Multivariate Time Series for Data-driven Endpoint Prediction in the Basic Oxygen Furnace

Davi Alberto Sala*, Azarakhsh Jalalvand*, Andy Van Yperen-De Deyne† and Erik Mannens*
* *IDLab, Ghent University–imec, Belgium*
Email: DaviAlberto.Sala@UGent.be
†*ArcelorMittal Belgium, Ghent, Belgium*

***Abstract*—Industrial processes are heavily instrumented by employing a large number of sensors, generating huge amounts of data. One goal of the Industry 4.0 era is to apply data-driven approaches to optimize such processes. At the basic oxygen furnace (BOF), molten iron is transformed into steel by lowering its carbon content and achieving a certain chemical endpoint. In this work, we propose a data-driven approach to predict the endpoint temperature and chemical concentration of phosphorus, manganese, sulfur and carbon at the basic oxygen furnace. The prediction is based on two distinct datasets. First, a collection of static features is used which represent a more classic data-driven solution. The second approach includes time-series data that provide a better estimate of the final endpoint and enable further tuning of the process parameters, if necessary. For both approaches, model-based feature selection is used to filter the most relevant information. Results obtained by both models are compared in order to estimate the added value of including the time series data analysis on the performance of the BOF process. Results show that a simple feature extraction approach can enhance the prediction for phosphorus, manganese and temperature.**

***Keywords*-Time-series data analysis, sensors, prediction, basic oxygen furnace, steel industry**

I. Introduction

Steel consumption in Europe has recovered less than expected after the crisis in 2008. This has led to a tough competition in this industry, in which process optimization and product quality play a crucial role. Thanks to the revolutionary advances in computational hardware and data analysis, a large amount of stored data can be used to enhance this process. This work presents the preliminary results of a joint research for data-driven process optimization that is being conducted within ArcelorMittal Gent (AMG).

The Basic Oxygen Furnace (BOF), also known as oxygen converter process, is a mechanism to transform hot metal into liquid steel. Blowing oxygen through the hot metal reduces its carbon concentration and results in low-carbon steel. During this process, data from more than 40 sensors are collected along with more than 300 static variables.

In each complete BOF process, on average 255 tons of hot metal is converted to liquid steel in less than 16 minutes. The quality of the steel is usually determined by the concentration of certain elements such as carbon, phosphorus, sulfur, silicon and manganese as well as the temperature reached at the endpoint. The precise prediction of the endpoint state plays a major role for an effective BOF process control, as it has direct impact on the steel quality and cost of the production.

Prediction of the BOF process is mainly carried out using either model-driven or data-driven approaches.

The model-driven prediction approach is based on the chemical and physical properties of the process by expressing the relations mathematically, while the data-driven solution aims to learn the relations among the available data to predict the desired output.

The focus of this work is the data-driven approach for endpoint prediction. First, we investigate the influence of BOF static variables that might be relevant to the process. This is accomplished by creating a dataset composed of features from the analytical (model-driven) approach and complemented by the most relevant static data features, which are selected via feature importance through model selection. The data is processed by classic prediction models such as Ridge Regression, Random Forest and Gradient Boosted Regression Trees. The predictive model based on the static variables (called *SF* dataset) provides an estimate of the endpoint prior to the measurements, which results in a more accurate calculation of secondary additions and corrections, such as addition of phosphorus or sulfur, addition of coolant if target temperature was exceeded, among others.

In the second phase, we apply feature extraction on the time-series signals recorded by the sensors during oxygen blow and create a second dataset (called *TS* dataset). We develop a time-series model to more accurately estimate the endpoint near the end-of-blow, allowing further chemical and temperature corrections to be taken, if necessary.

In this work we present preliminary results obtained on the prediction of endpoint steel compositions of phosphorus, carbon, manganese, sulfur as well as the final temperature obtained based the two aforementioned datasets, with the main objective of verifying if time-series signals contain relevant information to enhance predictive systems.

The rest of the paper is organized as follows: Section II introduces the BOF process for a better understanding of the problem and Section III gives an overview of the related work found in the literature. Section IV describes

the proposed methodology for feature selection as well as the regression models. In Section V, we present the obtained results which are followed by a conclusion and some ideas for further investigation.

II. An overview of the BOF process

A. Basic Oxygen Furnace Steel-making process

In the blast furnace, iron ore (which is mainly composed of iron and oxygen) is melted in a reducing atmosphere by lowering its oxygen content. The process begins with loading the furnace with coke and sinter. Hot blast is used to turn the coke into the gases necessary for reducing the iron oxides within the ore into metallic iron, according to the equation: $FeO_3 + 3CO \rightarrow Fe + 3CO_2$. The exothermic reaction provides the necessary heat for melting the reduced ores. The liquid metal is then transported to the steel plant, in the Basic Oxygen Furnace, which is the main focus of this study.

