European Commission

Research Programme of the Research Fund for Coal and Steel

Technical Group: TGA5

# Smart consideration of actual ladle status monitored by novel sensors for secondary metallurgy process parameters and ladle maintenance strategies

## SmartLadle

**Public**

Asier Arteaga, Diana Mier Vasallo
Sidenor Investigacion Y Desarrollosa SA (SID),
Basauri, Spain

Grant Agreement Number:101034017

01.07.2021 – 30.06.2024

## Deliverable 3.3 – Set-up of ML models

## Due 03 / 2022
## Lead beneficiary: SID

## Table of contents

## Project summary

What is the effect of the actual ladle status -new to worn- on steel bath properties? How do e.g. temperature or fluid flow vary with ladle conditions? When is the optimal moment for relining?

SmartLadle will provide a solution for online monitoring and dynamic incorporation of actual ladle status for process control. A soft sensor for ladle status shall be developed, supported by a smart sensor for detecting refractory wear and thermal status. Measurement data, models and advisory tools shall provide information for decision making to operators to adapt ladle metallurgy process parameters to actual ladle status and decide about maintenance actions.

### Definition of terms used in the project

Soft sensor: Mathematical calculation of value of a process parameter that is difficult or so far impossible to measure directly and online, based on other process values, measurements, models and smart sensor data.

Smart sensor: Combination of a pure sensor for the acquisition of a measured value, e.g. refractory temperature, and a small computing unit with implemented simplified models, e.g. for refractory wear.

ML model: Data-driven model that analysis data and detects relationships (linear or non-linear) among variables based on real-world data using Machine Learning (ML) Techniques.

## 1. Introduction

The deliverable D3.3 describes the work of task 3.3 that aims the preparation of the ML modelling approach for the rest of the project in Sidenor use case. The work is an initial version and will be further improved and completed throughout the project. It is based on previous works ([P2-4]) but intends to go one step beyond by the use of more available data and improved modelling approaches. The data to be used are described in D1.1 and include different sources and data granulometry, thus, data treatment is an important and time-consuming job before modelling itself. In fact, some of the models are providing elaborated data to others, doing a data preparation and transformation work.

**Figure 1** shows the models developed and the connections among them. The two main models are the refractory wear model and the liquid steel temperature model; each of them with their own data sources. There are four other models that exert a role of helper models for the main models. The refractory wear model is at the same time a helper model for the liquid steel temperature model and the refractory thermal model too. The main idea of the helper models is that they provide "elaborated" data for the next models.

Data sources are displayed as well in the figure, they provide the basis for the model development although in some cases the raw data need different treatment for the different models. The tundish data are distinguished in the figure as they are going to be obtained in the sensing tasks and, at this stage, are quite unripe.

Next chapters describe each of those models with more detail.

