



Enhanced Temperature Prediction in Ladle Furnace Steel Refining: Hybrid Process Modeling Based on Computational Thermodynamics and Statistical Learning Methods

DANIEL KAVIĆ, MICHAEL BERNHARD, ROMAN RÖSSLER, LENA SCHALK,
and CHRISTIAN BERNHARD

The ladle furnace (LF) is the central refining aggregate in modern secondary steelmaking, enabling the admixture of alloying elements, chemical homogenization of the melt, and removal of non-metallic inclusions (NMIs). In addition to metallurgical tasks, precise temperature control and superheating are central to the subsequent casting processes, with optimized temperature management significantly reducing energy consumption and the associated CO₂ footprint. The work's first part deals with a comprehensive literature study of existing mechanistic ("white-box"), data-driven ("black-box"), and hybrid ("gray-box") temperature models for secondary steelmaking, providing a detailed overview of the model's scope, methodology, used boundary conditions, and approach to validate the predictions. Based on the literature study, an evaluation highlights the strengths and weaknesses of the different models, focusing on their steel temperature predictability and the process variables they incorporate. The work's second part outlines the methodology and scope of the developed hybrid LF process model. It is based on the effective equilibrium reaction zone (EERZ) method, coupling fundamental computational thermodynamics with the ongoing kinetics within the steel ladle. Consequently, the heat effects of chemical reactions and mixing enthalpies of alloying materials can be quantified. Statistical learning techniques are applied to LF process data of an annual production at voestalpine Stahl GmbH. This data were extensively cleaned and analyzed using descriptive statistics. Finally, the influences of relevant process parameters on the temperature, *e.g.*, heating power *via* electrodes, cooling in the ladle due to radiation and convection, and purging flow rate *via* bottom plugs, were quantified by applying multiple linear regression. In the work's final part, the findings from the data-driven analysis are presented, and the hybrid temperature model is thoroughly tested using several case studies. With a deviation of less than 5 °C for 94 pct of the predicted temperatures compared to the actual measured ones, the model can be considered precise and reliable. The developed ladle furnace model can be employed to simulate various process sequences, identify potential optimization measures, and make appropriate adaptations during the steel refinement.

<https://doi.org/10.1007/s11663-025-03715-4>
© The Author(s) 2025

I. INTRODUCTION

LADLE furnace process treatment is of the utmost importance for a modern steel plant, offering various treatment steps such as final alloying adjustments, deoxidation, desulphurization, and purity improvements regarding non-metallic inclusions. In addition, liquid steel can be set to the required temperature for the subsequent degassing or casting process using electrical energy *via* carbon electrodes. Precise control of the casting temperature is critical. If it is too high, it can cause severe center segregation or even strand shell breakouts due to slow solidification. If the temperature is too low toward the end of the secondary metallurgy

DANIEL KAVIĆ is with the K1-MET GmbH, Stahlstrasse 14, 4020 Linz, Austria and also with the Chair of Ferrous Metallurgy, Technical University of Leoben, Franz-Josef-Strasse 18, 8700 Leoben, Austria. Contact e-mail: daniel.kavic@k1-met.com MICHAEL BERNHARD and CHRISTIAN BERNHARD are with the Chair of Ferrous Metallurgy, Technical University of Leoben. ROMAN RÖSSLER is with the voestalpine Stahl GmbH, voestalpine-Strasse 3, 4020 Linz, Austria. LENA SCHALK is with the K1-MET GmbH and also with the voestalpine Stahl GmbH.

Manuscript submitted May 14, 2025; accepted July 9, 2025.

Article published online August 1, 2025.

(SecMet) treatment, severe nozzle clogging may occur during casting. In its most pronounced form, this phenomenon can lead to massive steel contamination. With the steel temperature falling below the liquidus of the cast material, then casting may be interrupted because of freezing steel in the submerged entry nozzle or tundish. Moreover, temperature control during the LF treatment is of great importance, as late heating can lead to a new oxygen pick-up by the steel bath, resulting in increased aluminum burn-off and the formation of new inclusions. Recently, the transition toward greener steel production *via* the electric arc furnace (EAF) underscores the critical role of secondary steelmaking in optimizing the refinement of this new raw material to ensure high-quality products, comply with the steel plant logistics, and still achieve the required temperatures. Given the potential negative impacts of inaccurate temperature control and the pressing challenges of transitioning to green steel production, enhanced temperature models have been developed in recent years based on three different approaches: (i) mechanistic, physics-based (“white-box”) models,^[1–11] (ii) data-driven (“black-box”) models that utilize large amounts of process data,^[12–25] and (iii) hybrid (“gray-box”) models combining both approaches.^[26–32] Here, a novel hybrid approach is proposed by coupling statistical learning methods with computational thermodynamics to develop a model for a continuous temperature description during LF treatment. The work is structured according to the methodological approach to address the challenge of developing the hybrid temperature model and, afterward, validating it. It begins with comprehensively reviewing the existing literature for temperature models within the secondary steelmaking processes, aiming to evaluate applied methods and identify the key factors influencing temperature development. Based on this foundation, the concept of the newly developed methodology originated, combining thermodynamic-kinetic principles with advanced statistical methods to create a robust framework. Central to this approach is the effective equilibrium reaction zone (EERZ)^[33–36] methodology. It couples computational thermodynamic databases with a kinetic description of the metallurgical process unit, defining the released or consumed energy associated with chemical reactions occurring within the ladle during treatment. External influences, such as heating *via* carbon electrodes, are addressed by applying multiple linear regression (MLR). The results of the statistical investigation are discussed in great detail, and the information gained from the MLR is used to define process boundary conditions for the thermodynamic model. Once the model was finalized, it was thoroughly tested with case studies to validate its predictive power. While the ultimate goal is to develop an extensive model for secondary metallurgy that predicts not only temperature evolution but also accounts for chemical compositions and incorporates diverse process boundary conditions, the current work represents a significant first step toward achieving this overarching objective.

II. LITERATURE REVIEW

In recent years, various approaches and possibilities have been developed to predict the temperature evolution during SecMet treatment. The different prediction methods can be divided into (i) mechanistic models, which can also be interpreted as physically based models, (ii) purely data-driven modeling based on large amounts of process data, and (iii) hybrid models that attempt to combine both concepts. The subsequent chapters provide a detailed examination of the various approaches, outlining the scope of each model, the applied methodology, the defined boundary conditions, and the approach to results validation. Finally, a concise evaluation of the models’ advantages and limitations will be presented, with a particular focus on their applicability to temperature modeling and their considered process variables.

A. Mechanistic Modeling

Mechanistic models leverage differential equations and numerical methods rooted in thermodynamics and heat conduction principles to comprehensively represent a real system’s conditions, behaviors, and interactions among its functional structures and elements. As all parameters in this approach are “visible” to the user, the term “white-box” modeling is frequently used.^[37]

Çamdali *et al.*^[1,2] described analytically the energy and exergy efficiency at the LF. To achieve this, the initial energy of liquid steel, electrical energy inputs, heat losses by conduction and radiation from the LF and electrodes, a mass balance of the incoming and outgoing materials in the ladle, and relevant chemical reactions, *e.g.*, deoxidation and desulphurization, were considered. Further, the authors assumed that all chemical reactions are reversible, the process itself is steady-state, and the enthalpy is T -dependent only. This methodological approach led to the result that the exergy efficiency was found to be 81 pct. In a follow-up study, Çamdali and Tunç^[3] investigated the heat losses based on a detailed description of the thermal boundary conditions, analyzing losses through the LF wall and bottom by conduction. Additionally, the radiation and convection losses from the LF surface and radiation losses *via* electrodes were implemented into the calculations. This detailed investigation was achieved by applying fundamental heat transfer equations. The result revealed a distribution of energy losses, with radiation through the electrodes and convection at the outer wall of the ladle causing the highest dissipation rates.

Similarly, Belkovskii and Kats^[4] published a temperature model again based on fundamental heat transfer equations to determine the temperature change during tapping at the electric arc furnace (EAF), the transportation of the steel ladle to the LF, and the casting process. Thermal radiation from the surface of the tapping stream and the bath surface, as well as the heat losses through conduction *via* the refractory material, were accounted for to describe the tapping process. For the transportation and casting phase, the thermal losses

through the ladle walls and the surface were defined by applying classical heat transfer equations. Moreover, the temperature changes due to alloy additions were considered based on thermodynamic literature compilation.

Kathait^[5] studied heat losses *via* conduction through the ladle wall and radiation from the bath surface by applying Fourier's heat conduction equation and Stefan-Boltzmann law. Physical and thermal properties of ladle walls, molten steel and slag and four different emissivity coefficients (hot surface refractory material, outer ladle shell surface, slag surface and molten steel surface) were specified. The author provided a basic model concept but no critical comparison with plant data was available.

A computational model to define thermal losses with three different types of refractory material (doloma, alumina, high alumina) of a steelmaking ladle during the holding period was formulated by Duarte *et al.*^[6] It applied numerical integration of the defined physics-based equations over the process time. Heat losses due to heat conduction through the refractory and radiation from the slag surface to the surrounding environment were analyzed. Once again, the classic approaches of thermal engineering, the Fourier's heat conduction equation and the Stefan-Boltzmann equation, were employed. Reasonable agreement with previous findings from literature^[38,39] was obtained.

For solving a two-dimensional transient heat transfer problem within a steel ladle, a numerical discretization methodology based on the explicit finite difference method was presented by Zabadal *et al.*^[7] Here, the individual layers of the ladle lining from Aços Finos Piratini, Brazil were described in detail (thickness and thermal diffusivity). The thermal radiation at the bath surface was also taken into account. The model covers the entire production cycle from tapping to casting and was validated using process data from Aços Finos Piratini.

Mukhopadhyay *et al.*^[8] modeled the melt's flow and temperature field and its heat conduction within the ladle *via* a two-dimensional, quasi-single-phase turbulent backflow model, which was realized through the solution of Reynolds-averaged Navier-Stokes equation (RANS) and a $k-\epsilon$ model. The RANS was discretized using the finite volume method (FVM) in cylindrical coordinates and solved by applying the Simple Algorithm. Literature data^[40-42] were used to validate the flow and heat transfer model.

Building on the research of Mukhopadhyay *et al.*,^[8] Deb *et al.*^[9] expanded the model to incorporate heat transfer through the refractory, radiation at the bath surface, and the influence of alloy additions. A semi-implicit FVM was again used to solve the coupled governing equations. The model was validated using literature^[43] data and temperature measurements from LD-2 shop of Tata Steel (Jamshedpur, India).

