A Machine Learning Approach to Early Detection of Incorrect Anode Positioning in an Aluminum Electrolysis Cell

Patrice Côté[1] and Sébastien Guérard[2]
1. Principal Advisor, Data Science and Artificial Intelligence
2. Reduction Data Scientist / Modeller
Rio Tinto – Aluminum Technology Solutions, Jonquière, Canada
Corresponding author: patrice.cote@riotinto.com

Abstract

The Hall-Héroult process involves positioning very large anodes, weighing several tonnes, within a few millimeters of a given reference. Despite the continuous improvement of anode setting technologies and practices, there are still several cases where incorrect anode positioning impedes good cell performance. There is currently no reliable way to detect these abnormal situations before they become a problem.

The presented approach uses new sensors developed by Rio Tinto to continuously monitor current in individual anodes. Data from these sensors, together with machine learning techniques, have been used to elaborate an algorithm that predicts the anode current pickup which, in turn, is used to evaluate anode positioning. These sensors and model are currently being deployed on industrial cells and have been shown to greatly improve operators' decision making.

Keywords: Aluminum electrolysis cells, anode setting, anode positioning, continuous anode current monitoring, machine learning.

1. Introduction

Operating a Hall-Héroult cell typically involves replacing one or two anode assemblies every 24 – 36 hours. This operation is critical because it has the potential to significantly affect short-term cell performance, namely current efficiency (CE) and specific energy consumption. In fact, one of the most important challenges of any aluminum smelter is to keep the disturbance from the anode change to a minimum; thermal balance, alumina dispersion and current distribution are all directly affected by anode change to some extent. When it comes to changing an anode, there are several critical tasks, a crucial one being the positioning of the anode at the appropriate height where a precision of less than 5 mm is often expected.

Even with the latest technology, positioning large carbon blocks weighing several tons within a few mm of a reference is a difficult mission. Most smelters use the old anode as a reference; the goal is to set the bottom of the new anode at the same altitude as the bottom of the old one (plus a certain amount to account for the behavior of the new anode). This process, called gauging, relies on a number of references where many errors can add up: it is possible for the anode-cathode-distance (ACD) of the old anode to be incorrect in the first place, the stem of the new anode might be slightly angled resulting in a wrong measurement, the measurement equipment could be de-calibrated or include mechanical looseness leading to less-than-perfect results. Even with excellent operational practices, there are cases where anodes end up at an incorrect ACD after setting (5 – 10 % is typical).

It has been shown by many studies [1 – 4] that CE is related to the average ACD of a cell (Figure 1).

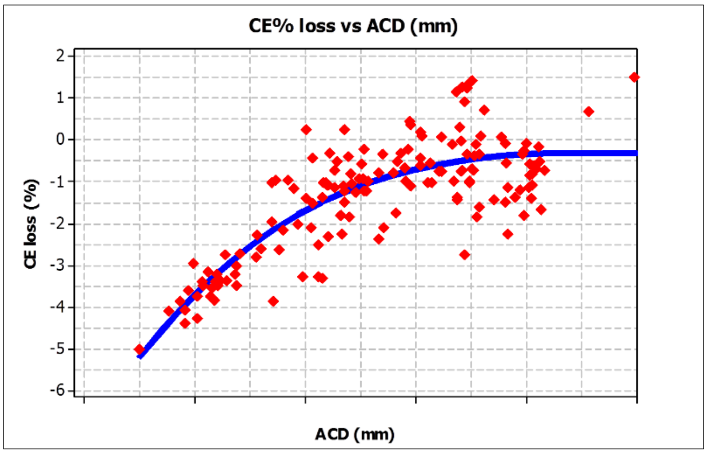

Figure 1. CE loss (%) as a function of ACD (mm) [4].

It is also believed, as studied by Tarcy [5] (Figure 2) that CE is strongly penalized by the anode drawing the most current – often the lowest one. As shown in Figure 3, this is coherent with calculations done with our own model. In addition, as stated by Segatz [6] *"the MHD stability strongly depends on the distribution of the ACD"*, which in turn also affects overall cell performance through other mechanisms [7].

