Data Science Applied to Anode Baking Furnace Environmental Footprint Reduction

Olivier Forato[1], François Ordronneau[2], Hervé Pedroli[3], David Brismalein[4], Gregor Jossé[5], Eloise Sorin[6], Maxime Redstone Leclerc[6], Julian Cooper[6], Christophe Bouché[7], Nicolas Fiot[8] and Louna Alsouki[9]
1. Technical & Business Improvement Director
2. Carbon Department Manager
3. Carbon Process Superintendent
4. Carbon Process Technician
Aluminium Dunkerque, Loon-Plage, France
5. Principal Data Scientist
6. Senior Data Scientist
Boston Consulting Group (BCG), Paris, France
7. Product Director
8. Firing Control System Product Manager
9. Data Scientist
Fives Solios, Givors, France
Corresponding author: christophe.bouche@fivesgroup.com
https://doi.org/10.71659/icsoba2024-el009

Abstract

Gas consumption and related CO_2 emissions become a critical challenge for all aluminium smelters, particularly in Europe. Aluminium Dunkerque, Boston Consulting Group, and Fives Solios have jointly embarked on a project to reduce gas consumption in the anode baking furnace by employing data-driven techniques. Beyond the objective of reducing gas consumption, this project was an opportunity to exploit production data to improve an industrial process. An approach that has been successful in other industries, but which is a premiere for the anode baking furnace.

The project cleaned and consolidated historical data from various sources to derive interrelations across the entire furnace, across time, and across adjacent departments (e.g., Gas treatment center, Electrolysis). In close collaboration with process experts, predictive models were developed which link the actionable baking furnace process parameters with the key process outcomes of gas consumption and anode quality, under consideration of other causal factors such as refractory state. The models were developed and tested on historical data during the Proof-of-concept phase to evaluate and confirm the potential benefits. During the subsequent Minimum-viable-product phase, for which dedicated user interfaces are co-designed with the end users, the models are evaluated in an online fashion to demonstrate gains under operating conditions.

This paper presents the project methodology, the data analysis done and the first results on ABF gas consumption and CO_2 emission reduction.

Keywords: Data science, Anode baking furnace, Gas consumption.

1. Introduction

1.1 Aluminium Dunkerque Smelter

Aluminium Dunkerque (ADK) specializes in the manufacture of rolling slabs and foundry ingots in a wide variety of alloys for high value-added applications in the automotive, transport and packaging sectors. Built in 1991, it is one of the largest primary aluminium smelter in Europe with more than 300 kt Al/y. It operates 264 AP technology pots running at 390 kA and powered by low carbon nuclear power plant.

Aluminum is a strategic metal for the environmental transition, and ADK is one of the world leaders in the production of low-carbon aluminum. The company has reduced its emissions (scope 1 & 2) by 17 % since 2013 and emits four times less greenhouse gases than the global average for the sector. On the strength of these assets, ADK intends to play a major role in the European production of low-carbon aluminum for the benefit of the customers and communities. This is why, in line with the objectives of COP21, ADK is accelerating its energy and environmental transition with an ambitious low carbon trajectory in three phases.

By 2025, efforts to optimize operations and increase recycled aluminium use aim to cut emissions by 5 %. **The Anode Baking Furnace (ABF) project will help, as it uses 50 % of the site's gas**. By 2030, ADK targets a 30 % emission reduction with increased energy flexibility and carbon capture technology. By 2050, ADK aims to reduce emissions by over 70 % by deploying a new smelting line with inert, non-carbon-emitting anodes and further increasing recycling activities.

1.2 Why Data Driven Techniques Applied to ABF?

ADK, Boston Consulting Group, and Fives Solios jointly embarked on this project, each party having already experience and expertise in either smelter operation or data-driven techniques.

Over the years ADK has developed, at smelter level, data driven practices to support operation performance with the deployment of a plant-wide data historian and data visualization tools based, in particular but not only, on PI platform from AVEVA (Figure 1). On ABF side, it encompasses process parameters, measured KPIs, baked anode quality as well as all the refractory field survey data on a flue wall and headwall conditions. The use of the data science techniques like machine learning, parametric or Bayesian modelling was not new for ADK [1] and therefore, ADK wished to investigate these techniques again in order to further optimize the smelter operations, their use being a premiere for such application to ABF.

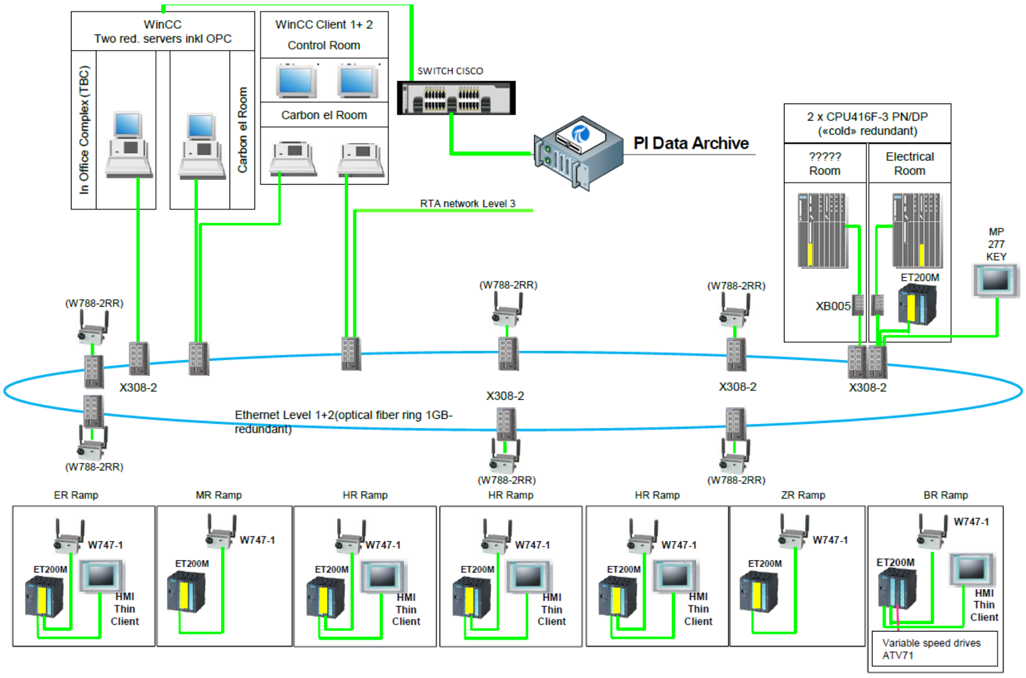


Figure 1. Data flow architecture at ADK.