Tripathi *et al.*^[10] applied computational fluid dynamics (CFD) using FLUENT for predicting ladle and steel temperature during the whole process cycle of secondary steelmaking. The work focused on the influence of various process parameters, *e.g.*, initial tapping

temperature, slag thickness, and ladle age, on the steel's temperature evolution. The model applied governing equations for heat conduction considering the radiation and natural convection from the outer surface of the ladle wall. Chill factors for the different alloying materials (SiMn, FeSi, etc.) were also implemented. The model's prediction power was confirmed using plant data comparison. A maximum deviation of 4 pct was achieved.

A third modeling technique for continuous temperature prediction in the field of secondary metallurgy utilizes the effective equilibrium reaction zone^[33] approach. This technique employs computer-aided thermodynamic databases to comprehensively describe the influence of chemical reactions and the associated released enthalpies. Van Ende *et al.*^[11] presented a ladle furnace model based on the EERZ method. Besides the thermodynamic description of all occurring reactions within the system, the process-dependent temperature influences (heating *via* electrodes, purging *via* bottom plug, etc.) were included. The model assumed two different temperatures for the steel and slag, with heat exchange between the two bulks based on a simple heat conduction equation. In addition, losses to the environment due to radiation and convection were described with a simple boundary condition ranging from 1 to 2 °C min⁻¹. The model's result comparison with plant data published by Graham and Irons^[44] showed great agreement. Finally, it should be briefly mentioned that the EERZ method has the significant advantage of also providing a continuous prediction of the chemical composition of the steel, slag, and non-metallic inclusions.^[11,33-35,45]

B. Data-Driven Modeling

Data-driven temperature prediction is playing an increasingly important role in secondary metallurgy due to the ever-increasing amount of data, more accurate measuring devices in recent years, and sophisticated regression modeling. This approach uses "black-box" modeling to describe only the behavior but not the functional structure and physical-based relationships between the parameters within the system.^[37] A significant advantage of data-driven modeling is that physically based parameters, which are often difficult to capture in real time, are not required.

Using multiple linear regression (MLR), Wang *et al.*^[12] investigated for two different operating modes (with and without oxygen blowing) the influence of the initial steel temperature, molten steel quantity, treatment time, oxygen blowing amount, baking temperature in vacuum chamber, scrap and alloy addition on the final temperature during the RH (Ruhrstahl-Heraeus)-TOP blowing refining process for interstitial-free steels. In each mode, 600 heat treatments were made available by their industrial partner, Qian-an Steel Company, China to train the temperature model. Another 180 observations were provided to test the developed model. Despite the simplicity of the MLR, 96 pct of the predictions for the mode without oxygen

blowing could be predicted within a 10 °C tolerance. The second mode's value was even slightly higher at 97 pct.

To predict the final temperature of the molten steel in the secondary refining and the target temperature in the tundish, a model was developed by Matos *et al.*^[13] that uses MLR for the linear relationships combined with Random Forest (RF) for the non-linear relationships between the independent and dependent variables. The two models were combined *via* a boosting ensemble technique, in which the models are sequentially fitted using the residuals of the previous model. The used data set contained 2500 observations provided by Gerdau Ouro Branco, Brazil. The parameters that were taken into account in the statistical analysis as influencing factors for the temperature prediction were the heating time at LF, the preheating time of the tundish, the number of casting strands, the age of the ladle, the casting time, the preheating time of the ladle, the circulation time of the ladle and the current number of the casting sequence. The combined model achieved an accuracy of 94.5 ± 0.02 pct within a ± 7 °C range. However, the simple MLR also achieved a predictive accuracy of 93.1 ± 0.02 pct for a maximum temperature deviation of 7 °C.

Wang *et al.*^[14,15] selected the ensemble regression algorithm, RF, to quantify the final temperature at the LF by the concepts of bootstrapping and aggregating. A data set with 1714 observations was utilized to train and test the regression model, which was normalized before the tenfold cross-validation. Finally, the process was repeated 20 times and averaged for all estimators. The weight of the steel, state and temperature of the ladle, treatment duration, initial temperature, the time interval of temperature measurements, amount of heating and stirring gas, and alloy additions were considered as input parameters. The *RMSE* of the final model was 2.4 °C with a maximum error of 6.7 °C.

A statistical model for predicting the steel temperature during the entire secondary metallurgical treatment was presented by Sonoda *et al.*^[16] using a bootstrap filter with predictive distributions. Remarkable is the model's capability of considering not only a single melt's temperature prediction but also its distribution and, thus, the possible deviation. 2840 observations given by their industrial partner were ready for the modeling, where 20 pct of it was used as a test set to validate the model's accuracy. Important modeling parameters included the initial temperature, chemical composition of the steel, and process-related influences, *e.g.*, waiting time, alloy additions, the thermal state of the ladle, etc. Absolute values for the model's accuracy were not given.

Besides the classic regression models, artificial neural networks (ANN) have also been established for temperature prediction in the SecMet. Duarte *et al.*^[17] applied a predictive ANN to determine the temperature loss of the ladle starting from LF treatment until placing it in the ladle turret of the continuous caster. Only seven independent variables were selected as input parameters for the ANN, *e.g.*, carbon content, heating duration, treatment duration, run number in the production sequence, etc. A data set with 1399 observations was

ready for the network's training, whereas 334 were used for validating the model. The developed ANN delivered great predictability with 96.1 pct of the forecasts being within ± 10 °C deviation. Duarte *et al.*^[17] compared his ANN with a linear regression model, which showed great predictability, too, with a value of 94.9 pct for ± 10 °C.

Sztangret *et al.*^[18] deployed a decision tree, an MLR and an ANN to estimate the cooling rate in the ladle during continuous casting. Each of the 7300 observations contained the steel grade information, the initial temperature after SecMet, the time difference between the measurements and the steel weight. CMC, Poland provided the data. Absolute error values for the models were unavailable; however, no big difference between the models can be established from the provided quality measures.

A Deep Neural Network (DNN) combined with a Multi-process Operation Simulation (MOS) was established by Yang *et al.*^[19] for the end-point temperature prediction of the steel ladle at the CC. In parallel, an extreme learning machine (ELM) and a multivariate polynomial regression (MVPR) were implemented. The steel mass, ladle weight, refractory temperature before tapping, transportation time between SecMet and CC, and the initial temperature were considered as potentially important input variables. Their industrial partner provided 1800 observations for training and another 450 for testing the model. The DNN was also compared to a multivariate polynomial regression, where the former achieved a 2 to 3 pct higher hit ratio.

Tian *et al.*^[20] published an incremental learning ensemble modeling method with multiple submodels utilizing an ELM. The influencing parameters for the modeling have been the refining power amount, ladle states, initial temperature, argon purging volume, weight of molten steel, refining time, and heat effects of additions. The exact values for the last one have been provided by their industrial partner Baoshan, China. 550 observations were made available for developing the incremental learning model. The developed model delivered an accuracy of 90 pct for an error margin of ± 5 °C between the measured and predicted values.

Another possibility for temperature modeling at the LF was shown by Tian *et al.*^[21] using back-propagation NN and a further development of the ensemble algorithm called AdaBoost.IR. 300 observations from Baosteel, China were used to train and test the developed model. The influencing parameters on the target temperature are the same as in Reference 20. The back-propagation NN showed a lower accuracy with 71 pct within ± 5 °C deviation than the AdaBoost.IR algorithm with 93 pct.

Another data-driven temperature model for the ladle furnace, published by Xin *et al.*^[22] utilized an isolation forest (IF), zero-phase component analysis (ZCA) whitening, and a DNN. The influencing parameters for the model prediction are the ladle's turnover time, the molten steel's weight and initial temperature, the refining time, the heating duration, the argon consumption, and the temperature change caused by additions. The model's training was based on a tenfold

cross-validation on a data set with 8264 observations. A further 1500 observations were available for the evaluation of the model. The IF-ZCA-DNN model achieved a 99 pct hit rate for a ± 10 °C deviation.

Lv *et al.*^[23] developed a data-driven temperature prediction model whereby the mechanisms of influencing variables (*e.g.*, heating, purging, alloying, treatment duration, etc.) during the LF treatment were taken into account in a Partial Linear Regularization Network with Sparse Representation (PLRN-SR). A data set with 400 heats was applied to train and another 200 heats to test the developed model. The model achieved an *MAE* of 3.6 °C and an *RMSE* of 4 °C.

He *et al.*^[24] carried out a case-based reasoning (CBR) method for the temperature prediction at the LF. In addition to the influencing parameters as in Reference 20, here, the refractory properties were additionally divided into six, and the thermal status of the ladle was divided into eight possible states, which were also included in the modeling as independent variables. A total of 3847 heats formed the database for the CBR modeling. The method delivered a 97.4 pct accuracy for ± 10 °C deviation from the measured temperatures.

Singh *et al.*^[25] went one step further by choosing a Reliable Machine Learning (RML) framework for predicting the temperature at the end of BOF treatment, at the start and end of LF, at the end of RH, and at the start of the casting process in the tundish. Eight regression models were combined, including K-nearest neighbor (KNN), support vector regression (SVR), least absolute shrinkage and selection operator (LASSO), ridge regression, elastic net regression, random forest regression, and decision tree regression for a temperature prediction across the entire SecMet process. Depending on the process phase, the Rourkela steel plant (India) provided between 1199 and 3796 observations for training and testing the RML. The independent variables were different for each process phase. However, the steel's chemical composition, the alloy material type and weight, the steel bath weight, and the initial temperature were always the standard parameters. The coefficient of determination, *i.e.*, *R*, was for every process phase prediction higher than 95 pct, with error distributions lower than 5 °C for all test observations.

C. Hybrid Modeling

The combined use of data-driven and physics-based modeling approaches results in a “gray-box” model that unites the advantages of both areas.^[37] On the one hand, physical process parameters, *e.g.*, temperature loss to the surroundings or temperature gain from heating *via* carbon electrodes, which are difficult to measure, are quantified with values defined by data-driven algorithms; on the other hand, the exact thermophysical data can be taken from thermodynamic databases for modeling, thus clarifying other unknown data, *e.g.*, released chemical reaction enthalpies during alloying or deoxidation.

A hybrid incremental learning model using an ELM algorithm and applying the AdaBoost method in combination with a conventional mechanism-based thermal model for temperature prediction at the LF was

designed by Tian *et al.*^[26] The ELM was first used to define process boundary conditions, which the mechanism model then utilizes to predict the temperature. The mechanism model is based on fundamental heat balance equations. Independent variables influencing the coefficients are the weight of molten steel, the initial temperature of molten steel, the area of molten slag, and the refining power. For the modeling, process data of 250 heat treatments were provided by Baoshan, China. Testing the developed hybrid model resulted in 90 pct of the predictions within a 5 °C error margin.