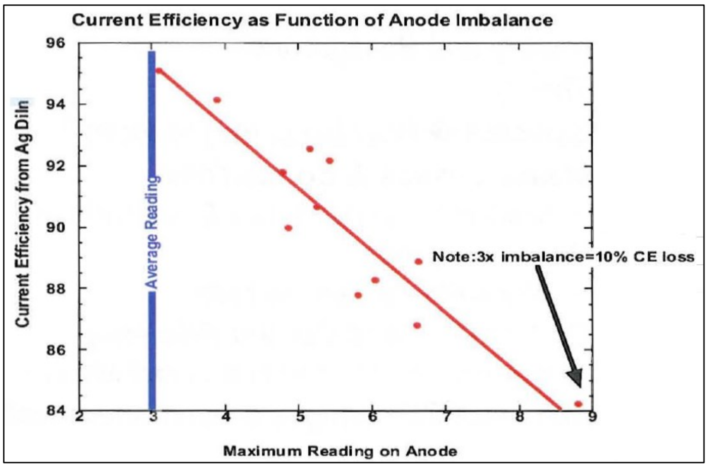

Figure 2. CE as a function of maximum anode current draw in a cell [5].

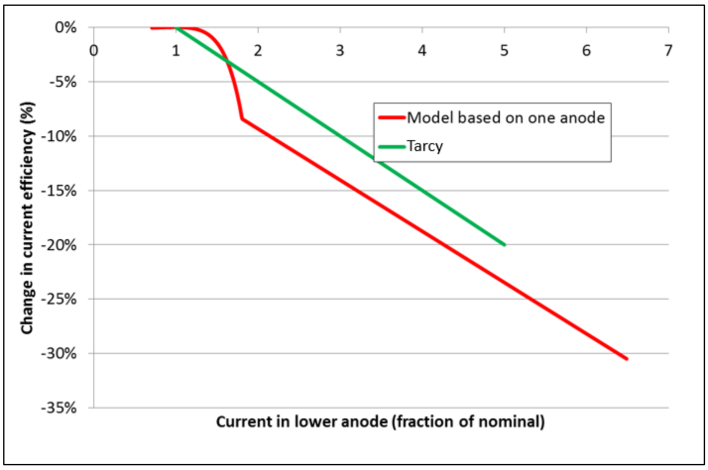

Figure 3. Superimposition of Rio Tinto model results and experimental curve (Figure 2).

With that in mind, it appears critical to detect anode positioning problems as soon as possible before they can even have an adverse effect on cell performance. A machine learning (ML) algorithm that does just that was built using data from new anode current sensors under development by Rio Tinto.

2. Detection of Incorrect Anode Positioning: the Current Process

There is no easy way to measure the ACD of an individual anode in an industrial context. Most Rio Tinto smelters use an indirect evaluation based on the relationship between ACD and anode current draw. In order to do accomplish this, a measurement is taken on a daily basis using a manual tool that reads the voltage drop between two points on every anode stem. This voltage drop being (more or less) proportional to the current flowing through the anode stem, a relative current (%) can be calculated for each anode. Figure 4 shows the behavior of the relative current over a complete anode lifecycle, for a specific cell technology.

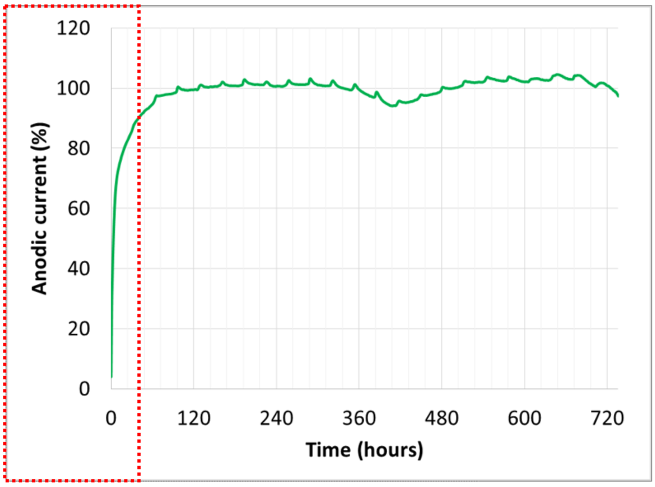

Figure 4. Average anodic current curve over a complete lifecycle.

When it comes to the assessment of anode positioning, the interest lies in the relative current of 24-hour old anodes. For any specific cell technology (and sometimes specific anode position) there is a 24-hour current target to aim for, in order to minimize current as well as ACD distributions over the long run. For instance, Figure 5 shows a target of 85 % for 24-hour old anodes; anything significantly higher or lower than this would translate into a performance loss.