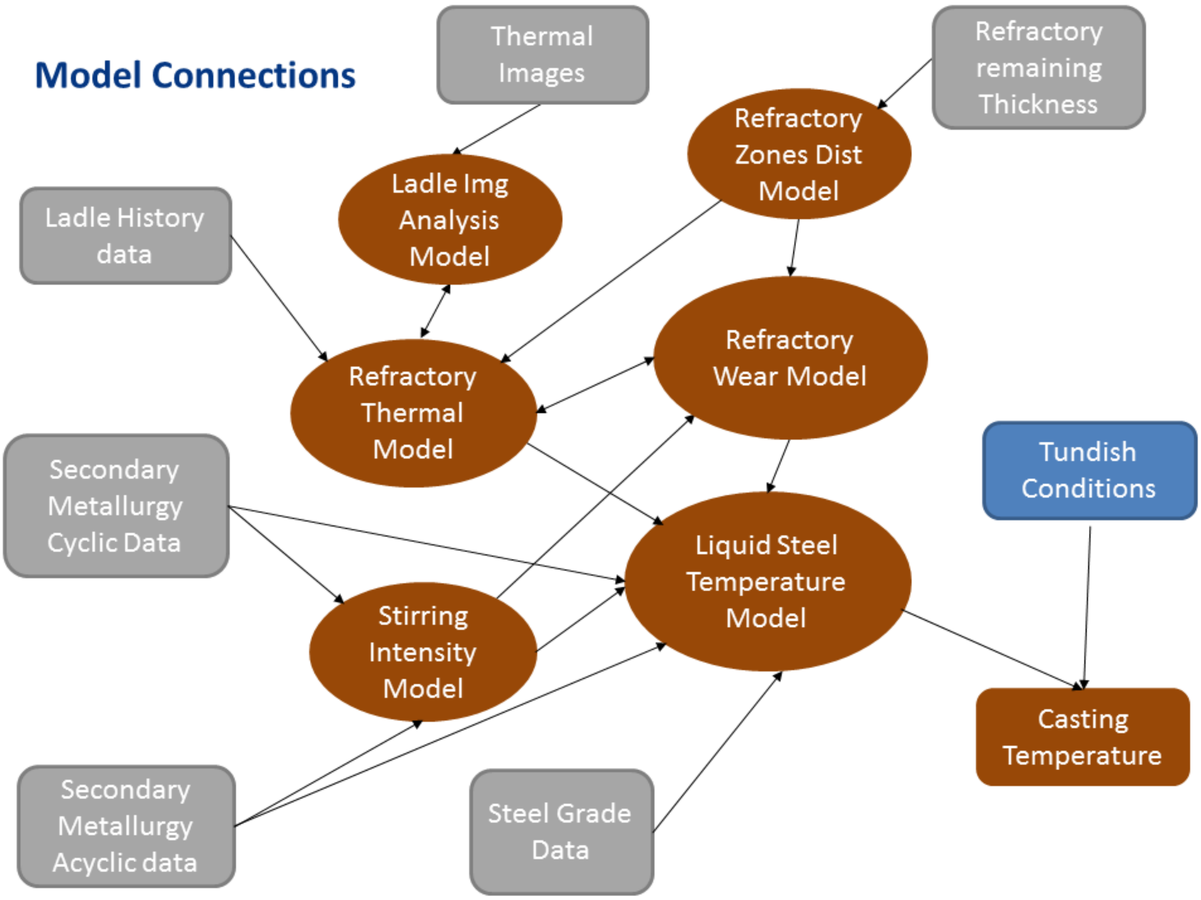


**Figure 1:** Schema of model connections

One general aspect is that the ML techniques or approaches will be used in many of them. In the supervised models that means that the data will be divided in training and testing sets; an error metric will be decided in function of the data; several models will be tried and based on the testing error the best one chosen; influence of the input variables will be estimated to check the variables themselves and the model too. This approach is not completely new for some of those models, but regarding the development of the techniques and the systematization of them, a good advance is expected.

## 2. Refractory Zones Distribution Model

**Concept**: The refractory wear rate is different from zone to zone of the ladle. From previous projects experience and practical knowledge this is a clear fact. The main distinction made in P3, for example, is slag line and steel part (**Figure 2**). Nevertheless, the intention in this project is to go beyond this approach and obtain numerically the zones from the wear data. This is a good opportunity to better understand wear data and at the same time provide a better output number for wear models.