Scrap and hot metal are charged into the converter vessel. In the BOF process pure oxygen is blown on the metal bath by means of a water-cooled lance. Burned lime and other additives are added during blowing to regulate the process and achieve targeted composition. The blowing phase takes up to sixteen minutes. Meanwhile, an inert gas is injected via the bottom of the BOF vessel to maintain the mixture homogeneous.

The blow partly oxidizes the carbon, silicon, manganese, phosphorus and iron in the bath. These transformations liberate a huge amount of heat, which melts the scrap and raises the bath temperature. The impure elements are converted into gas or slag, floating on the top of the liquid bath. By the end of the blowing stage, the bath temperature is about 1650^oC, and the steel mass is about 300t.After the blowing phase, the vessel is tilted, steel is tapped into a steel ladle and the slag is tapped into a slag pot. Temperature is measured and steel sample is taken at this stage, secondary analytical models are used to provide a more accurate endpoint prediction in order to add the correct amount of additions as soon as possible (if necessary). The converter is ready for the next batch, while the molten steel is further refined and alloyed at the ladle metallurgy and cast into slabs in continuous casting. Afterwards, these slabs are further processed to a hot rolled coil and possibly finished in the cold strip mill and coating installations. Fig. 1 illustrates the main steps of the BOF process. More details about the chemical reactions that occur in a BOF can be found in [1].

The individual or combined effects of carbon (C), phosphorus (P), sulfur (S), nitrogen (N), hydrogen (H), manganese (Mn) and oxygen (O) compositions have a large impact on many aspects of the steel properties, such as tensile strength, formability, toughness, weldability, resistance to cracking, corrosion and fatigue. Optimizing the BOF process is crucial to obtain the correct measurement of these targeted chemical compositions. Additionally, the temperature of the heat has a significant impact on the supply chain since the casting is highly dependent on it.

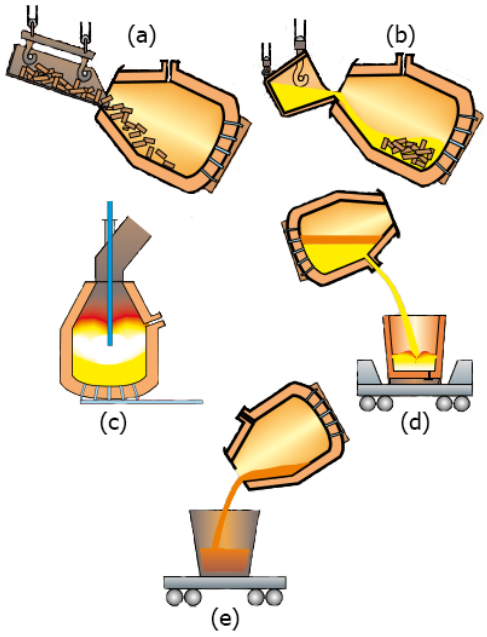


Figure 1: A diagram of the main stages in the BOF process: (a) scrap metal and (b) hot metal are charged into the ladle; (c) oxygen is blasted in the bath while an inert gas is injected through the bottom to keep the mixture homogeneous; (d) liquid steel is tapped; (e) the remaining slag is removed.

Regarding the BOF, a precise estimation of the endpoint values of chemical composition, weight and temperature are crucial for maintaining the steel quality and on the supply chain. The predictions are conducted in two phases; prior blow, for process control and is optimized around such values, to achieve the desired endpoint. Subsequently a secondary endpoint prediction is realized after blow, also accomplished by a model based on analytical features. This leads to a more accurate prediction than the estimated at prior blow, and happens within a time window where corrections to the steel composition are still possible.

Usually the composition of the molten steel is determined in laboratory, but liquid steel samples are only acquired with a success rate of roughly 70%. Such a low hit rate is caused by the very high temperature and the slag layer on top of the steel, which interferes with the sampling procedure. This adds to the justification of a more accurate predictive system approach, next to the direct availability of an accurate value to start injecting additions. The focus of this research is on the endpoint prediction of C, P, Mn and S concentrations as well as the final bath temperature.

B. Collecting data during the BOF process

The desired quality of the final steel product determines the weight and type of scrap, quantities of hot metal,

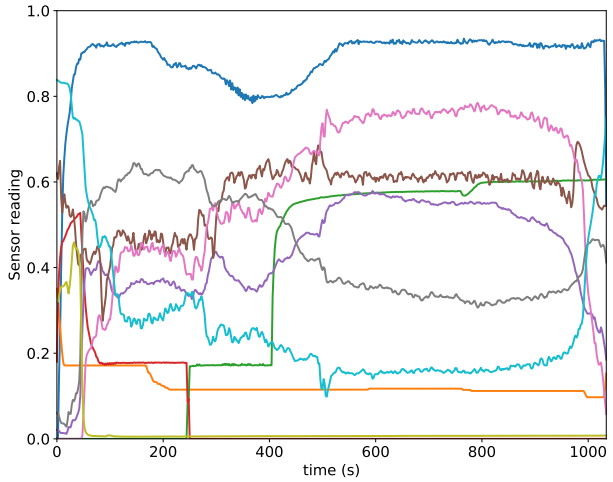


Figure 2: Re-scaled time-series of 10 signals from one batch. Signals are related to controllable variables such as oxygen flow, lance position and inert gas flow, and non-controllable ones such as amount of off-gases in the chimney, decarbonization curve and acoustic levels.

iron ore and lime additions, blow time, and many other BOF controllable inputs and process variables. These are calculated using the models for carbon, iron and temperature.

