applied sciences

MDPI

Article

Automated Machine Learning for Nitrogen Content Prediction in Steel Production: A Comprehensive Multi-Stage Process Analysis

Jaroslav Demeter *, Branislav Buľko, Peter Demeter, Martina Hrubovčáková, Slavomír Hubatka and Lukáš Fogaraš

Faculty of Materials, Metallurgy and Recycling, Institute of Metallurgical Technologies and Digital Transformation, Technical University of Košice, Letná 1/9, 04200 Košice, Slovakia; branislav.bulko@tuke.sk (B.B.); peter.demeter@tuke.sk (P.D.); martina.hrubovcakova@tuke.sk (M.H.); slavomir.hubatka@tuke.sk (S.H.); lukas.fogaras@tuke.sk (L.F.)

* Correspondence: jaroslav.demeter@tuke.sk; Tel.: +421-55-602-2407

Abstract

Nitrogen control in steel production critically influences mechanical properties and product quality, yet traditional mechanistic models struggle to capture complex multivariable interactions across the complete steelmaking chain. This study developed and validated automated machine learning (AutoML) models using Microsoft Azure Machine Learning Studio to predict nitrogen content at four critical stages: desulfurization of pig iron (Stage 1), basic oxygen furnace prior to tapping (Stage 2), secondary steelmaking initiation (Stage 3), and secondary steelmaking finishing (Stage 4). Industrial data from 291 metal samples across 76 heats were collected and processed, with stage-specific models employing stack ensemble architectures combining 4–7 algorithms with feature sets ranging from 12 to 35 variables. The models achieved normalized root mean squared errors between 0.112–0.149, mean absolute percentage errors of 14.6–21.1%, and Spearman correlations of 0.310–0.587, with secondary steelmaking models demonstrating superior performance due to more controlled thermodynamic conditions. All models achieved sub-second prediction latencies suitable for real-time industrial implementation. This research demonstrates that AutoML effectively captures complex physicochemical relationships governing nitrogen behavior throughout the steelmaking process, providing practical solutions for Industry 4.0 applications in steelmaking process control and quality optimization.

Keywords: automated machine learning; nitrogen prediction; steelmaking process; ensemble modeling; process optimization; digitalization

Academic Editors: Andrea Carpinteri, Vytautas Bucinskas, Andrius Dzedzickis and Janis Arents

Received: 6 November 2025
Revised: 9 December 2025
Accepted: 29 December 2025
Published: 31 December 2025

1. Introduction

Nitrogen control in steel production represents one of the most challenging aspects of modern metallurgy, significantly influencing the mechanical properties, corrosion resistance, and overall quality of steel products [1–3]. The precise prediction and control of nitrogen content throughout the steelmaking process chain is critical for achieving desired steel specifications while optimizing production efficiency and minimizing material waste [4,5]. Traditional approaches to nitrogen prediction have relied predominantly on mechanistic models based on thermodynamic equilibrium calculations and empirical relationships derived from extensive plant trials [6,7]. While these methods provide valuable insights into the fundamental physics and chemistry of nitrogen behavior, they

often struggle to capture the complex, multivariable interactions characteristic of industrial steelmaking processes [8,9]. The increasing availability of high-quality process data and advances in computational capabilities have created unprecedented opportunities for data-driven modeling approaches [10,11]. Automated machine learning (AutoML) has emerged as a transformative technology in metallurgical process modeling, offering significant advantages over traditional manual machine learning approaches [12,13]. AutoML frameworks automate the entire machine learning pipeline, including feature selection, algorithm selection, hyperparameter optimization, and ensemble construction, thereby reducing the requirement for extensive machine learning expertise while potentially achieving superior predictive performance [14,15]. Recent studies have demonstrated the effectiveness of AutoML in various metallurgical applications. Chen et al. [1] developed an automated ML program for predicting electrochemical reaction energies in metal-catalyzed nitrogen reduction reactions, achieving excellent performance without manual intervention. The program successfully identified key descriptors such as metal charge variance and electronegativity through automated feature selection processes. Zhang and Yang [2] conducted a comprehensive review of machine learning applications in steelmaking process modeling, highlighting that artificial neural networks, support vector machines, and case-based reasoning represent the most frequently employed algorithms. Despite these advances, a significant gap remains in the literature regarding the comprehensive application of AutoML across the complete steelmaking production chain. To the best of our knowledge, no previous studies have systematically applied AutoML methodology to predict nitrogen content across four consecutive production stages of steel manufacturing, thereby enabling holistic coverage of the entire production sequence from pig iron desulfurization processing to final molten steel formation within secondary steelmaking. Furthermore, the present work addresses a methodological advancement in AutoML implementation by developing a framework specifically designed for industrial deployment in steel production environments. This contribution represents a novel integration of AutoML technology with the practical requirements and constraints of real-world metallurgical operations, bridging the gap between academic research and industrial applicability in process optimization and quality control systems. The application of machine learning techniques to nitrogen prediction in steelmaking has shown considerable promise across various process stages. Patra et al. [6] developed a mathematical model for predicting nitrogen content in steel melts during stainless steel production using argon oxygen decarburization (AOD) converters. Their model effectively predicted nitrogen content with respect to blowing time, considering bath temperature, composition variations, and nitrogen absorption–desorption kinetics. Yoon et al. [4] presented a nitrogen prediction model for a 320-tonne converter, combining thermodynamic and kinetic approaches with real-time process data. Their model considered heat supply from oxidation reactions, CO evolution, and scrap melting dynamics. The study emphasized the importance of understanding nitrogen behavior in the context of reduced hot metal ratios and carbon neutrality initiatives in steelmaking. Pitkälä et al. [14] investigated nitrogen behavior in AOD processes during nitrogen-alloyed stainless steel production. They developed kinetic models integrated with thermodynamic equilibrium equations, successfully predicting nitrogen contents in the range of 0.150–0.400% for both nitrogenation and denitrogenation stages.

AutoML frameworks have demonstrated significant potential in steel production applications beyond nitrogen prediction. Wu et al. [5] implemented an AutoML approach using AutoGluon-Tabular for inclusion prediction and process optimization, achieving an 8.85% reduction in inclusion rates. Their quantitative causal analysis framework effectively identified critical process parameters and their interactions. Conrad et al. [8] conducted a comprehensive benchmarking study of AutoML for regression tasks on small tabular

datasets in materials design. They found that AutoML approaches were highly competitive with manual model optimization, with data sampling strategies being crucial for reliable results. Zhang et al. [9] evaluated TPOT and Auto-learn frameworks for mechanical property prediction in steel, demonstrating that AutoML was competitive with manual optimization while providing efficient model selection capabilities.

Ensemble methods have proven particularly effective for metallurgical process modeling due to their ability to capture diverse aspects of complex physicochemical systems [16,17]. Liu et al. [7] developed a stacked autoencoder with sparse Bayesian regression for end-point prediction problems in steelmaking, achieving improved accuracy and providing uncertainty estimates through error bars. The stack ensemble approach combines multiple base learners through a meta-learner, enabling the capture of different types of relationships within the data [18,19]. This architecture is particularly well-suited for steelmaking applications, where linear thermodynamic relationships coexist with non-linear kinetic processes and complex interfacial phenomena [20].

Data quality represents a critical factor in the success of machine learning applications in steelmaking. Zhang and Yang [2] emphasized that collected data in steelmaking plants are frequently faulty, making data processing, particularly data cleaning, crucially important for model performance. The detection of variable importance can be used to optimize process parameters and guide production decisions. Xiao et al. [13] demonstrated that AutoML platforms produce more reliable nanotoxicity prediction models than conventional ML algorithms, with better data quality significantly enhancing performance. This finding extends to metallurgical applications, where sensor noise, measurement uncertainties, and process disturbances can significantly impact model reliability.

Despite the promising developments in AutoML and machine learning for metallurgical applications, there remains a significant gap in comprehensive studies that systematically evaluate nitrogen prediction across multiple stages of the steel production process. Most existing studies focus on individual process units or specific steel grades, limiting the understanding of nitrogen evolution throughout the complete steelmaking chain.

This study addresses this gap by developing and evaluating AutoML models to predict nitrogen content at four critical stages of steel production. The research focuses on predicting nitrogen in pig iron after desulfurization, in crude steel prior to tapping from a basic oxygen furnace, at the initiation of secondary metallurgy and at the completion of secondary metallurgy. The research objectives center on developing and evaluating stage-specific AutoML models using the Microsoft Azure platform. This involves the comparative evaluation of model performance across different steel production stages, analysis of feature importance and the optimization of ensemble architecture. Finally, the project aims to assess model interpretability and industrial applicability while investigating the relationship between model complexity and predictive performance.

The study contributes to the growing body of knowledge on Industry 4.0 applications in metallurgy and provides practical insights for implementing automated nitrogen control systems in modern steelmaking operations.

2. Materials and Methods

The following is a brief outline of the production process of the material used to generate the data for this article. Pig iron that has been produced in a blast furnace is subject to pretreatment, i.e., desulfurization, using a vertical refractory lance (a Scandinavian lance). A mixture based on CaO and Mg was used in the desulfurization process. The mixture was injected into the pig iron using nitrogen as a carrier gas. In the next step, the desulfurized pig iron was charged in BOF onto pre-charged steel scrap. The BOF vessel had a maximum charging weight of 170 tonnes. An oxygen lance was then inserted into the BOF vessel and

pure oxygen was blown in at supersonic speed for approximately 17 min. Blowing oxygen involving the oxidation of impurities in the molten metal, such as Si, C, Mn, and P, increased the physical heat and reduced the content of individual elements to the required values specified by the steel grade being produced. If the chemical composition or temperature did not meet specifications, reblow of oxygen was performed. The crude steel was then tapped into a ladle containing carburizers (coke) at the bottom to promote mixing of the melt by generating of CO_2 bubbles. During tapping, aluminum blocks were added to the steel stream to deoxidize it. Subsequently, ferroalloys were added to the ladle during tapping. Throughout its time in secondary metallurgy, fine ferroalloys were added to finish the steel. First, the melt was stirred in the ladle using argon, and then it was softly bubbled with argon blown through a porous plug located at the bottom of the ladle. Steel treated in this way was then prepared for casting on a continuous casting machine. None of the monitored heats were processed using an RH vacuum degasser.

Between 17 May 2025 and 22 May 2025, a total of 291 metallic specimens from 76 distinct heats were systematically gathered and investigated for nitrogen concentration. The methodology was designed to encompass samples from all four stages of production (Figure 1). This approach facilitated the tracking of nitrogen levels within each specific heat, along with its fluctuations. Specifically, 76 samples were sourced from pig iron post-desulfurization, 68 from crude steel prior to tapping, 75 from molten steel in the ladle at the start of secondary metallurgy and 72 at its conclusion. Nevertheless, due to inadequate and unreliable evaluations, it was still not possible to gather 76 samples from each phase.

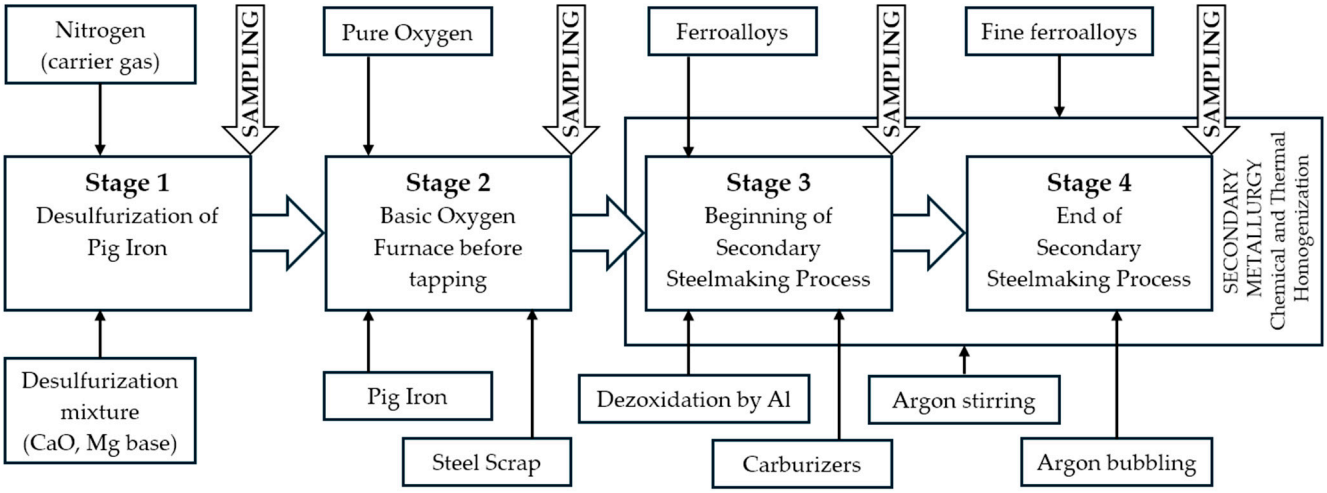

Figure 1. Schematic representation of BOF steel production and processing, showing the sequence of individual additions and operations within the monitored stages. The moment of sample collection for chemical analysis is indicated.

The study utilized industrial data from four distinct stages of the steel production process, each representing a critical control point for nitrogen content management. Data collection spanned a comprehensive range of process parameters, chemical compositions, and operational variables across the complete steelmaking chain (Figure 1).

The collected samples included two steel grades. The first was a structural steel characterized by over 0.80% manganese and guaranteed aluminum content (Grade 1) and the second grade analyzed was a deep-drawing, aluminum-killed (Al-killed) steel (Grade 2). Grades meeting the prescribed chemical composition listed in Table 1.

Table 1. Prescribed chemical composition of the analyzed steel grades.

