



<http://www.diva-portal.org>

This is the published version of a paper published in *Steel Research International*.

Citation for the original published paper (version of record):

Kuthe, S., Alonso Oña, I., Glaser, B. (2025)

Adaptive Neuro#Fuzzy Inference System#Long Short#Term Memory Hybrid Model to
Forecast Castability of Al#Killed Steel Prior to Continuous Casting

Steel Research International, 96(8): 2400220

<https://doi.org/10.1002/srin.202400220>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-354805>

Adaptive Neuro-Fuzzy Inference System-Long Short-Term Memory Hybrid Model to Forecast Castability of Al-Killed Steel Prior to Continuous Casting

Sudhanshu Kuthe,* Izaskun Alonso Oña, and Björn Glaser

Dedicated to Prof. Em. Seshadri Seetharaman on the occasion of his 80th birthday

Continuous casting of aluminum (Al) deoxidized steels demands careful inspection due to the occurrence of submerged entry nozzle (SEN) clogging, leading to unexpected production stops. Recognizing the castability of a specific “cast” by monitoring the condition of the SEN is essential for uninterrupted casting. With this information prior to casting, operators can take preventive action against possible clogging occurrences, thus reducing unplanned downtimes. In response to the severe implications of SEN clogging, this work introduces a novel way to forecast castability of Al-killed steels. A hybrid model is proposed that integrates the adaptive neuro-fuzzy inference system (ANFIS) and long short-term memory (LSTM) networks. The output of the model helps to anticipate the event of clogging by analyzing both the past condition of the SEN and changes in the steel chemistry during the transport of the steel ladle from refining to the casting process. A comprehensive analysis of 150 casts helped to build the ANFIS algorithm for estimating the castability index (CI) parameter from steel chemistry. LSTM algorithm is used as a subsequent step to forecast castability in the next 20–25 min. Discrepancies between the predictive response and the actual conditions are reported. Although the real-time implementation of the proposed model is the ultimate goal, the focus of this work was to present the methodology and demonstrate its potential.


1. Introduction

Operators in a steel plant need to monitor various production events associated with steelmaking. Predicting stochastic notorious events in accordance with production planning is extremely important to avoid undesired errors causing unnecessary production halts. However, predicting these stochastic events in advance is challenging as steelmaking involves the interaction of complex metallurgical, physical and chemical reactions. In this regard, developing efficient process monitoring tools and operator-assisted decision support systems becomes evident. Such advancement is crucial for the robustness and reliability of production, ensuring uninterrupted and efficient production in the steel industry. In recent times, the use of artificial intelligence (AI) assisted predictive process models to predict the anomalies during steelmaking has become a popular area of research.^[1,2] To develop and implement these robust predictive models in production, an in-depth knowledge of factory physics, as well as a foundational understanding of the metallurgical domains are essential.^[3] This knowledge is particularly vital considering the different processing routes for steelmaking. Most steel plants around the world utilize either “blast furnace-basic oxygen furnace-ladle furnace-vacuum degassing” (BF-BOF-LF-VD) or “electric arc furnace-ladle furnace-vacuum degassing” (EAF-LF-VD) processes for manufacturing high quality liquid steel melts.^[4,5] After these processes, continuous casting (CC) is a widely adopted technology for casting of the produced liquid steels melts due to better control of macrostructure, surface properties, and increased productivity.^[6,7] Today, more than 95% of world’s total crude steel is processed via CC route.^[8,9] Once proper refining is achieved and the required target composition is reached, the liquid steel gets transferred from the secondary steelmaking ladle to the tundish and subsequently from the tundish to the mold. This transfer of liquid steel from tundish to mold is primarily carried out through submerged entry nozzle (SEN). The SEN is a tube-shaped device composed of refractory material mostly oxides,

standing of the metallurgical domains are essential.^[3] This knowledge is particularly vital considering the different processing routes for steelmaking. Most steel plants around the world utilize either “blast furnace-basic oxygen furnace-ladle furnace-vacuum degassing” (BF-BOF-LF-VD) or “electric arc furnace-ladle furnace-vacuum degassing” (EAF-LF-VD) processes for manufacturing high quality liquid steel melts.^[4,5] After these processes, continuous casting (CC) is a widely adopted technology for casting of the produced liquid steels melts due to better control of macrostructure, surface properties, and increased productivity.^[6,7] Today, more than 95% of world’s total crude steel is processed via CC route.^[8,9] Once proper refining is achieved and the required target composition is reached, the liquid steel gets transferred from the secondary steelmaking ladle to the tundish and subsequently from the tundish to the mold. This transfer of liquid steel from tundish to mold is primarily carried out through submerged entry nozzle (SEN). The SEN is a tube-shaped device composed of refractory material mostly oxides,

S. Kuthe, B. Glaser
Department of Materials Science and Engineering
KTH Royal Institute of Technology
SE-10044 Stockholm, Sweden
E-mail: kuthe@kth.se

I. A. Oña
Research and Development
Sidenor Aceros Especiales
Basauri ES-48970, Spain

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/srin.202400220>.

© 2024 The Author(s). Steel Research International published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/srin.202400220

which serves to protect the liquid melt from oxidation and to stabilize the casting process. The refractory device is suitable for casting various types of steel grades. However, low-carbon aluminum deoxidized steel grades usually face a major issue of SEN clogging leading to abrupt production stops at the time of CC operation.^[10] The clogging interrupts the melt flow and the transfer of the melt from the tundish to the copper mold, thereby affecting the “castability” of a produced steel grade.^[11–13] Reduced castability due to SEN clogging leads to loss of productivity.^[14,15] These issues significantly increase the frequency of interruptions in the CC operation leading to nozzle exchange and tundish maintenance.^[8,10,16,17] Therefore, it is important to carefully monitor the “castability” of steels during production. In addition, it is also important to maintain the right flow of melt into the mold to keep a constant steel bath level.^[18] The level of the melt in the tundish and mold is controlled with a sensor and a slide gate.^[8,19] The condition and level of the melt in the mold are controlled by the positions of a regulating device called a “stopper rod”.^[20] The device regulates the level of melt in the tundish in conjunction with the size of the opening between the stopper rod and the nozzle. The size of this opening varies over time due to the deposition of impurities where monitoring of the stopper rod movements can provide a probabilistic indication of an SEN clogging event. However, the issue of SEN clogging is complex and influenced by several factors such as steel composition and process parameters in relation to CC operations.^[15,21–23] Recent studies in modeling have significantly enhanced understanding of SEN clogging. The mathematical modeling of the early stages of SEN clogging has shed light on the initiation processes influenced by titanium nitrides and oxides.^[24] A transient model has been developed that describes the formation and subsequent peeling of clogs, providing a dynamic perspective on clogging occurrences.^[25] Further, the impact of air aspiration in slide-gate nozzles has been explored, revealing how this process can exacerbate clogging through the promotion of oxidative environments conducive to non-metallic inclusion (NMIs) formation.^[26] Additionally, recent work has focused on the numerical modeling of clog fragmentation, offering insights into the disintegration processes that affect continuous casting operations.^[27] In this work, a pioneering method is introduced that leverages a hybrid adaptive neuro-fuzzy inference system (ANFIS)^[28] and long short-term memory (LSTM)^[29] network to forecast castability with sufficient lead time, facilitating preventive actions to mitigate potential SEN clogging incidents in the future. The aim of this work was to demonstrate the potential of using data-centric methods to enhance the operations during the continuous casting process of Al-deoxidized SAE1055 steel grade. In recent times, the use of LSTM networks model to forecast dynamic casting variable has been prevalent, as it has ability to remember patterns over long sequences, making this model ideal for time-series forecasting.^[30,31] However, the use of an LSTM networks model alone may not fully cater to the complexities of the steel casting process, particularly in understanding and predicting the castability parameters solely based on changes of steel chemistry from the ladle and tundish. Steel chemistry in real steel production is not measured as a continuous time-series data stream but sampled at specific intervals. This makes it unsuitable for LSTM networks models to excel in learning from sequences. In this context, the ANFIS model showed promising results. ANFIS, a special kind of fuzzy inference system (FIS),^[28] can be

designed to make use of data that isn’t strictly numerical. Its adaptive nature allows it to learn complex, nonlinear relationships between inputs and outputs, something especially pertinent when handling the uncertainties and complexities tied to steel chemistry from the ladle and tundish.