Once all the inputs are defined and before the blow begins, the analytical predictive models are used to estimate the endpoints of the blow, taking into account the chemical and physical characteristics and reactions for each endpoint target. Each predictive model in this phase, takes into account around 10 static variables used in the analytical models (based on the physical and chemical behavior of the BOF process). Based on the prediction, further chemical additions can be made during the blow to achieve the desired endpoint values.

The blowing phase lasts about 16 minutes in which several time-series signals are recorded during the blow with the frequency of $0.5Hz$, from which 10 are selected as most relevant by the process engineers and are re-scaled to the range of [0-1]. Fig. 2 shows the scaled signals of one heat (batch).

III. Related work

In the latest works from Han et al. [2], a hybrid model based on Fuzzy C-Means clustering (FCM) and Support Vector Machine (SVM) is trained to predict the endpoint temperature and carbon content of a BOF steel making process. It was shown that the proposed method performs better than the conventional solutions such as multi-layer perceptron (MLP), classical SVM and Case-based reasoning (CBR). The experiments are conducted on a small and closed dataset that contains only static variables information from each batch, such as, amount of scrap and iron, carbon content and temperature after the first blow, coolant.

A multivariate solution is presented in [3], where a search algorithm is used to select relevant static features (from a total of 52 features) for predicting the endpoint concentrations of Fe, Mn, P and S within the BOF process. The authors trained a distinct model for each of the seven possible steel mixes and endpoint target. The performance of the models are reasonable but a high variance is observed which is explained by the high uncertainty in the input variables.

Also, a model for predicting the endpoint phosphorus content on the BOF process is presented in [4]. The study employs 21 input variables, which are used to cluster the training data using k-Means (k is empirically determined). Each cluster is then considered to create a Polynomial Neural Network regressor. Results show that clustering the data can have a positive impact on the results when compared with a single regressor model.

Data-driven models have been widely applied on other stages of steel production, beside the BOF process. In [5], a Boosting Ensemble method is presented for predicting the endpoint temperature on the Ladle Furnace, based on the static variables. Also, the authors of [6] propose an online multivariate based model for steel casting monitoring. As other predictive use-cases in steel industry, yield prediction [7] and Steel Milling [8], [9] can be pointed out. All these works contribute with improved data-driven models over analytical approaches in steel production process.

Time-series in other industry sectors are also vastly applied on several applications, such as quality control, soft sensing and predictive maintenance. A Neurofuzzy model for temperature prediction on a polymerization reactor is presented at [10]. The approach uses several time-series and static variables to predict the temperature forecast within the next five minutes.In [11], a soft sensor-based regression ensemble is used to estimate the concentration of penicillin during the fermentation process on a batch basis. The soft sensor consists of 10 regression models, one for each time-series input, and is able to predict the current penicillin concentration (on each time step) with a low error rate. Soft sensors for continuous processes such as gas and liquid flow estimation also had recent developments in the industrial applications. AL-Qutami et al. [12] introduce an ensemble of diverse neural networks to learn a soft sensor for multiphase flow metering on continuous process. In [13], a soft sensor based on State Dependent Parameter model, with the goal to estimate concentration of gas on a stirring tank is used for quality prediction of the continuous process.

The literature study reveals that the steel industry heavily relies on analytical models or data-driven approaches that do not take advantage of all the process information data. Our aim is to utilize the data-driven techniques to enhance the prediction models taking into account all available information. In our approach, we evaluate the impacts of using extra static features selected from the process data

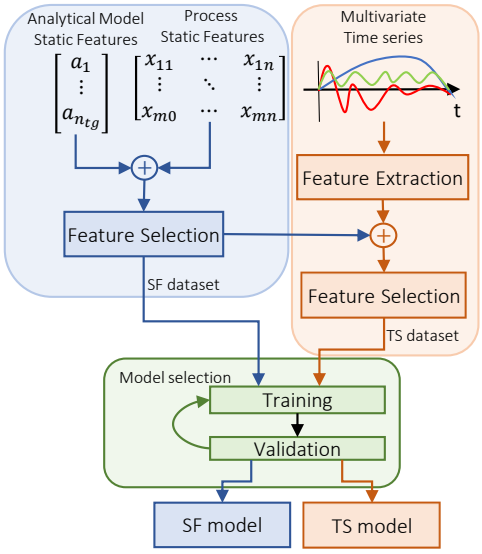


Figure 3: Workflow for dataset creation and model selection. The *SF* dataset is composed by the n_{tg} features from the analytical model and the most relevant features from the process selected via GBRT. For the *TS* dataset, several features are extracted from the signals and joined with the n_p features from the *SF* dataset, the n_{ts} most important features are selected (via GBRT) to complement the n_p.

and dynamic features extracted from time series, training classic prediction models to further reduce the error when compared to the current approach.