He *et al.*^[27] developed two hybrid temperature models based on a NN to investigate the influence of the ladle heat status on the steel temperature for the entire secondary metallurgical treatment. The first model, starting at the end of the converter process, estimated the temperatures at the end of the argon-blowing station, the starting and end temperatures at the LF, and the tundish temperature. The second model was a back-propagation of the process and predicted the temperatures starting from the continuous casting machine and going all the way back to the converter end temperature. Parameters for the models are the initial temperature, the process duration, the mass of molten steel, the amount and type of addition, and the ladle life. Depending on the process step, the database size was between 163 and 2588 observations. The process model showed good accuracy, but the further away the prediction point was, the worse the predictions became. The last predicted temperature had a forecast accuracy of 70 pct for a temperature deviation of ± 20 °C.

Gupta and Chandra^[28] developed a temperature prediction model to determine the tap temperature at the BOF and the resulting tundish steel temperature. A one-dimensional heat transfer model coupled with a simple polynomial fourth-order regression model was implemented for this purpose. The heat transfer model covers the first process phase until the start of secondary metallurgy. Subsequently, after the secondary metallurgical treatment the non-linear and complex cooling conditions are better described using statistical relationships. The validation of the model was done for the different process sequences using several hundreds of observations provided by Tata Steel, India. Depending on the process phase, a 90 to 95 pct predictability for a temperature deviation of ± 10 °C was achieved.

An LF online reckoner for continuously predicting the steel temperature and composition was developed by Nath *et al.*^[29] The model was a conventional mechanism-based thermal model. The developed software considered the heat losses to the environment due to radiation and convection, as well as the influences of heating and stirring at the LF. Moreover, the alloy additions and their effect on the liquid melt temperature were specified as numerical values. However, multivariable statistical analysis was conducted to estimate the parameter coefficients of those various process phenomena. Tata Steel, India, provided process data of more than 200 heats for training and testing the model. For a discrepancy of 5 °C between the calculated and measured temperature, an accuracy of 90 pct was achieved.

Lv *et al.*^[30] were able to develop a hybrid first principle mechanistic model combined with fuzzy logic, an ELM, and an ensemble-based approach for temperature prediction at the LF. The data-driven method was applied to define parameters within the mechanistic model using comprehensive process data. The model's influencing parameters were the slag thickness, arc length, heating amount, stirring gas quantity, treatment duration, refractory temperature, alloy additions, and the last measured temperature. A total of 1365 observations were recorded, 900 of which were employed for training the model. The accuracy of the best model was 89.6 pct for a 5 °C deviation between measured and predicted temperature.

An inverse solution methodology (ISM) in combination with a genetic algorithm to estimate the heat transfer coefficient of a mechanistic-based thermal model and, subsequently, the continuous determination of the steel temperature in the LF was published by Srinivas *et al.*^[31] The influencing parameters are the same as in Reference 26 including the slag weight, carry-over slag amount, the chemical composition of the slag and the ladle furnace dimensions and operating conditions. For the last LF measurement, a hit rate of 92.6 pct was achieved within ± 10 °C.

Xin *et al.*^[32] developed a hybrid model based on a DNN in combination with an expert control system (combination of statistical analysis, metallurgical mechanism, and the experience of operators to define the temperature change of molten steel through the addition of different materials) for the final temperature prediction of the steel in the LF. Stochastic gradient descent (SGD) coupled with momentum and l2 regularization was applied to optimize the DNN. The material additions, treatment duration, heating duration, purge gas quantity, steel weight, duration of the ladle cycle, and the initial temperature were accounted for in the modeling. 2000 observations of production data were available for the training and testing phase. In addition, the developed model was compared to an MLR. The former reached an impressive hit ratio of 99.4 pct for ± 5 °C, whereas the latter achieved 85.4 pct.

D. Conclusions on Relevant Methods and Process Variables to be Considered

The examined mechanistic models demonstrate that the analytical and numerical description of temperature evolution during secondary steelmaking processes using thermal boundary conditions is frequently associated with significant challenges in model validation. This complexity primarily arises due to the difficulties in experimentally capturing these boundary conditions with sufficient accuracy. An alternative approach is the EERZ methodology, which provides an efficient and systematic framework for developing a temperature model that requires only a limited number of thermal boundary conditions while simultaneously accounting for all chemical reactions within the system. A key advantage of this methodology is that it allows

boundary conditions to be defined in terms of global parameters, such as cooling and heating rates, thereby simplifying model implementation and interpretation.

A comprehensive analysis of data-driven models has further demonstrated that interpretable techniques, such as an MLR, have yielded robust results in numerous studies, where they were benchmarked against complex, non-linear “black-box” models. Those non-linear methods (*e.g.*, ELM, RF, DNN, etc.) have the limitation that the predictor's coefficients are not easily extractable nor interpretable. Therefore, applying MLR to large datasets presents a viable strategy for defining thermal boundary conditions, which can subsequently be integrated with the EERZ methodology. This hybrid approach facilitates the development of a “gray-box” model, a concept that has been extensively explored in the literature. Such a framework enables the strategic application of statistical methods for determining thermal boundary conditions in “white-box” models, thereby reducing dependency on experimentally challenging boundary condition measurements. Based on these findings, the approach will be presented in more detail in the following work.

Finally, despite methodological differences, all models exhibit high similarity in the key process variables that must be considered. The literature identifies several critical parameters for the steel temperature during ladle furnace treatment, including initial temperature, treatment duration, stirring duration, heating amount, and the impact of alloy additions. Given their significance, the present study also focuses on these factors and their influence.

III. HYBRID MODELING CONCEPT AND METHODS USED

A. LF Operation at Voestalpine Stahl GmbH and Its Associated Heat Flows

Figure 1 shows a schematic representation of the LFs installed in the steel plant at voestalpine Stahl GmbH in Linz with the associated process influences considered in the present work. As mentioned in the introduction, the tasks of the ladle furnace are to comply with the specified alloy limits of the respective steel grade and to reduce or modify the non-metallic inclusions as far as possible. Alloy additions can be introduced into the system in two different formats. Lumpy alloys are transported from hopper systems *via* conveyor belts and fed into the system. On the other hand, it is possible to inject CaFe, CaSi and/or aluminum wire into the steel ladle. The latter is primarily added as a deoxidizing agent to remove the dissolved oxygen in the steel completely. However, this procedure forms high-melting aluminum oxides, whereby an inclusion modification (calcium treatment) can be carried out toward the end of the secondary steelmaking process using the first two materials mentioned. This results in the formation of deep-melting calcium aluminates, which can achieve improved clogging properties during CC.^[46] Moreover,

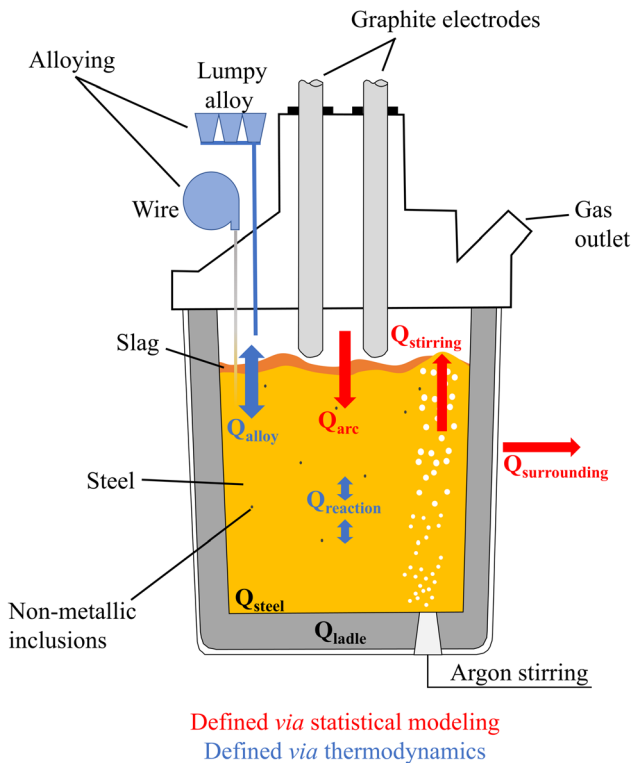


Fig. 1—Schematic illustration of a ladle furnace, its process influences, and the corresponding heat phenomena during refining.

the ladle furnace represents an important aggregate for logistics in steel plants, as the target temperature for the subsequent vacuum treatment or casting process can be set precisely by electrical heating. Hence, accurate knowledge of heat input into the steel ladle system and heat loss to the surroundings is essential in practical operation. The heat balance of the LF system can be estimated considering the following main contributions:

- (i) The conventional method of increasing the molten steel's temperature is based on the electrical heating process using carbon electrodes (Q_{arc}). In the heating phases, the applied voltage during the heating process at a ladle furnace with 180 tons ranges from 450 to 600 kV and has an electric current of approx. 40 kA. Since applied voltage and current have nearly constant values, it is reasonable to just consider the heating amount in kWh and not the heating duration as a first influencing parameter due to their strong collinearity.
- (ii) Potential energy input into the system arises from alloy additions (Q_{alloy}). This process can be exothermic, particularly when introducing aluminum or FeSi^[47] into the melt. However, alloy additions typically result in a temperature drop because they require substantially more energy to melt and heat the material to the temperature of liquid steel than they release during mixing and dissociation processes. These materials' quantities and chemical compositions are required to model the influence of the addition processes. Based on this information, Q_{alloy} can be determined.

- (iii) Consideration of ongoing chemical reactions ($Q_{reaction}$), *e.g.*, deoxidation, desulphurization, etc., as they lead to an energy change within the system.
- (iv) The heat loss caused by argon stirring using a bottom plug ($Q_{stirring}$) is taken into account. This heat loss can be mapped *via* the required gas volume in Nm³ or the stirring duration in min. Once again, there is a certain collinearity between those two parameters, so the focus was placed on the purge gas duration in this study.
- (v) $Q_{surrounding}$ represents a further heat loss. This value combines all the radiation and convective heat losses from the ladle system to the surroundings. Since those values are difficult to measure, especially radiation losses at the top of the ladle as they largely depend on the unknown slag temperature and thickness, $Q_{surrounding}$ mainly reflects the temperature drop due to treatment duration, measured in minutes, at the LF.
- (vi) The energy loss due to the thermal state of the ladle (Q_{ladle}) is primarily determined by the turnover cycle time. Therefore, it can be assumed that the temperature of the refractory lining decreases with increasing turnover cycle time. Thus, more energy can be extracted from the Q_{steel} . However, in the present work, only melts in a continuous process cycle and, therefore, in a stable thermal state were examined. For scenarios involving heats directly after the installation of a new refractory lining, or extended turnover intervals between treatments due to logistical delays, the ladle is in a thermally transient state. Under such conditions, temperature losses exhibit substantial variability and cannot be reliably predicted using data-driven methods, as the available data is insufficient to characterize the thermal behavior. Q_{steel} describes the heat balance of the steel bath within the steel ladle.

