ADVANCED DATA-DRIVEN METHODS FOR PREDICTING ELECTRIC STEELMAKING SLAGS

Ismael MATINO[1], Marco VANNUCCI[1], Alice PETRUCCIANI[1], Antonella ZACCARA[1], Valentina COLLA[1], Tommaso CUCINOTTA[1], Maria FERRER PRIETO[2]

[1] Scuola Superiore Sant'Anna, TeCIP Institute - ICT-COISP, 56124 Pisa, Italy
[2] Sidenor I+D, 48970 Basauri, Spain

i.matino@santannapisa.it, m.vannucci@santannapisa.it, a.petrucciani@santannapisa.it, a.zaccara@santannapisa.it, v.colla@santannapisa.it, t.cucinotta@santannapisa.it, maria.ferrer@sidenor.com

Introduction

Circular economy and industrial symbiosis are strategic pillars for European growth, economy, sustainability, and competitiveness, as emphasised in the Strategic Research Agenda of the European Steel Technology Platform (ESTEP)[1, 2] and of the Clean Steel Partnership[3]. In this background, by-products are acquiring ever more importance, and the steelmaking industry (where about 90% in mass of the by-products are slags[4]) is investigating solutions allowing their optimal management in terms of reuse and recycling[5-7]. In electric steelmaking, slags are produced in Electric Arc Furnaces (EAF) and in Ladle Furnaces (LF), and sometimes their management is not optimised since their features (*e.g.* their composition) are not continuously monitored. Knowing in real-time the slag composition with standard analytical techniques is challenging. The procedure often requires long times and presents difficulties connected to harsh environment (for liquid slag) and high heterogeneity of slags (especially at solid state), that would require costly and lengthy multiple sampling to ensure representativeness. The sample preparation, whose complexity depends on the adopted analytical method, also affects measurements accuracy, and sometimes only semi-quantitative methods are used. On the other hand, a more precise and anticipated knowledge of slags composition provides opportunities from the point of view of both process optimisation and slag valorisation. Slag conditioning and metallurgical processes control, as well as slag handling, can be improved using information on slag features, by enabling implementation of suitable practices for slag treatment, recycling and valorisation.

Therefore, besides the requirements of new sensing devices[8, 9], digital tools can be used to estimate slags composition. Different models can be found in the literature on the electric steelmaking process[10-13] but they are generally not focused only on the slags simulations and/or are too complex for real-time use in decision support systems finalised to the slag monitoring and management. For this reason, among the different

activities, during the European RFCS project “iSlag” (https://www.islag.eu/), different models were developed for both offline analyses (*e.g.* a flowsheet model considering the whole electric steelmaking route[10], and allowing computing, among others, steel and slags amounts, compositions and temperature, electric energy requirments, *etc.*) and online estimate of slag compositions. For this last goal, data-driven models were developed in their pure or hybrid configurations, by employing techniques based on Artificial Intelligence and particularly on Deep Neural Networks also combined with phyisics-based model. This paper focuses on them.

Materials and Methods

The models used for online estimate of slags composition exploit a Deep MultiLayer Perceptron (DMLP) architecture. It was trained and fed in its pure configuration with input data coming from a real industrial plant (referring to about 1500 heats), after an *ad-ho*c variable selection process.

The list of the inputs are the following:

- EAF slag models: amounts of most scrap qualities used in the charge mix, amounts of most charged and added non-metallic materials, consumed electricity, duration of activated electric arc, consumed oxygen and natural gas, tapping temperature;
- LF slag models: amounts of most additions at tapping and during the different secondary metallurgy steps.

While in hybrid form (Hybrid Net – HN) a Physics Informed Neural Network (PINN) approach[14] was followed, which is designed to incorporate knowledge of the underlying physical laws and constraints of a system into the standard neural networks. Therefore, the DMLP are fed also with results of mass-balance physics-based model obtained from the previously mentioned flowsheet model[10]. The mass-balances equations concern the physics/chemical-based calculation of the different slag compounds content, and have been *ad-hoc* tuned/adjusted with specific parameters derived from the flowsheet model. Therefore, these equations can be considered a sort of approximation and translation of the flowsheet model. The equations refer to the following aspects considered in the flowsheet model: direct transfer of a specific compound from the raw materials to the slags without transformation, reactions of some raw materials compounds to become components of the slags, partition of some compounds between slags and liquid steel, slag entrainment, differences in process phenomena and reactions depending on produced steel family (*i.e.* group of similar steel grades). In addition, they include some correction factors to counteract errors deriving from data (*e.g.* related to scraps variability, analytical methods and calibration of measurement devices). The

inputs of the mass-balance equations are amounts and compositions of all the input materials (*i.e.* scraps, metallic and non-metallic materials).

