Application of Artificial Neural Networks to Forecast Technological Process Parameters in Aluminum Production*

Anton Mikhalev1[0000-0002-8986-5953], Nina Lugovaya1[0000-0002-2939-0298], and Tatiana Penkova2[0000-0002-0057-0535]

1 Siberian Federal University, 26, Kirenskogo str., Krasnoyarsk, 660074, Russia,
2 Institute of computational modelling of the Siberian Branch of the Russian Academy of Sciences, 50/44 Akademgorodok, Krasnoyarsk, 660036, Russia
asmikhalev@yandex.ru

Abstract. The study is aimed at methods of machine learning as it relates to forecasting technological process parameters. The forecasting tools are developed in two main stages: analysis and preprocessing of input data, elaboration of a math model and validation of the solution. Forecasting relies on recurrent neural networks. The method of maximum accuracy was used to elicit the neural network architecture, and calculate the metrics of MSE, MAPE, the coefficient of determination and Theil coefficient. The results obtained in the tests run on the suggested model of forecasting the cell voltage are deemed acceptable in terms of predicting the technological process indicators. The identified errors will ensure that preventive measures are taken in a timely manner to avoid process disruptions and increase overall efficiency of aluminum production.

Keywords: Neural Network, Forecasting, Process Disruptions, Technological Process Parameters, Voltage, Aluminum Production.

1 Introduction

Of all non-ferrous metal industries, aluminum production has the world's biggest share in manufacturing and consumption [1]. The industry develops within the lines of enhancing productivity of the main unit, electrolysis cell, therefore one of the key tasks is to control low-duty cells. Some of such cells are easily identifiable (shutdown cells, those under localized repairs), so they are controlled based on the current technical condition. Other are harder to identify, as deterioration in technology does not manifest itself directly and can only be determined through indirect parameters. Their number varies depending on supplied raw materials, occurring troubles, operational activities, etc., which may cumulatively lead to a greater number of cells operating at lower capacities and consequently to a considerable decrease in technical and economic indexes [2, 3]. Timely detection of errors in the technological process can be ensured in case the performance parameters of the complex of aluminum production are analyzed using modern intelligent technologies.

The technical condition of cells is controlled across a number of parameters that are continuously measured and stored in the data base of the computer-aided process control system: cell voltage, anode current, modes of automatic alumina consumption, adjustable anode block position. Making sure that these parameters are properly controlled and identified is critical in timely detection of process disruptions in the course of cell operation.

Values of parameters that need to be predicted are predominantly described as time series, that is, in sequences of values taken at certain instants of time. Forecasting time series normally entails using regression and autoregression methods, exponential smoothing, neural networks, etc. [4-5]. The forecasting model in this study is represented by artificial neural networks [6]. This technology has the following key strengths: solving problems with unknown patterns, resistance to noises in input data, and potential high-speed response. Neural network topologies are selected depending on the input data and type of tasks to be solved. This study looks at the application of artificial neural networks to forecast one of the most crucial among the controllable parameters – cell voltage. Recurrent neural networks (RNNs) were chosen for the purpose. The elements in RNNs form a directed graph which allows for processing series of events in time or consecutive spatial sequences. Unlike multilayer perceptrons, RNNs can use their internal memory to process variable length sequences of inputs.

The predictive tools are developed in two stages: 1) analysis and preprocessing of input data; 2) elaboration of a math model and validation of the solutions. The main body of the article is structured based on this logic. Section 2 spells out the objectives for time series forecasting. Section 3 describes the inputs. Section 4 elaborates on the applied methods of preprocessing of input data. Section 5 presents the result of selecting an optimal neural network architecture. Section 6 gives the results of voltage forecasting.

2 Research Objective

The aim of the time series forecasting is set as follows. Let us assume that the values of the time series are the following:

$$X = \{x(t), t \in T, x(t) \in R\}, T = \{1,2, \dots , N\} \quad (4)$$

where $x(t)$ is the value of the analyzed parameter registered at a given instant in time.

