Predicting Quantitative Indices for SEN Clogging in Continuous Casting Using Long Short-term Memory Time-series Model

Rubin WANG,1) Heng LI,1) Fernando GUERRA,2) Chad CATHCART2) and Kinnor CHATTOPADHYAY1)*

1) Process Metallurgy Research Lab, Department of Materials Science and Engineering, University of Toronto, 184 College Street, Toronto, ON, M5S3E4 Canada.
2) STELCO Inc., 2330 Regional Road #3, Nanticoke, ON N0A 1L0, Canada.

(Received on April 26, 2022; accepted on June 28, 2022; J-STAGE Advance published date: August 18, 2022)

The clogging of submerged entry nozzles is a critical issue during continuous casting that adversely affects final product quality and process productivity. In order to impose effective monitoring and control over the continuous casting process, a quantitative index was formulated to quantify the magnitude of SEN clogging and erosion for a production dataset consisting of ultra-low carbon, low carbon, medium carbon, and calcium treated grades. Three critical index values are defined to represent the clogging event, erosion event, and critical casting condition. Long short-term memory network was established based on the quantitative index in the past four minutes to predict that in the future 48 seconds. The networks are found to be capable of predicting the overall trend in quantitative index, with the lowest normalized root mean squared error at 0.323 for medium carbon grade, followed by that at 0.340, 0.342, and 0.453 for low carbon, calcium-treated carbon, and ultra-low carbon grades respectively. The models can also identify most of the critical casting conditions and erosion incidents for all steel grades. Operators can take corresponding actions when critical conditions are predicted by the models in order to prevent the possible occurrence of clogging. Model precision could be improved with larger production datasets that consist of multiple number of clogging and erosion events.

KEY WORDS: steelmaking; continuous casting; submerged entry nozzle; clogging; erosion; time-series; LSTM.

1. Introduction

During the continuous casting process, clogging of the submerged entry nozzle (SEN) is one of the most disruptive phenomena caused by the accumulation of non-metallic inclusions. Clogging not only introduces unexpected operation downtime but also adversely affects downstream processes and product steel quality.1–4) The deposition of non-metallic particles attached to the SEN inner wall can build up into large clusters with time, which can be broken off with the high flow rate casting process and entrapped in the bulk or surface of the steel product. Such particle entrapment in the solidified steel is detrimental to its strength and toughness.5) Due to the adverse effect of clogging, it has been the subject of extensive research on understanding the formation mechanism, development of detection methods, and prevention of the clogging phenomenon. In general, clogging can be classified into four categories based on its formation mechanism: movement of oxides from steel to the nozzle wall, the transport of oxides present in the steel to the nozzle wall, air ingress into the nozzle, the chemical reaction between the refractory and the liquid steel, as well as steel solidified in the nozzle due to insufficient preheat of the nozzle.2,5,6) For aluminum-killed steel grades, aluminate is the common cluster material due to the transport of oxides from the tundish to the nozzle wall and from the interaction of nozzle refractory with molten steel.7–11) Calcium treatment is one of the measures taken to minimize the tendency of clogging aluminum-killed steel. The addition of calcium silicide transforms pure aluminate into calcium aluminate inclusions which exist in the liquid state at 1873 K, and depositions can be partially liquefied with the diffusion of CaO and flushed away with the flow. However, the inappropriate addition of calcium silicide can cause slow erosion of the stopper rod tip and SEN well block.12–14) Some others have tried to prevent nozzle clogging by injecting argon gas through circumferential porous slits inside the nozzle bores and stopper rod tip.1,2,12) On the other hand, many researchers have investigated the clogging by focusing on the fluid dynamic of the inclusion deposition to the refractory wall.

