Materials Science and technology* 6 (11) (1990) 1072–1081.

- [16] F. Siciliano Jr, K. Minami, M. TM, J. JJ, Mathematical modeling of the mean flow stress, fractional softening and grain size during the hot strip rolling of c-mn steels, *ISIJ international* 36 (12) (1996) 1500–1506.
- [17] W. Roberts, Microstructure evolution and flow stress during hot working, in : *Strength of Metals and Alloys (ICSMA 7)*, Elsevier, 1986, pp. 1859–1891.
- [18] C. Roucoules, P. Hodgson, S. Yue, J. Jonas, Softening and microstructural change following the dynamic recrystallization of austenite, *Metallurgical and Materials transactions A* 25 (1994) 389–400.
- [19] M. TM, J. JJ, H. PD, Spreadsheet modelling of grain size evolution during rod rolling, *ISIJ international* 36 (6) (1996) 720–728.
- [20] H. JC, B. Donnay, V. Leroy, Precipitation kinetics of microalloying additions during hot-rolling of hsla steels, *ISIJ international* 32 (6) (1992) 779–785.
- [21] P. Uranga, F. AI, B. López, R.-I. JM, Modeling of austenite grain size distribution in nb microalloyed steels processed by thin slab casting and direct rolling (tsdr) route, *ISIJ international* 44 (8) (2004) 1416–1425.
- [22] E. Orowan, *Internal stress in metals and alloys*, 451, London, The Institute of Metals (1948).
- [23] T. Gladman, *Physical metallurgy of microalloyed steels* (1997).

- [24] S. Vervynckt, K. Verbeken, B. Lopez, J. Jonas, Modern hsla steels and role of non-recrystallisation temperature, *International Materials Reviews* 57 (4) (2012) 187–207.
- [25] R. Cochrane, Phase transformations in microalloyed high strength low alloy (hsla) steels, in : *Phase transformations in steels*, Elsevier, 2012, pp. 153–212.
- [26] J. W. Cahn, The kinetics of grain boundary nucleated reactions, *Acta metallurgica* 4 (5) (1956) 449–459.
- [27] S. Yue, Thermomechanical processing of ferrous alloys, in : S. Semiatin (Ed.), *Metalworking : Bulk Forming*, Vol. 14A of *ASM Handbook*, ASM International, Materials Park, Ohio, 2005, pp. 286–296.
- [28] J. Rodriguez-Ibabe, Thin slab direct rolling of microalloyed steels, in : *Materials Science Forum*, Vol. 500, *Trans Tech Publ*, 2005, pp. 49–62.
- [29] C. Klinkenberg, Niobium in microalloyed structural and engineering steels, in : *Materials science forum*, Vol. 539, *Trans Tech Publ*, 2007, pp. 4261–4266.
- [30] L.-y. Sun, X. Liu, X. Xu, S.-w. Lei, H.-g. Li, Q.-j. Zhai, Review on niobium application in microalloyed steel, *Journal of Iron and Steel Research International* 29 (10) (2022) 1513–1525.
- [31] M. Arribas, B. López, J. Rodriguez-Ibabe, Additional grain refinement in

- recrystallization controlled rolling of ti-microalloyed steels processed by near-net-shape casting technology, *Materials Science and Engineering : A* 485 (1-2) (2008) 383–394.
- [32] T. Baker, Titanium microalloyed steels, *Ironmaking & Steelmaking* 46 (1) (2019) 1–55.
- [33] G. Xu, X. Gan, G. Ma, F. Luo, H. Zou, The development of ti-alloyed high strength microalloy steel, *Materials & Design* 31 (6) (2010) 2891–2896.
- [34] T. Baker, Processes, microstructure and properties of vanadium microalloyed steels, *Materials Science and Technology* 25 (9) (2009) 1083–1107.
- [35] R. Lagneborg, B. Hutchinson, T. Siwecki, S. Zajac, The role of vanadium in microalloyed steels, Swerea KIMAB Stockholm, Sweden, 2014.
- [36] M. I. Jordan, T. M. Mitchell, Machine learning : Trends, perspectives, and prospects, *Science* 349 (6245) (2015) 255–260.
- [37] K. P. Murphy, Machine learning : a probabilistic perspective, MIT press, 2012.
- [38] P. Cunningham, M. Cord, S. J. Delany, Supervised learning, in : Machine learning techniques for multimedia : case studies on organization and retrieval, Springer, 2008, pp. 21–49.
- [39] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, R. Tibshirani, J. Friedman, Unsupervised learning, *The elements of statistical learning : Data mining, inference, and prediction* (2009) 485–585.

- [40] T. Mueller, A. G. Kusne, R. Ramprasad, Machine learning in materials science : Recent progress and emerging applications, *Reviews in computational chemistry* 29 (2016) 186–273.
- [41] H. Abdi, L. J. Williams, Principal component analysis, *Wiley interdisciplinary reviews : computational statistics* 2 (4) (2010) 433–459.
- [42] M. Ahmed, R. Seraj, S. M. S. Islam, The k-means algorithm : A comprehensive survey and performance evaluation, *Electronics* 9 (8) (2020) 1295.

Chapter 3

Neural Network Prediction of the Effect of Thermomechanical Controlled Processing on Mechanical Properties

*The current chapter addresses the **first objective of the thesis** i.e. to develop machine learning models to predict mechanical properties of microalloyed steel. To add more confidence in the developed ML models, explainability was introduced in combination with analysis based-on physical metallurgy.*

This paper is published in *Machine Learning with Applications*, appeared as :

Sinha, S., Guye, D., Ma, X., Rehman, K., Yue, S., & Armanfard, N. (2024). Neural network prediction of the effect of thermomechanical controlled processing on mechanical properties.

Machine Learning with Applications, 100531.

Abstract

The as-rolled mechanical properties of microalloyed steels result from their chemical composition and thermomechanical processing history. Accurate predictions of the mechanical properties would reduce the need for expensive and time-consuming testing. At the same time, understanding the interplay between process variables and alloy composition will help reduce product variability and facilitate future alloy design. This paper provides an artificial neural network methodology to predict lower yield strength (LYS) and ultimate tensile strength (UTS). The proposed method uses feature engineering to transform raw data into features typically used in physical metallurgy to better utilize the artificial neural network model in understanding the process. SHAP values are used to reveal the effect of thermomechanical controlled processing, which can be rationalized by physical metallurgy theory.

3.1 Introduction

Microalloyed steels constitute about 12% of global steel production (approximately 200 million tonnes) and play a dominant role in industries such as oil and gas extraction, construction, and transportation. They remain an attractive material for these industries due to their low cost, good combination of mechanical properties, and weldability (1). Microalloyed steels possess high strength through the micro-additions of, e.g. niobium and

titanium, which form metal carbonitride precipitates that lead to grain refinement (2). Since these steels are typically used in their as-rolled condition, an improved understanding of alloying and the response to hot rolling have resulted in substantial strength increases over the past 40 years (3). To control the hot rolling process, considerable efforts have been made to develop mathematical models.

For a given steel composition, two mathematical models will be required for the prediction of mechanical properties from hot rolling schedules : (i) a process-structure model that uses kinetic and thermodynamic models to predict the as-hot-rolled microstructure ; (ii) a structure-property model that uses the predicted microstructure to determine the corresponding mechanical properties. Such models are primarily empirical, developed through experimental studies, with authors frequently proposing new equations or equation modifications to account for the addition of new alloying elements. Moreover, the laboratory models need to be tuned to fit industrial operations, indicating gaps in the industrial process's metallurgical understanding or modelling. Thus, typical mathematical models comprise many submodels, each containing errors and assumptions. A reduction of these submodels would likely be an improvement. Moreover, if laboratory experiments can be eliminated using data from industrial processing, this would also be beneficial. An obvious way to do this is to develop a model that directly predicts properties from the process without resorting to a microstructure-property model. However, going directly from processing to property without predicting the microstructure is not possible for a model

based on metallurgical principles. Therefore, a prediction not based on metallurgical principles is the logical alternative, with machine learning (ML) based models being an obvious choice.

Figure 3.1 shows the typical configuration of a direct strip processing complex (DSPC). The as-cast thin slab (70-80 mm) enters the tunnel furnace for a short interval (15-20 mins) directly from the caster primarily to obtain temperature homogeneity before rolling. Then the as-cast slab is passed through the edger for width control and subsequently to the roughing stand where it undergoes large deformation to form a transfer bar (35-40 mm). Hence, the transfer bar travels to the finishing stands where it is sequentially deformed in several stages, in this case, six, to achieve the final thickness and tight tolerances of the final geometry. In the final stage, the rolled sheet undergoes accelerated cooling before being coiled. In conventional hot strip mills (HSM) the cast slab is much thicker (~ 200 mm) and is cooled after casting and then fed to reheating furnaces where the slabs spend a longer time (3-3.5 hrs) before being fed to the roughing mill. The higher slab thickness in conventional HSM necessitates higher overall reduction especially multiple passes in the roughing stage to shape the transfer bar for subsequent finishing stages. Hence, the key distinction in thermomechanical processing between the DSPC route and conventional HSM lies in the reduced opportunities for grain refinement due to the fewer number of passes in the DSPC route (4).

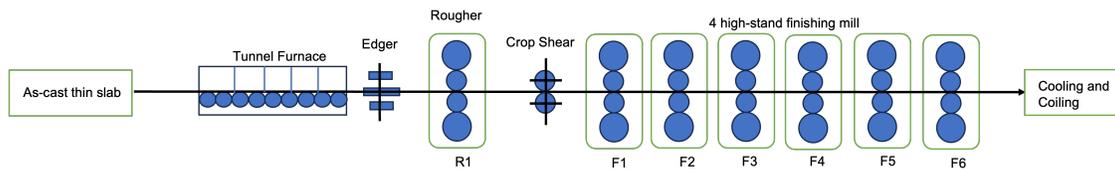


Figure 3.1 – Schematic of Direct Strip Processing Complex

3.2 Neural Network Approaches

Given various non-linear and complex mechanisms contributing to the strengthening via thermomechanical processing (5), ANN has been a successful method to model these processes. Singh et al. (6) used neural networks to predict mechanical properties after hot plate rolling utilizing 108 input variables, including factors such as slab reheat temperature, slab length, slab gauge (thickness), composition, and rolling parameters (“screw settings,” the delay between passes, time of each pass). The dataset consisted of 1892 examples. A Committee (or ensemble) model was applied. Instead of using the single best model obtained from the training procedure, the average of multiple good models was used to make a prediction and yielded a lower error. This is similar to the approach taken by Yang et al. (7), where an ensemble of 10 neural networks was used as it allowed for the determination of error bounds. Korczak et al. (8) applied neural networks to a hot plate rolling process, where 14 nodes were used to input chemical composition, microstructure, and rolling parameters with four output nodes to give ferritic grain size, hardness, yield strength, and tensile strength. However, in this case, only five examples were mentioned in

the article, with no mention of a validation set to test for overtraining of the model. Prediction of properties over the length of the coil during thermo-mechanical processing using deeper neural networks has also been performed (9). Using data from the run-out table, an average prediction of four DNN models was used to predict mechanical properties at different points along the coil's length. Hwang et al. (10) used artificial neural networks to predict roll force using Steckel mills. Here, the model is proposed as a hybrid model as it computes typical parameters from classical mathematical models and uses these as input features. Improvement of the model is expected by using domain-specific knowledge to provide essential insights into the processes. Deeper neural networks have been applied to the entire hot rolled plate production process, from continuous casting unit operation to leveling (11). Using 11,101 data points, models were trained to predict mechanical properties like lower yield strength, ultimate tensile strength, and elongation. Xie et al. (11) performed a detailed investigation of pre-processing and gradient descent optimization algorithm, concluding that z-score pre-processing and the Adam optimization algorithm were most suitable for their application. Here, the finish cooling temperature after equalization and the niobium content were determined as important factors for predicting yield strength, while the finish cooling temperature after equalization and Mn content were the two most influential parameters in predicting UTS. In terms of elongation, niobium, chromium, and vanadium were the three most influential compositional parameters. Mohanty et al. (12) used artificial neural networks and the NSGA-II optimization

algorithm for optimal microalloyed steel design and then compared these with thermodynamic simulations for the formation of strengthening precipitates. Multiobjective optimization was used to optimize Nb, Ti, V, and N levels to provide better mechanical properties.

3.3 Modeling Methods

While there is a wealth of literature discussing the application of neural network (NN) methods to analyze the mechanical properties of hot-rolled strips, as discussed in the previous section, there is a notable dearth of studies specifically addressing the use of these methods for the Direct Strip Processing Center (DSPC) route. Additionally, in the existing body of literature utilizing NN methods, there is a noteworthy absence of efforts to interpret and explain the models, based on our current knowledge. This work aims to produce a suitable model for predicting UTS and LYS at a reasonable level of accuracy and use a combination of model interpretation and metallurgical analysis to provide insight into the combined importance of deformation and chemical composition in thermomechanical controlled processing (TMCP). Building on the success of NN-approaches in modeling similar thermomechanical processes, the methodology employed in this work involved utilizing a neural network for predicting mechanical properties based on alloy composition and TMCP inputs. The neural network model has no predefined relationships between input data and mechanical properties and relies on learning relationships by exposure to

training data. If complex relationships exist between input data and mechanical properties that are not typically modelled in empirical microstructural models, a large enough neural network could capture them to produce a better prediction. The data-driven approach here utilizes A) Preprocessing, B) Prediction, and C) Model interpretation.

3.3.1 Preprocessing

The approach outlined in this study involves preprocessing the features to obtain strain and strain rate, which are commonly used in physical metallurgy. This preprocessing step enables a direct comparison of observations with existing physical metallurgy theory and aligns with previous research emphasizing the performance improvement achieved by applying domain-specific knowledge (10).

To compute the Von Mises effective strain (ε_{eff}) for each rolling stand, Equation 3.1 is utilized (13) :

$$\varepsilon_{\text{eff}} = \frac{2}{\sqrt{3}} \ln \left(\frac{H}{h} \right) \quad (3.1)$$

Here, H represents the entry thickness at each stand (in millimeters) and h denotes the exit thickness at each stand (in millimeters).

The average strain rate ($\dot{\varepsilon}$) can be expressed using Equation 3.2, which calculates the ratio of the effective strain to the time of strain application (t_{def}) (13) :

$$\dot{\epsilon} = \frac{\epsilon_{\text{eff}}}{t_{\text{def}}} \quad (3.2)$$

The value of t_{def} is determined using Equation 3.3 :

$$t_{\text{def}} = \frac{\alpha}{2\pi \times 60/U} \quad (3.3)$$

Here, U represents the roll speed at each stand (in revolutions per minute), and α signifies the portion of the roll in contact with the strip, i.e., the roll/strip interface where reduction occurs in the roll bite. The unit of t_{def} is seconds. The calculation of α is performed using Equation 3.4 :

$$\alpha = \cos^{-1} \left(1 - \frac{H - h}{2R} \right) \quad (3.4)$$

In this equation, R denotes the roll radius of each stand (in millimeters).

The physical deformation parameters obtained are then combined with the chemical composition and temperature of the strip at each stand to form the set of features used for training the artificial neural network. Therefore, the input data for the artificial neural network model comprise the strain and average strain rate at each pass of deformation, the temperature at each pass of deformation, and the chemical composition of each strip. This study uses 50 and 60 Ksi Nb-based HSLA data from the DSPC rolling mill of Algoma Steel Inc (Ontario, Canada). A total of 5930 instances were available and 20% of data were

randomly split as test set for evaluating the performance of the fully trained model. In total, there are thirty-eight features. To ensure consistent scaling, the features are normalized using mean and standard deviation. The mechanical properties of interest in this study are the ultimate tensile strength (UTS) and lower yield strength (LYS). It should be noted that any examples with missing or null values have been excluded from the analysis.

3.3.2 Prediction

Acknowledging the various non-linearities that exist in the literature regarding the mechanisms that act to reduce grain size and the prior application of neural networks to the prediction of mechanical properties in steel, artificial neural networks (ANN) present an exciting tool for TMCP strategy design. Graupe (14) likens the ability of neural networks to make predictions to the decision-making process in animals, where complex mathematical problems are not solved; rather, a collection of neurons with a simple structure is able to adapt to fit the problem at hand. An artificial neural network is a collection of algorithms loosely modeled on the animal brain. The nodes are arranged in layers, with each type of layer having a specific function. Feedforward neural networks are the simplest type of neural network and consist of an input layer, one or many hidden layers, and an output layer. The input layer takes in attributes descriptive of the process. Information from the input layer is passed to the nodes in the hidden layer, where a mathematical function is applied to them. A weight is assigned as the variable is passed

between layers to control the strength of the influence of each input (14). The output of hidden nodes can be represented as Equation 3.5 :

$$h_1^i = f(W_1^T x + b_0) \quad (3.5)$$

Here, h_1^i is the output of the i -th node of the first hidden layer. f represents the activation function. The activation function determines whether a node is active or not. W_1^T is the weight vector associated with the first hidden layer, and b_0 is the bias of the first hidden layer. In the case of multiple hidden layers, the output of the first hidden layer is passed to the next hidden layer. The output from the second layer of hidden nodes can be calculated using Equation 3.6 :

$$h_2^i = f(W_2^T \times f(W_1^T x + b_0) + b_1) \quad (3.6)$$

In this equation, h_2^i represents the output of the i -th node of the second hidden layer. W_2^T is the weight vector associated with the second hidden layer, and b_1 is the bias of the second hidden layer. The output from the output layer is computed in a similar manner to the output of the hidden layers.

Individual neural networks were used to predict the ultimate tensile strength (UTS) and lower yield strength (LYS) using single output neurons. In a supervised context, weights are assigned between nodes to determine the contribution of each input to the final output

value. These weights are initially randomly assigned and tuned during the recursive training process by exposing the network to training data. The modeling was performed using the open-source Python library PyTorch (15).