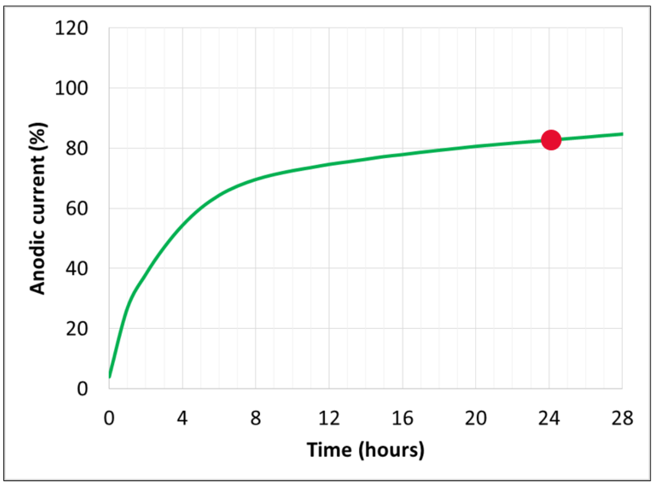

Figure 5. Zooming in on Figure 4, 24-hour target current for a correct positioning.

This system has been used for decades and works fairly well to identify outliers but it does have its limitations. As shown in Figure 6, those manual measurements are not completely up to the task when compared with data from continuous current sensors. For one thing, the measurement is much faster (a few seconds) than the period of the cell MHD noise (30 – 60 seconds) so a lot of this noise is inherently captured in the data. In addition, it is not adjusted for anode stem temperature so it cannot correctly be converted to current. Finally, this is a manual measurement which takes operator time and is prone to errors (wrong cell or anode numbering etc.).

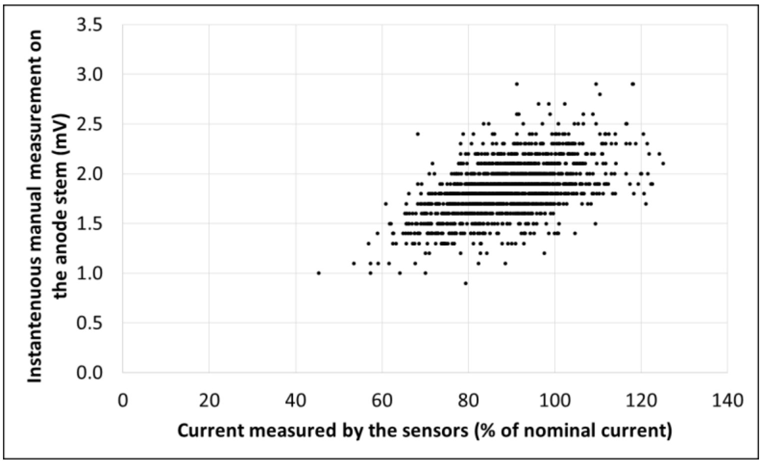

Figure 6. Comparison of manual stem voltage drop measurements and data from continuous anodic current sensors.

Once data are available, absolute thresholds are used in order to identify incorrect anodes. These problems can now be corrected but it is often late because cell performance has been hindered already. Figure 7 shows a retroactive evaluation of the performance of this system using data from continuous sensors as a reference. We can see that less than 50 % of incorrect anodes are currently detected by this approach (*precision* and *recall* in Figure 7 are explained in section 4.2).

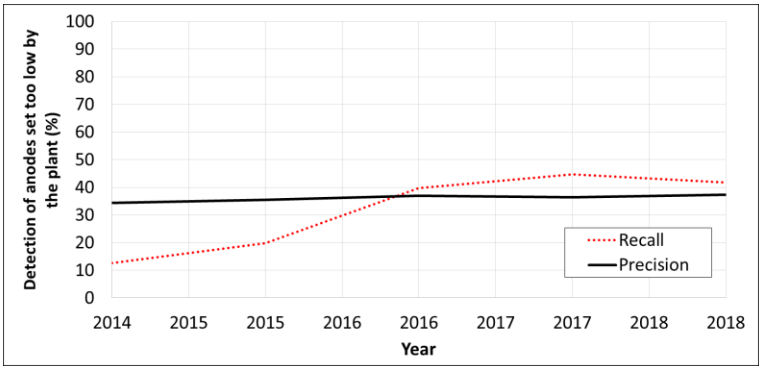

Figure 7. Retroactive evaluation of current system performance.

3. New Challenges and the Need for More Data

Historically, only two sources of continuous data were collected at a relatively high frequency and used in order to control the cells, namely the line current and the cell voltage, which were in turn converted to only one number: the cell resistance. This is not lot of data variety, considering that the cells are now 4 – 5 times larger than they were when the first control strategies were put in place (Figure 8). These new, larger cells are much more prone to local phenomena; parts of the cell can behave differently because of a less homogeneous alumina concentration, current density or ACD. Therefore, it makes less and less sense to represent and control the cell by using only one number. When combining this new reality with the rise of data science and artificial intelligence, it becomes obvious that there is a need for more data, both in terms of variety and volume.