Steel Grade	C (%)	Mn (%)	Si (%)	Al (%)	P (%)	S (%)	Nb (%)
Grade 1	0.07–0.21	0.8–1.6	0.03–0.6	min 0.02	max. 0.025	max. 0.020	-
Grade 2	0.02–0.1	0.1–0.55	max. 0.08	0.02–0.07	0.01–0.07	max. 0.020	0.004–0.0075

The examination of nitrogen concentration in pig iron and steel specimens was performed at the Quantometric Laboratory of U. S. Steel Košice (Labortest, s.r.o.) utilizing an ELTRA ON 900 (ELTRA GmbH, Haan, Germany) combustion analyzer. This apparatus functions based on the thermal conductivity detection methodology (complying with the ASTM E-1019 standard), featuring an extensive measurement range of 0.0001–0.03% N and a commendable error range of ±0.1 ppm or ±1% of the nitrogen content [21]. Quality assurance is upheld through a systematic approach that includes manufacturer calibration, annual service evaluations, and hourly assessments conducted by the laboratory technician utilizing a standard sample. To ascertain the trustworthiness of the results gathered, two analyses (primary and control) were executed for every specimen. Based on ELTRA manufacturer instructions [22], daily analysis of certified reference materials (minimum 3 replicates) with certified concentrations. The software calculates a new calibration factor to adjust for changes in reagent condition and instrument response. Drift correction shall be performed prior to sample analysis and repeated as necessary during measurement sequences. Under proper operation with regular maintenance and drift correction, ELTRA inert gas fusion analyzers achieve standard deviations of 1–3%, meeting ASTM E1019 requirements [23]. Typical inter-day precision for nitrogen determination ranges from ±0.1–2 ppm (low concentrations) to <1.5% relative standard deviation (higher concentrations) [21].

The quantification of nitrogen content, denoted by heat and sample IDs, were harmonized with pertinent databases. These databases encompassed records of the chemical composition of metal, thermal conditions of metal, weight, and additional operational parameters pertinent to that particular stage of processing. This extensive, harmonized repository of parameters was initially assembled using Microsoft Excel 365 (version 2510, build 16.0.19328.20178) in conjunction with the Lumivero XLSTAT 2019 (Lumivero Inc., Denver, CO, USA; version 2019.2.2) statistical add-in. Gretl 2025a (build 20 March 2025), a sophisticated statistical tool, was used to create the graphs. The training of the AutoML models was facilitated by Azure Machine Learning Studio, which employed a computing solution in the form of a virtual machine labelled as Standard_DS3_v2 and characterized by the following technical specifications (Table A1).

2.1. Data Collection

2.1.1. Desulphurization of Pig Iron (Stage 1)

The desulfurization stage involved the treatment of pig iron to reduce sulfur content before steelmaking. The dataset comprised 15 input features including chemical composition parameters (C, Mn, Si, P content), sulfur levels before and after desulfurization, temperature measurements, mass balance parameters, and process timing variables. The target variable was nitrogen content percentage after desulfurization treatment.

2.1.2. Basic Oxygen Furnace Before Tapping (Stage 2)

The BOF stage represented the primary steelmaking process where molten pig iron is transformed to crude steel through oxygen blowing. This dataset included 32 input features encompassing chemical composition of steel and slag, temporal parameters, temperature and oxygen activity measurements, mass balance parameters for raw materials and fluxes,

and operational variables. The target variable was nitrogen content percentage in crude steel before tapping from the BOF.

2.1.3. Beginning of Secondary Steelmaking Process (Stage 3)

This stage captured the beginning of secondary steelmaking processes in the ladle, focusing on initial composition adjustment and temperature control. The dataset featured 12 input features, including chemical composition of the steel after tapping, initial secondary steelmaking temperature, tapping duration, mass parameters, and tapping geometry factors. The target variable was nitrogen content percentage in a steel at the beginning of secondary steelmaking process.

2.1.4. End of Secondary Steelmaking Process (Stage 4)

The final Stage 4 represented the completion of secondary steelmaking processes, including final alloying, deoxidation, and argon stirring operations. This comprehensive dataset included 35 input features covering deoxidation additives (aluminum in various forms), complete chemical composition, process timing parameters, argon stirring parameters, mass balance variables, and ferroalloy addition strategies. The target variable was final nitrogen content percentage in steel at the end of secondary metallurgy.

2.2. Data Preprocessing and Quality Control

Data preprocessing followed a systematic approach to ensure data quality and model reliability. The preprocessing pipeline included several critical steps implemented within the Microsoft Azure AutoML framework.

Initial data cleaning involved the identification and treatment of missing values, outliers, and erroneous measurements. Missing values were handled through multiple imputation strategies, with the specific approach selected based on the missingness pattern and variable type. Outlier detection employed statistical methods including interquartile range (IQR) analysis and modified z-score calculations, with outliers either removed or adjusted based on process knowledge and physical constraints.

The AutoML framework implemented automated feature engineering, including the creation of interaction terms, polynomial features, and temporal differences where appropriate. Feature normalization employed multiple scaling strategies optimized for each algorithm within the ensemble. Continuous variables were typically standardized (zero mean, unit variance) or normalized to [0, 1] range depending on the algorithm requirements. Categorical variables, where present, were encoded using one-hot encoding or target encoding based on cardinality and predictive power.

All models implemented 10-fold cross-validation with stratified sampling to ensure representative training and validation splits. The cross-validation strategy was designed to maintain temporal order where relevant and ensure that each fold contained representative samples from different operational conditions and time periods.

2.3. AutoML Model Development

Model development utilized Microsoft Azure Machine Learning Studio, which provided a comprehensive AutoML framework with automated algorithm selection, hyperparameter optimization, and ensemble construction capabilities. The AutoML framework evaluated multiple algorithm families, which are described in short below.

- Linear Models: Elastic Net is a linear regression technique that combines the penalties of L1 (Lasso) and L2 (Ridge) regularization into a single loss function to improve model performance [24];

- Tree-Based Methods: Random Forest is an ensemble learning method for classification and regression that uses decision trees as base models, leveraging bootstrap aggregating (bagging) to improve stability and accuracy [25];
- Boosting Algorithms: LightGBM and XGBoostRegressor are high-performance, open-source gradient boosting frameworks designed for speed, efficiency, and scalability, particularly with large datasets. LightGBM's speed and memory efficiency are due to its advanced optimizations and leaf-wise growth strategy [26];
- Instance-Based Methods: K-Nearest Neighbors (KNN) with local interpolation is a method used for regression and spatial data analysis, where the target value for a query point is predicted by interpolating the values of its k nearest neighbors in the training set [27];
- Regularized Methos: LassoLars is a method that combines the Least Angle Regression (LARS) algorithm with L1 penalization, effectively merging the efficiency of forward feature selection with the sparsity-inducing properties of L1 regularization [28].

The final models employed stack ensemble architecture where multiple base learners were combined through a meta-learner (1). Nomenclature is listed at the end of the paper.

$$\hat{y} = g(f_1(x),\ f_2(x), \cdots, f_n(x)) \quad (1)$$

The meta-learner training utilized 20% of the ensemble data to prevent overfitting while maintaining sufficient training samples for reliable learning.

Hyperparameter optimization employed Bayesian optimization techniques to efficiently explore the hyperparameter space. The optimization process considered both individual algorithm performance and ensemble diversity to maximize overall predictive capability.

2.4. Models Configuration Settings

In the context of training models in individual monitored stages (Stage 1–Stage 4), the most successful models in Microsoft Machine Learning Studio exhibited the following job settings (Table 2).

Table 2. Task settings of the most successful models in Microsoft Machine Learning Studio for predicting the amount of nitrogen in metal.

Model Stage	Input Features	Allowed Models	Metric Score Threshold	Number of Cross Validation	Percentage Validation of Data (%)	Percentage Test of Data (%)
Stage 1	15	LightGBM, XGBoostRegressor, RandomForest, ElasticNet	0.1	10	20	20
Stage 2	32	GradientBoosting, ElasticNet, DecisionTree, KNN, LassoLars, RandomForest, LightGBM	0.08	10	20	20
Stage 3	12	ElasticNet, GradientBoosting, DecisionTree, KNN, LassoLars, RandomForest	0.08	10	20	20
Stage 4	35	ElasticNet, GradientBoosting, DecisionTree, KNN, LassoLars, RandomForest, LightGBM	0.1	10	20	20

2.5. Model Evaluation Metrics

Model performance was evaluated by using multiple complementary metrics to provide comprehensive assessment of predictive capability:

Primary metrics

- Normalized Root Mean Squared Error (*NRMSE*): is a scale-independent metric used to assess model performance, allowing for fair comparisons across datasets with different scales (2) where Root Mean Squared Error (*RMSE*) is calculated as (3) [29].

$$NRMSE = \frac{RMSE}{y_{max} - y_{min}} \quad (2)$$

$$RMSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (3)$$

- Coefficient of Determination (R^2): measuring the proportion of variance explained by the model [30] and is calculated by Equation (4).

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (4)$$

- Mean Absolute Error (*MAE*): is a widely used metric for evaluating the accuracy of regression models, calculated using the Equation (5). This computes the average of the absolute differences between predicted and actual values, providing a straightforward measure of the average magnitude of errors [31].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (5)$$

- Mean Absolute Percentage Error (*MAPE*): is a measure of prediction accuracy for forecasting methods, commonly used in statistics and regression analysis [32]. *MAPE* is widely used as a loss function in regression problems because it provides a relative measure of error and is expressed by the Equation (6).

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| [\%] \quad (6)$$

Secondary metrics

- Spearman Rank Correlation: Non-parametric measure of monotonic relationship strength [33];
- Explained Variance: Variance proportion explained by the model [34];
- Normalized Median Absolute Error: Robust central tendency error measure [35].

2.6. Model Interpretability and Explainability

Model interpretability was enhanced through the integration of SHAP analysis [36], providing both global and local explanation capabilities [37]. SHAP values were calculated for each feature (monitored parameter) to understand their contribution to individual predictions and overall model behavior [38]. Feature importance rankings were generated for each base learner and the ensemble as a whole [39], enabling the identification of critical process parameters and their relative influence on nitrogen content prediction [40]. This information supports both model validation and process optimization initiatives [41].

2.7. Industrial Validation and Deployment Considerations

Model validation included assessment of industrial applicability, computational requirements, and real-time deployment feasibility. Performance evaluation considered both accuracy metrics and practical constraints such as prediction latency, model update frequency, and integration with existing process control systems [42,43]. The study evaluated

the trade-offs between model complexity and interpretability, considering the requirements for industrial deployment where model transparency and operator understanding are critical for successful implementation [44]. Further research with an industrial partner is required to test the model on real data in an industrial environment.

3. Results

3.1. Overall Model Performance Comparison

The comprehensive evaluation of AutoML models across the four steelmaking stages revealed distinct performance characteristics and varying degrees of predictive capability. Table 3 presents the complete performance metrics for all developed models.

Table 3. Performance Metrics for All Four Nitrogen Prediction Models.

Model Stage	Model Name	NRMSE	MAE	MAPE (%)	Spearman Correlation
Stage 1	NitroML-DeS	0.14878	0.00053011	14.608	0.31017
Stage 2	NitroML-BOF	0.12735	0.00046536	20.635	0.48132
Stage 3	NitroML-SMB	0.12699	0.00063443	21.081	0.58554
Stage 4	NitroML-SME	0.11239	0.00063450	20.283	0.58692

The results demonstrate considerable variation in model performance across different process stages, reflecting the varying complexity of nitrogen behavior and the availability of predictive features at each stage.

Validity ranges of the models listed in Table 3 are exhibited in Tables A2–A5.

As demonstrated by the flow charts (Figures A1–A4) presented in the Appendix A, this research framework is outlined in comprehensive detail with the subsequent implementation of the model across the specified stages.

3.2. Stage-Specific Model Performance Analysis

3.2.1. NitroML-DeS Model Performance (Stage 1)

The desulfurization model NitroML-DeS achieved an *NRMSE* score of 0.14878, representing the most challenging prediction environment among all stages. The model demonstrated a mean absolute percentage error of 14.608%, which was the lowest among all models, and a Spearman correlation of 0.31017.

These values can be attributed to the complex physicochemical interactions during the desulfurization process, where nitrogen behavior is influenced by multiple simultaneous reactions including sulfur removal, temperature variations, and slag-metal equilibria. Despite the lower correlation coefficient, the model achieved the lowest *MAPE*, suggesting consistent relative accuracy across the nitrogen content range.

The NitroML-DeS model configuration indicated that polynomial relationships and power transformations were critical for capturing the non-linear behavior characteristic of high-temperature metallurgical processes. The model's performance suggests that nitrogen content changes during desulfurization follow complex kinetic patterns that require sophisticated ensemble approaches to predict accurately.

The NitroML-DeS model analysis was facilitated by graphical representations, including line graphs, histogram and scatter plots. The model was evaluated based on residuals, representing the difference between the measured and predicted nitrogen content of desulphurised pig iron. Figure 2a shows the residual variance for the NitroML-DeS model. Figure 2b shows the blue dots arranged close to the line, demonstrating the normal distribution of the residuals. This is also confirmed by the histogram in Figure 2c. Figure 2d shows a graphical comparison of the measured and predicted nitrogen values using the NitroML-DeS model.