The described influences on the steel temperature are summarized once again in the following Table I. By means of the law of energy conservation and the phenomena outlined previously, the following relationship can be defined to determine the heat balance of the steel bath:

$$Q_{steel} = Q_{arc} \pm Q_{addition} \pm Q_{reaction} - Q_{argon} - Q_{surroundings} - Q_{storage} \quad [1]$$

B. Hybrid Process Modeling Concept for LF Operation

A concept based on a novel hybrid approach was developed to predict the temperature and characterize the process phenomena. Figure 2 illustrates the architecture of this new model. Two pillars generally define the hybrid temperature modeling. On the one hand, all chemical reactions that take place within the steel ladle system are described utilizing the thermodynamic software products FactSageTM and ChemAppTM Python. The FactSageTM databases FactPS, FStel and FTOxid are used to explain the chemical reactions fully. The phases selected from the databases were the ones that

could form under the given process conditions and proved to be stable. To enable the Python program to use these predefined databases, ChemSage™ files (*.cst) containing all the information are exported using the FactSage software. ChemApp Python provides an application programming interface to read the extracted files into the developed software. The kinetics within the system is defined macroscopically using the EERZ^[33] method. The model's second pillar deals with the detailed description of the process boundary conditions. Depending on the type of boundary condition, these are defined using statistical learning methods or empirical equations. The description of the required process boundary conditions (e.g., influence of heating using carbon electrodes, argon purging using bottom purging on the steel bath temperature) *via* statistical learning methods generally requires large amounts of data. These were provided by the industrial partner voestalpine Stahl GmbH and included all process measures, treatment steps, and temperature measurements during LF treatment. The data also contain the starting and boundary conditions required for simulation and temperature prediction.

To gain a better understanding of the LF treatment as performed by voestalpine Stahl GmbH and the possible process influences, Figure 3 shows the schematic process sequence. Process phases in which only heating and purging were carried out and no alloy additions were conducted are used for the statistical evaluation in this study. The temperature influences due to various alloy additions and the related chemical reactions are covered by the EERZ methodology.

Based on the abovementioned architecture, a systematic process flow diagram was developed, as shown in Figure 4. Starting with step I, the steel bath and slag have an initial temperature (T_{current}), chemical composition, and quantity entering the next iteration step. Depending on the different mass transfer coefficients, a certain proportion of the steel and slag bulk is extracted (step II) and made available for the EERZ reaction. The reaction materials also have the same initial temperature and chemical composition as the bulk material of the respective medium. In subsequent step III, the addition of light alloying agents (aluminum granules or carbon carriers) is also considered, taking into account the alloying schedule. It is assumed that the added materials have a medium composition as measured by voestalpine Stahl GmbH and enter the reaction zone at room temperature ($T_{\text{room}} = 20\text{ °C}$). The thermodynamic equilibrium for the reaction zone is calculated, providing the amount of individual stable phases and the enthalpy change and temperature (T_{equilib}). In step IV, the phases formed in the equilibrium calculation are now divided into the steel and slag phases accordingly. These quantities are then returned to the bulk material. Again, alloys (e.g., FeMn, FeSi, etc.) and slag formers (e.g., alumina, lime, etc.) can be added depending on the schedule. The composition of these alloy materials also corresponds to the average measured analysis values provided by voestalpine Stahl GmbH. The addition has the same temperature as the surroundings. This fifth process step then includes the calculation of the thermodynamic equilibria of the two

bulks, which provides the new chemical composition for the next time step. Moreover, calculating the thermodynamic equilibria provides the temperatures (T_{steel} and T_{slag}), specific heat capacities, quantities, and enthalpy changes. The new temperature of the system is determined in the next step, VI, based on the information obtained from the previous equilibrium calculations and the associated process influences according to the process schedule.

1. Thermodynamically based approach

Using the thermodynamic databases and considering the incoming reaction quantities, their dissolution time, and the subsequent chemical reaction, it is possible to determine whether the system requires energy (endothermic reaction) and thus heat or whether an exothermic process occurs and the steel experiences an increase in temperature. The effective equilibrium reaction zone^[33,45,48] method is employed to quantify the temperature effects of alloy additions and chemical reactions within the steel ladle. The method makes it possible to link the kinetic processes occurring within the system with fundamental thermodynamics. Overall, it can be stated that the EERZ method attempts to simplify the complex kinetics at the interface. One of these assumptions is that, in contrast to the coupled reaction models,^[44,49] the mass transfer coefficient k_i is assumed to be identical for all species of a specific bulk (steel or slag). The reaction quantity RA of the respective bulk material can be calculated with the following relationship:

$$RA = (k\rho A)\Delta t, \quad [2]$$

where k (m s^{-1}) describes the overall mass transfer coefficient common to all species, ρ (kg m^{-3}) the density of the respective medium, and A (m^2) the reaction area between the steel bath and the slag. Δt indicates the simulation time step. The mass transfer coefficient k_{Steel} was determined based on empirical investigations by Graham and Irons.^[44,50] It was shown that it primarily depends on the effective stirring rate per ton of steel ε (W t^{-1}), whereby the relationship was defined using Eq. [3]. As the slag has a significantly higher viscosity and its components show lower diffusion coefficients, it is suggested that the slag's mass transfer coefficient k_{slag} is one-twentieth of the steel, see Eq. [4].

$$k_{\text{Steel}} = (0.006 \pm 0.002)\varepsilon^{1.4 \pm 0.09} \quad [3]$$

$$k_{\text{Slag}} = \frac{k_{\text{Steel}}}{20}. \quad [4]$$

The purging rate can be calculated based on the following empirical equation:

$$\varepsilon = \left(\frac{nRT}{m}\right) \ln\left(\frac{P_t}{P_0}\right), \quad [5]$$

where n (mol s^{-1}) describes the molar gas flow rate, R ($\text{J mol}^{-1} \text{K}^{-1}$) is the ideal gas constant, T (K) is the absolute temperature, and m (kg) is the steel mass.

Table I. Process Influences for Temperature Prediction of Molten Steel

Variable	Description	Units
x_{Heating}	electrical heating amount <i>via</i> carbon electrodes	kWh
x_{Stirring}	argon stirring via bottom purge	min
x_{Duration}	ladle furnace refining duration	min
x_{Weight}	weight of the molten steel	kg
$x_{\text{Initial_Temp}}$	initial temperature of molten steel at LF	°C
x_{Turnover}	turnover cycle of steel ladle	min
x_{Alloy}	addition amount of alloy and slag former	kg

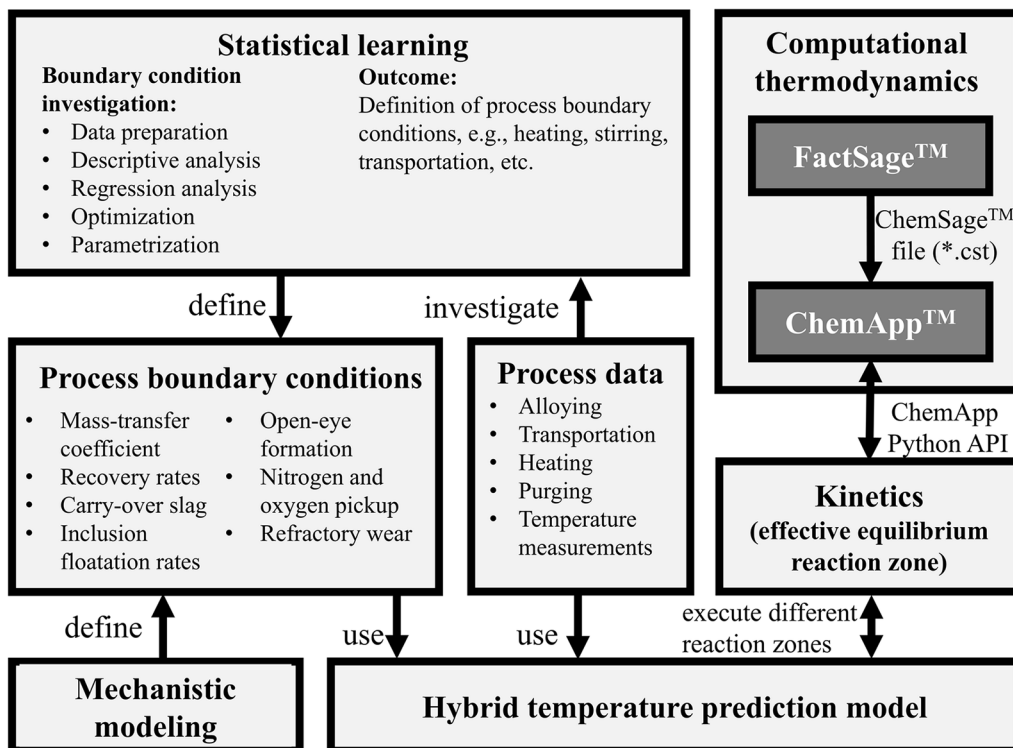


Fig. 2—Concept of the hybrid data-driven thermodynamically based temperature prediction model.

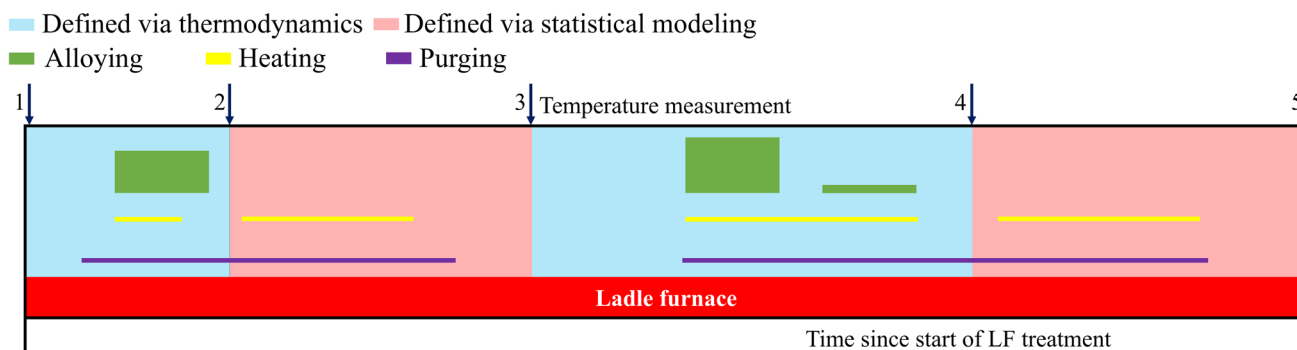


Fig. 3—Schematic representation of LF treatment, including all temperature measurements.

The pressure P_t (atm) describes the prevailing pressure at the bottom of the steel ladle. P_0 (atm) is the surrounding atmospheric pressure, respectively. The EERZ method considers all reactions between steel

bath and slag, steel bath and atmosphere, based on the modeling of Zhang *et al.*,^[51] and within the bulk materials. The EERZ method was modeled using FactSage Version 8.3 and ChemApp Version 3.9.7.