Boston Consulting group is a leading strategy consulting firm with a strong footprint in designing, developing, and deploying analytics solutions via its tech build unit BCG X. With over 3 000 data scientists, engineers, designers, etc., BCG X has extensive experience in building end-to-end AI-powered solutions for industrial processes.

Fives Solios supplies Firing Control Systems (FCS) for decades. They are controlled by computers to integrate advanced control modules [2] ensuring complete and safe combustion of the volatile and injected fuel. Large historian data base were embedded in the solution and more recently dedicated digital monitoring modules were developed for enhanced ABF monitoring [3].

From a range of initially scoped use cases, the ABF application was jointly prioritized by ADK, BCG, and Fives based on potential gains, data readiness, and operational feasibility. Motivation and objective of the use case:

- In the anode baking process, higher quality is achieved through more intense baking via gas injection at substantial cost and emission impact.
- ABFs are typically operated steadily without taking active trade-offs between gas consumption and anode quality.
- The objective of the use case was therefore to develop an algorithmic solution to enable such trade-offs by identifying optimal baking laws taking baking parameters, furnace state, raw materials etc. into account.

Solution approach:

- Leverage Machine learning & Bayesian models to link baking laws to process outcomes (gas consumption and anode quality).
- Embed these models into a recommendation engine to generate baking laws based on the current furnace state.
- Co-develop an intuitive user interface to provide baking law recommendations to technicians and facilitate data-driven decision-making.

Value potential:

- Target of up to 10 % gas savings translating into substantial cost relief and reduction of CO_2 emissions.
- Potential increase in the lifetime of refractories which was eventually considered as difficult to evaluate within the time frame of such project.

2. The ADK Anode Baking Furnace

The ABF comprises 72 sections, 7 flue walls and 4 production fires running on natural gas. 432 anodes are produced every day to supply the 264-pots reduction line (Figure 2). It was rebuilt in 2013 to enable longer anodes to be baked. Only the original concrete shell has been preserved. A centralized control system [4] manages the process, enabling to achieve an average annual consumption of 2.6 GJ/t with a quality criterion of less than 5 % of core samples with Lc < 31 Å. Refractory maintenance is carried out daily, and flue walls are replaced according to aging and condition monitoring. An internal flue wall has a service life of at least 140 firing cycles. The ABF is actually connected directly to the Gas Treatment Center (GTC) which treats both the fumes from the ABF and the exhaust gas from the pots.

Figure 2. View of the ADK anode baking furnace.

3. Project Description

The project is divided into three phases:

- **Readiness Phase (2 weeks)**:
  - Conduct in-depth interviews to evaluate the ADK process, focusing on:
    - Historical and live data availability, connectivity, and quality,
    - Status of existing FCS and instrumentation,
    - Operational practices and control philosophy.
  - Build a data cube aggregating relevant data,
  - Conduct initial data exploration to quantify potential value pockets.
- **Proof-of-Concept (POC) Phase (10 weeks)**:
  - Analyze two years of data using data science techniques to identify key model features and outputs,
  - Develop the first model and optimization scenarios,
  - Evaluate these scenarios on historical data to:
    - Validate technical feasibility,
    - Refine value pocket quantification,
    - Define an initial implementation strategy considering production constraints and current processes.
- **Minimum Viable Product (MVP) Phase (planned for 16 weeks)**:
  - Conduct initial tests and refinements,
  - Integrate the solution into local tools, establish live data connections, and deploy on local servers,
  - Implement user interfaces,
  - Provide operator training.

This paper details the Readiness and POC phases, with the MVP phase just beginning.

4. Methodology and Results Step by Step

This section will detail the steps to development of an algorithmic solution, providing a comprehensive overview of the process. The key stages of development for the algorithmic core are summarized schematically in Figure 3.

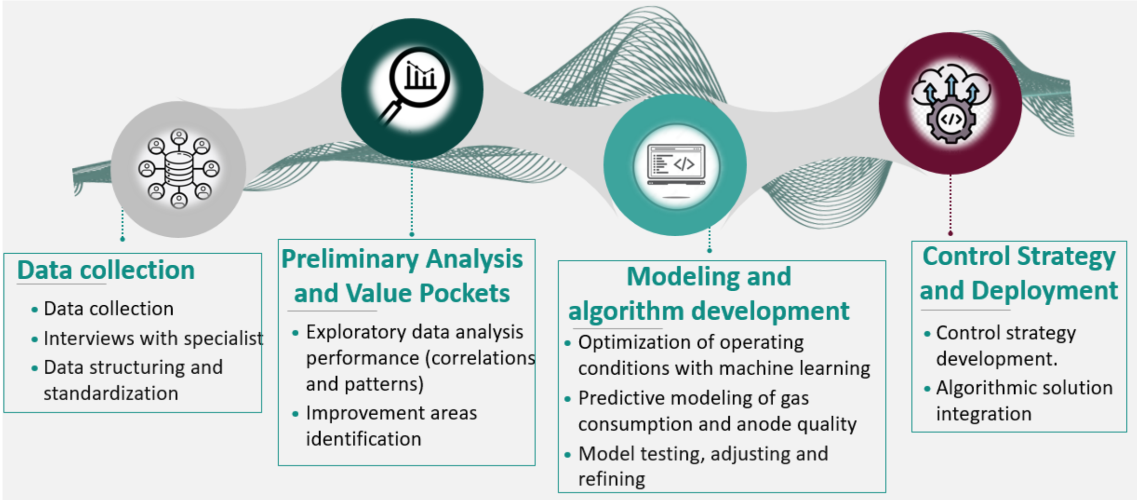


Figure 3. Key stages of development for the algorithmic core.