Further approaches were tested especially for LF slags. This is because more uncertainties derived from data (*e.g.* the LF slag amount is not available for each heat as in the case of EAF slag) and from higher complexity of this process (due to more additions and steps with respect to EAF) are transferred to the flowsheet model and, consequently, to the derived mass balance equations that are less accurate. Therefore, a Limited Hybrid-Net (LHN) was tested with the exploitation only mass-balances of the physics-based model having higher performances, related to the following LF slag compounds: SiO_2, Al_2O_3 and TiO_2. Furthermore, training was done with and without inclusion of the steel family information (models named with the suffix SF), and considering a higher or a reduced number of variables (suffix RV is used in this last case for identifying the models). Although these approaches were developed especially for LF slags, for the sake of scientific completeness, they were tested also for EAF slags.

The inputs in the case of reduced number of variables are the following:

- EAF slag models: amounts of reduced number of scraps qualities used in the charge mix, amounts of reduced number of the charged and added non-metallic materials, and consumed natural gas;
- LF slag models: amounts of reduced number of additions at tapping and during the different secondary metallurgy steps.

Results and Discussion

This section presents and discusses the results obtained by using the proposed techniques. Results in terms of Mean Absolute Error (MAE) related to the best model approach (*i.e.* the one with lower MAE) for the estimate of the mass fraction of EAF and LF slags compounds are reported in Table 1. This includes also the MAE obtained by the flowsheet model (MAE FM) to show the improvements obtained via data-driven approaches. MAE refers to the test carried out on 30% of the global samples (randomly selected). It is noteworthy that the performance of each data-driven model is affected by the combination of all inputs. Their capability to “capture” the unknown synergy among inputs enables better performances compared to the flowsheet model.

MAE is normally used to compare the models, but is difficult to interpret as it is on absolute basis. Therefore, Table 2 shows the Mean Absolute Percentage Error (MAPE) for both best data-driven approach and flowsheet model (MAPE FM), and the Percentage Variation Coefficient (PVC), which is the percentage ratio of the standard

deviation over the mean value, to better show the models' accuracy. As MAE, also MAPE and PVC refer to the test dataset.

Table 1: Test MAE of best data-driven model for the estimate of slags compounds mass fraction compared with flowsheet model MAE

EAF Slag	Best Model	MAE	MAE FM	LF Slag	Best Model	MAE	MAE FM
SiO_2	HN_RV_SF	$1.28 \cdot 10^{-2}$	$2.17 \cdot 10^{-2}$	SiO_2	DMLP	$1.89 \cdot 10^{-2}$	$1.94 \cdot 10^{-2}$
FeO	HN_RV_SF	$4.94 \cdot 10^{-2}$	$6.12 \cdot 10^{-2}$	FeO	HN_RV_SF	$1.34 \cdot 10^{-3}$	$1.45 \cdot 10^{-3}$
Al_2O_3	HN_RV_SF	$1.09 \cdot 10^{-2}$	$1.42 \cdot 10^{-2}$	Al_2O_3	HN	$1.23 \cdot 10^{-2}$	$1.23 \cdot 10^{-2}$
CaO	HN_RV_SF	$3.11 \cdot 10^{-2}$	$3.51 \cdot 10^{-2}$	CaO	DMLP_SF	$2.00 \cdot 10^{-2}$	$2.08 \cdot 10^{-2}$
MgO	HN	$6.27 \cdot 10^{-3}$	$1.10 \cdot 10^{-2}$	MgO	DMLP_SF	$1.56 \cdot 10^{-2}$	$1.63 \cdot 10^{-2}$
MnO	HN	$8.98 \cdot 10^{-3}$	$8.98 \cdot 10^{-3}$	MnO	DMLP_RV_SF	$3.29 \cdot 10^{-4}$	$3.88 \cdot 10^{-4}$
Cr_2O_3	DMLP	$4.69 \cdot 10^{-3}$	$9.52 \cdot 10^{-3}$	Cr_2O_3	HN_RV_SF	$8.07 \cdot 10^{-5}$	$1.07 \cdot 10^{-4}$

