DEVELOPMENT OF DECISION SUPPORT SYSTEM BASED ON MACHINE LEARNING AND DIGITAL TWIN MODELS FOR ALUMINIUM MELTING FURNACES

Zeynep Tutku Özen[a] , Hasan Basri Taşkın[a] , İlyas Artunç Sarı[a] , Başar Arkan[a] ,Görkem Özçelik[a] , Andrea Fernández Martínez[b], Santiago Muiños Landín[b] , Matteo Pozzi[c] , Angelo Gordini[c] , Luca Ferrari[c] , Jaroslaw Hercog[d], Aleksandra Kiedrzyńska[d], Beata Glot[d]
[a] ASAŞ Alüminyum San. Tic.A.Ş., İstanbul, TURKİYE , tutku.ozen@asastr.com
[b] Technological Centre AIMEN, Porrino - Pontevedra, Spain
[c] OPTIT Srl, Bologna, Italy
[d] Institute of Power Engineering-Research Institute, Warsaw, Poland

Abstract

The sustainability vision that gained importance with the European Green Deal, has also affected the Aluminum sector, which is considered as one of the most energy-intensive industries in Europe. Starting from this vision, the European Union called for projects, leveraging the Horizon 2020 program, with the concept of using existing raw materials more efficiently and enabling the use of alternative raw materials within various industrial sectors. RETROFEED is one of the projects supported by European commission under Horizon 2020 program. Among the solutions under development for the aluminum sector there are: (1) the usage of alternative raw materials in ASAS' aluminum melting furnaces, (2) equipment retrofitting allowing the existing resources to be processed more energy-efficiently, and (3) the design of decision support strategies in order to use the existing raw materials for producing less waste material during production.

Strategical and operational decisions are made by the factory personnel, at times causing slowdown in production and/or inefficient use of resources. Within the scope of the project, a Decision Support System (DSS) will be developed integrating machine learning methods, for predicting billet quality according to different raw material types used in ASAŞ Aluminum Melting Furnaces, and a furnace Digital Twin to simulate furnace operations under different conditions for efficiency improvements and production simulation. By means of correlations and algorithms established as a result of the study, the output of the billet chemical composition, the furnace setpoints and the production plan can be adjusted according to different inputs in 6060 aluminum alloy, in accordance with the principles of zero waste and the new green deal guidelines.

Keywords: Decision Support System, Machine Learning, Digital Twin, Aluminium Industry, Automation

1. Introduction

The aluminium billet casting process start with the loading of the raw material into the melting furnace. In ASAS, the total charge consists aluminium ingot material (70-80%) and scrap (20-30%). The amount of external scrap is around 8%. The furnace operates at a temperature of 950ºC and a pressure of 0-15 bar. The temperature of the molten aluminium is 720ºC. The liquid aluminium is checked after 30 minutes for alloying. If necessary, some alloy elements are fed into the furnace, including Si, Mg, C, Fe, Cr. The alloy content is usually around 4% of the total charge, and it is programmed depending on the final products requirements. This iteratvie process ir run until the proper composition for the aluminium billet is achieved. The liquid aluminium is ready to cast roughly after 1 hour melting in furnace.

The molten aluminium obtained must comply with standard alloy composition in order to be suitable for casting production. In this sense, the variability of the feedstock particularly affects energy efficiency, time duration to detect proper composition in spectral analysis and emission rate. Therefore, increasing the amount of contaminated scrap that can be fed into the melting furnace would augment the time needed for aluminium billet production, therefore reducing the productivity of the proces. In this regard, if the use of contaminated scrap increased to 50-60% in ASAS facilities, this implies an increased in yearly production costs of around 15M€, due to the increased energy consumption and the loss of productivity in the process. ASAS has a 90.000 ton/year casting capacity with use of low-cost raw material after R&D work and improvement studies. If this amount of primary aluminium could be replaced, ASAS would make a profit approximately 0.5M euro in a year, since scrap aluminium is 5% cheaper than primary aluminium.