2. Background

CC process despite its broad adoption in steel production, has been persistently plagued by the challenge of SEN clogging events. These disruptive and stochastic events, specifically for low-carbon aluminum deoxidized steels such as Al-killed steel grades, stems from complex interactions of metallurgical, physical, and chemical parameters.^[32] SEN clogging results in “bad” castability, hampering the overall quality of the casting process. Conventional sensors and prevention measures implemented in the factory as flow meter and temperature sensors, although somewhat useful, often fall short to forecast clogging events in advance due to its intricate and random appearance. Consequently, numerous studies in the past attempted to predict the castability from metallurgical and thermodynamic aspects for Al-killed steels, acknowledging the complexity and crucial necessity of resolving this challenge.^[19,32,33] It is crucial to note that “bad” castability due to stochastic appearance of SEN clogging^[34,35] can cause loss of productivity and unwanted wastage of precious energy resources. An extensive body of literature has explored the primary causes of SEN clogging. One prominent outcome from these studies indicates that the formation of alumina clusters, a result of the steel melt’s deoxidation and subsequent reoxidation of aluminum within the melt, significantly contributes to SEN clogging.^[5–7] This research was focused on understanding the root cause of the undesirable castability condition, studying the role of clusters of solid oxide alumina generated from deoxidation and reoxidation processes within the melt.^[13,14] Simultaneously, other investigations, while instrumental in understanding castability were focused on developing transient mathematical models for clogging detections.^[21] In recent times, attempts were made to detect clogging using novel methods by exploring use of graphical models and identification of anomalies in casting parameters.^[36] The researchers also used LSTM neural networks for classifying the occurrence of clogging from historical data.^[30,37] Previous research studies have also ventured into the development of predictive models, largely centered on classical statistical or machine learning (ML) techniques. However, these models struggled to capture the nonlinear relationships and temporal dependencies intrinsic to the continuous casting process, limiting their predictive accuracy and generalizability with real industrial data. With the recent advancements in AI, more sophisticated tools for understanding, predicting, and circumventing clogging events in steel production become feasible. Hybrid models, combining techniques such as ANFIS and LSTM networks, show potential in handling the inherent complexity and dynamism of steelmaking. Despite this potential, applications of such hybrid models in predicting clogging events, particularly in the continuous casting of SAE1055 grade steel, remain largely unexplored, indicating a clear gap in the existing literature. This steel grade was manufactured using the EAF-LF-VD-CC processing route. The process

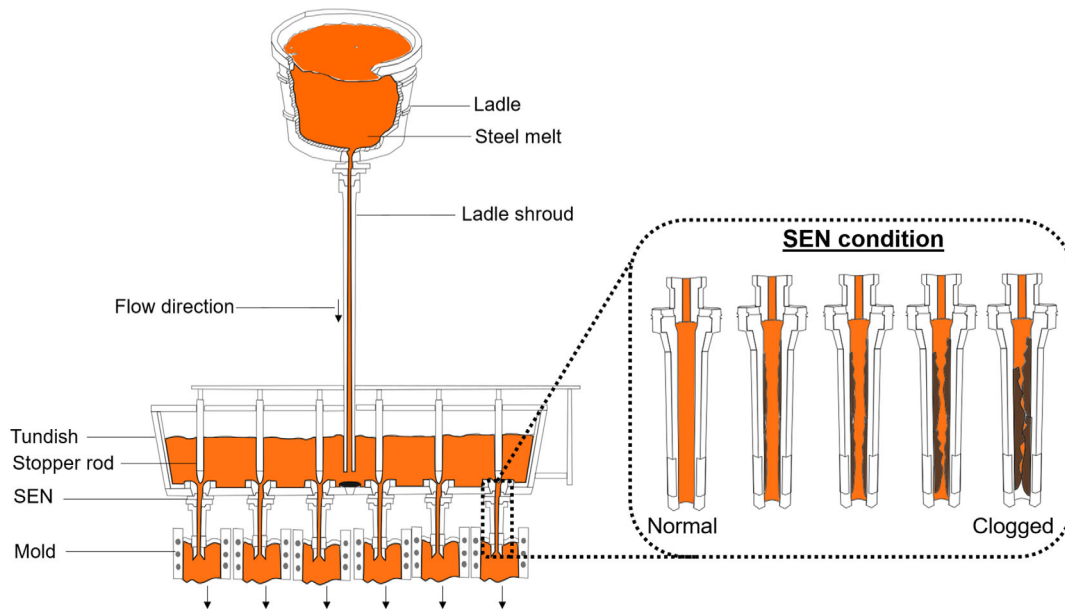


Figure 1. The melt is transferred from the ladle to a six-strand tundish, and then to a copper mold. The inset shows conditions inside the submerged entry nozzle (SEN).

began with the production of the initial steel melt in an EAF, followed by refinement stages at the LF and VD during the steel refining process. The manufacturing process then concluded with casting of refined steel through a six strand CC machine. As a standard practice, the steel melt is transferred from the ladle to the tundish, and then poured into a copper mold to produce steel products with the designated shapes, as shown in **Figure 1**. In the context of the CC machine, the casting of one batch referred to as a “cast”, and one “cast” typically lasts between 45 and 70 min. The casting duration of studies steel grade was contingent on the casting conditions to the specific cast.

3. Methods

3.1. Data Retrieval and Preprocessing Steps

The necessary datasets comprising the production data of SAE1055 steel grade were obtained from the data repositories and collected in comma separated files. The target compositions of the steel grade are given in **Table 1**. To protect the company’s interest, only the main elements were mentioned. The dataset comprising production data of ≈ 150 casts was used for the formulation of the ANFIS-LSTM hybrid model. The production data from casting machine incorporated with casting speed, stopper rod positions, and temperature was also extracted for each casts. The associated values were organized chronologically in a time

Table 1. The target steel chemistry of SAE1055 steel grade used for model formulation in wt%.

	C	Si	Mn	Cr	Al	S	P
Minimum	0.52	0.07	0.6	0	0.02	0	0
Maximum	0.6	0.6	0.9	0.2	0.04	0.008	0.020

series manner, with each data point collected every second. Spot measurements of the steel chemistry from refining ladles and tundish were also obtained. The detailed and systematic data collection provided comprehensive information on the casting of SAE1055 steel grade forming the foundation for the hybrid model.

3.2. Castability Measurement

To effectively monitor the conditions of SEN, especially for the detection of “bad” castability caused by clogging conditions, it was crucial to quantify the steel melt’s ability to flow through the SENs. This measure was encapsulated by a parameter known as the “Castability Index” (CI). The CI served as a metric that measured the ease with which liquid steel melt travelled through the six SEN nozzles in the continuous casting machine. The “CI” value provided a crucial quantifiable insight into the casting process and complements the monitoring of CC parameters particularly in detecting potential clogging incidents. A standard procedure was adopted by measuring the rate of change in the average slope of the stopper rod movement measured against time with realization of existence of peaks within the same graphical depiction. The CI value was then subsequently calculated by averaging the movement of the stopper rods of all six nozzles, which modulated the flux of the liquid steel from the tundish to the mold. The Equation (1) depicts the formula to calculate the CI value with ‘ n ’ strands.

$$\text{Castability Index (CI)} = \frac{1}{n} \sum_{i=1}^n \left(\frac{S_i(t_{\text{end}-5}) - S_i(t_{\text{start}+5})}{(t_{\text{end}-5}) - (t_{\text{start}+5})} \right) \quad (1)$$

where;

n : represents the total number of strands or nozzles involved in the casting process. In this case, n is equal to 6, corresponding to a six-strand continuous casting machine.

