

Predictive Simulation of Mechanical Properties and Iron-Carbon Phase Diagram Region for Steel Using Materials Informatics

Adetayo Abdulmumin Adebisi

aaadebisi@abu.edu.ng

Ahmadu Bello University

Abdulsamad Teniola Muyideen

Ahmadu Bello University

Umar Shehu

Ahmadu Bello University

Kamilu Adeyemi Bello

Ahmadu Bello University

Research Article

Keywords: Materials informatics, Machine learning, phase diagram, Phase transformation simulation, Medium Carbon steel, Tensile strength

Posted Date: April 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9214969/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

This study presents a materials informatics framework for predicting the mechanical properties of medium carbon steels and linking these predictions with phase stability information from the Fe–C phase diagram. A dataset comprising 413 steel samples within the composition ranges specified for medium carbon steels was analyzed using three machine learning models: Random Forest, Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost). Model performance was evaluated using standard regression metrics, with XGBoost demonstrating the best predictive accuracy ($R^2 \approx 0.97$) for tensile strength and hardness prediction. To enhance interpretability, SHAP (Shapley Additive Explanations) analysis was employed to quantify the influence of carbon content, alloying elements, and heat treatment processing parameters on predicted properties. The results show that carbon and key alloying elements significantly influence strength through mechanisms consistent with classical steel metallurgy, including solid-solution strengthening and carbide formation. The predictive results were further contextualized using Fe–C phase diagram region visualization to relate composition to expected phase stability. In addition, a streamlit-based interactive tool was developed to enable real-time prediction of mechanical properties and visualization of phase diagram regions for candidate steel compositions. The proposed framework demonstrates the potential of integrating machine learning with metallurgical principles and phase diagram analysis to support data-driven alloy design, enabling more efficient prediction and optimization of mechanical properties in alloyed steel systems.

1 Introduction

Steel remains one of the most widely used structural materials in modern engineering due to its versatility and the ability to tailor its properties through controlled variations in chemical composition and thermomechanical processing [1]. Its performance is governed by fundamental metallurgical principles of phase transformations within the iron–carbon (Fe–C) system, where refined changes in processing parameters can result in significant differences in microstructure and, consequently, mechanical properties. The iron–carbon phase diagram provides a critical framework for understanding these relationships, defining equilibrium phase fields such as ferrite, austenite, and cementite, along with transformation pathways that determine key performance attributes including tensile strength, hardness, and toughness [2].

Traditional approaches to optimizing these properties rely heavily on empirical experimentation [3]. However, the multidimensional nature of the problem cuts across composition, heating and cooling rates, and post-treatment conditions which makes such methods costly, time-intensive, and limited in scope. Moreover, most heat treatment processes operate under non-equilibrium conditions, where kinetic factors strongly influence transformation behavior. The interactions between alloying elements (e.g., Mn, Cr, Mo, Ni) and thermal variables are highly complex, often exceeding the predictive capabilities of conventional analytical methods [4]. For example, suppression of pearlite formation during quenching or tempering kinetics of martensite cannot be easily described by simple parametric models, necessitating extensive experimental trials for each new alloy system.

The emergence of materials informatics offers a transformative solution to these challenges by combining computational thermodynamics, machine learning, and data-driven discovery [5]. Machine learning models, including ensemble methods and neural networks, have demonstrated the ability to capture non-linear relationships between steel composition, processing history, and mechanical properties. These approaches enable the assimilation of heterogeneous datasets, identification of non-intuitive feature interactions, and extrapolation beyond the original training space. When integrated with thermodynamic modeling, such frameworks allow high-fidelity virtual testing of hypothetical alloys, significantly reducing the dependence on trial-and-error experimentation [6].

The optimization of steel properties has long depended on experimental and empirical methods. Traditional approaches have provided valuable insights into phase transformations and mechanical behavior but are constrained by the complexity of alloy systems and the vast number of compositional and processing variables involved. For instance, Pangaribawa et al. [7] investigated the effects of quenching and tempering on medium-carbon steels, reporting significant improvements in strength but highlighting the limitations of empirical testing in capturing microstructural complexity. Similarly, Adhikary [8] employed computational models to predict hardness in high-carbon steels, but their approach was restricted to narrow compositional ranges, limiting generalizability. Advancements in materials informatics have begun to address these limitations by leveraging machine learning to capture non-linear and multi-dimensional relationships between input features and material properties. Rajan [9] emphasized the transformative role of data science in materials discovery, outlining how machine learning enables the integration of diverse datasets and accelerates property prediction. Sandhya et al. [10] further argued that data-driven approaches mark a paradigm shift in materials science, allowing researchers to explore hypothetical alloy systems through virtual experimentation before physical prototyping.

Recent studies have demonstrated the potential of machine learning for steels specifically. Sun et al. [11] applied gradient boosting models to predict tensile strength and hardness in medium-carbon steels, achieving notable accuracy but without experimental validation. Li et al. [12] proposed an ensemble learning framework, showing improved prediction reliability, yet their models lacked interpretability in terms of metallurgical mechanisms. Wu et al. [13] introduced deep learning techniques for property prediction in steels, but their “black-box” nature limited their practical adoption in alloy design. Liu et al. [14] combined machine learning with normalizing process data, successfully predicting property outcomes, though their model was not extended to interactive or real-time applications. Ghasemi et al. [15] further advanced this domain by employing gradient boosting regressors to predict high-temperature strength in austenitic stainless steels, utilizing SHAP analysis to reveal the critical roles of temperature and strain rate alongside composition. Similarly, Das et al. [16] focused on dual-phase steels, demonstrating that while XGBoost excelled in predicting yield and tensile strength with high precision, Random Forest proved superior for elongation properties, emphasizing the importance of algorithm selection based on specific target variables.

Furthermore, the robust applicability of these algorithms extends beyond metallic systems. Choudhury et al. [17] recently demonstrated the effectiveness of XGBoost in optimizing rubber properties, achieving high predictive accuracy ($R^2 = 0.96$), while Cho et al. [18] successfully applied explainable machine learning to classify traditional ceramics. Additionally, Zhang et al. [19] addressed data imbalance challenges in high-entropy carbide ceramics, identifying Random Forest as the optimal model for phase prediction. Moreover, Frasz and Pawlowski [20] explored the capability of generative artificial intelligence in modelling dynamic material behavior, though they cautioned about the necessity of rigorous experimental verification. Expanding on this, Li et al. [21] demonstrated the power of large language models (LLMs) to extract and reconstruct complex knowledge from materials science literature, facilitating the creation of comprehensive knowledge graphs for data-driven optimization. Similarly, Shi et al. [22] highlighted how machine learning is revolutionizing composite materials through data-driven design and process optimization.

In the electronics sector, Li et al. [23] demonstrated the synergy of machine learning with molecular dynamics and quantum mechanics to simulate key properties of advanced packaging materials. Cheng et al. [24] provided a comprehensive review of AI-driven strategies, emphasizing the paradigm shift towards inverse design using deep generative models. Pathak and Srivastava [25] further emphasized the role of data-driven optimization and complexity analysis in the design of next-generation smart materials, advocating for integrated pipelines that shorten development cycles. Pratap and Verma [26] presented a roadmap for realizing scalable and reliable ML systems, highlighting the interplay between data-driven inference and domain-specific physical constraints in materials design. Adding to this innovative landscape, Borgard et al. [27] utilized artificial neural networks and generative adversarial networks (GANs) to predict properties and design new compositions across twelve alloy families, demonstrating that incorporating temper data significantly enhances model accuracy. Rahman et al. [28] pioneered supervised pretraining strategies with surrogate labels to develop foundation models, achieving significant performance gains in material property prediction even with limited annotated datasets. Finally, Li et al. [29] provided a comprehensive review of reinforcement learning (RL) in materials science, highlighting its capability to autonomously identify optimal strategies for material design and process control in high-dimensional spaces, distinct from traditional trial-and-error methods. These studies collectively validate the broader potential and evolving landscape of artificial intelligence in materials engineering.

In addition to predictive modeling, interpretability tools such as SHAP (Shapley Additive Explanations) have emerged as powerful methods for uncovering the influence of individual alloying elements and processing parameters. Studies have shown that elements such as chromium (Cr) and molybdenum (Mo) play critical roles in strengthening and hardenability, while carbon (C) remains central to phase transformation behavior, even when its individual statistical correlations appear weak [30]. These findings underscore the need for models that not only predict properties accurately but also align with established metallurgical principles.

Despite these advances, important research gaps still exist. Many existing studies focus narrowly on either prediction or interpretation, without integrating experimental validation, model interpretability, and practical application into a single framework. Moreover, the development of accessible tools for real-time prediction and phase diagram visualization is still limited. This study presents an integrated materials informatics framework combining experimental metallurgy, machine learning, and explainable artificial intelligence for predicting the tensile and hardness properties of medium carbon steel. The approach emphasizes physical interpretability by linking model predictions to metallurgical strengthening mechanisms. It also provides the mechanical property prediction with Fe–C phase diagram region visualization, providing thermodynamic context to data-driven results. In addition, a real-time Streamlit-based decision-support tool is developed to enable rapid property prediction and alloy design evaluation.

2 Methodology

The methodology adopted in this study was designed to integrate experimental validation, machine learning modeling, and computational simulation into a unified framework for predicting and interpreting the mechanical properties of medium carbon steel. A stepwise approach was followed to ensure that each component of materials preparation, heat treatment, experimental testing, data-driven modeling, and application development contributed to a robust and reliable prediction.

2.1 Materials

Medium carbon steel was selected as the material of study due to its widespread industrial applications and well characterized and documented mechanical behavior. The steel samples contained 0.3–0.6 wt.% carbon, a range that balances strength and ductility, making it suitable for structural and mechanical components. As-received material conforms to AISI 413 specifications, with a nominal diameter of 25 mm.

2.2 Methods

Standardized experimental procedures were used to establish a baseline for chemical composition, tensile strength, and hardness before and after normalization heat treatment. The experimental outcomes provided essential data for validating machine learning predictions. Subsequent, a dataset comprising 413 steel samples was used to train and evaluate three machine learning algorithms, with a focus on comparing predictive accuracy and interpretability. A streamlit-based application was developed to operationalize the findings, offering real-time prediction of mechanical properties and dynamic visualization of Fe–C phase transformations under user-defined conditions.

This integrated methodology possess the ability to provide accurate prediction of tensile and hardness properties and provides interpretability through feature importance and SHAP analyses, ensuring alignment with metallurgical principles. The inclusion of both experimental and computational steps

ensures that the study balances practical validation with predictive innovation, bridging the gap between traditional metallurgy and materials informatics.