Based on the values of the analyzed parameter at preceding moments in time $x(t), x(t-1), x(t-2), \dots x(t-k+1), k \leq N$ we must predict (assess the values with highest precision) the analyzed parameter as it should appear at points in time $t+1, t+2, \dots, t+l$, i.e. build a sequence of forecasted values:

$$\hat{X} = \{\hat{x}(t+1), \hat{x}(t+2), \dots, \hat{x}(t+l)\} \tag{5}$$

To calculate the values in the time series at future moments in time, we must determine the functional relationship that shows the connection between the past and future values of this time:

$$\hat{x}(t+\tau) = \hat{f}\big(x(t-k+1), x(t-k+2), \dots, x(t+\tau-1)\big) \tag{6}$$

The presented functional relationship (3) represents the prediction model.

Therefore, the task of time series forecasting is fulfilled through creating a forecasting model that will satisfy the relevant criteria of forecasting quality control. Figure 1 illustrates the idea behind the objective of time series forecasting.

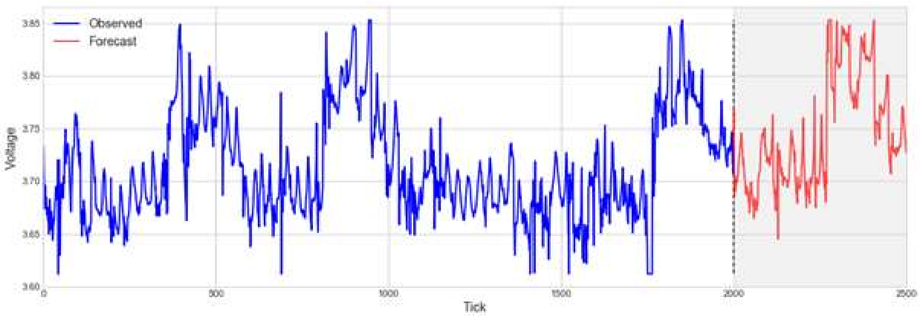

Fig. 1. Illustrated objective of time series forecasting.

Currently the accuracy of time series modelling is commonly estimated using the following two indicators:

– mean squared error, MSE:

$$MSE = \left[\frac{1}{l}\sum_{i=t+1}^{t+l}\big(x(i) - \hat{x}(i)\big)^2\right]^{\frac{1}{2}} \tag{7}$$

– mean absolute percentage error, MAPE, mean average percentage deviation (mean relative forecast error):

$$MAPE = n^{-1}\sum_{i=t+1}^{t+l}\frac{|x(i)-\hat{x}(i)|}{x(i)} * 100\% \tag{8}$$

In addition, apart from the given evaluation characteristics, this study estimates the accuracy of forecasts made to the elaborated prediction model using the coefficient of determination and Theil inequality coefficient:

– the coefficient of determination:

$$R^2 = \frac{\sum_{i=t+1}^{t+l}(\hat{x}(i)-\overline{x})^2}{\sum_{i=t+1}^{t+l}(x(t)-\overline{x})^2} \tag{9}$$

The coefficient of determination characterizes the strength of association of inputs and forecasts, so the closer it gets to 1, the better is the quality of the prediction model.

— Theil inequality coefficient:

$$v = \sqrt{\frac{\sum_{i=t+1}^{t+l}\left(x(i)-\hat{x}(i)\right)^2}{\sum_{i=t+1}^{t+l}\left(x(i)\right)^2+\sum_{i=t+1}^{t+l}\left(\hat{x}(i)\right)^2}} \tag{10}$$

The Theil index shows the strength of association in time series, so the closer it is to zero, the more strongly associated the series are that are compared.

3 Description of Inputs

The basic time series presents the data on the cell voltage registered by system detectors in the experimental area of the Khakas aluminum smelter. The voltage time series contain three-minute values of voltage for the period from January 3, 2020 to January 31, 2020. The time series parameters are demonstrated in Table 1.

Table 1. Parameters of voltage time series.

Cell	Series length	Mean value	Min value	Max value	Standard error
No.1	14320	3.737383	3.192000	4.023000	0.067460
No.2	14320	3.737383	3.192000	4.023000	0.067460
No.3	14320	3.692608	2.959000	4.198000	0.062725
No.4	14320	3.712393	0.000000	4.123000	0.107927
No.5	14320	3.702826	2.367000	4.469000	0.061439
No.6	14320	3.739327	3.483000	4.201000	0.081022
No.7	14320	3.701671	0.000000	4.224000	0.090101
No.8	14320	3.716235	0.000000	4.450000	0.083461

The overall sample volume contains about 115,000 entries. To set up the prediction model and evaluate the quality of the model itself, the sample volume was broken down into three parts: training (voltage at cells No.1-6), validating (voltage at cell No. 7), and testing (voltage at cell No.7).