IV. Methodology

Our approach is composed of three parts; Firstly, we evaluate all the static features that are available at the end of the oxygen blow and select the n_p most relevant ones through model selection, among these features are the n_{tg} variables used on the analytical model-based approach for each target. This data forms the *SF* dataset. On a second stage, features from 10 time-series signals are extracted and the n_{ts} most important ones are joined with the n_p static features. The number of features n_p and n_{ts} are selected by evaluating the performance of several models on different configurations. This procedure leads to the *TS* dataset. In the last step, regression models are trained and evaluated on both datasets and a model selection is made after hyper-parameter optimization. This work-flow is presented in Fig. 3.

The recorded data in the interval consists of the most recent 7199 heats. The usable dataset after filtering corrupted or incomplete data, is composed of 6868 heats. This data was split into 70% for training and 30% for testing. With the aim of keeping the temporal order, the data points are not shuffled. This way we respect the realistic setup, in which a model would be deployed into production for predicting the next heat batches. We studied the application of Ridge Regression, Random Forest Regression and Gradient Boosted Regression Trees for each of the five endpoint targets, P, C, Mn, S and Temperature at each of the two blow phases. This leads to a total of 30 models. The important features are selected separately for each target. The hyper-parameters of the models are tuned via random search and a 5-fold cross validation on the training set. The rest of this section provides more details on feature extraction procedure and the developed prediction models.

A. Feature Selection for SF dataset

In order to select the input features from the analytical models, we followed the recipe in the BOF steel production literature. For each endpoint target (P, C, Mn, S and temperature) n_{tg} features are suggested by the experts. These so-called model variables are usually used for prediction models as noted in Section III. Using a large amount of irrelevant variables may deteriorate model behavior and increase the risk of overfitting [14]. Therefore, feature selection is an important asset to avoid such issues.

For each endpoint target, n_{tg} model variables are joined with all the available static features and are used to optimize a Gradient Boosted Regression Trees (GBRT) by 5-fold cross-validation. We adopted Gradient Boosting for feature selection because it can reliably select relevant features and identify non-linear feature interactions [15]. Another advantage of this method is the scalability to larger datasets. The feature importance order obtained from the optimized model is then used for feature selection.

n_p is optimized based on the trade-off between complexity and performance during a 5-fold cross validation on GBRT for all targets. For instance, 8 features have been suggested for the target phosphorus ($n_{tg} = 8$) in the literature. If the optimum number of features for the targets is 100 ($n_p = 100$), the 92 most relevant $staticfeatures$ will be added. This will be studied in Section V.

B. Feature Extraction for TS dataset

With the aim of extracting the most informative features from the available time-series data, we utilize the python package *TSFresh* as presented in [16]. These features include frequency components, statistic values, entropy, among many others.

The feature selection and optimizing the final number of features (n_{ts}) is performed in the same vein as in the *SF* dataset by using a 5-fold cross validation training on a GBRT algorithm.

C. Prediction Models

The prediction models in the BOF process are not only used to tune the parameters, but also to have a better understanding of the conditions which is crucial to control and optimize the process. Therefore in this work, we opted for the interpretable regression models (also known as white-box models) as opposed to black-box models. Apart from the

prediction, the white-box techniques provide a framework to measure and interpret the importance of input features, which can lead to expand the knowledge on improving the steel production process. The chosen models are presented below.

1) Ridge Regression: The Ordinary least squares fits a model with coefficients $w = (w_1, w_2, ...w_p)$ by solving $\min_w ||\mathbf{X}w - Y||_2^2$, with p being the number of features in the input matrix, $p = n_p$ and $p = n_p + n_{ts}$ for the *SF* and *TS* datasets, respectively. Matrix $\mathbf{X}$ contains the input feature matrix and Y the target values of the desired endpoint.

When data are correlated, the columns of $\mathbf{X}$ are approximately linearly dependent [14]. Consequently, the least square estimate becomes highly sensitive to random errors, producing a large variance. The Ridge regression imposes a penalty ($\alpha \geq 0$) on the size of the coefficients, which constrains the growth of the weights, as can be seen in (1).

$$\hat{\beta}^{ridge} = \min_w ||\mathbf{X}w - Y||_2^2 + \alpha||w||_2^2 \quad (1)$$

2) Random Forest (RF): Also known as Random Decision Forests, consists of an ensemble of decision trees. This method was chosen for being robust to outliers in the input space [14]. RF is computationally scalable as it can be trained on large datasets with several features without significant impact on computational time. Another advantage of RF for industrial applications is the white box architecture, enabling a more interpretable approach towards feature importance.