* Corresponding author: E-mail: kinnor.chattopadhyay@utoronto.ca

© 2022 The Iron and Steel Institute of Japan. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs license (https://creativecommons.org/licenses/by-nc-nd/4.0/).
Thomas and Bai reported that the clog materials tend to accumulate in the more stagnant regions right above the slide gate opening.\(^1\) Mukai \textit{et al.} developed a mathematical model to explain the mechanism of non-metallic inclusion deposition, and the author found that about 30% of the inclusions smaller than 20 \textmu m in diameter are forced to move towards the refractory wall.\(^5\) Long \textit{et al.} reported that inclusion particles with diameters larger than 100 \textmu m could not be attached to the nozzle wall, and the entrapment probability increases with decreasing size of inclusion.\(^16\) Ni \textit{et al.} demonstrated that deposition across the SEN is not uniform based on a three-dimensional CFD simulation.\(^17\) Gutierrez \textit{et al.} created a mathematical model and found two preferred deposition zones: upper tundish nozzle down the entry radius and the internal upper port surface, which are identified as low static pressure zones. Sambasivam reported that a redesign of the nozzle geometry could improve the clogging during continuous casting. The author reported that a nozzle with a parabolic curve-shaped bottom could create a better flow profile and exit characteristics and thus, reduces the possibility of clogging.\(^18\)

Many preventive techniques for clogging have been discussed above, and most researchers approach the issue either by using physiochemical approaches or computational modeling to understand the flow and design a corresponding solution. However, detecting the possibility of clogging is crucial by monitoring the continuous casting process in real time. The occurrence of clogging cannot be detected by a single process parameter, but rather a combination of different parameters such as argon back pressure, mold level fluctuations, stopper rod position, and casting speed etc.\(^1\) Direct measurement of clogging in the caster is difficult due to extreme temperature and environment. The flow pattern and liquid steel and non-metallic inclusions cannot be directly observed due to the same reason. The modeling of the casting process based on fluid dynamics and chemical reactions can be computationally expensive. As a result, a deep learning technique was implemented in this paper to monitor and predict the occurrence of clogging in real time. In order to effectively monitor the casting process and detect the clogging phenomenon in real time, a quantitative index needed to be developed to represent the magnitude of clogging or erosion of the SEN. There are three characteristic behaviors in casting: stable casting, clogging, and erosion.\(^19\) There has been some research on the development of quantitative indices for the casting process. McKague and Kemeny developed a quantitative index based on the stopper rod movement. The clogging severity was described as the ratio of the number of increments in stopper rod position to the number of decrements in stopper rod position.\(^2,20\) However, this index is developed solely based on stopper rod movement and has limited capability in distinguishing between the effects of other operating parameters and clogging buildup. Girase developed a clogging severity factor, the ratio between the actual and theoretical flow rates with a baseline of zero clogging buildup.\(^21\) This factor is developed based on tundish geometry, tundish weight, and casting speed. In 2016, Rajendra formulated a clogging index calculated as the ratio of the theoretical stopper rod position to the actual stopper lift in the tundish to the maximum opening.\(^22\)

This study proposes a novel method to monitor the casting condition using long short-term memory (LSTM) networks based on the clogging index. The clogging index is developed from various process parameters to effectively represent the occurrence of clogging, erosion, and critical situation when corrective actions need to be imposed. The deep learning model in this paper was developed by using the time-series data for continuous casting collected from the industrial partner.

2. Data and Methodology

2.1. Time-series Data for Modeling

In this paper, four steel grades were considered for model formulation: calcium-treated grade (Ca-treated), medium-carbon grade (MC), low-carbon grade (LC), and ultra-low carbon grade (ULC). During continuous casting, molten steel flows from a ladle through a tundish into the mold to produce steel products with designated shapes. The casting of one ladle can take 40 to 90 minutes, depending on the steel grade and casting condition, etc. For the present study, production data were obtained over various number of heats for each steel grade as shown in Table 1, and their data was sorted chronologically in a time-series manner. Each data point was sampled every four seconds and contains process parameters listed in Table 2.

As mentioned before, it is rather difficult to visualize the presence of clogging based on a single process parameter. Figures 1(a) and 1(b) correspond to two casting processes with stable casting condition and clogging occurrence respectively. For the stable casting, all process parameters are relatively steady throughout the heat, and only argon pressure fluctuates between 20 and 60 SLPM. On the other hand, all process parameters tend to fluctuate at a higher magnitude for the casting process with a fully clogged nozzle, as shown in Fig. 1(b). However, clogging cannot be directly identified based on the fluctuations of parameters. Therefore, it is necessary to formulate a quantitative index.