Table 2: Test MAPE of best data-driven model for the estimate of slags compounds mass fraction compared with flowsheet model test MAPE and test dataset PVC

EAF Slag	Best Model	MAPE	MAPE FM	PVC	LF Slag	Best Model	MAPE	MAPE FM	PVC
SiO_2	HN_RV_SF	16.5 %	26.7 %	23.2 %	SiO_2	DMLP	8.7%	11.9 %	12.0 %
FeO	HN_RV_SF	12.7 %	14.7 %	16.7 %	FeO	HN_RV_SF	48.6 %	83.9 %	52.5 %
Al_2O_3	HN_RV_SF	24.1 %	31.7 %	31.8 %	Al_2O_3	HN	22.1 %	38.1 %	33.5 %
CaO	HN_RV_SF	17.0 %	17.2 %	28.3 %	CaO	DMLP_SF	3.5%	5.1%	4.7%
MgO	HN	8.5%	14.7 %	13.8 %	MgO	DMLP_SF	30.3 %	31.6 %	49.4 %
MnO	HN	11.6 %	12.4 %	17.5 %	MnO	DMLP_RV_SF	63.1 %	84.5 %	97.8 %
Cr_2O_3	DMLP	15.3 %	18.2 %	30.7 %	Cr_2O_3	HN_RV_SF	28.2 %	83.0 %	29.8 %

Minor compounds that together constitute on average less than 5 wt% of slags (*i.e.* TiO_2, P_2O_5, V_2O_5 and Na_2O) are not reported in the tables as the models do not accurately estimate them for both the slags. Among these minor compounds, only for the following acceptable results were obtained: P_2O_5 in EAF slag (MAPE of 22.0% with DMLP_RV_SF), V_2O_5 in EAF slag (MAPE of 28.0% with HN) and TiO_2 in LF slag (MAPE of 19.9% with DMLP_RV_SF).

The example results shown in Figure 1 compare the values of the measured mass fractions of some slag compounds (on the X axes) to the estimates provided by the best

modelling approach (on the Y axes). For the reported compounds (*i.e.* CaO and FeO for EAF slags, and SiO_2 and Al_2O_3 for LF slags), the estimates provided by the best data-driven approach are comparable with measured values, such as shown by the fact that the points clouds are close to the bisector line.

As shown in Table 1 and 2, data-driven approaches always provide better performances compared to the flowsheet model; and, generally, HN (in one of the tested configurations) provides best results for most EAF slag compounds, demonstrating that physics-based model introduces significant information to the data-driven system improving its performances. On the other hand, usually DMLP (in one of the tested configurations) provides the best results for the estimate of most compounds of LF slag. This demonstrates that, in this case, the mass-balance physics-based model introduces uncertainties in HN and decreases its accuracy.

Observing the MAPEs in Table 3, the best modelling approach can characterise most major compounds of EAF slags with MAPE values lower or equal to 17% (Al_2O_3 is the only exception). For LF slags, the models are less performant; actually, most compounds that together represent in average more than 95% of LF slags (*i.e.* SiO_2, Al_2O_3, CaO and MgO) are estimated with MAPE values lower than 23% with the only exception of MgO. Moreover, as expected, MAPE follows the behaviour of the dataset PVC: higher error values correspond to variables showing higher PVC, highlighting that the models performances can be considered suitable enough.

In effects, model performances are affected by data quality. For instance, although inputs consider real mixtures of different scraps, only average composition of each scrap type were available, thus their variability through time and the related effects on slags cannot be considered. Furthermore, although outliers were removed, some intrinsic errors remained in the provided data, and these affected the models training. Some of them can be related to errors introduced by operators, while other ones are connected to the measurement devices calibration or the analytical method. Furthermore, in case of slag analyses, errors can be related to the fact that the samples are not completely representative of the slag due to its heterogeneity. Therefore, it is expected that better performances can be achieved when at least some of the above listed issues are solved.