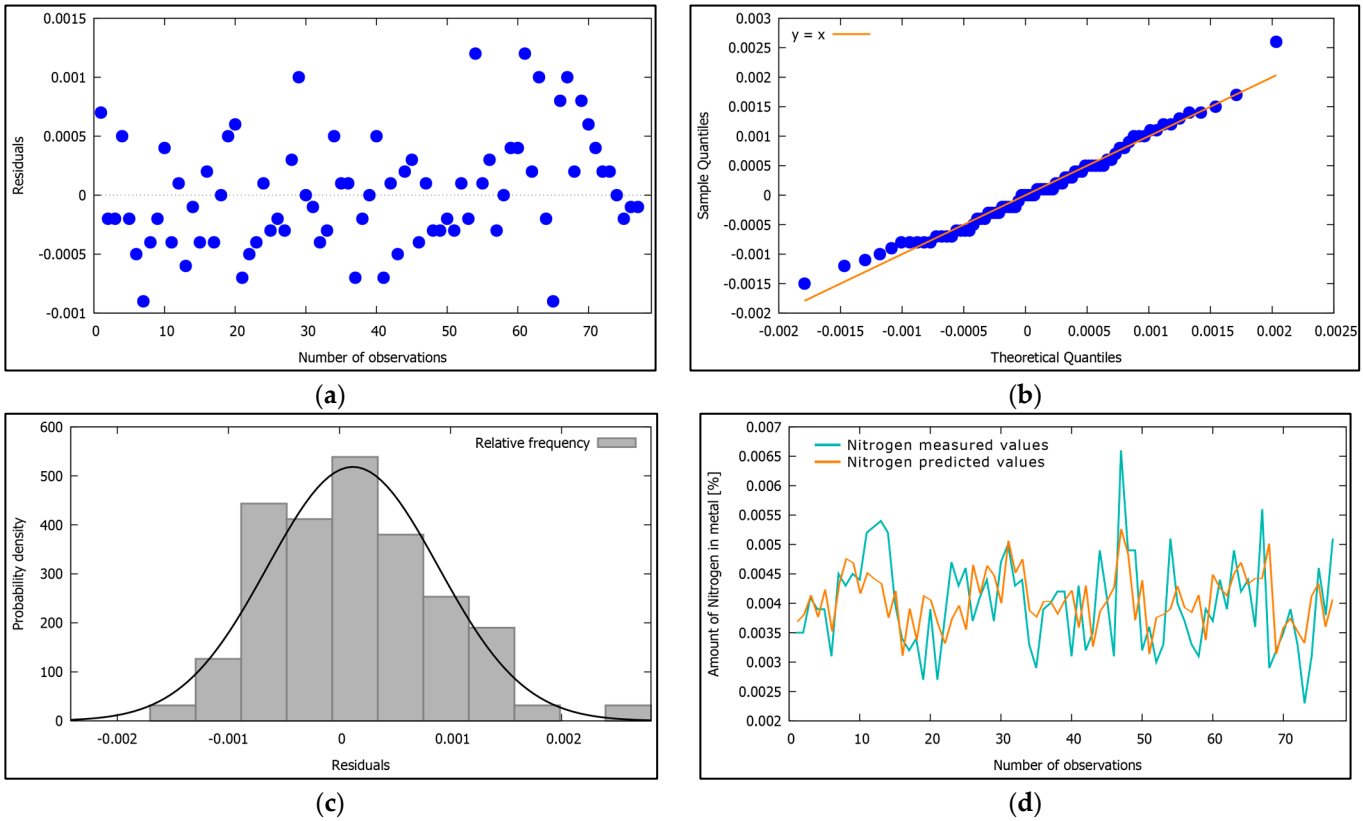

Figure 2. Graphical analysis of the NitroML-DeS model: (**a**) Residual dispersion for NitroML-DeS model; (**b**) Q-Q plot of residuals for NitroML-DeS model; (**c**) Residual frequency distribution for the NitroML-DeS model; (**d**) Comparison of measured and predicted amount of nitrogen in molten metal by using NitroML-DeS model.

3.2.2. NitroML-BOF Model Performance (Stage 2)

The NitroML-BOF model demonstrated improved performance with an *NRMSE* of 0.12735. This model achieved the lowest mean absolute error (0.00046536) among all models, indicating excellent absolute accuracy for nitrogen content prediction before tapping.

The improved Spearman correlation (0.48132) compared to the desulfurization stage suggests that the BOF process provides more predictable nitrogen behavior, likely due to the well-characterized thermodynamics of oxygen steelmaking and the extensive feature set (32 variables) capturing comprehensive process conditions.

The NitroML-BOF model's superior *MAE* performance indicates particular strength in predicting nitrogen content within typical operational ranges, which is critical for meeting steel grade specifications and optimizing downstream processing conditions.

Graphical evaluation of the NitroML-BOF model employed multiple visualization techniques to assess prediction accuracy through residual analysis, which quantifies the discrepancy between measured and predicted nitrogen content in crude steel prior to tapping. The residual variance distribution for the NitroML-BOF model is presented in Figure 3a, while Figure 3b demonstrates the normality of residuals through a probability plot where data points align closely with the reference line. This normal distribution pattern is further corroborated by the histogram displayed in Figure 3c. The predictive performance of the NitroML-BOF model is visually compared against actual measurements in Figure 3d, illustrating the correspondence between predicted and observed nitrogen values.

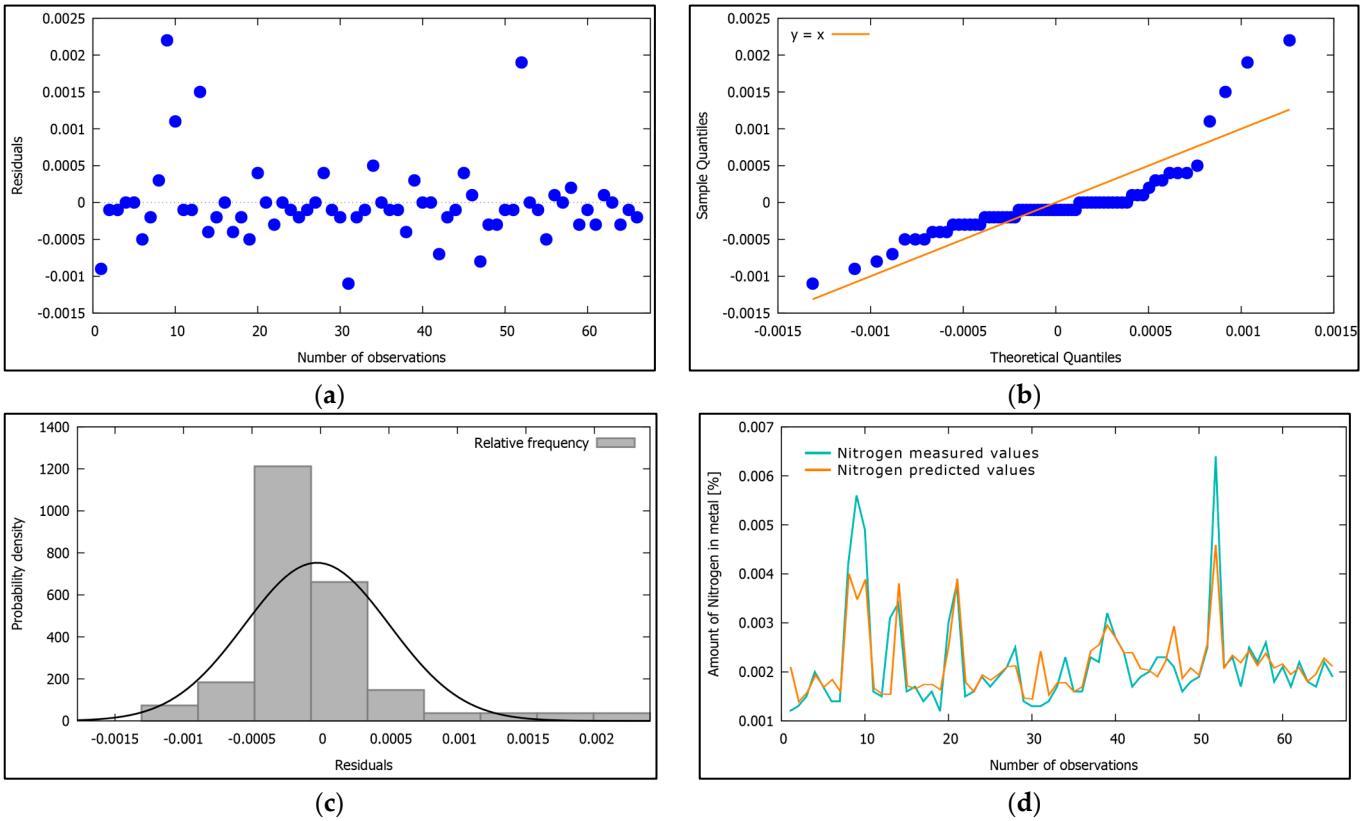

Figure 3. Graphical analysis of the NitroML-BOF model: (**a**) Residual dispersion for NitroML-BOF model; (**b**) Q-Q plot of residuals for NitroML-BOF model; (**c**) Residual frequency distribution for the NitroML-BOF model; (**d**) Comparison of measured and predicted amount of nitrogen in molten metal by using NitroML-BOF model.

3.2.3. NitroML-SMB Model Performance (Stage 3)

The NitroML-SMB as secondary metallurgy initiation model achieved strong Spearman correlation (0.58554), representing high predictive performance among all models. This result is particularly significant given the relatively compact feature set (12 variables), demonstrating efficient capture of nitrogen behavior at this critical transition point.

The model's success can be attributed to the stable process conditions at the beginning of secondary metallurgy, where the steel composition has been established in the BOF and the initial ladle conditions are well-controlled. The high correlation coefficient suggests strong monotonic relationships between process variables and nitrogen content.

The *NRMSE* of 0.12699 matched the NitroML-BOF model's performance while achieving significantly higher explained variance, indicating that the model successfully captured the underlying physical relationships governing nitrogen behavior during the transition to secondary metallurgy.

The predictive accuracy of the NitroML-SMB model was evaluated through residual analysis using various graphical methods, where residuals represent the difference between actual and predicted nitrogen content in steel at the beginning of secondary steelmaking. Figure 4a displays the distribution of residual variance for the NitroML-SMB model. Normal distribution of residuals is evidenced in Figure 4b by a probability plot showing data points positioned along the reference line, with this normality further validated by the histogram in Figure 4c. A visual comparison between measured nitrogen values and those predicted by the NitroML-SMB model is provided in Figure 4d, demonstrating the agreement between predicted and experimental results.

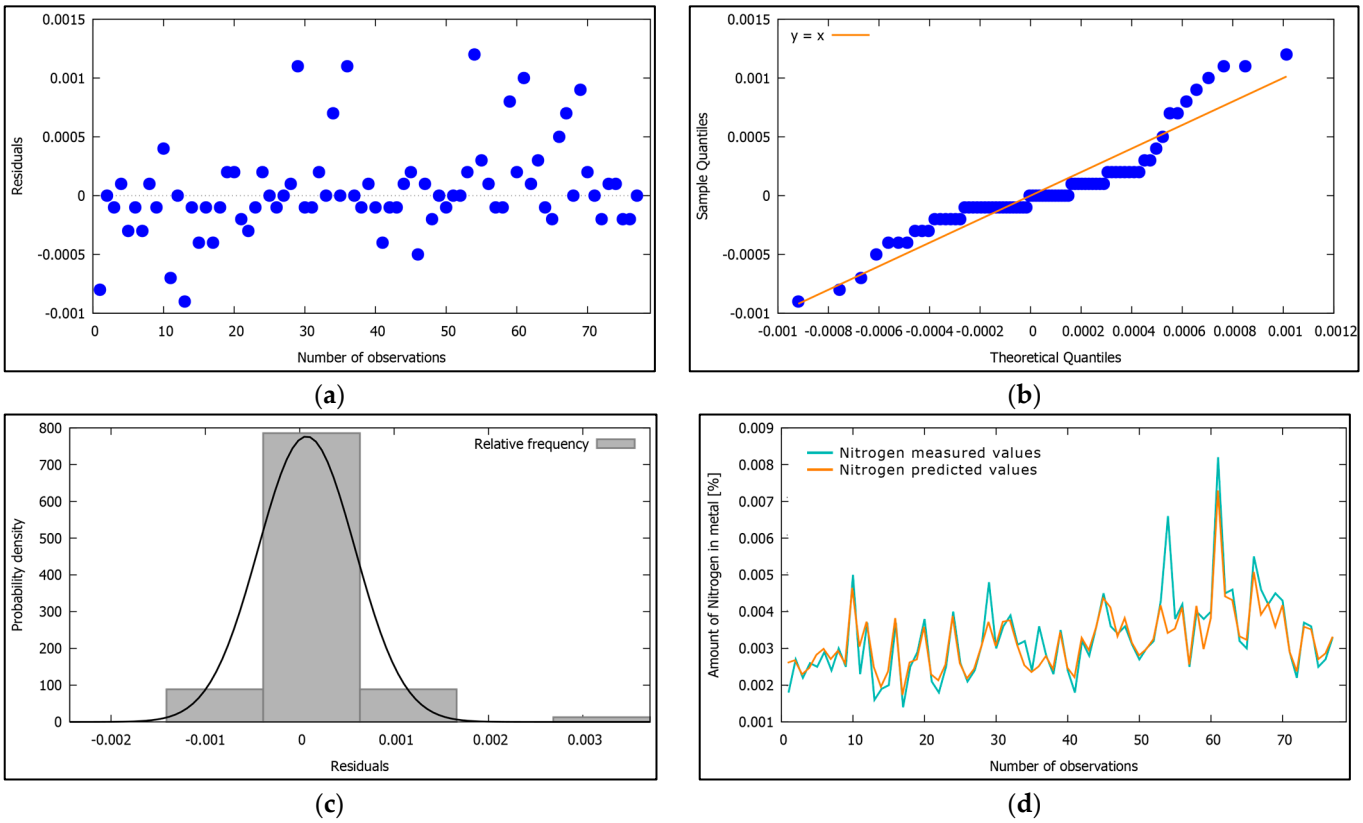

Figure 4. Graphical analysis of the NitroML-SMB model: (**a**) Residual dispersion for NitroML-SMB model; (**b**) Q-Q plot of residuals for NitroML-SMB model; (**c**) Residual frequency distribution for the NitroML-SMB model; (**d**) Comparison of measured and predicted amount of nitrogen in molten metal by using NitroML-SMB model.