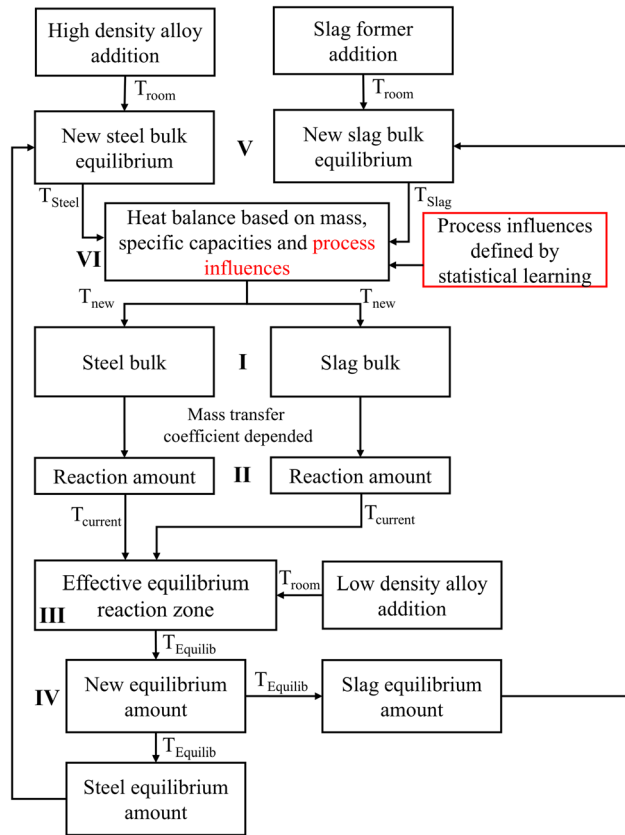


Fig. 4—Flow chart of hybrid temperature model.

2. Statistical learning methods

Using statistical learning methods often requires extensive data preparation at the beginning. This includes removing or imputing outliers and missing values, normalization, transformation, and identifying possible background noise and feature extraction.^[52] Yet, not all of these methods have been performed during the course of the current work. The Isolation Forest^[53] algorithm and a visual examination of the data *via* scatterplots were applied for outlier detection. For the Isolation Forest algorithm, the default parameter values as defined in scikit-learn version 1.2.0 were employed. Scatterplots were also used to visualize the relationship between the independent and target variables. Furthermore, the Pearson correlation matrix was selected to identify a possible linear relationship or collinearity between the investigated variables. This method is based on the following Eq. [6], where x_i and y_i describe the values of the respective property of an observation and \bar{x} or \bar{y} the mean value of the properties across all observations.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}. \quad [6]$$

The data used were not normalized nor transformed. Missing values were removed from the data set in a listwise manner due to the high number of observations. Once the missing values and the outliers were removed

from the dataset Little's^[54] missing completely at random (*MCAR*) test was performed and it was shown that the missing values were in the *MCAR* category. Thus, the listwise deletion also does not influence the distribution of the features and does not lead to any distortion of the statistical parameters. A feature extraction algorithm was not applied, as only three independent variables were investigated in greater detail after the scatterplot and Pearson correlation analysis were conducted. After the data preparation and the descriptive examination of the provided data, a classic 70-15-15 pct train validation test split^[55] was applied. The first data set was used to train the statistical model. The best coefficients for the respective model could thus be identified. The validation dataset was utilized to select the best model. Ultimately, the test data set should represent a new set of data independent of the modeling, with which all result diagrams are created. Several evaluation metrics^[56] were used to validate the performance of the statistical regression models, including the root-mean-square error (*RMSE*), the mean absolute error (*MAE*), and the coefficient of determination (R^2). The *RMSE*, see Eq. [7], is the square root of the *MSE*, which reflects the average sum of the differences e_i between predicted and measured values divided by the number of observations n .

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |e_i|^2}. \quad [7]$$

The mathematical formulation of the *MAE* is as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i|. \quad [8]$$

The definition of R^2 can be expressed by the Sum of Squares Total (*SST*) and Sum of Squares Error (*SSE*) or Sum of Squares Residuals (*SSR*), given by Eq. [9]. The description of *SST*, *SSE* and *SSR* can be taken from Eqs. [10] through [12]. y_i represents the actual measured value, \bar{y}_i the mean of all observed values and \hat{y}_i the predicted value.

$$R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST} \quad [9]$$

$$SST = \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad [10]$$

$$SSE = \sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2 \quad [11]$$

$$SSR = \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad [12]$$

In order to predict the model and estimate the coefficients of the process influences, multiple linear regression^[57] was performed. The MLR can be formulated according to Eq. [13].

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e} = \hat{\beta}_1 \mathbf{x}_1 + \hat{\beta}_2 \mathbf{x}_2 + \dots + \hat{\beta}_n \mathbf{x}_n + \mathbf{e}, \quad [13]$$

where \mathbf{y} is the actual measured value of the target variable, $\hat{\mathbf{y}}$ is the estimated target vector, $\hat{\beta}_n$ is the estimated regression coefficient, \mathbf{x}_p is the vector of an independent variable p and \mathbf{e} is the residuals vector. At this point, it should be mentioned that the model predictions using MLR were made without an intercept variable. This is because the target variable is the temperature difference ΔT ($^{\circ}\text{C}$) between two consecutive temperature measurements. For example, in Figure 3, this would correspond to the temperature difference between samples 2–3 and 4–5. Equation [14] describes the explained relationship mathematically.

$$\Delta T = T_{\text{target}} - T_{\text{initial}}. \quad [14]$$

In addition to the multiple linear regression model, a Random Forest regression optimized *via* a random search algorithm was employed for comparative analysis. This aimed to assess whether a more complex model could better capture potential non-linear relationships within the data. Since the Random Forest served primarily as a benchmark to evaluate the validity of the linearity assumption, its implementation is not described in detail here. For a comprehensive overview of the algorithm, one is referred to Reference 57.

Python version 3.9.7, numpy version 1.23.5, pandas version 1.5.2, matplotlib 3.9.2, and scikit-learn version 1.2.0 were used to analyze the data.

3. Other boundary conditions

The model also includes the refractory wear material from the steel and slag zones as an additional input stream into the system. Based on measurements carried out by the industry partner voestalpine Stahl GmbH, an average of 50 kg of refractory material is worn out per melt. Two-thirds of the material is from the steel bath area, and one-third is from the slag zone. Such refractory wear influences the composition of the slag and, thus, the thermodynamic behavior of the system and, subsequently, temperature development, even if this fraction is negligible. The alloy yield during ladle furnace treatment corresponds to the results presented by Kavić *et al.*^[45]

IV. RESULTS AND DISCUSSION

A. Statistical Learning from LF Process Data

The industrial partner provided a data set containing the entire annual process information of voestalpine Stahl GmbH's four ladle furnaces. Two data sets were available after various data preparation steps were performed, as described in Section III-B-2. The first data set (12088 observations) contained process sequences where an electrical heating operation occurred between two consecutive temperature measurements. The second data set, in which no electrical heating period took place, provided a data set with 2537 observations and was therefore still comprehensive enough to reflect a high statistical significance. All

observations represent process sequences where no alloy additions were conducted, and thus, only the process influences and their effects on the steel bath temperature could be investigated (see Figure 3). In the following, the results of the first data set are discussed in detail. To understand the interaction between the process influences and the steel bath temperature, the Pearson correlation results, see Figure 5(a), and the scatterplot analyses in Figures 5(b) through (d) were carried out to make the findings visually interpretable. As expected, Figure 5(a) shows that, in particular, the heating amount has a strong positive correlation with the temperature change, resulting in a Pearson correlation coefficient of 0.9. As expected, a negative correlation ($r = -0.4$) exists between the sample time difference and the temperature change measured. These relationships can be seen more clearly by looking at the scatterplots, which show linear trends between the three variables (see Figures 5(b) and (c)). However, the Pearson correlation matrix shows that the stirring duration positively correlates with the temperature change. The scatterplot in Figure 5(d) illustrates why this unnatural result was obtained. The high correlation between the heating quantity and the temperature change overshadows the less strong correlation between the stirring duration and the temperature change. Nevertheless, the color gradients in the diagram indicate that the temperature change decreases constantly with increasing stirring gas treatment. The steel and slag masses showed no significant correlation, as indicated by the Pearson correlation (< 0.2), and the scatterplot likewise revealed no clear relationship; therefore, its explicit visual representation is omitted. This lack of correlation originates from only slight deviations in the reported steel and slag quantities. Based on the findings from the descriptive statistics, a model with excellent predictive power can be obtained using the MLR and taking into account the three significant temperature-influencing parameters (the influence of the heating by carbon electrodes, the stirring amount, and the treatment duration). When comparing the calculated and measured temperature values, the model provides an R^2 of 0.93, an $RMSE$ of 4.27 $^{\circ}\text{C}$, and a value of 3.36 $^{\circ}\text{C}$ for the MAE (see Figure 5(e)) using the test data set. It should be mentioned at this point that the temperature measurements have a standard deviation of approximately 2 $^{\circ}\text{C}$; therefore, each measurement is associated with a certain degree of uncertainty. Nevertheless, almost 77 pct of the calculations are within ± 5 $^{\circ}\text{C}$ error tolerance. Within ± 10 $^{\circ}\text{C}$, 97 pct of the observations can be reconstructed. From Figure 5(f), it is observed that the error distribution between the calculated and measured temperature values is approximately normally distributed. Due to the excellent predictive power of the MLR model, the regression coefficients can now be extracted and used for the hybrid temperature model as a process boundary condition for the respective process phase. The coefficient for a heating quantity of 1000 kWh is + 18 $^{\circ}\text{C}$. A purging treatment leads to a temperature reduction of -0.48 $^{\circ}\text{C min}^{-1}$. The temperature losses due to the measurement time difference