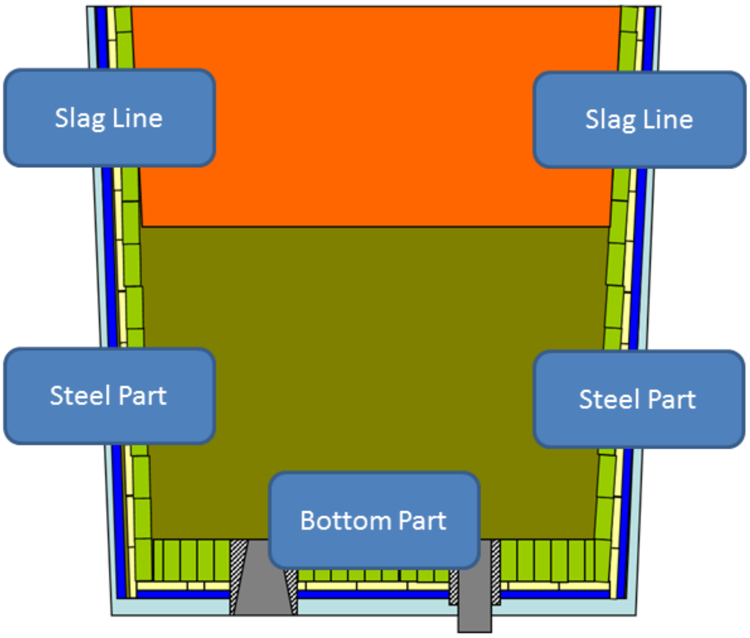


**Figure 2:** Simple Refractory Zones distribution

**Data**: Data collected from demolished ladles in cold conditions allow the measurement of residual length of the bricks. Joining these residual lengths with the initial length of the bricks and the number of heats the ladle has been working, it can be obtained the average wear rate per brick. It is important to note that only a representative brick per row has been measured and the only measurement is post-mortem in this case. So it refers to the remaining size after (residual length) after ladle reparation or demolition. This data requires a heavy pretreatment as they come from manual measurements and annotations, particularities in the cycles,…This aspect has been one of the most time consuming from the initial work, but, at the same time, it is the basis for the use of the richness they include. The calculation itself of the wear rate requires consideration as if the data come from partial demolition or final demolition or if they are in final demolition which bricks were changed in partial demolition. This data richness is an improvement from previous works as they were based on averaged values or just significant ones.

**Model Approach**: The main idea of the model is to use a clustering algorithm to differentiate the zones by the measured wear rate, it is an unsupervised approach. The first version tried was using a K-Means algorithm for each of the ladle sides (they are different due to the wear produced by the stirring gas) and the result for both of them was that 3 clusters classify well the wear rate data (**Figure 3**). 3 zones for each of the two sides of the ladle. The wear rates are well distinguished by those groups giving three distinct populations. Nevertheless, this is an initial analysis that will be further improved and enriched with new data.

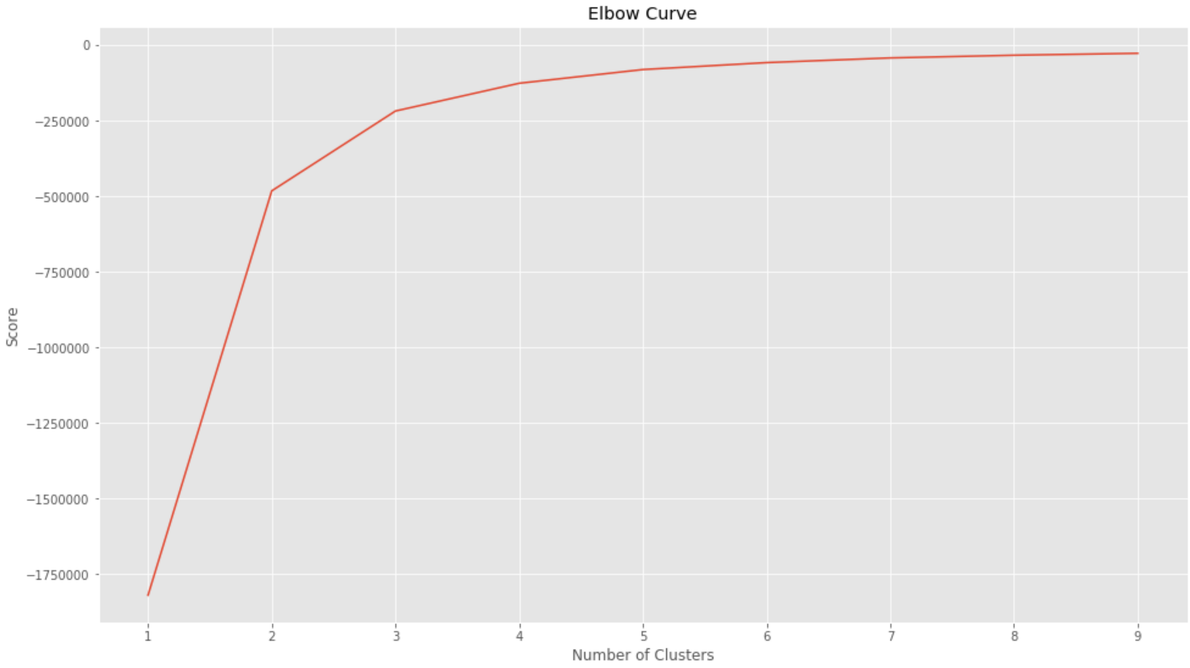


**Figure 3:** Elbow curve to check number of clusters for one of the ladle sides, the results for the other side are very similar regarding number of clusters for each ladle side

## 3. <u>Ladle Thermal Images Analysis Model</u>

**Concept**: Thermal images have been acquired since some years for the steelmaking ladles. For example, in P2 several were used to assess thermal state of the ladles and check a thermal model. For this project the same approach is expected but adding some innovations. The main concept to try is to use modern neural network image analysis to check ladle numbers and automatize ladle detection in the image. The output has to be treated to be able to check the output of refractory thermal model.