1) Implementation Details : K-fold cross-validation was used to optimize the hyperparameters of the model. The data were shuffled and divided into k folds, where k is an integer greater than 1. Each fold contains the same number of examples. $k - 1$ folds were used for model training, and 1 fold was held out as validation data. The fold held out for validation changes sequentially until all folds have been used for validation (16). In this work, 5-fold cross-validation was performed. 500 epochs and an exponential scheduler with step size 0.995 were used for training each network. The loss function used was mean squared error (MSE) along with ridge regularization represented by Equation 3.7 :

$$L(\theta) = \frac{1}{2N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^p \theta_j^2 \quad (3.7)$$

Here, $L(\theta)$ is the loss function, N is the number of data points, y_i is the observed value, \hat{y}_i is the predicted value, θ_j are model parameters, p is the parameter count, and λ is the regularization parameter. The Adam optimization algorithm was applied as the optimization algorithm. For the prediction of UTS and LYS, the activation function was ReLU. The number of nodes were varied between 100 to 4000, and regularization parameter(λ) was varied between 0 to 0.1 to select the optimized hyperparameters.

2) Evaluation Metrics : In addition to mean squared error (MSE), the predictive ability of

the model (PAM) is also selected as the quantification metric for the prediction performance and is defined by Equation 3.8 :

$$\text{PAM} = \frac{\text{NCP}}{\text{NP}} \times 100\% \quad (3.8)$$

Here, NP is the number of predictions, and NCP is the number of ‘correct’ predictions. A ‘correct’ prediction is one in which the difference between the predicted value and the measured value is ‘small’. Two levels of ‘small’ are chosen : (i) less than 5% of the true value or (ii) less than 2.5% of the true value. The number of ‘correct’ predictions (NCP) is calculated using Equation 3.9 :

$$\text{NCP} = \left| \frac{y_i - \hat{y}_i}{y_i} \right| < 5\% \text{ or } 2.5\% \quad (3.9)$$

Here, \hat{y}_i represents the predicted mechanical property, and y_i is the true (i.e., final measured) mechanical property value. The individual metrics produced using these equations are referred to as PAM 5% and PAM 2.5%.

3.3.3 Model Interpretation

For a given alloy composition, TMCP controls the strip’s microstructure and subsequent mechanical properties. In addition to the final mechanical properties, the strips produced must achieve the desired final product thickness (gauge). Recall that the rolling

process illustrated in Fig 3.1 consists of sequential rolling stands to provide the desired shape control. The consequence of this sequential rolling process is the coupled nature of TMCP inputs in the model. For example, increasing strain at the earlier processing stands must be compensated for by reducing the strain at later processing stands. Such changes in strain will lead to changes in temperatures and strain rates at all subsequent stands. Assessing the effect of processing on mechanical properties is essential if the processing is to be used to compensate for variations in the alloy composition. As the dataset does not contain an abundance of examples of different TMCP parameters with the same alloy composition, the effect of TMCP must be separated from the influence of alloy composition.

In this work, the SHAP (Shapley Additive exPlanations) method (17) is used as an explanation model to isolate the effect of inputs on the model output. An explanation model is a simpler model approximating the original model and is interpretable by humans (17). The SHAP method belongs to a class of additive feature attribution methods that rely on simplifying the inputs such that a mapping function is used to map them to the original inputs (17). The explanation model is then proposed based on these simplified inputs to approximate the original model. Using such model interpretation methods allows for understanding the predictions made by the neural network model. Kernel SHAP explainer was used which is a model-agnostic method that uses a special weighted linear regression to calculate the contribution of each feature. The outputs of the SHAP method are SHAP values, which are

approximations of Shapley values and are the attributed change in the fully trained model's prediction for each feature of each instance. The explanation model approximates the neural network prediction for each instance, starting with a base value. The base value, in this case, determined using the training data, is the predicted value if no features are known for the current output.

3.4 Results

The results of varying hyperparameters are presented here, and models for predicting UTS and LYS are selected. Initially, a grid search was performed for a single hidden layer model with various nodes and activation functions. ReLU performed better than tanh, logistic, and sigmoid. Two hidden layer models were used with the ReLU activation function to improve the model performance further.

Figures 3.2 and 3.3 show the initial results of varying model complexity (log-scale) and regularization strength for two hidden layer neural network models. Model complexity is the number of parameters in the network. 5-fold cross-validation reports model performance, and error bars represent the standard deviation. In the context of the conducted search, it was found that increasing model complexity did not lead to overfitting, as evidenced by the convergence of both the validation error and the PAM metrics to a plateau. Including regularization did not yield a significant difference in performance, indicating that the models could generalize well to the data without overfitting. Hence, all

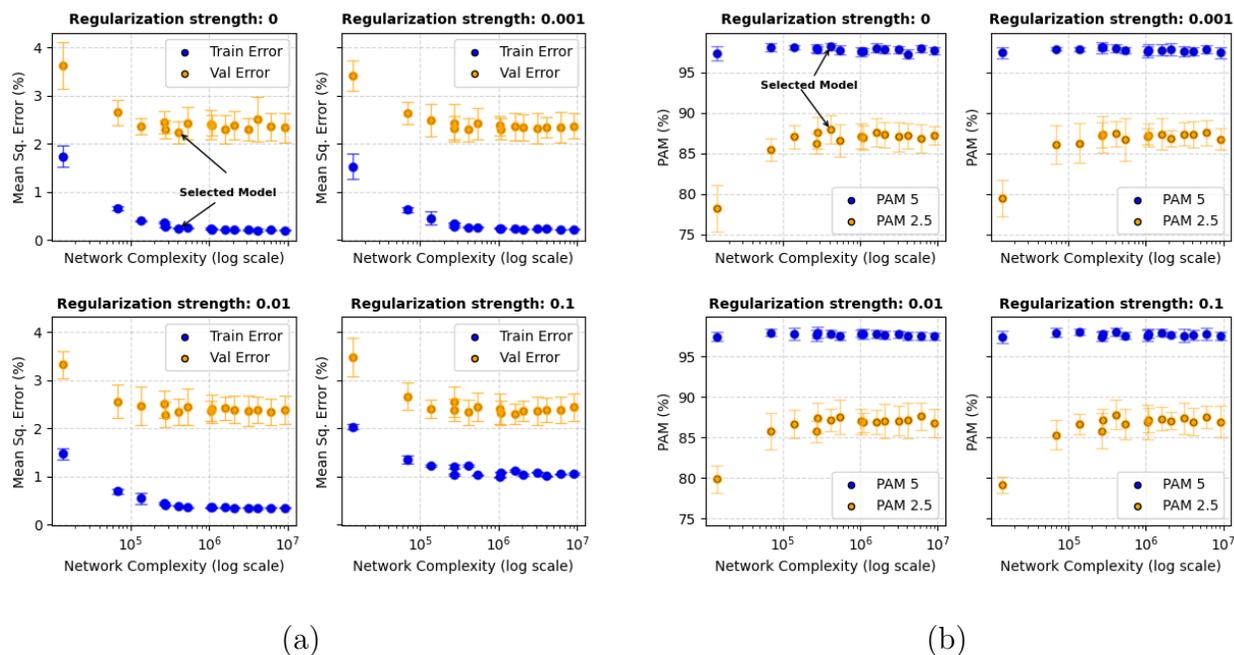


Figure 3.2 – Hyperparameter grid search for the neural network model with varying model complexity and regularization strength for predicting UTS. (a) Training Error and Validation Error, (b) PAM 5% and 2.5%. The model selected is the neural network trained with (3000, 100) nodes and no regularization.

trained models were considered to evaluate the best model for UTS and LYS. The evaluation process involved sorting the models according to their performance in the three metrics : minimizing the validation error, maximizing the PAM 5% metric, and maximizing the PAM 2.5% metric. The neural networks with selected hyperparameters were trained on the entire training data, and the performance was observed on the test set. The model chosen for UTS scored validation MSE of 5.2%, PAM 5% of 99%, and PAM 2.5% of 89%. The selected model for LYS achieved validation MSE of 7.6%, PAM 5% of 94%, and PAM 2.5% of 81%. The performance of the selected neural network models for UTS and LYS for

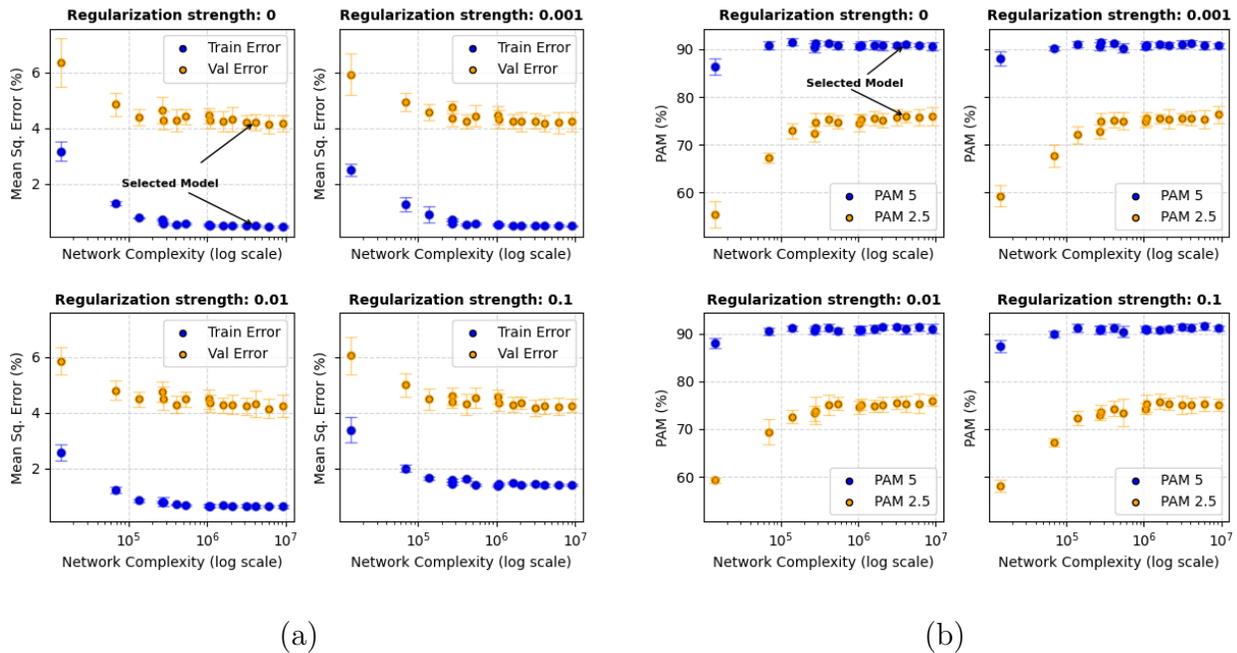


Figure 3.3 – Hyperparameter grid search for the neural network model with varying model complexity and regularization strength for predicting LYS. (a) Training Error and Validation Error, (b) PAM 5% and 2.5%. The model selected is the neural network trained with (3000, 2000) nodes and no regularization.

the test data is shown in Figure 3.4. On the x-axis is the measured mechanical property from plant data, while the y-axis is the neural network predicted mechanical property.

Figure 3.5 shows the residual (%) for the prediction by the neural network model of (a) UTS and (b) LYS. None of the predictions have a residual that exceeds $\pm 10\%$ for both UTS and LYS. Observing both figure 3.4 and figure 3.5, it is apparent that the neural networks can predict the strengths of hot-rolled sheets with reasonable accuracy compared to other similar works in literature (11; 18; 19). Given the absence of a standardized dataset and the utilization of diverse thermomechanical processes in literature, direct comparisons

of accuracy are challenging. However, our evaluation criteria are notably stringent and we confidently assert that this work stands on par with recent advancements in this field.

The SHAP method calculates the effect of features observed in the dataset. The output of the SHAP method are SHAP values. The SHAP values give the impact of each feature on model output for each instance. The SHAP values of individual features are not studied in this work but rather the combination of SHAP values of groups of features. The combined effects of the groups of interest are calculated by Eq 3.10 and Eq 3.11. The combined impact of TMCP and chemistry is interpreted here as the influence of a set of parameters on strength. Combining the SHAP values(s) in this way allows for observing mechanical property variation predicted by the model against the plant data.

$$s(\text{TMCP}) = \sum_{i=R1}^{F6} s(\varepsilon_{\text{eff}})_i + \sum_{i=R1}^{F6} s(\dot{\varepsilon})_i + \sum_{i=R1}^{F6} s(T)_i \quad (3.10)$$

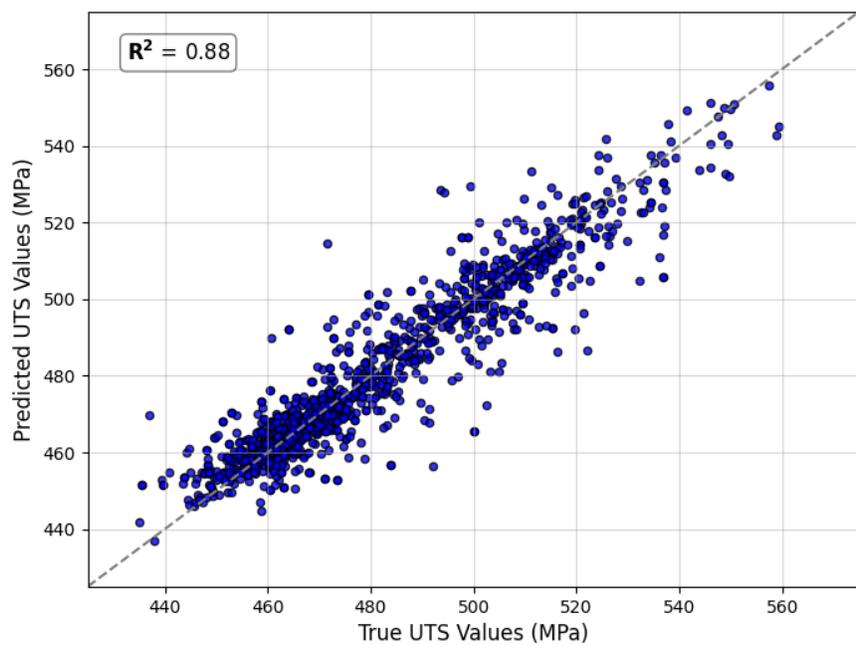
In this equation 3.10, $s(\text{TMCP})$ represents the combined effect of thermomechanical parameters, $s(\varepsilon_{\text{eff}})$ represents the SHAP values of strain, $s(\dot{\varepsilon})$ represents the SHAP values of strain rate, and $s(T)$ represents the SHAP values of temperature.

$$s(\text{chemistry}) = \sum_{i=1}^{16} s(\text{element})_i \quad (3.11)$$

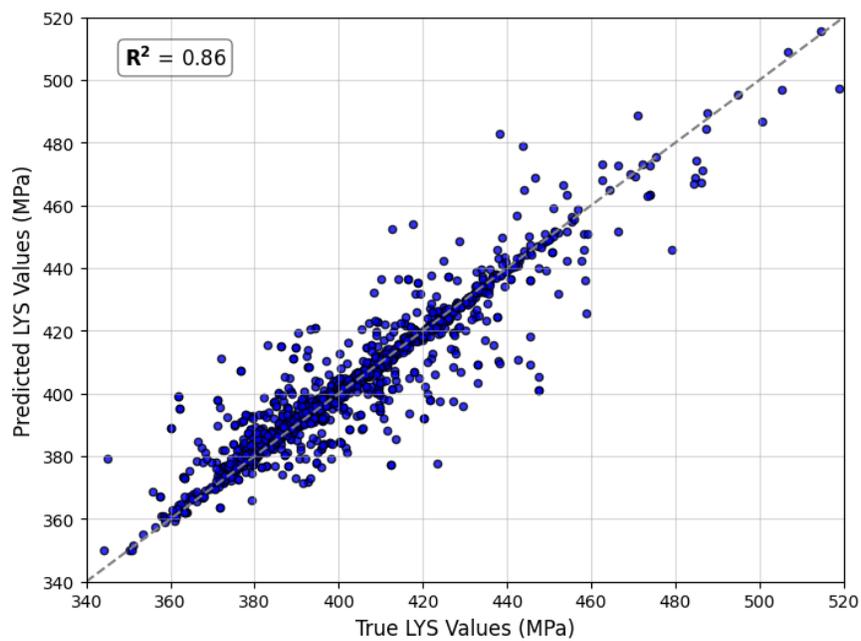
In the equation 3.11, $s(\text{chemistry})$ represents the combined effect of all the chemical composition, and $s(\text{elements})$ is the shap values of an element.

Figure 3.6 (a) shows the combined effect of TMCP on UTS on the y-axis. On the x-axis, the gauge is shown. The gauge here has been normalized between 0 and 1 such that the smallest gauge has a value of 0 and the largest gauge has a value of 1. Measured UTS is indicated as the color scale. The combined effect of TMCP on UTS has a range of 200 MPa.

In Figure 3.6 (a), it can be observed that the change in UTS decreases with an increasing gauge that can be attributed to TMCP. From a metallurgical point of view, this may be rationalized by increased pancaking related to the amount of strain. Figure 3.6 (b) shows the combined effect of chemistry on UTS on the y-axis. The combined effect of chemistry on UTS ranges from about 100 MPa. On comparing both Figure 3.6 (a) and 3.6 (b), one can observe that for thin gauges where the effect of TMCP is high, the effect of chemistry is relatively less, while for thick gauges where there is not enough reduction, the effect of chemistry is higher. This iterates that HSLA steel is relatively richly alloyed in industrial practice to obtain a similar strength for thicker gauges, with insufficient opportunity for microstructure strengthening through TMCP. It is also worth noting from figure 3.6 (b) that the highest strength is obtained for steel where the alloying is high and has a high reduction. Figure 3.7 shows a similar effect for LYS as explained for UTS. The combined effect of chemistry on LYS appears to be greater than that on UTS, especially for thicker gauges.