<table>
<thead>
<tr>
<th>Steel grade</th>
<th>Number of heats</th>
<th>Equivalent production time (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ULC</td>
<td>245</td>
<td>184–368</td>
</tr>
<tr>
<td>LC</td>
<td>454</td>
<td>340–680</td>
</tr>
<tr>
<td>MC</td>
<td>587</td>
<td>440–880</td>
</tr>
<tr>
<td>Ca-treated</td>
<td>480</td>
<td>360–720</td>
</tr>
</tbody>
</table>

Table 1. Steel grade and respective number of heats.

<table>
<thead>
<tr>
<th>Process parameters</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mold level depth</td>
<td>mm</td>
</tr>
<tr>
<td>Mold width</td>
<td>mm</td>
</tr>
<tr>
<td>Argon pressure</td>
<td>kPa</td>
</tr>
<tr>
<td>Argon gas flowrate</td>
<td>Standard Litre Per Minute (SLPM)</td>
</tr>
<tr>
<td>Stopper rod position</td>
<td>mm</td>
</tr>
<tr>
<td>Timestamp</td>
<td>Date/hour: minute: second</td>
</tr>
<tr>
<td>Heat length</td>
<td>minute</td>
</tr>
</tbody>
</table>

Table 2. Process parameters included in the time-series dataset for all steel grades.
based on these process parameters to effectively monitor the casting process and detect the presence of clogging phenomena.

2.2. Quantitative Index for Continuous Casting

Since the magnitude of deposition buildup cannot be directly monitored by a single process parameter, a quantitative index (QI) is necessary to be developed to effectively monitor the casting process. One of the requirements for this index is to distinguish between the presence of clogging events and mold fluctuation caused by changes in operating parameters such as casting speed. In the present study, the clogging index developed by Rajendra is utilized, and it can be expressed as:

$$QI = \frac{\text{Actual lift of stopper} - \text{theoretical lift of stopper}}{100 - (\text{theoretical lift of stopper} + \text{Zeroing Calibration})}$$

where the actual lift of stopper can be directly accessed from the dataset, and the theoretical lift of the stopper represents the stopper rod position for a specific casting speed when zero deposition is present in the SEN. In order to calculate this value, an equation to calculate the theoretical flow rate ($Q_{\text{theoretical}}$) needs to be created based on the geometry of the nozzle, which can be expressed as:

$$Q_{\text{theoretical}} = C_d \times \left( \frac{1}{2} 	imes (2g \times (0.0567 + (2.57 \times 10^{-3} \times W))^1.751 \times 60 \times \left( (192.988 \times N + (0.620 \times N^2 - 4.22) \times 10^6 \right) - D_s \right)$$

where $C_d$ is the discharge coefficient, $W$ is the tundish weight, $N$ is the theoretical lift of the stopper, and $D_s$ denotes the density of liquid steel. The theoretical stopper rod lift ($N$) in Eq. (2) can be determined by equating $Q_{\text{theoretical}}$ to the actual flow rate ($Q_{\text{actual}}$), and $Q_{\text{actual}}$ can be calculated through:

$$Q_{\text{actual}} = V \times A_{\text{ring}} \times D_{\text{steel}}$$

where $V$ represents the casting speed of the machine, $A_{\text{ring}}$ denotes the cross-sectional area between the stopper rod and the SEN.

2.3. Determine Critical Quantitative Index

One can effectively monitor the casting condition and detect the presence of clogging with the developed QI. The magnitude of the index is directly correlated with the severity of clogging deposition in the SEN. In theory, its value ranges from zero to one, where zero corresponds to no deposition present in the nozzle, and one represents a fully clogged SEN.