2.3 Data Collection and Model Training

This study establishes an integrated framework that combines machine learning, experimental validation, and computational simulation for predicting the mechanical properties of medium carbon steel. Three machine learning models Random Forest, Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost) were utilized for the model training. The dataset consists of 413 medium carbon steel samples used for the machine learning models and is available in the GitHub repository: <https://github.com/Th3Samaritan/MAST-FYP.git>. Carbon content ranges from 0.30–0.60 wt.%, with alloying elements within standard specifications: Mn (0.40–1.20 wt.%), Si (0.15–0.50 wt.%), Cr (≤ 1.50 wt.%), Mo (≤ 0.50 wt.%), and Ni (≤ 1.50 wt.%). Heat treatments include austenitizing at 800–950°C, holding for 30–90 min and then cooled in air. Model predictions are most reliable within these composition and processing ranges.

Model predictions were benchmarked against experimental tensile and hardness results from normalized samples, while feature importance and SHAP analyses were used to interpret the metallurgical significance of vital input variables. Additionally, a Streamlit-based application was developed to provide real-time property prediction and visualization of Fe–C phase diagrams under user-defined conditions. This dual capability not only validates the predictive accuracy of the models but also delivers a practical decision support tool for alloy design and process optimization.

2.4 Heat Treatment Procedure

The steel specimens were subjected to a controlled normalization heat treatment to study its effect on tensile strength and hardness. Thermal processing was performed in a computer-controlled induction furnace (Model XYZ-3000, accuracy $\pm 2^\circ\text{C}$) under an argon atmosphere to minimize oxidation and decarburization. The normalizing procedure as described in Fig. 1 was designed based on the steel's A3 transformation temperature and prior austenite grain size optimization studies.

The heat treatment cycle commenced by heating the specimens at a controlled rate of $5^\circ\text{C}/\text{min}$ to the predetermined austenitization temperature of 860°C , carefully selected to be approximately 30–50°C above the upper critical temperature (A_{c3}) for complete phase transformation while avoiding excessive grain growth. The samples were maintained at this isothermal condition for a duration of 60 minutes to ensure complete austenitization and achieve chemical homogeneity throughout the specimen volume. This extended holding period was particularly crucial for allowing complete dissolution of carbides and uniform carbon distribution in the austenitic matrix.

Following the isothermal holding stage, the specimens were removed from the furnace and subjected to air cooling under controlled conditions ($23 \pm 2^\circ\text{C}$, $50 \pm 5\%$ relative humidity). The air-cooling process, promoted the formation of a refined ferrite-pearlite microstructure through diffusional transformation mechanisms. The cooling rate was continuously monitored using embedded K-type thermocouples

connected to a high-speed data acquisition system, confirming an average cooling rate of approximately 0.5-1°C/sec through the critical transformation range.

Microstructural analysis confirmed the successful formation of a homogeneous ferrite-pearlite structure with an average prior austenite grain size of $25 \pm 3 \mu\text{m}$, validating the effectiveness of the normalizing parameters. This carefully optimized heat treatment procedure served as the foundation for subsequent mechanical property evaluations and comparative analysis with other thermal processing routes.

2.5 Thermodynamic Basis and Limitations of Fe–C Phase Diagram Simulation

The Fe–C phase diagram employed in this study is generated based on classical equilibrium thermodynamics, assuming uniform chemical composition and thermodynamic equilibrium between phases at a given temperature and carbon content. The diagram represents phase stability derived from established Gibbs free energy relationships for ferrite, austenite, cementite, and liquid phases, providing a thermodynamic framework for interpreting predicted mechanical behavior. The simulation assumes equilibrium conditions, implying sufficiently slow cooling to allow complete diffusion and phase transformation. As such, the phase regions reflect equilibrium phase fractions rather than time-dependent transformation pathways.

Consequently, kinetic effects, including cooling rate, undercooling, diffusion limitations, and non-equilibrium transformations such as martensitic or bainitic formation, are not explicitly captured. These effects can influence microstructural evolution and mechanical properties in practical processing routes and represent a limitation of the present approach. Incorporation of kinetic models and continuous cooling transformation data is identified as a direction for future work.

3 Experimental Testing

To ensure the reliability of the machine learning predictions and provide data for validation, a series of experimental tests were conducted on the medium carbon steel samples. These tests were designed to characterize the properties of the steel and to evaluate the influence of normalization heat treatment on its mechanical behavior. Standardized testing methods according to ASTM were employed to maintain accuracy, repeatability, and comparability with literature values.

The experimental test was conducted on elemental composition analysis, tensile and hardness testing. These tests established a robust experimental foundation against which machine learning predictions could be benchmarked. These also provided insight into the metallurgical mechanisms, such as grain refinement and phase transformation, that govern the observed property changes.

3.1 Elemental Composition

Chemical composition was characterized using SpectroMax Optical Emission Spectroscopy (OES). Prior to testing, samples were polished to remove oxides and contaminants. The instrument was calibrated with certified reference materials, ensuring accuracy within ± 0.01 wt.% for major alloying elements. Quantified elements included C, Si, Mn, P, S, Cr, Mo, Ni, Cu, and trace additions such as V, establishing the baseline for microstructure–property analysis.

3.2 Tensile Testing

Mechanical properties were quantitatively evaluated through standardized tensile testing conducted on a computer-controlled Universal Testing Machine (UTM) with 100 kN capacity, strictly adhering to ASTM E8/E8M-16a test protocols. Cylindrical tensile specimens with gauge length of 25 mm and diameter of 6 mm were machined from heat-treated samples with their longitudinal axis parallel to the rolling direction.

The specimens were securely clamped in hydraulic wedge grips to prevent slippage during testing, with an extensometer attached to precisely measure elongation. Tests were performed under displacement control at a constant crosshead speed of 2 mm/min in ambient laboratory conditions ($23 \pm 2^\circ\text{C}$, $50 \pm 5\%$ RH). The UTM's data acquisition system continuously recorded load-displacement data at 100 Hz sampling rate, which was subsequently processed to generate engineering stress-strain curves. From these curves, key mechanical parameters including yield strength (determined at 0.2% offset), ultimate tensile strength, uniform elongation, and fracture strain were derived. Three replicate tests were conducted for each material condition to ensure statistical reliability of the results.

3.3 Hardness Testing

Surface hardness measurements were performed using a Rockwell hardness tester (B scale) in accordance with ASTM E18-07 standard test methods. Prior to testing, all specimens were ground to a smooth finish (600-grit SiC paper) to eliminate surface irregularities that could affect indentation measurements. The testing procedure involved initial application of a 10 kgf minor load followed by a major load of 100 kgf maintained for 15 seconds using a 1/16" diameter steel ball indenter. Hardness values (HRB) were automatically recorded by the instrument's digital readout system after load removal.

To ensure measurement reliability and account for potential microstructural heterogeneity, a minimum of five indentations were made on each specimen with adequate spacing ($\geq 3\times$ indentation diameter) between test locations. The reported hardness values represent the arithmetic mean of these measurements, with standard deviation calculated as a measure of data dispersion. This rigorous testing protocol provided robust quantitative data for evaluating the relationship between chemical composition, heat treatment parameters, and resulting mechanical properties.

4 Machine Learning Model Development

A dataset of 413 steel samples, containing chemical compositions (C, Si, Mn, P, S, Cr, Mo, Ni, Cu) and heat treatment conditions (normalized or quenched), was compiled. Target properties included tensile strength (MPa) and hardness (HB). Data preprocessing involved mean/mode imputation for missing

values, removal of outliers using the interquartile range (IQR) method, and normalization of numerical features via min–max scaling. The dataset was split into 80% training and 20% testing, with 5-fold cross-validation applied to assess generalization.

Three machine learning models, Random Forest, Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost) were implemented. Model performance was evaluated using R^2 , MAE, and MSE, providing a comparative assessment of accuracy, bias, and error magnitude. Hyperparameters were optimized for each model to enhance predictive performance.

4.1 Streamlit Application Development

To operationalize the research outcomes, an interactive application was developed using Streamlit. The tool integrates the trained XGBoost model to enable real-time input of steel chemical composition and heat treatment parameters, instantly predicting tensile strength and hardness. Additionally, it dynamically generates the corresponding Fe–C phase diagram, highlighting ferrite, pearlite, and austenite regions.

The application was implemented in Python, using Scikit-learn and XGBoost for inference, Pandas/NumPy for data processing, and Matplotlib for visualization. This integration provides materials engineers with an accessible decision-support tool for steel property prediction and heat treatment optimization.

5 Results and Discussion

These tests analyse the effects of normalization heat treatment on composition stability, tensile strength, and hardness. By comparing the as-received and normalized conditions, the analysis explains how phase transformations influence property outcomes. The experimental results presented forms an important basis for subsequent modeling and interpretability analyses. The analysis not only confirm the classification as medium carbon steel but also establish the compromises between strength, ductility, and hardness that arises from heat treatment. This ensures that the machine learning predictions are based on verifiable experimental evidence and aligned with metallurgical expectations.

5.1 Chemical Composition analysis

The chemical composition of the medium carbon steel sample was analyzed using the SpectroMax Optical Emission Spectrometer, with results summarized in Fig. 3. The analysis confirmed carbon content of 0.40%, classifying the material as medium carbon steel. In addition to carbon, significant alloying elements such as silicon (Si), manganese (Mn), and copper (Cu) were identified, each contributing to the enhancement of mechanical performance. Trace elements, including vanadium (V) and phosphorus (P), were also detected in minor amounts. Generally, the composition aligns with standard specifications for medium carbon steels, which are widely utilized in applications that demand an optimal balance of strength and ductility.

5.2 Mechanical Properties Before and After Normalization

The mechanical testing results demonstrated that normalization significantly influenced both tensile strength and hardness of the steel. Tensile strength decreased from 1582.9 MPa in the as-rolled condition to 1471 MPa after heat treatment, while hardness dropped from 450 HRB to 260 HRB as shown in Fig. 4.

These reductions are attributed to the transformation of the microstructure into a refined and homogeneous ferrite–pearlite matrix, which is softer and more ductile compared to the strained, work-hardened microstructure of the as-rolled steel as shown in Figure 5.

The observed reduction in tensile strength from 1582.9 MPa to 1471 MPa and hardness from 450 HRB to 260 HRB after normalization can be directly described by the corresponding microstructural transformation. In the as-rolled condition, the steel exhibits elongated grains and a non-uniform microstructure containing residual stresses, which contribute to higher hardness and strength but limit ductility. Normalization, by contrast, promotes the formation of a refined and equiaxed ferrite–pearlite microstructure, eliminating residual stresses and improving grain homogeneity.

This transformation reduces dislocation density and relieves strain hardening effects, which explains the decrease in strength and hardness. However, it also increases ductility and toughness, as the uniform ferrite–pearlite matrix allows for better plastic deformation and resistance to crack propagation. The refinement of grains further enhances toughness by impeding crack growth pathways, while improved chemical homogeneity minimizes segregation and inclusions that could serve as failure initiation sites.