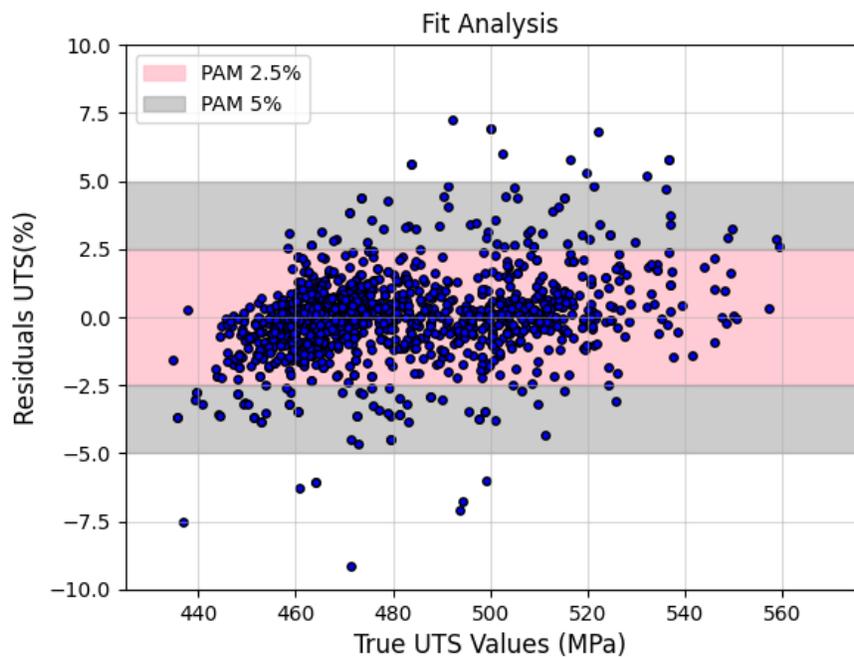


(a)

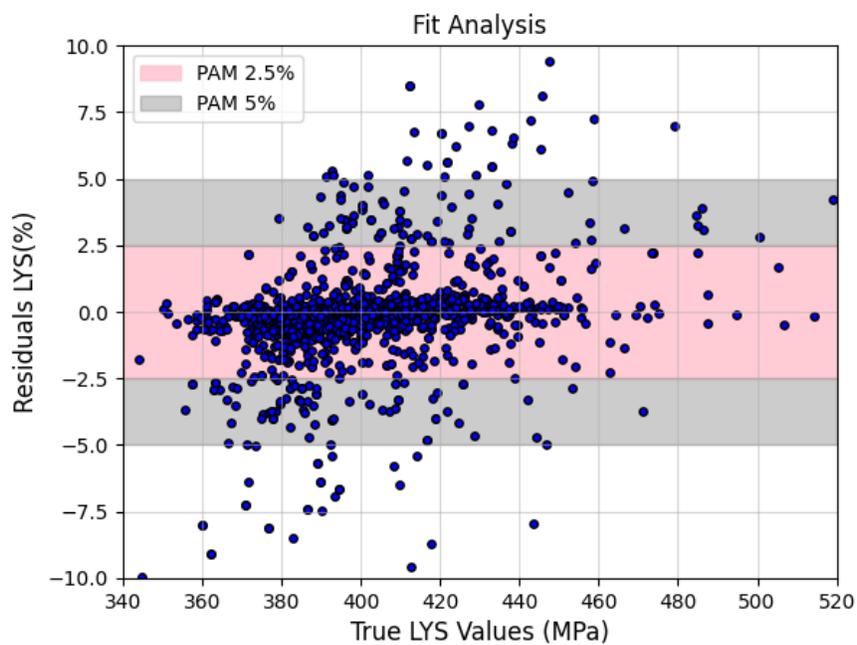


(b)

Figure 3.4 – Prediction by neural network models with selected hyperparameters of (a) UTS and (b) LYS

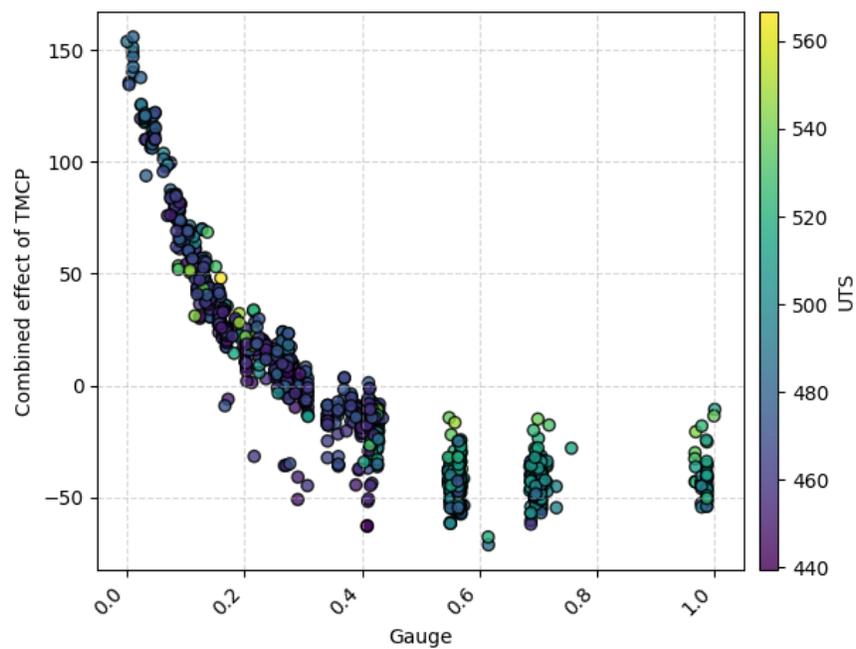


(a)

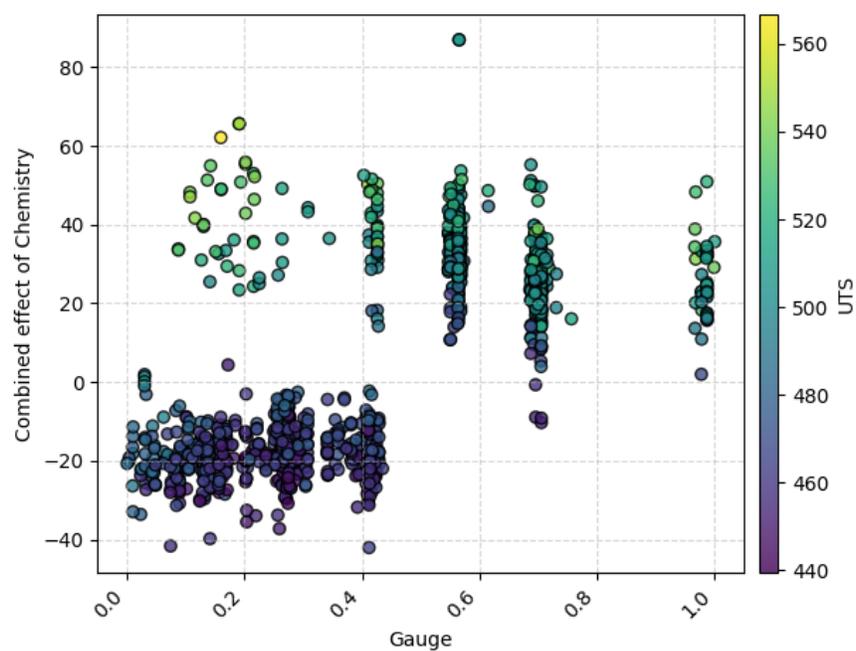


(b)

Figure 3.5 – Residual plot for the prediction of selected Neural Networks. (a) Residual plot for UTS, (b) Residual plot for LYS.

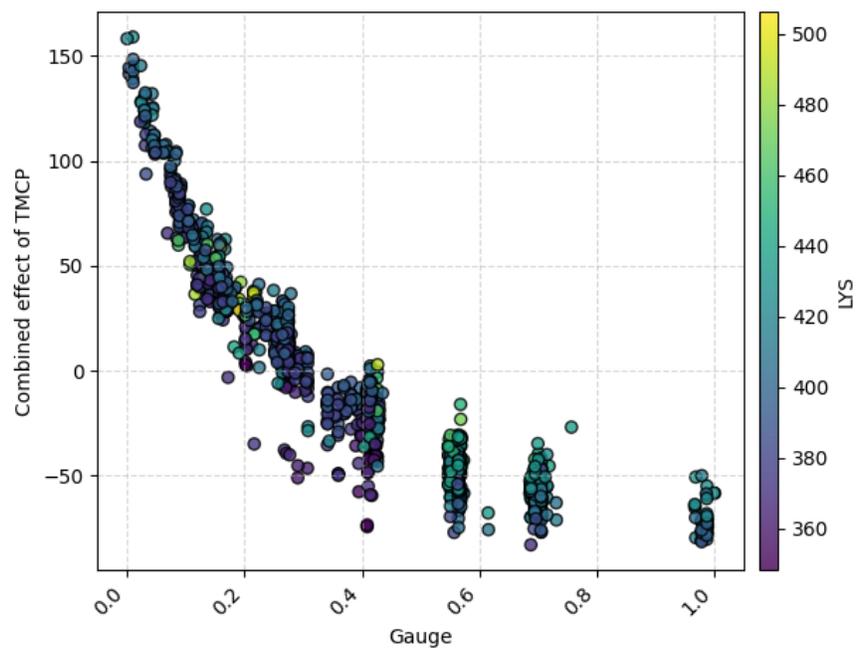


(a)

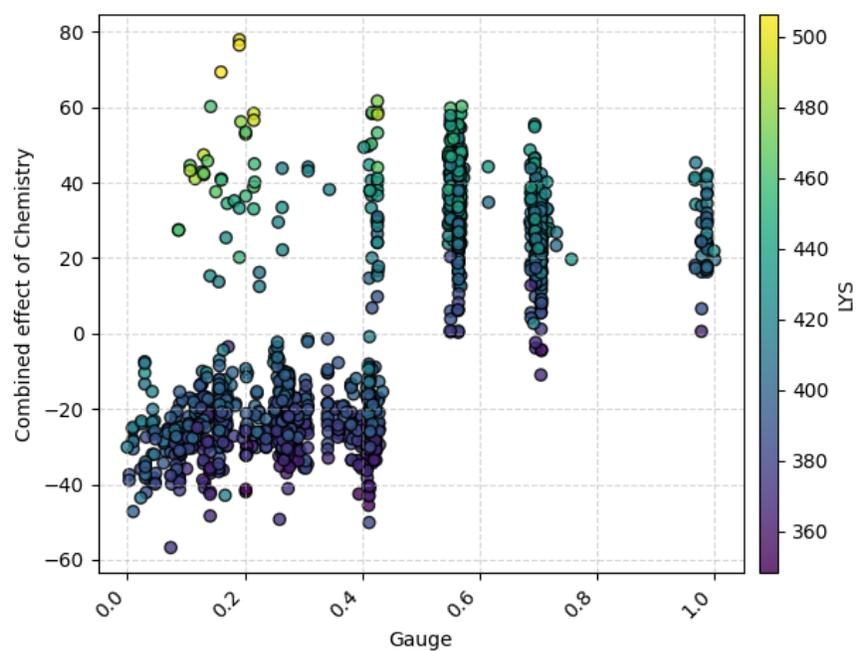


(b)

Figure 3.6 – Effect of TMCP parameters on UTS. (a) Shows the combined effect of TMCP on UTS against the relative gauge, and (b) shows the combined effect of chemistry on UTS against the relative gauge



(a)



(b)

Figure 3.7 – Effect of TMCP parameters on LYS. (a) Shows the combined effect of TMCP on LYS against the relative gauge, and (b) shows the combined effect of chemistry on LYS against the relative gauge

3.5 Discussion

The combined effect of TMCP and chemical composition on the model is explained using SHAP, which shows a good opportunity for process control and alloy design. But for a steelmaker, it is necessary to understand metallurgical influence on the mechanical property. We discuss the effect of TMCP and chemistry using mean flow stress analysis to add confidence in neural network prediction and model explainability.

The mean flow stress at each stand is essentially the 'average' value flow stress of the steel as it flows through the roll gap. The mean flow stress (MFS) during hot rolling can be obtained from roll force by Equation 3.12, given by Sims (20) :

$$\text{MFS}_{\text{Sims}} = \frac{P}{\frac{2}{\sqrt{3}}w\sqrt{R'(h_0 - h_f)Q}} \quad (3.12)$$

where P is roll force (N), w is the width of the strip (m), h_0 and h_f are the entry and exit strip thicknesses respectively (m), R' is the radius of curvature due to work roll flattening (m), and Q is a geometric factor for which details of computation can be found by consulting Sims (20).

The subsequent data points in the figure 3.8, corresponding to varying thicknesses, denote the SIMS mean flow stress at successive rolling stands. For simplicity in qualitative analysis, the data is divided into two parts based on Nb composition, i.e., Low Nb ($\leq 0.03\%$) and

High Nb (0.03-0.06%). To avoid repeatability only the Low Nb case is presented here for three thickness to present different processing conditions. In direct strip production of Nb-based HSLA steels, the strength is primarily influenced by three main mechanisms : solid solution strengthening, grain refinement, and work hardening (21). For a fully recrystallized austenite, the MFS determined from the rolling loads will depend on solute strengthening and grain size. If recrystallization is stopped, then the MFS will have contributions from other strengthening mechanisms. The observed slope deviation in the MFS plots can serve as indicative measures for distinct microstructural transformations occurring within the feed during rolling. A pronounced incline in the MFS slope can suggest a strengthening effect, such as work hardening or precipitation. While, a negative change in slope can be inferred as softening mechanism, such as recrystallization (22).

A clear transition is discernible between finishing stands 3 and 4 for gauges 3.4 mm and 4.5 mm. In these instances, a clear sign of softening is apparent, while such an effect is not observed for the 6.2 mm gauge. This difference may signify a missed opportunity for grain refinement in higher thicknesses, attributed to the limited deformation available.

Figure 3.9(a) illustrates the flow stress behaviour of two different alloy compositions with a mean gauge of 3.0-3.5 mm. The low Nb composition exhibited average Lower Yield Strength (LYS) and Ultimate Tensile Strength (UTS) values of 388 MPa and 465 MPa, respectively, while the high Nb composition displayed average LYS and UTS values of 450 MPa and 522 MPa, respectively. The flow stress behaviour observations indicate that the

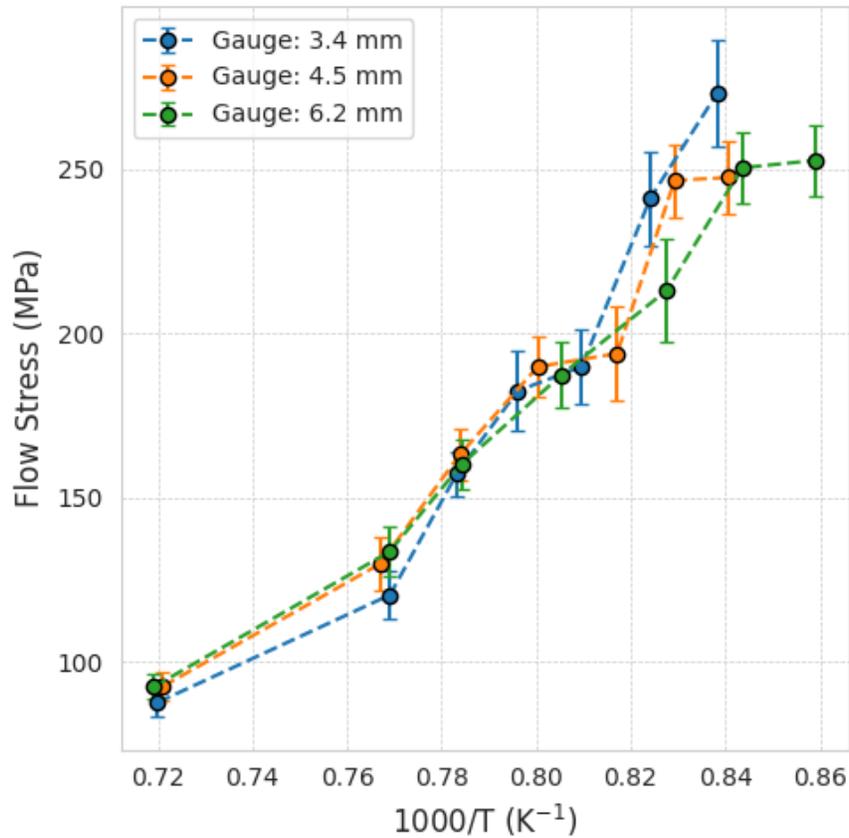
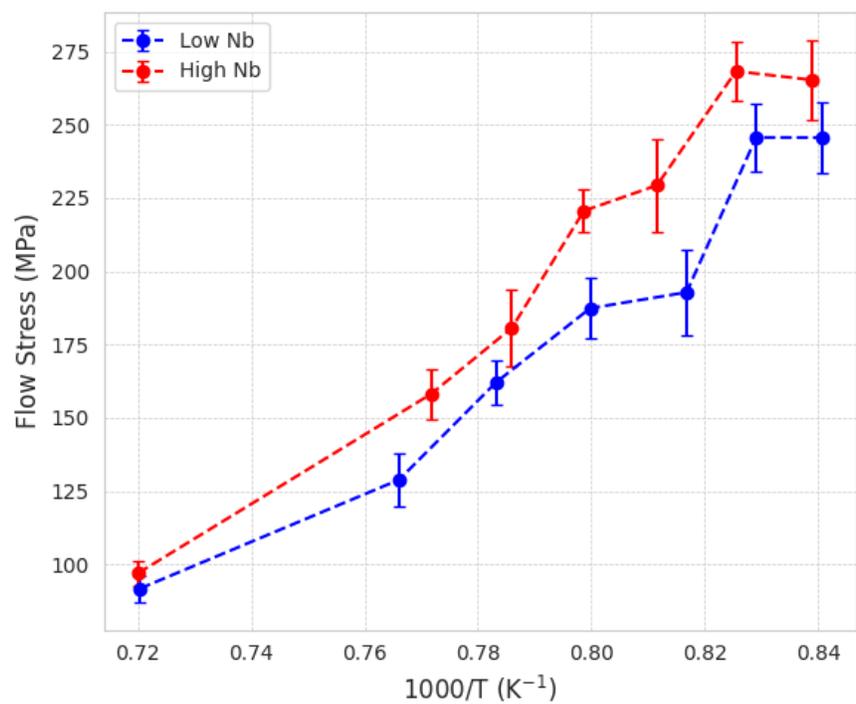