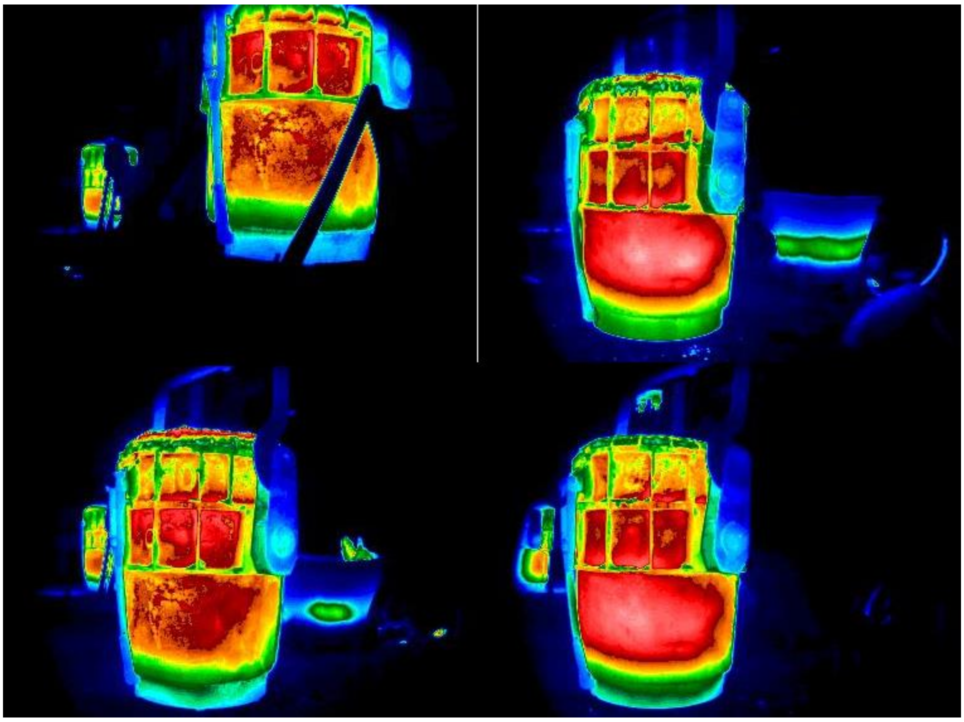

**Figure 4:** Examples of recent ladle thermographs

**Data**: Some thousands of new thermal images has been collected during 2022 to use current conditions and ladle status (**Figure 4**). Those images are the starting point for the analysis. The first activity was to adapt the camera acquisition procedure to increment the number of images acquired. After acquisition, a first data cleaning (elimination of low quality images, or images difficult to treat) has been the next step to leave less images, but better, to treat and classify. In the first stage of the analysis many images have been manually classified (ladle number) to have a number of annotated data. In some cases the annotation is not possible as ladle number is not visible in the image. In others it is clearly seen, as it is displayed in the images top right and bottom left in Figure 4.

**Model Approach**: The main concept is to use pretrained CNN models and retrain them with annotated images in order to obtain a reliable model that can detect the ladle in the image and "read" the ladle number.

## 4. Refractory Thermal Model

**Concept**: The refractory thermal model was elaborated and updated in projects like P2 and P3. It was a Finite Different model in 1D. A much better modelling and checking has been done in other projects, using more elaborated models and precise thermocouple measurements, but real-life conditions require simplification: for one part, calculation time has to be fast to cope with the long lives of the ladles; in the other hand, the uncertainties in the industrial conditions are much greater and a rough calculation with better data can be more reliable than a great calculation with wrong data. In any case, this is a helper model in this project, designed to provide data to the wear model and liquid steel temperature model. In the case of liquid steel temperature model the implication is clear; the total amount of heat accumulated in the refractory is a key factor for the temperature loses of the liquid. Regarding the wear model, strong thermal stress estimations are sought as a cause of refractory wear. However, this is a bilateral relationship, as remaining size of the refractory is one of the important inputs for the thermal state of the ladle.