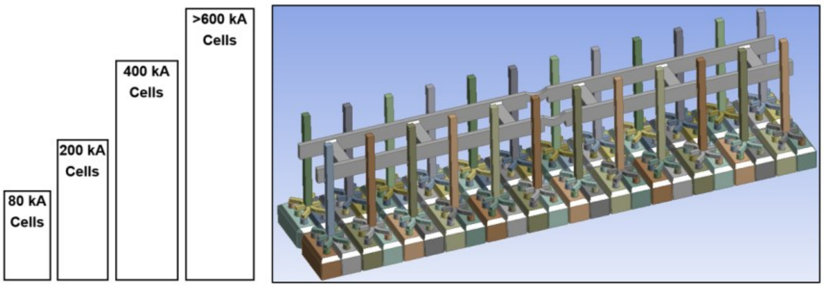

Figure 8. The relative dimensions of cells, from an old 80 kA cell to a modern 600+ kA cell.

One way to characterize the cell in a more accurate manner is to measure the current flowing through each anode stem. This allows the monitoring of specific parts of a cell, enabling new control strategies. Rio Tinto has been quite active in that field since 2012; our own sensors are currently collecting such data on a growing number of cells on three different technologies in order to feed R&D needs including modeling and, more recently, data analytics.

The industrialization of such an extensive monitoring system requires the capacity to store, analyze and process much more data. Ultimately, 1500 times more data will be collected, considering a frequency of 1 Hz for each data point. This is illustrated in Figure 8.

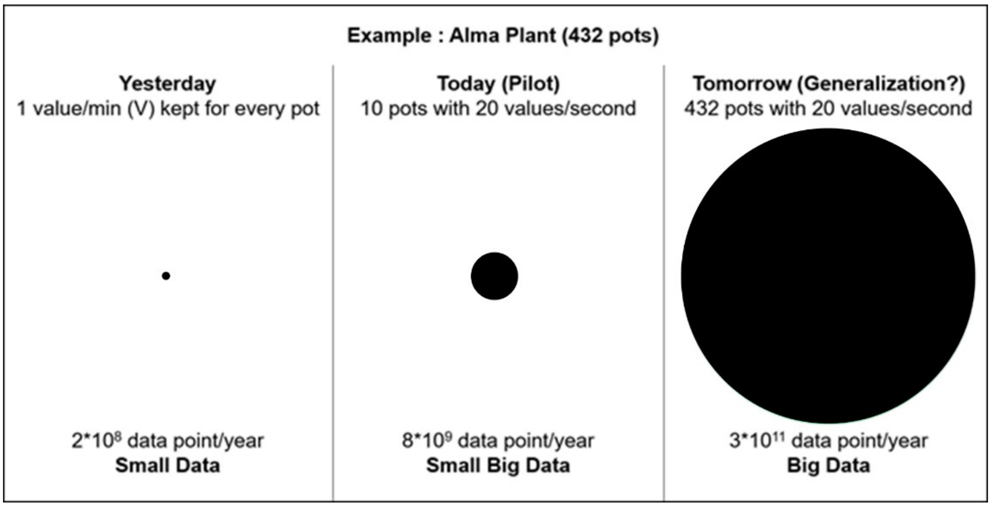

Figure 9. Massive volume of data collected by anodic current sensors at high frequency.

4. Machine Learning Algorithms Applied to the "Anode Positioning Problem"

Machine learning (ML) is a technique by which a computer can "learn" some relationship between input and output variables using existing data. The aim of ML is usually not to understand the physics behind these relationships because most algorithms are too complex in nature to be easily interpretable. The main interest lies in the fact that, somehow, a good ML model is able to capture the underlying relationships between the variables and make predictions that generalize well to new data. A thorough validation process has to take place in order to characterize the performance of the model and to guarantee it, up to a certain point.

During this validation process the model has to be tested against a "test dataset" that is completely independent from the data used for its training.

There is an ever-growing number of ML algorithms to choose from, and selecting the one that best matches our specific problem is not always obvious (in fact, it is never obvious). Table 1 lists some criteria to consider when selecting an ML algorithm.

Table 1. Criteria and considerations when selecting a ML algorithm.

Criteria	Considerations
Performance	Are the results good enough and useful in the context? Two important metrics for classification problems – like the one described in this work – are *precision* and *recall* (see section 4.2)
Complexity	Would a simpler approach work? When choosing between 2 approaches with similar performance, always go with the simplest one.
Speed and Calculation power	Is the calculation power and data infrastructure available to predict results at the required pace? The algorithm complexity together with the required volume of data can be very demanding technology-wise.
Maintainability	How stable is the algorithm in time? There has to be a control plan or a way to detect drifts and to correct (often re-train) the model.

4.1. General Machine Learning Approach to the Problem

In this specific case, our goal was to use the current ramp-up of a new anode to predict its final state: current after 24-hour. Some 7000 anode lifecycles – gathered from anodic current sensors – were used to train, test and validate the model.