Figure 3.8 – SIMS MFS Plot for low Nb composition in different thickness

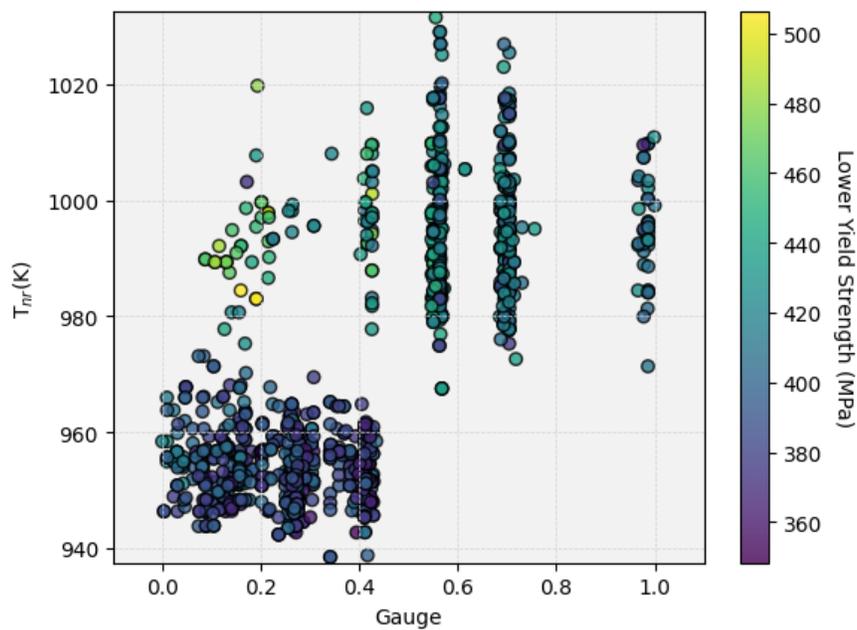
higher MFS of the high Nb composition can be attributed to solid solution strengthening. Figure 3.9(b) presents the variation of the estimated no recrystallization temperature (T_{nr}) based on changes in chemical composition with mean gauge. T_{nr} , obtained from the empirical relationship proposed by F. Boratto and Jonas (23), provides a rough estimate based on an empirical relation of chemical composition but it is also influenced by factors such as grain size, strain, and strain rate. Interestingly, Figure 3.9(b) exhibits similarities to the SHAP plots depicting the combined effect of chemical composition. A higher T_{nr} facilitates increased

dislocation generation and work hardening of austenite grains, serving as nucleation sites for ferrite during transformation and leading to a finer grain structure. For a more comprehensive understanding of the importance of T_{nr} in modern HSLA steels, further insights can be found in the work of Vervynckt et al. (24).

Strength variations can arise because of variations in alloy composition and rolling schedule. Regarding alloy composition control, the variations seen in the dataset represent the best control that can be achieved in the melt shop (i.e., in steelmaking). However, the rolling schedule is not designed to account for any variations in alloy composition; the rolling schedule is chosen based on the target alloy composition and the gauge. Therefore, there is an opportunity to reduce the effect of alloy variation by customizing the rolling schedule to compensate for alloy variations.

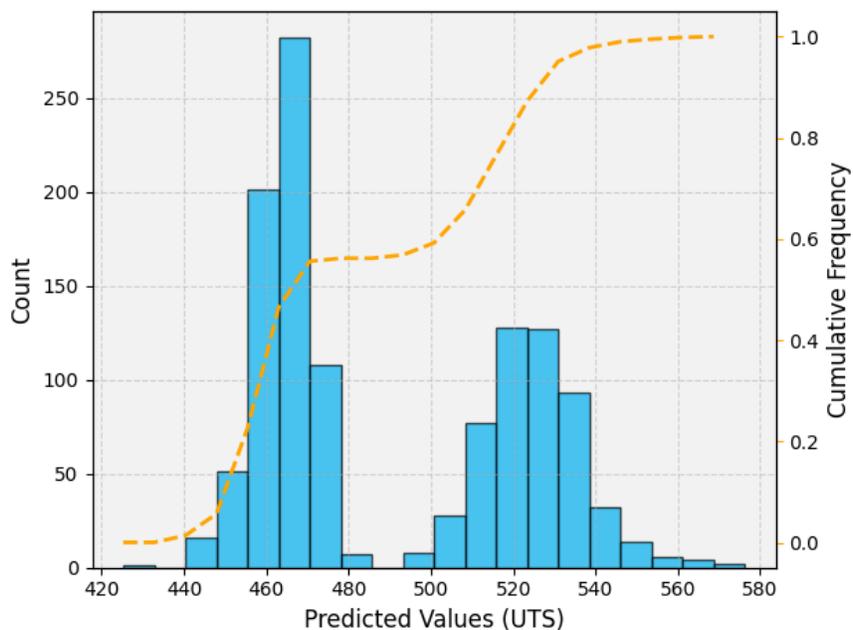


(a)

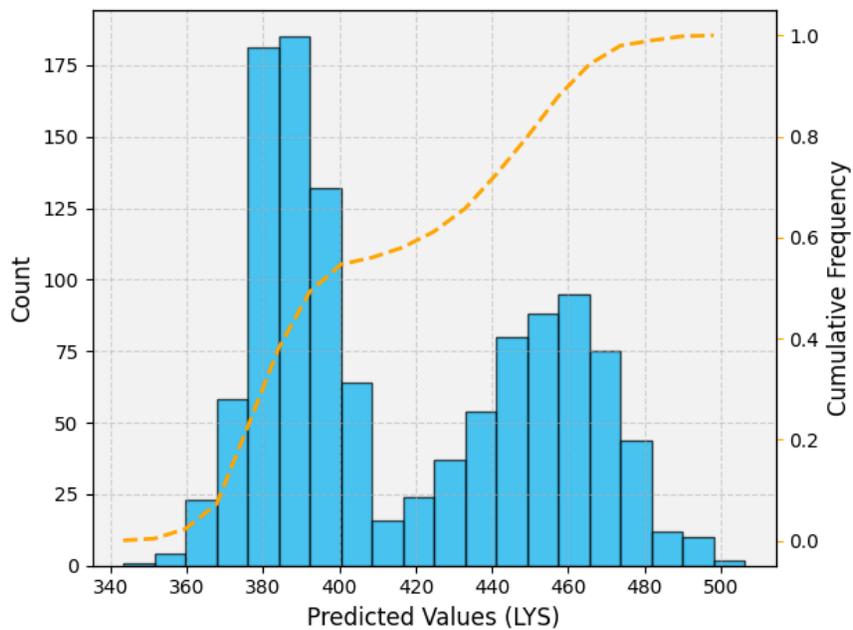


(b)

Figure 3.9 – (a) Flow stress behaviour of two different Nb compositions with mean gauge 4.5 mm. (b) Variation of estimated no recrystallization temperature (T_{nr}) with mean gauge based on changes in chemical composition.



(a)



(b)

Figure 3.10 – Predicted effect of variation in strength due to chemical compositions observed in the dataset using a single deformation schedule on (a) UTS and (b) LYS. Note that the axes in (a) and (b) are different.

Figure 3.10 assesses the variation in strength produced by actual alloy composition variation observed in the dataset over a single deformation schedule using the process-property model. The process-property model predicts UTS and LYS using the alloy compositions from all instances paired with a single rolling schedule for the most frequently produced gauge (4.5 mm). The bimodality in the distribution comes from different chemical compositions of 50 and 60 Ksi grades. The chemical compositions observed in the dataset produced an approximately 100 MPa variation in the prediction of both UTS and LYS. TMCP was shown to be responsible for an effect of 200 MPa for UTS and LYS, so alloy compositional variations could be offset by designing the appropriate deformation schedules. However, recall that this variation due to processing can only be achieved by changing the gauge and the effect of changing the deformation pattern; the variation for a given gauge is much smaller, depending on the gauge.

3.6 Conclusion

Thirty-eight features representing strain at each rolling stand, strain rate at each rolling stand, the temperature at each rolling stand, and the chemical composition of each strip are used to predict ultimate tensile strength (UTS) and lower yield strength (LYS). The selected model for UTS predicted 99% of data points within $\pm 5\%$ error and 89% of data points within $\pm 2.5\%$ while the model chosen for LYS predicted 94% of data points within $\pm 5\%$ error and 81% of data points within $\pm 2.5\%$. SHAP analysis shows that for UTS and

LYS, the combined effect of TMCP is around 200 MPa and the combined effect of chemical composition is around 100 MPa. Mean flow stress analysis shows that for thinner gauges there is enough opportunity for grain refinement and hence results in a finer ferrite grain size. But for thick gauges, SHAP plots indicate that there is a good opportunity for process control to obtain consistent properties. The effect of chemical composition is to increase the solid state strengthening and increase the no-recrystallization temperature to promote a finer ferrite grain size.

Beyond property predictions and process understanding, these models hold promise for applications in alloy design as well as process optimization. However, the concern about certain predictions deviating significantly from true values raises concern and underscores the need for a focused effort to understand these deviations.

Acknowledgements

This work was supported by NSERC grant (CRDPJ 543584-19). The authors express their gratitude to Algoma Steel Inc., Canada, for their financial support, data assistance, and permission to publish this work. Additionally, the authors acknowledge the MITACS Graduate Fellowship awarded to Sushant Sinha during his research at McGill University.

References

- [1] A. DeArdo, M. Hua, K. G. Cho, C. Garcia, On strength of microalloyed steels : an interpretive review, *Materials Science and Technology* 25 (9) (2009) 1074.
- [2] W. Morrison, Microalloy steels-the beginning, *Materials Science and Technology* 25 (9) (2009) 1066.
- [3] T. Baker, Microalloyed steels, *Ironmaking & Steelmaking* 43 (4) (2016) 264–307.
- [4] R. Kaspar, Microstructural aspects and optimization of thin slab direct rolling of steels, *steel research international* 74 (5) (2003) 318–326.
- [5] I. Tamura, H. Sekine, T. Tanaka, Thermomechanical processing of high-strength low-alloy steels, Butterworth-Heinemann, 2013.
- [6] S. Singh, H. Bhadeshia, D. MacKay, H. Carey, I. Martin, Neural network analysis of steel plate processing, *Ironmaking and Steelmaking* 25 (5) (1998) 355–365.
- [7] Y. Yang, D. Linkens, A. Trowsdale, J. Tenner, Ensemble neural network model for steel properties prediction, *IFAC Proceedings Volumes* 33 (22) (2000) 401–406.
- [8] P. Korczak, H. Dyja, E. Łabuda, Using neural network models for predicting mechanical properties after hot plate rolling processes, *Journal of Materials Processing Technology* 80 (1998) 481–486.

- [9] I. Mohanty, R. Banerjee, A. Santara, S. Kundu, P. Mitra, Prediction of properties over the length of the coil during thermo-mechanical processing using dnn, *Ironmaking & Steelmaking* 48 (8) (2021) 953–961.
- [10] R. Hwang, H. Jo, K. S. Kim, H. J. Hwang, Hybrid model of mathematical and neural network formulations for rolling force and temperature prediction in hot rolling processes, *IEEE Access* 8 (2020) 153123–153133.
- [11] Q. Xie, M. Suvarna, J. Li, X. Zhu, J. Cai, X. Wang, Online prediction of mechanical properties of hot rolled steel plate using machine learning, *Materials & Design* 197 (2021) 109201.
- [12] I. Mohanty, A. R. Chintla, S. Kundu, Design optimization of microalloyed steels using thermodynamics principles and neural-network-based modeling, *Metallurgical and Materials Transactions A* 49 (2018) 2405–2418.
- [13] T. Maccagno, J. JJ, S. Yue, M. BJ, R. Slobodian, D. Deeks, Determination of recrystallization stop temperature from rolling mill logs and comparison with laboratory simulation results, *ISIJ International* 34 (11) (1994) 917–922.
- [14] D. Graupe, *Principles of artificial neural networks*, Vol. 7, World Scientific, 2013.
- [15] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin,

- N. Gimelshein, L. Antiga, et al., Pytorch : An imperative style, high-performance deep learning library, *Advances in neural information processing systems* 32 (2019).
- [16] G. Vanwinckelen, H. Blockeel, B. De Baets, B. Manderick, M. Rademaker, W. Waegeman, On estimating model accuracy with repeated cross-validation, in : *BeneLearn 2012 : Proceedings of the 21st Belgian-Dutch conference on machine learning*, 2012, pp. 39–44.
- [17] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, *Advances in neural information processing systems* 30 (2017).
- [18] S. Wu, G. Cao, X. Zhou, N. Shi, Z. Liu, High dimensional data-driven optimal design for hot strip rolling of c-mn steels, *Isij International* 57 (7) (2017) 1213–1220.
- [19] W.-g. Li, L. Xie, Y.-t. Zhao, Z.-x. Li, W.-b. Wang, Prediction model for mechanical properties of hot-rolled strips by deep learning, *Journal of Iron and Steel Research International* 27 (2020) 1045–1053.
- [20] R. Sims, The calculation of roll force and torque in hot rolling mills. proceedings of the institution of mechanical engineers, *Proceedings of the Institution of Mechanical Engineers*, 168 (1954) 191.
- [21] R. Wang, G. CI, M. Hua, K. Cho, H. Zhang, D. AJ, Microstructure and precipitation

behavior of nb, ti complex microalloyed steel produced by compact strip processing, ISIJ international 46 (9) (2006) 1345–1353.

- [22] F. Siciliano, J. J. Jonas, Mathematical modeling of the hot strip rolling of microalloyed nb, multiply-alloyed cr-mo, and plain c-mn steels, Metallurgical and materials transactions A 31 (2000) 511–530.
- [23] S. Y. F. Boratto, R. Barbosa, J. J. Jonas, Physical metallurgy of thermomechanical processing of steels and other metals' (thermec-88), THERMEC-88 Conf. Proc. (1988) 383–390.
- [24] S. Vervynckt, K. Verbeken, B. Lopez, J. Jonas, Modern hsla steels and role of non-recrystallisation temperature, International Materials Reviews 57 (4) (2012) 187–207.

Chapter 4

An Interpretable and Reliable Framework for Alloy Discovery in Thermomechanical Processing

*In this chapter, we conduct experiments using various supervised ML algorithms to create models for mechanical property prediction. This research further expands to address the **second objective of the thesis** i.e., to develop a ML-framework for alloy design. To increase trust in the use of ML-frameworks in industries, uncertainty quantification is incorporated along with model explainability.*

This manuscript is under *preparation* :

Sinha, S., Ma, X., Rehman, K., Armanfard, N., & Yue, S., (2024). An Interpretable and Reliable Framework for Alloy Discovery in Thermomechanical Processing.

Abstract

In thermomechanical controlled processing, both the alloy composition and the processing strategy shape the mechanical properties of metals. In this study, we present a data-driven approach to discover alloys with optimized strength-ductility trade-off in the thin slab direct rolling process. We evaluate seven different supervised machine learning algorithms to predict two mechanical properties, namely Ultimate Tensile Strength and % Elongation. SHapely Additive exPlanations (SHAP) augments interpretability to the best performing models. NSGA-II, an evolutionary genetic algorithm, is employed with ML models as objective functions to obtain the optimal Pareto Front solutions. Further, we incorporate manifold learning and unsupervised clustering to screen the Pareto Front and to select a few unique solutions which can facilitate added analysis for implementation. Furthermore, we introduce the application of conformal predictions for uncertainty quantification, ensuring reliability of the framework. Overall, the proposed approach enables interpretable and reliable property prediction, thus accelerating alloy design in thermomechanical processing.

4.1 Introduction

Thermomechanical controlled processing (TMCP) of metals is crucial in modern manufacturing processes. Its success largely stems from its ability to shape metals into

desired forms and simultaneously achieve their mechanical properties (1). In response to escalating demands for cost-effective high-quality products in today's global economy, achieving precise and error-free manufacturing has become more important than ever (2). The high-temperature and rapid-paced nature of TMCP industries restricts in-situ control, requiring post-production testing and corrections, escalating costs. In line with the 'right-first-time' approach, it becomes necessary to predict and optimize the variables in advance (3). The conventional semi-empirical approach in such large-scale industries encounters significant challenges. The sheer magnitude of parameters involved along with the complex, non-linear relationships between them hinders meaningful analysis through isolated parameter variations during trials.

Machine-learning approaches have emerged as the obvious choice to address the issue (4; 5; 6). ML methods are designed to automatically detect patterns in data, unlike traditional computational and experimental approaches in materials science, which are mostly studied empirically. This shift breaks away from the traditional composition-process-structure-property paradigm by directly relating data—specifically, composition and process variables—with labels representing properties. Supervised ML algorithms excel in solving such relationships by accurately learning underlying patterns, making them relatively trivial. For example, in hot rolling, they have been effectively used to predict the mechanical properties of as-rolled products (7; 8; 9; 10). Other applications also involve predicting roll force (11; 12; 13), bending force (14; 15), microstructure

evolution (16), rolling schedule (17), width deviation (18; 19), crown (20; 21), defect detection (22; 23), oxide scale thickness (24), among others.