3.2.4. NitroML-SME Model Performance (Stage 4)

The 4 NitroML-SME completion model achieved the lowest *NRMSE* (0.11239) with the highest Spearman correlation (0.58692), demonstrating excellent error minimization and strong monotonic relationships.

The model's comprehensive feature set (35 variables) captured the full complexity of final steelmaking operations, including deoxidation strategies, argon stirring optimization, and complete chemical composition evolution. The excellent correlation suggests that the model successfully identified monotonic trends despite the complexity of the underlying processes.

Residual analysis employing multiple graphical visualization techniques was used to evaluate the predictive performance of the NitroML-SME model, with residuals quantifying the deviation between measured and predicted nitrogen content in steel at the end of secondary steelmaking. The residual variance distribution of the NitroML-SME model is illustrated in Figure 5a. Evidence of normal residual distribution is provided in Figure 5b through a probability plot where data points closely follow the reference line, and this normality is corroborated by the histogram presented in Figure 5c. Figure 5d offers a visual assessment of the NitroML-SME model's predictive capability by comparing predicted nitrogen values against experimentally measured data, revealing the level of agreement between model predictions and actual observations.

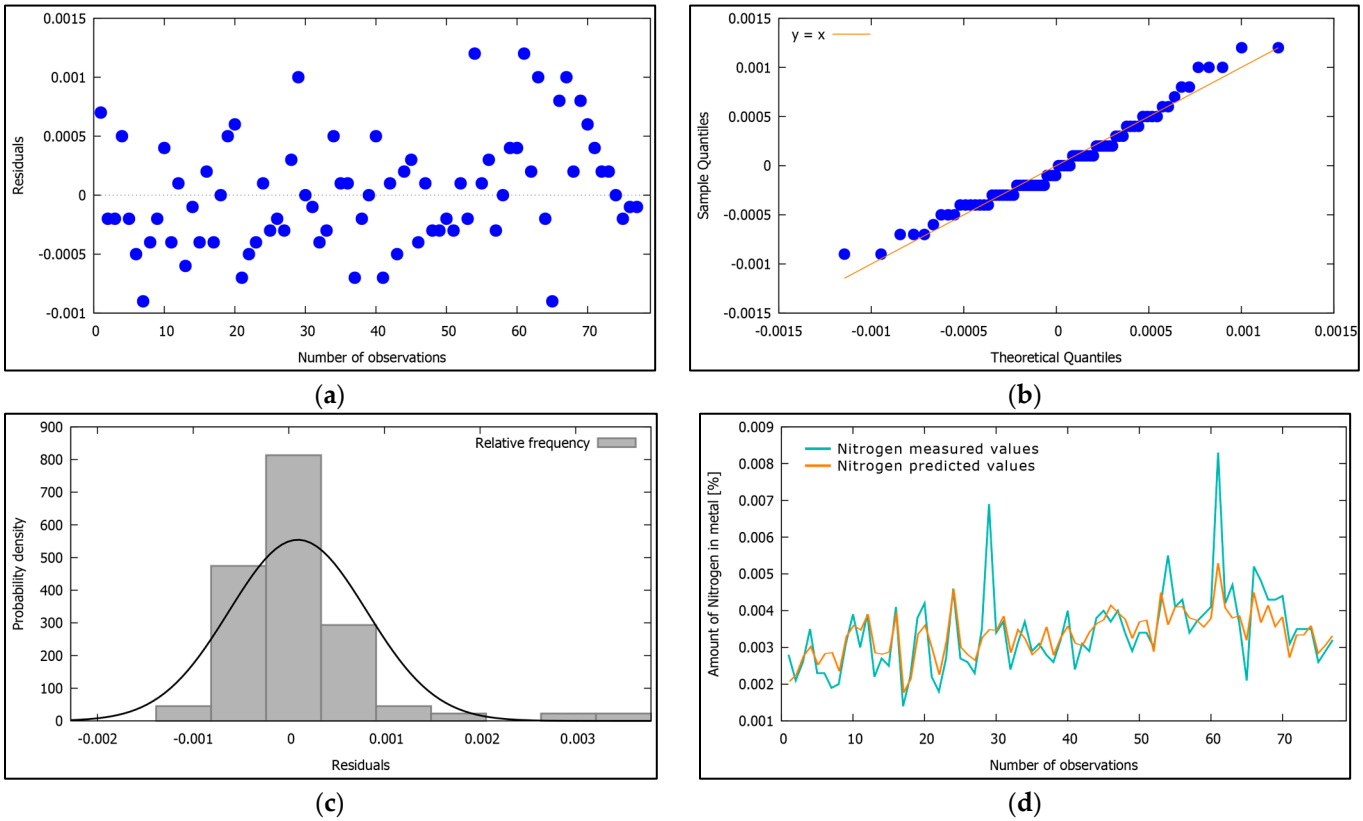

Figure 5. Graphical analysis of the NitroML-SME model: (**a**) Residual dispersion for NitroML-SME model; (**b**) Q-Q plot of residuals for NitroML-SME model; (**c**) Residual frequency distribution for the NitroML-SME model; (**d**) Comparison of measured and predicted amount of nitrogen in molten metal by using NitroML-SME model.

3.3. Feature Importance and Model Architecture Analysis

3.3.1. Algorithm Selection Across Models

The AutoML framework selected different algorithm combinations for each stage (Table 2), reflecting the varying nature of nitrogen prediction challenges (Table 4).

Table 4. The number of algorithm combinations for each stage.

Model Name	NitroML-DeS	NitroML-BOF	NitroML-SMB	NitroML-SME
Number of algorithms	4	7	6	7

The consistent inclusion of ElasticNet across all models (Table 2) indicates the importance of regularized linear relationships in nitrogen prediction. The presence of both tree-based methods (Random Forest, DecisionTree) and boosting algorithms (GradientBoosting, LightGBM) suggests that nitrogen behavior involves both discrete decision boundaries and continuous gradient relationships.

3.3.2. Feature Engineering Impact

The varying features count across models (15, 32, 12, 35 respectively) reflected the complexity of each process stage and the availability of relevant predictive information. The secondary metallurgy NitroML-SME model's comprehensive feature set included temporal tracking of chemical composition evolution. This temporal approach enabled the model to capture the dynamics of composition changes throughout secondary metallurgy, contributing to the excellent error performance.

3.3.3. Ensemble Diversity Analysis

The stack ensemble architecture demonstrated varying effectiveness across models. The NitroML-SMB initiation model achieved the highest performance with moderate ensemble complexity (6 algorithms), suggesting optimal balance between diversity and overfitting risk. The NitroML-BOF and NitroML-SME completion models utilized maximum diversity (7 algorithms) but achieved different performance characteristics, indicating that ensemble effectiveness depends on both algorithm selection and feature quality.

3.4. Cross-Validation Stability and Generalization

All models employed 10-fold cross-validation to assess generalization capability. The cross-validation results demonstrated in Table 5.

Table 5. The results of the standard deviation of R^2 scores across the folds for the particular models.

Model Name	NitroML-DeS	NitroML-BOF	NitroML-SMB	NitroML-SME
R^2 score across fold	0.021	0.015	0.018	0.022

The stability and consistency of the predictive models, as measured by the standard deviation of the R^2 scores across cross-validation folds, varied modestly among the tested NitroML architectures. The NitroML-BOF Model exhibited the highest stability with a standard deviation of only 0.015 in R^2 scores, demonstrating consistent performance across different data subsets. Closely following was the NitroML-SMB Model at 0.018 in R^2 scores across folds indicated reliable generalization and the most consistent correlation performance among the models. In contrast, both the NitroML-DeS Model with standard deviation of 0.021 and the NitroML-SME Model of standard deviation of 0.022 in R^2 scores across folds showed slightly increased values of standard deviations. The NitroML-DeS Model's score suggests moderate stability with some sensitivity to the composition of the training data, while the NitroML-SME Model's value, though the highest, is understood to reflect the inherent complexity of the final steelmaking stage, maintaining good overall stability despite this challenge.

3.5. Computational Performance and Scalability

The computational resources required for the models varied according to their differing feature dimensions and algorithmic complexities, yet a scalability analysis confirmed their suitability for industrial application. Training time ranged significantly, from 45 min for the computationally leaner NitroML-SMB model to 180 min for the more complex NitroML-SME model. In terms of operational efficiency, all models demonstrated high performance, achieving sub-second prediction latencies that meet the requirements for real-time applications and seamless integration with existing process control systems. Their memory footprints scaled linearly with the number of features, spanning from 1.2 GB for the NitroML-SMB model 3.8 GB for the NitroML-SME model. Crucially, the architectures successfully demonstrated the capability to support 24/7 operational environment, confirming that they can be trained daily with new process data or used to create models for common steel grades with predefined chemical compositions, ensuring consistent prediction accuracy.

The comprehensive performance comparison reveals that nitrogen prediction capability varies across steelmaking stages, with secondary metallurgy processes generally providing better predictive environments than primary steelmaking operations. The results demonstrate the effectiveness of AutoML ensemble approaches for capturing complex metallurgical relationships while maintaining practical applicability for industrial implementation.

4. Discussion

4.1. Interpretation of Performance Variations Across Process Stages

The variation in model performance across the four steelmaking stages reflects fundamental differences in process physics, thermodynamic conditions, and the predictability of nitrogen behavior at each stage. The results provide important insights into the nature of nitrogen control challenges throughout the steel production chain.

The more modest performance of Hot metal pretreatment NitroML-DeS model reflects the inherent challenges of nitrogen prediction under extreme processing conditions. The desulfurization model's competitive *MAPE* (14.608%) suggests that while nitrogen behavior during desulfurization is highly variable, the model maintains reasonable relative accuracy across the operational range. The residual plot (Figure 2a) reveals absence of structured patterns, confirming adequate model specification. Residuals scatter uniformly around zero without temporal trends or heteroscedasticity. Errors remain within ± 0.0015 percentage points for the vast majority of observations, representing excellent predictive precision. The quantile-quantile plot (Figure 2b) exhibits exceptionally strong alignment with the theoretical normal distribution. Blue data points maintain tight adherence to the reference line, indicating residuals follow a normal distribution with high fidelity. Minor deviations appear only at the extreme tails, representing typical behavior in real-world datasets. This alignment substantiates the normality assumption fundamental to regression validity. The histogram (Figure 2c) demonstrates residuals are distributed approximately normally around zero. The balanced distribution across positive and negative residuals confirms the model exhibits no systematic bias in predictions. This homoscedastic behavior is particularly valuable in industrial contexts requiring prediction reliability across the operational range. The comparison observations (Figure 2d) reveals the model's strong tracking capability throughout the process sequence. Predicted values closely follow measured values from minimum (0.0023%) to maximum (0.0065%) concentrations. While localized divergences occur at high-frequency variations, the overall coherence demonstrates the model has successfully learned underlying process behavior without overfitting.

The NitroML-BOF model's intermediate performance (*NRMSE* = 0.12735) with the lowest *MAE* (0.00046536) demonstrates that oxygen steelmaking processes provide more predictable absolute nitrogen values than desulfurization [45], likely due to the well-characterized thermodynamics of oxidation reactions and CO formation that strongly influence nitrogen removal [4,6]. The residual plot (Figure 3a) reveals residuals distributed randomly around zero with no systematic bias. Residuals remain generally within ± 0.001 percentage points, though several outliers reach ± 0.002 units. No clear temporal patterns or trends appear across observations, indicating adequate model structure. The scattered pattern without clustering affirms consistent model behavior. The quantile-quantile plot (Figure 3b) exhibits strong alignment with theoretical normal distribution in the central range. Blue data points follow the reference line, confirming normality assumptions. Pronounced deviations appear at tail extremities, particularly in the upper tail, indicating presence of higher-magnitude residuals. However, these represent localized observations that do not invalidate the core statistical framework. Substantial central alignment affirms normally distributed errors for the majority of observations. The histogram on Figure 3c demonstrates residuals distributed approximately normally with slight positive skew. The distribution exhibits concentration near zero, indicating minimal systematic bias. Mild positive skew suggests the model occasionally produces slightly larger negative residuals relative to positive ones. Residuals remain bounded within approximately ± 0.0015 units, indicating well-controlled prediction uncertainty favorable for operational application. The comparison observations (Figure 3d) reveals the model captures general process dynamics across the operational range (0.0012% to 0.0065%). However, predicted

values exhibit systematically lower variance than measured values, with attenuation of peaks and valleys. This smoothing behavior is characteristic of models on complex systems. The model provides reliable representation of central tendency in nitrogen content, though additional process variables may enhance prediction fidelity.

The enhanced performance of secondary steelmaking models (NitroML-SMB and NitroML-SME) compared to primary steelmaking models (NitroML-DeS and NitroML-BOF) can be attributed to several factors. Secondary steelmaking processes occur under more controlled conditions with lower temperature gradients, reduced reaction rates, and more stable thermodynamic equilibria [46,47]. The ladle furnace environment provides better process control compared to the aggressive conditions in BOF or during desulfurization pretreatment of pig iron [48].