S_i : represents the stopper rod position in the i^{th} nozzle $i = [1, 2, 3, 4, 5, 6]$.

t_{start} : refers to the time at which the casting process starts for each cast.

t_{end} : refers to the time at which the casting process ends for each cast.

This method ensured a holistic perspective that took into account the behavior of all nozzles, thereby providing a comprehensive assessment of casting process's efficiency. Extreme variations in the stopper rods' movements could suggest the descent of inclusions into the liquid steel within the mold. Conversely, a persistent upward trajectory might indicate the agglomeration of inclusions within the nozzle. The final computation of the index was the mean of all observations in six nozzles. The complex behavior was then condensed into a single numerical value on a normalized scale from 0 to 10, with a higher value indicating better castability. A low CI value suggests a high likelihood of inclusions in the liquid steel clumping together, potentially leading to SEN clogging. Even though CI indirectly measures castability, it serves as a reliable proxy for operators. Greater quantities of impurities typically equated to more severe castability issues. Therefore, CI effectively represented the castability of the SAE1055 steel grade, giving an indication of its ability to be cast without complications.

In the formulation of the CI, a crucial time modification was introduced in the analysis of the casting process. Specifically, temporal adjustments were applied at the beginning and end of the process, as these periods often encompass transient phenomena that do not accurately mirror the representative behavior of the effective casting duration. This was addressed by implementing offsets at these critical junctures. At the outset of the casting process, an offset of five minutes was added to the initial start time, referred to as $t_{\text{start}+5}$. This adjustment was undertaken to circumvent the initial instabilities and oscillations inherent in the onset of the casting process. It aided in eliminating the influence of these fluctuations that may distort the calculated CI and not accurately reflect the standard behaviour of the process. Similarly, toward the conclusion of the casting process, an offset is introduced by subtracting five minutes from the end time, denoted as $t_{\text{end}-5}$. This was primarily to avoid the impact of variations that occur toward the tail end of the process, which might not be representative of the standard operational behaviour. It was often observed that the concluding stages of casting might be subject to irregularities that could potentially skew the CI values. Therefore, the time points $t_{\text{start}+5}$ and $t_{\text{end}-5}$ effectively represented the start and end times of the casting process for the measurements of the CI. These adjustments enabled the avoidance of the non-representative behaviour that typically occurs at the commencement and termination of the casting process. Consequently, these adjusted time points served as the reference for calculating the change in stopper rod position and the total duration of the casting process, both essential parameters in the computation of the castability. By doing so, the CI measurements became indicative of the standard behaviour of the casting process, providing a more accurate and reliable assessment of castability. The advantage of employing the CI in the proposed hybrid model was its universal applicability to scrutinize each "cast" produced, offering a consistent measure for assessing process performance and identifying potential challenges. In this study, for each manufactured cast, both castability and nozzle

conditions were meticulously inspected and quantified through CI. Therefore, the resulting CI values for each "cast" were merged with the corresponding process data. Here, CI value greater than nine was consider good for an excellent fluidity without clogging, whereas CI value between six and nine consider to have sufficient fluidity without clogging. CI value less than six consider to have bad fluidity with serious clogging problems. These datasets integrated with steel chemistry and stopper rod positions served as the foundation inputs for training the hybrid model equipped decision support system.

3.3. ANFIS Model Development

ANFIS model was developed using five inputs and one output architecture. Specifically, the model embodies two input membership functions (MFs) for each input variable and one output membership function for output variable. Each input variable was associated with two membership functions 'inputmf1' and 'inputmf2', which were interpreted as linguistic labels low and high making model able to capture complex relationships. The structure of five layers model is highlighted in **Figure 2**. In the structure, a fixed node was represented as a circle while an adaptive node was indicated by a square. As a simple structure, two inputs with x and y , and one output with f were considered. Two fuzzy if-then rules based on the first-order Sugeno fuzzy model was expressed as follows in Equation (2) and (3)

$$\text{Rule 1: if } x \text{ is } A_1 \text{ and } y \text{ is } B_1, \text{ then } Z_1 = p_1x + q_1y + r_1 \quad (2)$$

$$\text{Rule 2: if } x \text{ is } A_2 \text{ and } y \text{ is } B_2, \text{ then } Z_2 = p_2x + q_2y + r_2 \quad (3)$$

where " A_i " and " B_i " expressed as the fuzzy clusters in the data, while " p_i ", " q_i ", and " r_i " defined the design parameters in the training process. Layer 1 defined an input variable for each of the appropriate fuzzy sets.

The functions by which membership degrees created using MFs were indicated by the nodes in this layer. The node function of a node " i " was defined as follows in Equation (4)–(5)

$$\mu_i^1 = \mu_{A_i(x)} \quad i = 1, 2 \quad (4)$$

$$\mu_i^1 = \mu_{B_{i-2}(y)} \quad i = 3, 4 \quad (5)$$

where μ_{A_i} and μ_{B_i} are the MFs. $\mu_{A_i(x)}$ is bell-shaped being 1 or 0, like the generalized bell function or the Gaussian function, stated in Equation (6,7)

$$\mu_{A_i(x)} = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i} \right)^2 \right]^{b_i}} \quad (6)$$

$$\mu_{A_i(x)} = \exp \left[- \left(\frac{x - c_i}{a_i} \right)^2 \right] \quad (7)$$

where $\{a_i, b_i, c_i\}$ is the parameter set. The parameters in this layer were introductory. In the nodes of this layer, incoming signals were duplicated and the product out was sent. Each node in this layer 2 calculated the firing strength of a rule by multiplying MFs, stated in Equation (8).

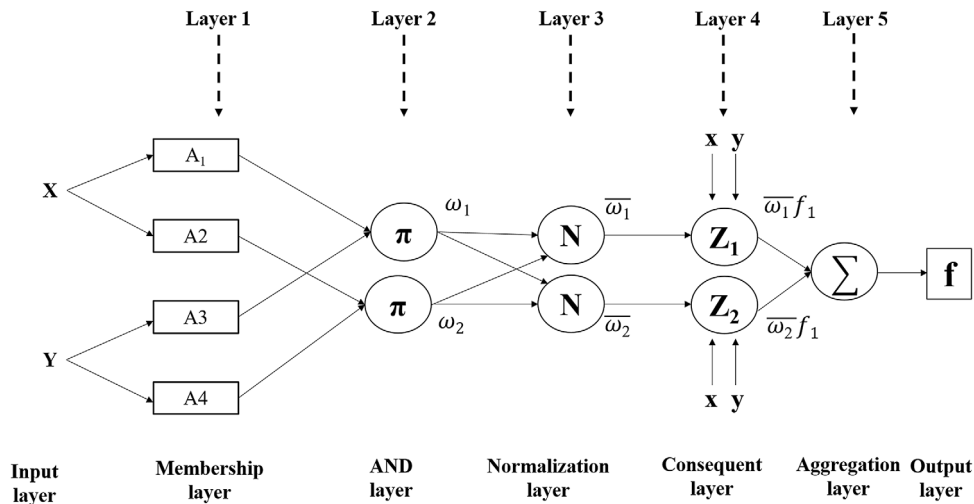


Figure 2. ANFIS architecture used in this work.

$$Q_i^2 = \omega_i = \mu_{A_i(x)}\mu_{B_i(y)}, i = 1, 2 \quad (8)$$

The i^{th} node in the layer 3 computed the ratio of the i^{th} rule's firing strength to the sum of all rules firing strengths of all rules, as stated in Equation (9).

$$Q_i^3 = \bar{\omega}_i = \frac{\omega_i}{\omega_1 + \omega_2}, i = 1, 2 \quad (9)$$

where $\bar{\omega}_i$ is the normalized firing strengths. In the layer 4, node “ i ” calculates the effect of the i^{th} rule on the model output with the node function shown in Equation (10).

$$Q_i^4 = \bar{\omega}_i f_i = \bar{\omega}_i(p_i x + q_i y + r_i), i = 1, 2 \quad (10)$$

where $\bar{\omega}_i$ and $\{p_i x + q_i y + r_i\}$ denoted the output of layer 3 and parameter set respectively. The parameters in the layer 4 were consequent. The layer 5 was addressed as the output nodes where the single node calculated the overall output as the total of all incoming signals expressed in Equation (11).

$$Q_i^5 = \sum_{i=1}^2 \bar{\omega}_i z_i = \frac{w_1 z_1 + w_2 z_2}{w_1 + w_2} \quad (11)$$

The output f was written as in Equation (12)–(14);

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \quad (12)$$

$$f = \bar{\omega}_1 f_1 + \bar{\omega}_2 f_2 \quad (13)$$

$$f = (\bar{\omega}_1 x) p_1 + (\bar{\omega}_1 y) q_1 + (\bar{\omega}_1) r_1 + (\bar{\omega}_2 x) p_2 + (\bar{\omega}_2 y) q_2 + (\bar{\omega}_2) r_2 \quad (14)$$

The single output membership function ‘outputmf’ was designed for generating the output of the system based on the weighted average of the input rules. MATLAB 2022a (The Mathworks, Inc., Natick, Massachusetts, United States) was used in conjunction with a custom-built programming setup in Python version 3.9.1 for formulating the model architecture highlighted in Figure 2.