However, the real challenge lies in tackling inverse problems, wherein ML is leveraged to discover alloys based on desired properties (25). Some of the major challenges in inverse design problems are multiple objectives, non-unique solutions, and large search space leading to combinatorial explosion (26; 27). The common practice among researchers involves employing a trained machine learning model as the optimization objective function, integrated with evolutionary algorithms such as Genetic Algorithm (GA) (28), Particle Swarm Optimization (PSO) (29; 30), Strength Pareto Evolutionary Algorithm 2 (SPEA2)(31), Non-dominated Sorting Genetic Algorithm II (NSGA-II) (32; 33), Multi-objective Evolutionary Algorithm Dominance and Decomposition (MOEA/DD) (34; 35) etc. In this work, we employ a similar approach which involves ML models as objective functions coupled with a multi-objective optimization algorithm to optimize the strength-ductility trade-off of hot-rolled steels. Adding to the optimization framework, we also analyze the Pareto front to select unique alloy solutions using a combination of manifold learning and unsupervised clustering.

Machine learning frameworks often face skepticism from metallurgists regarding their interpretability and reliability. Interpretability, or the ability to understand how a model makes decisions, is crucial for transparency. Simple models like linear regression or low-depth decision trees are easy to understand and offer high transparency but might not

effectively capture data patterns. More complex models, such as neural networks or gradient boosted trees, can handle intricate data but are harder to interpret. According to Lipton et al.(36), a machine learning model is interpretable if it meets three criteria : it must be transparent across the entire model (simulatability), in its individual components (decomposability), and in its training algorithm (algorithmic transparency). Hence, there is a trade-off between model complexity and model interpretability (37). Typically, researchers categorize interpretable methods into two types : ante-hoc and post-hoc. Ante-hoc methods involve models that are intrinsically interpretable, like linear regression, where the rationale behind each decision is clear from the beginning. Post-hoc methods focus on deciphering the decision-making processes of "black-box" models after training. For a more comprehensive overview of these methods in materials science, readers can refer to recent studies detailed in the literature (38; 39). The aim here is to implement methods that preserve the representation power of ML models while improving their transparency. Shapely Additive exPlanation (SHAP) (40) is one such post-hoc tool which has garnered widespread attention, including its successful application in steel alloy design (41; 42; 43). We propose the use of TreeSHAP, a method in the SHAP library, to add a layer of interpretability to our models. Section 4.2.5 presents a detailed discussion of the rationale and theoretical underpinnings of using TreeSHAP.

As mentioned above, another key concern in machine learning frameworks is their reliability. Reliability refers to the trustworthiness or confidence in the predictions made by

ML algorithms, which is critical for real-world applications where understanding the associated risk is essential. Providing a single metric of predictive performance, like RMSE or MAE, is insufficient for such applications. Instead, one needs access to uncertainty quantification (UQ) to make well-informed decisions (44). A common approach to achieve this would be through probabilistic machine learning methods, such as Gaussian Processes and Bayesian Neural Networks. These methods inherently learn probabilistic distributions and can make predictions with associated uncertainty. Unfortunately, they are often computationally expensive, particularly working with high-dimensional industrial datasets. Alternative computationally feasible methods have been developed within neural network architectures such as variational inference, deep ensembles, and Monte Carlo (MC) dropouts. Lately, these UQ methods have caught the attention of metallurgists as well (45; 46; 47; 48). An avid reader can find a more comprehensive explanation of these methods in a recently published tutorial (49). To highlight, these methods are model-specific, and hence, puts additional constraints on the choice of modelling method. We propose the use of conformal predictions using Jackknife+(50) which is a model-agnostic framework for UQ with a mathematical guarantee of marginal convergence on predictive intervals (51). Recently, researchers have successfully applied the conformal predictions based UQ in domains such as cyclone forecasting (52), rockburst assessment (53), and various other reliability systems (54; 55). To the best of the author's knowledge, it has not been tried in materials discovery or property predictions. We have described the

theoretical background and implementation in section 4.2.4

In essence, we propose an interpretable and reliable framework designed for both property prediction and inverse alloy design. Additionally, we aim to implement the framework using computationally inexpensive, open-access and compatible Python libraries, which will allow users to implement it with minimal effort.

4.2 Methodology

4.2.1 Dataset Background

This study uses historical data from V-based grades rolled in the Thin Slab Direct Rolling (TSDR) mill of Algoma Steel Inc. (Ontario, Canada). In conventional hot rolling, slabs (~ 200 mm) are reheated to be fed into the rolling mill, which includes 5-7 roughing passes and subsequently 6-7 finishing passes to get the desired strip thickness. TSDR employs a streamlined continuous process where high-speed cast thin slabs (~ 70 -80 mm) are directly hot rolled into strips involving 1 roughing pass and subsequent finishing pass, six here. This eliminates the need for intermediate cooling and reheating, as compared to traditional hot rolling process. It simplifies the material flow, significantly reduces energy consumption, and enhances the production efficiency (56). TMCP variables, along with the alloy composition, determine the final mechanical property of the strip. In this work, strength-ductility optimization is achieved through alloy composition optimization.

4.2.2 Pre-processing

Features were chosen based on domain knowledge to accurately reflect the causal relationship with properties. From the raw dataset, 8,313 data instances could be compiled after removing duplicate, erroneous, and missing data values and incorrect data types. The *3-sigma rule* criterion with scatter plot visualization was used to remove global outliers from the dataset. Mathematically, it is expressed as :

$$\text{Outliers are points where } x < \mu - 3\sigma \text{ or } x > \mu + 3\sigma$$

Finally, 6995 data instances were obtained, the descriptive statistics of which is presented in table 4.1. The dataset was shuffled and split into train and test sets in the ratio of 3 :1. The test data was never seen by any models during hyperparameter optimization or training. *Min-Max Normalization* was applied to both features and target variables :

$$x_{\text{scaled}} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (4.1)$$

where x is an original value, $\min(x)$ is the minimum value in the dataset, $\max(x)$ is the maximum value, and x_{scaled} is the normalized value. The transformation scales the values of x to a fixed range of $[0, 1]$. The normalization statistics were computed from the training dataset which were then used to scale both the training and testing datasets to prevent

information leakage.

4.2.3 Predictive Modelling

We propose the use of eXtreme Gradient Boosting (XGBoost) (57) as the primary modeling method, which is detailed in Section 4.2.3.1. In the past, XGBoost has been successfully used by several researchers in predictive modeling of similar datasets (58; 59; 60). We also compare the performance of the proposed method with six other standard supervised machine learning algorithms, namely, k -Nearest Neighbors (KNN), Kernel Ridge Regression (KRR), Random Forest (RF), Light Gradient Boosting Machine (LightGBM), Support Vector Regression (SVR), and Multi-Layer Perceptron (MLP).

4.2.3.1 XGBoost

XGBoost (eXtreme Gradient Boosting) is a decision-tree-based model designed to improve speed and performance in predictive modeling. It builds decision trees sequentially in an ensemble, with each tree correcting the errors of the previous ones. The final prediction model is a linear combination of several trees. Mathematically, the prediction for any data point x_i at iteration t can be written as :

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i), \quad (4.2)$$

where f_k denotes the k -th decision tree in the sequence.

XGBoost optimizes an objective function that combines a differentiable loss function like Mean Squared Error (MSE) and a regularization term to control overfitting :

$$\mathcal{L}(\phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k), \quad (4.3)$$

where $l(y_i, \hat{y}_i)$ measures the difference between the actual values y_i and the predictions \hat{y}_i , and $\Omega(f_k)$ represents the regularization term for the k -th tree.

The regularization term is important to prevent overfitting and promotes generalization :

$$\Omega(f_k) = \gamma T_k + \frac{1}{2} \lambda \sum_{j=1}^{T_k} w_j^2, \quad (4.4)$$

where T_k is the number of terminal nodes in tree k , w_j are the weights on the leaves, γ adds a penalty for the number of leaves, and λ is the L2 penalty on the leaf weights.

For training, XGBoost uses an additive approach by adding new trees to minimize the errors of the existing model, using a second-order Taylor expansion of the loss function. The expansion includes both the gradient and the Hessian (second derivative), allowing for better optimization :

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(x_i), \quad (4.5)$$

where η is the learning rate (shrinkage parameter) that controls the impact of each new tree.

When training, the model uses a metric called gain to evaluate possible data splits and

determine the extent of improvement achieved through each split, as measured by loss reduction :

$$\text{Gain} = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G^2}{H + \lambda} \right] - \gamma, \quad (4.6)$$

where G_L , G_R , and G are sums of gradients in the left, right, and entire node respectively, and H_L , H_R , and H are the corresponding sums of Hessians. If the gain is less than γ , it's best to not include the branch i.e. pruning. The overall quality of a tree can be assessed using the structure score, which calculates how well it fits the data compared to its complexity :

$$S(f) = -\frac{1}{2} \sum_{j=1}^T \left[\frac{G_j^2}{H_j + \lambda} + \gamma \right], \quad (4.7)$$

where G_j and H_j are the gradient and Hessian calculations for each leaf j .

In addition to the components discussed above, XGBoost includes several hyperparameters that can influence model performance. A brief summary of all the hyperparameters used in this work is presented in the table 4.2.

Features	Description	Mean	Std Dev	Min	Max
Al	Aluminium	0.025283	0.003414	0.0142	0.0366
B	Boron	0.000024	0.000055	0.00	0.0002
C	Carbon	0.053050	0.003265	0.0423	0.0641
Ca	Calcium	0.003958	0.001020	0.0007	0.0073
Cr	Chromium	0.026005	0.006908	0.0103	0.0535
Cu	Copper	0.030594	0.010539	0.0017	0.0684
Mn	Manganese	0.770181	0.359122	0.3527	1.5531
Mo	Molybdenum	0.013385	0.018543	0.00	0.0584
Nb	Niobium	0.004771	0.002559	0.00	0.0099
Ni	Nickel	0.015506	0.005445	0.0043	0.0453
P	Phosphorus	0.009391	0.002491	0.0031	0.0179
S	Sulphur	0.002791	0.001142	0.0001	0.0069
Si	Silicon	0.077896	0.098426	0.0036	0.2971
Sn	Tin	0.001136	0.000632	0.00	0.0035
Ti	Titanium	0.000808	0.000435	0.00	0.0024
V	Vanadium	0.064459	0.035584	0.0243	0.1545
N	Nitrogen	0.013504	0.002844	0.00805	0.02205
CS	Avg. Casting Speed	3.28	0.15	2.78	3.61
ST	Superheat	18.16	4.02	5.87	30.19
Width	Final Width	1348.27	150.82	1031.0	1616.0
Gauge	Final Thickness	3.29	0.996	1.58	6.20
HT	Holding Time	18.65	2.75	9.10	56.90
F6RS	F6 Roll Speed	3.73	0.967	1.66	6.67
RET	Roughing Entry Temp.	1118.71	6.33	1096.0	1140.0
FET	FM Entry Temp.	1021.82	9.07	992.0	1050.0
FT	Finishing Temp.	879.35	19.89	833.0	906.0
CT	Coiling Temp.	603.71	4.99	578.0	629.0
UTS (MPa)	Ultimate Tensile Strength	542.13	74.82	422.66	734.02
Elong (%)	2" Elongation	27.68	3.47	18.00	38.00

Table 4.1 – Descriptive statistics of the dataset

Hyperparameter	Description
lambda	L2 regularization term on weights, helps control overfitting by penalizing large weights.
alpha	L1 regularization term, encourages sparsity in the leaf weights, useful for feature selection.
colsample_bytree	Fraction of features used per tree. A lower value provides more regularization.
subsample	Fraction of instances used per tree. A technique to reduce overfitting and variance.
learning_rate	Step size shrinkage used to prevent overfitting. Lower values require more boosting rounds.
n_estimators	Number of trees. More trees can improve accuracy but may lead to overfitting.
max_depth	Maximum depth of a tree. Controls overfitting as higher depth will allow model to learn relations very specific to a particular sample.
min_child_weight	Minimum sum of instance weight (hessian) needed in a child. Higher values prevent a model from learning overly specific patterns, thus controlling overfitting.

Table 4.2 – Summary of XGBoost hyperparameters tuned in this work.

4.2.3.2 Evaluation Metrics

To assess the performance of the regression models, three different metrics are employed which include Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2).

Root Mean Squared Error (RMSE) RMSE is a widely accepted measure for evaluating the accuracy of models in predicting numerical data. Formally, it is defined as the square root of the average of the squares of the differences between predicted and actual values. Mathematically, RMSE is given by :

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (4.8)$$

where y_i are the actual values, \hat{y}_i are the predicted values, and n is the number of observations.

Mean Absolute Error (MAE) MAE calculates the average magnitude of errors in predictions, regardless of their direction. It assigns equal importance to all individual differences, making it a linear score. It is calculated as :

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (4.9)$$

where $|y_i - \hat{y}_i|$ denotes the absolute error between the actual and the predicted values.

Coefficient of Determination (R^2) The R^2 metric gauges the model's goodness of fit

and predicts how accurately unseen samples will be predicted. The optimal score is 1.0 and can be negative if the model is arbitrarily worse. If a model predicts the average y without considering the input features, its R^2 score would be 0.0. R^2 is defined as :

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (4.10)$$

where \bar{y} is the mean of the observed data :

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (4.11)$$

4.2.4 Uncertainty Quantification

Conformal Prediction (CP) is a statistical method that adds a layer of probabilistic interpretation to machine learning predictions. The prediction intervals guarantee coverage of the true response variable with a probability of $1 - \alpha$ or higher. The mathematical foundation of CP is based on the notion of exchangeability, assuming future data is exchangeable with training data. For every instance, the method computes a conformity score α_i that measures its deviation from typical instances observed in the training data. For regression, the conformity scores can be calculated based on residuals. For a new prediction \hat{y}_{new} , the prediction interval is constructed so that it covers the true target value y_{new} with a confidence level of $1 - \alpha$, based on the empirical distribution of the conformity

scores calculated from a calibration set (50). While there are various methods for constructing predictive confidence intervals within CP, we focus on the Jackknife+ (51) method in this work.

Jackknife+ builds upon the Jackknife resampling technique in conformal prediction to generate more robust prediction intervals. Unlike the classical Jackknife, Jackknife+ takes into account the dependence between the training set and the test instance to avoid uncertainty underestimation. The approach requires calculating predictions for each data point in the training set by leaving one out. A prediction is made for each omitted data point by training a new model on the remaining data. This results in an ensemble of predictions from which uncertainty estimates can be derived :

$$\hat{y}_{-i} = f_{-i}(x_i), \quad (4.12)$$

where \hat{y}_{-i} is the prediction for the i -th instance when it is left out during training, and f_{-i} represents the model trained on the dataset excluding the i -th instance.

The prediction interval for a new instance x_{new} is then adjusted according to the variability observed in these leave-one-out predictions. Specifically, the bounds of the interval are set to capture the proportion $1 - \alpha$ of these adjusted predictions :

$$\text{PI} = \left[Q_{\alpha/2}(\{\hat{y}_{-i} - \epsilon_i\}), Q_{1-\alpha/2}(\{\hat{y}_{-i} + \epsilon_i\}) \right], \quad (4.13)$$

Algorithm 1 Conformal Prediction with Jackknife+

- 1: Initialize :
 - 2: Split D_{train} into K subsets (using K-fold cross-validation or leave-one-out strategy)
 - 3: **for** $k = 1$ to K **do**
 - 4: $D_{\text{train}_k} = D_{\text{train}} \setminus k\text{-th subset}$ ▷ Train set excluding the k-th subset
 - 5: Train the model M on D_{train_k}
 - 6: Predict the response for instances in the k-th subset using model M
 - 7: Calculate absolute residuals for each prediction :
 - 8: $\text{residuals}_k[i] = |y_i - \hat{y}_i|$
 - 9: **end for**
 - 10: Aggregate residuals :
 - 11: Combine residuals from all K subsets into a single list : $\text{all}_{\text{residuals}}$
 - 12: Determine the empirical quantile of the residuals :
 - 13: Compute the $(1 - \alpha)$ quantile, q , of $\text{all}_{\text{residuals}}$
 - 14: Retrain the model M on the entire D_{train}
 - 15: Predict the response \hat{y}_{new} for x_{new} using the fully trained model M
 - 16: Construct the prediction interval :
 - 17: $y_{\text{lower}} = \hat{y}_{\text{new}} - q$
 - 18: $y_{\text{upper}} = \hat{y}_{\text{new}} + q$
 - 19: **return** Prediction interval $[y_{\text{lower}}, y_{\text{upper}}]$
-

where Q denotes the quantile function, and ϵ_i are the residuals $y_i - \hat{y}_{-i}$. Another approach, based on cross-validation (CV+), can be used to reduce computational time instead of the leave-one-out approach. Analogous to the Jackknife+ method, the training set is split into K -disjoint subsets. The Jackknife+ method can be seen as a special case of the CV+ method, with K being equal to the number of samples. In cases where the Jackknife+ method is not computationally feasible, the CV+ method provides a conservative yet reasonable compromise for large datasets. A pseudocode for UQ calculation using Jackknife+ is presented in algorithm 1.

4.2.5 TreeSHAP : Model Interpretability

Before discussing TreeSHAP, we must mention that there are three different metric options i.e. by weight, cover and gain to calculate feature importance given by the XGBoost library itself. But these methods are inconsistent which implies that a model can shift its reliance towards a specific feature, even if the assigned importance estimate for that feature decreases (61).

Shapley values are derived from cooperative game theory, which assigns a fair payout to players based on their contribution to the total game outcome. In a predictive model, the features function as players, and the payout is the deviation of the prediction from the mean.

The Shapley value for feature i is formally defined as :

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} [f(S \cup \{i\}) - f(S)], \quad (4.14)$$

where N is the set of all features, S is a subset of features excluding i , n is the total number of features, $f(S)$ is the model prediction with features in S , and $f(S \cup \{i\})$ is the prediction with feature i added to S .