**Data**: Ladle History data are the key ones to assess the thermal state. The times for each heat in which the ladle has been with liquid steel, in burners, empty or covered determine how much heat has entered in the refractory and how much has been lost. Refractory configuration is also a key factor, as different configurations leave to different heat capabilities. Data coming from wear calculation are also needed.

**Model Approach**: The former running model is the basis for the calculation. But two developments are ongoing: one is to get an appropriate output for the liquid steel temperature model and wear model, the other one is to check the option of substituting the finite element model with a data based model. This possibility can be interesting if a low error is obtained and the data based model is faster to run. For this purpose, a series of models runs will be done to create a training data set, some others for a testing data set and the errors of the data models will be calculated over them.

## 5. Stirring Intensity Model

**Concept**: The stirring rate is at the same time an important aspect of the secondary metallurgy and a value difficult to measure. In several projects (P4, P7, P9 for example) it has been addressed and different techniques have been developed to measure or estimate it better. The basic idea would be that a well measured flow rate of the stirring gas would be good enough to measure the stirring power but, this is not reliable in industrial conditions. Among the developments to measure it better are open-eye images, vibration measurements, electric arc stability and the classical flow rate vs counter pressure checking. This project is not devoted at Sidenor to deepen in the stirring measurement but to use some of the gained knowledge to provide more reliable stirring data to the other models. Concretely, it is reasonable to expect that different stirring conditions will affect both: the thermal behaviour of liquid steel and the wear of the refractory. With the data and knowledge acquired in previous projects a stirring "value" is being elaborated to feed the rest of the models.

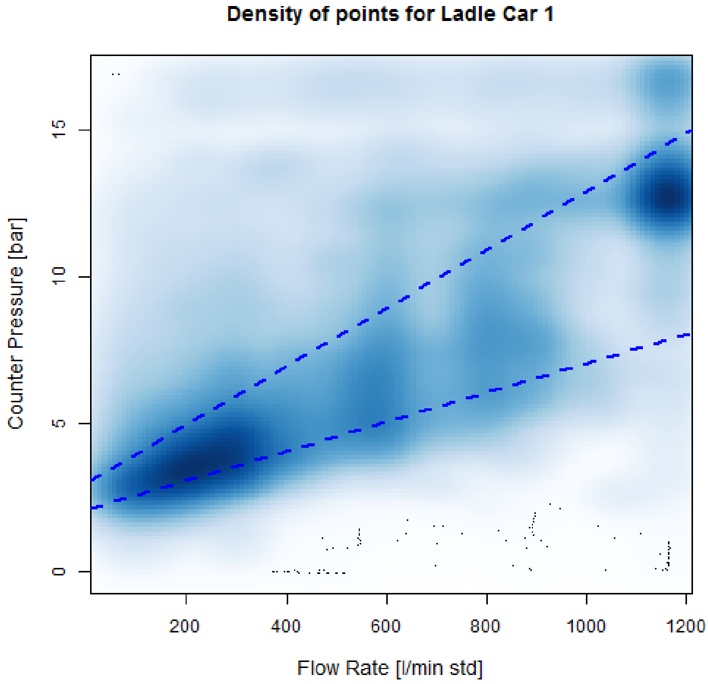


**Figure 5:** Smooth Scatter plot of the flow rate vs pressure in data including first 3 months from 2022.

**Data**: The data to be used come mainly from cyclic secondary metallurgy data collected in the two ladle cars. They have to be prepared and filtered and also enriched with refractory data (porous plug type). The raw data consist of flow rate, counter pressure, vacuum pressure, vibration power, and valve state. In **Figure 5** a good example of those data is displayed for the initial months of year 2022.