Equipments retrofitting requires a proper settlement in order to avoid large industry's process to shut down for a long period of time. In order to realize these achievements, implementation of a Furnace system forxygen level measurement and control is one of the key elements of higher efficiency melting process with higher scrap rates and better environmental solutions. The Installation of a O_2 injection system will also enable to control the oxygen level inside the furnace chamber better. A delacquering burner mode will enable contaminated scrap which generally contains 2-5% of contamination such as paint, lacquers, oil and various plastics to be used as a raw material. Implementation of a chimney TOC analysis system which will be integrated on the furnace exhaust system, control and analyse the chimney gas organic compounds concentration. Control system modifications are targeted to automate the furnace for better flexibility. A delacquering drum furnace will be integrated on the existing furnace for better efficiency on delacquering process.

2. Decision Support System

A Decision Support System (DSS) is a tool used to support teams, companies and individuals in dealing with complex processes which decisions require the analysis of multiple variables, constraints and objectives like the ones explained in the introductory chapter. A modern DSS, described in Figure 1, is built upon a complex IT architecture which requires different tasks such as models' execution, with different models' types, data management, user interaction and a Graphical User Interface (GUI) to interact with data and model results. RETROFEED's objective is to build a level 2 smart system that could be used, next to the company's MES or DCS platforms, for performing advanced analysis and support operators and managers during the decision-making process [1]. ASAS is currently developing solutions to reduce the overall environmental impact of its operations through a series of changes that will impact both processes and operations. Therefore, a DSS is becoming necessary to reduce both the impact of the changes made and to improve overall operations.

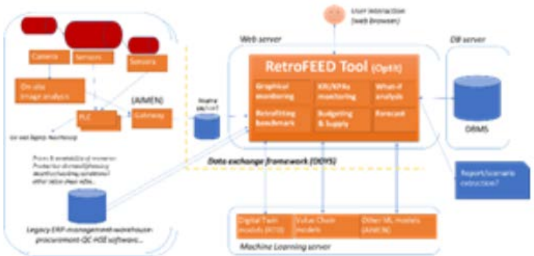

Figure 1. IT Architecture of Decision Support System

The models developed for the tool are divided in 3 main areas: (1) Machine Learning models for process validation and control, (2) Digital Twin Simulators build with CFD and (3) Input Optimization Models. The models could be run both in simulation mode, evaluating beforehand the impact of certain decisions, or as day-by-day support tools with certain analysis and calculations scheduled to run automatically to predict future process output. A model for validating the results and to monitor current process is being developed by the Spanish Technology center AIMEN using various machine learning techniques to predict how input materials could influence the product quality. A digital twin, developed by the Polish Institute of Power Engineering (IEN), will be used to simulate furnace's operations giving the user the possibility to run various scenarios changing a series of key inputs which could impact both environmental footprint and economics. Feedstock materials for the process could be optimized using an optimization model developed by Optit which takes into account variables like scrap availability in the warehouse, costs of feedstock and chemical composition in order to create the optimal mix to produce the product selected. The 3 models could be run, through the DSS, as standalone application or integrated together creating a digital value chain which could help the users take decision from input mix to output quality going through the process detailed evaluation. In the next paragraphs each model will be described in detail, also showing initial results based on offline data shared by ASAS.

2.1. Machine Learning

The billet chemical composition of the aluminium process is essential to guarantee high-quality products that comply with current standards, while also ensuring an optimal use of resources and the overall process efficiency. The billet quality depends mostly on the type of materials used during the aluminium process, which can come from different sources and whose chemical composition might slightly vary over time. These two factors challenge the possibility to have detailed estimations of the chemical composition of billets, which is of especial importance to avoid rare out-of-specifications products as well as to gain insight knowledge about the influence of the different feedstock used in the final product.