For regression tasks and given a joint distribution $(\mathbf{X}, Y)$, the prediction can be given by the unweighted average over the collection of trees:

$$h(\mathbf{x}) = \frac{1}{K}\sum_{k=1}^{K} h(\mathbf{x}; \theta_k) \quad (2)$$

where, $\mathbf{x}$ represents the observed input covariance vector associated with random vector $\mathbf{X}$ and the θ_k is the independent and identically distributed random vector associated with the k-st decision tree ($h(\mathbf{x}; \theta_k)$) with $k = 1...K$.

3) Gradient Boosted Regression Trees (GBRT): Hypothesis boosting, or simply "boosting", is an ensemble method for improving model predictions of a given algorithm in supervised learning, the method combines many "weak" classifiers to achieve a final strong classifier with reduced bias and variance. Boosting can be applied to many classification and regression methods and they can be implemented with different learner types.

The GBRT uses decision trees of fixed size as weak learners [14]. Then the additive model is built in a forward stage mode, $F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$. At each training stage $m = 1...M$, the decision tree $h_m(x)$ is chosen to minimize the loss function L given the last model F_{m-1} and its prediction $F_{m-1}(x)$. Steepest descent is used to solve this minimization problem:

$$F_m(x) = F_{m-1}(x) + \gamma_m \sum_{i=1}^{n} \nabla_F L(y_i, F_{m-1}(x_i)) \quad (3)$$

The descent step γ_m is chosen using line search. In this work, we used least squares as the loss function.

Individual decision trees internally execute feature selection when selecting split points. This information can be used to measure the importance of each feature; the more often a feature is used in the split points of a tree, the more significant that feature is. This approach can be expanded to decision tree ensembles by simply averaging the feature importance of each tree.

V. Results and Discussions

In this section we evaluate the described regression models the *SF* and *TS* datasets, containing data of 6868 heats, with a split of 4808 samples for training and 2060 heats for testing. We also compare the obtained results with the performance of the models that are currently being used in the factory for predicting the aforementioned five targets in the BOF metallurgical process.

The evaluation metric is the standard Root Mean Squared Error, given by $RMSE = \sqrt{\frac{1}{q}\sum_{i=1}^{q} e_i^2}$, where q is the number of test samples and e_i is the difference between the predicted and desired values for the i-th sample. Due to confidentiality reasons, all results presented here are normalized with respect to the measured values (training targets).

A. SF Models

A GBRT model is trained for each of the five target endpoints, on the training dataset of 265 static features and the parameters are optimized using random search [17]. A set of n_p variables is selected, containing the static variables from the analytical model of the desired output and the $n_p - n_{tg}$ most relevant features according to GBRT. These final n_p are used to train the models and optimize them based on random search.

Figure 4 shows the RMSE curves of GBRT as a function of n_p during the training and validation. The training was performed for 1000 boosting iterations of GBRT models targeting the phosphorus composition. Each iteration adds a new tree to the ensemble, as mentioned in IV-C. Reducing the number of features increases both training and validation performance, while removing too many features may result worse accuracy, as relevant process information might be left out. The phosphorus validation curve suggests an optimal range between 50 and 100 for n_p. The same behavior was observed for the other targets. With respect to a trade-off between complexity and performance, $n_p = 50$ was selected as the optimal value.

Table I lists the performance of the three regression models trained on the 50 selected features for the five

Figure 4: RMSE curves of GBRT for phosphorus as a function of n_p during the training and validation. The training was performed for 1000 boosting iterations.

aforementioned targets. The values are re-scaled based on the classifier with lowest performance.

Table I: RMSE obtained with the three regression models in the static features approach.

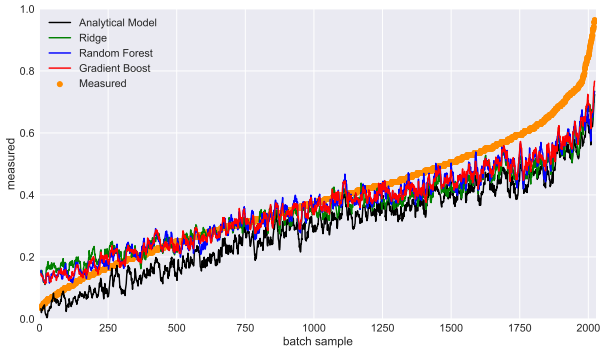
	Ridge	RF	GBRT
P	1.00	1.00	**0.92**
C	0.94	1.00	**0.89**
Mn	1.00	1.00	**0.94**
S	**0.93**	1.00	**0.93**
Temp.	**0.79**	1.00	0.90

Fig. 5 compares the prediction of the data-driven models with the analytical model for sulfur and final blow temperature obtained on the *SF* test dataset. The batch samples are sorted based on their measured target values. Considering the fact that the target values are normally distributed in their given range, it is not surprising to observe that all models perform better for the samples in the mid-range target values. Although all the models also struggle with the samples in the high-end of the range, the data-driven models seems to be more successful in predicting the low-end target values.