**Model Approach**: In previous projects a "corrected flow rate" gas elaborated, using mainly porous plus measured flow rate vs pressure data. The measured cyclic data were interpreted using a linear formula. In this development a similar but more elaborated and updated approach is being followed: This time the production cyclic data are the main input. They need to be correctly filtered and interpreted to feed a linear model. The greater amount of data and the incorporation of vibration data and more precise ML error measurements are being used to check an updated and more reliable "corrected flow rate". This amount has to be easy to calculate as the value to be used in the other models has to be simple enough and possible to calculate between concrete timestamps or in a complete heat, and in different conditions (vacuum, LF).

## 6. Refractory Wear Model

**Concept**: As worked in previous project (P3), refractory wear rate is valuable by itself to optimize ladle performances, but at the same time it is a value needed for ladle thermal calculation, and, this way, liquid steel temperature model. The main idea behind this kind of model is that the wear of the refractory is not uniform in all the heats and significant differences can be found depending on several factors.

**Data**: This model follows a supervised model approach. One variable has to be explained (wear rate) depending on many others. And this is, at the same time, the difficulty and the opportunity of this model version. The difficulty is that counting only aggregated data at the end of ladle life, the effect of the many explaining variables can be disguised, the opportunity is to try to improve the feature engineering with respect previous experiences and the help of the helper models. Concretely, the Y variable will be calculated using the aggregation of data from the ladle zones model; in each zone an average or median wear rate will be tried. In the other hand, refractory thermal model and stirring data model will provide additional data to enrich and improve the wear rate data explanations divided by zones. Apart from those models' inputs more standard inputs will be used as steel compositions, time with liquid steel in ladle, vacuum time, time in burners, slag formers,...

**Model Approach**: The refractory wear rate is different from zone to zone of the ladle as defined in the refractories zones definition model. In consequence, two approaches are possible: one model for each zone or a general model in which categorical variables distinguish the zones. In a linear model, the multiple model approach is almost needed to cope with possible strong influence differences from factors, for example, slag formers can be critical in slag line but less important in steel parts. However, more complex models can cope better with those differences. Both approaches will be tried.

The model is still not in the first stage as a minimum of one year of data are desired for this model as the ladle campaigns are not so many and the first months of the project have still been difficult for production regularity due to several external factors (Covid, energy costs, supply chain difficulties). These aspects affect all the models but this one especially due to the lower amount of data, as they are aggregated.

## 7. Tundish Conditions

Liquid steel temperature evolution during secondary metallurgy is closely connected with the ladle. But once casting conditions are considered another vessel has to be considered: The tundish (**Figure 6** shows a schematics). In P2 it was treated in a very simple way, so to improve it, this is a key aspect of Sidenor approach to SmartLadle, but it will be developed in the activities related to sensors in WP2 and WP4. From those measurements, there will be deduced if specific modelling is needed to understand and preview the thermal evolution of the steel temperature in the tundish and how to cope with it in the Soft Sensor and advisory tool.

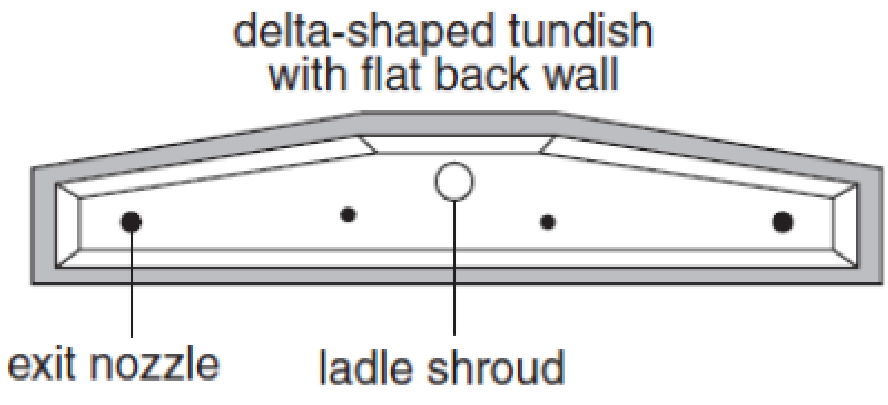


**Figure 6:** Tundish Schematics

In any case, the last step from the ladle to the tundish always includes the ladle effect too, as the liquid steel stays some minutes there. In consequence, both aspects have to be considered to make a good prediction of the casting temperature.