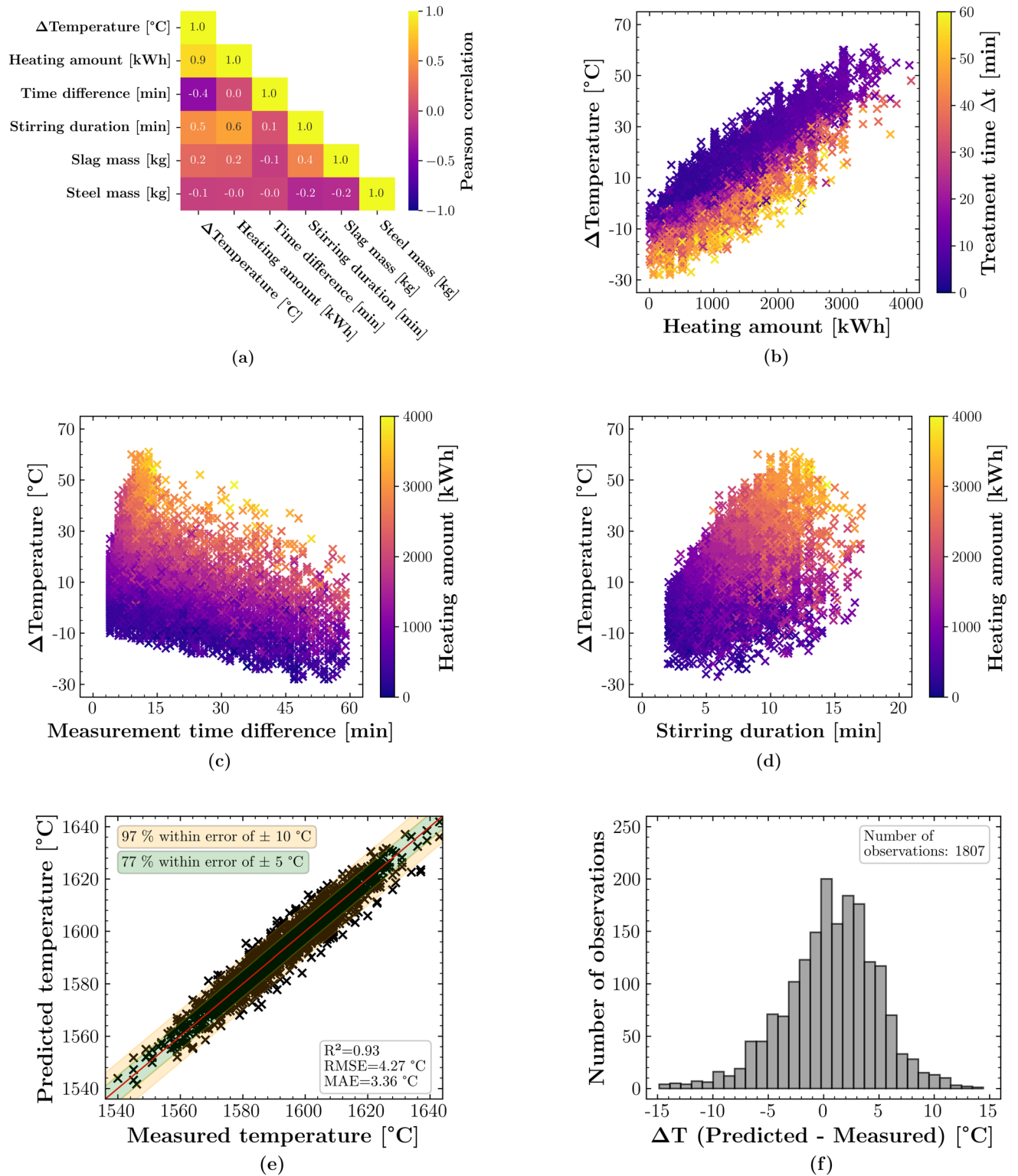


Fig. 5—(a) Pearson correlation matrix between the temperature change and diverse process influence parameters; (b) temperature change plotted against the heating rate, (c) the time span between the two temperature measurements, and (d) the stirring gas treatment time; (e) comparison of the calculated and measured temperatures; (f) error distribution between the calculated and measured temperature values.

are $-0.56\text{ }^\circ\text{C min}^{-1}$, representing the overall cooling of the ladle temperature during LF treatment. Compared to the multiple linear regression, the Random Forest

model achieved an R^2 value of 0.94, an $RMSE$ of $4.06\text{ }^\circ\text{C}$, and a MAE of $3.12\text{ }^\circ\text{C}$. Since the performance improvement is marginal, this indicates that the

relationship between the independent and dependent variables is predominantly linear, with no significant hidden non-linearities present.

The second data set was examined following a similar procedure. After analyzing the scatterplots, only the time between the two temperature measurements and the stirring duration negatively influences the steel bath temperature. An MLR with these two parameters gives an R^2 of 0.94, an $RMSE$ of 4.34 °C, and an MAE of 3.33 °C (see Figure 6(a)). 96 pct of the predictions are within ± 10 °C. The ± 5 °C interval is achieved by 79 pct of the forecasts. Again, the errors between the predicted and measured temperatures are normally distributed, as illustrated in Figure 6(b). The optimized regression model provides the following coefficients: stirring leads to a temperature reduction of -0.79 °C min^{-1} , and the temperature loss due to the measurement time difference is -0.53 °C min^{-1} . Also, for the second data set a Random Forest model was applied, yielding an R^2 of 0.94, an $RMSE$ of 4.16 °C, and an MAE of 3.15 °C, showing only a slight improvement over the multiple linear regression. This marginal gain suggests that the data's structure is largely linear, and complex non-linear effects have little influence.

B. LF Process Simulation by the Hybrid Model

The following section provides a detailed discussion of the temperature model results for selected case studies conducted at voestalpine Stahl GmbH's steel plant in Linz. Comprehensive process data and initial conditions are required to simulate the temperature development using the hybrid temperature model during ladle furnace treatment. Unfortunately, the process data recorded during standard operation cannot be used for the hybrid temperature model. This is because no slag sample is taken at the beginning of the ladle furnace treatment during regular operation. Therefore, small-scale trials

must be carried out in which the samples are extracted from the slag bulk. A steel sample is collected at the end of the tapping process and used as the initial composition for the simulation.

The analyzed melts belong to a steel grade within the high-strength micro-alloyed structural steel group. The composition range of the respective elements before LF treatment is shown in Table II.

The slag sample is taken at the start of the ladle furnace process and provides the initial values for the slag composition. As shown in Table III, the range of the initial compositions is significantly wider than for the steel samples. Partly, this broader range can be attributed to the inhomogeneity of the slag at the beginning of the ladle furnace process as well as to the difficulties of sampling itself.

The heats have been deoxidized with Al during the tapping process. Besides the compositions, the initial masses of the two bulks are also important. These values are determined by lifting the steel ladle by crane after tapping and considering all slag formers, carry-over slag, etc., during tapping; then, the quantity of steel and slag can be recalculated. These values provided by voestalpine Stahl GmbH were applied as initial quantities. The average amount of steel was nearly 182 tons, and the amount of slag ranged between 2 and 3 tons. During the LF process, different treatment schedules are carried out. The treatment duration is between 30 and 90 minutes, depending on the performed operation. Table IV contains relevant information about the case studies (I to V).

The predicted temperature curves of the five case studies, presented in the following Figure 7, clearly show excellent agreement with the temperatures measured during the actual operation. As expected, the model reacts with a temperature increase during heating and increased temperature loss during stirring phases. Moreover, the alloying processes generally lead to a

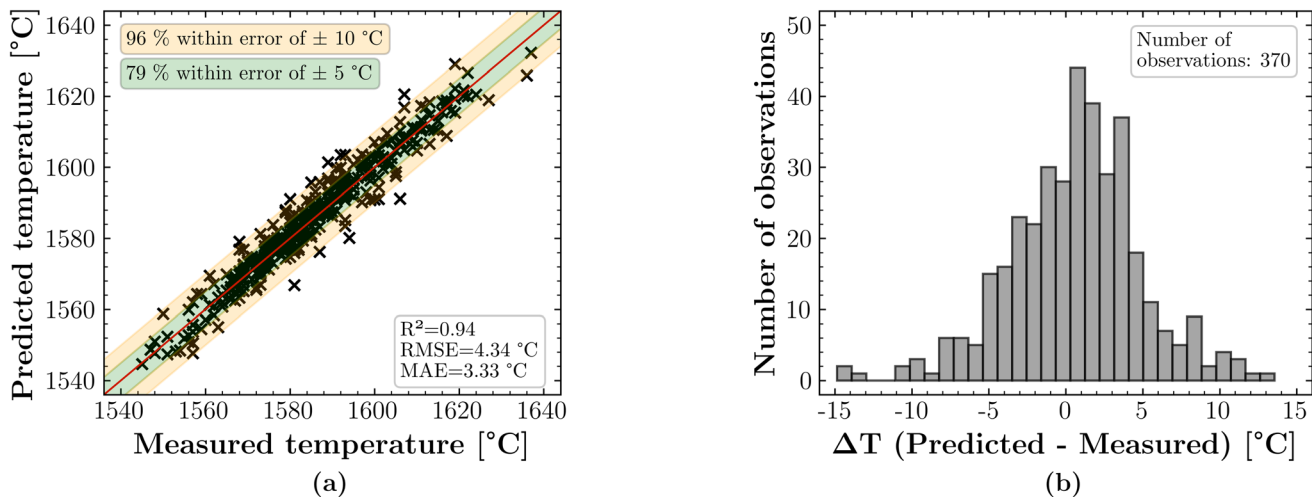


Fig. 6—(a) Comparison of the calculated and measured temperatures and (b) error distribution between the calculated and measured temperature values for the heating-free process treatment phase.

temperature loss. Only aluminum and FeSi (Si content of 75 pct) led to a temperature increase, but this is not visible due to the low addition amounts. In the following, the individual case studies are discussed in detail. Note that all the investigated melts have a short turnover time, so the ladle refractory is in thermal equilibrium.

Starting at a temperature of 1550 °C, the melt from case study I undergoes a significant temperature increase during the treatment and has a final temperature of 1609 °C at the time of the last temperature measurement. Considering all temperature measurements of this case, an *MAE* for the temperature deviation of 1.7 °C is obtained. Case study II starts with a temperature of 1581 °C and rises to 1626 °C by minute 26 due to high heating power at the beginning of the treatment. Subsequently, the injection of quicklime and the purging treatment led to a significant decrease in temperature, reaching 1590 °C at $t = 40$ minutes. Finally, a slight heating process is made to maintain the temperature. The simulation ultimately results in a deviation of 4.8 °C for the last measurement. The *MAE* for the second case is 1.91 °C. Case study III also shows a significant temperature increase of more than 50 °C in the beginning. Quicklime is injected from minute 34, and the melt is homogenized *via* bottom purging. This leads to a significant drop in the temperature, which is then raised again to a final temperature of 1638 °C by large amounts of heating during further treatment. At 2.28 °C, the *MAE* is again in a very small range. Case study IV shows a relatively moderate increase at the beginning of the treatment. With quicklime and subsequent purging, a temperature of 1562 °C is reached. The melt is then heated to a temperature of over 1600 °C. At the end of the process, the temperature deviation between the measured and calculated value is 4.8 °C. With an *MAE* of 3.63 °C, case study IV shows the highest temperature deviation. The last case study examined shows a similar trend in the first section of the treatment. In contrast to the other case studies, this treatment experiences an interruption where no process influences have taken place for more than half an hour starting at minute 28. Reasons for this could be the steel mill logistics, *e.g.*, because there may have been delays at the continuous casting plant, and the melt is therefore not required until later. At the 60th minute of the simulation, the ladle furnace treatment started again, and a short heating and stirring gas treatment phase for thermal homogenization of the steel bath was conducted. The last two temperature measurements at the ladle furnace show a deviation of less than 5 °C from the predicted simulation results. This allows the statement to be made that the temperature prediction can also reproduce longer treatment times with an interim interruption very well. Ultimately, case study V delivers an *MAE* of 1.70 °C.