4.1 Data Collection and Interviews

Optimizing the operation of the ABF requires various data, collected from diverse sources. Highly granular, near real-time FCS data, notably from exhaust, heating, and blowing ramps is the central source of data collected in an AVEVA PI system. The system continuously monitors more than 1 000 process parameters such as temperature, pressure, flow rates, and injection levels, providing a detailed snapshot of the furnace's operation at any given moment. Additionally, manual and laboratory measurements of physical parameters and technician-conducted measurements play an important role in data collection, from green anode composition (mass, raw material, and pitch content) to refractory conditions (section age, flue wall condition, maintenance actions) as well as gas consumption data and gas supplier's data, such as Gross/Net Calorific Value (GCV/NCV). Furthermore, data on baked anodes, such as mass, coke calcination level (Lc), electrical resistance, and reactivity provide insights into final product characteristics. These data are sourced from AQ Manager, a LIMS (Laboratory Information Management System) software.

In parallel of the data collection process, various interviews with ADK experts were conducted (operators, engineers, Gas Treatment Center (GTC) and electrolysis specialists, IT personnel etc.) to synthesize constraints and requirements of the operation and to contextualize the collected data. Regular meetings with Carbon process engineers are paramount to project progress.

4.2 Data Aggregation and Data Cube

The data collected so far spans ~ 2 years to date, amounts to ~ 50 GB, and is continuously amended by the latest extracts. In order to facilitate these regular updates at arbitrary intervals, we built a modular and automatable data processing pipeline which aggregates all the data in a single *data cube*. Common keys and granularity levels facilitate the analyses and model development. Consolidating these vast datasets into a structured format provides a comprehensive and multi-dimensional view of furnace operations.

The data cube comprises eight key data elements representing distinct aspects of furnace operations linked by common keys such as date, fire, section, flue wall, and ramp number and type (see Figure 4).

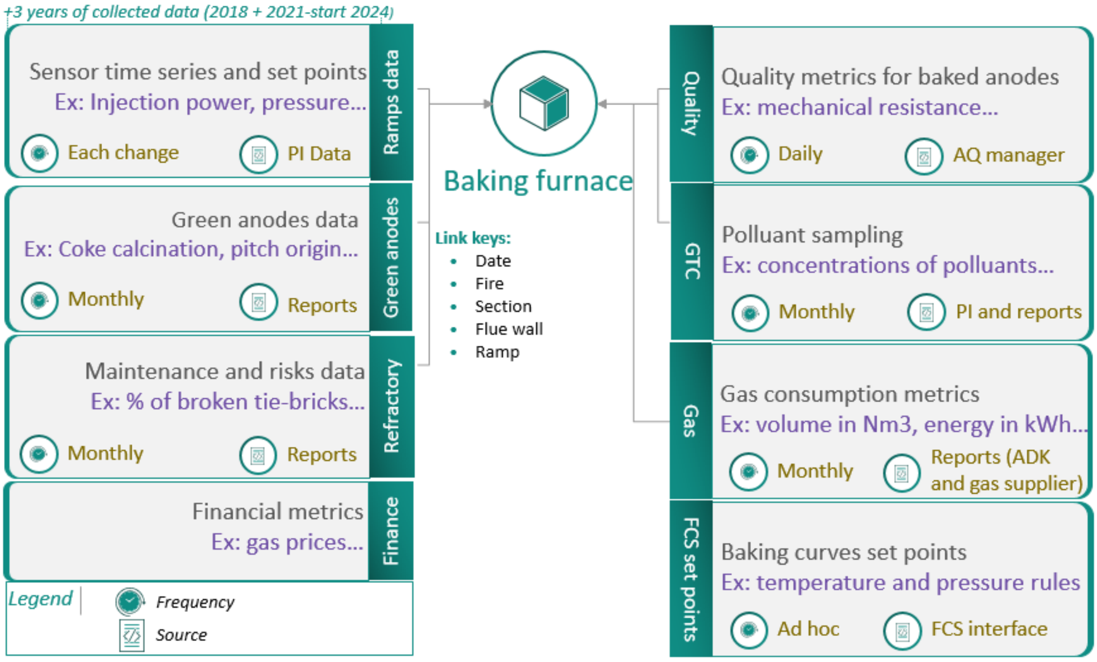


Figure 4. Project data cube.

In addition to structuring the data, various preprocessing techniques were applied to ensure data quality and usability. For instance, removing outliers, filtering out partial baking, and establishing baking IDs to effectively integrate and link data, particularly quality metrics and baking parameters. Using Python and versioning in GitHub, the data were organized and cleaned to facilitate correlation evaluation and crosschecking of information. Notably, the most voluminous dataset, the ramp data, required special attention. This dataset contains variables such as temperature and pressure for every ramp, fire, section, and flue wall. With each alteration in any value of these variables, a data point is generated. To manage the volume effectively, the data were aggregated to hourly or five-minutes frequency, with flexibility to adjust the frequency to adapt to specific analysis requirements.

4.3 Preliminary Analyses

After collecting the data, the next phase of the project focused on generation of hypotheses grounded in data, like KPI analysis and correlations between variables. Goal was to uncover patterns between variables, particularly those influencing gas consumption and to confirm experts’ intuition about its main drivers. This was done through two main types of analysis. First, we visualized key KPIs (e.g., quality, injection level, refractory defects) across the furnace. For example, the heat map in Figure 5 shows the average over-consumption across the baking furnace derived from heating ramp data. Crossover sections as expected consume more gas than others. Additional over-consuming sections were identified (e.g., Section 17) and subsequently visually inspected.

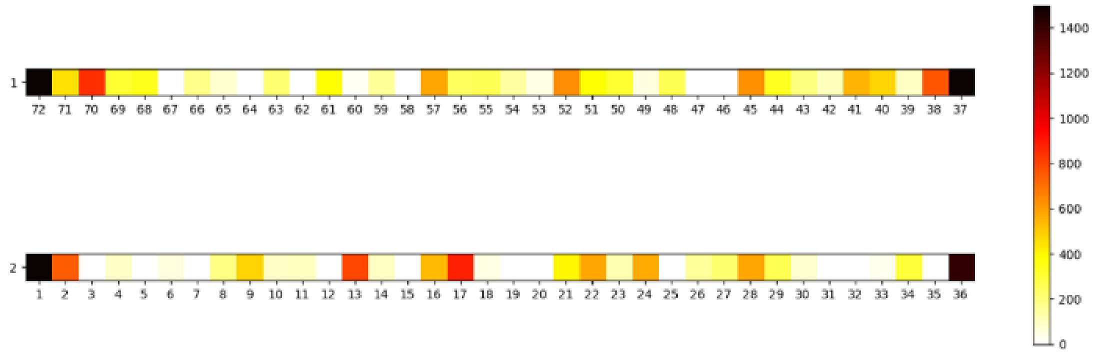


Figure 5. Average overconsumption per section.

