Building of a digital model of aluminum electrolysis modes

E A Vereschagina¹, A L Zolkin², E A Lavrov³, D I Fakhertdinova⁴ and V V Dragulenko⁵

¹ Information security department, Far Eastern Federal University, Russky Island, Campus D, Vladivostok, Primorsky Kray 690922, Russia
² Computer and Information Sciences Department, Povolzhskiy State University of Telecommunications and Informatics, L.Tolstogo Street 23, Samara 443010, Russia; Natural Sciences Department, Private institution of higher education "Medical University" Reaviz, ChapaevskayaStreet227, Samara 443099, Russia
³ Computer science department, Sumy State University, Rymskogo-KorsakovaStreet2, Sumy 40007, Ukraine
⁴ Department of Natural Sciences, Service and Tourism, Kazan Cooperative Institute of the Russian University of Cooperation (KCI RUC), N.Ershova Street, 58, Kazan 420081, Russia; Department of Natural and Physics-mathematical Sciences, International preparatory school, Kazan (Volga region) Federal University (KFU), Kremlevskaya Street 18, Kazan 420008, Russia
⁵ Department of tractors, automobiles and technical mechanics, Federal State Budgetary Educational Institution of Higher Education “Kuban State Agrarian University named after I.T. Trubilin”, Kalinina Street13, Krasnodar 350044, Russia

E-mail: everesh@mail.ru

Abstract. The article discusses the issues of constructing a digital model of aluminum electrolysis modes. This model allows to achieve the specified criteria of the production process on the basis of scientific processing of historical data accumulated during the industrial production of aluminum at factories using a mathematical apparatus as well as to obtain recommendations on the values of control parameters. The study is aimed at achieving the business goals of the company's management in particular on increasing of the profits by improving the quality and quantity of products with reduced production costs.

1. Rationale
Metal production is divided into three main stages: extraction of bauxites (aluminum-containing ores), their processing into alumina (aluminum oxide), obtaining of the pure metal using the electrolysis process.

The main raw material for the electrolysis process is alumina (aluminum oxide Al2O3) which is a white friable powder. To create an environment in which this process will take place, one more component is needed and this component is cryolite. NaF:AIF3 is used as cryolite, the proportions of the components are called as the cryolite ratio, it has an important effect on the temperature of the process.
The aluminum electrolysis process takes place in electrolyte pots. The pot is filled with molten cryolite, which creates an electrolytic (conductive) environment at a temperature of about 950°C. The bottom of the pot plays the role of the cathode, and the coal blocks immersed in cryolite are the anode.

At set time intervals, a new portion of raw material is loaded into the pot using an automatic alumina supply system. Under the influence of an electric current, the bond between aluminum and oxygen is broken - aluminum is deposited on the bottom of the pod, forming a layer with a sickness of 10-15 cm, and oxygen combines with carbon, which is part of the anode blocks, and forms carbon dioxide [1].

Once a day, aluminum is removed from the pot using vacuum crucibles. In the electrolyte crust frozen on the surface of the pot, a hole is punched. The pipe is inserted into this hole. Liquid aluminum is sucked through it into the crucible. Before doing this the air is evacuated from the crucible. On average, about 2.5 tons of metal are pumped out from one pot, and about 7 tons of molten aluminum can be stored in one crucible. This crucible is then sent to the foundry.

During the production of each ton of aluminum, 280,000 m³ of gases are emitted. Therefore, each electrolysis cell, regardless of its design, is equipped with a gas collection system, which captures gases released during electrolysis and directs them to the gas treatment system [2]. At the described plant, alumina is used in the gas treatment system to trap harmful fluoride compounds. Before being used for the production of aluminum, alumina is involved in the purification of gases that have been formed earlier in the metal production process. Alumina is enriched with aluminum fluoride. And it reduces the consumption of cryolite.

Molten aluminum in crucibles is delivered to the foundry.

The research objectives are:

- preliminary research of the subject area and provided data;
- study of processes and algorithms of aluminum electrolysis;
- study of production systems that control electrolysis processes;
- development of a preliminary digital model of aluminum electrolysis.

2. Determination of control points for the technological process of aluminum electrolysis

The work identified the following points of control of the technological process, as well as the scheduled time frames of the process that are significant for solving the project's tasks.

The main point of impact on the technological process is the place of aluminum production itself, namely the aluminum electrolysis workshop. The shop includes 144 electrolyzers (pots) of the first stage and 144 electrolyzers (pots) of the second stage. The stages of pots differ from each other only by the manufacturer of electrolyzers and the terms of commissioning. The technological parameters of all electrolyzers are the same [3]. For the purposes of the study, data on the first stage pots have been used.