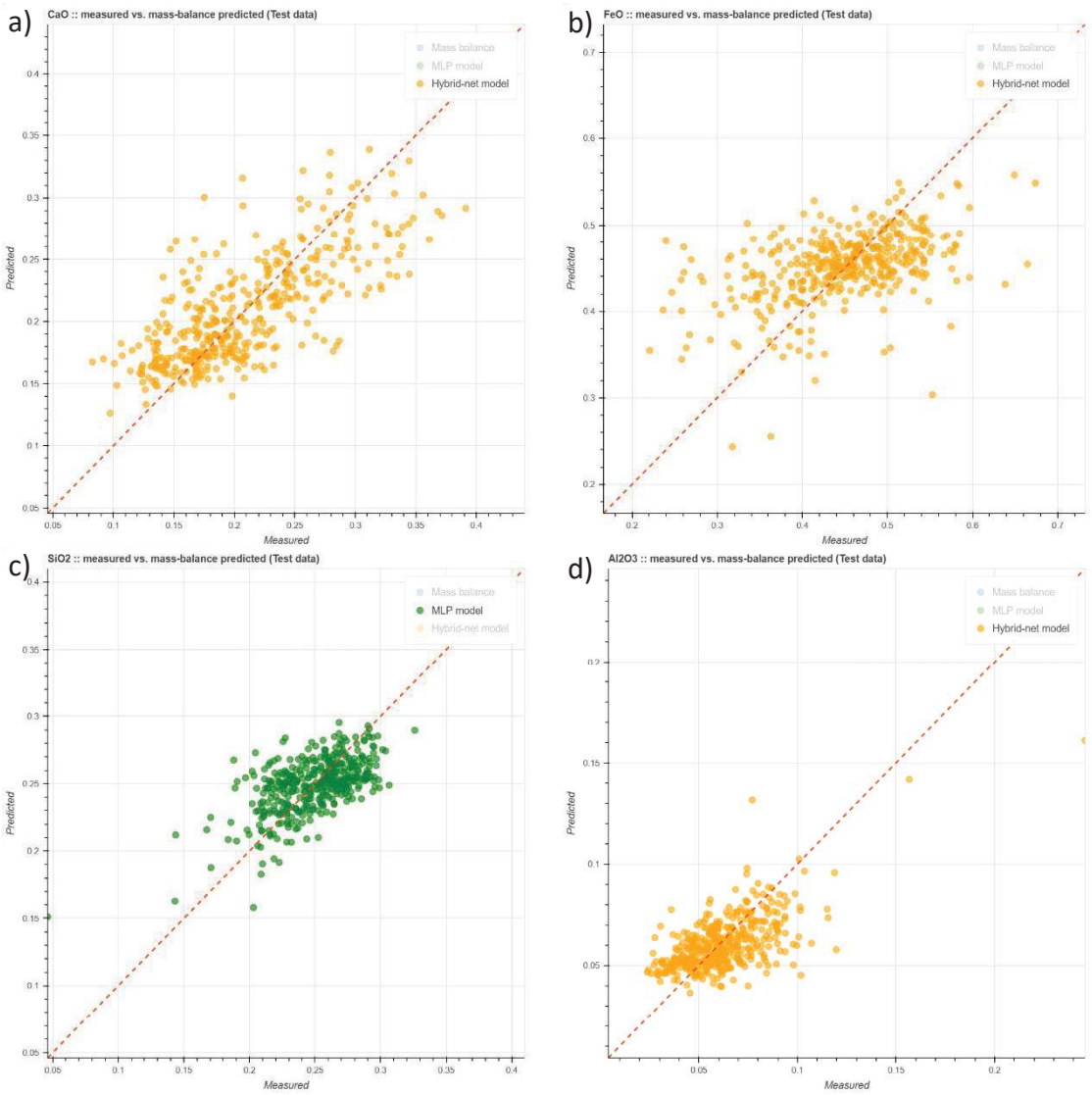

Figure 1: Example of comparison of measured mass fractions with predicted values with the best approach of: a) CaO content in EAF slag by using HN_RV_SF; b) FeO content in EAF slag by using HN_RV_SF; c) SiO_2 content in LF slag by using DMLP; d) Al_2O_3 in LF slag by using an HN