In parallel, multiple correlation analyses were performed between the collected features to identify the main drivers of gas consumption and to explore temporal and spatial effects. For example, we investigated how the end temperature of the exhaust ramp (best proxy of position of degassing) on previous sections affected the injection of the first heating ramp, or how the initial temperature of this first ramp affected its consumption (Figure 6).

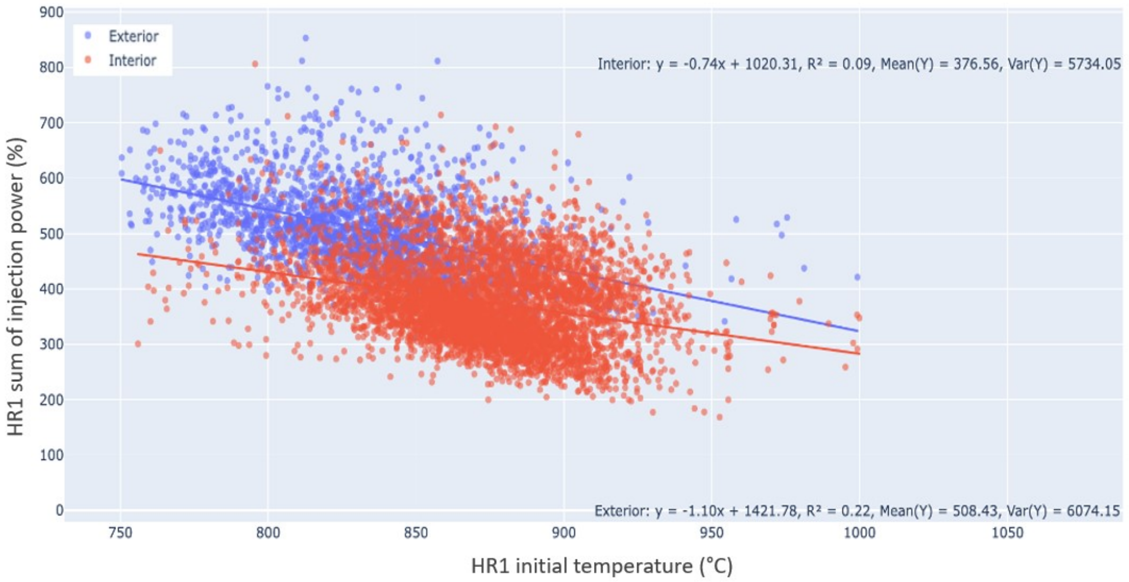


Figure 6. Example correlation plot: Initial temperature of the first heating ramp (HR1) vs its total gas consumption expressed as the sum of the two injectors' power (in % of nominal power capacity) from HR1 over 24 hours.

4.4 Value Pocket Identification and Quantification

Subsequently, to orient all further efforts, key areas for improvement across ABF operations were formalized and quantified based on historical data. Figure 7 provides a summary of the potential in terms of reduction in gas consumption. As shown, the most significant impact can be expected from baking law optimization and modification of draft. Although the other areas are also important, their potential in reducing gas consumption is low given ADK's point of departure. While the methodology is universal, it would likely surface different results when applied at other smelters.

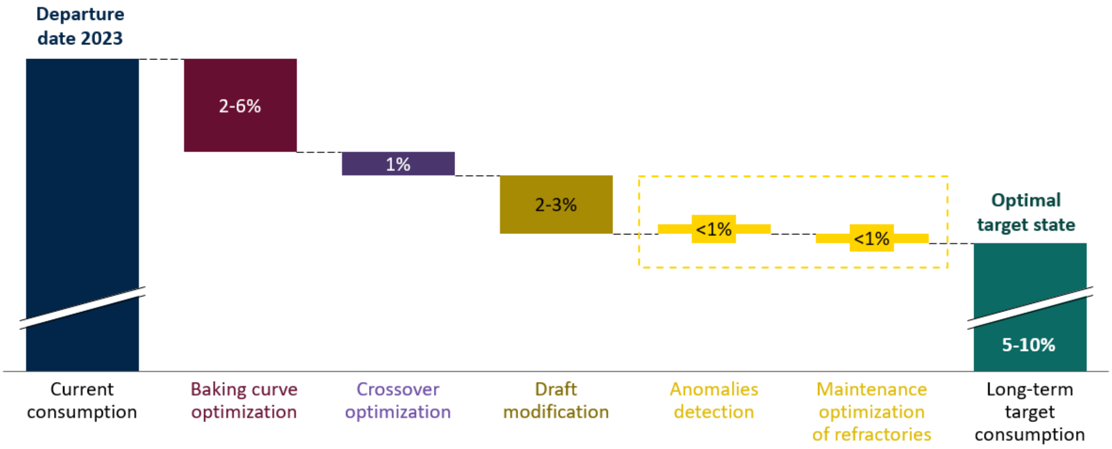


Figure 7. Potential gains on baking furnace gas consumption per identified area for improvement.

Optimization of baking laws relates to defining alternative baking profiles at the heating ramp by modifying three key criteria: soaking time (duration at highest temperature), soaking temperature (highest temperature), and rate of temperature increase until soaking temperature (with potential inflection points along the way). Based on extrapolation from historical data, we assessed the impact of changing said criteria, determining a potential reduction in gas consumption between 2 % and 6 %.

As in every baking furnace, the crossover sections play a significant role. The eight sections considered as crossover sections account for 13 % of the total gas consumption, with a consumption 20 % higher than straight-line sections. Enhancing refractory maintenance in these areas or refining furnace control to optimize degassing can yield an 8 % reduction in gas consumption translating to a 1 % reduction in total gas consumption across the furnace.

Draft modification relates to the optimization of the draft at the exhaust ramp to reduce energy lost to the GTC while maintaining steady combustion of volatiles and gas. By reducing excess air flow, which is associated with surplus energy, we have determined a potential reduction of gas between 2 % and 3 %.