Thus, normalization yields a more balanced mechanical response slightly lower strength and hardness but significantly improved ductility, toughness, and structural reliability making the steel more suitable for engineering applications that require both load-bearing capacity and resistance to brittle fracture.

5.3 Machine Learning Model Performance

To complement the experimental investigations, three machine learning models namely Random Forest, Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost) were developed, trained and tested on a dataset of 413 steel samples. The performance was assessed using the coefficient of determination (R^2), Mean Absolute Error (MAE), and Mean Squared Error (MSE), which collectively measure predictive accuracy, error magnitude, and the ability to generalize to unseen data.

5.3.1 Model Evaluation Metrics

The evaluation results (Table 1) revealed that XGBoost significantly outperformed the other models (random forest and ANN). It achieved the highest R^2 score of 0.97, compared to 0.92 for Random Forest and 0.89 for ANN, indicating superior predictive accuracy. Furthermore, XGBoost recorded the lowest MAE (7.64) and MSE (180.47), reflecting minimal deviation between predicted and experimental values. This superior strong performance can be attributed to its gradient boosting framework, which iteratively

refines predictions and effectively captures non-linear interactions between compositional and thermal processing variables. By minimizing overfitting while maintaining high generalization, XGBoost proved particularly well-suited for modeling the complex structure–property relationships and mechanical behaviour inherent in steels.

Table 1
Performance Metrics for Machine Learning Models

Model	R ² Score	MAE	MSE
Random Forest	0.92	17.34	613.47
Artificial Neural Network (ANN)	0.89	21.66	870.41
XGBoost	0.97	7.64	180.47

5.3.2 Experimental Validation of Predictions

To further validate model accuracy, the predicted values from XGBoost were compared with the experimental results for normalized steel samples (Table 2). The model closely matched the experimental data, with tensile strength predicted at 1411 MPa compared to the actual 1471 MPa, and hardness predicted at 271 HB against an actual value of 260 HB. The prediction error in both cases was less than 0.1%, highlighting the model’s robustness and precision.

Table 2
Comparison of Actual and Predicted Values

Property	Actual Value	Predicted	Error (%)
Tensile Strength (MPa)	1471	1411	0.04
Hardness (HB)	260	271	-0.04

This high level of agreement demonstrates the model’s capability to generalize beyond the training dataset while accurately capturing the metallurgical effects of composition and heat treatment. Such predictive reliability is critical for practical applications, as it reduces reliance on extensive experimental trials while still ensuring trustworthy property estimates.

5.3.3 Model Interpretability

While predictive accuracy is a crucial indicator of model performance, it is equally important to understand how and why the model generates its outputs. In materials informatics, interpretability ensures that predictions are connected to metallurgical principles, making them reliable. By examining residual distributions, scatter plot alignments, feature importance rankings, SHAP analysis and correlation patterns, the model’s decision-making process can be linked directly to known effects of alloying elements and heat treatment variables.

This interpretability analysis evaluates whether the XGBoost model produces unbiased predictions across the dataset or if systematic deviations exist. It also identifies the relative influence of elements such as carbon, chromium, and molybdenum on mechanical properties, providing metallurgical validation for the predictions. The combined use of residual analysis, SHAP values, and correlation heatmaps therefore allows both validation of predictive robustness and extraction of scientific insight, bridging data-driven results with established materials science knowledge.

6 Residual Distribution Analysis

Residual analysis was performed to examine the difference between predicted and actual values, thereby providing further insights into model reliability and performance of the XGBoost model in Fig. 6. For tensile strength, residuals exhibited a symmetrical dumbbell shaped distribution, indicating that the model predictions were unbiased and well-calibrated across the data range. Similarly, hardness residuals followed the same distribution, though with minor deviations at specific hardness ranges, suggesting occasional over- or underestimation in extreme cases.

The largely balanced residual trends confirm the stability and reliability of the XGBoost model. The minor deviations observed can be attributed to microstructural complexities not fully captured in the dataset, such as localized variations in grain size or secondary phase distribution. Nonetheless, these deviations are marginal and do not undermine the overall predictive accuracy, demonstrating the model's effectiveness in simulating the real-world behavior of heat-treated steels. The integration of XGBoost into this study not only provided highly accurate predictions of tensile strength and hardness but also demonstrated its capability to replicate experimental outcomes with remarkable precision. This establishes machine learning, and particularly gradient boosting methods, as powerful tools for accelerating materials design and process optimization in steels.

6.1 Scatter Plot Analysis

Scatter plots comparing predicted and actual values for tensile strength and hardness provided a visual assessment of the XGBoost model's accuracy in Fig. 7. The data points clustered closely around the diagonal line, confirming a strong agreement between predicted and experimental results. This alignment demonstrates the model's ability to reliably capture the complex, non-linear relationships between steel composition, heat treatment, and resulting mechanical properties.

Minor deviations from the diagonal were observed in samples with extreme alloying compositions, particularly in hardness predictions. These discrepancies can be attributed to microstructural factors not fully represented in the dataset, such as localized grain refinement, carbide morphology, or secondary phase distributions. Despite these small variations, the overall consistency of the scatter plots reinforces the robustness of the model, validating its capability to generalize beyond the training data while maintaining high predictive accuracy. By visually demonstrating the close correlation between actual and

predicted outcomes, the scatter plots complement the residual analysis, collectively highlighting the reliability of XGBoost as a predictive tool for steel property optimization.

6.2 Feature Importance and SHAP Analysis

To interpret the predictive behavior of the XGBoost model and identify significant contributors to tensile strength and hardness, feature importance and SHAP (Shapley Additive Explanations) analyses were performed. These methods provide a metallurgically meaningful understanding of how input variables influence model predictions. The feature importance results revealed chromium (Cr), carbon (C), and molybdenum (Mo) as the most significant predictors, consistent with established metallurgical principles as indicated in Figs. 8 and 9.

Chromium strongly enhanced both hardness and tensile strength by promoting carbide formation and grain refinement, while molybdenum improved hardenability and stabilized strength during heat treatment. Carbon, although exhibiting a weaker direct correlation in simple statistical analysis, emerged as a critical factor in combination with other alloying elements particularly in controlling matrix microstructure and enabling phase transformations. This highlights the superiority of SHAP analysis over conventional correlation methods, as it captures complex non-linear and synergistic effects that govern metallurgical behaviour performance.

6.3 Prediction Uncertainty and Model Robustness

Prediction uncertainty in this study was evaluated using cross-validation variance across repeated training and testing cycles. The relatively low variance observed in key performance metrics (R^2 , RMSE, and MAE) indicates that the developed models exhibit stable and repeatable predictive behavior, with limited sensitivity to data partitioning. This consistency suggests that the models capture robust composition–processing–property relationships rather than dataset-specific artifacts.

Analysis of the error distributions across the test dataset reveals a near-symmetric spread around zero, with the majority of prediction errors confined within a narrow range. This behavior indicates the absence of systematic bias and suggests that prediction inaccuracies are primarily associated with localized compositional or processing variations rather than global model deficiencies.

While the present study provides a deterministic assessment of uncertainty, advanced probabilistic uncertainty quantification methods, such as Bayesian machine learning or ensemble-based confidence interval estimation, were not implemented. These approaches will be explored in future studies to provide explicit prediction confidence bounds and further enhance model reliability.

7 Heatmap Correlation and Metallurgical Justification

The correlation heatmap in Fig. 9 provided valuable insights into the relationships between input features (chemical composition and heat treatment parameters) and target variables (tensile strength and hardness). The heatmap displayed perfect correlation coefficient values (1.0) along the diagonal, as

expected, indicating that each feature is perfectly correlated. The feature target correlation for hardness shows that chromium exhibited a strong positive correlation of 0.70 primarily due to its role in forming carbides that enhance deformation resistance. Molybdenum showed a moderate correlation (0.57), attributed to its ability to improve hardenability and retain hardness post-heat treatment. Carbon had a weaker but still notable influence (0.36), as it contributes to hardness by influencing the matrix microstructure.

For tensile strength, chromium demonstrated a strong positive correlation of 0.68, likely due to its grain-refining effect, which enhances strength, while also showed a moderate correlation of 0.54, as it inhibits grain coarsening during heat treatment, thereby maintaining strength. Carbon (-0.01) displayed a negligible to slightly negative correlation of -0.01, possibly because its effect was overshadowed by stronger alloying elements like Cr and Mo. The metallurgical significance of heatmap analysis emphasizes the critical influence of chromium (Cr), molybdenum (Mo), and carbon (C) in governing the mechanical properties of the material, reinforcing their selection as significant predictive features [10]. Chromium and Molybdenum exhibit strong correlations with both hardness and tensile strength, primarily due to their roles in carbide formation, grain refinement, and hardenability enhancement. These elements are pivotal in optimizing strength and wear resistance in alloy systems. Carbon, while displaying weaker individual correlations, remains metallurgically significant. Its subdued influence in the heatmap likely stems from competing interactions with dominant alloying elements (Cr, Mo).

However, SHAP analysis confirms carbon's substantial contribution in multi-feature interactions, particularly in microstructural phase stability and solid-solution strengthening. These findings align with established metallurgical principles, where synergistic effects between alloying elements often dictate material performance more than individual contributions. Thus, while Cr and Mo are primary drivers of mechanical properties, carbon's role becomes more pronounced in complex, interdependent feature relationships.

8 Phase Diagram Simulation

The predictive framework was extended through a Streamlit-based tool capable of dynamically simulating iron-carbon phase diagrams under specified compositions and heat treatment conditions.

8.1 Fe-C Phase Region Diagram Visualization

The Streamlit-based simulation tool dynamically generated an iron-carbon phase diagram based on user-specified compositional and thermal parameters. As demonstrated in Fig. 10, the model accurately predicted phase regions including ferrite, pearlite, and austenite for a 0.4 wt.% carbon steel undergoing normalizing heat treatment.

By integrating property prediction with phase diagram visualization, the simulation tool provides an interactive decision-support platform for materials engineers. It enables rapid evaluation of

composition–processing–property relationships, significantly reducing reliance on costly trial-and-error experiment while supporting optimization of heat treatment processes.

8.2 Comparison with Experimental Observations

The integration of experimental testing, predictive modeling, and interpretability analyses establishes a robust framework for property prediction in medium carbon steels. Experimental results validated the trends predicted by the machine learning models, while residuals, scatter plots, SHAP, and correlation analyses demonstrated both predictive accuracy and metallurgical validity.

Moreover, the incorporation of a streamlit application extended the study into a practical decision-support tool, enabling real-time prediction of tensile strength and hardness alongside Fe–C phase diagram visualization. This capability reduces reliance on costly, time-intensive experimental trials and empowers engineers to explore alloy design and heat treatment strategies interactively. Together, the residual patterns, SHAP-based feature contributions, and correlation trends establish both the predictive robustness and the scientific validity of the model. This integrated analysis not only ensures confidence in the results but also strengthens their relevance for guiding alloy design and heat treatment optimization.