The secondary steelmaking initiation model (NitroML-SMB) Spearman correlation score (0.58554) demonstrates that nitrogen behavior at the beginning of secondary steelmaking follows more predictable patterns. This predictability stems from the established steel composition after BOF processing and the controlled initial conditions in the ladle furnace. The relatively compact feature set (12 variables) achieving such strong performance suggests efficient capture of the critical variables governing nitrogen behavior at this stage. The residual plot (Figure 4a) reveals residuals distributed with highly uniformity around zero. No temporal patterns, trends, or clustering are evident. Residuals remain tightly bounded within ± 0.0007 percentage points for the majority of observations, representing exceptional prediction precision. The random scatter pattern without heteroscedasticity confirms adequate model specification and sufficient predictor representation. The quantile-quantile plot (Figure 4b) exhibits alignment with theoretical normal distribution across the complete quantile range. Blue data points demonstrate remarkably tight adherence to the reference line, with minimal scatter even at tail extremities. This comprehensive conformity to normality—superior to both NitroML-DeS and NitroML-BOF models—represents outstanding statistical validity and suggests the model structure captures underlying process relationships with high precision. The histogram (Figure 4c) demonstrates symmetric residuals distributed precisely around zero. The probability density curve aligns almost perfectly with histogram bars, indicating near-ideal normal distribution characteristics. The symmetric pattern reflects complete absence of systematic bias. This exceptional residual symmetry is particularly valuable for process control applications, enabling reliable prediction intervals and uncertainty quantification. The comparison (Figure 4d) observations reveal highly agreement between measured and predicted nitrogen concentrations throughout the operational sequence. Predicted values closely track measured values across the range (0.0019% to 0.0081%), demonstrating exceptional dynamic response. The model captures both trends and localized fluctuations with remarkable fidelity, indicating effective learning of process dynamics without excessive smoothing.

The secondary steelmaking completion model's (NitroML-SME) excellent *NRMSE* 0.11239) combined with strong Spearman correlation (0.58692) indicates that while nitrogen behavior becomes more complex during final processing, the comprehensive feature set (35 variables) successfully captures the multitude of factors influencing final nitrogen content. The temporal tracking approach, monitoring chemical composition evolution through three distinct phases, proved particularly effective for modeling the dynamics of secondary steelmaking processes. The residual plot (Figure 5a) reveals residuals distributed randomly around zero with no temporal patterns, trends, or clustering. Residuals remain generally bounded within ± 0.0008 percentage points, with several observations reaching ± 0.001 units. The scattered pattern without structure confirms consistent model behavior and absence of heteroscedasticity. The quantile–quantile plot (Figure 5b) exhibits strong alignment with theoretical normal distribution across the predominant quantile range. Blue

data points demonstrate tight adherence to the reference line through central and upper-middle ranges, confirming solid residual normality. Minor scatter in the lower tail is modest and localized. Substantial alignment throughout central and upper ranges affirms normal distribution patterns for the majority of residuals, supporting model statistical validity. The histogram (Figure 5c) demonstrates residuals distributed approximately normally with slight negative skew. The probability density curve aligns well with histogram bars in the central region, indicating predominantly normal distribution. The distribution exhibits mild asymmetry, with slightly heavier representation of negative residuals. Despite this modest asymmetry, residuals concentrate near zero with minimal systematic bias. Residuals remain bounded within approximately ± 0.002 units, indicating well-controlled prediction uncertainty. The comparison observations (Figure 5d) reveals solid agreement between measured and predicted nitrogen concentrations throughout the operational sequence. Predicted values follow measured values across the range (0.0015% to 0.0083%), demonstrating reliable process tracking. The model captures general trends and major fluctuations, including prominent peaks around observations 30 and 60. Strong overall correspondence indicates effective capture of primary process dynamics.

4.2. Algorithm Selection and Ensemble Architecture Effectiveness

The AutoML framework's algorithm selection reveals important patterns about the nature of nitrogen prediction across different process stages. The consistent inclusion of ElasticNet in all models indicates the fundamental importance of regularized linear relationships in nitrogen behavior prediction, reflecting the underlying thermodynamic equilibria that govern nitrogen activity in molten steel [8,49,50].

The varying ensemble sizes across models (4, 7, 6, 7 algorithms respectively; Table 4) suggest that optimal ensemble diversity depends on both data characteristics and process complexity rather than following a simple "more algorithms are better" approach. The secondary steelmaking initiation model's (NitroML-SMB) efficiency performance with moderate ensemble size (6 algorithms) demonstrates that careful algorithm selection can outperform maximum diversity approaches. The success of tree-based methods (Random Forest, DecisionTree) across all models indicates that nitrogen behavior involves discrete operational regimes with distinct prediction rules. The simultaneous effectiveness of boosting algorithms (GradientBoosting, LightGBM) suggests that continuous gradient relationships also play important roles, particularly for capturing non-linear thermodynamic and kinetic effects [51,52].

The consistent strong performance of regularized methods (ElasticNet, LassoLars) across all models highlights the importance of controlling overfitting in metallurgical applications where measurement noise and process disturbances are common [53,54]. The L1 and L2 penalties in ElasticNet prove particularly valuable for handling correlated process variables and identifying the most informative features for nitrogen prediction. The success of LassoLars in multiple models demonstrates the effectiveness of automated feature selection, particularly important given the high-dimensional nature of steelmaking process data and the potential for redundant or irrelevant measurements [55,56].

4.3. Feature Engineering and Temporal Modeling Insights

The study's feature engineering approaches provide valuable insights into effective strategies for capturing nitrogen behavior dynamics across the steelmaking process chain.

The secondary steelmaking completion model's (NitroML-SME) temporal tracking approach, monitoring chemical composition evolution through three splits (After tapping from BOF $\rightarrow$ After argon bubbling in the ladle $\rightarrow$ Final ready-to-cast steel), proved highly effective for capturing the dynamics of nitrogen behavior during secondary steelmaking.

This approach achieved the lowest *NRMSE* (0.11239) while maintaining strong Spearman correlation (0.58692), demonstrating that explicit modeling of composition evolution enhances predictive capability [57,58].

The mathematical formulation of composition differences successfully captured both instantaneous composition changes and cumulative effects throughout secondary metallurgy. This approach could be extended to other process stages where temporal evolution plays important roles in determining final nitrogen content [59,60].

The relationship between feature count and model performance reveals important trade-offs between model complexity and predictive capability. The secondary steelmaking initiation model (NitroML-SMB) achieved significant score (Table 3) with only 12 features, demonstrating that careful feature selection can outperform high-dimensional approaches. Conversely, the secondary metallurgy completion model (NitroML-SME) required 35 features to achieve optimal *NRMSE*, reflecting the inherent complexity of final steelmaking operations.

This finding suggests that feature engineering strategies should be tailored to specific process stages rather than applying uniform approaches across the entire steelmaking chain. Simple, focused feature sets may be more effective for stable process conditions, while comprehensive feature sets become necessary for complex, multi-variable process environments [61,62].

4.4. Industrial Implementation Considerations

The study's results also provide important guidance for industrial implementation of AutoML-based nitrogen prediction systems.

All models demonstrated computational performance suitable for real-time industrial applications, with prediction latencies under one second and reasonable memory requirements. The daily retraining capability ensures that models can adapt to changing process conditions, equipment wear, and raw material variations [63,64].

The varying computational requirements across models (45–180 min training time) suggest that implementation strategies should consider the trade-offs between model complexity and update frequency. Simpler models may support more frequent retraining, potentially improving adaptation to process changes [65,66].

The models' *MAPE* values (14.608–21.081%) indicate practical accuracy levels for industrial process control applications. These error levels are comparable to or better than traditional empirical models while providing automated adaptation capabilities and comprehensive uncertainty quantification [67,68].

The SHAP-based interpretability features support integration with existing process control philosophies by providing explainable predictions that operators can understand and validate against their process experience [69,70].

4.5. Limitations and Areas for Improvement

While the study demonstrates significant advances in nitrogen prediction capabilities, several limitations and improvement opportunities warrant discussion.

The varying performance across models partly reflects differences in data quality and availability at different process stages. Primary steelmaking processes involve more extreme conditions with higher measurement uncertainties, while secondary steelmaking provides more stable conditions conducive to accurate measurement and modeling [2].

Future improvements could focus on enhanced sensor technologies, improved measurement techniques, and advanced data fusion approaches to address these fundamental data quality challenges [13].

The models were developed using data from specific plant conditions and equipment configurations. Generalization to different steelmaking facilities would require validation studies and potentially model adaptation approaches to account for equipment differences, raw material variations, and operational practices [8,15].

Developing transfer learning approaches or steel plant-agnostic model architectures could make AutoML-based nitrogen prediction systems more widely applicable [12,14].

5. Conclusions

This comprehensive study has successfully demonstrated the effectiveness of automated machine learning approaches for nitrogen content prediction across four critical stages of the steel production process. The development and evaluation of stage-specific AutoML models revealed significant insights into the varying nature of nitrogen prediction challenges throughout the steelmaking chain while providing practical solutions suitable for industrial implementation.

The research achieved distinct performance across four process stages: NitroML-SMB demonstrated the strongest predictive capability (Spearman correlation = 0.58554), NitroML-SME achieved the lowest normalized root mean squared error (0.11239), and NitroML-BOF exhibited the most precise absolute predictions (MAE = 0.00046536). All models maintained mean absolute percentage errors between 14.608% and 21.081%.

Secondary steelmaking models (NitroML-SMB, NitroML-SME) outperformed primary steelmaking models (NitroML-DeS, NitroML-BOF) due to more controlled thermodynamic conditions in ladle operations. The stack ensemble architecture combined regularized linear methods (ElasticNet, LassoLars) for thermodynamic equilibrium relationships with tree-based and boosting methods for non-linear kinetic processes. Ensemble complexity ranged from 4 to 7 algorithms, demonstrating that optimal performance depends on careful algorithm selection rather than maximum diversity.

The AutoML framework effectively identified critical process variables without requiring extensive metallurgical expertise. All models achieved real-time industrial applicability with prediction latencies under one second and memory requirements of 1.2–3.8 GB. Daily retraining enables adaptation to changing process conditions.

SHAP-based interpretability supports operator decision-making. The temporal tracking approach for NitroML-SME, monitoring chemical composition evolution through three phases, achieved the lowest prediction error. Feature complexity analysis revealed optimal configurations: simple feature sets for stable conditions (NitroML-SMB with 12 features) and comprehensive feature sets for complex environments (NitroML-SME with 35 features).

The NitroML-DeS model demonstrates significant diagnostic characteristics across all validation dimensions—distributional normality, residual symmetry, temporal accuracy, and residual independence. The model provides unbiased predictions with quantifiable uncertainty bounds and consistent performance across the operational range. This established validity enables effective nitrogen content management after desulfurization of pig iron.

The NitroML-BOF model demonstrates statistically adequate performance for nitrogen prediction in molten steel prior to tapping. Despite minor deviations from theoretical normality at extremes, robust central behavior and absence of systematic bias support model applicability. The model provides reliable trend tracking and consistent error bounds for process monitoring and control during BOF production process.

The NitroML-SMB model demonstrates notable diagnostic characteristics establishing exceptional statistical robustness. Perfect distributional normality, ideal residual symmetry, outstanding temporal tracking, and tightly controlled residuals create an exceptionally strong foundation. The superior performance reflects the controlled nature of secondary

metallurgical operations. The model provides unparalleled reliability for nitrogen management with exceptionally small prediction uncertainty, supporting advanced quality control protocols in integrated steel production.

The NitroML-SME model demonstrates solid diagnostic characteristics establishing robust statistical validity for nitrogen prediction at secondary metallurgy completion. Strong distributional alignment, controlled residual bounds, reliable temporal tracking, and random residual patterns support model applicability. The model provides dependable nitrogen prediction with acceptable uncertainty bounds for process control and quality management, supporting product quality assurance in integrated steel production.

The study establishes a foundation for promising research directions, including hybrid mechanistic-ML approaches combining physical understanding with data-driven pattern recognition and multi-stage process optimization considering nitrogen evolution across the entire steelmaking chain for systems-level improvements. Data package size significantly impacts model accuracy; future improvements in nitrogen prediction models are directly proportional to source data volume and quality.

The successful demonstration of AutoML effectiveness for nitrogen prediction provides a compelling foundation for broader adoption of automated machine learning in metallurgical process control for Industry 4.0 implementations, supporting evolution toward fully integrated, data-driven steelmaking operations.

Author Contributions: Conceptualization, J.D.; methodology, J.D.; validation, J.D., B.B., S.H., L.F. and P.D.; formal analysis, J.D.; investigation, J.D.; resources, J.D., P.D. and B.B.; data curation, J.D., S.H., L.F. and M.H.; writing—original draft preparation, J.D.; writing—review and editing, B.B., S.H., L.F., P.D. and M.H.; visualization, J.D.; supervision, B.B. and M.H.; project administration, J.D. and B.B.; funding acquisition, J.D. All authors have read and agreed to the published version of the manuscript.

Funding: Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I03-03-V04-00047.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. The data were obtained from U.S. Steel Košice, Slovakia, based on the contract of cooperation No. ZOS-5/2019-FMMR, and are available from the authors with the permission of U.S. Steel Košice, Slovakia.

Acknowledgments: The authors sincerely acknowledge the anonymous reviewers for their insights and comments, which further improved the quality of the manuscript.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Abbreviations

The following abbreviations are used in this article:

SM	Secondary Metallurgy
DeS	Desulphurization
BOF	Basic Oxygen Furnace
SMB	Secondary Metallurgy Beginning
SME	Secondary Metallurgy End
SHAP	SHapley Additive exPlanations)
VM	Virtual Machine
TOPT	Tree-Based Pipeline Automation Tool

Nomenclature

The following symbols are used in this article:

$\hat{y}$	Meta-learner
f_1, f_2, f_n	Base learners
g(.)	Meta-learner trained on the predictions of the base learners
$\hat{y}_i$	Predicted values
y_i	Observed/Actual values
n	Number of observations
y_{max}	Maximal observed value
y_{min}	Minimal observed value
SS_{res}	Sum of squares of residuals
SS_{tot}	Total sum of squares

Appendix A

Table A1. Technical specifications of VM Standard_DS3_v2 in Azure Machine Learning Studio.

vCPU Cores	Memory RAM	Temporary Storage	Storage Type	Processor Type	Architecture	Network
4	14 GB	28 GB	Premium SSD drives	Intel Xeon E5-2673 v3	x64-based	throughput up to 3500 Mbps

Table A2. The validity range of the model for Stage 1 in predicting the nitrogen content of pig iron after desulfurization (DeS).