4 Preprocessing of Inputs

The stage of building a prediction model is preceded by the stage of analysis and preprocessing of the time series. The preprocessing of the time series entails identifying outliers and smoothing the series. Certain discrepancies in the quality of measurements occur in various time series of data characterizing the production process. The outliers may be caused by technical errors in data collection, processing, and transfer.

Sifting out the outliers from the rest of data is a specific mechanism to identify and delete obvious discrepancies and other possible errors in inputs and make sure further forecasts are accurate. In the study, outliers were isolated by the isolation forest algorithm [7]. The isolation forest is a method to detect outliers that is mainly centered around constructing a forest of decision trees during training and forecast output. When it comes to detecting outliers, this method relies on the fact that outliers have values that are decidedly different from the norm and only make up a small proportion of the whole set of data. The results of detected outliers for voltage in cell No. 1 are presented in Figure 2.

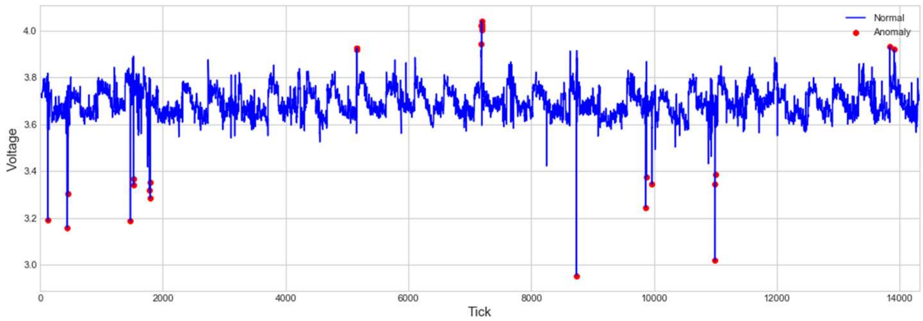

Fig. 2. Example of the isolation forest algorithm as it is applied to inputs.

Detected outliers are removed from the set and the resulting gaps in the data are recovered by the interpolation technique [8]. The view after the removal of outliers for the cell voltage can be seen in Figure 3.

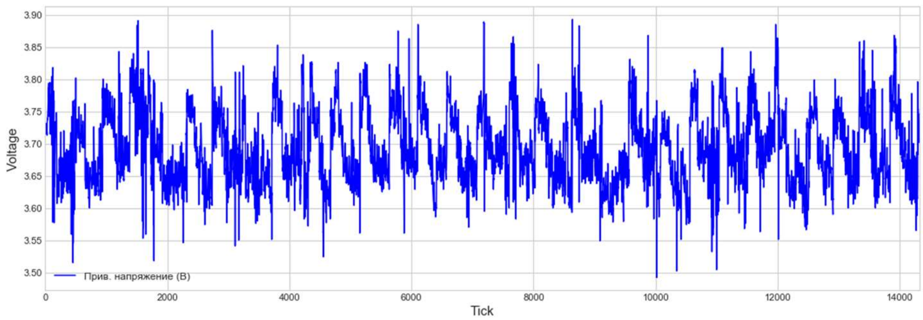

Fig. 3. Example of cleaned inputs for retention cell No. 1.

5 Selection of Neural Network Architecture

The efficiency of solving time series forecasting tasks that feature artificial neural networks is defined by their hyperparameters. The main hyperparameters underlying an artificial neural network are the number of layers and the number of neurons in each of the layers.

The neural network architecture was selected by iterating over the values of the number of layers/neurons. The number of LSTM-layers ranged from 1 to 3, whereas the number of neurons in each layer varied from 50 to 100 with the step size of 10. The Dropout technique was used to combat overfitting.

The results of the neural network architecture selection are presented in Table 2.

Table 2. Values of the forecast model quality evaluation characteristics for various neural network architectures.