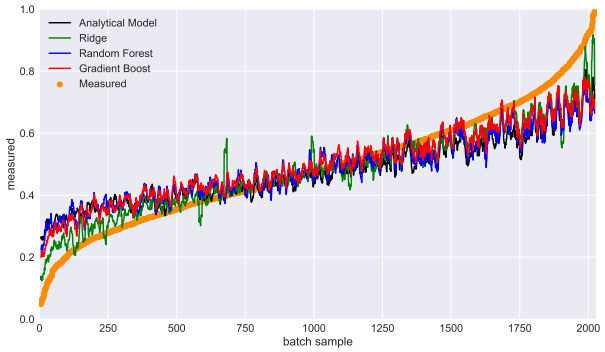
Gradient Boost is the overall winner between all models, except for the target temperature, where Ridge Regression presents a lower RMSE. The Ridge performs better at predicting lower temperatures (below 0.4) than the RF and BGRT models, as can be seen as Fig. 5(b).

Figure 6 presents the scatter distributions of predicted and measured values for the chemical endpoint composition and temperature of the molten steel, using the *SF* dataset. The employed models are the ones with lowest RMSE in Table I. The predictions present a higher variance at high target values in all cases, with P, C and Temperature having a negative bias at this region.

The Pierce correlation coefficients of 0.86, 0.85, 0.85, 0.82, 0.80 for S, P, Mn, Temperature, and C, respectively, reveal that carbon and temperature are the most difficult targets to predict.



(a) Sulfur



(b) Temperature

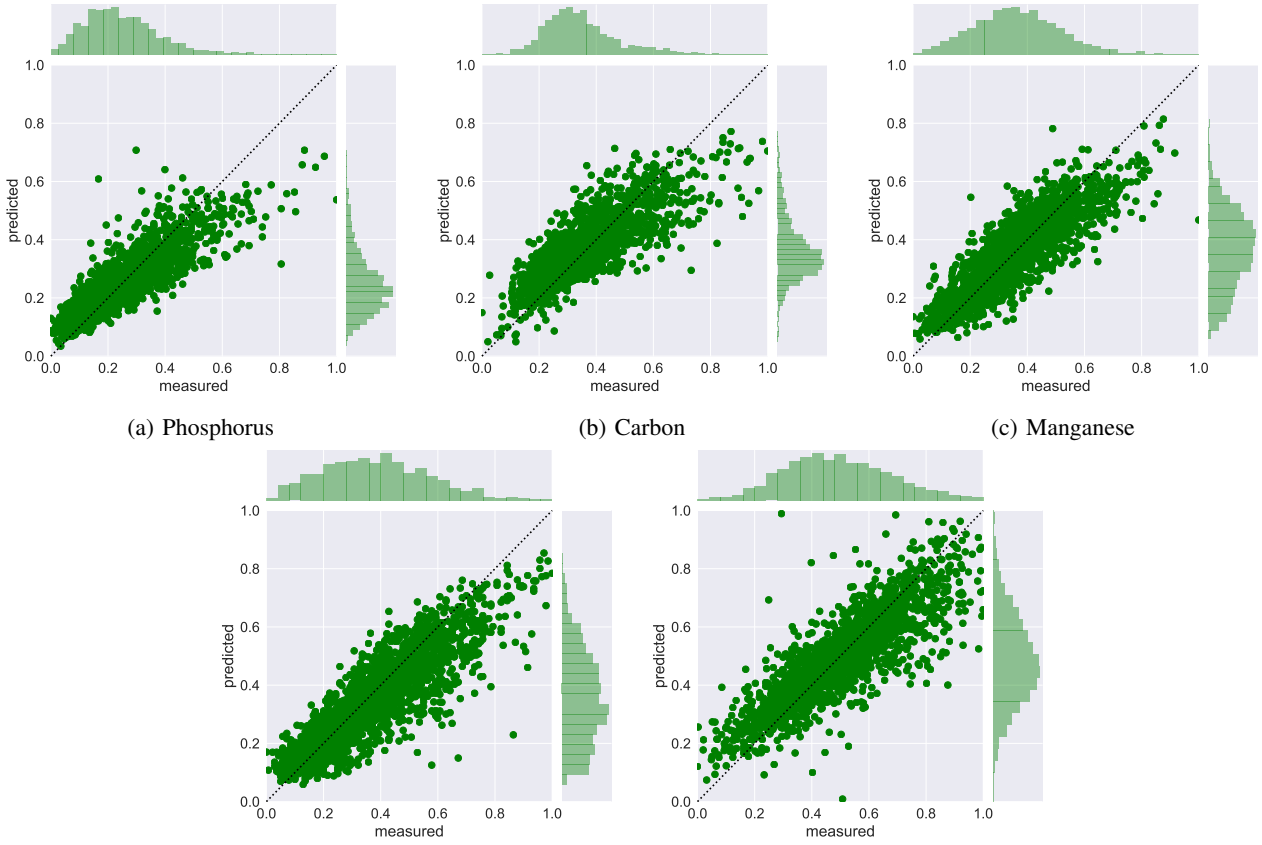
Figure 5: Prediction results for the endpoint concentration of sulfur and temperature. The orange dots are the measured values and the black, blue, green and red curves are the moving average of the results on the test set for the Analytical, Ridge Regression, Random Forest and Gradient Boost prediction models, respectively. The batch samples are ordered according to the crescent order of the test set target.