The steel chemistry comprising elements Al, Si, Ca, Mn, and S were chosen due to their integral roles in steelmaking, specifically their propensity to form solid non-metallic inclusions like Al_2O_3 , SiO_2 , CaO , CaS , and MnS . Given the distribution of steel chemistry, a “Gaussian” membership function was chosen in the FIS of ANFIS. This function effectively captured the characteristics of the input data, leading to improved model performance. A Gaussian membership function, with its bell-shaped curve, was assigned a membership value between 0 and 1 for each input value, indicating the degree to which the value belongs to a specific fuzzy set. The height of the curve signifies the degree of membership, with the highest point (value 1) representing full membership. Incorporating these key steel chemistry parameters and leveraging the Gaussian membership function, the ANFIS model was trained to predict the CI values. The model's training and testing were done by a K-fold cross-validation methodology, where “ K ” was set to five, effectively partitioning the data into five distinct subsets. This method was adopted to prevent overfitting and provide a more comprehensive evaluation of the model's predictive performance. For each iteration of the k-fold cross-validation, the ANFIS model was trained on 80% of the data, with the remaining 20% reserved for validation. This partitioning allowed the model to learn from a substantial dataset while reserving a sufficient portion for validation. The learning algorithm used in ANFIS was combined with the least squares method and back propagation for optimization. This precise parameter tuning was critical in ensuring that the model learned effectively from the fluctuating data and optimized its performance, refer **Figure 3**. It shows the systematic process, including data segregation, model tuning, and subsequent validation for achieving high predictive accuracy. During training and tuning of the hyper parameters, the ANFIS model's predictive performance was validated using the root mean square error (RMSE) metric, which quantifies the differences between the predicted values and the observed values. The performance of the optimized ANFIS model was evaluated on the evaluation dataset comprising 30 “casts” to gauge its ability to estimate the castability index accurately with a given elemental composition. This carefully optimized ANFIS forms the first constituent of the proposed hybrid model.

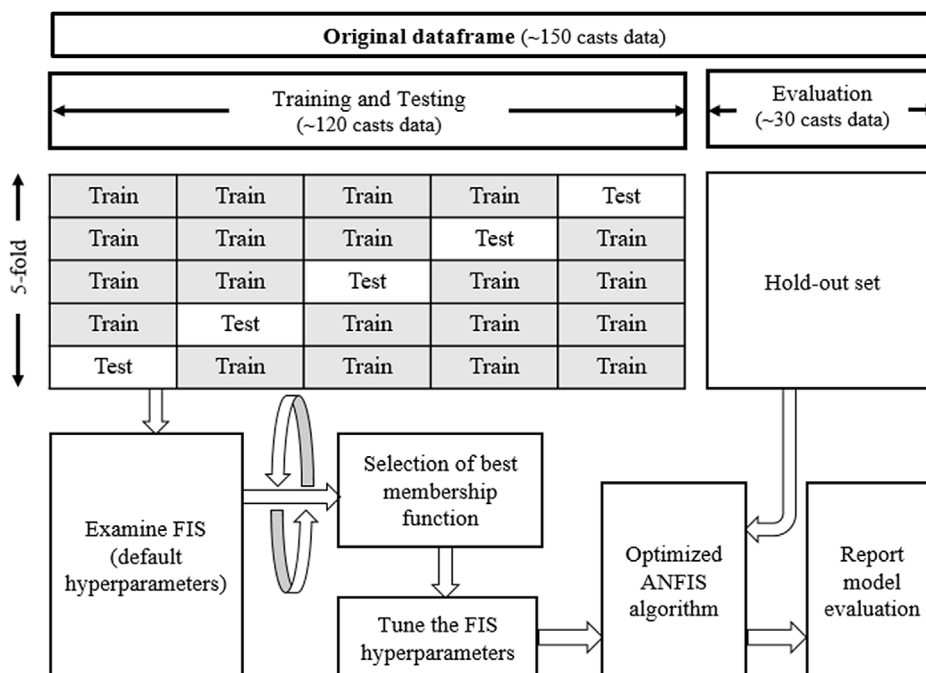


Figure 3. Schematic representation of process data division and the stepwise methodology for selection of the optimal membership function for the ANFIS model.

3.4. LSTM Network Model Development

The LSTM network model was developed using MATLAB 2022a (The Mathworks, Inc., Natick, Massachusetts, United States) and a custom-built Python programming setup. This strategy facilitated comprehensive model development and ensured robust and precise forecasting. The requisite Python libraries for the process, namely Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn, and Shap, were used extensively for data management, model development, visualization, and model interpretation. All computations were performed on a 64-bit Windows operating system equipped with 16 GB RAM and an Intel (R) Core (TM) i7-8665U CPU, which provided ample computing power for the development and execution of the models. Before the actual model development of the model, CI value was processed with sequencing function.^[30] The sequencing function was the number of historical CI values needed to predict a value in the future in a time-series manner refer Table 2. While modeling, the stopper rod positions of each SENs were taken every two second from each individual cast; thus at a specific timestamp t , one lag corresponds to timestamp “ $t - 1$ ”, which was equivalent to two second

Table 2. Time-Series features processed with sequence function.

Input features [X]	Timestamp	Step	Target variable [y]
$t-1200, t-1199, t-1198, \dots, t-2, t-1$	t	0	$t+720$
$t-1199, t-1198, t-1197, \dots, t-1, t$	$t+1$	1	$t+721$
.....
.....
$t+n-1200, t+n-1199, \dots, t+n-1$	$t+n$	n	$t+n+720$

in the past. In contrast, recording in future timestamp “ $t + 1$ ” equals two second ahead of the current timestamp. The LSTM network model used 1200 lags to predict the 720 timestamp ahead in the future, equivalent to predicting the castability index in 1440 s (≈ 24 min) based on the past data from 40 min. The input was a 3D array with dimensions containing samples, time steps, features. Here, samples represent the number of data points, time steps represent the number of time steps in each input sequence, and features represent the number of variables or series used for forecasting. The input vectors were prepared by applying this sequence function. Hence, input feature required to contain CI values from “ $t - 1200$ ” to “ $t - 1$ ”, and the output variable containing CI values at “ $t + 720$ ” for a timestamp at t . The same is described in Table 2. In subsequent timestamp, all the input and output variables were adjusted by adding step ‘ n ’ timestamp. The original “casts” data were transformed into multiple smaller training sequences of fixed time steps. The sliding window approach was used to predict the next value in the series. Here, a multiple-multiple LSTM forecasting was preferred, as multiple past steps were used to predict multiple future steps.

Figure 4 shows the LSTM layer architecture used in this study, indicating the flow of an X time series with S -length C properties (channels) across an LSTM layer. h_t was the output (also known as the hidden state) and c_t was the cell state at time step “ t ”. In order to calculate the first output and the updated cell state, the first LSTM part was utilized for the first state of the network and the early time step of the series. At time step “ t ”, this block uses the current state of the network (c_{t-1}, h_{t-1}) and the next time step of the sequence to compute the output and the updated cell state c_t .