The main parameters affecting the aluminum production process are the current strength, the temperature inside the cell, the composition and amount of the electrolyte, and the composition of alumina (See table 1). Since it is not possible to directly influence the voltage supplied to the pot, as well as the quality of the feedstock, in order to obtain the maximum possible amount of aluminum of a certain quality, it is necessary to maintain the specified temperature of the technological process in the range of 951-956 degrees Celsius. Ideal temperature is 953 degrees Celsius. This temperature allows to perform optimal melting of the raw material and perform separation of the pure aluminum from impurities [4]. The required temperature is maintained by adjusting the electrolyte composition by means of weighed portions of fluoride salts.

Table 1. Process control points.

| Item No. | Impact                      | Frequency                                      |
|----------|-----------------------------|------------------------------------------------|
| 1        | Decision making and supply of alumina | Decided by the process engineer               |
| 2        | Decision making and supply of fluoride salts | Decided by the process engineer. Decision is made once per day, |
supply is carried out in equal portions during the day

| No. | Task Description                                      | Frequency   |
|-----|------------------------------------------------------|-------------|
| 3   | Temperature measurement                              | 1 day       |
| 4   | Level measurement                                    | 1 day       |
| 5   | Aluminum tapping                                     | 1 day       |
| 6   | DC supply                                            | Continuously|
| 7   | Provision of data on electrolyte composition         | 14 days     |
| 8   | Provision of data on the composition of aluminum in the pot | 7 days     |
| 9   | Enrichment of alumina                                | Continuously|
| 10  | Overhaul of pots and their start-up after overhaul   | 5.5 years   |
| 11  | Anodes replacement                                   | 20 days     |

3. Collection, processing and cleaning of extended data
Machine learning is a class of methods for data based automatic generation of predictive models. Learning from precedents, or inductive learning, is based on identifying general patterns from particular empirical data. Machine learning algorithms turn a dataset into a model.

There is no such thing as pure data. To be useful for machine learning, the data must be cleaned up. The following actions are performed during data cleaning:

- transformation of datasets into tabular form, where the columns will contain data values, grouped by attributes, and rows will contain data values, grouped by time;
- excluding columns with a lot of missing data;
- selection of columns to be used for forecasting;
- excluding rows with missed data in the remaining columns;
- correcting obvious typos and combining equivalent values;
- excluding rows with data that are out of required range.

Due to the fact that the data at the electrolysis plant are in different databases (DB), different departments and information systems are responsible for their maintenance, ETL has been used for data collection, processing and cleaning. ETL are enterprise-class methods that are used to lead to a single consolidated view and load data from several different accounting systems into a single database for further use.

The consequence of data fragmentation is the appearance of two types of inaccuracies in the data that must be taken into account and corrected:

- random errors arising at the level of input, data transfer, or due to bugs;
- differences in reference books and data detailing between adjacent information systems.

At the same time, if the first type can be corrected by cleaning the data, then the second type is for the most part not an error - controlled differences in the data structure, this is normal optimization for the purpose of a particular system. In order to consolidate such data special processing procedures shall be created.

In addition, it is necessary to carry out special data processing in order to obtain datasets that are ready to be used in models.

In the course of the work, the features of the data have been identified (in particular, the different discreteness of obtaining data in different sources).

Data preparation includes selection and consolidation of data into a temporary table for subsequent cleaning and compilation of a dataset for use in the model [5,6]. During data preparation, NaN values that appear during merging of different tables are removed.
In order to clean the data, confidence ranges for the data must be defined. Data that are out of confidence range are removed in order to avoid the penetration of outliers in the dataset. It is a standard practice used during preparation of samples for machine learning. There are no gaps in the database on the state of the pots, data on castings, weighed portions, weather and data on alumina supplies.

There are dozens of machine learning algorithms, ranging in complexity from linear regression and logistic regression to deep neural networks and ensembles (combinations of other models) [7].

For the purposes of this study, the following of the most common algorithms are suitable:

- linear regression;
- decision trees;
- random forest and bagging methods;
- boosting methods.

Machine learning algorithms are trained on data to find the best set of weights for each independent variable that affects the predicted value or class. The algorithms themselves have variables called hyperparameters. They are called hyperparameters because, unlike parameters, they control the operation of the algorithm, but are not determined by the weights [8].