In recent years, Machine Learning (ML) algorithms have proved to be successful analyzing high-dimensional data showing complex, non-linear relationships, which would otherwise be infeasible using conventional statistical approaches or simulations due to their incapability to deal with missing information and uncertainty. Supervised Learning (SL) is a sub-field of ML in which the desired target is known and used during the training procedure in order to find the relations between inputs and outputs, in our case, between the feedstock and chemical composition, respectively. There is a wide range of ML algorithms for SL in manufacturing processes, ranging from Ridge Regression, Support Vector Machines (SVMs), and Random Forests (RFs), to the most advanced Artificial Neural Networks (ANNs)[2]. Ridge Regression is a multi-linear-based regression algorithm widely used to estimate regression coefficients under the assumption of multicollinear (highly-correlated) independent variables. The algorithm was originally developed to solve the imprecision of Ordinary Least Square estimators of Linear Regression models and is especially interesting due to its simplicity and results understanding.
Although mostly used in classification problems, SVMs can also be extended to solve regression problems with the so-called Support Vector Regressors (SVRs). This algorithm aims to find the best hyperplane (decision boundary) to fit the data, going beyond linear-based models by introducing the kernel trick to create high-dimensional feature spaces.
RFRs are ensemble algorithms that build upon, and extend the capabilities, of Decision Trees. By introducing the processes of bagging and feature randomness, RFs can outperform conventional models in a simply, hierarchical-based approach.
Finally, and initially inspired by biological neural networks, ANNs are based on a sequence of connected layers, each of them consisting in several units known as neurons, that analyses the data in a feedforward approach exploiting its backpropagation self-adjusting capability. Despite its well performance, one of the major drawbacks of ANNs is its intrinsic "black-box" behavior.

2.2. Digital Twin

A digital twin is basically a virtual model of a product, process or service. It is the transfer of physically existing situations to digital media. The digital twin can be generally referred to as a technology solution. With the transfer of processes to digital media, decision-making mechanisms and follow-up mechanisms will work actively. In connection with this, more precise decisions can be made. The fact that the control and follow-up will be done with artificial intelligence will directly contribute to the efficiency. Thanks to the digital twin, the life cycle of the processes will be kept under full control, and the analysis and improvement of the parts where human access is difficult will be carried out much easier [3].

The digital twin can be specified as the head runner of industry 4.0 and digitization provides simple access between virtual machines, data and physical environments [4]. The data collected in processes digitized by the digital twin will be able to easily solve many problems by bringing them together and making sense of them. It will be very easy and useful to make benchmark studies with the collected data. At this point, ensuring data security becomes also critical.

The aim of the digital twin for core process is to deeply characterise the behavior of this equipment and their interaction with upstream and downstream processes, as well as to analyse the impact of several retrofitting scenarios and the modification of the process conditions without affecting the product quality.

To ensure a modular approach whereby statistical/machine learning modules and the processes' digital twins can be leveraged upon in a modular and integrated way.

Digital Twin results is a Reduced Order Model (ROM) of the melting furnace. ROMs are simplifications of high-fidelity, complex models that capture the behavior of full models so that the system can be effectively analysed using minimal computational resources.

2.3. Additional Models

ASAS's raw material management process is currently not automatized. The DSS is expected to support such plant operation by keeping track of the raw material flow, taking into account feedstock levels, depending on data availability, as well as economic aspects, so as to select the most convenient mixture of raw materials to be fed to the melting furnace to achieve the proper product quality. The model is written in Python code and run through and open-source solver.

Such decision should not compromise the energy efficiency of the whole process, nor increase the emission of pollutants. ASAS objective to increase the amount of scrap used as raw material should not determine a decrease in quality of the molten aluminum, which is assessed by means of laboratory analyses carried out on samples taken at the exit of the melting furnace, which are used to measure product chemical composition. The model's output is the Optimal raw material mix given the above list of inputs and parameters. It could be run for a single batch or to reflect a production plan formed by multiple batches scheduled over a certain period of time (e.g., optimization of weekly feedstock level based on production plan). As described in previous chapters, the model could be run as standalone application or be integrated with the digital twin to evaluate not only the optimal feedstock mix but also simulate the process with such mix of materials evaluating possible changes to support and improve production.

3. Results

3.1. Preliminary Results of Machine Learning

Considering ASAS 6060 aluminium alloys, seven different scrap types and three ingots that are fed to the furnace for alloying were analyzed with the aim of predicting the final chemical composition of billets in terms of seven chemical elements, namely Si, Fe, Cu, Mg, Zn, Ti, and Cr, having as reference EN573-3 standards. The dataset comprised 729 casting processes from January to November 2020. For modelling purposes, 80% of the data was used for training (583 castings) and 20% for testing (146 castings). All data was previously pre-process and scaled within their feature range using min-max normalization. The models explored include Ridge Regression, RFR, SVR and a ANN. A hyper-parametrization tuning process with a cross-validated (5) Grid-Search approach was followed in order to find the best parameters for fitting the data for the Ridge, RFR and SVRs models, which were implemented using the open-source implementations available in Scikit-Learn library. For the ANN, the Tensorflow and Keras libraries were used.