During the course of this work several models were explored. We started out with a simple approach which was a list of twenty three regression equations (one for each hour of anode life up to 24). It actually worked, but had its limitations: for instance, it could only be ran once an hour, which was not flexible enough for our use case. After some iterations the results improved significantly thanks to a better algorithm, which can output predictions at high frequency. Predictions are also made on average one hour earlier considering the same error rate. The final model is referred to as *RBF model* in Figure 10.

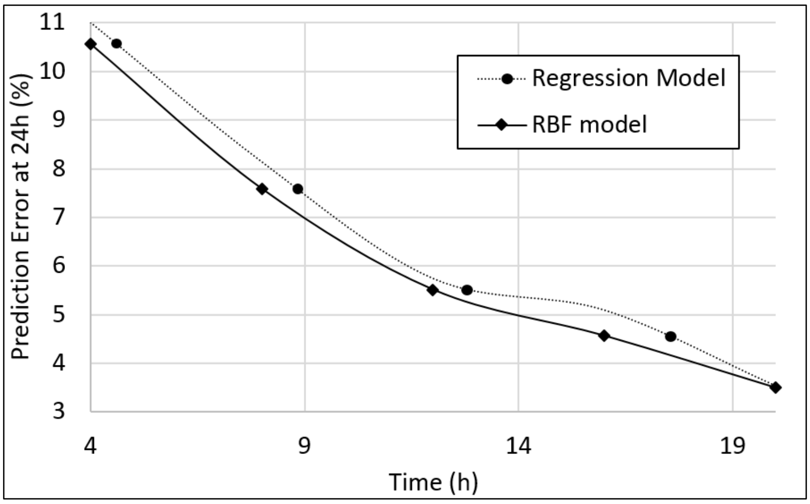

Figure 10. Improved model: predictions are made on average 1 hour earlier.

4.2. The Chosen Algorithm

The chosen algorithm is relatively complex but presents two major advantages over most other algorithms that were tested: the performance is better, and it also automatically adapts to changes. Where most ML algorithms would be trained on historical data and become static thereafter, this one continuously updates its references when new data becomes available.

Let's have a look at a summary of the algorithm, in three main steps:

1. A database of several anode lifecycles is built and constantly kept up-to-date (7000 in our case).
2. Similarities are calculated between the anode we are trying to predict (the subject anode) and all anodes in the database. The *N* most similar anodes will be kept for further steps.
3. A prediction is made, along with a confidence interval, based on the average final state of the *N* closest anodes and their statistical distribution.

4.2.1 Step 1: Building a Database of Anode Lifecycles

The algorithm needs to have access to several anode lifecycles, because it relies heavily on similarities between the subject anode and all the anodes in the bank. Therefore, the more anodes in the database, the more likely it is to find anodes with a similar behavior, although there is a cost in terms of calculation time. We have observed that the system works best when integrating recent anode lifecycles, most probably because of trends in the anode properties, work practices or cell conditions in general. Therefore, we chose to keep a constant number of anode lifecycles in the database, and to update it with new information at a regular interval.

4.2.2 Step 2: Finding the *N* Most Similar Anodes

The root-mean-square-deviation is used to estimate the similarity between two anodes:

$$\Delta I = \sqrt{\frac{\sum_t \left(I^t_{Anode\,1} - I^t_{Anode\,2}\right)^2}{n}} \tag{1}$$

where:

ΔI	Difference between the two anodes, %
t	Time, h
n	The number of time steps dt used to average the current
$_1$	Current of first anode being compared at time t, %
$_2$	Current of second anode being compared at time t, %

The current in Equation (1) is averaged over periods dt, itself a parameter of the model. We observed that better results could be obtained by giving more weight to values closest to the time up to which the anodes are compared:

$$\Delta I = \sqrt{\frac{\sum_t \varepsilon^{(T-t)}\left(I^t_{Anode\,1} - I^t_{Anode\,2}\right)^2}{\sum_t \varepsilon^{(T-t)}}} \tag{2}$$

where:

ε	Weighting factor, $0 \le \varepsilon \le 1$
T	Age up to which the two anodes are compared, h

4.2.3 Step 3: Predicting Anodic Current at 24-hour

In the case of this model, the prediction is very straightforward as it is directly established from the current at 24-hour of the most similar anodes. In fact, the average and standard deviation over the N values could have been used as the prediction, but in order to give more weight to the most similar anodes, a Gaussian radial basis function was used instead:

$$W_i = e^{-\frac{\Delta I_i^2}{2\sigma^2}} \tag{3}$$

where:

W_i	Weight given to anode i in the prediction
σ	Standard deviation of the Gaussian function

Finally, those weights are used to calculate the predicted current at 24-hour according to Equation (4):

$$I_{24h} = \frac{\sum_{i=1}^{N} W_i I_{24h_i}}{\sum_{i=1}^{N} W_i} \tag{4}$$

4.3. Validation and Some Words on the Model

As a final step, the four parameters of the model *(dt,ε, N, σ)* were tuned using a grid search of more than one thousand configurations.