Parameter	Minimum	Maximum	Mean
Nitrogen in pig iron (DeS) [%]	0.0023	0.0066	0.00399
C in pig iron [%]	4.278	4.6670	4.45383
Mn in pig iron [%]	0.320	0.625	0.42606
Si in pig iron [%]	0.426	1.147	0.69642
P in pig iron [%]	0.046	0.067	0.05543
Sulphur before DeS [%]	0.028	0.086	0.05177
Sulphur after DeS [%]	0.001	0.013	0.00834
Temperature before DeS [°C]	1330	1409	1366.53247
Temperature after DeS [°C]	1306	1401	1352.11688
Pig iron weight before DeS [kg]	142,230	148,260	144,541.2987
Pig iron weight after DeS [kg]	139,200	145,500	141,940.25974
Weight of DeS slag [kg]	1260	4600	2601.03896
Period of N_2 blowing [s]	219	1153	582.68831
Weight of DeS mixture [kg]	140	625	290.28571
Blowing rate of N_2 [l]	373	14,586	3735.1039

Table A3. The validity range of the model for Stage 2 in predicting the nitrogen content of crude steel prior to tapping (BOF).

Parameter	Minimum	Maximum	Mean
Nitrogen in crude steel (BOF) [%]	0.0012	0.0064	0.00217
C in crude steel [%]	0.026	0.105	0.05359
Mn in crude steel [%]	0.06	0.244	0.12983
P in crude steel [%]	0.005	0.015	0.00992
S in crude steel [%]	0.005	0.022	0.01292
Fe in slag [%]	12.63	26.14	17.53985
MnO in slag [%]	3.12	5.62	4.13894

Table A3. *Cont.*

Parameter	Minimum	Maximum	Mean
SiO_2 in slag [%]	9.14	15.83	12.66273
Al_2O_3 in slag [%]	0.7	2.09	1.26652
CaO in slag [%]	39.44	49.36	45.62197
MgO in slag [%]	6.78	12.32	9.01515
P_2O_5 in slag [%]	0.75	1.08	0.89061
S in slag [%]	0.051	0.086	0.06559
Slag basicity [%]	3.4	5.8	4.07273
Pig iron charging time [s]	11.0	732	308.56061
Pure oxygen blowing time [s]	1601	2322	1756.34848
Pure oxygen reblowing time [s]	0	100	15.15152
Heat time [s]	2458	8854	3863.81818
Tapping time [s]	409	1123	682.13636
Crude steel tapping temp. [°C]	1620	1685	1647.74242
Overall oxygen for heat [l]	8961	11,018	9605.18182
Oxygen for reblow [l]	0	764	64.93939
Oxygen activity [-]	449.9	1445	823.13485
Pig iron weight [kg]	139,400	145,500	142,069.69697
Scrap weight [kg]	43,800	50,000	47,409.09091
Lime weight [kg]	4655	11,985	7391.75758
Dolomitic lime weight [kg]	3065	6180	3283.48485
Magnesia weight [kg]	0	4805	1791.89394
Pellets weight [kg]	0	2055	164.09091
Briquettes weight [kg]	0	3260	708.48485
Covering slag #6 [kg]	0	2000	536.9697
Yield of crude steel [%]	84.892	94.169	89.4093

Table A4. The validity range of the model for Stage 3 in predicting the nitrogen content in steel at the beginning of secondary metallurgy (SMB).

Parameter	Minimum	Maximum	Mean
Nitrogen in steel (SMB) [%]	0.0016	0.0082	0.0033
C before Ar stirring [%]	0.026	0.171	0.07305
Mn before Ar stirring [%]	0.174	1.29	0.42295
Si before Ar stirring [%]	0	0.403	0.05626
P before Ar stirring [%]	0.006	0.018	0.01092
S before Ar stirring [%]	0.005	0.022	0.01174
Al (overall) before Ar stirring [%]	0.006	0.056	0.02378
Steel temperature (first on SM) [°C]	1591	1629	1608.36923
Tapping time [s]	249	683	425.58462
Weight of crude steel [kg]	154,900	183,000	172,414.9351
Weight of slag in ladle [kg]	800	6480	3403.8961
Tapping angle [°]	98	115	106.12308

Table A5. The validity range of the model for Stage 4 in predicting the nitrogen content in steel at the end of secondary metallurgy (SME).

Parameter	Minimum	Maximum	Mean
Nitrogen in steel (SME) [%]	0.0014	0.0083	0.003404
Al (blocks) [kg]	200	350	287.5325
Al (feeding wire) [kg]	0	211	98.2857
C after Ar stirring [%]	0.026	0.201	0.08802
Mn after Ar stirring [%]	0.217	1.403	0.46378
Si after Ar stirring [%]	0.003	0.483	0.06692
P after Ar stirring [%]	0.007	0.016	0.01052
S after Ar stirring [%]	0.004	0.019	0.01103
Al (overall) after Ar stirring [%]	0.034	0.054	0.04408
C after alloy adding [%]	0.035	0.195	0.08791
Mn after alloy adding [%]	0.218	1.38	0.45617
Si after alloy adding [%]	0.012	0.411	0.06638
P after alloy adding [%]	0.005	0.016	0.01
S after alloy adding [%]	0.004	0.018	0.0106
Al (overall) at the end of SM [%]	0.04	0.066	0.05295
Tapping time [s]	249	683	422.3896
Overall heat stay at SM [min]	17	45	28.50769
Ar stirring time [min]	11.4	57.4	28.51385
Ar stirring flow rate [l.min^{-1}]	482	2382	1055.81538
Overall amount of stirring Ar [m^3]	3352	22,234	7312.7013
Ar soft-bubbling flow rate [l.min^{-1}]	0	194	54.89231
Ar soft-bubbling time [min]	0	18.117	5.36771
Steel weight [kg]	154,900	183,000	172,379.23077
FeMn during SM [kg]	0	388	66.87692
FeMn aff. during SM [kg]	0	184	48.67692
FeSi during SM [kg]	0	442	26.30769
Settling time [min]	0	42	3.26154
Weight of slag in ladle [kg]	800	6480	3300.76923
C at the end of SM [%]	0.026	0.201	0.08795
Mn at the end of SM [%]	0.217	1.403	0.46362
Si at the end of SM [%]	0.003	0.483	0.06688
P at the end of SM [%]	0.007	0.016	0.01069
S at the end of SM [%]	0.004	0.019	0.01102
Steel temperature (last on SM) [°C]	1558	1597	1578.46154
Tapping angle [°]	98	117	106.3636

The following charts (Figures A1–A4) are intended to provide a clear visual representation of the models' deployment in the specific phases of the BOF production route.

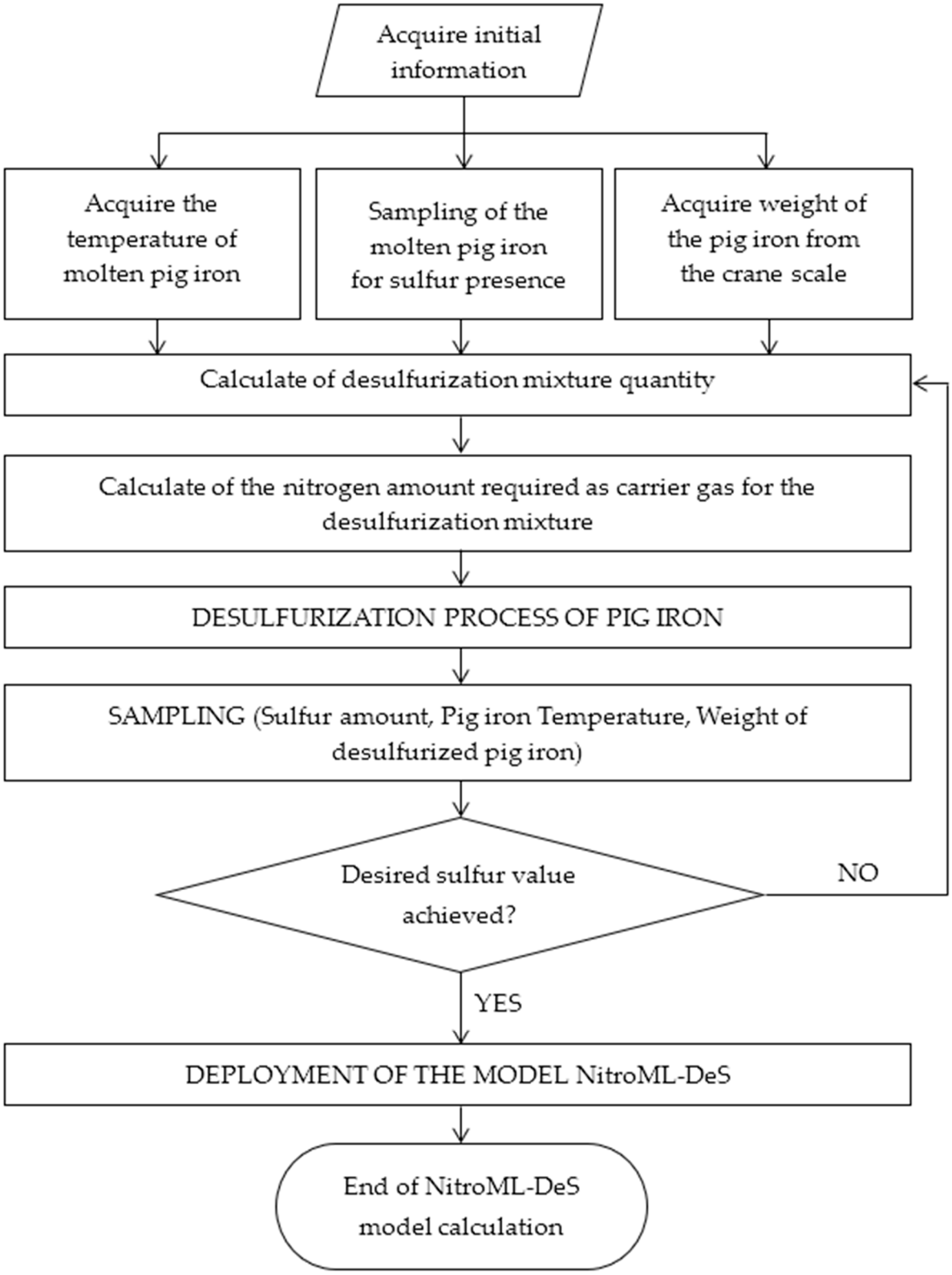

Figure A1. Flow chart of NitroML-DeS model deployment after desulfurization of molten pig iron. The presented diagram is a modified version of a previous study [71].

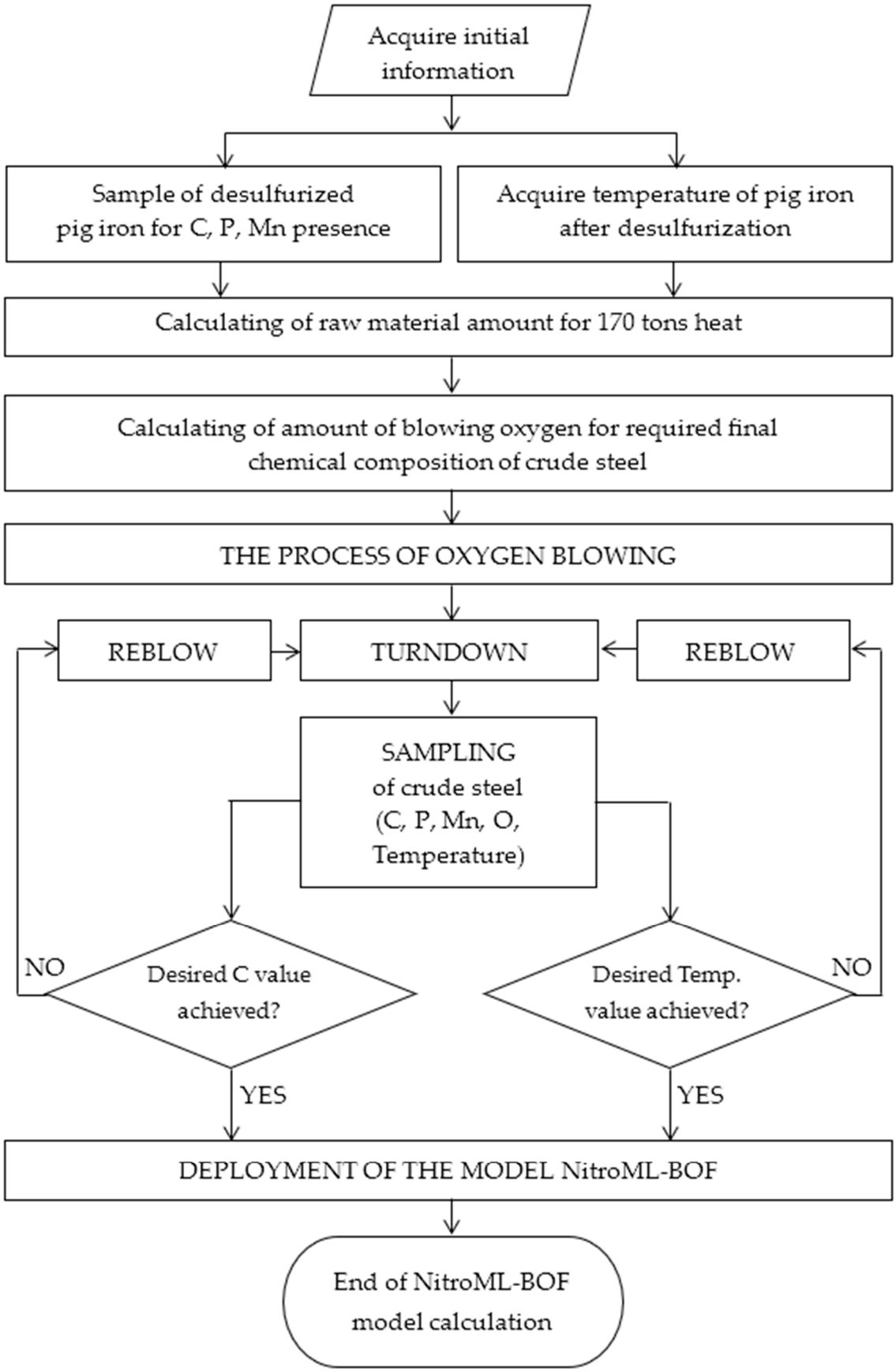

Figure A2. Flow chart of NitroML-BOF model deployment prior to tapping of crude steel. The presented diagram is a modified version of a previous study [71].