SHAP uses Shapley values to explain the impact of features in ML models ensuring local accuracy and consistency in attributions across different predictions. TreeSHAP is a model-specific variant of SHAP method which optimizes the evaluation of Shapley values for tree-based models by utilizing their structural properties. The calculation involves conditional expectations based on the paths through the decision trees unlike model-agnostic approaches like KernelSHAP which rely on marginal expectations. As TreeSHAP is a local explanation technique, it comes with the theoretical guarantees of local accuracy and consistency. These local explanation can be further used to build consistent global explanations as well. For example, the mean |SHAP| value, a measure of overall feature importance across all predictions can be defined as :

$$\text{mean|SHAP|} = \frac{1}{n} \sum_{i=1}^n |\phi_i|, \quad (4.15)$$

where ϕ_i are the SHAP values for feature i across all instances. This provides an average

magnitude of impact, indicating how much, on average, each feature shifts the model output from the base value.

4.2.6 Multiobjective Optimization

Multiobjective optimization (MOO) involves simultaneously optimizing multiple objectives, often conflicting, subject to certain constraints. It aims to find a set of solutions called Pareto-optimal solutions, rather than a single optimal solution like single-objective optimization (62). Mathematically, a multiobjective optimization problem can be expressed as follows :

$$\min_{\mathbf{x}} \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x})), \quad (4.16)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)$ represents the vector of decision variables within the feasible region defined by the constraints :

$$g_j(\mathbf{x}) \leq 0, \quad j = 1, 2, \dots, m, \quad (4.17)$$

and

$$h_k(\mathbf{x}) = 0, \quad k = 1, 2, \dots, p. \quad (4.18)$$

Generating the exact pareto set is computationally intractable, given the combinatorially

large search space. Hence, we employ a search strategy based on evolutionary algorithm (EA). In this approach, a population of candidate solutions is maintained, and selections are made from it for mating. The combinations of these selections create new potential solutions. This process mimics natural evolution, with candidates resembling individuals in a population. EAs use mutation and crossover to introduce variability and recombination. Selection pressures gradually enhance the population towards optimal solutions, guided by fitness evaluations ranking individuals based on objective fulfillment (63).

NSGA-II (Non-dominated Sorting Genetic Algorithm II) is an evolutionary algorithm developed by Deb et al. (64; 65) for MOO problems. It is known for its efficient non-dominated sorting method that categorizes the population using dominance criterion. Each subsequent front consists of individuals dominated by the preceding front. The algorithm incorporates two primary genetic operators : crossover and mutation. Crossover creates new offspring by merging elements from two parent solutions, promoting population diversity while retaining the qualities of the top solutions. Mutation helps maintain genetic diversity and prevents premature convergence to local optima by introducing random changes to individual solutions. NSGA-II also incorporates a crowding distance mechanism to maintain diversity in the Pareto front. The mechanism calculates solution density around an individual and favors less crowded regions, promoting an even distribution across the Pareto front. A pseudocode of the NSGA-II algorithm is presented in algorithm 2.

Algorithm 2 NSGA-II : Non-dominated Sorting Genetic Algorithm II

- 1: **Input** : Population size N , number of generations G
 - 2: **Output** : Pareto-optimal set
 - 3: Initialize population P_0 randomly
 - 4: Evaluate fitness of P_0
 - 5: **for** $g = 1$ to G **do**
 - 6: Perform non-dominated sorting on P_g
 - 7: Calculate crowding distance for each individual in P_g
 - 8: Perform selection based on fitness and crowding distance
 - 9: Apply crossover and mutation to generate Q_g
 - 10: Combine P_g and Q_g to form R_g
 - 11: Perform non-dominated sorting on R_g
 - 12: Select the top N individuals from R_g to form P_{g+1}
 - 13: Evaluate fitness of P_{g+1}
 - 14: **end for**
 - 15: **return** Non-dominated individuals from P_G
-

4.2.7 Post-Pareto Analysis

4.2.7.1 t-SNE

t-SNE (66) is a commonly used method for dimensionality reduction and visualizing high-dimensional datasets. It transforms data point similarities into joint probabilities and seeks to minimize the Kullback-Leibler divergence between the joint probabilities of the low-dimensional embedding and the high-dimensional data. This technique can effectively capture the local structure and global patterns in high-dimensional data. Mathematically, t-SNE begins by calculating the conditional probability $p_{j|i}$ that represents the probability of selecting x_j as a neighbor of x_i :

$$p_{j|i} = \frac{\exp(-\|x_i - x_j\|^2/2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2/2\sigma_i^2)}, \quad (4.19)$$

where σ_i is the variance of the Gaussian centered at data point x_i . This incorporates a notion of directional similarity of point j to i . The symmetrized probabilities p_{ij} are then defined as :

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n}, \quad (4.20)$$

where n is the number of data points. In the low-dimensional space, t-SNE uses a Student's t-distribution to model the pairwise affinities q_{ij} :

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}}, \quad (4.21)$$

where y_i and y_j are the low-dimensional representations of x_i and x_j respectively. The Kullback-Leibler divergence between the distribution P in the high-dimensional space and Q in the low-dimensional space is minimized :

$$C = \text{KL}(P||Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}}. \quad (4.22)$$

There are three main parameters related to t-SNE algorithm optimization. *Perplexity* can be regarded as an estimate of the number of nearby neighbors for each point. It plays

an important role in the balance between local and global structure of the data, influencing σ_i . *Learning rate* determines the size of each step taken during iteration to minimize the cost function. If too high, the map may become chaotic ; if too low, the optimization might get trapped in local minima. *Iterations* directly influence the cost function optimization. Improved convergence stability can be achieved by increasing the number of iterations, although this comes at the expense of computational time.

4.2.7.2 K-Means Clustering

K-Means (67) is a widely used unsupervised ML method to partition observations into k clusters of equal variance. The algorithm approaches this problem by first randomly selecting k centroids, one for each cluster. The next step involves assigning each data point to the nearest centroid. After all points have been assigned, the positions of the k centroids are recalculated. This process is repeated until the centroids are stable, indicating convergence. The objective function in K-means aims to minimize the total intra-cluster variance by summing the squared distances between data points and their nearest cluster center :

$$J = \sum_{i=1}^n \min_{\mu_j \in C} (\|\mathbf{x}_i - \mu_j\|^2) \quad (4.23)$$

where μ_j is the centroid of cluster C_j . Proper initialization of the centroids is crucial for the algorithm to converge. In order to handle this issue, we can opt for the K-means++ initialization scheme instead of random initialization. Better convergence is achieved in K-

means++ by initializing distant centroids. K-means is not purely unsupervised since the algorithm requires the optimal value of k to be given (68). Two commonly used methods to determine the optimal k are the Silhouette Score (69) and the Elbow Method (70).

The Silhouette Score evaluates an object's resemblance to its cluster in comparison to other clusters. The Silhouette Score for a set of samples is given by :

$$s = \frac{b - a}{\max(a, b)} \quad (4.24)$$

where a is the mean distance between a sample and all other points in the same cluster, and b is the mean distance between a sample and all other points in the nearest cluster that the sample does not belong to. s can take a value between -1 and +1, with a high value suggesting a good match to its own cluster and a poor match to neighboring clusters. The final score is the mean of Silhouette Score calculated for each sample.

The Elbow Method involves plotting the total Within-Cluster Sum of Squares (WSS), also known as inertia, as a function of the number of clusters :

$$WSS(k) = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (4.25)$$

where C_i is the set of points in cluster i and μ_i is the centroid of cluster i . Inertia measures how tightly packed the clusters are and represents the overall variance within them. When plotting inertia ($WSS(k)$) against k , we usually see an arm like curve, and the “elbow”

(inflection point on the curve) is seen as a sign of optimal number of clusters. The underlying assumption here is that it's not beneficial to increase the number of clusters unless there is a substantial increase in explained variance.

4.2.8 Implementation Details

This study was carried out using a suite of open-source Python libraries, ensuring that the proposed framework is easily replicable. We employed XGBoost (57) and LightGBM (71) by their respective libraries. All other machine learning models, including KNN, KRR, RF, SVR, and MLP, were implemented using the Scikit-learn library (72). For Bayesian hyperparameter optimization, we used Optuna (73). Uncertainty quantification was conducted through the MAPIE library (74). SHAP library (61) provided feature importance analysis for model interpretability. Multiobjective optimization was performed using Pymoo (75), which provides easy access to various optimization algorithms including NSGA-II. For post pareto data visualization, t-SNE and K-means were also implemented via Scikit-learn. Plots were generated using Matplotlib and Seaborn library in Scikit-learn library. All computations were performed on a MacBook Pro with an Apple M1 Pro chip, equipped with 16GB of RAM and a 512GB hard disk.

4.3 Results and Discussion

4.3.1 Predictive Modelling

The comparison of the performance of seven ML models on the test dataset can be found in table 4.3. The hyperparameters for all models were tuned on train dataset using 5-fold cross validation strategy and Bayesian optimization. The ML models will serve as the objective functions in the optimization process. Beginning with the initial question : How do we determine the suitable model ? In the existing literature, there is no fixed guide to choose

Model	UTS			Elongation		
	RMSE	MAE	R2	RMSE	MAE	R2
KNN	13.1616	9.9097	0.9682	1.3382	1.0147	0.8467
KRR	13.6417	10.4024	0.9658	1.3814	1.0409	0.8366
RF	13.2910	10.0665	0.9675	1.3240	0.9980	0.8499
LightGBM	12.8392	9.7332	0.9697	1.3115	1.0079	0.8528
XGBoost	12.0169	9.2157	0.9738	1.2927	0.9823	0.8613
SVR	14.1967	10.8506	0.9629	1.4389	1.0861	0.8228
MLP	14.8649	11.5660	0.9594	1.3975	1.0436	0.8328

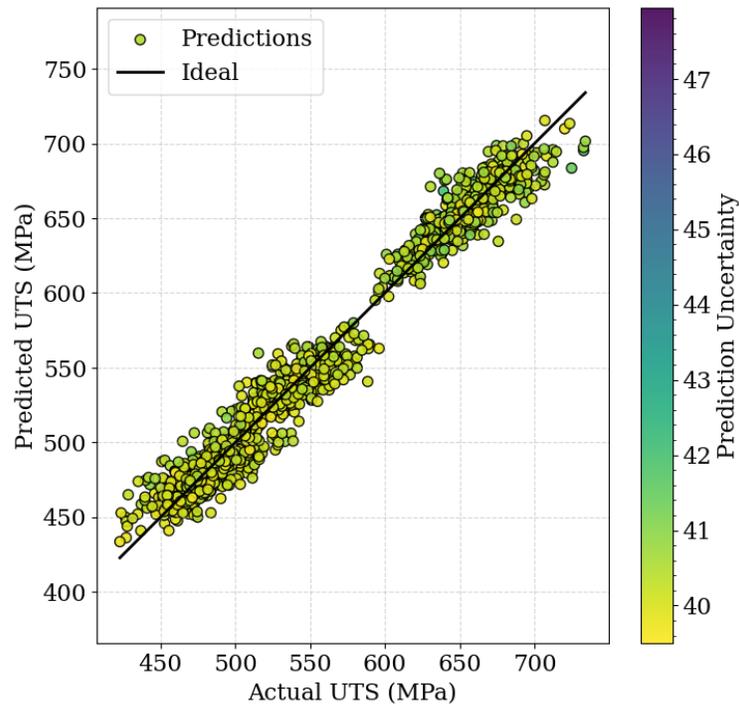
Table 4.3 – Performance Metrics for various ML models for UTS and Elongation (Test Set).

the modeling method for objective functions in MOO. Selecting the best machine learning algorithm is non-trivial because of the "No Free Lunch" Theorems (76; 77), which state that no model is universally superior across all applications. On average, all models have similar

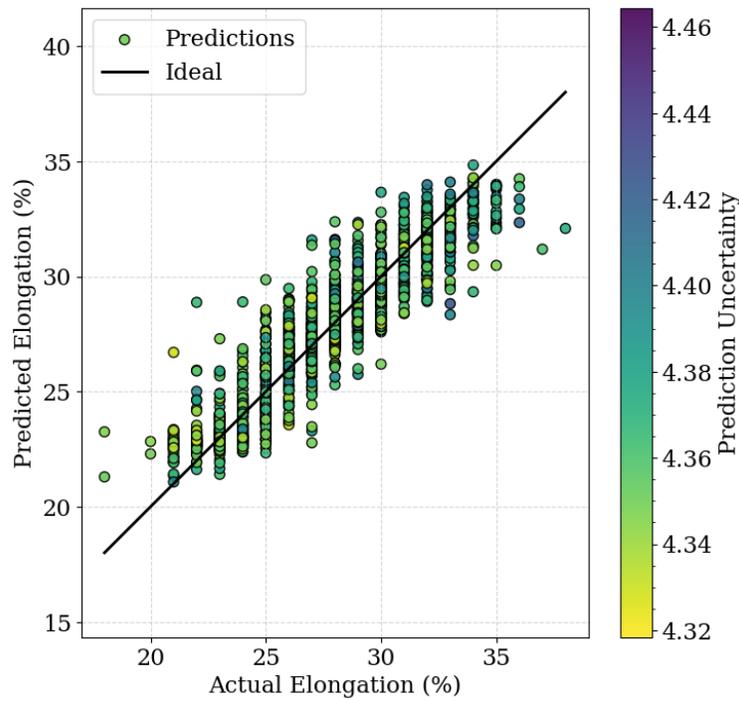
performance across various data distributions, and no single model performs the best in every scenario. As a result, we find it fitting to designate the best performing models as the objective functions. Recently, studies have demonstrated that decision tree-based algorithms are superior to other algorithms for medium-sized tabular datasets (less than 10k instances), such as ours, due to their distinctive characteristics (78; 79). This could clarify why tree-based models like RF, LightGBM, and XGBoost perform exceptionally well. The XGBoost model is clearly superior to all other models for both, UTS and Elongation, and thus our choice for objective functions. It could be easily noted that predicting UTS is simpler than Elongation. The Elongation values, being pseudo-continuous and limited to integer values, pose a challenge for regression algorithms designed for continuous data prediction.

4.3.2 Uncertainty Quantification

The predictions of the selected XGBoost models are displayed in Figure 4.1. The colour contrast in the scatters reflects the uncertainty derived from conformal predictions using Jackknife+ strategy. To digress from the flow, it would be apt to discuss another suitable application of these conformal predictive models. The testing protocol in the rolling mill can be streamlined by using uncertainty-aware predictive models. A strategy could be developed to prioritize high-risk specimens of rolled products and minimize random sampling using these models. As a result, excessive manual testing would be reduced, leading to cost, materials and man-hours saving and thus improved production efficiency. It



(a)



(b)

Figure 4.1 – Performance of predictive models with uncertainty as colour mapping. (a) True vs Predicted for UTS , (b) True vs Predicted for Elongation.

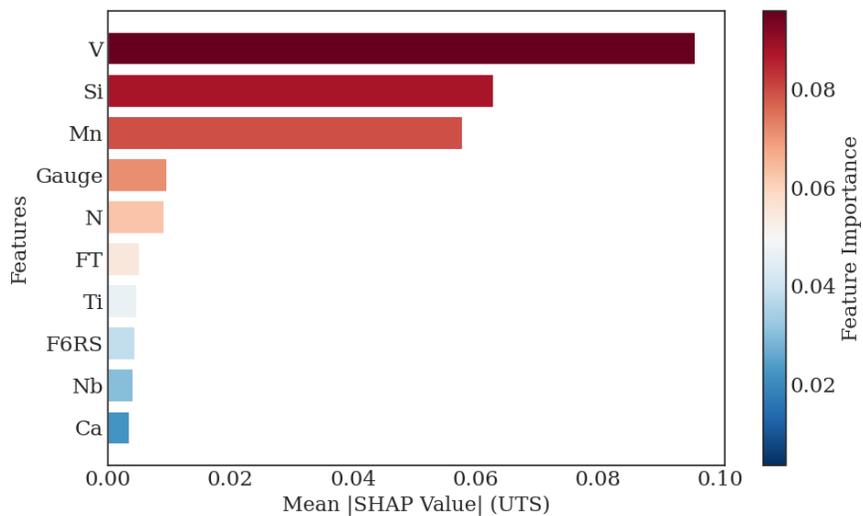
should be noted that this method does not distinguish between different types of uncertainty, such as aleatoric and epistemic, which is not the focus of this research.

4.3.3 Model Interpretability

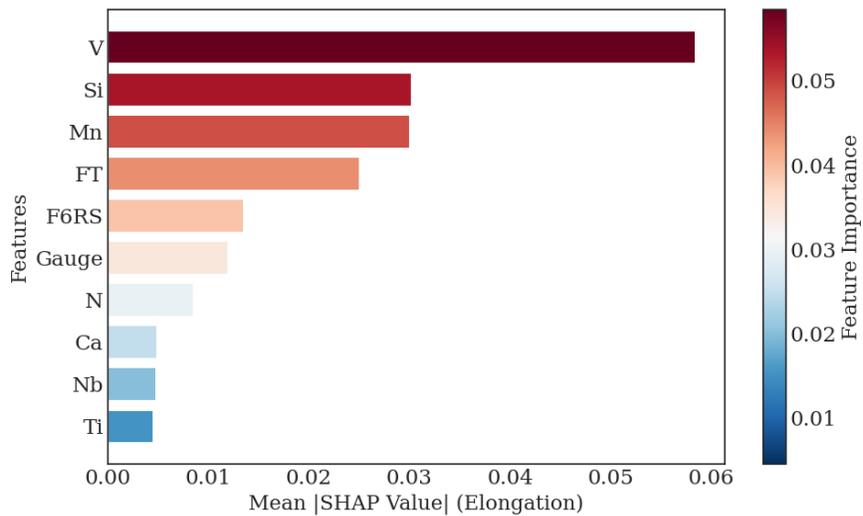
Figure 4.2 displays the top 10 features according to their contribution. The plots are presented on a normalized scale for proper comparison. In both the properties, Vanadium, Silicon, and Manganese comes out as top three features but on different scales of impact and have opposing effect on UTS and Elongation. This gives an impetus to this work showing ample opportunity for alloy composition optimization. While we concentrate on global significance using mean SHAP values, other local and global attributions can be effortlessly visualized with the SHAP library.