Clogging phenomena can be observed for ULC, LC, and MC grades when QI reaches or exceeds one. On the other hand, SEN erosion can be identified for Ca-treated grade when the index value drops below zero. To prevent production downtime due to the fully clogged nozzle, corrective actions should be taken when the casting condition is heavily affected by non-metallic deposition. Such flow conditions can be defined with a critical QI value. Operators can incorporate measures such as flushing the nozzle to remove deposition accumulation when the critical value is met. In this study, QI equal to 0.5 is considered the critical point because it is equivalent to a situation where half of the SEN is blocked by deposition. In conclusion, QI equal to or above one corresponds to the occurrence of a fully clogged nozzle, below zero represents the SEN erosion specific for Ca-treated grades, and equal to or above 0.5 is the critical condition when corrective actions need to be taken.

2.4. LSTM Network

LSTM is a specific type of recurrent neural network (RNN) that overcomes the problems of vanishing and exploding gradients of traditional RNNs. Different from the conventional neural networks, where the features only progress in the forward direction, RNN has recurrent looping connections where information can be fed back into the input. These connections in RNNs enable it to capture the sequential information hidden in the input data, thus making RNNs excel in time-series prediction and natural
language processing. However, the traditional RNNs suffer from exponentially growing gradients as training progresses, known as the exploding gradient. As a result, conventional RNNs cannot effectively learn the underlying relation within time-series data when the input lags are greater than 5 to 10 steps.

The LSTM can address the limitation of exploding gradients by having self-connected memory blocks in the hidden layer, which allows the capture of long-term dependencies. As a result, LSTM has state-of-the-art performance in various fields such as speech recognition, handwriting recognition, and time-series modeling. Inside the LSTM network are internal units called gates that regulate the flow of information. Three types of gates are available in each memory block, including input gate (i_t), output gate (o_t), and forget gate (f_t) as expressed as:

\[ i_t = \sigma (w_i [m_{t-1}, x_t] + b_i) \] 
\[ o_t = \sigma (w_o [m_{t-1}, x_t] + b_o) \] 
\[ f_t = \sigma (w_f [m_{t-1}, x_t] + b_f) \]

where \( \sigma \) is the sigmoid activation function, \( w_{i/o/f} \) are the weight parameters for each gate, \( m_{t-1} \) is the output from the previous timestep, \( x_t \) is the input at the current timestep, and \( b_{i/o/f} \) are the bias vectors for each gate. The input gate is in charge of the activation flows into the memory, and the output gates control the information flow from the block into the rest of the network. The forget gate bypasses information or resets the cell state when it is not needed.

The information from the previous state and the current state input are scaled through a sigmoid function; then the output result determines whether the information is to be thrown out or kept.

### 2.5. Feature Processing and Model Validation

LSTM networks were modeled based on the QI developed from section 2.2. Because the production dataset was organized chronologically in a time-series manner, the corresponding index value for each datapoint was also sorted in the same order. Before the development of the LSTM network, the quantitative index needs to be processed with the sequencing function for time-series modeling. The sequencing function defines the number of historical data points needed to predict a value in the future in a time-series manner. As mentioned before, the production datapoints are recorded every four seconds; thus, at a specific timestamp \( t \), one lag corresponds to timestamp \( t-1 \), which is equivalent to four seconds in the past.

On the other hand, datapoint in one future timestamp \( (t+1) \) equals four seconds ahead of the current timestamp. In this study, LSTM networks use 60 lags to predict the 12 timestamps in the future, equivalent to predicting the clogging index in 48 seconds based on historical data from the past four minutes. The input vectors for the networks are prepared by applying the sequencing function, each datapoint needs to contain an input feature consisting of QI from \( t-60 \) to \( t-1 \), and the target variable of QI at \( t+12 \) for a timestamp at \( t \) as shown in Table 3. Moving towards the following timestamp \( t+1 \), all input features and the target variable must be adjusted by adding one timestamp. The sequencing function was conducted throughout the dataset for each steel grade.