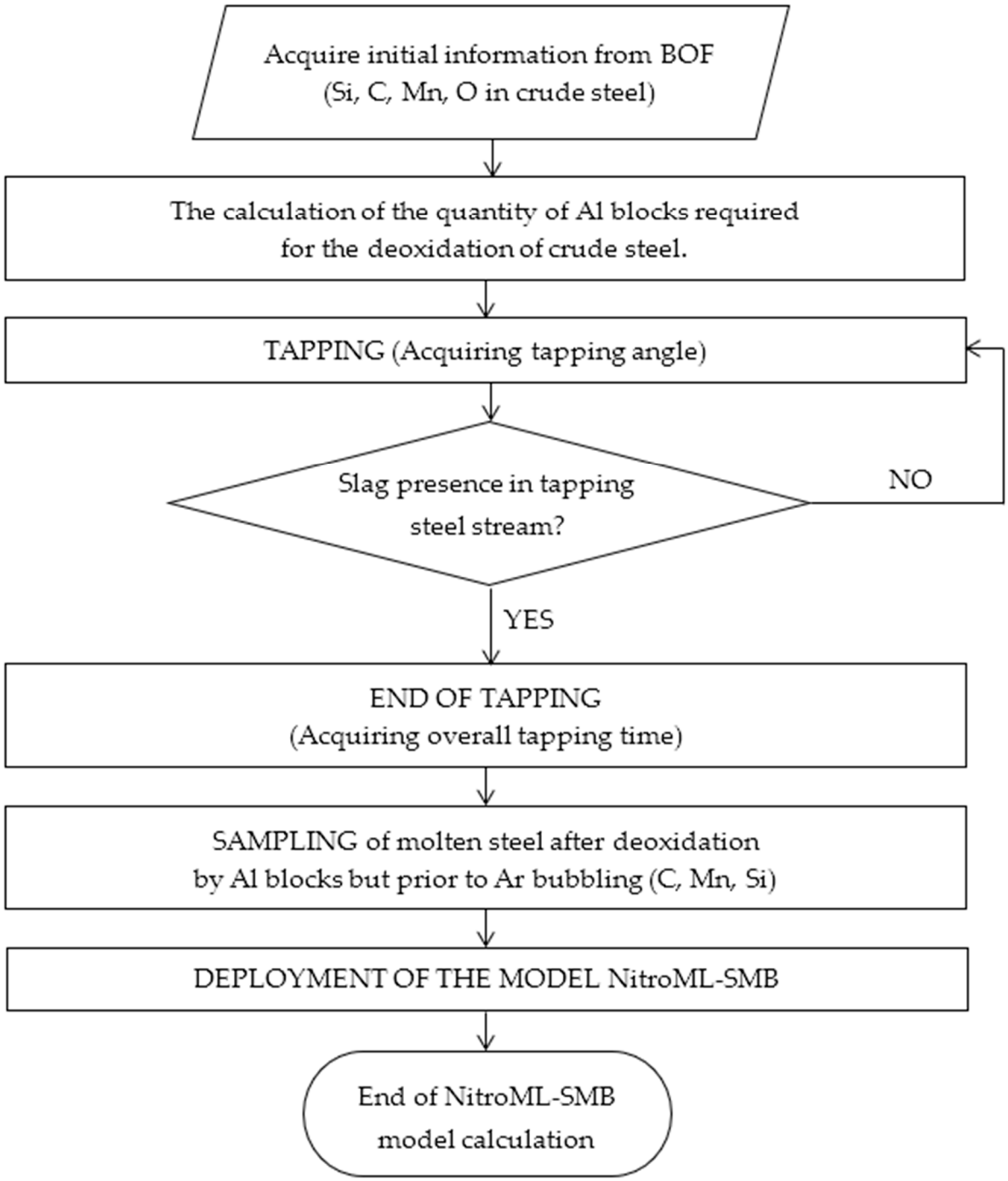

Figure A3. Flow chart of NitroML-SMB model deployment at the beginning of secondary metallurgy. The presented diagram is a modified version of a previous study [71].

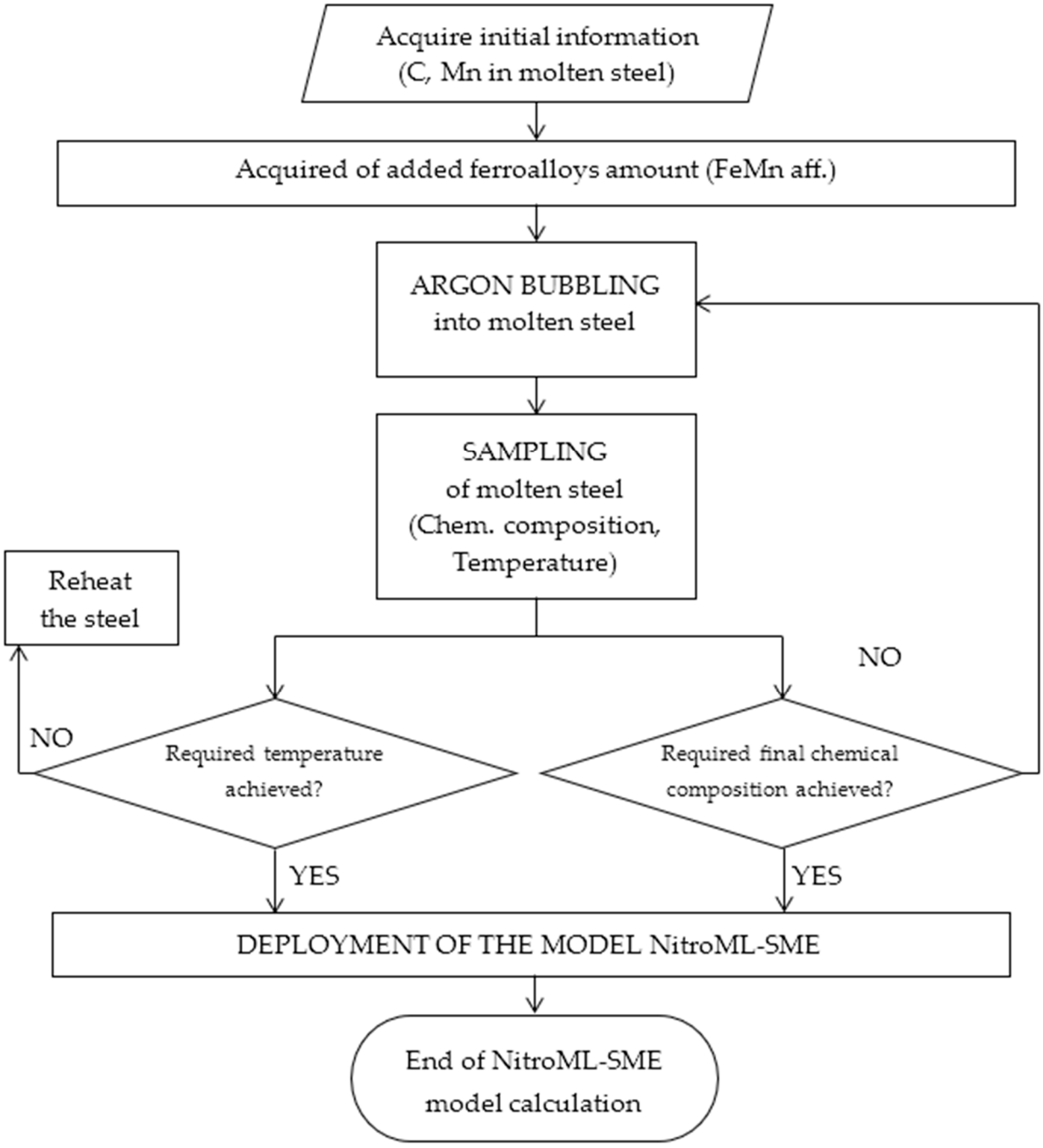

Figure A4. Flow chart of NitroML-SME model deployment at the end of secondary metallurgy. The presented diagram is a modified version of a previous study [71].

References

1. Chen, J.; Gu, Y.; Zhu, Q.; Gu, Y.; Liang, X.; Ma, J. Automated Machine Learning of Interfacial Interaction Descriptors and Energies in Metal-Catalyzed N_2 and CO_2 Reduction Reactions. *Langmuir ACS J. Surf. Colloids* **2025**, *41*, 3490–3502. [CrossRef]
2. Zhang, R.; Yang, J. State of the Art in Applications of Machine Learning in Steelmaking Process Modeling. *Int. J. Miner. Metall. Mater.* **2023**, *30*, 2055–2075. [CrossRef]
3. Deng, Z.; Sarkisov, L. Engineering Machine Learning Features to Predict Adsorption of Carbon Dioxide and Nitrogen in Metal–Organic Frameworks. *J. Phys. Chem. C* **2024**, *128*, 10202–10215. [CrossRef]
4. Yoon, C.; Eom, C.; Jeon, Y.; Kim, K. Development of a Nitrogen Prediction Model for 320 Tonne Converter. In Proceedings of the 12th International Conference of Molten Slags, Fluxes and Salts (MOLTEN 2024), Brisbane, Australia, 17–19 June 2024. [CrossRef]
5. Wu, Y.; Zhang, H.; Jian, L.; Lv, Z. A Quantitative Causal Analysis and Optimization Framework for Inclusions of Steel Products. *Adv. Eng. Inform.* **2024**, *62*, 102629. [CrossRef]