4.3.4 Multiobjective Optimization

Now that we have chosen our ML models, we reformulate the MOO description for our problem in hand : We aim to minimize the objective functions (ML Models) $f_1(\mathbf{x}) = -\text{UTS}(\mathbf{x})$ and $f_2(\mathbf{x}) = -\text{Elongation}(\mathbf{x})$. The NSGA-II algorithm, being designed for minimization, requires us to reformulate the objectives as their negatives. So, we minimize $\mathbf{F}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x})]$, with the objective space defined by \mathbb{R}^2 . The decision variables $\mathbf{x} \in \mathbb{R}^{27}$ represent the features of our ML models i.e., chemical compositions and processing parameters, each typically bounded within a defined range $x_{\min,i} \leq x_i \leq x_{\max,i}$ for



(a)



(b)

Figure 4.2 – Feature Importance plot based on the selected XGBoost Models and mean shap values (a) For Ultimate Tensile Strength (b) For Elongation.

$i = 1, 2, \dots, 27$. It's important to mention that TMCP variables in the decision space have narrower bounds to accommodate the rolling mill schedule. The rolling schedule is determined according to the target thickness of the strip. Since ML models struggle to generalize beyond the distribution, the general constraints ensures that the predicted properties stay within acceptable limits :

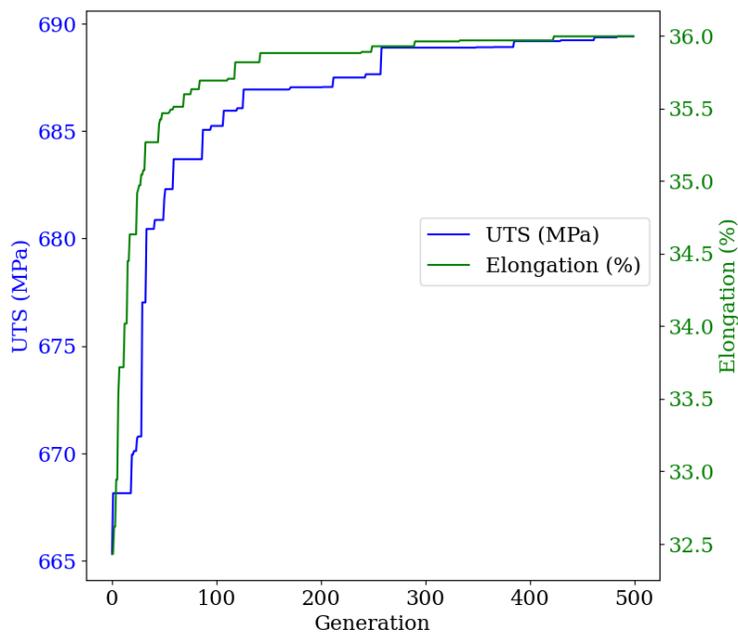
$$g_1(\mathbf{x}) = -\text{UTS}(\mathbf{x}) \leq 0, \quad (\text{UTS minimum}) \quad (4.26)$$

$$g_2(\mathbf{x}) = \text{UTS}(\mathbf{x}) - 1 \leq 0, \quad (\text{UTS maximum}) \quad (4.27)$$

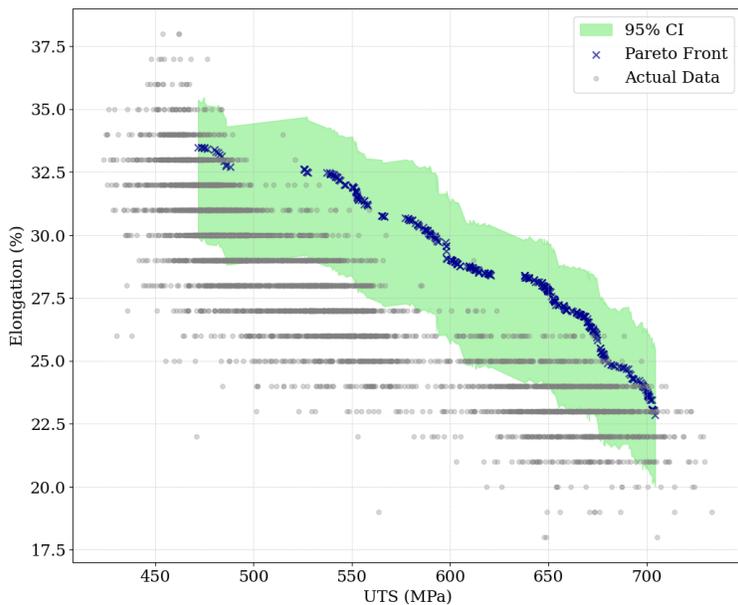
$$g_3(\mathbf{x}) = -\text{Elongation}(\mathbf{x}) \leq 0, \quad (\text{Elongation minimum}) \quad (4.28)$$

$$g_4(\mathbf{x}) = \text{Elongation}(\mathbf{x}) - 1 \leq 0. \quad (\text{Elongation maximum}) \quad (4.29)$$

The initial population consisted of 1000 individuals and each generation had 200 offspring. For crossover, Simulated Binary Crossover (SBX) operator and for mutation, Polynomial Mutation (PM) operator were used with their default hyperparameters from the Pymoo library. To ensure convergence, the algorithm was executed for 500 generations. Convergence was tracked using objective functions and hypervolume. The convergence of the NSGA-II algorithm with respect to the objective functions is shown in Figure 4.3a. Figure 4.3b displays the pareto front resulting from the above formulation with confidence interval (in green) obtained by applying the Jackknife+ UQ framework to the optimized solution population.



(a)



(b)

Figure 4.3 – NSGA-II results (a) Convergence of Objectives over generations, (b) Optimized Pareto Front with 95% Confidence Interval. The actual distribution is also presented in the background for comparison.

The optimized solution with CI incorporates risk assessment which is invaluable for process engineers to make informed decisions. Consider an alloy chosen from the pareto front, which shows a strength-elongation combination of 476 MPa UTS and 33% Elongation. We can assert, with 95% certainty, that it surpasses 446 MPa UTS and 30% Elongation as determined by lower bound in the CI. Since optimized solutions are out-of-distribution samples, there is no theoretical guarantee on the coverage, but it can be considered a conservative estimate. For future work, it would be beneficial to perform a study comparing the coverage of Jackknife+ with various probabilistic approaches for out-of-distribution optimized solutions.

4.3.5 Post-Pareto Analysis

Our aim now is to visually represent the pareto-optimal solutions and specifically identify a small subset of unique solutions from the resulting pareto set for further analysis. High dimensional datasets, like ours, make it nearly impossible to understand patterns or structure within data. The initial solution could be to incorporate multiple 2D or 3D plots, which is not easily summarized. Next strategy could be to try Principal Component Analysis (PCA), but it cannot preserve the non-linearity in the data structure. Therefore, we utilize a mix of manifold learning (t-SNE) and unsupervised clustering (K-means), which is explained briefly in the section 4.2.7. We start by experimenting the t-SNE embedding on our train dataset to check if it works for our case. We follow an approach similar to Kobak et. al (80) to optimize the t-SNE hyperparameters. Learning rate is $\sim n/12$, early exaggeration is

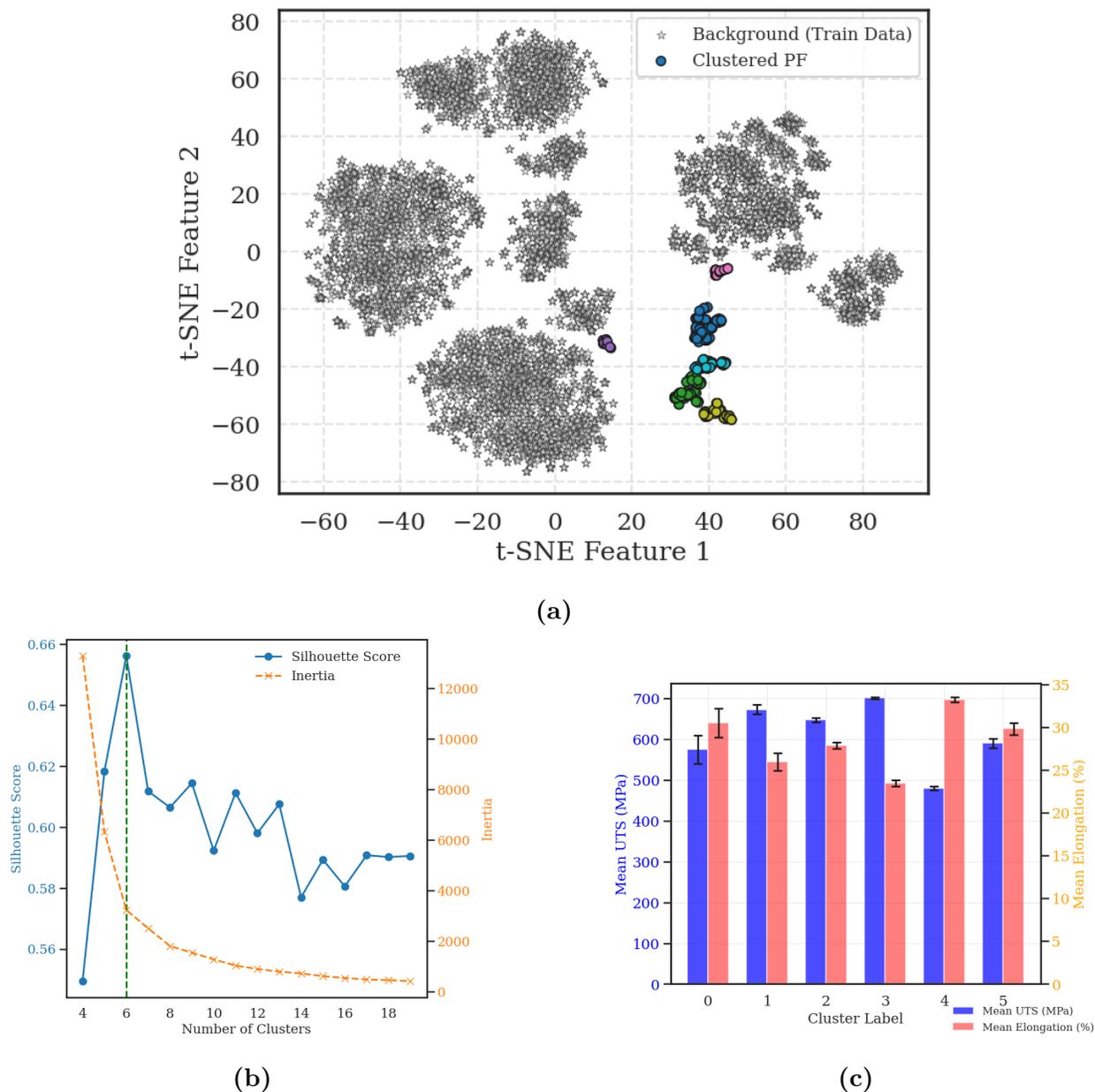


Figure 4.4 – Overview of clustering results and analysis : (a) Clustered Pareto fronts using t-SNE and K-means, background is t-SNE compressed train data, (b) Selection of optimal number of clusters using elbow method for inertia, (c) Cluster statistics to understand the distinct characteristics of each cluster.

4, perplexity is $\sim n/100$, and with PCA initialization. The PCA initialization, rather than random, helps in reproducibility and preserving global aspects of the data. We use the Barnes-Hut approximation to reduce the computational complexity of t-SNE. By employing the above approach, the train dataset could be effectively organized into clusters that correspond to different rolling practices utilized in the rolling mill. This gives us the confidence that the optimized t-SNE can effectively capture the global structure the data. Consequently, we implement t-SNE embedding on the combined dataset, which consists of train data and our pareto-optimal solutions. The figure 4.4a shows the combined dataset (train + PF) in two-dimensional embedding using t-SNE. The train dataset is represented in light grey. The PF solutions are colour-coded to represent different clusters obtained using K-means algorithm. The optimal number of clusters, five, are obtained using the silhouette score and inertia as shown in figure 4.4b. The statistics of the clusters are presented in figure 4.4c. Here, we can make the inference that the clusters truly represent unique solutions with different levels of UTS and Elongation. From the clusters we can select a small set of unique solutions which represent the entire PF. We employ an approach which selects the individual nearest to each cluster centroid. An alternative strategy would be to choose solutions with a lean composition, minimizing costly alloying elements like Vanadium. These unique solutions with risk assessment can be further analyzed thermodynamically and with domain experience for implementation.

4.4 Conclusion

In this work, we have systematically implemented data-driven methods to discover alloys with optimized mechanical properties. Our conclusions are as follows :

- (I) Tree-based prediction algorithms, including XGBoost, demonstrated superior performance for our dataset compared to SVM, KNN, KRR, and MLP. XGBoost achieved an R^2 of approximately 0.97 for UTS and 0.86 for % Elongation. The prediction of elongation was particularly difficult due to its pseudo-continuous nature.
- (II) TreeSHAP, in combination with XGBoost, provided a model-specific explanation method that enabled the global feature importance, adding a layer of interpretability.
- (III) To increase the trustworthiness of our ML framework, we integrated conformal predictions using the Jackknife+ approach. The same can also be implemented to reduce the need for extensive manual testing by focusing on strategic testing of high-risk specimens.
- (IV) The decision variable space of the NSGA-II algorithm was utilized to encode the process schedule for obtaining geometry-specific optimized compositions. The resultant Pareto Front, coupled with conformal predictions, can facilitate informed decision-making in alloy design.
- (V) Manifold learning techniques such as t-SNE, in combination with K-means clustering,

were employed for post-Pareto analysis. This approach allowed us to extract a few representative unique solutions that can be further analyzed for implementation.

- (VI) The proposed framework is not only applicable to the steel industry but can also be translated to other TMCP industries. We have ensured the reproducibility of our methods by utilizing computationally inexpensive, open-access and compatible libraries, encouraging adaptation within the broader scientific and industrial community.
- (VII) This work has primarily focused on discovering alloys, and it has displayed promising potential for improvement. For future work, it would be intriguing to explore both composition and process schedule together. The incorporation of a process schedule would involve adopting a hierarchical modeling and optimization approach, given the continuous nature of the process dependent on principles of mass and energy flow.

Acknowledgements

This work was supported by NSERC grant (CRDPJ 543584-19). The authors express their gratitude to Algoma Steel Inc., Canada, for their financial support, data assistance, and permission to publish this work. Additionally, the authors acknowledge the MITACS Graduate Fellowship awarded to Sushant Sinha during his research at McGill University.

Declaration of Competing Interest

The authors declare that they have no known competing interests that could have appeared to influence this work.

Data availability

The data required to reproduce these findings cannot be shared, due to confidentiality. No details of the plant have been disclosed to satisfy proprietary issues.

References

- [1] B. Verlinden, J. Driver, I. Samajdar, R. D. Doherty, Thermo-mechanical processing of metallic materials, Elsevier, 2007.
- [2] W. Sun, Q. Wang, Y. Zhou, J. Wu, Material and energy flows of the iron and steel industry : Status quo, challenges and perspectives, *Applied Energy* 268 (2020) 114946.
- [3] A. Choudhury, The role of machine learning algorithms in materials science : A state of art review on industry 4.0, *Archives of Computational Methods in Engineering* 28 (5) (2021) 3361–3381.

- [4] C. H. Chan, M. Sun, B. Huang, Application of machine learning for advanced material prediction and design, *EcoMat* 4 (4) (2022) e12194.
- [5] J. E. Saal, A. O. Oliynyk, B. Meredig, Machine learning in materials discovery : confirmed predictions and their underlying approaches, *Annual Review of Materials Research* 50 (2020) 49–69.
- [6] X. Liu, P. Xu, J. Zhao, W. Lu, M. Li, G. Wang, Material machine learning for alloys : Applications, challenges and perspectives, *Journal of Alloys and Compounds* 921 (2022) 165984.
- [7] Q. Xie, M. Suvarna, J. Li, X. Zhu, J. Cai, X. Wang, Online prediction of mechanical properties of hot rolled steel plate using machine learning, *Materials & Design* 197 (2021) 109201.
- [8] P. Saravanakumar, V. Jothimani, L. Sureshbabu, S. Ayyappan, D. Noorullah, P. Venkatakrishnan, Prediction of mechanical properties of low carbon steel in hot rolling process using artificial neural network model, *Procedia Engineering* 38 (2012) 3418–3425.
- [9] S. Sinha, D. Guye, X. Ma, K. Rehman, S. Yue, N. Armanfard, Neural network prediction of the effect of thermomechanical controlled processing on mechanical properties, *Machine Learning with Applications* (2024) 100531.

- [10] Z. Shi, L. Du, X. He, X. Gao, H. Wu, Y. Liu, H. Ma, X. Huo, X. Chen, Prediction model of yield strength of v-n steel hot-rolled plate based on machine learning algorithm, *JOM* 75 (5) (2023) 1750–1762.
- [11] R. Hwang, H. Jo, K. S. Kim, H. J. Hwang, Hybrid model of mathematical and neural network formulations for rolling force and temperature prediction in hot rolling processes, *Ieee Access* 8 (2020) 153123–153133.
- [12] W.-g. Li, C. Liu, N. Feng, X. Chen, X.-h. Liu, Friction estimation and roll force prediction during hot strip rolling, *Journal of Iron and Steel Research International* 23 (12) (2016) 1268–1276.
- [13] Q. Wang, L. Song, J. Zhao, H. Wang, L. Dong, X. Wang, Q. Yang, Application of the gradient boosting decision tree in the online prediction of rolling force in hot rolling, *The International Journal of Advanced Manufacturing Technology* 125 (1) (2023) 387–397.
- [14] X. Li, F. Luan, Y. Wu, A comparative assessment of six machine learning models for prediction of bending force in hot strip rolling process, *Metals* 10 (5) (2020) 685.
- [15] Z.-H. Wang, D.-Y. Gong, X. Li, G.-T. Li, D.-H. Zhang, Prediction of bending force in the hot strip rolling process using artificial neural network and genetic algorithm (ann-ga), *The International Journal of Advanced Manufacturing Technology* 93 (2017) 3325–3338.