Conclusions and Future Works

Different data-driven modelling approaches were tested for the prediction of EAF and LF slags composition. Hybrid-AI approaches proved to be suitable to increase data-driven model performances, if uncertainties deriving from unreliable data are not transferred to physical-based models. The proposed digital solutions can be valuable tools for real-time characterisation of electric steelmaking slags to improve their management and valorisation, as they can provide continuous and fast estimate of slag composition with respect to the discontinuous analyses that are currently done.

In order to improve the models' accuracies, ongoing and future work focuses on better understanding the phenomena linked with the compounds that are estimated with higher errors (*e.g.* minor compounds) to provide more phenomenological-based information to the HN models. In addition, solutions for obtaining more reliable data on both inputs (*e.g.* scraps) and outputs (*e.g.* slags) are under investigations.

Finally, the assessment of model robustness will be carried out with "out-of-time" dataset, referring to heats produced in a period well after the ones considered in the training set.

Acknowledgements

The work described in the present paper was developed within the project entitled "Optimising slag reuse and recycling in electric steelmaking at optimum metallurgical performance through on-line characterisation devices and intelligent decision support system - iSlag" (G.A. 899164) co-funded by the Research Fund for Coal and Steel of the European Union, which is gratefully acknowledged. The sole responsibility of the issues treated in the present paper lies with the authors; the Union is not responsible for any use that may be made of the information contained therein.

References

1. ESTEP, Strategic Research Agenda (SRA), available online: https://www.estep.eu/assets/Uploads/SRA-Update-2017Final.pdf
2. K. Peters, E. Malfa and V. Colla, "The European steel technology platform's strategic research agenda: A further step for the steel as backbone of EU resource and energy intense industry sustainability", *Metall. Ital*, **111**(5), 5-17 (2019).
3. ESTEP, Clean Steel Partnership, available online: https://www.estep.eu/clean-steel-partnership
4. worldsteel association, Fact sheet - Steel industry co-products, available online: https://worldsteel.org/wp-content/uploads/Fact-sheet-Steel-industry-co-products.pdf
5. L. Kieush, J. Rieger, R. Attrotto, A. Sorino, W. van der Stricht, H. Oterdoom, E. P. Heikkinen, G. Dall'Osto G., C. Mapelli, D. Mombelli, L. Di Sante, F. Cirilli, V. Colla, T. A. Branca, I. Matino, A. Petrucciani, A. Zaccara, C. Brondi, E. Mousa, E. Nylund, E. Sandberg, M. Guzzon, E. Malfa, A. Schröder and I. Bellemans, "Roadmap for recycling practices and resource utilization in the iron and steelmaking industry: a case studies", *Matériaux & Techniques*, **112**(5), 503 (2024).
6. J. Rieger, V. Colla, I. Matino, T. A. Branca, G. Stubbe, A. Panizza, C. Brondi, M. Falsafi, J. Hage, X. Wang, B. Voraberger, T. Fenzi, V. Masaguer, E. Faraci, L. di Sante, F. Cirilli, F. Loose, C. Thaler, A. Soto, P. Fritella and D. Snaet, "Residue valorization in the iron and steel industries: Sustainable solutions for a cleaner and more competitive future europe", *Metals*, **11**(8), 1202 (2021).
7. T. A. Branca, V. Colla, D. Algermissen, H. Granbom, U. Martini, A. Morillon, R. Pietruck and S. Rosendahl, "Reuse and recycling of by-products in the steel sector: Recent achievements paving the way to circular economy and industrial symbiosis in Europe", *Metals* **10**(3), 345 (2020).
8. J. Petersson, M. Gilbert-Gatty and A. Bengtson, "Rapid chemical analysis of steel slag by laser-induced breakdown spectroscopy for near-the-line applications", *Journal of Analytical Atomic Spectrometry*, **35**(9), 1848-1858 (2020).