6. Patra, S.; Nayak, J.; Singhal, L.; Pal, S. Prediction of Nitrogen Content of Steel Melt during Stainless Steel Making Using AOD Converter. *Steel Res. Int.* **2017**, *88*, 1600271. [CrossRef]
7. Liu, C.; Tang, L.; Liu, J. A Stacked Autoencoder With Sparse Bayesian Regression for End-Point Prediction Problems in Steelmaking Process. *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 550–561. [CrossRef]
8. Conrad, F.; Mälzer, M.; Schwarzenberger, M.; Wiemer, H.; Ihlenfeldt, S. Benchmarking AutoML for Regression Tasks on Small Tabular Data in Materials Design. *Sci. Rep.* **2022**, *12*, 19350. [CrossRef] [PubMed]
9. Zhang, T.; Zhang, J.; Peng, G.; Wang, H. Automated Machine Learning for Steel Production: A Case Study of TPOT for Material Mechanical Property Prediction. In Proceedings of the 2022 IEEE International Conference on e-Business Engineering (ICEBE), Bournemouth, UK, 14–16 October 2022; pp. 94–99. [CrossRef]
10. Feng, L.; Zhao, C.; Li, Y.; Zhou, M.; Qiao, H.; Fu, C. Multichannel Diffusion Graph Convolutional Network for the Prediction of Endpoint Composition in the Converter Steelmaking Process. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 1–13. [CrossRef]
11. Chumanov, I.; Sedukhin, V. Analysis of methods for predicting the limiting nitrogen concentration in duplex steels. *Ferr. Metall. Bull. Sci. Tech. Econ. Inf.* **2022**, *78*, 598–604. [CrossRef]
12. Laha, D.; Ye, R.; Suganthan, P. Modeling of Steelmaking Process with Effective Machine Learning Techniques. *Expert Syst. Appl.* **2015**, *42*, 4687–4696. [CrossRef]
13. Xiao, X.; Trinh, T.; Gerelkhuu, Z.; Ha, E.; Yoon, T. Automated Machine Learning in Nanotoxicity Assessment: A Comparative Study of Predictive Model Performance. *Comput. Struct. Biotechnol. J.* **2024**, *25*, 9–19. [CrossRef]
14. Pitkälä, J.; Holappa, L.; Jokilaakso, A. Production of Nitrogen-Alloyed Stainless Steels in Argon Oxygen Decarburization Converter: Kinetics and Modeling of Nitrogenation and Denitrogenation. *Steel Res. Int.* **2023**, *95*, 2300597. [CrossRef]
15. Tsamardinos, I.; Fanourgakis, G.; Greasidou, E.; Klontzas, E.; Gkagkas, K.; Froudakis, G. An Automated Machine Learning Architecture for the Accelerated Prediction of Metal-Organic Frameworks Performance in Energy and Environmental Applications. *Microporous Mesoporous Mater.* **2020**, *300*, 110160. [CrossRef]
16. Luo, J.; Luo, Y.; Cheng, X.; Liu, X.; Wang, F.; Fang, F.; Cao, J.; Liu, W.; Xu, R.-Z. Prediction of Biological Nutrients Removal in Full-Scale Wastewater Treatment Plants Using H_2O Automated Machine Learning and Back Propagation Artificial Neural Network Model: Optimization and Comparison. *Bioresour. Technol.* **2023**, *390*, 129842. [CrossRef]
17. Sheik, S.; Mohammed, R.; Teeparthi, K.; Raghuvamsi, Y. Machine Learning-Based Prediction of Intergranular Corrosion Resistance in Austenitic Stainless Steels Exposed to Various Heat Treatments. *J. Inst. Eng. India Ser. D* **2025**, *106*, 491–504. [CrossRef]
18. Kateb, M.; Safarian, S. Machine Learning-Driven Predictive Modeling of Mechanical Properties in Diverse Steels. *Mach. Learn. Appl.* **2025**, *20*, 100634. [CrossRef]
19. Shinde, P.P.; Shah, S. A Review of Machine Learning and Deep Learning Applications. In Proceedings of the 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 16–18 August 2018; pp. 1–6.
20. Ghalati, M.K.; Zhang, J.; El-Fallah, G.M.M.M.; Nenchev, B.; Dong, H. Toward Learning Steelmaking—A Review on Machine Learning for Basic Oxygen Furnace Process. *Mater. Genome Eng. Adv.* **2023**, *1*, e6. [CrossRef]
21. ELTRA GmbH. Basic Application Information. Available online: https://www.eltra.com/files/446383/expert-guide-application-information.pdf (accessed on 8 December 2025).
22. ELTRA GmbH. Effective Quality Control of Steel and Iron Products with Combustion Analysis. Available online: https://www.eltra.com/files/14146/effective-quality-control-of-steel-and-iron-products-with-combustion-analysis.pdf (accessed on 28 December 2025).
23. *ASTM E1019-18*; Standard Test Methods for Determination of Carbon, Sulfur, Nitrogen, and Oxygen in Steel, Iron, Nickel, and Cobalt Alloys by Various Combustion and Inert Gas Fusion Techniques. ASTM International: West Conshohocken, PA, USA, 2018. Available online: https://store.astm.org/e1019-18.html (accessed on 8 December 2025).
24. Elastic Net Regularization. Available online: https://web.archive.org/web/20250814115617/https://questdb.com/glossary/elastic-net-regularization/ (accessed on 14 August 2025).
25. Tanner, G. Random Forest. Available online: https://ml-explained.com/blog/random-forest-explained (accessed on 22 October 2025).
26. Hossain, M.M. Mastering LightGBM: An In-Depth Guide to Efficient Gradient Boosting. Available online: https://medium.com/@mohtasim.hossain2000/mastering-lightgbm-an-in-depth-guide-to-efficient-gradient-boosting-8bfeff15ee17 (accessed on 22 October 2025).
27. Ni, K.S.; Nguyen, T.Q. Adaptable K-Nearest Neighbor for Image Interpolation. In Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 1297–1300.
28. Hesterberg, T.; Choi, N.H.; Meier, L.; Fraley, C. Least Angle and L1 Penalized Regression: A Review. *Stat. Surv.* **2008**, *2*, 61–93. [CrossRef]
29. Lifesight Normalized Root Mean Square Error (NRMSE). Available online: https://lifesight.io/glossary/normalized-root-mean-square-error/ (accessed on 22 October 2025).

30. Draper, N.R.; Smith, H. *Applied Regression Analysis*; John Wiley & Sons: New York, NY, USA, 1998; ISBN 978-0-471-17082-2.
31. Lee, S. MAE Mastery: Your Guide to Mean Absolute Error. Available online: https://www.numberanalytics.com/blog/mae-mastery-guide-mean-absolute-error (accessed on 22 October 2025).
32. Mean Absolute Percentage Error (*MAPE*): What You Need To Know. Available online: https://arize.com/blog-course/mean-absolute-percentage-error-mape-what-you-need-to-know/ (accessed on 22 October 2025).
33. Bhandari, P. Correlation Coefficient | Types, Formulas & Examples. Available online: https://www.scribbr.com/statistics/correlation-coefficient/ (accessed on 24 March 2025).
34. Lane, D. Proportion of Variance Explained. Available online: https://stats.libretexts.org/Bookshelves/Introductory_Statistics/Introductory_Statistics_(Lane)/19%3A_Effect_Size/19.04%3A_Proportion_of_Variance_Explained (accessed on 22 October 2025).
35. Borgogno Mondino, E.; Farbo, A.; Novello, V.; Palma, L. A Fast Regression-Based Approach to Map Water Status of Pomegranate Orchards with Sentinel 2 Data. *Horticulturae* **2022**, *8*, 759. [CrossRef]
36. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems 30 (NIPS 2017)*; Curran Associates, Inc.: Red Hook, NY, USA, 2017.
37. Van den Broeck, G.; Lykov, A.; Schleich, M.; Suciu, D. Tractability SHAP Explanations. *arXiv* **2021**, arXiv:2009.08634. [CrossRef]
38. Nohara, Y.; Matsumoto, K.; Soejima, H.; Nakashima, N. Explanation of Machine Learning Models Using Shapley Additive Explanation and Application for Real Data in Hospital. *Comput. Methods Programs Biomed.* **2022**, *214*, 106584. [CrossRef]
39. Lundberg, S.M.; Lee, S.-I. Consistent Feature Attribution for Tree Ensembles. *arXiv* **2018**, arXiv:1802.03888. [CrossRef]
40. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef]
41. Sudjianto, A.; Knauth, W.; Singh, R.; Yang, Z.; Zhang, A. Unwrapping The Black Box of Deep ReLU Networks: Interpretability, Diagnostics, and Simplification. *arXiv* **2020**. [CrossRef]
42. Covert, I.; Lundberg, S.; Lee, S.-I. Understanding Global Feature Contributions With Additive Importance Measures. *arXiv* **2020**. [CrossRef]
43. Ponce-Bobadilla, A.V.; Schmitt, V.; Maier, C.S.; Mensing, S.; Stodtmann, S. Practical Guide to SHAP Analysis: Explaining Supervised Machine Learning Model Predictions in Drug Development. *Clin. Transl. Sci.* **2024**, *17*, e70056. [CrossRef]
44. Mohsin, M.T.; Nasim, N.B. Explaining the Unexplainable: A Systematic Review of Explainable AI in Finance. *arXiv* **2025**. [CrossRef]
45. Truong, A.; Walters, A.; Goodsitt, J.; Hines, K.; Bruss, B.; Farivar, R. Towards Automated Machine Learning: Evaluation and Comparison of AutoML Approaches and Tools. In Proceedings of the 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), Portland, OR, USA, 4 November–6 November 2019; pp. 1471–1479. [CrossRef]
46. Goldstein, D.A.; Fruehan, R.J. Mathematical Model for Nitrogen Control in Oxygen Steelmaking. *Met. Mater. Trans. B* **1999**, *30*, 945–956. [CrossRef]
47. Han, H.; Shaker, B.; Lee, J.H.; Choi, S.; Yoon, S.; Singh, M.; Basith, S.; Cui, M.; Ahn, S.; An, J.; et al. Employing Automated Machine Learning (AutoML) Methods to Facilitate the In Silico ADMET Properties Prediction. *J. Chem. Inf. Model.* **2025**, *65*, 3215–3225. [CrossRef]
48. Shamsuddin, M. Secondary Steelmaking. In *Physical Chemistry of Metallurgical Processes*, 2nd ed.; Shamsuddin, M., Ed.; Springer International Publishing: Cham, Switzerland, 2021; pp. 293–351. ISBN 978-3-030-58069-8.
49. Wu, H.; Zhang, B.; Li, Z. Small Sample-Oriented Prediction Method of Mechanical Properties for Hot Rolled Strip Steel Based on Model Independent Element Learning. *IEEE Access* **2024**, *12*, 197300–197311. [CrossRef]
50. Zhang, C.-J.; Zhang, Y.-C.; Han, Y. Industrial Cyber-Physical System Driven Intelligent Prediction Model for Converter End Carbon Content in Steelmaking Plants. *J. Ind. Inf. Integr.* **2022**, *28*, 100356. [CrossRef]
51. Yang, Q.; Fan, Y.; Rong, D.; Bao, R.; Zhang, D. An Auto-configurable Machine Learning Framework to Optimize and Predict Catalysts for CO2 to Light Olefins Process. *AIChE J.* **2024**, *70*, e18437. [CrossRef]
52. De Oliveira, V.; Komati, K.; Andrade, J. Implementing Neuroevolution for Gas Consumption Forecasting in the Steel Industry. In Proceedings of the 2024 L Latin American Computer Conference (CLEI), Buenos Aires, Argentina, 12–16 August 2024; pp. 1–10. [CrossRef]
53. Bender, J.; Trat, M.; Ovtcharova, J. Benchmarking AutoML-Supported Lead Time Prediction. *Procedia Comput. Sci.* **2022**, *200*, 482–494. [CrossRef]
54. Hariri-Ardebili, M.; Mahdavi, P.; Pourkamali-Anaraki, F. Benchmarking AutoML Solutions for Concrete Strength Prediction: Reliability, Uncertainty, and Dilemma. *Constr. Build. Mater.* **2024**, *423*, 135782. [CrossRef]
55. Liu, G.; Lu, D.; Lu, J. Pharm-AutoML: An Open-source, End-to-end Automated Machine Learning Package for Clinical Outcome Prediction. *CPT Pharmacomet. Syst. Pharmacol.* **2021**, *10*, 478–488. [CrossRef]
56. Kwon, N.; Comuzzi, M. Genetic Algorithms for AutoML in Process Predictive Monitoring. In *Process Mining Workshops*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 242–254. [CrossRef]

57. Luo, C.; Zhang, Z.; Qiao, D.; Lai, X.; Li, Y.; Wang, S. Life Prediction under Charging Process of Lithium-Ion Batteries Based on AutoML. *Energies* **2022**, *15*, 4594. [CrossRef]
58. Denkena, B.; Dittrich, M.; Lindauer, M.; Mainka, J.; Stürenburg, L. Using AutoML to Optimize Shape Error Prediction in Milling Processes. *SSRN* **2020**. [CrossRef]
59. Hadi, R.; Hady, H.; Hasan, A.; Al-Jodah, A.; Humaidi, A. Improved Fault Classification for Predictive Maintenance in Industrial IoT Based on AutoML: A Case Study of Ball-Bearing Faults. *Processes* **2023**, *11*, 1507. [CrossRef]
60. Musigmann, M.; Akkurt, B.; Krähling, H.; Nacul, N.G.; Remonda, L.; Sartoretti, T.; Henssen, D.; Brokinkel, B.; Stummer, W.; Heindel, W.; et al. Testing the Applicability and Performance of Auto ML for Potential Applications in Diagnostic Neuroradiology. *Sci. Rep.* **2022**, *12*, 13648. [CrossRef]
61. Li, P.; Yang, Y.; Chen, C. Research on Fatigue Crack Propagation Prediction for Marine Structures Based on Automated Machine Learning. *J. Mar. Sci. Eng.* **2024**, *12*, 1492. [CrossRef]
62. De S'a, A.; Ascher, D. Auto-ADMET: An Effective and Interpretable AutoML Method for Chemical ADMET Property Prediction. *arXiv* **2025**. [CrossRef]
63. Mubarak, Y.; Koeshidayatullah, A. Hierarchical Automated Machine Learning (AutoML) for Advanced Unconventional Reservoir Characterization. *Sci. Rep.* **2023**, *13*, 13812. [CrossRef] [PubMed]
64. Kaftantzis, S.; Bousdekis, A.; Theodoropoulou, G.; Miaoulis, G. Predictive Business Process Monitoring with AutoML for next Activity Prediction. *Intell. Decis. Technol.* **2024**, *18*, 1965–1980. [CrossRef]
65. Sousa, A.; Ferreira, L.; Ribeiro, R.; Xavier, J.; Pilastri, A.; Cortez, P. Production Time Prediction for Contract Manufacturing Industries Using Automated Machine Learning. In *Artificial Intelligence Applications and Innovations*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 262–273. [CrossRef]
66. Kim, G.; Steller, M.; Olson, S. Modeling Watershed Nutrient Concentrations with AutoML. In *Proceedings of the 10th International Conference on Climate Informatics*; ACM: New York, NY, USA, 2020. [CrossRef]
67. Wang, J.; Xue, Q.; Zhang, C.; Wong, K.K.L.; Liu, Z. Explainable Coronary Artery Disease Prediction Model Based on AutoGluon from AutoML Framework. *Front. Cardiovasc. Med.* **2024**, *11*, 1360548. [CrossRef]
68. Kulkarni, G.; Ambesange, S.; Vijayalaxmi, A.; Sahoo, A. Comparision of Diabetic Prediction AutoML Model with Customized Model. In Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 25–27 March 2021; pp. 842–847. [CrossRef]
69. Kačur, J.; Flegner, P.; Durdán, M.; Laciak, M. Prediction of Temperature and Carbon Concentration in Oxygen Steelmaking by Machine Learning: A Comparative Study. *Appl. Sci.* **2022**, *12*, 7757. [CrossRef]
70. Conrad, F.; Mälzer, M.; Lange, F.; Wiemer, H.; Ihlenfeldt, S. AutoML Applied to Time Series Analysis Tasks in Production Engineering. *Procedia Comput. Sci.* **2024**, *232*, 849–860. [CrossRef]
71. Demeter, J.; Buľko, B.; Demeter, P.; Hrubovčáková, M. Prediction Models for Nitrogen Content in Metal at Various Stages of the Basic Oxygen Furnace Steelmaking Process. *Appl. Sci.* **2025**, *15*, 9561. [CrossRef]