In general, the optimization problem in this study can be formulated as follows. A finite set of precedents (objects, situations) is given, for each of them some data are collected (measured). The collection of all available case descriptions is called as the training sample. It is required, to identify general dependencies, patterns, relationships inherent using these particular data. It shall be done not only in this particular sample, but in general to all precedents, including those that have not yet been observed.

During build up of a digital model of aluminum electrolysis, the boosting algorithm will be mainly used on decision trees, since it is a fairly effective machine learning method that allows to get an acceptable result quite quickly [9]. Also, in the model for predicting major repairs, the running principle is used (the models make decisions independently of each other, and then the decision is made by weighted voting). This approach allows to keep the routine logic of making a decision on sending of the pot for overhaul. To determine the recommended weight, we will use optimization methods over a temperature prediction model built on boosting decision trees.

4. Working out hypotheses
In this work, hypotheses are put forward and tested on the factors that allow influencing the processes of optimizing aluminum production, increasing the volume of commercial output, reducing production costs and improving product quality:

- Anode wearing forecasting;
- Pots overhaul forecasting;
- Determination of pots with abnormal electricity indicators;
- AlF3 supply recommendation;
- Determination of pots with abnormal ALF3 flow rates depending on the service life;
- Reduction of the amount of aluminum that is involved into the reverse reaction;
- Aluminum yield forecast. Determination of controllable factors affecting the aluminum yield;
- Determination of the factors affecting the metal waving;
- Determination of factors affecting artifacts;
- Determination of controllable factors affecting the quality of aluminum.

Hypothesis 1. Anode wearing forecasting. Formulation: prediction of anode wearing can lead to more efficient replacement and savings due to increased anode life. In order to confirm the result, it is necessary to obtain a depreciation forecasting model.
Research status: rejected by process engineer.

Cause: Violation of the schedule of process work in terms of negative impact. It exceeds the economic effect. Changing the mode of anodes replacement entails big changes in the schedule of the plant.

Hypothesis 2. Pots overhaul forecasting. Formulation: prediction of the overhaul of pots will allow to send for overhaul, first of all, those pots that need it mostly. It lets pots to work as long as it is possible (at least six months).

Research status: further development is needed.

Model type: predictive model.

As a confirmation, it is planned to consider the statistics of overhauls, as well as a theoretical calculation of the amount of aluminum that the pots would produce if they worked more for at least a month. It allows to avoid situation when pots have been sent for overhaul without exhausting their resource to the end. Also, a saving factor will be the fact that the number of overhauls per year will be reduced.

Hypothesis 3. Determination of pots with abnormal electricity rates. Formulation: Identifying pots with abnormal electricity consumption can reduce overall electricity consumption.

Research status: completed.

Model type: descriptive model.

As a confirmation, it is planned to show a comparison of the graphs of pots with abnormal and normal energy consumption, on the basis of which the delta of unnecessarily spent energy can be calculated.

Hypothesis 4. AlF3 supply recommendation. Formulation: it is possible to optimize the AlF3 flow rate by stabilizing the process temperature.

Research status: further development is needed.

Model type: recommendation model.

It is planned to show a comparison of the graphs of the current and routine AlF3 consumption as well as energy consumption to calculate the economic effect.

Hypothesis 5. Determination of pots with abnormal AlF3 flow rates depending on the service life.

Formulation: there is a dependence of the AlF3 consumption on the time since the pot overhaul. As a result, a clear numerical dependence will be obtained.

Research status: further development is needed.

Model type: Descriptive model.

Research status: baths with different specific consumption of fluoride salt have been analyzed and divided into 4 categories. Each category corresponds to how the tapping and sample weight differ from the average:

- First grade - high tapping and low weighted material.
- Second grade - high weighted material and low tapping.
- Third grade - low tapping and low weighted material.
- Fourth class - high weighted material and high tapping.

Hypothesis 6. Reduction of the amount of aluminum that is involved into the reverse reaction.

Formulation: reducing of the amount of aluminum that is involved into the reverse reaction can increase the current efficiency of aluminum.

Research status: rejected.

Model type: recommendation model.

Result: temporarily rejected, since the degree of influence of the factor on the increase in the Al yield has not been studied, and there is an uncertainty in the calculation of the aluminum that reacts. It is proposed to introduce it as a factor in the hypothesis on dependence of yield on the current.
Hypothesis 7. Aluminum yield forecast. Determination of controllable factors affecting the aluminum yield. Formulation: the definition of controllable factors affecting the aluminum yield that can lead to an increase in the current yield.
   Model type: predictive model.
   Research status: suspended.