Regarding anomaly detection, algorithms were developed to detect alterations in thermocouple signals and identify defects as well as to recognize irregular fluctuations with unexpected variation in gas consumption that could indicate flooding situations. These algorithms enabled the tracking of the number of suspected anomalies throughout the study period. However, in the case of ADK, the impact on gas consumption is minimal. Despite gas consumption being higher during anomalies, and faster detection helping prevent overconsumption, the low frequency of anomalies per year results in only minor overconsumption.

Assessing the condition of refractory materials used in the furnace algorithmically and implementing strategies to optimize maintenance schedules also represent opportunities for improvement. Gas consumption increases by 30 % in flue walls with the most worn refractories. However, in the context of ADK, with a fixed budget, the trade-offs would be made between highly worn and slightly less worn refractories where we determined the delta in gas overconsumption to be minimal.

4.5 Control Strategies Definition and Associated Algorithms

As outlined previously, the two aspects contributing most to gas reduction are baking law optimization and modification of draft. The proposed methodologies to improve these aspects will be explored in this section.

4.5.1 Baking Law Optimization

Baking furnaces are typically operated in an almost static fashion applying the same baking laws across the furnace with few variations for example to differentiate crossover from straight line sections. Our analyses indicate significant gas reduction potential in moving from applying static baking laws to selecting baking laws dynamically taking into account the refractory state, potential defects, historical performance, and more. This paradigm shift of steering baking furnaces dynamically and at higher granularity is the motivation for the algorithmic solution laid out in the following.

The goal is to develop a dynamic model that takes into account this information to propose the best baking law to minimize gas consumption while maintaining quality thresholds. To achieve this, two main models were built: one predicting gas consumption and one predicting anode quality. These models are linked via a recommendation engine. This engine generates all possible baking law candidates, each of which is evaluated in terms of expected quality and ranked in terms of expected gas consumption. Candidates that did not meet the quality criteria were ruled out. The baking law that minimized gas consumption was kept as the optimal solution. The procedure is illustrated in Figure 10 and explained in detail in the following.

Explicative variables: Inputs included fire number, section type (straight line or crossover), flue wall (internal or external), average historical consumption, refractory age and main type of defects as well as previous baking parameters (e.g., ER – Exhaust Ramp - final temperature providing a preliminary indication of the degassing position). In addition, the baking law parameters as depicted in Figure 8 are included in the model.

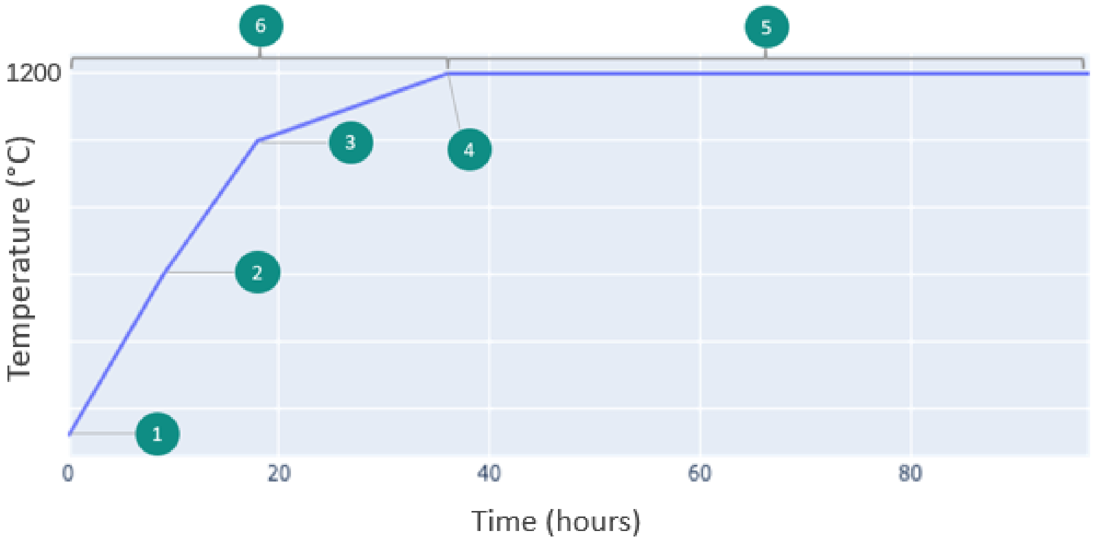


Figure 8. Key baking law parameters: 1) initial temperature, 2) quarter slope duration temperature 3) midway slope duration temperature 4) maximum temperature 5) duration of plateau 6) duration of slope; y-axis shows baking duration in hours.

Pre-processing: Outliers, which are data points that deviate significantly from the typical pattern or distribution, are removed based on business considerations and physical constraints. For instance, any data point suggesting an initial heating ramp temperature below 600 °C would be considered anomalous and thus removed. In addition to typical statistical features (e.g., mean, max, median), more complex features were created (e.g., delayed versus theoretical cycle time start, initial heating ramp temperature when starting the injection, ratio between flue-wall past injections). For each feature, there is a clear causal link to the respective target variable, and only features with significant impact on model performance were retained.

Gas consumption model: The algorithm chosen for this model is XGBoost's Gradient Boosted Decision Tree [9]. It builds an ensemble of decision trees that work together to improve prediction accuracy. It is known to be efficient in capturing non-linear relationships and robust to overfitting via various means of regularization. It has been trained on two years of data (2022–2023) to learn the relationship between the features described above and the response variable gas consumption, as given by the power of the heating ramp injectors. Performance is back-tested on the first three months of 2024, and achieves a R-squared value of 0.72. This metric indicates how well the model's predictions match the actual data. The features were analyzed using SHapley Additive exPlanations (SHAP values) [10], a framework to quantify the contribution of each feature to the final prediction (Figure 9). For example, we can see that a high HR (Heating Ramp) start temperatures leads to less gas consumption which matches process understanding.