B. TS Model

The time-series data is composed of the recording from 10 sensor during the BOF blow phase. The feature extraction described in Section IV-B provides 388 features per signal per batch. Filtering the always-zero features and concatenating the remaining to the $n_p = 50$ features extracted in the previous stage, leads to a set of 3088 features per batch. A GBRT model is trained on this data and the resulting feature importance is used for feature selection.

To selected a set of n_{ts} variables, several models were trained in different configurations varying between 150 to 3088 features. Fig. 7 shows the training and validation curves for optimizing the number of features with the GBRT model on a 5-fold cross-validation training for phosphorus. Based on these results, we opted for $n_{ts} = 300$ as the configurations with more than 300 features only contribute for the models to improve the performance on training data while the validation score remains almost constant.

The Ridge, RF and BGRT approaches are then trained on the selected 300 features Late Blow dataset for all endpoint targets. The hyper-parameters were optimized using a 5-fold



(a) Phosphorus (b) Carbon (c) Manganese

(d) Sulfur (e) Temperature

Figure 6: Scatter distribution of predicted vs measured values for all the BOF endpoints, calculated by the proposed models. Graphs at the top and right sides of each image present the distribution of each variable (Predicted and Measured).

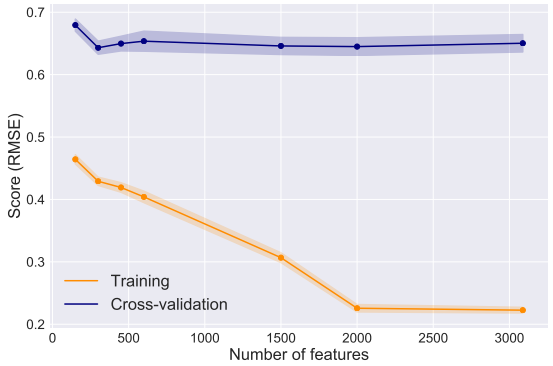


Figure 7: RMSE curves as a factor of boosting iterations and number of selected features to train a GBRT model for predicting the phosphorus composition in the time series dataset. The blue and orange lines indicate the performance on the training and validation, respectively.

cross validation on the training set.

The RMSE results obtained on the test set are listed in Table II, the *Best SF* column presents lists the best results from the *SF* models. Overall, the GBRT model performed better and there is a considerable improvement over the previous models. While the predictions of P, Mn and Temperature are relatively enhanced by more than 8%, the prediction of C and S did not show a significant gain. The lower gain on carbon prediction can be explained by the fact that the whole BOF process is practically designed based on optimizing the measure of carbon concentration.

Table II: RMSE obtained with time-series models for predicting the endpoint targets on the test set.

	Best SF	Ridge	RF	GBRT	Imp.
P	0.92	0.97	0.95	**0.83**	10%
C	0.89	0.93	1.05	**0.88**	1%
Mn	0.94	0.90	0.99	**0.83**	12%
S	0.93	1.03	1.01	**0.89**	4%
Temp.	0.79	**0.73**	0.92	0.79	8%

VI. Conclusion and Future Work

A data-driven prediction approach for basic oxygen furnace (BOF) was presented in this paper. The goal was to predict the endpoint concentrations of phosphorus, carbon, manganese and sulfur, as well as the final temperature of the liquid steel. Two sets of features were created using feature selection on available data during the process. The first set is aimed for the prediction in the late blow phase based on

the analytical features currently in use by the model-driven prediction approach, complemented by the most relevant static features from the process, selected in a data-driven procedure. The second dataset contains of features extracted from the time-series data, recorded by the sensors during the blowing process, aiming for a more accurate prediction of the steel composition at the end of the process. The improved prediction at this stage can optimize the addition of coolants and provide a better composition estimate.

Three regression models are evaluated for predicting the targets; one classical linear approach with Ridge Regression and two non-linear multivariate models based on decision trees, Random Forest and Gradient Boosted Regression Trees. The obtained results on the first set of features showed improvements over the analytical models currently used in the steel production pipeline, and further investigation will be conducted to employ them.