Number of neurons/Number of layers	1	2	3
50	MSE 0.0003195597 MAPE 0.231096479 v 0.0034003583981 R^2 0.891027504246	MSE 0.000321653 MAPE 0.23555911 v 0.003410919610 R^2 0.89031356842	MSE 0.0003247026 MAPE 0.253191952 v 0.0034285759926 R^2 0.88927375558
60	MSE 0.0003215470 MAPE 0.228269917 v 0.0034105846637 R^2 0.890349819008	MSE 0.000335930 MAPE 0.25498155 v 0.0034843430721 R^2 0.885444884519	MSE 0.0003606315 MAPE 0.300183535 v 0.0036151698427 R^2 0.877021687371
70	MSE 0.0003359085 MAPE 0.261688514 v 0.0034879700818 R^2 0.885452433313	MSE 0.0003257920 MAPE 0.239746288 v 0.0034340883833 R^2 0.888902250687	MSE 0.0003255863 MAPE 0.237674537 v 0.0034330130443 R^2 0.888972387867
80	MSE 0.0003399340 MAPE 0.265887808 v 0.0035089207105 R^2 0.884079714165	MSE 0.0003398388 MAPE 0.256396719 v 0.0035081461294 R^2 0.884112165713	MSE 0.0003229711 MAPE 0.234390351 v 0.0034188288637 R^2 0.889864189599
90	MSE 0.0003275380 MAPE 0.238560829 v 0.0034412220481 R^2 0.888306851186	MSE 0.0003216582 MAPE 0.248235352 v 0.0034113558002 R^2 0.890311907914	MSE 0.0003421464 MAPE 0.263077962 v 0.0035161051010 R^2 0.883325278229
100	MSE 0.0003231217 MAPE 0.241065737 v 0.0034181660805 R^2 0.889812833874	MSE 0.0003437367 MAPE 0.277972490 v 0.0035288340033 R^2 0.882782972649	MSE 0.0003504917 MAPE 0.278388048 v 0.0035634811089 R^2 0.880479436508

The training data showed a similar result for all possible architectures. The eventually selected architecture consisted of 1 LSTM-layers with 50 neurons and one fully connected layer.

Other hyperparameters of the model were set using the random-walk method with cross-validation. The parameters for model construction were selected based on the principle of maximum accuracy (Table 3).

Table 3. Setting the hyperparameters of the model.

Parameter name	Description	Value
Optimizer	Parameter that shows how the model is updated based on inputs and the loss function	adam
Loss function	Parameter that measures the model accuracy during training	mean_absolute_error
Metrics	Parameter that is used to monitor training and testing of the model	accuracy
Number of epochs	Number of training algorithm runs across the entire set of training data	50
Mini-batch size	Number of sets that must be processed before the model parameters are updated	50

6 Forecasting Voltage

Process disruptions in the retention cell operation build up over time and undetected errors may spiral into serious accidents. Timely detection of deviation will entail long-term forecasting.

The long-term voltage forecasting is carried out through the iterative approach. The iterative approach in forecasting involves a few forecasting runs performed one step ahead, though using the values in the preceding stage. The general diagram of long-term forecasting is given in Figure 4.

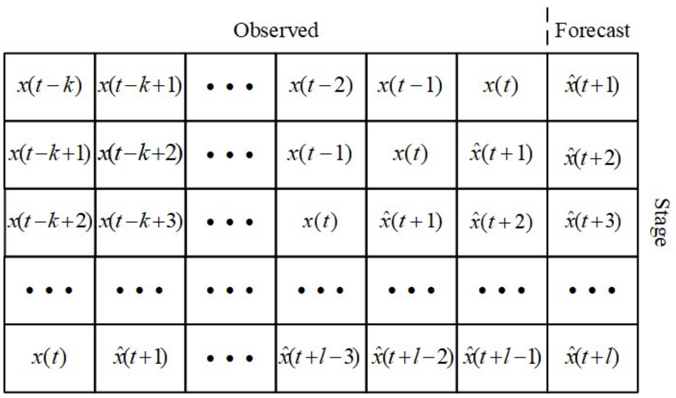

Fig. 4. Diagram of long-term forecasting.

The forecasting sequence is chosen to have a length of 10. The forecasts were fulfilled 10 steps ahead, which translates into 30 minutes. The forecasting results for the test sets are presented in Figure 5.