# 8. Liquid Steel Temperature Model

**Concept**: During the secondary metallurgy process, liquid steel temperature is maintained within a reasonable range considering the metallurgical processes. Nonetheless, there are many factors affecting it as melting of ferroalloys and slag formers, heating to compensate loses, thermal loses to the refractory or vacuum degassing. In P2 data acquisition was performed and a linear model obtained to be able to calculate the thermal behaviour, optimize it, and, finally, be precise in the most important temperature loss: The one from secondary metallurgy to casting, as casting temperature is a critical parameter for production and quality. This approach is not a novelty, but in this project the tundish part will be much better characterized and modelled as mentioned in Chapter 7 and the secondary metallurgy part will be improved with richer data and modelling capabilities.

**Data**: The data have been collected from liquid steel temperature measurement to the next one. In each of those time periods the influential events have been collected (vacuum, additions, arc heating,...) but also significant variables calculated as thermal state or stirring power. But it is important to note that this is not for the whole heat, it is for each of the periods between measurements. The last period is from the last secondary metallurgy measurement and the continuous casting. Those data are not obtained only from process data, they require the helper model outputs.

**Model Approach**: The initial data collection has been completed as many data are got for this model; some data per heat allows values in the thousands easily. The “holes” due to lack of helper models (wear rate model) have been completed with reasonable data (for example uniform wear per heat). This is not considered a big problem in this initial setup, as this is a first step that will be improved with the measurements in the tundish and other developments throughout the project.

The error type used is Root Mean Squared Error (RMSE) as it is usual in continuous output variables. Several models have been compared: Linear Model, Random Forest, XGBoost, LightGBM, K Nearest Neighbours, Support Vector Machines and a Neural Network. The initial best model has been XGBoost and the training error (RMSE) close to 8 ºC in training set and 15ºC in testing set. Those values suppose a $R^2$ superior to 80%.

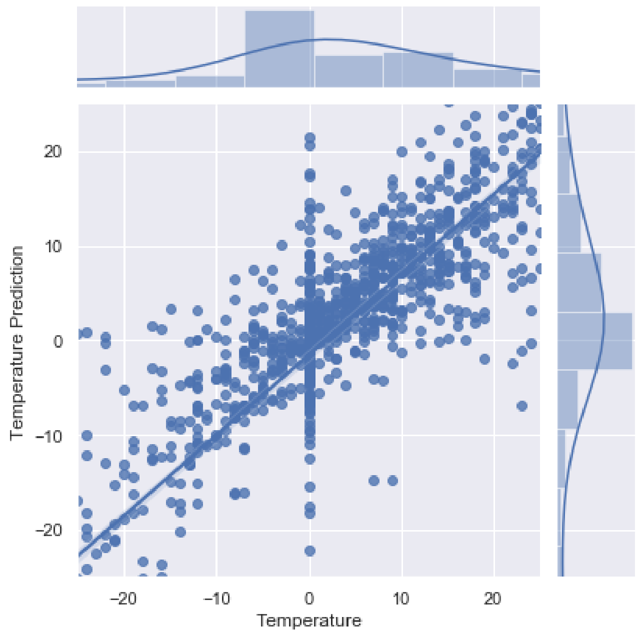


**Figure 7:** Prediction vs measurement in the modelled population

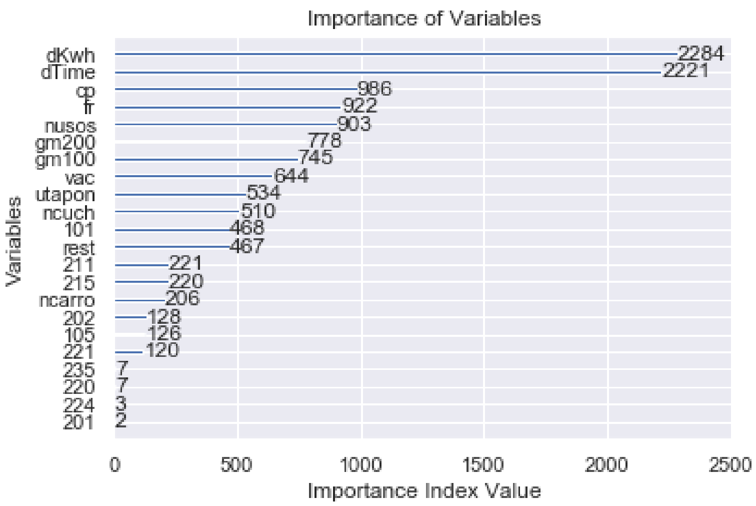


**Figure 8:** Importance of the input variables in the LightGBM model.

**Figure 7** and **Figure 8** show two results from two of the models regarding the predicted vs measured values and the ranking of variable importance. None of those results is especially critical as the main objective of this stage is to establish the founding of the models and data pipelines; nonetheless, the first results show promising perspectives.