This model can be summarized using an existing type of neural network called normalized Radial Basis Function (RBF) network, which is a RBF network where the output of the hidden layer is normalized. In the current application, the inputs to the network are the current at each time t for the subject anode, the hidden neurons correspond to the anodes in the database, and the final weights are the current at 24-hour of each anode.

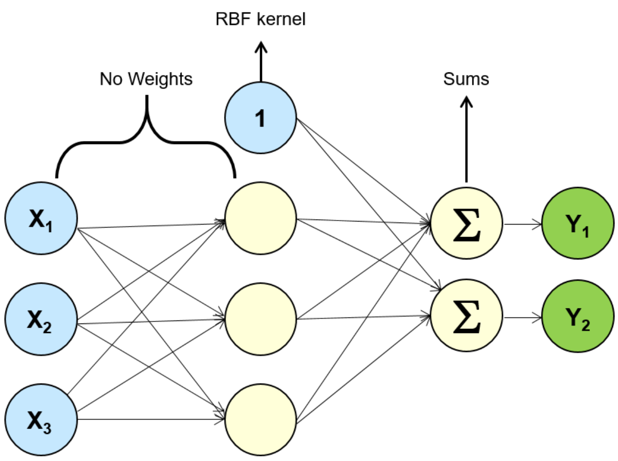

Figure 11. Graphical representation of a Radial Basis Function (RBF) network.

4.4. Results and Application

As can be seen in Figure 12, the predictions after 1 – 4 hour are far from perfect but they improve dramatically afterwards. During the first few hours, only anodes drawing very much – or very little – current will trigger an alarm, hence the most problematic anodes will be corrected quickly while the ones that are just a little bit off will be identified later on.

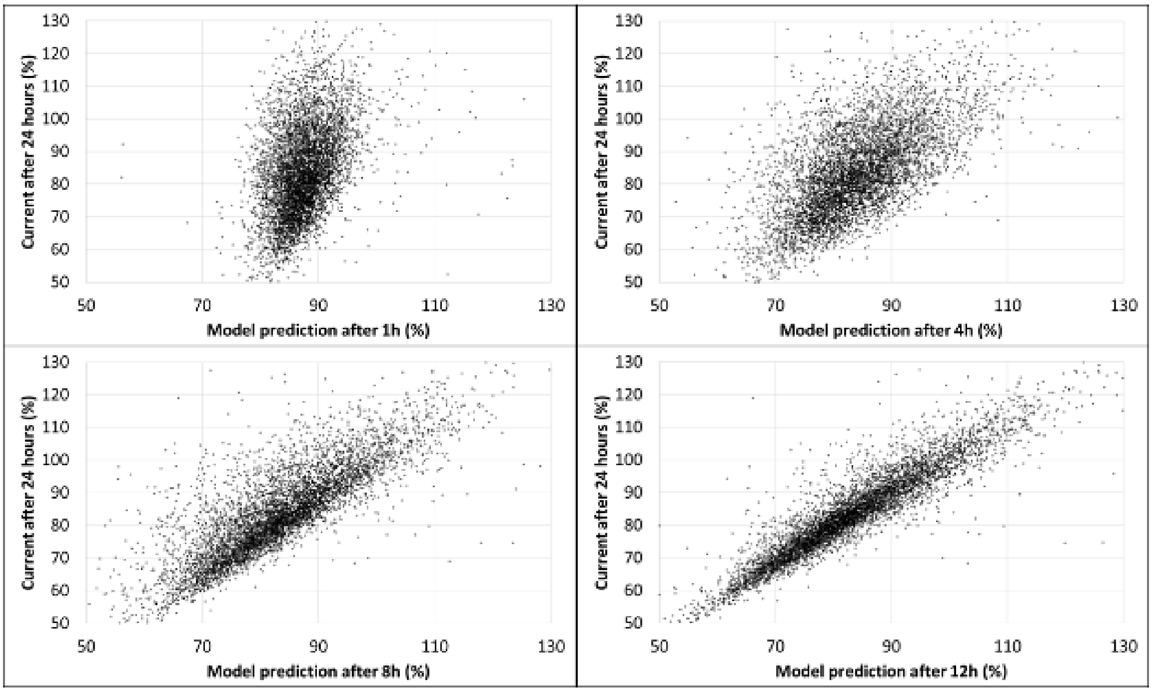

Figure 12. Actual vs Predicted 24h-current after 1, 4, 8 and 12 hours.