In addition to the five case studies presented, another four heats corresponding to the same steel grade were sampled and afterward examined using the developed hybrid temperature model. All results of the investigated cases are shown in Figure 8. The predicted temperatures

were compared with the actual measurements, and it was found that a maximum temperature deviation of ± 2.5 °C could be achieved for 63 pct of the measurements. 94 pct of the measurements were within a ± 5 °C range. Considering all observations, the *MAE* is 2.30 °C, the *RMSE* is 2.85 °C and the R^2 is 0.97.

V. CONCLUSIONS

This work first provides a comprehensive overview of various temperature modeling approaches for secondary steelmaking developed and published over the last two decades. The considered models can be divided into three categories: (i) mechanistic, physics-based (“white-box”) models, (ii) data-driven (“black-box”) models, which are based on extensive process data, and (iii) hybrid (“gray-box”) models, that combine elements of the two previously mentioned approaches. Subsequently, the considered process-related temperature influences (*e.g.*, heating through carbon electrodes, stirring *via* bottom plugs, alloying, etc.) are discussed, whereby a novel hybrid approach for temperature prediction at the ladle furnace is proposed. The hybrid model was implemented by applying the EERZ method coupled with the definition of thermal boundary conditions *via* statistical learning methods. The EERZ method quantifies the required or released heat amounts for all chemical reactions within the steel ladle and ongoing dissolution processes of the added alloying materials. To define the process influences using data-driven methods, extensive amounts of data were handled, outliers were removed using different techniques, and multiple linear regression was performed to obtain the coefficients of the thermal boundary condition parameters. For the process phases where heating was applied according to the process schedule, it was shown that 1000 kWh results in a temperature increase of 18 °C, whereas stirring results in a temperature reduction of -0.48 °C min⁻¹, and for the treatment duration between two consecutive samples, a value of -0.56 °C min⁻¹ was defined. During the process phases where no heating occurred, the temperature drop for the treatment duration was -0.53 °C min⁻¹. Stirring during this process phases lead to a temperature reduction of -0.79 °C min⁻¹. The final model was validated based on nine case studies; five of them were explained in greater detail. The hybrid model has the ability to reproduce more than 94 pct of the temperature measurements with a maximum deviation of 5 °C and delivers an *MAE* of 2.30 °C, an *RMSE* of 2.85 °C and an R^2 of 0.97. The proposed hybrid temperature prediction approach demonstrates excellent performance and represents a valuable extension of the previously published model by Kavić *et al.*,^[45] which focused on detailed predictions of steel and slag composition. Moreover, the approach can be readily adapted to other refining processes, such as the Ruhrstahl-Heraeus (RH) vacuum degassing stage.

Table II. Typical Steel Composition of the Investigated Steel Grade Before LF Treatment

C [Wt Pct]	Si [Wt Pct]	Mn [Wt Pct]	S [Wt Pct]	P [Wt Pct]	Al [Wt Pct]	Cr [Wt Pct]	N [Wt Pct]
0.011 to 0.014	0.15 to 0.17	1.08 to 1.3	0.008 to 0.019	0.005 to 0.014	0.02 to 0.09	0.014 to 0.045	0.0036 to 0.0061

Table III. Typical Slag Composition of the Investigated Steel Grade

CaO [Wt Pct]	SiO ₂ [Wt Pct]	Al ₂ O ₃ [Wt Pct]	MgO [Wt Pct]	Mn _{Total} [Wt Pct]	Fe _{Total} [Wt Pct]
45 to 51	2.8 to 10.2	25.6 to 35.1	9 to 16	0.07 to 2	0.5 to 1.2

Table IV. Process Influences During Ladle Furnace Treatment

Case	Heating		Stirring		Type	Addition	
	Time Range [min]	Amount [kWh]	Time Range [min]	Rate [Nm ³ /h]		Time [min]	Amount [kg]
I	0 to 3.2	866	0 to 4.2	64	Al-granulate	11	40
	8 to 10.6	667	8 to 29.4	76	FeSi	11	44
	11 to 17.8	2000			alumina	12	98
	18 to 19.7	519			FeMn	12	254
	21 to 23.9	802			lime	13	300
	24 to 26.3	705			Alumina	17	100
					lime	17	400
					lime	24	401
					Alumina	24	97
					Al-granulate	11	150
II	3 to 4	193	3 to 4.8	12.7	Al-granulate	11	150
	7 to 10.8	985	7 to 23.3	61.5	FeSi	12	44
	13 to 26	3209	25 to 46.3	61.6	FeMn	12	169
	40 to 43	813	56 to 61.9	98.7	lime	12	106
	59 to 62	549			alumina	13	300
					alumina	16	95
III	0 to 8	2392	0 to 6	17.7	Lime	17	299
	18 to 21.1	800			quicklime	26	500
	23 to 33	2699	6 to 8.8	54.1	Al-granulate	23	122
	47 to 50.8	1011	12 to 13.2	35	FeSi	24	43
	52 to 62.4	2938	14 to 21	11.7	FeMn	24	388
	74 to 77.9	1011	21 to 28.6	59.6	lime	25	602
	80 to 83.8	1012	28.6 to 62	59.6	alumina	25	200
			62 to 83	24.4	quicklime	34	500
IV	3 to 5.3	604	1 to 31.6	87	Al-granulate	73	75
	6 to 7	198			alumina	73	92
	8 to 12	1200			Al-granulate	3	100
	21 to 26	1522			alumina	4	202
	27 to 32	1522			lime	6	801
					FeSi	12	83
V	3 to 6	966	2 to 6.1	54.1	FeMn	12	223
	8 to 10	469	8 to 30.3	59.4	quicklime	13	500
	12 to 18	1839	60 to 71	69	Al-granulate	12	77
	21 to 24.5	1016			FeSi	12	55
	25.6 to 28.6	717			FeMn	12	100
	61 to 64.3	642			lime	13	800
	70 to 72.25	408			alumina	14	210
					quicklime	22	118
				lime	23	400	

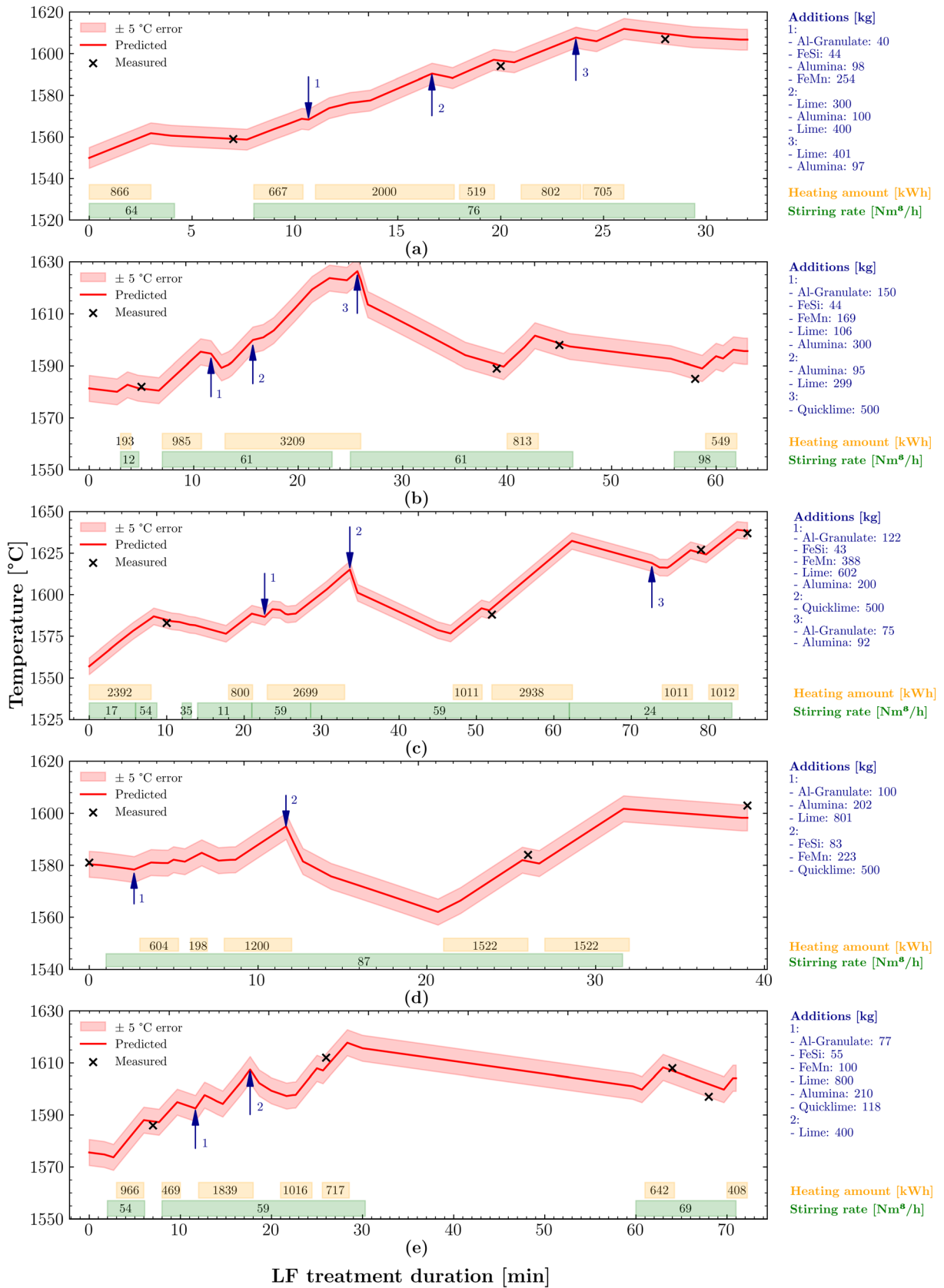


Fig. 7—Calculated temperature evolution of test cases listed in Table IV along with plant measurements. (a) Case study I, (b) case study II, (c) case study III, (d) case study IV, and (e) case study V.