Hypothesis 8. Determination of factors affecting power consumption. Formulation: the definition of controllable factors affecting the power supply that can lead to an increase in the current yield.
   Model type: descriptive model.
   Research status: transformed, re-combined with sub-hypotheses.

Hypothesis 9. Determination of the factors affecting the metal waving. Formulation: the definition of controllable factors affecting waves on the metal that can lead to an increase in the current yield.
   Model type: Descriptive model.
   Research status: transformed, re-combined with sub-hypotheses.
   Results:
   - Data on the average duration of waves have been obtained.
   - Distributions of artifacts and waves values are obtained.
   - Information on the change in the mean values of waves in time has been obtained.

Hypothesis 10. Determination of factors affecting artifacts. Formulation: the definition of controllable factors affecting artifacts that can lead to an increase in the current yield.
   Model type: Descriptive model.
   Research status: transformed, re-combined with sub-hypotheses.
   Results:
   - Data on the average duration of waves have been obtained.
   - Distributions of artifacts and waves values are obtained.
   - Information on the change in the mean values of waves in time has been obtained.

Hypothesis 11. Determination of controllable factors affecting the quality of aluminum. Formulation: the definition of controllable factors affecting the quality of aluminum can increase the amount of A8 aluminum.
   Model type: predictive model.
   It is necessary to build a model predicting the grade of aluminum and determine the best input parameters.
   Research status: suspended by the decision of process engineers.

5. Findings
The following have been done in the work on building of a digital model of aluminum electrolysis modes: research aimed on study of the subject area has been carried out. Provided data have been studied. Processes and algorithms of aluminum electrolysis. The production systems that control the electrolysis process have been studied.
   Primary models have been identified and preliminary studies have been carried out for the selected hypotheses.
   The results at this stage make it possible to determine that the right direction is chosen. Thus, further areas of research will include:
   - elaboration of target hypotheses, definition of patterns;
   - development of control systems for instruments of influence, determination of control parameters;
   - models introduction into production processes.
References

[1] Mintsis M Ya, Polyakov P V and Sirazutdinov G A 2001 Aluminum electrometallurgy (Novosibirsk: Nauka) 368 p

[2] Isaeva L A and Polyakov P V 2000 Alumina in the production of aluminum by electrolysis method (Krasnoturyinsk: OJSC BAZ publishing house) 199 p

[3] Kolesnikov A S 2015 Kinetic investigations into the distillation of nonferrous metals during complex processing of waste of metallurgical industry Russian Journal of Non-Ferrous Metals 56(1) pp 1-5 https://doi.org/10.3103/S1067821215010113

[4] Kolesnikov A S, Sergeeva I V, Botabaev N E, Al’zhanova A Zh and Ashirbaev Kh A 2017 Chemical and phase transitions in oxidized manganese ore in the presence of carbon Steel in Translation 47 605-9 https://doi.org/10.3103/S0967091217090078

[5] Khudyakova T M, Kolesnikov A S, Zhakipbaev B E, Kenzhibaeva G S, Kutzhanova A N, Iztleuov G M, Zhanikulov N N, Kolesnikova O G and Mynbaeva E 2019 Optimization of Raw Material Mixes in Studying Mixed Cements and Their Physicomechanical Properties Refract Ind Ceram 60 pp 76–81 https://doi.org/10.1007/s11148-019-00312-2

[6] Taimasov B T, Sarsenbayev B K, Khudyakova T M, Kolesnikov A S and Zhanikulov N N 2017 Development and testing of low-energy intensive technology of receiving sulfate-resistant and road Portland cement Eurasian Chemico-Technological Journal 19(4) 347-55 DOI:https://doi.org/10.18321/ectj683

[7] Ling V V and Yumashev A V 2018 Estimation of worker encouragement system at industrial enterprise Espacios 39(28)

[8] Akhmetshin E M, Kovalenko K E, Mueller J E, Khakimov A K, Yumashev A V and Khairullina A D 2018 Freelancing as a type of entrepreneurship: Advantages, disadvantages and development prospects Journal of Entrepreneurship Education 21(2) https://www.abacademies.org/articles/Freelancing-as-a

[9] Yumashev A, Slusarczyk B, Kondrashev S and Mikhaylov A 2020 Global indicators of sustainable development: Evaluation of the influence of the human development index on consumption and quality of energy Energies 13(11) 2768. doi: 10.3390/en13112768 https://www.mdpi.com/1996-1073/13/11/2768