9 Metallurgical Interpretation of SHAP and Feature Importance in Relation to Fe–C Phase Stability

The SHAP and feature importance results (Figs. 8 and 9) provide a physically interpretable framework for linking machine learning predictions with metallurgical principles governing strength development in steels. The SHAP summary plot shows that carbon content is the most influential feature affecting predicted mechanical properties. This observation is consistent with the Fe–C phase diagram, where carbon controls the equilibrium phase fractions of ferrite (α), austenite (γ), and cementite (Fe_3C), thereby strongly influencing microstructural evolution and mechanical behaviour.

Within the composition range of medium carbon steels considered in this study, increasing carbon content shifts equilibrium phase stability from ferrite-dominated regions toward ferrite–pearlite mixtures with higher cementite fractions. This transition increases strength through multiple mechanisms. First, carbon atoms dissolved in ferrite contribute to solid solution strengthening, where lattice distortions impede dislocation motion. The strengthening contribution can be approximated by

$$\Delta\sigma_{ss} = kc^n$$

where c represents solute concentration and k and n are material constants. The positive SHAP values associated with increasing carbon concentration therefore reflect the direct strengthening contribution of interstitial carbon.

Second, the Fe–C phase diagram predicts that higher carbon contents increase the fraction of pearlite, a lamellar mixture of ferrite and cementite. The presence of hard cementite lamellae provides significant

resistance to plastic deformation. This microstructural strengthening mechanism is reflected in the SHAP dependence plot for carbon, which shows a monotonic increase in predicted hardness and strength as carbon content approaches the upper range of the dataset.

Alloying elements such as chromium and molybdenum exhibit secondary but meaningful SHAP contributions. These elements influence phase stability by promoting the formation of alloy carbides and modifying transformation temperatures. Their strengthening effect is associated with precipitation and dispersion strengthening, which can be described through the Orowan bypass mechanism, where dislocations bow around finely dispersed carbide particles. The corresponding strengthening increment is approximately

$$\Delta\sigma_{\text{orowan}} \propto Gb/\lambda$$

where G is the shear modulus, b is the Burgers vector, and λ is the spacing between precipitates. The positive SHAP values for chromium and molybdenum therefore reflect their role in stabilizing carbides and improving precipitation strengthening during heat treatment.

Microstructural refinement effects are also captured in the SHAP analysis. Alloying elements that modify transformation behavior can influence the scale of ferrite–pearlite structures and prior austenite grain size. Strengthening associated with grain refinement follows the Hall–Petch relationship, which links yield strength to grain size:

$$\sigma_y = \sigma_0 + k_y d^{1/2}$$

where d is the average grain diameter. Alloying elements that retard grain growth or refine transformation products therefore contribute indirectly to strengthening, which is reflected in their SHAP feature contributions.

The strong correspondence between SHAP-derived feature importance and established strengthening models including solid solution strengthening, precipitation strengthening, and grain boundary strengthening demonstrates that the machine learning model captures physically meaningful composition–phase–microstructure–property relationships. The alignment of ML predictions with phase stability trends predicted by the Fe–C phase diagram therefore supports both the interpretability and metallurgical validity of the developed predictive framework.

10 Conclusion

This study developed a materials informatics framework for predicting the mechanical properties of medium carbon steels by integrating machine learning modeling, metallurgical interpretation, and phase diagram analysis. Among the evaluated algorithms, the XGBoost model provided the highest predictive accuracy for tensile strength and hardness, demonstrating strong capability in capturing complex composition–property relationships within the studied dataset.

Interpretability analysis using SHAP revealed that carbon content and key alloying elements exert the most significant influence on predicted mechanical properties. These findings are consistent with established metallurgical mechanisms, including solid-solution strengthening, carbide precipitation, and phase fraction variations predicted by the Fe–C phase diagram. The integration of machine learning predictions with phase diagram region visualization provides a physically interpretable metallurgical framework that provides reliable composition–phase–property relationships for accelerating alloy design in steel systems.

To enhance practical applicability, a streamlit-based decision-support tool was developed to enable real-time property prediction and phase region identification for candidate steel compositions. The developed approach demonstrates how materials informatics can complement traditional metallurgical analysis in alloy design and property optimization.

However, the predictive capability of the models is primarily constrained to medium carbon steels within the composition and processing ranges represented in the dataset. Future work will focus on incorporating uncertainty quantification, expanding the dataset to additional steel grades, and integrating kinetic transformation models to better account for non-equilibrium microstructural evolution.

Declarations

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflict of Interest: The authors have no relevant financial or non-financial interests to disclose.

Availability of Data and Materials: The dataset of 413 steel samples used for training the machine learning models is available in the GitHub repository: <https://github.com/Th3Samaritan/MAST-FYP.git>.

Code Availability: The Streamlit application and source code developed for this study are available at <https://github.com/Th3Samaritan/MAST-FYP.git>.

Authors' Contributions:

A. A. Adebisi, A.T. Muyideen and U Shehu – Conceptualization and design methodology.

A. A. Adebisi, and A.T. Muyideen – Data collection and experimental analysis.

A. A. Adebisi, and A.T. Muyideen – Analysed and interpreted the result.

A. A. Adebisi, and A.T. Muyideen – Drafted the main manuscript

A. A. Adebisi, and U. Shehu – Supervision, assessment and conceptual guidance

A. A. Adebisi, U Shehu and K. A. Bello – Review and editing

All authors have read and approved the final manuscript.

Ethics Approval: Not applicable.

Consent to Participate: Not applicable.

Consent for Publication: Not applicable.

References

1. Callister WD, Rethwisch DG. (2020). Materials Science and Engineering: An Introduction (10th ed.). John Wiley & Sons. <https://www.google.com/url?sa=E&source=gmail&q=https://doi.org/10.1002/9781119405404>
2. Bhadeshia HKDH, Honeycombe R. (2017). Steels: Microstructure and Properties (4th ed.). Butterworth-Heinemann. <https://www.google.com/search?q=https://doi.org/10.1016/C2014-0-03097-3>
3. Krauss G. (2015). Steels: Processing, Structure, and Performance (2nd ed.). ASM International. <https://www.google.com/search?q=https://doi.org/10.31399/asm.tb.spsp2.9781627082716>
4. Totten GE. Steel Heat Treatment: Metallurgy and Technologies. CRC; 2006. <https://www.google.com/search?q=https://doi.org/.10.1201/9781420017414>.
5. Agrawal A, Choudhary A. Perspective: Materials informatics and big data: Realization of the fourth paradigm of science in materials science. APL Mater. 2016;4(5):053208. <https://doi.org/10.1063/1.4946894>.
6. Mueller T, Kusne AG, Ramprasad R. Machine learning in materials science: Recent progress and emerging applications. Rev Comput Chem. 2016;29:186–273. <https://doi.org/10.1002/9781119148739.ch4>.
7. Pangaribawa MR, Purboputo PI, Wicaksono BF et al. (2024) Effects of quenching and tempering medium carbon steel AISI 1045 to microstructure and hardness. AIP Conf Proc. <https://doi.org/10.1063/5.0179637>
8. Adhikary A. Predicting Mechanical Properties of Low-Alloy Steel Using Machine Learning. Int J Adv Eng Manag; 2021.
9. Rajan K. Materials informatics: The data science of material discovery. Mater Today. 2015;18(10). [https://doi.org/10.1016/S1369-7021\(05\)71123-8](https://doi.org/10.1016/S1369-7021(05)71123-8).
10. Sandhya N, Sowmya V, Bandaru CR, Babu GR. Prediction of Mechanical Properties of Steel Using Data Science Techniques. Int J Recent Technol. 2019;8(3). <https://doi.org/10.35940/ijrte.C3952.098319>.
11. Sun Y, Bai Y, Xu Y. Mechanical property prediction of medium carbon steel using machine learning models. J Mater Res Technol. 2021;15:158–68.

12. Li X, Zhang Y, Zhao P. Prediction of mechanical properties of hot-rolled carbon steel using machine learning. *Mater Des.* 2020;189:108502.
13. Wu Z, Guo W, Zhang J. Prediction of mechanical properties of steel strips based on deep learning. *Ironmak Steelmak.* 2021;48(9):1022–30.
14. Liu Y, Chen W, Zhang J. Machine learning prediction of the outcomes of normalizing in medium carbon steels. *Comput Mater Sci.* 2021;189. <https://doi.org/10.1016/j.commatsci.2021.101243>.
15. Ghasemi A, Eivani AR, Banapour Ghaffari O, et al. Machine learning for predicting high-temperature strength of austenitic stainless steels based on composition and deformation parameters. *J Mater Sci.* 2025;60:24665–83. <https://doi.org/10.1007/s10853-025-11805-6>.
16. Das B, Hulipalled P, Algur V, et al. Enhancing Prediction Accuracy of Machine Learning Models for Materials Informatics Problems in Alloy Design: A Case Study on Dual-Phase Steel. *J Mater Eng Perform.* 2025. <https://doi.org/10.1007/s11665-025-12401-0>.
17. Choudhury PR, Senthilkumar R, Ebin A, Thangarasu H. Artificial intelligence in the optimisation of rubber properties: experimental validation and testing of predicted models. *J Rubber Res.* 2026. <https://doi.org/10.1007/s42464-025-00331-4>. <https://www.google.com/search?q=https://doi..>
18. Cho YE, Sim S, Choi J, et al. Explainable machine learning-based classification of traditional Korean ceramics using XRF chemical composition data. *npj Herit Sci.* 2026;14:28. <https://doi.org/10.1038/s40494-026-02301-4>.
19. Zhang X, Sun J, Zhang Y, et al. Machine learning for phase prediction of high entropy carbide ceramics from imbalanced data. *npj Comput Mater.* 2026;12:22. <https://doi.org/10.1038/s41524-025-01873-2>.
20. Fras T, Pawlowski P. On the Applicability of Generative Artificial Intelligence in Modelling the Dynamic Behaviour of Materials–Potentials and Pitfalls. *J Dyn Behav Mater.* 2026. <https://doi.org/10.1007/s40870-025-00506-5>.
21. Li S, Wei S, Huang C, et al. Extracting and reconstructing knowledge in materials science literature using large language models. *Commun Mater.* 2026;7:31. <https://doi.org/10.1038/s43246-025-01043-3>.
22. Shi J, Gao Y, Lu Z, et al. Advances in the Application of Machine Learning in Composite Materials. *Appl Compos Mater.* 2026;33:42. <https://doi.org/10.1007/s10443-025-10415-4>.
23. Li Y, Kim SY, Shaw S, et al. Material Property Simulation for Advanced Packaging. *J Electron Mater.* 2026;55:1664–73. <https://doi.org/10.1007/s11664-025-12636-6>.
24. Cheng M, Fu CL, Okabe R, et al. Artificial intelligence-driven approaches for materials design and discovery. *Nat Mater.* 2026. <https://doi.org/10.1038/s41563-025-02403-7>.
25. Pathak P, Srivastava S. Smart Material Design: Integrating Data-Driven Optimization and Complexity Analysis for Next-Generation Materials. In: Ahmadian A, Srivastava S, Yadav AK, Kumar V, Kumar Srivastava P, editors. *Smart Materials Engineering*. Cham: Springer; 2026. https://doi.org/10.1007/978-3-032-09540-4_3.