The Figure 5 highlighted the flow of process data at time step “ t ”. The state of the layer comprised the hidden state and the cell state. The hidden state at time step “ t ” included the output of the

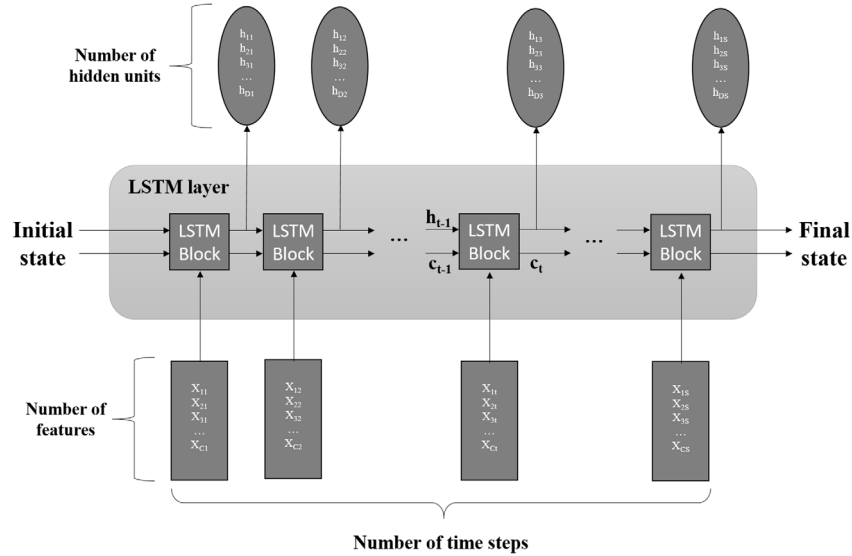


Figure 4. LSTM layer or cell adopted in the proposed study.

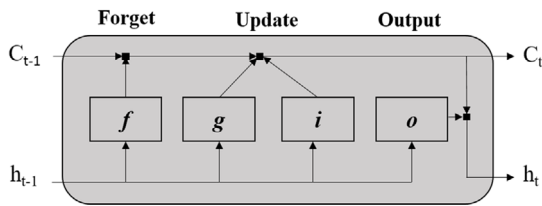


Figure 5. The flow of data at time step “t”.

LSTM layer for this time step. The cell state covered knowledge elicited from previous time steps. At each time step, the layer performs the task of adding or removing information from the cell state. These updates were controlled by means of logic gates. Some components in the LSTM layer architecture were used to control the cell state and the hidden state of the layer. For example, input gate (*i*) and output gate (*o*) control the level of cell state update and level of cell state added to the hidden state, respectively. Forget gate (*f*) checked the level of cell state reset (forget). In contrast, cell candidate (*g*) added the information to the cell state. The input weights *W*, the recurrent weights *R*, and the bias *b* were called as the learnable weights of an LSTM layer. The matrices *W*, *R*, and *b* are concatenations of the input weights, recurrent weights and bias of each component, respectively. These matrices were concatenated as follows in Equation (15);

$$W = \begin{bmatrix} W_t \\ W_f \\ W_g \\ W_o \end{bmatrix}, R = \begin{bmatrix} R_t \\ R_f \\ R_g \\ R_o \end{bmatrix}, b = \begin{bmatrix} b_t \\ b_f \\ b_g \\ b_o \end{bmatrix} \quad (15)$$

Where “*i*” is the input gate, “*f*” is the forget gate, “*g*” is the cell candidate and “*o*” is the output gate. The cell state at time step “*t*” is indicated as in Equation (16);

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (16)$$

where \odot is the Hadamard product. The hidden state at time step “*t*” was given by Equation (17);

$$h_t = o_t \odot \sigma_c c_t \quad (17)$$

Where σ_c indicates the state activation function. Hyperbolic tangent function is used to compute the state activation function for the LSTM layer function by default. The components at time step “*t*” are described by the following formulas stated in Equation (18)–(21).

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \quad (18)$$

$$f_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \quad (19)$$

$$g_t = \sigma_c(W_g x_t + R_g h_{t-1} + b_g) \quad (20)$$

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \quad (21)$$

where σ_g is the gate activation function. The LSTM layer in this study, used the sigmoid function to compute the gate activation function, and it was given by Equation (22).

$$\sigma(x) = (1 + e^{-x})^{-1} \quad (22)$$

The number of LSTM layers and the number of neurons in each layer were chosen based on the input data analysis. After testing different architectures, it was found that a model with two LSTM layers provided the best results. In a similar vein to the development of ANFIS model, the LSTM model was trained and validated using the K-fold cross-validation technique, with *k* set to five. This approach provided a robust mechanism to prevent model overfitting and helped ensure that the trained model would generalize well to unseen data. The training data from 120 casts and validation data from 30 casts was used respectively for training-testing and evaluation of the LSTM network. The model evaluation performance was examined using the RMSE metric. The LSTM network model was carefully

optimized, with a particular focus on optimizing the number of hidden layers and units, the learning rate, the dropout rate, and the batch size. Adam optimizer, known for its effectiveness in training deep learning models was used for the optimization. In this case, the optimal hyper parameters were selected based on the combination that achieved the lowest RMSE on the training-testing set. In this work double LSTM-layer networks were created with the same number of neurons. The optimized LSTM network model then formed the second stage of the hybrid ANFIS-LSTM model.

4. Results

The results illustrate the effectiveness of hybrid model in estimating and forecasting the CI values. It was found during the preprocessing that the elemental chemistry in steel melt observed in refining ladles and tundish follow Gaussian distribution. Some interesting insights about manufacturing of SAE1055 process were found in the data analysis. The mean casting speed was found to be ≈ 57 min with deviation of ≈ 7 min suggesting the reasonable consistency in casting time for the analyzed 150 casts. The data analysis showed that the mean casting process was operated at an average temperature of 1527.78 ± 25.24 °C. Concentration of Al in the analyzed datasets averages at 0.019% by weight. Al concentration exhibits a relatively small standard deviation of 0.004%, signifying stable consistency in its distribution throughout the steel transfer from ladle refining to tundish. Even though the concentration ranges from 0.012% to 0.034%, its mean value and minimal variance highlight its steady contribution. The mean sulfur concentration was found to be 0.016% by weight. This variability, coupled with the wide range of sulfur concentration and calcium concentration of 5.40 ppm might potentially influenced the castability of the steel grade. These observed results were highlighted in **Table 3**.

The outcomes of ANFIS optimization processes, which were conducted using backpropagation with least square estimation is described in **Table 4**. The best cross validation RMSE was obtained for the ANFIS model with 32 sets of fuzzy rules and 50 training epochs. The estimated CI values for a given steel chemistry were obtained from the ANFIS model. In Figure 3a–d, surface plots were shown describing the relationship between the estimated value of CI from ANFIS with the input steel chemistry comprising features Al, Ca, Si and S. The surface plots clearly present the direct relationship between steel chemistry and estimated CI value.

Table 3. Descriptive statistics of casting temperature, castability index, casting time, and elemental composition (measured after VD in ladle and tundish) for the analyzed datasets of 150 heats.

Measure	Casting temperature [°C]	Castability index (Dimensionless)	Casting time [min]	Steel chemistry from ladle and tundish [wt%]				
				Al	S	Mn	Ca	Si
Mean	1527.78	8.75	57.13	0.019	0.016	0.757	0.0005	0.29
Std Dev	25.24	1.14	7.09	0.004	0.012	0.040	0.0001	0.01
Min	0	4.37	47.65	0.012	0.001	0.67	0.0001	0.25
Max	1650	10	78.61	0.034	0.071	0.85	0.0009	0.35

Table 4. ANFIS model optimization parameters and the model evaluation results.

Number of membership	Number of fuzzy rules	Number of training epochs	Cross-validation RMSE
2	32	50	0.871
2	32	50	0.048
3	243	50	0.715
3	243	50	0.602
2	32	100	0.544
2	32	100	0.423
3	243	100	0.645
3	243	100	0.437

Table 5. The best obtained hyper parameters for LSTM network model.