Hyper-parametrization tuning

For the Ridge Regression model, a regularization strength of alpha:0.5 was found to give the best results. The SVRs implementations, namely SVR1 and SVR2, use a Radial Basis Function and a Polynomial Function as kernel, respectively. Due to the intrinsic single-output behaviour of these models, parameters C:1 and epsilon:0.1 were chosen as default. Regarding the RFR, a maximum depth:4 and number of estimators:100 were chosen.

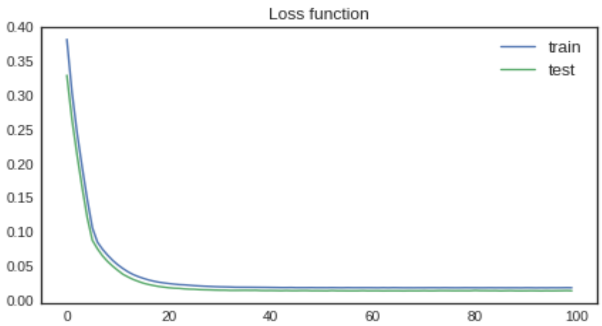

Figure 2. Loss Function in the training set (blue) and test set (green) during ANN training

Due to the limited amount of data for the development of an ANN, a 1-deep layer network with 8 neurons was implemented, with an output layer with 7 neurons, each for the prediction of the 7 target chemical elements. The loss function used was Mean Absolute Error, with an Adam optimizer. The model was trained for 100 epochs. The loss curve in the training and test sets is shown in Figure 2.

Model performance

The Mean Absolute Error (MAE) of each of the models for the prediction of Si, Fe, Cu, Mg, Zn, Ti, and Cr, elements to assess the billet quality is shown in Table 1. The performance of the ANN, RFR, and Ridge

(Regression), is generally comparable regarding the MAE values obtained. There is one element, Cu, which seems to be challenging for prediction, only the ANN has a MAE:0.007%, followed by the RFR with MAE:0.029% and Ridge with MAE:0.092%, significantly greater than the latter.

Table 1. Mean Absolute Error (MAE) comparison in the prediction of Si, Fe, Cu, Mg, Zn, Ti, and Cr, considering the ANN, Ridge Regression, RFR and SVR1 and SVR2 models.

	ANN	Ridge	RFR	SVR1	SVR2
Si	0.029	0.029	0.030	0.042	0.042
Fe	0.019	0.019	0.020	0.023	0.024
Cu	0.007	0.092	0.029	1.075	1.276
Mg	0.027	0.028	0.029	0.031	0.033
Zn	0.005	0.005	0.005	0.030	0.029
Ti	0.003	0.003	0.003	0.037	0.037
Cr	0.003	0.003	0.003	0.005	0.005

To deepen in the behaviour of the three best-performance models in the prediction of Cu, that is, ANN, RFR and Ridge, Figure 2 shows the ground truth test set (green) and the predicted sets by the ANN (blue), RFR (yellow) and Ridge (orange) models within the range of true values to facilitate visualization

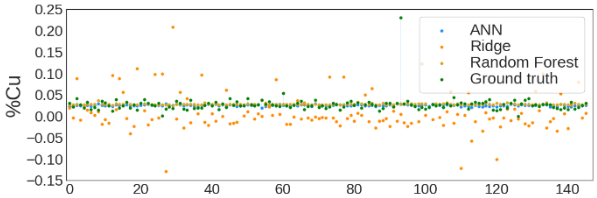

Figure 2. Comparison of test set ground truth values (green) and predicted sets using an ANN (blue), Ridge Regression (orange), and Random Forest (yellow) models within the range of ground truth values for the prediction of Cu (%).