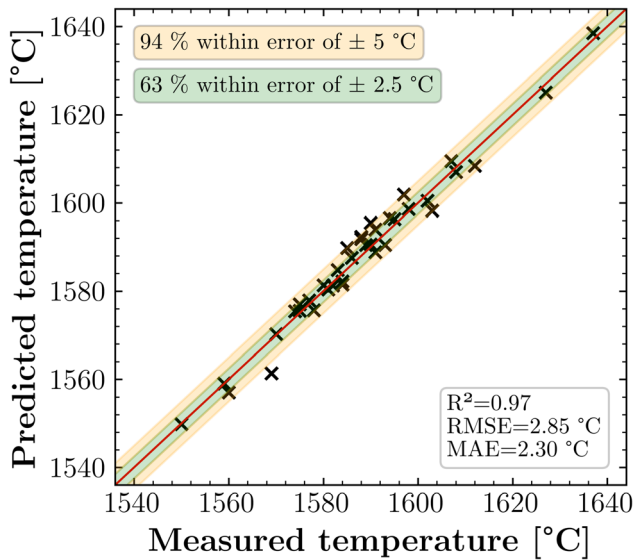


Fig. 8—Predicted vs measured temperatures for all investigated case studies.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the funding support of K1-MET GmbH, metallurgical competence center. The research program of the K1-MET competence center is supported by COMET (Competence Center for Excellent Technologies), the Austrian program for competence centers. COMET is funded by the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology, the Federal Ministry for Labor and Economy, the Federal States of Upper Austria, Tyrol and Styria as well as the Styrian Business Promotion Agency (SFG) and the Standortagentur Tyrol. Furthermore, Upper Austrian Research GmbH continuously supports K1-MET. Beside the public funding from COMET, this research project is partially financed by the Montanuniversität Leoben and the industrial partners Primetals Technologies Austria, RHI Magnesita, voestalpine Stahl, and voestalpine Stahl Donawitz.

FUNDING

Open access funding provided by Montanuniversität Leoben.

CONFLICT OF INTEREST

On behalf of the authors, the corresponding author states that there is no conflict of interest.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and

indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

1. Ü. Çamdali, M. Tunç, and F. Dikeç: *Appl. Therm. Eng.*, 2001, vol. 21, pp. 643–55.
2. Ü. Çamdali and M. Tunç: *Can. Metall. Q.*, 2003, vol. 42, pp. 439–46.
3. Ü. Çamdali and M. Tunç: *J. Iron. Steel Res. Int.*, 2006, vol. 13, pp. 18–20.
4. A.G. Belkovskii and Y.L. Kats: *Metallurgist*, 2009, vol. 53, pp. 261–73.
5. S.D. Kathait: *Int. Res. J. Eng. Technol.*, 2016, vol. 3, pp. 1627–31.
6. I.D. Duarte, C.A. Da Silva, I.A. Da Silva, E.R. Ferreira, and V. Seshadri: *Adv. Mater. Res.*, 2015, vol. 1125, pp. 166–70.
7. J. Zabadal, M. Vilhena, and S.Q. Bogado Leite: *Ironmak. Steelmak.*, 2004, vol. 31, pp. 227–34.
8. A. Mukhopadhyay, P. Deb, A. Ghosh, B. Basu, R. Dutta, and P. Kumar Mukhopadhyay: *Steel Res.*, 2001, vol. 72, pp. 192–99.
9. P. Deb, A. Mukhopadhyay, A. Ghosh, B. Basu, S. Paul, G. Mishra, J. Kumar Saha, R. Sharma, and S. Sarkar: *Steel Res.*, 2001, vol. 72, pp. 200–07.
10. A. Tripathi, J.K. Saha, J.B. Singh, and S.K. Ajmani: *ISIJ Int.*, 2012, vol. 52, pp. 1591–600.
11. M.-A. van Ende and I.-H. Jung: *Metall. Mater. Trans. B*, 2017, vol. 48B, pp. 28–36.
12. Y. Wang, Y. Bao, H. Cui, B. Chen, and C. Ji: *J. Iron. Steel Res. Int.*, 2012, vol. 19, pp. 1–5.
13. S.S. de Matos, C.A. da Silva, J.J. Mol Peixoto, E.N. de Almeida, W.J.C. da Conceição, and I.C. Lima: *Ironmak. Steelmak.*, 2023, vol. 50, pp. 1659–67.
14. X. Wang, P. Yuan, Z. Mao, and M. You: *Knowl. Based Syst.*, 2016, vol. 101, pp. 48–59.
15. X. Wang: *IEEE/CAA J. Autom. Sinica*, 2017, vol. 4, pp. 770–74.
16. S. Sonoda, N. Murata, H. Hino, H. Kitada, and M. Kano: *ISIJ Int.*, 2012, vol. 52, pp. 1086–91.
17. I.C.D. Duarte, G.M. de Almeida, and M. Cardoso: *J. Oper. Res. Soc.*, 2022, vol. 73, pp. 326–37.
18. Ł Sztangret, K. Regulski, M. Pernach, and Ł Rauch: *Coatings*, 2023, vol. 13, p. 1504.
19. J. Yang, J. Zhang, W. Guo, S. Gao, and Q. Liu: *ISIJ Int.*, 2021, vol. 61, pp. 2100–10.
20. H.-X. Tian, Z. Mao, and A. Wang: *ISIJ Int.*, 2009, vol. 49, pp. 58–63.
21. H.-X. Tian, Y.-D. Liu, K. Li, R.-R. Yang, and B. Meng: *ISIJ Int.*, 2017, vol. 57, pp. 841–50.
22. Z. Xin, J. Zhang, J. Zhang, J. Zheng, Y. Jin, and Q. Liu: *Metall. Mater. Trans. B*, 2023, vol. 54B, pp. 1181–94.
23. W. Lv, Z. Mao, and P. Yuan: *Steel Res. Int.*, 2012, vol. 83, pp. 288–96.
24. F. He, A. Xu, H. Wang, D. He, and N. Tian: *Steel Res. Int.*, 2012, vol. 83, pp. 1079–86.
25. M.K. Singh, A. Choudhury, D. Uikey, and S. Pal: *Trans. Indian Inst. Met.*, 2023, vol. 76, pp. 3365–77.
26. H.-X. Tian, Z.-Z. Mao, and Y. Wang: *ISIJ Int.*, 2008, vol. 48, pp. 58–62.
27. F. He, D. He, A. Xu, H. Wang, and N. Tian: *J. Iron. Steel Res. Int.*, 2014, vol. 21, pp. 181–90.
28. N. Gupta and S. Chandra: *ISIJ Int.*, 2004, vol. 44, pp. 1517–26.
29. N.K. Nath, K. Mandal, A.K. Singh, B. Basu, C. Bhanu, S. Kumar, and A. Ghosh: *Ironmak. Steelmak.*, 2006, vol. 33, pp. 140–50.

30. W. Lv, Z. Mao, P. Yuan, and M. Jia: *Steel Res. Int.*, 2014, vol. 85, pp. 405–14.
31. P.S. Srinivas, A.K. Kothari, and A. Agrawal: *ISIJ Int.*, 2016, vol. 56, pp. 977–85.
32. Z. Xin, J. Zhang, J. Zheng, Y. Jin, and Q. Liu: *ISIJ Int.*, 2022, vol. 62, pp. 532–41.
33. M.-A. van Ende, Y.-M. Kim, M.-K. Cho, J. Choi, and I.-H. Jung: *Metall. Mater. Trans. B*, 2011, vol. 42B, pp. 477–89.
34. D. You, C. Bernhard, P. Mayer, J. Fasching, G. Kloesch, R. Rössler, and R. Ammer: *Metall. Mater. Trans. B*, 2021, vol. 52B, pp. 1854–65.
35. D. You, C. Bernhard, A. Viertauer, and B. Linzer: *Crystals*, 2021, vol. 11, p. 893.
36. M.-A. van Ende and I.-H. Jung: in *Computational Materials System Design*, Springer, Cham, 2018, pp. 47–66.
37. H. Bossel: *Systeme, Dynamik, Simulation: Modellbildung, Analyse und Simulation komplexer Systeme*, Books on Demand, Norderstedt, 2004.
38. J.W. Hlinka, A.W. Cramb, and D.H. Bright: in *Steelmaking Conference 1985*, pp. 35–47.
39. O. Volkova and D. Janke: *ISIJ Int.*, 2003, vol. 43, pp. 1185–90.
40. A.H. Castillejos, M.E. Salcudean, and J.K. Brimacombe: *Metall. Mater. Trans. B*, 1989, vol. 20B, pp. 603–11.
41. J.S. Woo, J. Szekely, A.H. Castillejos, and J.K. Brimacombe: *Metall. Mater. Trans. B*, 1990, vol. 21B, pp. 269–77.
42. Y.Y. Sheng and G.A. Irons: *Metall. Mater. Trans. B*, 1993, vol. 24B, pp. 695–705.
43. T.P. Fredman: *Scand. J. Metall.*, 2000, vol. 29, pp. 232–58.
44. K.J. Graham and G.A. Irons: *Iron Steel Technol.*, 2009, vol. 6, pp. 164–73.
45. D. Kavić, M. Bernhard, R. Rössler, and C. Bernhard: in *AISTech 2024 Proceedings*, pp. 771–84.
46. S. Abraham, R. Bodnar, J. Raines, and Y. Wang: *J. Iron. Steel Res. Int.*, 2018, vol. 25, pp. 133–45.
47. M. Bernhard, D. Kavic, P. Presoly, T.-G. Wi, W.-B. Park, R. Rössler, A. Jungreithmeier, S. Ilie, C. Bernhard, and Y.-B. Kang: *Metall. Mater. Trans. B*, 2025, vol. 56B, pp. 2249–76.
48. D. You, S.K. Michelic, and C. Bernhard: *Steel Res. Int.*, 2020, vol. 91, p. 2000045.
49. A. Harada, N. Maruoka, H. Shibata, and S. Kitamura: *ISIJ Int.*, 2013, vol. 53, pp. 2110–17.
50. K.J. Graham: in *Proceedings AISTech2008*, Pittsburgh, 2008.
51. Y. Zhang, Y. Ren, and L. Zhang: *Metall. Res. Technol.*, 2018, vol. 115, p. 415.
52. S. Garcia, S. Ramirez-Gallego, J. Luengo, J. M. Benitez, and F. Herrera: *Big Data Anal.*, 2016, vol. 1(1), p. 9.
53. F. T. Liu, K. M. Ting, and Z.-H. Zhou: in *IEEE International Conference*, 2008, pp. 413–22.
54. R.J.A. Little: *J. Am. Stat. Assoc.*, 1988, vol. 83, pp. 1198–202.
55. Y. Xu and R. Goodacre: *J. Anal. Test.*, 2018, vol. 2, pp. 249–62.
56. A. Botchkarev: *IJKM*, 2019, vol. 14, pp. 45–76.
57. G. James, D. Witten, T. Hastie, R. Tibshirani, and J.E. Taylor: *An Introduction to Statistical Learning: With Applications in Python*, Springer, Cham, 2023.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.