Once the features are processed with the sequencing function, they need to be converted into a three-dimensional feature shape of \((n, t, f)\), where \( n \) is the number of datapoints, \( t \) represents the number of timesteps of the input feature, and \( f \) corresponds to the number of features. In this study, the number of datapoints varies with steel grade, number of timestamps equals 60, and \( f \) is defined to be one since clogging index is the only feature that is used in LSTM modeling.

Machine learning models tend to perform very well on datasets that they have seen but worse on unseen datasets, which is a commonly known issue called overfitting. In order to prevent the overfitting, during the development of LSTM networks, three-fold expanding cross-validation is conducted to select the optimal hyperparameters of LSTM networks and validate the generalization of predictions. In this study, the time-series datasets were split into a training set consisting of 70% of the data and a testing set with the rest of the data. When LSTM undergoes the training phase, the three-fold expanding cross-validation further divides the training set into three folds. Within each fold, the training set consists of a specific length of time-series history whose length grows over subsequent folds. On the other hand, the length of the validation set is always constant across all folds, as shown in Fig. 2. The training set is used to train the LSTMs by updating internal parameters, and they are tested by using the validation set after each fold. Finally, the hyperparameters with the lowest validation error are selected as the optimal hyperparameters for the model.

Once LSTM networks are created, models are evaluated based on root mean squared error (RMSE) which can be expressed as:

\[ RMSE = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \hat{x}_i)^2}{N}} \]

where \( x_i \) is the \( i^{th} \) observed value of \( x \) and \( \hat{x}_i \) is the \( i^{th} \) predicted value.

### Table 3. Example of time-series features processed with sequencing function.

<table>
<thead>
<tr>
<th>Input features (X)</th>
<th>Current timestamp</th>
<th>Target variable (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-60, t-59, …, t-2, t-1</td>
<td>t</td>
<td>t+12</td>
</tr>
<tr>
<td>t-59, t-59, …, t-2, t</td>
<td>t+1</td>
<td>t+13</td>
</tr>
<tr>
<td>t+1, t+2, …, t+n-1</td>
<td>t+n</td>
<td>t+n+12</td>
</tr>
</tbody>
</table>

Fig. 2. Three-fold expanding window cross-validation method. (Online version in color.)
predicted value of x. RMSE is a good measure of prediction accuracy, however, it is scale-dependent, and it can be challenging to compare models between for different target variables. In order to compare model performance independent of the scale, the normalized RMSE (NRMSE) is also estimated for each model, which can be expressed as:

\[
NRMSE = \frac{RMSE}{\sigma}
\]

where \(\sigma\) is the standard deviation of target variable. Lower value of RMSE and NRMSE correspond to a better predictive accuracy.

3. Results and Discussion

3.1. Visualization of Clogging Index

As discussed in Section 2.3, the magnitude of QI is directly related to the deposition accumulation within the SEN. The grade of steel is one of the major factors in determining the flow condition, in order to have a better understanding of the overall casting situation for each steel grade, quantitative indices are visualized in Fig. 3. From these figures, it can be observed that the clogging indices for ULC, LC, and MC experience relatively larger fluctuations than the Ca-treated grade, where index values for ULC, LC, and MC grades are mainly distributed between 0 and 1, and those for Ca-treated grade are mostly located between 0 and 0.4. Based on this observation, it can be reported that the overall casting process for Ca-treated is smoother than the other three grades due to the lack of clogging events. However, multiple erosion incidents are present for the Ca-grade as displayed when QI drops below zero.

3.2. Results of Critical Clogging Index, Clogging Phenomenon, and Erosion

The casting process for ULC, LC, and MC grades frequently fluctuate due to aluminate deposition in the nozzle. Based on the theoretical calculation of the QI, the nozzle is fully clogged when the index equals or exceeds one. However, there is an exception known as a caster event. Caster events are corrective actions imposed by operators when they believe the amount of deposition will be an issue if continuing production. General caster events can be flushing the SEN by injecting argon gas, and they are usually conducted solely on operators’ experience. During the caster events, the stopper rod is lifted to the maximum position, thus resulting in QI being larger than one.