26. Pratap A, Verma AS. Machine Learning's Emergence in Predictive and Prescriptive Materials Design Modeling. In: Ahmadian A, Srivastava S, Yadav AK, Kumar V, Kumar Srivastava P, editors. Smart Materials Engineering. Cham: Springer; 2026. https://doi.org/10.1007/978-3-032-09540-4_1.
27. Borgard O, Chomsaeng N, Wongtimnoi K, et al. Predicting and designing properties of twelve alloy families using artificial neural networks and generative adversarial networks. Bull Mater Sci. 2025;48:153. <https://doi.org/10.1007/s12034-025-03511-5>.
28. Rahman CMA, Romero AH, Gyawali PK. Supervised pretraining for material property prediction. Commun Mater. 2025;6:271. <https://doi.org/10.1038/s43246-025-00999-6>.
29. Li J, Zhang X, Shang C, et al. Reinforcement Learning in Materials Science: Recent Advances, Methodologies and Applications. Acta Metall Sin (Engl Lett). 2025;38:2077–101. <https://doi.org/10.1007/s40195-025-01934-x>.
30. Johnson T, Lee R. (2019) The role of chromium and molybdenum in corrosion resistance and mechanical properties of steels. J Metall Eng 29(2).

Figures

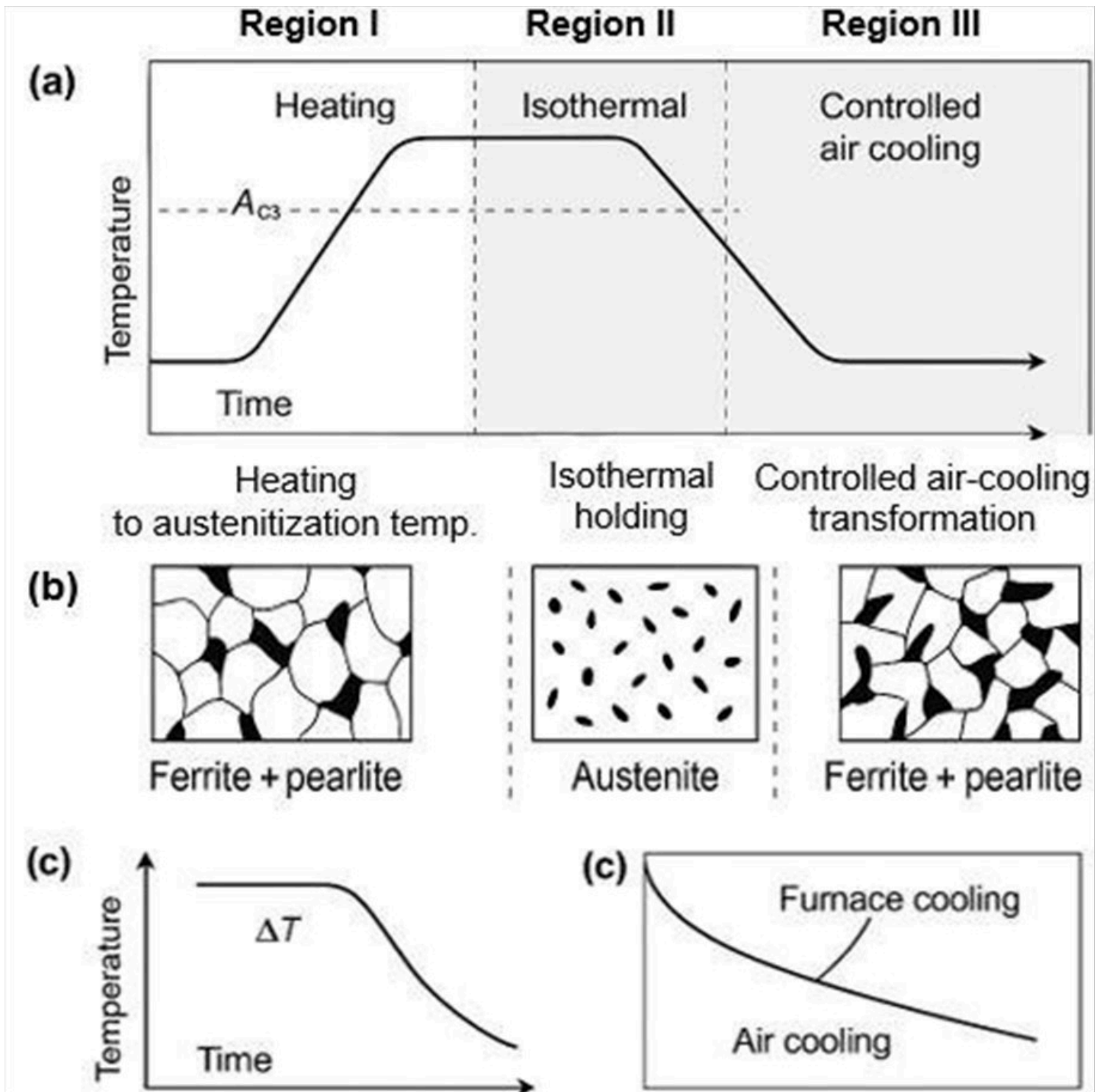


Figure 1

Normalizing heat treatment profile showing (a) temperature-time curve with critical transformation regions annotated, (b) schematic representation of microstructural evolution during processing, and (c) comparative cooling rate analysis between furnace and air-cooling methods. The diagram highlights three distinct phases: heating to austenitization temperature (Region I), isothermal holding period (Region II), and controlled air-cooling transformation (Region III), with metallographic insets showing corresponding microstructural changes at each stage.

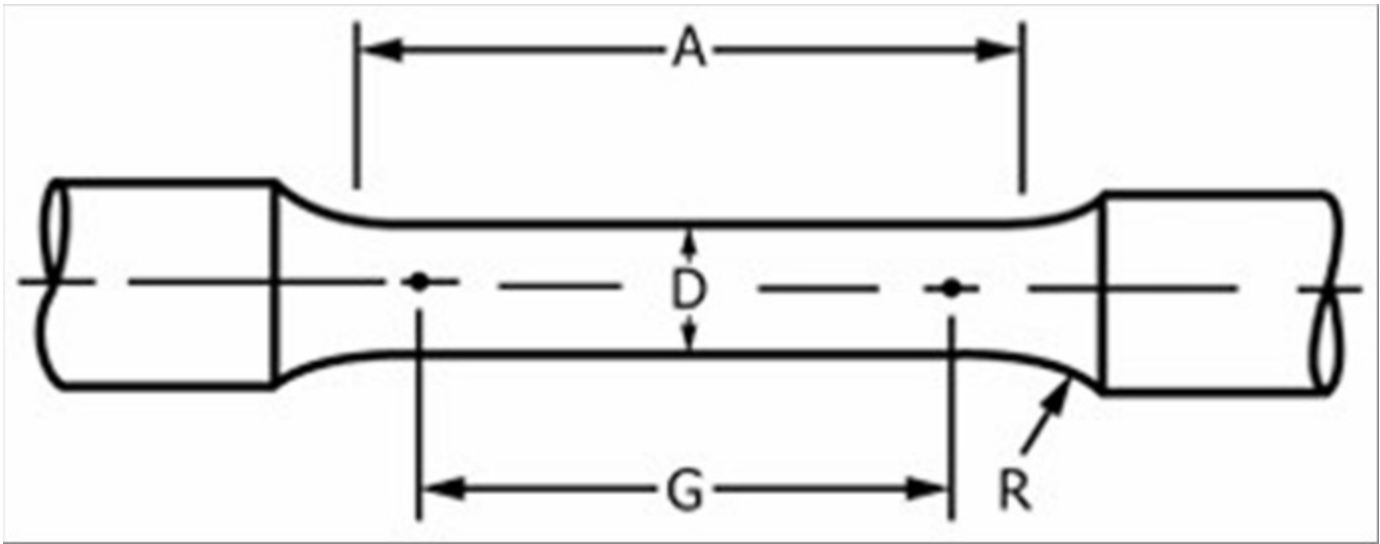


Figure 2

Schematic view of tensile test sample

A = Length of reduced parallel section = 16 mm; **D** = Diameter = 2.5 mm;

G = Gauge length = 10 mm; **R** = Radius of fillet = 2 mm

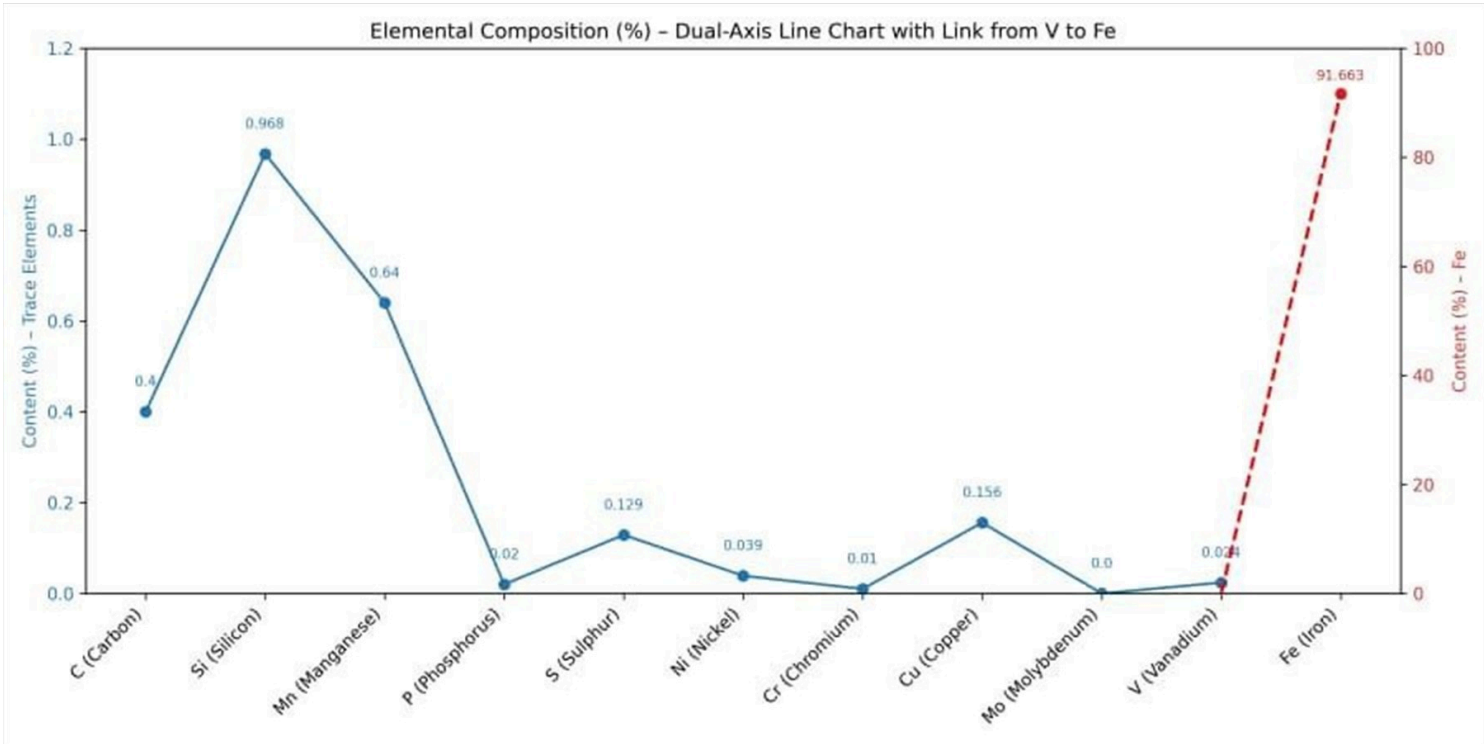


Figure 3

Dual axis line analysis of medium carbon steel elemental composition (%)