Employing time-series features, led to a 10% improvement for phosphorus, manganese and temperature. Which confirms the hypothesis that recorded sensor data during the blow contain important information relevant to the final prediction which can improve the current approach.

Optimizing chemical additions and blow time is crucial to enhance steel quality and guarantee the supply chain within the factory. The evidence that time-series can be used to improve endpoint prediction is valuable, as it compels the application of more complex approaches, which can further improve the prediction.

Based on the outcome of this research, the factory aims to deploy the proposed models into production. Time-series feature engineering is also an application to be discussed, as several signals have direct information related to some targets (e.g., decarbonization curves have correlation with C concentration). Another research venue to be investigated is the utilization of more advanced models such as conventional recurrent neural networks and Long-Short Term Memory networks to process the temporal information included in the recorded time-series data.

ACKNOWLEDGMENTS

The research activities described in this paper were funded by ArcelorMittal Belgium, Ghent University and imec.

REFERENCES

[1] Subagyo, G. A. Brooks, and K. S. Coley, "Interfacial area in top blown oxygen steelmaking," *Ironmaking Conference Proceedings*, no. 1, pp. 837–850, 2002.

[2] M. Han and Z. J. Cao, "An improved case-based reasoning method and its application in endpoint prediction of basic oxygen furnace," *Neurocomputing*, vol. 149, no. PC, pp. 1245–1252, 2015.

[3] J. Ruuska, A. Sorsa, J. Lilja, and K. Leiviskä, "Mass-balance Based Multivariate Modelling of Basic Oxygen Furnace Used in Steel Industry," *International Federation of Automatic Control*, vol. 50, no. 1, pp. 13 784–13 789, 2017.

[4] H. bing Wang, A. jun Xu, L. xiang Ai, and N. yuan Tian, "Prediction of Endpoint Phosphorus Content of Molten Steel in BOF Using Weighted K-Means and GMDH Neural Network," *Journal of Iron and Steel Research International*, vol. 19, no. 1, pp. 11–16, 2012.

[5] H.-x. Tian, Y.-d. Liu, K. Li, R.-r. Yang, and B. Meng, "A New AdaBoost . IR Soft Sensor Method for Robust Operation Optimization of Ladle Furnace Refining," *ISIJ International*, vol. 57, no. 5, pp. 841–850, 2017.

[6] Y. Zhang and M. S. Dudzic, "Online monitoring of steel casting processes using multivariate statistical technologies: From continuous to transitional operations," *Journal of Process Control*, vol. 16, no. 8, pp. 819–829, 2006.

[7] D. Laha, Y. Ren, and P. N. Suganthan, "Modeling of steelmaking process with effective machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 10, pp. 4687–4696, 2015.

[8] L. Ma, J. Dong, K. Peng, and K. Zhang, "A novel data-based quality-related fault diagnosis scheme for fault detection and root cause diagnosis with application to hot strip mill process," *Control Engineering Practice*, vol. 67, pp. 43–51, 2017.

[9] K. Zhang, J. Dong, and K. Peng, "A novel dynamic non-Gaussian approach for quality-related fault diagnosis with application to the hot strip mill process," *Journal of the Franklin Institute*, vol. 354, no. 2, pp. 702–721, 2017.

[10] F. Aller, L. F. Blázquez, and L. J. Miguel, "Neurofuzzy based temperature prediction of an industrial polymerization reactor in real time," in *2015 54th IEEE Conference on Decision and Control*, dec 2015, pp. 2743–2748.

[11] H. Jin, X. Chen, L. Wang, K. Yang, and L. Wu, "Dual learning-based online ensemble regression approach for adaptive soft sensor modeling of nonlinear time-varying processes," *Chemometrics and Intelligent Laboratory Systems*, vol. 151, pp. 228–244, 2016.

[12] T. A. AL-Qutami, R. Ibrahim, I. Ismail, and M. A. Ishak, "Virtual multiphase flow metering using diverse neural network ensemble and adaptive simulated annealing," *Expert Systems with Applications*, vol. 93, pp. 72–85, 2018.

[13] B. Bidar, J. Sadeghi, F. Shahraki, and M. M. Khalilipour, "Data-driven soft sensor approach for online quality prediction using state dependent parameter models," *Chemometrics and Intelligent Laboratory Systems*, vol. 162, no. September 2016, pp. 130–141, 2017.

[14] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, ser. Springer Series in Statistics. New York, NY, USA: Springer New York Inc., 2001.

[15] Z. E. Xu, K. Q. Weinberger, and A. X. Zheng, "Gradient Boosted Feature Selection," pp. 522–531, 2014.

[16] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh A Python package)," *Neurocomputing*, vol. 307, pp. 72–77, 2018.

[17] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, pp. 281–305, Feb. 2012.