In order to distinguish the caster events from actual clogging incidents, it has been found that the caster events usually take less than one minute in the dataset. In contrast, the clogging phenomenon can last much longer. As a result, caster events can be filtered from clogging events based on their duration. Based on the dataset for each steel grade, the number of clogging, erosion, critical, and caster events are summarized in Table 4.

3.3. LSTM Hyperparameter Tuning

Hyperparameters of machine learning models are parameters that cannot be derived through the training phase. They are crucial to predictive performance since they are directly related to learning time and model convergence. In this study, the hyperparameters are defined and tuned via the
trial-and-error approach accomplished with the three-fold expanding window cross-validation technique. As explained in Section 2.4, LSTM is a specific type of RNN that can receive and process data while propagating information forward as training progresses, and the structure of LSTM allows it to preserve information through backpropagation of the error for prediction in the future. This study created double LSTM-layer networks with the same number of neurons. Learning rate, number of neurons, dropout rate, and batch size were selected based on cross-validation, with the starting value, increments, and the end values displayed in Table 5. Learning rate is a configurable hyperparameter that controls how quickly a network learns. A small learning rate can consume extensive computational time for training, whereas a large rate can cause faster convergence outside of the global minimum. 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. In general, larger number of neurons can increase accuracy with proper regularization techniques. LSTM layers within the network are accompanied by dropout layers, which is a layer that helps reduce overfitting by randomly bypassing a certain number of neurons to reduce the effects of particular neurons. Finally, batch size defines the number of training examples for the LSTM to feedforward and backpropagate before updating internal parameters. A large batch size may require expensive computational resources such as the memory space, whereas a small batch size can adversely affect the training time. Four LSTM networks were developed for each steel grade. For each model, the optimal hyperparameters were selected based on the combination that achieved the lowest RMSE on the testing set.

### 3.4. Results of LSTM Networks

Once the LSTMs were created and tuned based on the training set, they were then validated against the testing set that consists of 30% of the datapoints. In order to save computational time and prevent overfitting, early stopping was deployed as a regularized method in addition to the dropout layers. RMSE values were recorded for each LSTM network when tested against the testing set as presented in Table 6. These evaluation metrics represents the goodness of time-series predictions on unseen datasets. To compare the time-series predictions among all four steel grades, it can be seen that the LSTM for MC grade achieves the lowest NRMSE at 0.323, followed by LC and Ca-treated grades at 0.340 and 0.342 respectively. Meanwhile, LSTM for ULC grade attained the highest NRMSE at 0.453. One of the reasons for which the LSTM for MC grade performed the best is because its dataset consists of the highest number of critical QIs among all grades, which is 21 in comparison with ULC grade that only has 11 within the dataset.

Predictions of LSTM against the actual QI for each steel grade are visualized in Fig. 4. It can be observed that the LSTMs are capable of predicting the overall trends of the quantitative index for all grades. However, for certain peaks that correspond to the clogging event, the LSTM is incapable of predicting the exact value, as shown in Fig. 5. In this case, a fully clogged nozzle occurs during the production, represented by the actual QI exceeding one, whereas the predictions generated by LSTM scatter around 0.6 throughout the clogging phenomenon. Despite the magnitude of error, the peak of this prediction can still be distinguished from peaks of non-clogging events, as shown in Fig. 4(a), because all other peaks within the testing set for ULC only reach 0.4. One of the possible causes of the magnitude of error could be the model’s inability to accurately capture the non-linear relationships present in the data.

<table>
<thead>
<tr>
<th>Steel grade</th>
<th>QI &lt; 0 or QI &gt; 1</th>
<th>Clogging</th>
<th>Caster event</th>
<th>Erosion</th>
<th>Critical QI (QI &gt; 0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ULC</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>LC</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>MC</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Ca-treated</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5. Selection of hyperparameters for LSTM models for each steel grade.