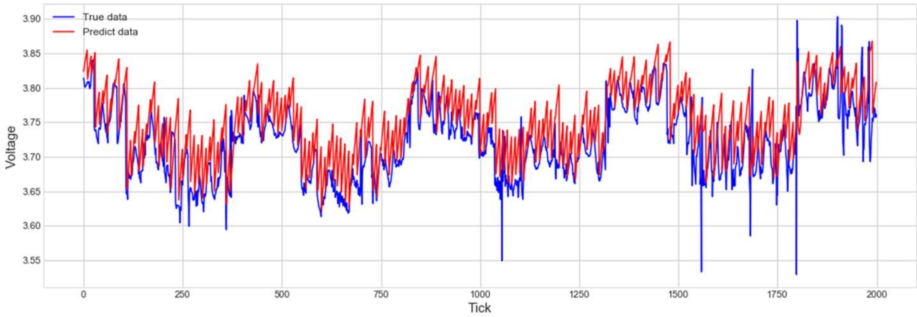

Fig. 1. Results of voltage forecasting.

It can be derived from the resulting graph that long-term forecasting performed with the iterative approach entails a value in every step that will differ from the real one, i.e. there will always be a certain error that will be growing with every new step. In its turn, the resulting prediction model makes it possible to reveal a tendency in how the controlled parameter is changing and identify the process disruptions in a timely manner.

7 Conclusion

The paper presents the results of artificial neural networks as they were applied to forecast the values of the technological process parameters in aluminum production. It looks at the mechanics of the prediction model construction aimed at one of the key controllable process parameters, namely the retention cell voltage. The elaboration of the forecasting tools is carried out in two main stages: analysis and preprocessing of inputs, construction of a math model and validation of the solution. Forecasting was chosen to be performed using recurrent neural networks. The method of maximum accuracy was used in the selection of an optimal neural network architecture, calculation of MSE, MAPE metrics, determination coefficient, and Theil coefficient. As it can be derived from the values of the selected metrics, the accuracy of the suggested model may be deemed appropriate.

The results obtained in the testing process are acceptable in terms of forecasting values of the process parameters. Timely detection of deviations in the forecasted parameter will allow for a quick response to prevent any process disruptions and thus increase aluminum production efficiency.

References

1. Abubakar, S.: Alyuminiyevaya promyshlennost' v sovremennom mire [The aluminum industry in the modern world]. Iinternational student research bulletin. **4-4**. 542–545 (2016)
2. Puzanov, I.I., Zavadyak, A.V., Klykov, V.A., Makeev, A.V., Plotnikov, V.N.: Continuous monitoring of information on anode current distribution as means of improving the process

of controlling and forecasting process disturbances. J. Sib. Fed. Univ. Eng. technol. **9(6)**. 788–801 (2016). doi: 10.17516/1999-494X-2016-9-6-788-801
3. Zavadyak, A.V., Puzanov, I.I., Tretyakov, Ya.A., Morozov, M.M., Makeev, A.V., Pianykh, A.A.: Mathematical modeling of the impact of anode bottom problems of the anode current distribution high current electrolyzer. J. Sib. Fed. Univ. Eng. technol. **10(7)**. 862–873 (2017). doi: 10.17516/1999-494X-2017-10-7-862-873
4. Montgomery, D.C., Jennings, C.L., Kulahci, M.: Introduction to Time Series Analysis and Forecasting. New Jersey: John Wiley and Sons (2008)
5. Hyndman, R.J., Athanasopoulos, G.: Forecasting: Principles and Practice. Australia: OTexts (2018)
6. Goodfellow, I., Bengio, Y., Courville, A.: Deep learning. Cambridge: MIT press (2016)
7. Liu, F.T., Ting, K.M., Zhou Z.-H.: Isolation forest. In: Proceedings of the 2008 Eighth IEEE International Conference on Data Mining. pp. 413–422 (2008)
8. Method for interpolating the Pandas library. https://pandas.pydata.org/pandasdocs/stable/reference/api/pandas.DataFrame.interpolate.html
9. Kolmykov, V.: The comparative analysis of the statistical model and neural network of the backpropagation in a forecasting problem. Applied Computer Science **6(30)**, 111–119 (2010) (in Russian)