Althought the Ridge model has a MAE comparable with RFR, a detailed study of the results show that the RFR predicted 2 extreme values (not shown for visualization purposes) that cause its MAE to increase, despite modelling the remaining 144 samples of the test set quiet well. From Figure 2, it can be seen that Ridge Regression predictions are dispersed and sometimes even negative in value, which does not fit our our context. On the other hand, the ANN proved successfully predicting the test samples with the lowest MAE, which results in some predictions being hindered from the figure due to its precise location on the truth value. Despite the good performance of the ANN in the prediction of Cu, it is still challenging to predict those values which are outliers from the distribution (i.e. %Cu>0.003%), being these values of special interest to forecast those aluminium casting processes in which the billet quality might be compromised.

3.2. Preliminary Results of Digital Twin

In order to create digital twin model ANSYS Twin Builder v2021R1 were used. In the digital twin; power in fuel, air temperature, air to gas ratio, total mass of feedstock, ratio of scraps in charge, ratio of contaminations and flow of O_2 are used as inputs. Digital twin model uses correlations regarding ROM (Reduced Order Model) of furnace with burners, Model of melting Al and Model of dross formation. As a result, digital twin system iterates outlet gasses, average and maximum temperature at the furnace, temperature at the roof of the furnace, average outlet temperature, average wall temperature for each furnace wall, NOx at outlet from furnace, fraction of molten Al and total mass of dross. Figure 3 demonstrates the relations of digital twin between inputs and outputs.

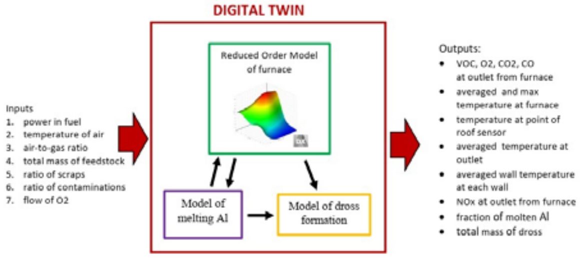

Figure 3. Digital Twin Model Development Strategy

4. Discussions

4.1. Machine Learning

ML has shown promising results in finding and exploiting complex relations from high-dimensional data spaces, outperforming advanced statistical approaches such as Ridge Regression. Table 1 reveals interesting results when it comes to the prediction of the chemical composition of billets based on the feedstock. For the Si, Fe, Mg, Zn, Ti, and Cr elements, the MAE obtained by the different models suggests a comparable performance between the ANN, Ridge and RFR models. This behaviour could be explained by a linear-based relationship between the inputs and outputs, what blocks even the most advanced ML models (ANN, RFR) tto perform better than Ridge Regression usually is; a potential incapability of the most advanced models to outperform the Ridge Regression due to insufficiently informative data from the feedstock; or a hindered bad performance of the Ridge Regression due to potential outliers that can cause the MAE values to be comparable in value. Focusing on the seeming most complex case, that is the prediction of Cu shown in Figure 2, the outperforming capabilities of the ANN and RFR models with respect to Ridge seems to be confirmed, having these former negative predictions for %Cu, which is impossible by the physical nature of the problem. Moreover, in the case of the SVR models, it can be concluded from the results in Table 1 that these models are not suitable for the problem in question since they seem to be unable to find a suitable hyperplane for the prediction of the chemical composition of billets, with the exception of Fe, Mg and Cr, in which the MAE can be comparable to its counterparts. Thus, in view of the results, the ANN seems to be the most suitable model for the prediction of Si, Fe, Cu, Mg, Zn, Ti and Cr elements, owing the lowest MAE in all cases with a robust behaviour avoiding extreme or out-of-range values as in the case of the RFR and Ridge models. Nevertheless, despite these results, more data would be needed to better evaluate and potentially improve the ANN, especially in cases where the real values do not follow the distribution which are in turn risky to the overall process and (billet) product quality.

4.2. Digital Twin

Developed Digital Twin was divided into 3 parts to take into account the different time scales of the processes observed during Al melting inside of furnace. The combustion process of natural gas which generated heat for melting takes below 1 second, while the melting process lasts about 4 hours. Detailed simulations of the