Hyperparameters	Starting value	Increment	End value	Selected parameter
Learning rate	0.005	0.005	0.5	0.01
Number of neurons	64	64	128	128
Dropout rate	0.1	0.1	0.5	0.2
Batch size	256	256	512	512

It was found while optimizing the LSTM model that a small learning rate could consume extensive computational time for training, whereas a large rate caused faster convergence outside of the global minimum value. The best obtained hyper parameters for LSTM network model is highlighted in **Table 5**.

The output CI values from hybrid model were encapsulated into actual values, ANFIS-estimated, and LSTM-forecasted values, providing a succinct yet comprehensive view of model's response. The model predictive response is highlighted in Figure 7. The output values of CI demarcates normal SEN conditions in green, warning conditions in yellow, and possible SEN clogging conditions in red, based on the dynamics of the casting process across the timescale of 'cast $n - 1$ ', 'cast n ', and 'cast $n + 1$ '. The grey area inside the Figure 7a shows that when the output CI value reach below specific value there could be possibility of ladle change or production halts due to SEN clogging. Figure 7b showed the results of estimated and forecasted CI values over a specific casting time of 100 min.

5. Discussions

The main results illustrate the effectiveness of this hybrid model in forecasting the castability of steel, a notoriously challenging task due to the complexity and dynamic nature of the steelmaking process. The developed ANFIS model leveraged historical steel composition data to estimate CI value from the input features. The FIS inside ANFIS was optimized with best hyper parameters mentioned in Table 4, and it showed quite good prediction when tested on the unseen evaluation dataset. The results showed the impact of manipulating the number of membership functions, fuzzy rules, and training epochs on the model's performance. It was noteworthy that an increase in training epochs typically corresponded to a reduction in RMSE, thus underscoring enhanced model performance. However, the interplay between the number of membership functions and fuzzy rules proved to be more nuanced, suggesting that optimizing these parameters requires careful consideration of the specific dataset

and research problem. In general, estimating the parameter associated with the continuous casting process solely based on steel chemistry is not often desirable from online production perspective. According to the results of the present investigations, castability of SAE1055 steel grade was majorly influenced by fluctuations in the Al, Ca and S content. Hence, to achieve the good castability a tight control of Al in the range of [0.02–0.03%], Ca [2–4 ppm] and S [0.03–0.06%] content in the tundish was important. The surfaces generated by plotting CI values against the steel chemistry shown in the **Figure 6** was helpful in understanding ideal range of steel chemistry associated with the 'good' castability and 'bad' castability. The surface plots described the interactions between two elements and their combined influence on the CI was inferred from the contour and texture of the surfaces. However, as each pair of elements and their interaction could uniquely affect CI, each plot should be carefully studied for its specific characteristics and impacts on CI. The color scale accompanying the surface plots provided

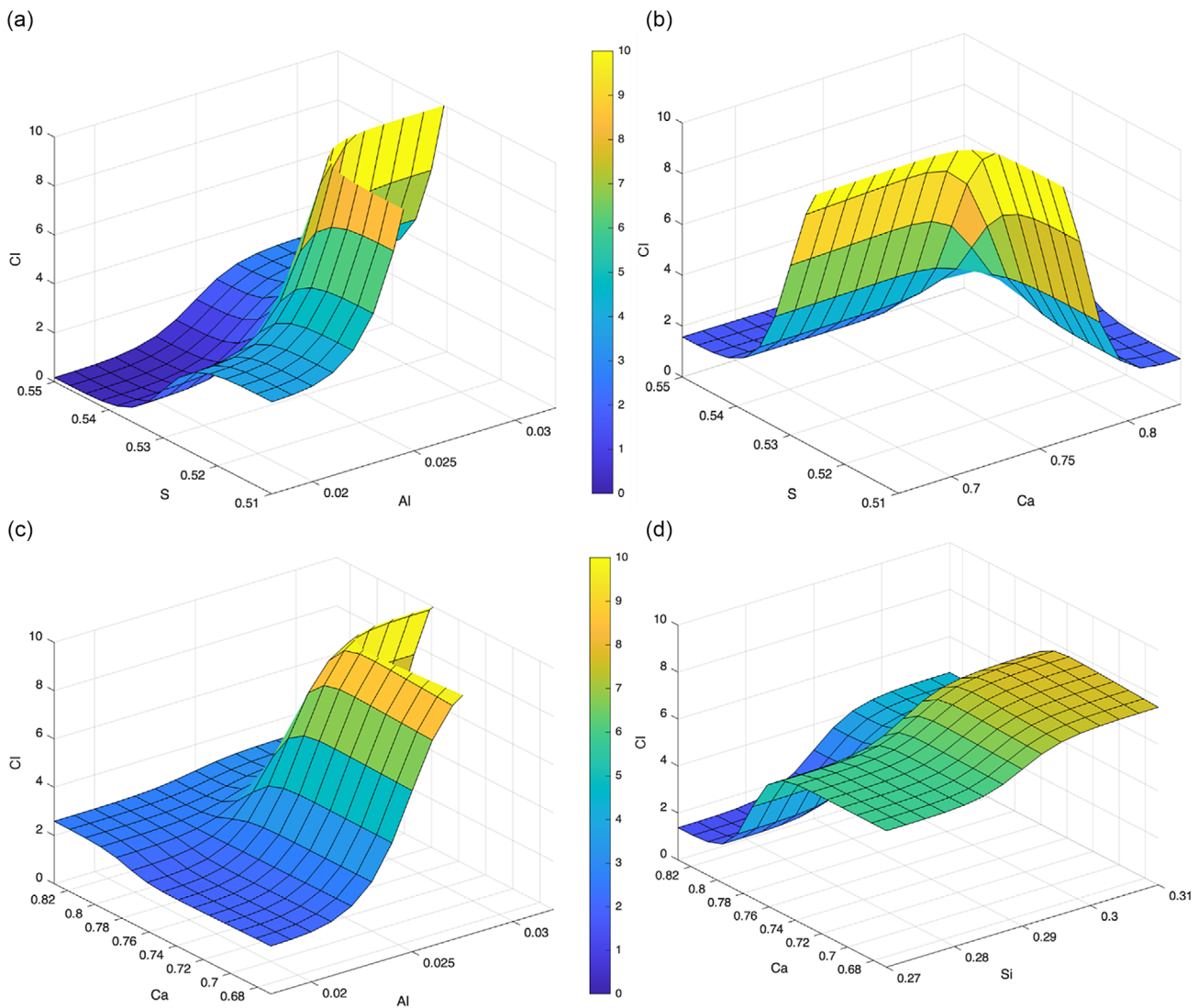


Figure 6. Surface plots highlighting the relationship between the predicted value of castability index from ANFIS with the input features Al, Ca, Si and S. a) CI-S-Al, b) CI-S-Ca, c) CI-Ca-Al, and d) CI-Ca-Si.

a more nuanced understanding of the relationships between the variables. This scale, spanning from 0 (depicted in blue) to 10 (depicted in yellow), signified the CI value. In these plots, blue areas pinpointed regions where the CI was low, suggesting that under these specific combinations of the two variables, the conditions were less than optimal. Conversely, yellow areas indicated high CI values, signifying preferable normal SEN conditions. This color gradation shown in Figure 6 not only clarified the behavior of the CI across steel chemistry ranges but also acted as a visual guide to identify areas of maximum or minimum CI values. When integrated with the contours and gradients of the surface plots, it offered a comprehensive visualization tool for interpreting the multi-dimensional relationships between the different variables and the CI. It was important to understand the effect of S values on the CI with different combination of Al and Ca. Figure 6a,b highlighted that CI value was quite dynamic when steel composition changes with minute change in sulfur. Further, surface plots highlighting the relation of CI-Ca-Al in Figure 6c and CI-Ca-Si, in Figure 6d showed that CI changes significantly due to slight changes in Al and Si concentration indicating requirement of tighter chemistry control. It also signified that from a metallurgical perspective, when steel flows from one refractory vessel to another, it could cause alterations in sulfur and calcium composition subsequently influencing the changes in CI value. In addition, during transfer from ladle shroud some amount of slag could entrap in the tundish. This entrapment of slag can change the steel chemistry. Further, ladle change can also cause the change in steel chemistry. Therefore, it was also justified to incorporate the model for estimating CI values, which take into account the elemental steel compositions.