## 9. First conclusions

This deliverable explains the modelling setup developed in Task 3.3. As it can be seen in Figure 1, the data collected in WP1 and described in D1.1 feed many different models, connected among them. Although many models share the ML approaches and techniques, they are quite different in objectives, complexity and expectations. In fact, the complete task, and the project work by extension, is an actualization of the developments in previous projects and the more systematic implementation of ML concepts updated with last years' improvements in the area.

The structure of the models, inputs, outputs and approach have been defined but the degree of maturity of model implementation is different in each case, depending on the data collection and dependency with other tasks.

However, there are some initial results available that show the potential of the models and the interest of the approach.

## 10. Future Work

As explained in the model descriptions and conclusions many of the models are still unfinished and some lack data or time to collect the data. In fact, this deliverable is a roadmap for the rest of the projects than a report of results.

It will also be discussed by the project partners, if and how some of the approaches can also be used for other use cases besides the one of Sidenor.

**Project References from the Deliverable 7.1 (comprehensive overview)**

*Projects of general nature*

[P1] **Valorisation and dissemination of technologies for measurement, modelling, and control in secondary metallurgy, DISSTEC** RFCS no.709740 (2016) (2016-2017)

*Modelling of temperature and stress in refractory material, liquid steel temperature and fluid flow in steelmaking ladle, and sensors at ladle*

[P2] **Improving steelmaking processes by enhancing thermal state ladle management, LADTHERM** RFSR-CT-2014-00006 (2014-2017)

[P3] **Enhanced steel ladle life by improving the resistance of lining to thermal, thermomechanical and thermochemical alteration, LADLIFE** RFSR-CT-2009-00003 (2009–2012)

[P4] **Stirring plug monitoring system for improvement of plug availability and stirring performance, PLUGWATCH** RFSR-CT-2012-00005 (2012-2015)

[P5] **Multi-criteria through-process optimisation of liquid steelmaking, TOTOPTLIS** RFSR-CT-2010-00003 (2010-2013)

[P6] **Continuous performance monitoring and calibration of model and control functions for liquid steelmaking processes, PERMONLIST** RFCS no.709620 (2016) (2016-2019)

[P7] **Dynamic stirring for improvement of energy efficiency in secondary steelmaking, DYNSTIR** RFSR-CT-2015-00004 (2015-2018)

[P8] **Improved control of inclusion chemistry and steel cleanness in the ladle furnace** ECSC-STEEL C7210-PR/331 (2002-2005)

[P9] **Development of advanced methods for the control of ladle stirring process** ECSC-STEELC 7210-PR/330 (2002-2005)

[P10] **Production of EAF steels with low content of N2 and S through vacuum treatment** ECSC-STEEL C7210-PR/135 (1999-2002)

[P11] **Control of inclusion, slag foaming and temperature in vacuum degassing** ECSC-STEELC 7210-PR/079 (1998-2001)

*Industry4.0*

[P12] **Development of a new automation and information paradigm for integrated intelligent manufacturing in steel industry based on holonic agent technology, I2MSTEEL** RFSR-CT-2012-00038 (2012-2015)

[P13] **Consistent ladle tracking for optimisation of steel plant logistics and product quality, TrackOpt** RFCS no.753592 (2017) (2018-2021)

[P14] **Virtual design of cyber-physical production optimization systems for long production factories, CYBER-POS** RFCS no.709669 (2016) (2016-2019)

[P15] **Predictive sensor data mining for product quality improvement, PRESED** RFSR-CT-2014-00031(2014-2017)

*Power supply*

[P16] **Power generation from hot waste gases using thermoelectrics, POWGETEG** RFSR-CT-2015-00028 (2015-2018)