# Key Indicators of Predicting the Aluminium Reduction Cells Based on a Transformer Model Driven by Multi-Source Time Series Data

**Qingtao Hu[1], Long Duan[2], Yanfang Zhang[3] and Qiaoyun Liu[4]**
1. Senior Engineer
Chalco Zunyi, Zunyi, China
2, 4. Engineers
3. Professor Level Senior Engineer
Zhengzhou Non-ferrous Metals Research Institute of Chalco (ZRI), Zhengzhou, China
Corresponding author: long_duan@chinalco.com.cn
https://doi.org/10.71659/icsoba2025-al034

**Abstract**

Aluminium electrolysis is regarded as a complex industrial process, and accurate prediction of its key indicators is crucial to optimize energy consumption and stabilize production. A key indicator prediction method based on a Transformer model is put forward in this research due to the fact that it is difficult to process the heterogeneous data and the dependency relationship of long-term time series data is mined insufficiently through traditional models. Through the integration of multi-source time series data of the aluminium reduction cell, the valid features of sequence data are captured based on the self-attention layer and multi-head attention mechanism of the Transformer architecture, and the synergistic effect of aluminium electrolysis process is mined, to ultimately build a key indicator prediction model for industrial scenarios. In the experiment, the Partial Least Square (PLS) and Convolutional Neural Network (CNN) models are compared based on the electrolytic cell data collected over three consecutive months of a smelter. The results indicate that the Transformer model has the optimal prediction accuracy, and that the average relative error is less than 3 % for the predicted molecular ratio and aluminium output of the test set. According to the results, the Transformer-based key indicator prediction model for electrolytic cells can be effectively implemented in production, supporting the process optimization and energy consumption management, and offering valuable insights into the intelligent upgrading of the aluminium electrolysis industry.

**Keywords:** Multi-source time series data, Attention mechanism, Indicator prediction, Transformer, Aluminium electrolysis process.

## 1. Introduction

As the digital economy becomes deeply integrated with the real economy, the traditional manufacturing industry is undergoing a paradigm shift from mechanized production to intelligent production. Integrated application of such new-generation information technologies as big data, artificial intelligence and the Industrial Internet has grown into a key engine for enhancing industrial production quality and driving efficiency improvements. Although a complete industrial system has been developed for the electrolytic aluminium, the industry still faces several key challenges in actual production and operation, including lagging dynamic assessment of key electrolytic cell parameters and reliance on experiential decision making for raw material feeding, which are restricting the precise regulation and energy efficiency enhancement during production.

Regarding process properties, the aluminium electrolysis process shows a strong multi-physics field coupling feature that, with the electrolytic cell as the core, complex electrochemical reactions occur among materials such as carbon anodes, alumina, electrolytes and catalysts under the coupling of the strong electric field, magnetic field, flow field and temperature field [1]. Harsh working conditions result in significant spatial-temporal heterogeneity of key process parameters, such as aluminium level, electrolysis temperature and material composition. It is hard to accurately collect all the parameters in real time by traditional sensor technologies, which further affects the dynamic assessment accuracy of key process indicators, including molecular ratio, current efficiency and aluminium output [2]. Regarding prediction modelling, most of the existing methods, which are mainly based on the linear mechanism assumption or the shallow machine learning framework, fail to capture the high-dimensional nonlinear association and spatio-temporal evolution law of the process parameters effectively, resulting in prominent lagging and systematic deviations in the predicted results [3, 4].

Accurate prediction of key indicators during aluminium electrolysis is crucial to optimize energy consumption and stabilize production [4, 5]. In researches on the prediction of aluminium electrolysis indicators and parameters, Yong Chen, et al. [3] have successfully predicted the feeding amount of aluminium fluoride through the deep belief network, Ziling Zhao, et al. [6] have predicted the molecular ratios by the least squares support vector machine and the autoregressive neural networks, and Xiaofeng Ni, et al. [7] have predicted the aluminium output based on the multilayer perceptron neural networks. These researches explored the complex relevance of aluminium electrolysis parameters, analysed the main factors affecting the aluminium electrolysis, and predicted the trends of related indicators. However, there are still opportunities for the enhancement of prediction accuracy. Moreover, the above researchers rely more on mass historical data and have not yet realized the rolling prediction of the indicators. They still need abundant historical references.