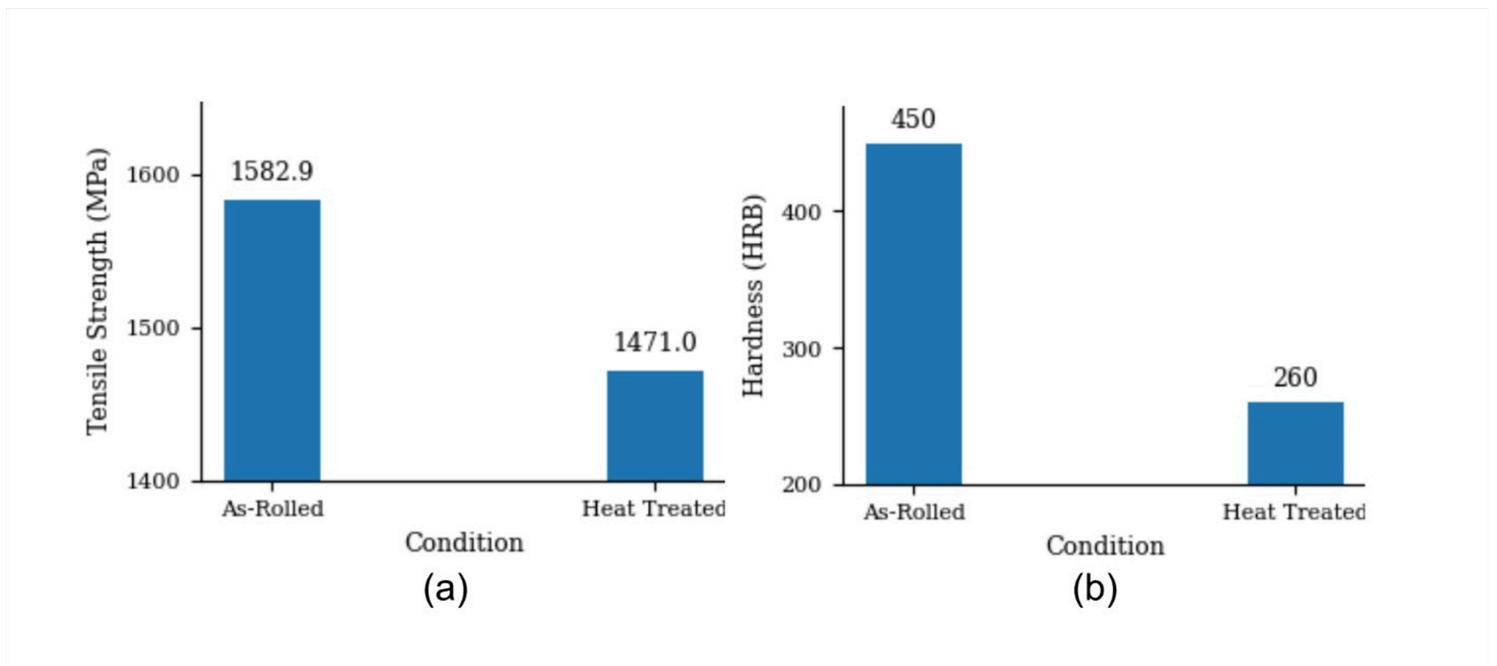


Figure 4

Normalization condition for before and after treatment of steel for (a) Tensile strength, and (b) Hardness

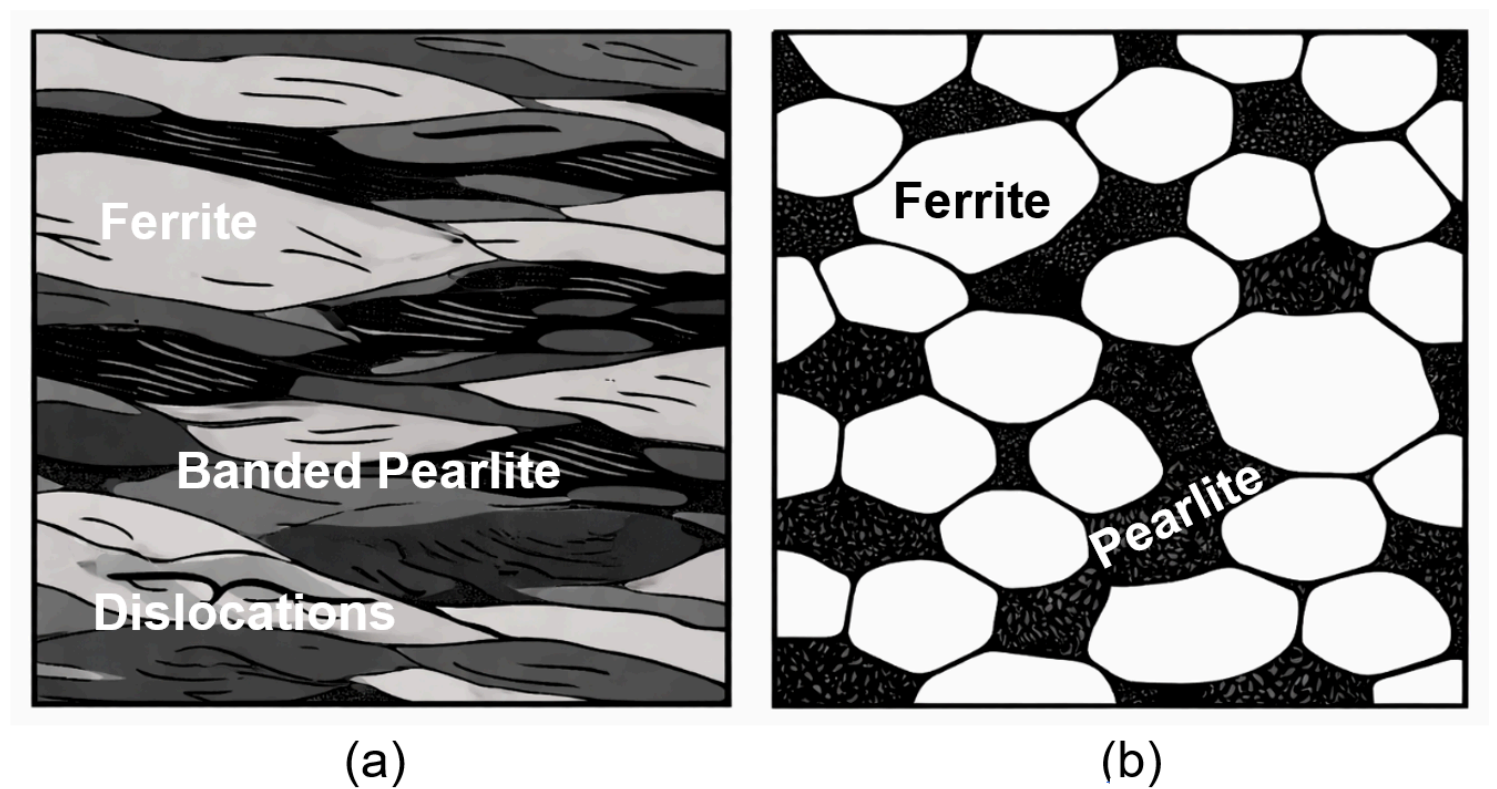


Figure 5

Schematic view of (a) as-rolled steel microstructure and (b) refined ferrite-pearlite microstructure