- [16] C. Cui, G. Cao, X. Li, Z. Gao, X. Zhou, Z. Liu, The coupling machine learning for microstructural evolution and rolling force during hot strip rolling of steels, *Journal of Materials Processing Technology* 309 (2022) 117736.
- [17] C. Idzik, A. Krämer, G. Hirt, J. Lohmar, Coupling of an analytical rolling model and reinforcement learning to design pass schedules : towards properties controlled hot rolling, *Journal of Intelligent Manufacturing* (2023) 1–22.
- [18] Y. Ji, S. Liu, M. Zhou, Z. Zhao, X. Guo, L. Qi, A machine learning and genetic algorithm-based method for predicting width deviation of hot-rolled strip in steel production systems, *Information Sciences* 589 (2022) 360–375.
- [19] Y. Zhong, J. Wang, J. Xu, J. Rao, K. Dang, Data-driven width spread prediction model improvement and parameters optimization in hot strip rolling process, *Applied Intelligence* 53 (21) (2023) 25752–25770.
- [20] J. Deng, J. Sun, W. Peng, Y. Hu, D. Zhang, Application of neural networks for predicting hot-rolled strip crown, *Applied Soft Computing* 78 (2019) 119–131.
- [21] J. Sun, J. Deng, W. Peng, D. Zhang, Strip crown prediction in hot rolling process using random forest, *International Journal of Precision Engineering and Manufacturing* 22 (2021) 301–311.
- [22] S. Ghorai, A. Mukherjee, M. Gangadaran, P. K. Dutta, Automatic defect detection on

- hot-rolled flat steel products, *IEEE Transactions on Instrumentation and Measurement* 62 (3) (2012) 612–621.
- [23] S. Youkachen, M. Ruchanurucks, T. Phatrapomnant, H. Kaneko, Defect segmentation of hot-rolled steel strip surface by using convolutional auto-encoder and conventional image processing, in : 2019 10th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES), IEEE, 2019, pp. 1–5.
- [24] C. Cui, H. Wang, X. Gao, G. Cao, Z. Liu, Machine learning model for thickness evolution of oxide scale during hot strip rolling of steels, *Metallurgical and Materials Transactions A* 52 (2021) 4112–4124.
- [25] J. Wang, Y. Wang, Y. Chen, Inverse design of materials by machine learning, *Materials* 15 (5) (2022) 1811.
- [26] M. Hu, Q. Tan, R. Knibbe, M. Xu, B. Jiang, S. Wang, X. Li, M.-X. Zhang, Recent applications of machine learning in alloy design : A review, *Materials Science and Engineering : R : Reports* 155 (2023) 100746.
- [27] J. Lee, D. Park, M. Lee, H. Lee, K. Park, I. Lee, S. Ryu, Machine learning-based inverse design methods considering data characteristics and design space size in materials design and manufacturing : a review, *Materials Horizons* (2023).
- [28] Z. Zhu, Y. Liang, J. Zou, Modeling and composition design of low-alloy steel's

- mechanical properties based on neural networks and genetic algorithms, *Materials* 13 (23) (2020) 5316.
- [29] J. Xiao, J. Cao, C. Song, C. Lv, G. Liu, Y. Wang, The collapse deformation prediction model of wide 7075 al-alloy intermediate slabs based on particle swarm optimization and support vector regression during the hot rolling process, *Journal of Materials Engineering and Performance* 33 (2) (2024) 1034–1050.
- [30] S. Xu, X. Xu, W. Jia, W. Liu, J. Li, D. Li, Microstructure-property mapping modeling for az31 alloy rolling deformation using improved pso-bp neural network, *Journal of Materials Research and Technology* 25 (2023) 2127–2139.
- [31] S. Wu, G. Cao, X. Zhou, N. Shi, Z. Liu, High dimensional data-driven optimal design for hot strip rolling of c–mn steels, *Isij International* 57 (7) (2017) 1213–1220.
- [32] J.-W. Lee, C. Park, B. Do Lee, J. Park, N. H. Goo, K.-S. Sohn, A machine-learning-based alloy design platform that enables both forward and inverse predictions for thermo-mechanically controlled processed (tmcp) steel alloys, *Scientific Reports* 11 (1) (2021) 11012.
- [33] I. Mohanty, D. Bhattacharjee, S. Datta, Designing cold rolled if steel sheets with optimized tensile properties using ann and ga, *Computational materials science* 50 (8) (2011) 2331–2337.

- [34] P. Mittal, A. Malik, I. Mohanty, K. Mitra, Comparative study of multi/many-objective evolutionary algorithms on hot rolling application, *Optimization in Industry : Present Practices and Future Scopes* (2019) 331–349.
- [35] P. Mittal, I. Mohanty, A. Malik, K. Mitra, Many-objective optimization of hot-rolling process of steel : A hybrid approach, *Materials and Manufacturing Processes* 35 (6) (2020) 668–676.
- [36] Z. C. Lipton, The mythos of model interpretability : In machine learning, the concept of interpretability is both important and slippery., *Queue* 16 (3) (2018) 31–57.
- [37] P. Barceló, M. Monet, J. Pérez, B. Subercaseaux, Model interpretability through the lens of computational complexity, *Advances in neural information processing systems* 33 (2020) 15487–15498.
- [38] X. Zhong, B. Gallagher, S. Liu, B. Kailkhura, A. Hiszpanski, T. Y.-J. Han, Explainable machine learning in materials science, *npj computational materials* 8 (1) (2022) 204.
- [39] G. Pilania, Machine learning in materials science : From explainable predictions to autonomous design, *Computational Materials Science* 193 (2021) 110360.
- [40] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, *Advances in neural information processing systems* 30 (2017).
- [41] J. Jeon, N. Seo, S. B. Son, S.-J. Lee, M. Jung, Application of machine learning algorithms

- and shape for prediction and feature analysis of tempered martensite hardness in low-alloy steels, *Metals* 11 (8) (2021) 1159.
- [42] F. Yan, K. Song, Y. Liu, S. Chen, J. Chen, Predictions and mechanism analyses of the fatigue strength of steel based on machine learning, *Journal of Materials Science* 55 (2020) 15334–15349.
- [43] J. Jeon, N. Seo, J.-G. Jung, H.-S. Kim, S. B. Son, S.-J. Lee, Prediction and mechanism explain of austenite-grain growth during reheating of alloy steel using xai, *Journal of Materials Research and Technology* 21 (2022) 1408–1418.
- [44] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, Machine learning in materials informatics : recent applications and prospects, *npj Computational Materials* 3 (1) (2017) 54.
- [45] C. Frie, A. Kolyshkin, C. Eberl, Analysis of data-driven models for predicting fatigue strength of steel components with uncertainty quantification, *Fatigue & Fracture of Engineering Materials & Structures* 47 (3) (2024) 1036–1052.
- [46] L. Li, J. Chang, A. Vakanski, Y. Wang, T. Yao, M. Xian, Uncertainty quantification in multivariable regression for material property prediction with bayesian neural networks, *Scientific Reports* 14 (1) (2024) 10543.

- [47] G. He, Y. Zhao, C. Yan, Uncertainty quantification in multiaxial fatigue life prediction using bayesian neural networks, *Engineering Fracture Mechanics* (2024) 109961.
- [48] B. Cui, H. Wang, Pipeline corrosion prediction and uncertainty analysis with an ensemble bayesian neural network approach, *Process Safety and Environmental Protection* (2024).
- [49] V. Nemani, L. Biggio, X. Huan, Z. Hu, O. Fink, A. Tran, Y. Wang, X. Zhang, C. Hu, Uncertainty quantification in machine learning for engineering design and health prognostics : A tutorial, *Mechanical Systems and Signal Processing* 205 (2023) 110796.
- [50] R. F. Barber, E. J. Candes, A. Ramdas, R. J. Tibshirani, Predictive inference with the jackknife+ (2021).
- [51] T. Cordier, V. Blot, L. Lacombe, T. Morzadec, A. Capitaine, N. Brunel, Flexible and systematic uncertainty estimation with conformal prediction via the mapie library, in : *Conformal and Probabilistic Prediction with Applications*, PMLR, 2023, pp. 549–581.
- [52] F. Meng, T. Song, Uncertainty forecasting system for tropical cyclone tracks based on conformal prediction, *Expert Systems with Applications* 249 (2024) 123743.
- [53] B. Ibrahim, A. Tetteh-Asare, I. Ahenkorah, A transparent and valid framework for rockburst assessment : Unifying interpretable machine learning and conformal prediction, *Rock Mechanics and Rock Engineering* (2024) 1–15.

- [54] N. Tamascelli, G. E. Scarponi, M. T. Amin, Z. Sajid, N. Paltrinieri, F. Khan, V. Cozzani, A neural network approach to predict the time-to-failure of atmospheric tanks exposed to external fire, *Reliability Engineering & System Safety* 245 (2024) 109974.
- [55] T.-T. Nguyen, V.-H. Dang, D.-M. Hoang, X.-D. Pham, T.-H. Nguyen, V.-T. Dinh, Robust active learning framework for structural reliability analysis using uncertainty quantification and flexible meta-model, in : *Structures*, Vol. 63, Elsevier, 2024, p. 106465.
- [56] C. Klinkenberg, B. Kintscher, K. Hoen, M. Reifferscheid, More than 25 years of experience in thin slab casting and rolling current state of the art and future developments, *steel research international* 88 (10) (2017) 1700272.
- [57] T. Chen, C. Guestrin, Xgboost : A scalable tree boosting system, in : *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
- [58] Z. Shi, L. Du, X. He, X. Gao, H. Wu, Y. Liu, H. Ma, X. Huo, X. Chen, Prediction model of yield strength of v–n steel hot-rolled plate based on machine learning algorithm, *JOM* 75 (5) (2023) 1750–1762.
- [59] J. Chen, F. Zhao, Y. Sun, L. Zhang, Y. Yin, Prediction model based on xgboost for mechanical properties of steel materials, *International Journal of Modelling, Identification and Control* 33 (4) (2019) 322–330.

- [60] Z. H. Wang, Y. F. Liu, T. Wang, J. G. Wang, Y. M. Liu, Q. X. Huang, Intelligent prediction model of mechanical properties of ultrathin niobium strips based on xgboost ensemble learning algorithm, *Computational Materials Science* 231 (2024) 112579.
- [61] S. M. Lundberg, G. G. Erion, S.-I. Lee, Consistent individualized feature attribution for tree ensembles, *arXiv preprint arXiv :1802.03888* (2018).
- [62] N. Gunantara, A review of multi-objective optimization : Methods and its applications, *Cogent Engineering* 5 (1) (2018) 1502242.
- [63] E. Zitzler, M. Laumanns, S. Bleuler, A tutorial on evolutionary multiobjective optimization, *Metaheuristics for multiobjective optimisation* (2004) 3–37.
- [64] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm : Nsga-ii, *IEEE transactions on evolutionary computation* 6 (2) (2002) 182–197.
- [65] K. Deb, S. Agrawal, A. Pratap, T. Meyarivan, A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization : Nsga-ii, in : *Parallel Problem Solving from Nature PPSN VI : 6th International Conference Paris, France, September 18–20, 2000 Proceedings* 6, Springer, 2000, pp. 849–858.
- [66] L. Van der Maaten, G. Hinton, Visualizing data using t-sne., *Journal of machine learning research* 9 (11) (2008).

- [67] M. Ahmed, R. Seraj, S. M. S. Islam, The k-means algorithm : A comprehensive survey and performance evaluation, *Electronics* 9 (8) (2020) 1295.
- [68] K. P. Sinaga, M.-S. Yang, Unsupervised k-means clustering algorithm, *IEEE access* 8 (2020) 80716–80727.
- [69] K. R. Shahapure, C. Nicholas, Cluster quality analysis using silhouette score, in : 2020 IEEE 7th international conference on data science and advanced analytics (DSAA), IEEE, 2020, pp. 747–748.
- [70] M. Syakur, B. K. Khotimah, E. Rochman, B. D. Satoto, Integration k-means clustering method and elbow method for identification of the best customer profile cluster, in : IOP conference series : materials science and engineering, Vol. 336, IOP Publishing, 2018, p. 012017.
- [71] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, Lightgbm : A highly efficient gradient boosting decision tree, *Advances in neural information processing systems* 30 (2017).
- [72] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn : Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.

- [73] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna : A next-generation hyperparameter optimization framework, in : Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2019, pp. 2623–2631.
- [74] V. Taquet, V. Blot, T. Morzadec, L. Lacombe, N. Brunel, Mapie : an open-source library for distribution-free uncertainty quantification, arXiv preprint arXiv :2207.12274 (2022).
- [75] J. Blank, K. Deb, Pymoo : Multi-objective optimization in python, Ieee access 8 (2020) 89497–89509.
- [76] D. H. Wolpert, The supervised learning no-free-lunch theorems, Soft computing and industry : Recent applications (2002) 25–42.
- [77] T. F. Sterkenburg, P. D. Grünwald, The no-free-lunch theorems of supervised learning, Synthese 199 (3) (2021) 9979–10015.
- [78] S. Uddin, H. Lu, Confirming the statistically significant superiority of tree-based machine learning algorithms over their counterparts for tabular data, Plos one 19 (4) (2024) e0301541.
- [79] L. Grinsztajn, E. Oyallon, G. Varoquaux, Why do tree-based models still outperform deep learning on typical tabular data?, Advances in neural information processing systems 35 (2022) 507–520.

- [80] D. Kobak, P. Berens, The art of using t-sne for single-cell transcriptomics, *Nature communications* 10 (1) (2019) 5416.

Chapter 5

Conclusion

This thesis developed data-driven frameworks to predict and optimize the mechanical properties of microalloyed steels. The general conclusion are presented in three sections below :

5.1 Global Discussion and Conclusions

(I) In Chapter 3, deep neural networks were utilized to model two mechanical properties—Ultimate Tensile Strength (UTS) and Lower Yield Strength (LYS)—for Nb-microalloyed steel. The models achieved R^2 accuracies of 0.88 and 0.86, respectively. It was noticed that certain instances proved more difficult to predict than others. The lack of data during casting and coiling in the mill could explain the observation.

Accordingly, this data was included in Chapter 4, where V-microalloyed steel was modeled. Though no direct quantitative comparisons can be made on separate datasets, R^2 values for UTS were 0.97 and 0.95 using XGBoost and Neural Network models respectively—a significant enhancement over the previous result of 0.88.

- (II) From chapter 4, another key observation is the lower R^2 scores for %Elongation models, ranging from 0.83 to 0.86 across all the seven evaluated ML models. The pseudo-continuous nature of the target i.e., % Elongation data, posed challenges for ML algorithms designed for continuous data prediction.
- (III) The combined effect of TMCP parameters, as analyzed through SHAP values and SIMS Mean Flow Stress (MFS) plots, indicates that thicker gauges have limited opportunities for grain refinement, thus receiving less strength contribution from TMCP parameters. However, the contribution of chemical composition is significantly noted, especially in thicker gauges, where richer alloying compensates for strength losses attributed to TMCP limitations. A correlation with the no-recrystallization temperature provides insights into the rolling strategy of the studied microalloyed steels.
- (IV) For V-based microalloyed steel, global feature importance identified Vanadium, Silicon, and Manganese as the top influencing elements for both UTS and Elongation, each having opposing effects on strength and ductility. This finding supports our strategy of optimizing mechanical properties through alloy design while maintaining the original processing schedules.
- (V) In chapter 4, a systematic approach is proposed for using ML models in conjunction with evolutionary genetic algorithms for alloy design in thermomechanical processing. By setting TMCP parameters as strict bounds in the decision variable space, the search

for alternative compositions that yield improved strength-ductility combinations can be streamlined.

- (VI) A novel approach utilizing manifold learning and unsupervised clustering facilitates post-Pareto analysis, which is otherwise challenging with high-dimensional industrial datasets. This method allows the selection of a few unique representative solutions for further analysis using thermodynamic calculations and experimental trials. An example strategy to select solutions could involve choosing cost-effective alloys with lower levels of expensive alloying elements such as Vanadium.

5.2 Proposed Applications

- (I) The ML models, enhanced with explainability tools, provide a deeper understanding of physical metallurgy processes and aid in process control.
- (II) These models enable reliable and explainable online predictions of the mechanical properties of as-rolled steels, potentially reducing the need for extensive manual testing and hence, enhanced operational efficiency.
- (III) The developed approach facilitates the discovery of new alloys within the existing compositional ranges with superior strength and ductility combinations, or allowing for the selection of leaner compositions to achieve current strength and ductility combinations.

5.3 Future Work

Building on the current work, the following are four possible areas of future work :

- (I) It would add confidence to the alloy design methodology if the finally chosen alloy solutions are validated based on thermodynamic calculations for critical temperatures like no-recrystallization temperature (T_{nr}) and phase transformation temperature (Ar_3).
- (II) It would be interesting to compare other UQ approaches like probabilistic models, ensemble methods etc. with conformal predictions used in this work. For online application it is required that the proposed methodology should be robust to out-of-distribution (OOD) data as for commercial industries mass produce a limit number of unique product types. The OOD samples can be highlighted as high-risk specimen for manual testing.
- (III) A Bayesian approach may be explored for multi-objective optimization as well. Among many advantages, the ability to naturally quantify uncertainty of the optimized solutions would be most important for high risk associated situations like in industries.
- (IV) The primary emphasis of our work has been on optimizing the chemical composition to achieve improvement. To take a more comprehensive approach, we may consider both composition and the process schedule. Given that process schedules are determined

by the flow of mass and energy in continuous processes, a different approach would be needed. To address this, one possible solution could be to integrate hierarchical modeling, enabling sampling based on physical principles.

To conclude, this thesis presents a strong case for the incorporation of ML-based tools in hot rolling and other TMCP industries. Additionally, this work emphasizes the need for reliable and trustworthy ML frameworks, crucial for wider adoption and successful implementation in industries.