Due to the fact that it is difficult to process the heterogeneous data and the dependency of long-term time series data is mined insufficiently through traditional models, a Transformer-based [8] key indicator prediction method is put forward in this research, enabling accurate and prompt predictions of aluminium output and molecular ratio. Besides, a rolling prediction solution has been designed, which reduces the subjective human intervention and keeps the indicators and parameters at a reasonably optimal level. They support the process optimization and energy consumption improvement of electrolytic cells with data available, raise the energy consumption efficiency, and offer valuable insights into the intelligent upgrading of the aluminium electrolysis industry.

## 2. Data and Methods

### 2.1 Data Acquisition

This research collected data from 34 sets of 200 kA cell control systems and process reports over three consecutive months (from February 15, 2025 to May 15, 2025). Thirty-four electrolytic cells were in the same electrolysis shop. Data including the voltage, anode stroke, blanking amount, electrolysis temperature, molecular ratio and aluminium output were collected for each cell. Detailed data is shown in Table 1, and the data collected from the first three cells on February 15 is taken as an example.

### 2.2 Prediction Model

The partial least squares regression (PLSR) [9] is a multivariate statistical method integrating principal component analysis and regression analysis. The covariance of two groups of variables is maximized with this method by iteratively optimizing the latent variables of independent variable and dependent variable spaces, making the dimension reduction and multicollinear processing of data available. This algorithm balances the variable interpretability and the model simplicity through the latent variable projection mechanism, and completes the prediction modelling of complex data structures effectively.

**Table 1. Detailed data.**

| | Name | Description | Electrolytic cell #1 | Electrolytic cell #2 | Electrolytic cell #3 |
|---|---|---|---|---|---|
| 1 | Cell age (days) | Continuous operation days of electrolytic cell | 649 | 1013 | 649 |
| 2 | Feeding amount (kg) | Total amount of alumina fed on the same day | 3130 | 3263 | 3187 |
| 3 | Feeding frequency | Frequency of alumina feeding on the same day | 850 | 901 | 866 |
| 4 | Set voltage (V) | Voltage set on the same day | 3.94 | 3.945 | 3.94 |
| 5 | Operating voltage (V) | Average of operating voltages on the same day | 3.944 | 3.937 | 3.942 |
| 6 | Noise (mV) | Average of noises on the same day | 12 | 37 | 12 |
| 7 | Anode stroke/day (mm) | Carbon anode displacement on the same day | 12 | 13 | 11 |
| 8 | Anode stroke/t Al (mm) | Displacement of carbon anode to produce each tonne of aluminium on the same day | 8 | 7 | 9 |
| 9 | Electrolysis temperature (°C) | Average of aluminium electrolysis temperatures on the same day | 941 | 942 | 940 |
| 10 | Molecular ratio | Ratio of sodium fluoride to aluminium fluoride in electrolyte | 2.38 | 2.42 | 2.32 |
| 11 | Set blanking amount of fluoride salt (kg) | Feeding amount of fluoride salt set on the same day | 18 | 18 | 16 |
| 12 | Actual blanking amount of fluoride salt (kg) | Actual feeding amount of fluoride salt on the same day | 16 | 18 | 15 |
| 13 | Aluminium level (mm) | Height of molten aluminium in electrolytic cell | 210 | 200 | 200 |
| 14 | Electrolyte level (mm) | Electrolyte height of electrolytic cell | 180 | 170 | 180 |
| 15 | Aluminium output (kg) | Aluminium volume removed from electrolytic cell on the same day | 1580 | 1610 | 1550 |
| 16 | Fe content (%) | % of Fe in molten aluminium | 0.093 | 0.063 | 0.093 |
| 17 | Si content (%) | % of Si in molten aluminium | 0.056 | 0.042 | 0.052 |