One way to characterize ML models performance is to calculate the *precision* and *recall*.

Precision, in our context, is defined as the proportion of **correctly predicted** “anodes too high or too low” (true positives or *TP*) out of **all predicted** “anodes too high or too low” (true positives *TP* and false positives *FP)* as per Equation (5). Precision is very important in our specific context because you do not want to waste precious operation time to adjust an anode that does not need to be adjusted.

Recall is simply the fraction of all **correctly predicted** “anodes too high or too low” (*TP)* over the total number of **actual** “anodes too high or too low” (correctly predicted *TP* and incorrectly predicted *FN),* as given in Equation (6).

(5)

(6)

Fundamentally, a trade-off always exists between precision and recall. The output of such a ML model being a probability, it is possible to adjust the threshold by which *TP, FP* and *FN* are categorized. This threshold adjustment will bias the model towards *more precision* or *more recall*. The correct adjustment depends very much on the specific application of the model. In some cases, we want *more precision* (for instance, to decide whether or not a patient needs a brain surgery – you don't want to do it if not really needed!) while there are also cases where we prefer *more recall* (for instance to detect early signs of cancer – you don't want to miss any!). Figure 13 illustrates that trade-off with three different thresholds applied to our model's probability output.

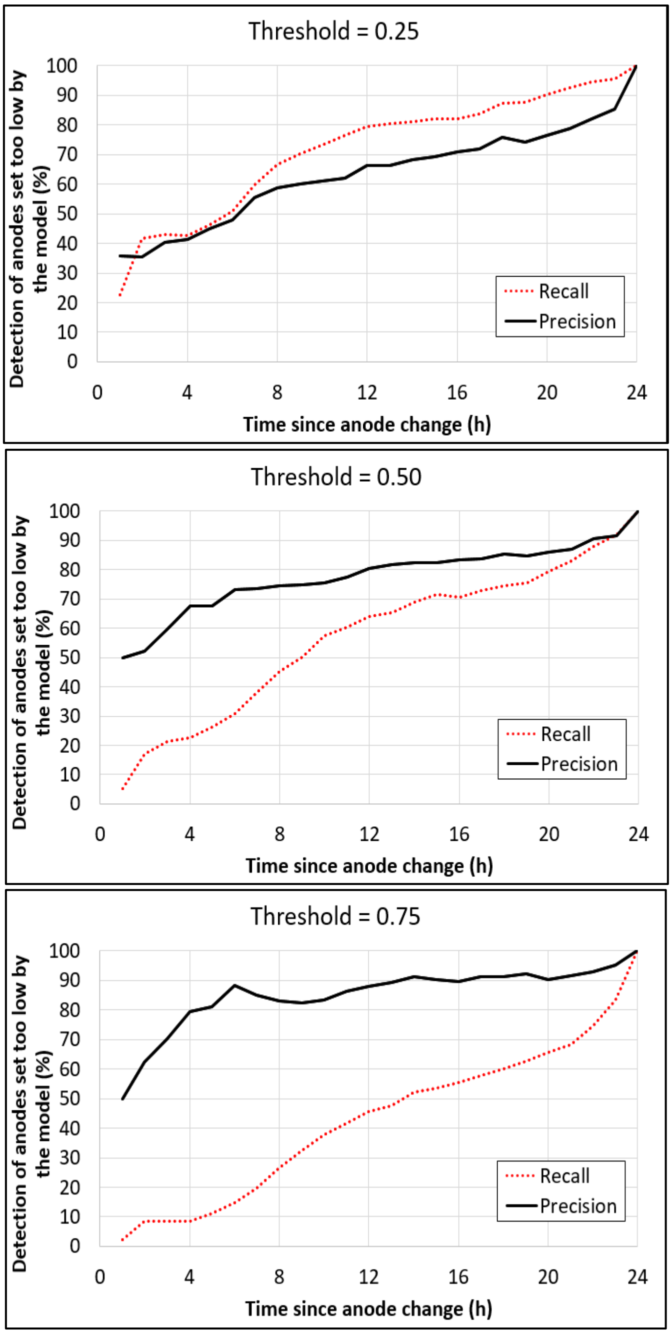

Figure 13. Three different probability thresholds applied to the model output.

Following the model validation, a prototype application was built in order to help cell operators visualize predictions and observe the result of their actions. Figure 14 is a screenshot taken from the user interface showing the target current curve in blue, the actual current in green, and the probability for the anode to be positioned *too high* or *too low* in the bar-chart at the bottom. It is currently available on cells equipped with anodic current sensors.