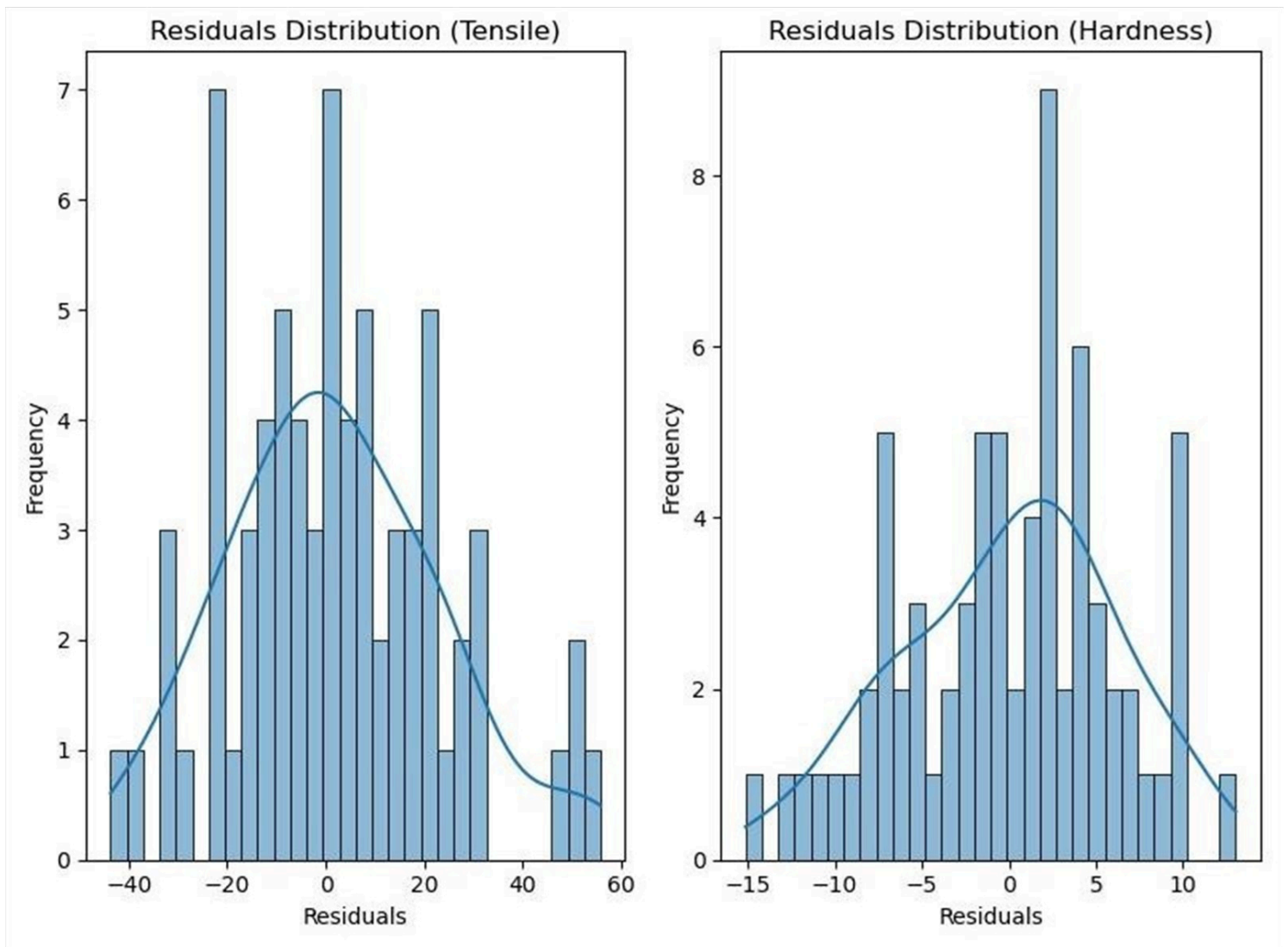


Figure 6

Residual Plot for tensile strength and hardness

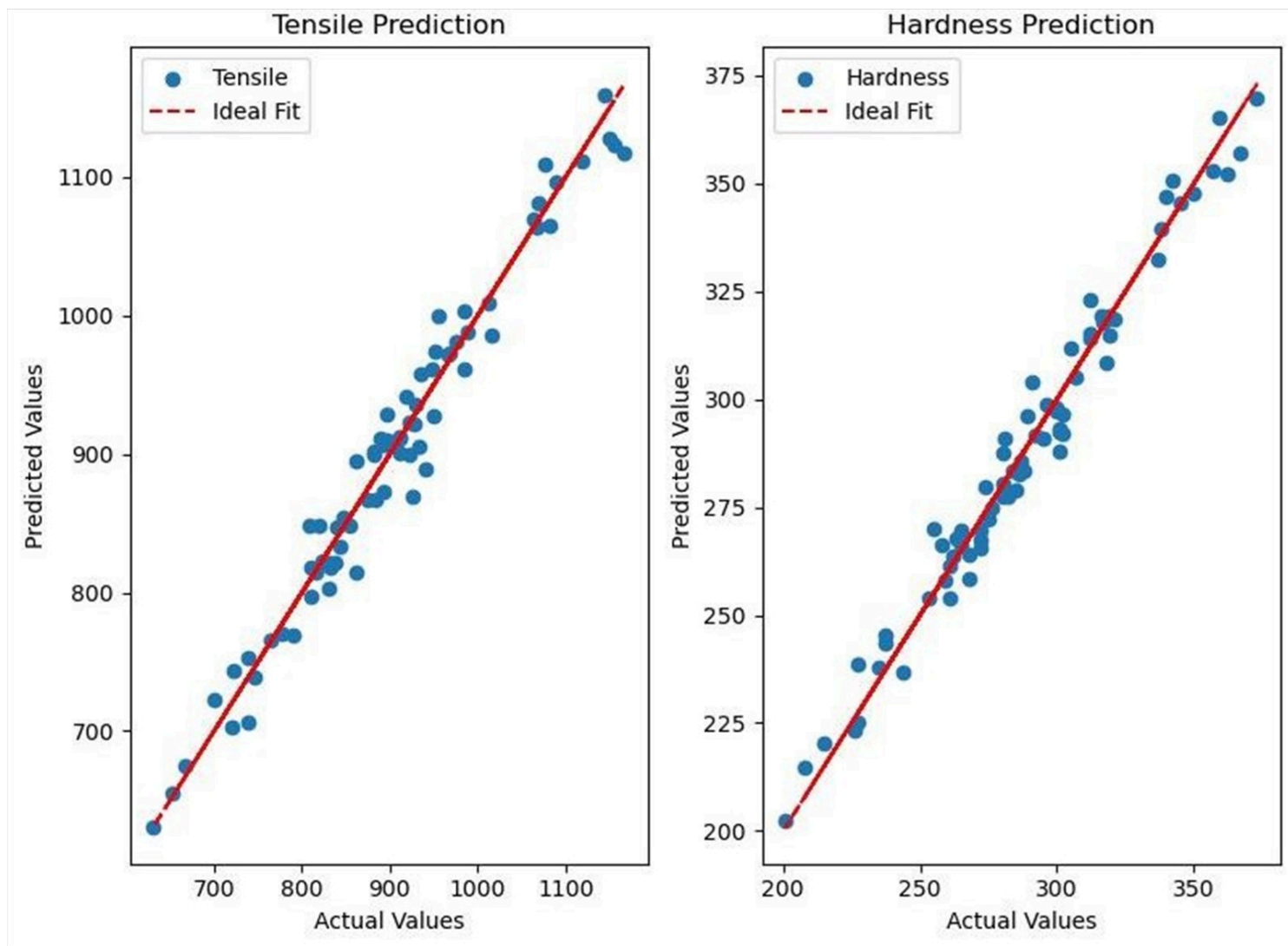


Figure 7

Scatter Plot for Predicted vs. Actual Tensile Strength and Hardness

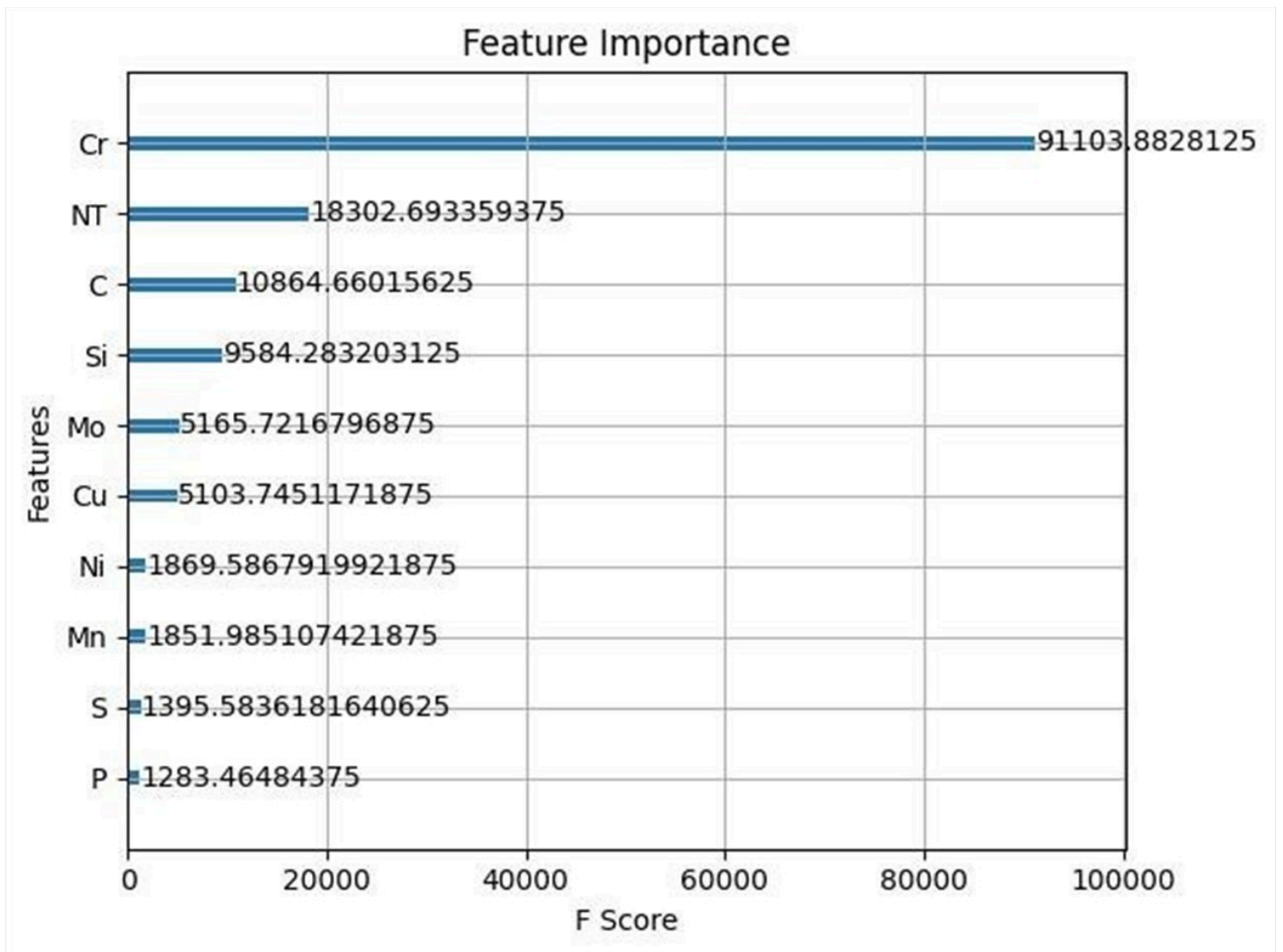


Figure 8

Feature Importance Plot from XGBoost Model

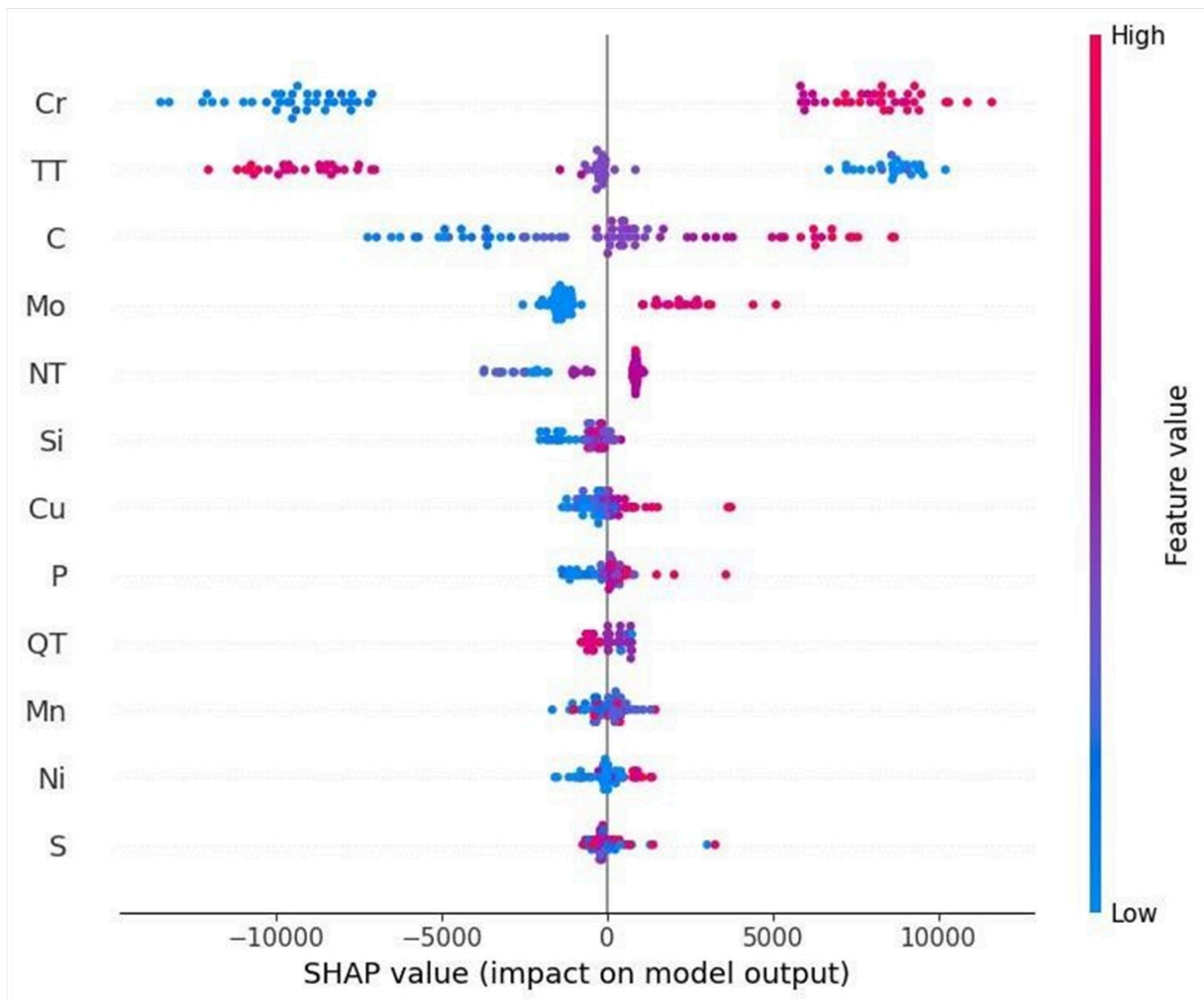