<table>
<thead>
<tr>
<th>Hyperparameters</th>
<th>Starting value</th>
<th>Increment</th>
<th>End value</th>
<th>Selected parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning rate</td>
<td>0.005</td>
<td>0.005</td>
<td>0.02</td>
<td>ULC 0.005 LC 0.005 MC 0.005 Ca-Treated 0.005</td>
</tr>
<tr>
<td>Number of neurons</td>
<td>64</td>
<td>64</td>
<td>128</td>
<td>128 128 128 128</td>
</tr>
<tr>
<td>Dropout rate</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3 0.2 0.1 0.2</td>
</tr>
<tr>
<td>Batch size</td>
<td>256</td>
<td>256</td>
<td>512</td>
<td>256 256 512 512</td>
</tr>
</tbody>
</table>

Table 6. RMSE and NRMSE of LSTM networks for each steel grade, and their predictions on critical casting conditions and erosion events within the testing set.

<table>
<thead>
<tr>
<th>Steel grade</th>
<th>RMSE</th>
<th>NRMSE</th>
<th>Number of critical QI</th>
<th>Predicted critical QI</th>
<th>Number of erosion</th>
<th>Predicted erosion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ULC</td>
<td>0.0572</td>
<td>0.453</td>
<td>1</td>
<td>1</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>LC</td>
<td>0.0385</td>
<td>0.340</td>
<td>5</td>
<td>5</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>MC</td>
<td>0.0345</td>
<td>0.323</td>
<td>7</td>
<td>5</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Ca-treated</td>
<td>0.0269</td>
<td>0.342</td>
<td>/</td>
<td>/</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

© 2022 ISIJ
this prediction error is the lack of clogging incidents in the training set. As recorded in Table 4, there is only one clogging event for the ULC dataset, which is included in the testing set, meaning the LSTM has never seen a clogging event during its training phase. This issue is expected to be resolved with a larger training set consisting of multiple clogging phenomena. On the other hand, there are also peaks associated with caster events that cannot be predicted, as shown in Fig. 4(b), because they are corrective measures directly imposed by operators. Therefore, to further evaluate the performance of the LSTM networks, the predicted and actual critical QIs and erosion incidents are recorded and compared in Table 6. The LSTM networks for ULC and LC can predict all of the critical QIs whose index value is equal to or above 0.5. For MC grade, 71% of the heats with critical QIs are predicted. And 67% of the erosion phenomenon is successfully forecasted for the Ca-treated grade. Based on the number of critical casting conditions and Fig. 4, it can be concluded that the critical QI is a good indicator of the possible clogging event presence. Thus, operators can take corresponding actions when critical QIs are achieved when implementing the LSTM models during production.

4. Conclusions

A novel method for predicting the clogging and erosion phenomena in the continuous casting process was proposed. The LSTM networks were developed with the quantitative index for ULC, LC, MC, and Ca-treated grades based on production data collected from the industrial partner. The following results were obtained:

(1) Quantitative index can effectively represent the cast-
ing condition with three critical values. Index value of one represents a fully clogged nozzle, below zero corresponds to the erosion of the nozzle, which is a special phenomenon for Ca-treated grade, and 0.5 is the critical casting condition where corrective measures should be imposed to avoid a clogging event.

(2) LSTM networks developed based on the processed feature by sequencing function can successfully predict the trend of the future quantitative index in 48 seconds based on 4 minutes of lag. Among all grades, the LSTM for MC grade achieves the lowest NRMSE at 0.323, followed by that for LC, Ca-treated, and ULC grades at 0.340, 0.342, and 0.453, respectively. In addition, the LSTMs for ULC and LC achieved 100% accuracy in predicting the critical QI, followed by 71% for MC grades and 67% in predicting erosion incidents for Ca-treated grade in the testing set.

(3) Operators can take corrective actions during the production when LSTM predicts the critical QI to prevent clogging.

(4) Prediction precision of the LSTM could be improved with a larger dataset consisting of multiple clogging or erosion incidents in the training set.

Acknowledgements
The authors would sincerely thank Stelco Inc. for providing all the necessary support towards this study.

Funding Sources
Funding: This work was supported by the NSERC CRD program in collaboration with Stelco Inc.

REFERENCES