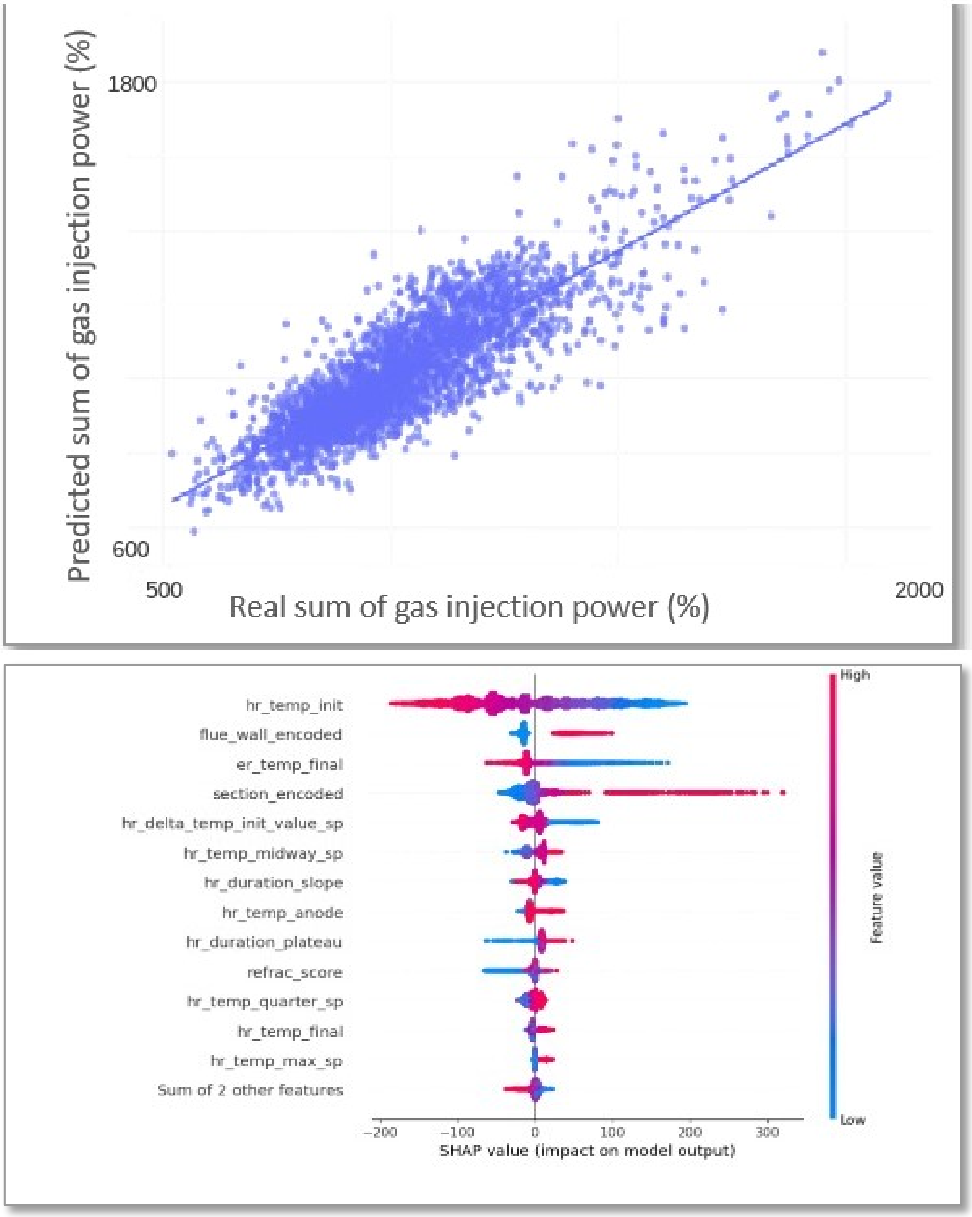


Figure 9. For gas consumption model, (Top) Predictions vs Actuals (Bottom) SHAP values.

Quality model: A similar approach was used to predict quality, except that the target variable was now the coke calcination measured by the crystallite size Lc in angströms. Similar training and evaluation was carried out, and after refinement of the data in scope with the Carbon team, an R-squared of 0.68 was achieved on the back-test. To inject more physical knowledge into the model, we also developed a custom Bayesian probabilistic model in the programming language STAN [11]. This allows to fully specify a mathematical model of the underlying process along with prior assumptions on the distribution of each parameter. The model parameters are subsequently updated based on data. In particular, this allows to estimate unobserved effects such as the heat exchange between flue wall and anode in a probabilistic way. Figure 10 illustrates how a probabilistic anode temperature profile (red line and gray shading) is estimated based on fixed flue wall temperatures (blue lines).