ANFIS's ability to handle the complexity and ambiguity of the steel chemistry data, combined with LSTM's capacity for time-series forecasting, results in a system with enhanced predictive capabilities. The number of neurons is a fundamental concept in building an LSTM network, and there is no rule of thumb in the optimal number to be selected. LSTM networks were optimized with dropout layers that helped reduce overfitting of the data.

Finally, batch size defines the number of training examples for the LSTM to feedforward and backpropagate before updating internal parameters. Table 5 described the best optimized learning rate, number of neurons, dropout rate and batch size. It was also found that a large batch size required expensive computational resources such as the memory space, whereas a small batch size can adversely affect the training time. The plot in Figure 7a dynamically shows the actual and estimated CI values in relation to the casting process's expected status. The color-coding allows operators to understand the SEN's current status at a glance and help taking necessary actions. The vertical time axis represents the progression of the casting process, specifically focusing on the 'cast n-1', 'cast n', and 'cast n + 1'. This allows operators to observe the trends and anticipate potential issues. The average casting time is approximately 56 min, but the plot displays up to 200 min to present a broader view and enable detailed inspection of each casting phase. On the contrary, Figure 7b describes the output of the LSTM network. The LSTM networks model detected sequential patterns in time-series data, incorporating a sliding window of approximately 24 min for its next forecasting. These forecasts notify about real-time decisions to prevent production stoppages and maintain optimal operations. The forecasting system proposed herein, delivers a novel, comprehensive approach for real-time decision-making. By monitoring SEN conditions by interpreting the CI, the decision support system (DSS) can forecast potential problems in the casting process, thus allowing operators to make informed decisions to avoid costly downtime and maintain the efficiency of the steel production process. The casting process's average duration under standard conditions was about ≈ 55 min, but the graph has been extended to 200 min to account for potential variations or extensions in casting time due to anomalies or equipment malfunctions. This time extension offers an even closer look into the future, aiding in the identification of potential disruptions well ahead of time.

The proposed model was industrially validated by comparing actual and predicted castability for additional heats of SAE1055

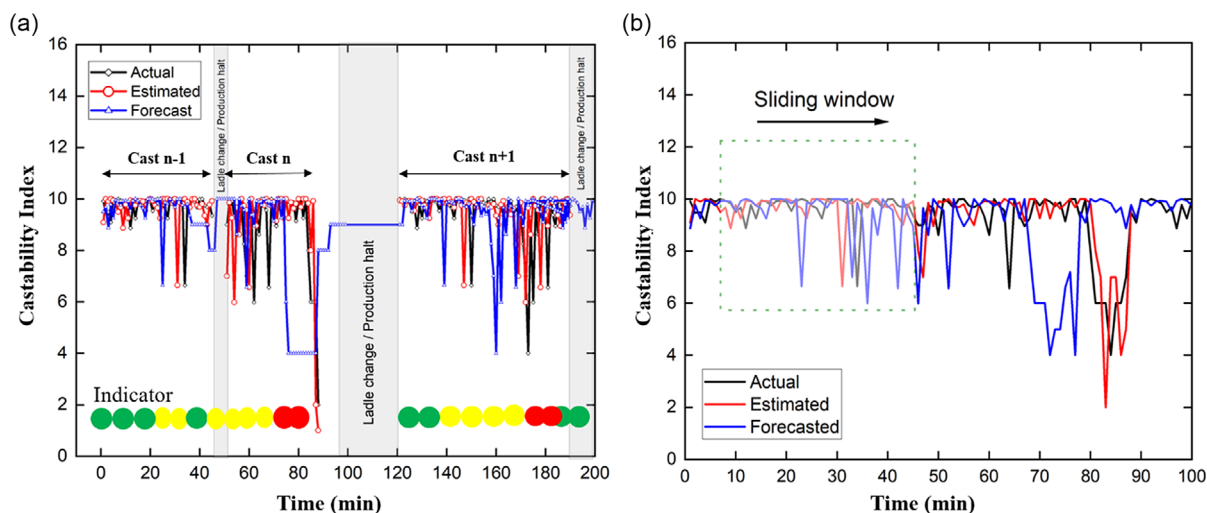


Figure 7. a) Model predictive response with estimated and forecasted CI values. b) Forecasting outcomes generated from LSTM network model with ≈ 24 min sliding window.

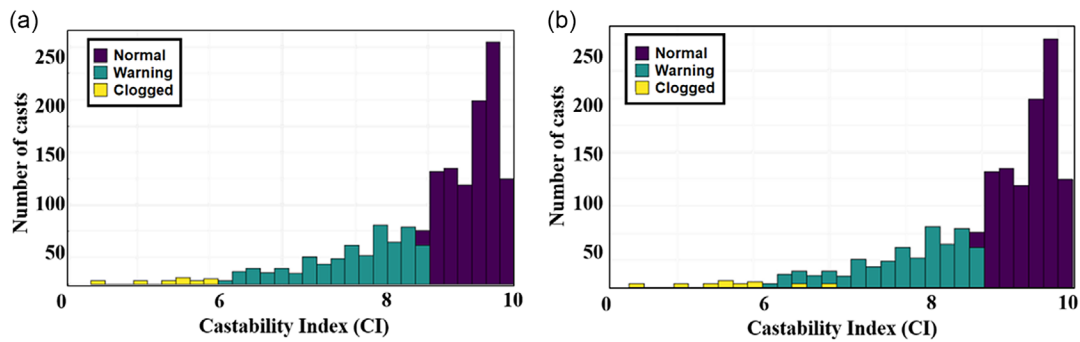


Figure 8. a) Distribution of actual castability index (CI) in unseen dataset for validation and classification in three considered categories, b) distribution of CI in test batch of data, and categories predicted by decision support system.

steel grade. The frontend interface named *Checkcast* has been implemented in steel plant for this purpose within the Horizon 2020 project *INEVITABLE*. **Figure 8a** shows the distribution of the real castability index corresponding to this tests dataset. The output from *Checkcast*, “the forecasted castability”, was not a number but a categorical variable. Thereby, for comparison of real castability against forecasted values, the numerical CI of each heat was translated to its corresponding category. This classification was shown by color scale in **Figure 8b**. Upon examining this validation results, it was found that the classification accuracy for both “Normal” and “Clogged” categories was exceptionally high. However, a marginal classification error was encountered where 1.5% of the “Warning” cases were erroneously classified as “Clogged” casts.

6. Conclusions

The proposed study introduced a novel DSS based on hybrid model, leveraging the combined power of ANFIS and LSTM networks to forecast and mitigate clogging events during the continuous casting of SAE1055 steel grade. The utilization of these models enabled real-time adaption to changes in steel chemistry and stopper rod changes, improving SEN monitoring and castability forecasting. Despite potential areas for refinement, the model demonstrated its effectiveness and robustness in forecasting castability, offering valuable insights into the complex steel-making process. The successful deployment of this hybrid model promises a paradigm shift in steel plants’ operations, emphasizing a data-centric approach in decision-making processes. Some of the main findings were described as: 1) The proposed system successfully forecasted castability index parameter prior to the continuous casting process, effectively addressing clogging events during the casting of SAE1055 steel grade; 2) The hybrid approach effectively adapted to variations in steel chemistry and stopper rod changes, enhancing real-time monitoring of SEN conditions; and 3) Despite areas requiring refinement, the decision support system showed robustness and accuracy, illuminating the potential of deploying AI-assisted, data-driven predictive models in the steelmaking process.