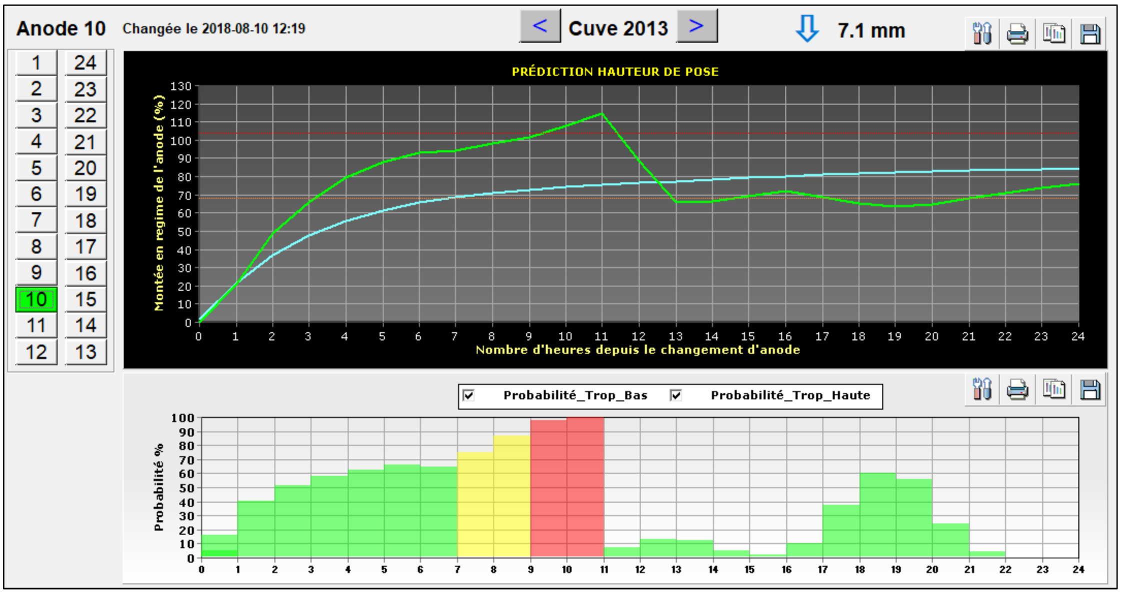

Figure 14. Example of a detection and correction after 9 – 10 h.

5. Conclusions

A model that detects incorrect anode positioning was presented. It uses the current pick-up after the anode change as an indicator of anode position and predicts the 24-hour old anode current, which is then compared to a target. The model was built using data from individual anode current sensors currently installed on industrial cells. These sensors are much more accurate in this context then the former approach using manual voltage drop measurements. These continuous measurements were a prerequisite to this work and could also be used in many other ways to tackle heterogeneity problems.

Machine learning was used to develop our model. Many approaches were tested and a final model was chosen based on its performance and maintainability. The chosen model, a normalized Radial Basis Function (RBF) network, automatically updates its references with recent information in order to adapt to changes. The maintainability of ML models is a potential pitfall of most machine learning algorithms.

The results are impressive; it is now possible to prevent a large proportion of incorrect positioning before it even affects the cell. A user interface was built to visualize the prediction in real-time and is currently used by cell operators where data from anodic current sensor are available.

6. References

1. Michel J. Leroy, Theodoros Pelekis, and Jean-Michel Jolas, Continuous measurement of current efficiency by mass spectrometry on a 280 kA prototype cell, *Light Metals* 1987, 177-180.
2. T.R. Alcorn, C.J. McMinn, and A.T. Tabereaux, Current efficiency in aluminum electrolysis by anode gas analysis, *Light Metals* 1988, 683-695.
3. R. Dorin and E. J. Frazer, Operational characteristics of laboratory scale alumina reduction cells with wettable cathodes, *Journal of Applied Electrochemistry* Vol.23 1993, 933-942.
4. Patrice Côté et al., Predicting instability and current efficiency of industrial cells, *Light Metals* 2017, 623-629.

5. Gary P. Tarcy, Current efficiency theory and practices for continuous improvement, *TMS 2012 Industrial Aluminum Electrolysis Course*, Saguenay, 2012.
6. Martin Segatz, Christian Droste, and Detlef Vogelsang, Magnetohydrodynamic effect of anode set pattern on cell performance, *Light Metals* 1997, 352-358.
7. Benoit Verreault and Renaud Santerre, Millivolt anodes – new technology opportunities for better pot control, *Proceedings of 35th International ICSOBA Conference,* 2 – 5 October 2017, Hamburg, Paper AL06, *Travaux* 46, 823-833.