Figure 9

SHAP summary plot illustrating feature contributions to tensile strength and hardness predictions

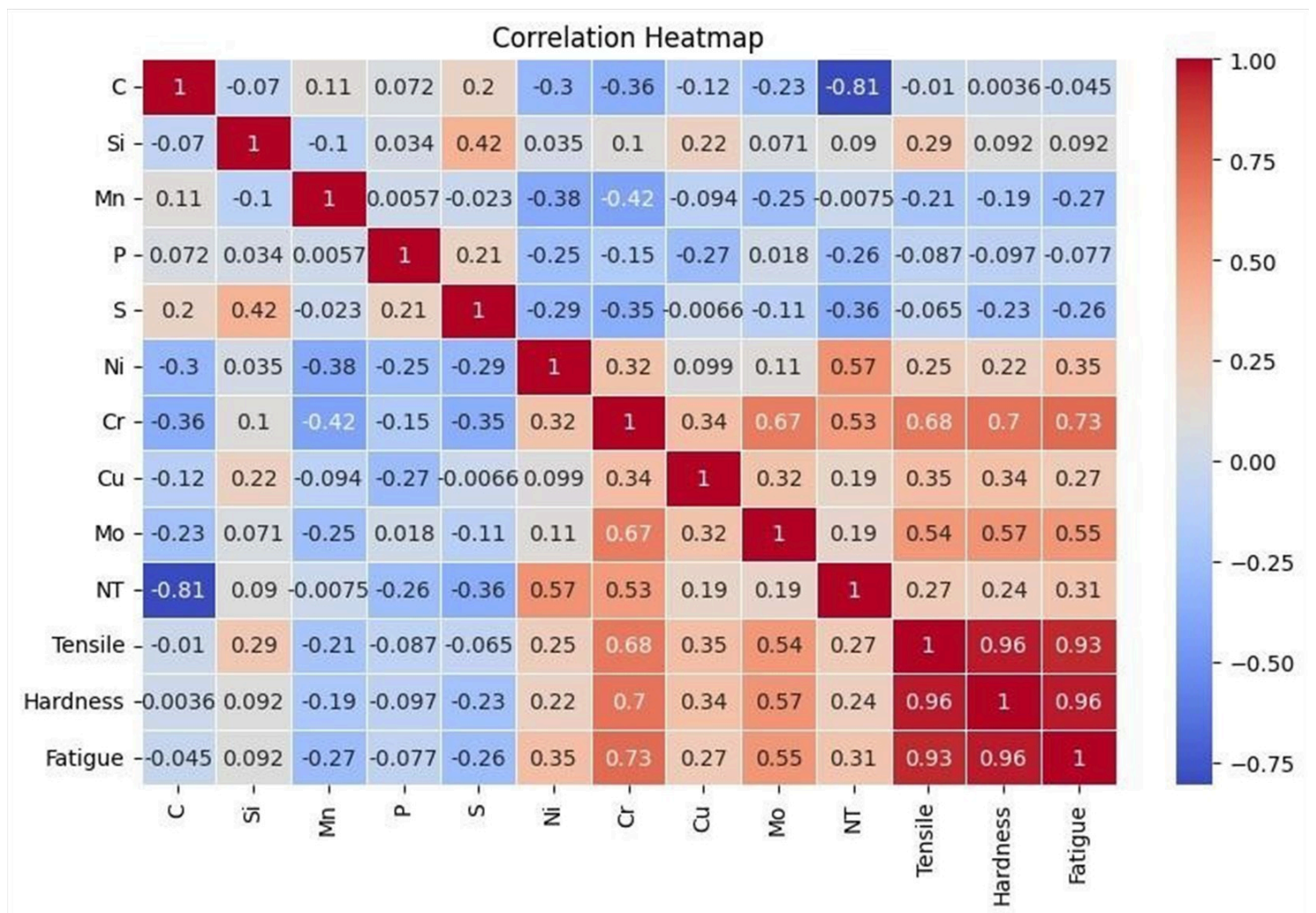


Figure 10

Correlation Heatmap Plot

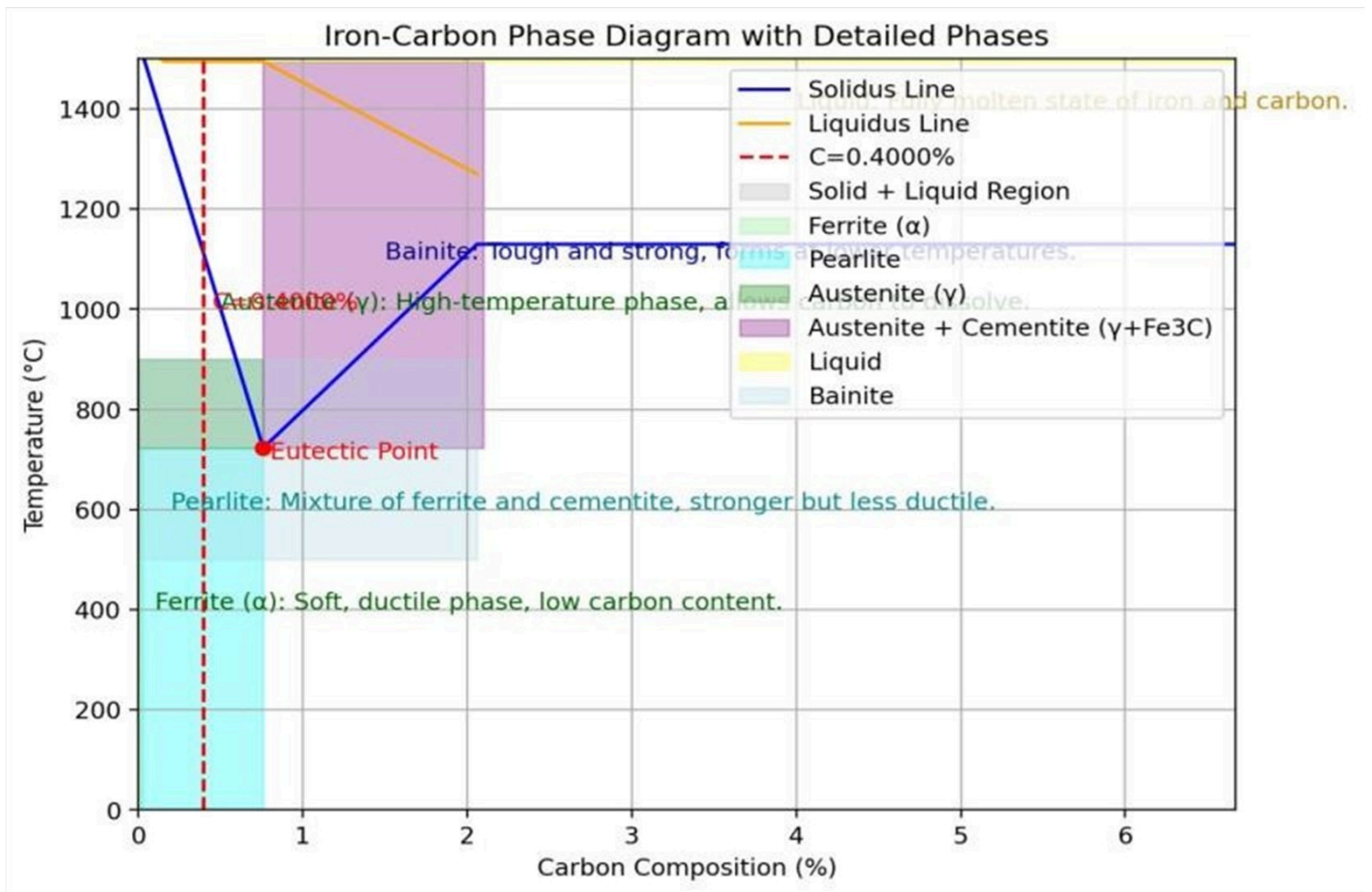


Figure 11

Simulated Fe-C phase diagram (0.4 wt.% C) under normalizing conditions, with experimental validation.