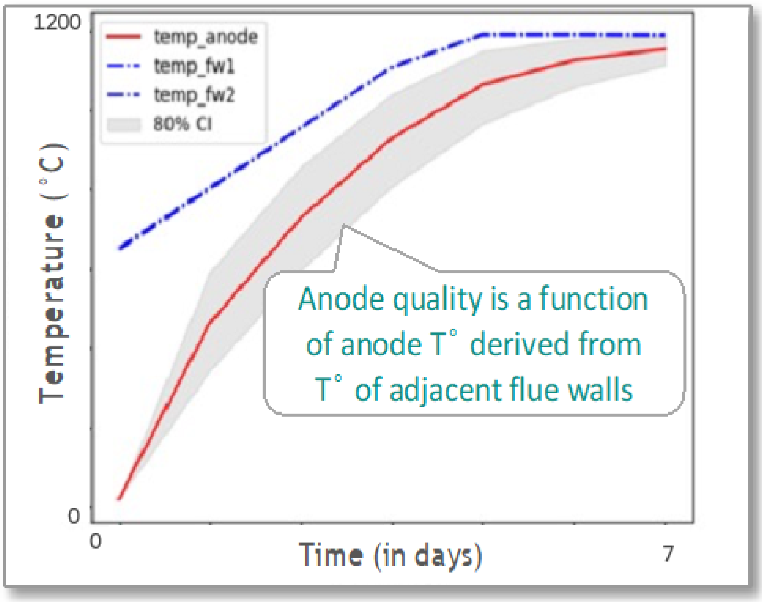
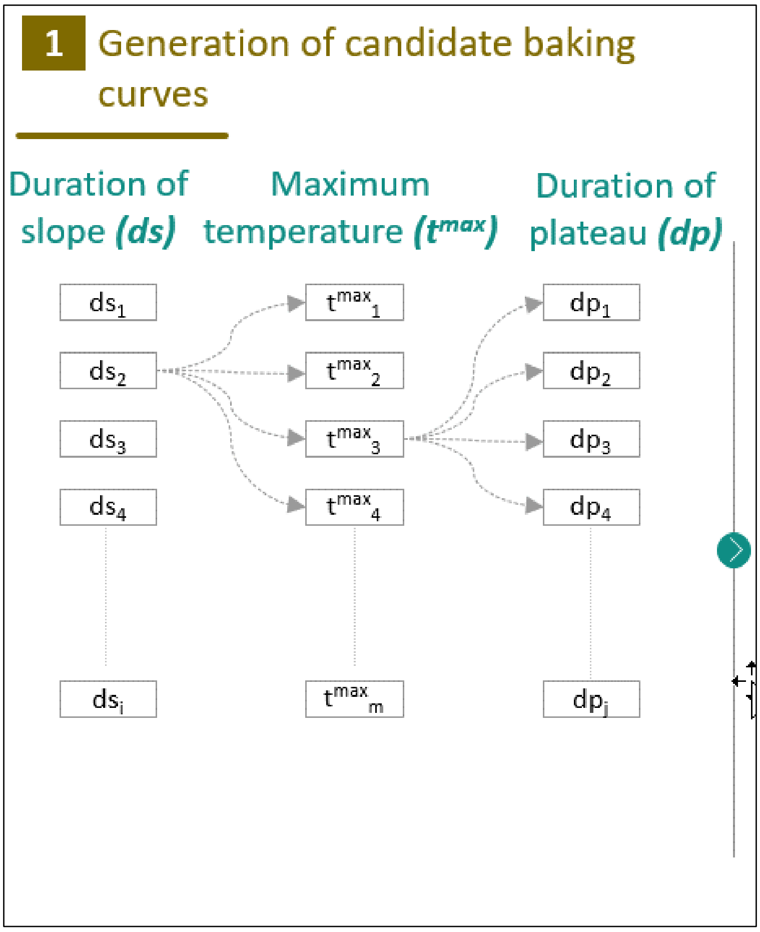
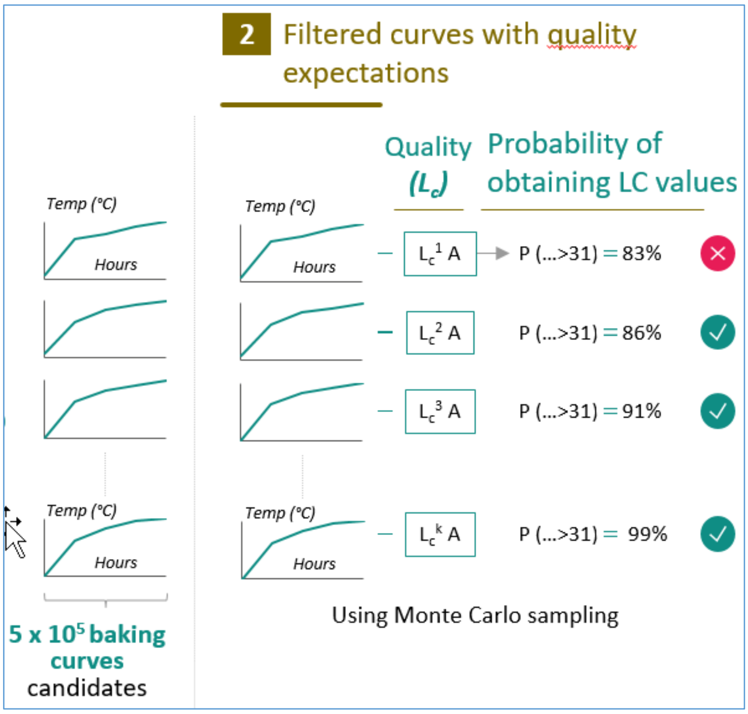


Figure 10. For quality model: Evolution of anode temperature based on adjacent flue walls ones.

Recommendation engine: This engine first generates all ~50 000 possible baking laws by combining all baking law parameters from their pre-defined ranges (sensible boundaries aligned with Carbon team). For each candidate curve, the described quality and gas consumption models are used to generate a prediction. The predicted baking law quality serves as a filtration to reflect the current operational threshold where no more than 5 % of anodes with Lc below 31 Å are tolerated. Among all baking laws in line with quality requirements, the one with minimal consumption is chosen. Figure 11 illustrates the steps of the recommendation engine.


1
Generation of candidate baking curves
Duration of slope (ds)
Maximum temperature (t^{max})
Duration of plateau (dp)
ds_1
ds_2
ds_3
ds_4
ds_i
t^{max}_1
t^{max}_2
t^{max}_3
t^{max}_4
t^{max}_m
dp_1
dp_2
dp_3
dp_4
dp_j


2
Filtered curves with quality expectations
Temp (°C)
Hours
5 x 10^5 baking curves candidates
Quality (L_c)
Probability of obtaining LC values
L_c^1 A
P (...>31) = 83%
L_c^2 A
P (...>31) = 86%
L_c^3 A
P (...>31) = 91%
L_c^k A
P (...>31) = 99%
Using Monte Carlo sampling