whole process using CFD (Computational Fluid Dynamics) method is extremely computationally consuming, therefore the model consists of:
1) ROM (Reduced Order Model) of furnace with burners
2) Model of melting Al
3) Model of dross formation
The ROM was build based on series of detailed CFD simulations of the furnace including the processes of heat and mass transfer inside the furnace together with mixing and reactions during combustion processes. The ROM model was developed using ANSYS Twin Builder v2021R1 software. The melting process is calculated using 0D model (assuming uniform temperature distribution). The data about heat exchange between flue gas and melting Al are passing from ROM to melting model returning values of heat adsorbed during melting process and VOC (Volatile Organic Compounds) released by contaminated scraps. The module of dross formation calculates the amount of dross based on temperature and oxygen concentration at the aluminium surface and scrap fraction in the feedstock.
The details about the inputs and outputs to and from the Digital Twin are listed below.
Inputs:
1. power in fuel dosed to burners
2. temperature of air dosed to burners (after heating)
3. air-to-gas ratio
4. total mass of feedstock
5. ratio of scraps (all types) to total mass of feedstock
6. ratio of contaminations
7. flow of 100% O2 dosed by O2 lance
Outputs:
1. VOC at outlet from furnace
2. O_2 at outlet from furnace
3. CO at outlet from furnace
4. CO_2 at outlet from furnace
5. average temperature at furnace
6. maximum temperature at furnace
7. temperature at point of roof sensor
8. average temperature at outlet from furnace o
9. averaged wall temperature at back wall
10. averaged wall temperature at bottom wall
11. averaged wall temperature at door wall
12. averaged wall temperature at left wall
13. averaged wall temperature at right wall
14. averaged wall temperature at top wall
15. NOx at outlet from furnace
16. fraction of molten Al
17. total mass of dross

5. Conclusion

The DSS under development will include all the m odels described in the article, integrated together to replicate as closely as possible the real process in a digital platform. The Digital twin will be the center of the analysis providing simulation capabilities of t he real furnace in order to test beforehand prossibl e possible process modifications or improvements without having to run physical and costly tests. The Machine Learning model will be integrated with the DT output evaluating billet composition of simulated batches supporting process' quality valuation.
At the same time, the model will be used also for day-by-day processes using fresh data, downloaded from ASAS' DCS system, to predict billet's quality of real batches in quasi realtime. Raw real-time will give the operator the opportunity to act on these predictions. For example, the raw material selection model may be used to optimize both digital twin and real process inputs taking into account not only physical and chemical aspects but also economics and process constraints such as availability of raw materials and environmetal environmental limits. The platform will be flexible enough to run different models as standalone appli cation or as an integrated service, opening the door to the integration of even more models in the future. All together, the digital platform of the process will be useful for ASAS to manage increasing process complexities, like the increase of scrap in the melting mix, and enhance productivity while reducing its environmental impact.
At the end of the RETROFEED project, these developments will be validated industrially by ASAS' furnaces and control-monitoring systems. Business exploitation plan is also studied within the scope of this project and applicability of the novel system will be investigated and validated also upon European Aluminium producer industries and other sectors which also could benefit from having an integrated modelling and digital platform.

6. Acknowledgement

This work was performed in the framework of the RETROFEED Project (Implementation of a Smart RETROfitting Framework in the Process Industry towards its Operation with Variable, Biobased and Circular FEEDstock). This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N.869939. It is a pleasure to thank allRETROFEED Project partners for their intensive support to this study.

7. References

[1] Ferrari L., Gordini A., Pozzi M. ,D4.1 - Benchmarking of available retrofitting support tools, https://retrofeed.eu/public-deliverables/
[2] Fahle, S., Prinz, C., Kuhlenkötter, B. (2020) Systematic review on machine learning (ML) methods for manufacturing processes – Identifying artificial intelligence (AI) methods for field application, Procedia CIRP, 93, 413-418. https://doi.org/10.1016/j.procir.2020.04.109.
[3] Leng, J., Wang, D., Shen, W., Li, X., Liu, Q., & Chen, X. (2021). Digital Twins-based Smart Manufacturing System Design in industry 4.0: A Review. Journal of Manufacturing Systems, 60, 119–137. https://doi.org/10.1016/j.jmsy.2021.05.011
[4] Onile, A. E., Machlev, R., Petlenkov, E., Levron, Y., & Belikov, J. (2021). Uses of the Digital Twins concept for Energy Se rvices, intelligent recommendation systems, and Dema nd Side Management: A Review. Energy Reports, 7, 997–1015. https://doi.org/10.1016/j.egyr.2021.01.090