The study demonstrated the potential for increased productivity, decreased production costs, and improved product quality using the AI-assisted predictive models. The implementation

of this system could spur a data-centric operational approach within steel plants, transforming traditional practices and increasing efficiency. Additional research is necessary to refine the model further, aiming for a more robust, efficient, and error-free steel production process. Authors believed that the promising results from this study mark an important step toward revolutionizing steelmaking operations.

The proposed model has demonstrated robust performance in inline investigations. However, it was not without its limitations. One significant limitation was the lack of involvement of caster events mentioned by previous work,^[30] potentially limiting its scope. The importance of incorporating these events in future research cannot be overstated as they often play a vital role in influencing castability. Another limitation involves the issue of data imbalance. The performance of trained models could potentially be affected by imbalanced datasets, which can create bias in the model toward the majority class. As such, future enhancements to the model should involve strategies for balancing the datasets to reduce this bias, such as resampling techniques or the use of alternative evaluation metrics. Despite its current limitations, the fusion approach has exhibited remarkable potential in forecasting SEN clogging events, facilitating proactive measures like timely replacement of the potential SEN nozzles. This advantage has been gained without incorporating the complexity of other casting parameters like mold level control and various sensor data. The exclusion of these parameters was driven by the intention to create a simple yet effective model that could concentrate on the core factors influencing SEN clogging and “bad” castability. The parameters were narrowed down to those directly related to SEN clogging, keeping the model streamlined and less prone to overfitting, thus enhancing its generalizability. This simplified approach has the added advantage of making the model more interpretable, easing the task of identifying the crucial factors leading to SEN clogging. While this approach has proven beneficial, the model’s predictive power might be further enhanced with the inclusion of additional casting parameters in future iterations, as long as this does not overly complicate the model or lead to overfitting. Furthermore, exploring other deep learning algorithms could also be beneficial. Deep learning CNNs and RNNs could potentially offer different perspectives on the data and might lead to even more accurate predictions. However, the choice of algorithm should be carefully considered, taking into account factors like computational complexity and

interpretability. Finally, while the model has proven effective in offline investigations, it has yet to be implemented in real-time, which is a crucial aspect for any decision support system.^[38] The future work should therefore focus on developing a real-time version of the ANFIS-LSTM decision-support system that can make quick and accurate predictions, contributing to improved operational efficiency and effectiveness. With these enhancements, the model is expected to become an even more powerful tool for understanding and predicting castability in the steel industry.

Acknowledgements

The work presented in this article is funded by the European Union's Horizon 2020 research and innovation programme, the SPIRE initiative, under Grant agreement no. 869815, the INEVITABLE project ("Optimization and performance improving in metal industry by digital technologies")

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from [PARTY]. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the authors with the permission of [PARTY].

Keywords

artificial intelligence, castability, clogging, forecasting, submerged entry nozzle

Received: March 19, 2024

Revised: June 7, 2024

Published online:

- [1] V. Colla, C. Pietrosanti, E. Malfa, K. Peters, *Mater. Tech.* **2020**, *108*, 507.
- [2] F. Boto, M. Murua, T. Gutierrez, S. Casado, A. Carrillo, A. Arteaga, *Metals* **2022**, *12*, 172.
- [3] L. S. Carlsson, P. B. Samuelsson, P. G. Jönsson, *Metals* **2019**, *9*, 959.
- [4] S. Dworak, H. Rechberger, J. Fellner, *Resour. Conserv. Recycl.* **2022**, *179*, 106072.
- [5] R. Bruckhaus, R. Fandrich, *Trans. Indian Inst. Met.* **2013**, *66*, 561.
- [6] H. V. Atkinson, G. Shi, *Prog. Mater. Sci.* **2003**, *48*, 457.

- [7] G. Rosegger, *J. Ind. Econ.* **1979**, *28*, 39.
- [8] Y. Zhou, K. Xu, F. He, Z. Zhang, *ISIJ Int.* **2022**, *62*, 689.
- [9] A. Vakhrushev, A. Kharicha, Z. Liu, M. Wu, A. Ludwig, G. Nitzl, Y. Tang, G. Hackl, J. Watzinger, *Metall. Mater. Trans. B.* **2020**, *51*, 2811.
- [10] S. K. Michelic, C. Bernhard, *Steel Res. Int.* **2022**, *93*, 2200086.
- [11] Y. Sahai, T. Emi, *ISIJ Int.* **1996**, *36*, 1166.
- [12] C. Yao, M. Wang, M. Xu Pan, Y. Ping Bao, *J. Iron Steel Res. Int.* **2021**, *28*, 1114.
- [13] D. Mazumdar, *Metall. Mater. Trans. B.* **2021**, *52*, 23.
- [14] S. Hore, S. K. Das, M. M. Humane, A. K. Peethala, *Trans. Indian Inst. Met.* **2019**, *72*, 3015.
- [15] C. Bernhard, G. Xia, A. Karasangabo, M. Egger, A. Pissenberger, in *7th European Continuous Casting Conf.* **2011**.
- [16] F. Tehovnik, J. Burja, B. Arh, M. Knap, *Metalurgija* **2015**, *54*, 371.
- [17] J. Ikäheimonen, K. Leiviskä, J. Ruuska, J. Matkala, in *IFAC Proc.* **2002**, Vol. 35, 143.
- [18] Y. Sahai, *Metall. Mater. Trans. B.* **2016**, *47*, 2095.
- [19] B. G. Thomas, H. Bai, in *ISS Steelmaking Conf. Proc.* **2001**, Vol. 84, p. 895.
- [20] L. Bergman, *Master of Science*, Luleå Technological University **2006**, <https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1020451&dswid=3113>.
- [21] H. Barati, M. Wu, A. Kharicha, A. Ludwig, *Powder Technol.* **2018**, *329*, 181.
- [22] J. H. Park, H. Todoroki, *ISIJ Int.* **2010**, *50*, 1333.
- [23] K. Miao, M. Nabeel, N. Dogan, *Metall. Mater. Trans. B.* **2022**, *53*, 2897.
- [24] H. Barati, M. Wu, S. Michelic, *Metall. Mater. Trans. B.* **2021**, *52*, 4167.
- [25] C. Hua, Y. Bao, M. Wang, *Metall. Mater. Trans. B.* **2022**, *53*, 3757.
- [26] H. Yang, H. Olia, B. G. Thomas, *Metals* **2021**, *11*, 116.
- [27] H. Barati, M. Wu, S. Ilie, A. Kharicha, A. Ludwig, *J. Powder Technol.* **2024**, *434*, 119307.
- [28] D. Karaboga, E. Kaya, *Artif. Intell. Rev.* **2019**, *52*, 2263.
- [29] Y. Yu, X. Si, C. Hu, J. Zhang, *Neural Comput.* **2019**, *31*, 1235.
- [30] R. Wang, H. Li, F. Guerra, C. Cathcart, K. Chattopadhyay, *ISIJ Int.* **2022**, *62*, 2311.
- [31] M. Bazarbaev, T. Chuluunsaikhan, H. Oh, G. A. Ryu, A. Nasridinov, K. H. Yoo, *Sensors* **2022**, *22*, 29.
- [32] D. Janke, Z. Ma, P. Valentin, A. Heinen, *ISIJ Int.* **2000**, *40*, 31.
- [33] K. Tshilombo, *Int. J. Miner. Metall. Mater.* **2010**, *17*, 28.
- [34] K. G. Rackers, B. G. Thomas, in *78th Steelmaking Conf. Proc.* **1995**, Vol. 78, p. 723.
- [35] W. F. Caley, *High Temp. Mater. Process.* **2006**, *25*, 157.
- [36] S. Yang, A. Rebmann, M. Tang, R. Moravec, D. Behrmann, M. Baird, B. W. Bequette, *J. Process Control* **2021**, *105*, 259.
- [37] A. P. M. Diniz, P. M. Ciarelli, E. O. T. Salles, K. F. Coco, *Decis. Mak. Appl. Manag. Eng.* **2022**, *5*, 154.
- [38] S. Kuthe, R. Rössler, A. Karasev, B. Glaser, *Metall. Mater. Trans. B.* **2024**, *55*, 1395.