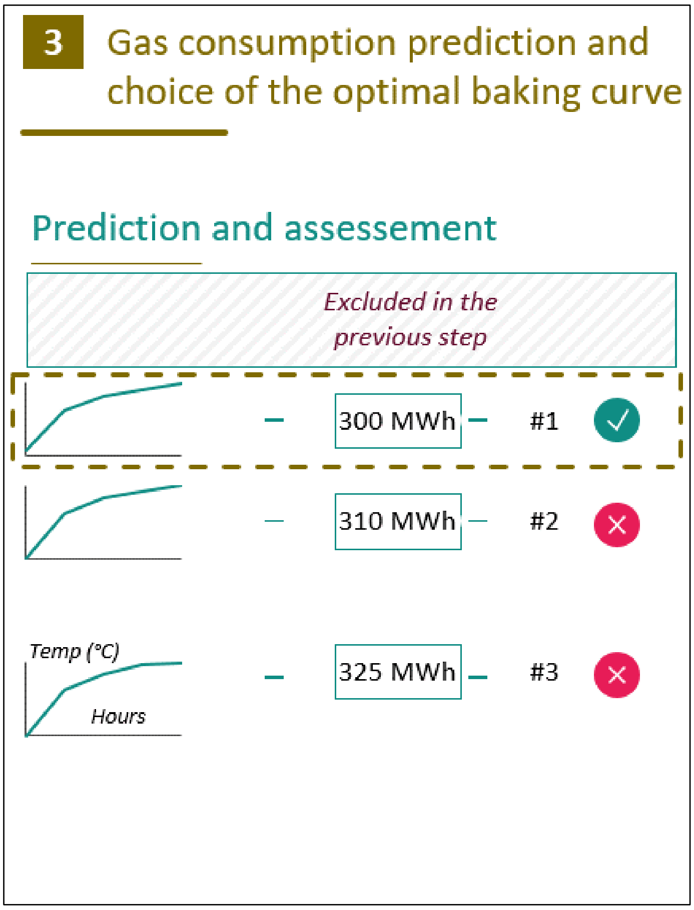


Figure 11. Recommendation engine steps to choose optimal baking laws.

4.5.2 Modification of Draft Pressure

Preliminary analysis of the furnace indicated potential improvements in draft pressure control. To refine this, we used a physical model to calculate:

- The theoretical flow required for optimal combustion. This was estimated using the gas flow required for combustion based on injection setpoints and the volatiles flow, derived from anode weight and cycle time.
- The actual flow in the flue wall. This was calculated using the flue wall parameters (temperature, pressure, damper opening), and calibrated against measurement of fumes flow exiting the furnace to the GTC.

By comparing these values (see Figure 12), we identified instances of excessive flow and gained insights on how to adjust the draft pressure. The next step involves translating this delta into a recommended range of draft pressure, to ensure that the exhaust ramp control allows to align more closely to the theoretical flow.

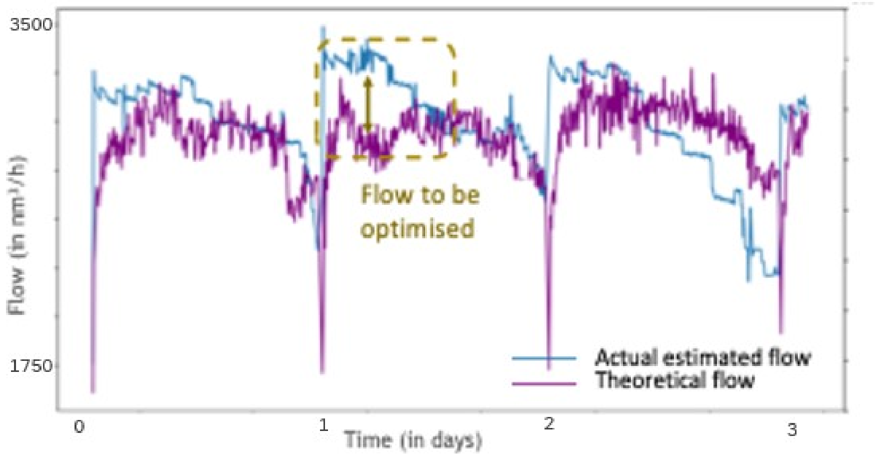


Figure 12. Exhaust flow model – Comparison between actual and theoretical air flow.

5. Conclusion

The use of data-driven techniques for real-time ABF control with such granularity of action is a first and a great challenge. The results obtained in the first two phases of this project look promising, with confirmation of over 5 % reduction in gas consumption while maintaining anode quality.

The combination of data science, OEM expertise and field operations, as well as the project methodology with frequent cross-checks with the operating team and validation of each step by management, are the key factors in the success of this project.

This proves that the use of granular operating data can be exploited with modern data science tools to enable detailed prediction of gas consumption and anode quality. Due to limitation of instrumentation in the preheating zone, we continue to rely on physical and theoretical modeling to characterize optimal flow rates in real time.

Although there are still improvements to be made in predicting the position of the degassing front, the results obtained so far are sufficient to deploy and test the new advanced control strategy and measure the reduction in gas consumption. This will be the next phase of the project.

The deployment of this future control strategy and the results of actual performance will be carried out progressively to build confidence and acceptance of the operation and will provide a decision-support tool on the operation side that will enable practices to be standardized.

Eventually, if pushed to the limit of granularity with actions potentially tailored for each pit and updated at each fire cycle, the existing conventional FCS will have to be upgraded.

6. References

1. Bilal Azennoud et al., Anode Quality Monitoring Using Advanced Data Analytics, *Light Metals* 2019, 1237–1245.
2. Hugues Vincent and Nicolas Fiot, Helios: A firing and Control System for Anode Baking, *Aluminium International Today Magazine*, January-February 2012, 34–36.
3. Christophe Bouché and Nicolas Fiot, Make your Anode Baking Furnace Smart and Greener, *Proceedings of 10th International Bauxite, Alumina & Aluminium Conference*, Raipur, India, 14–17 September 2022, 297–308.

4. Philippe Contard et al., Anode Baking Furnace Firing System Lean Engineering, *Light Metals* 2014, 1171–1174.
5. Felix Keller et al., Specific Energy Consumption in Anode Bake Furnaces, *Light Metals*, 2010, 1005–1010.
6. Pierre Mahieu et al., Anode Baking Process Improvement at Alro, *Light Metals* 2013, 1155–1161.
7. François Ordronneau et al., Meeting the challenge of increasing anode baking furnace productivity, *Light Metals*, 2011, 865–870.
8. Detlef Maiwald, Wolfgang Leisenberg, Flue condition Index - a new challenge to increase flue lifetime, operational safety and fuel efficiency in open pit anode baking furnaces, *Light Metals* 2006, 625–630.
9. Tianqui Chen and Carlos Guestrin, XGBoost: A Scalable Tree Boosting System, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, 785–794.
10. Scott M. Lundberg and Sun- In Lee, A Unified Approach to Interpreting Model Predictions, *Advances in Neural Information Processing Systems 30,* 2017, 4765–4774.
11. Bob Carpenter et al., Stan: A Probabilistic Programming Language, *Journal of Statistical Software*, 2017, 76(1), 1–32.