The convolutional neural networks (CNN) [10] perform better than the shallow networks in terms of feature selection, feature extraction, prediction and classification. There are two advantages of CNN: (1) Local perception: CNN only perceives local data elements and then integrates such local information into a higher-level network to gain all the data information; (2) Weight sharing: CNN lowers the network complexity and reduces the weight. With a wide range of application, it is usually composed of six layers, i.e., input layer, convolutional structure, activation layer, pooling layer, fully connected layer and output layer. The CNN model built in this research is shown in Figure 1.

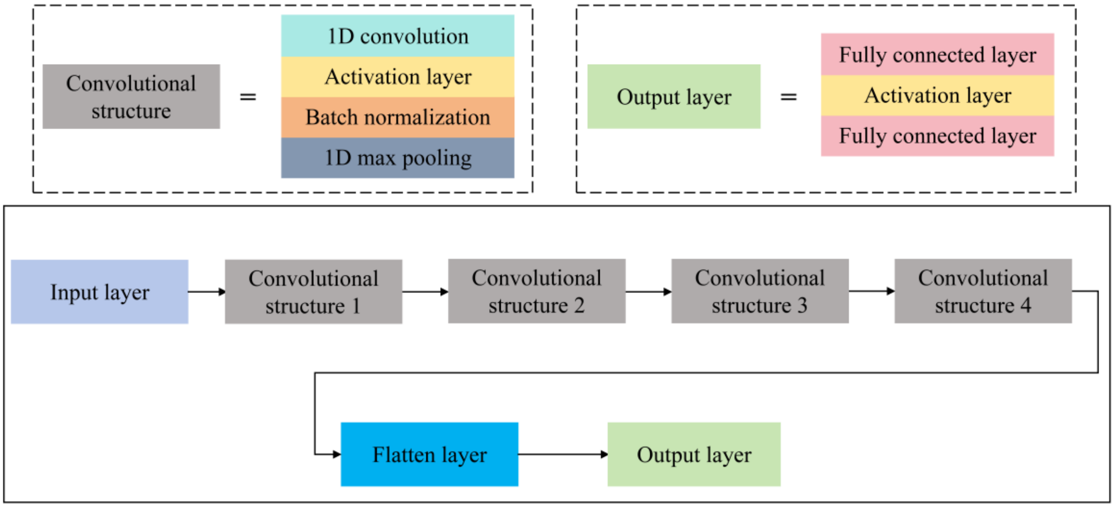


**Figure 1. CNN model structure.**

The Transformer model demonstrates its significant strengths in time series data processing by virtue of its unique self-attention mechanism and encoder-decoder architecture. Through the self-attention mechanism, this model dynamically calculates the global correlation weight at any position in the series, breaks the recursive limit of traditional recurrent neural networks, and directly captures the nonlinear relationships across time steps, such as voltage fluctuations and temperature variation, during the aluminium electrolysis process. With the multi-head attention mechanism adopted, this model allows multiple groups of attention heads to run in parallel, and focuses on the interactive features of indicators during the aluminium electrolysis at the same time. By combining the location encoding technology, this model integrates the periodic features of time series effectively, addressing the perceptual deficiency of self-attention mechanism for time sequences.

The structure of the Transformer model built in this research is shown in Figure 2. The entire model is composed of an encoder and a decoder. The encoder consists of input layer, location encoding layer and two feature encoding layers, while the decoder consists of input layer, two feature encoding layers and regression prediction layer. The encoder is connected to the decoder through the multi-head attention mechanism to fully explore significant features. The feature encoder contains residual structure, normalization structure, fully connected structure (feature extraction) and self-attention module, while the feature decoder contains residual structure, normalization structure, fully connected structure (feature extraction) as well as multi-head attention mechanism and self-attention mechanism modules. Feature information is extracted from multiple nodes in the previous series in the model training to predict the key indicators and parameters of the next node.

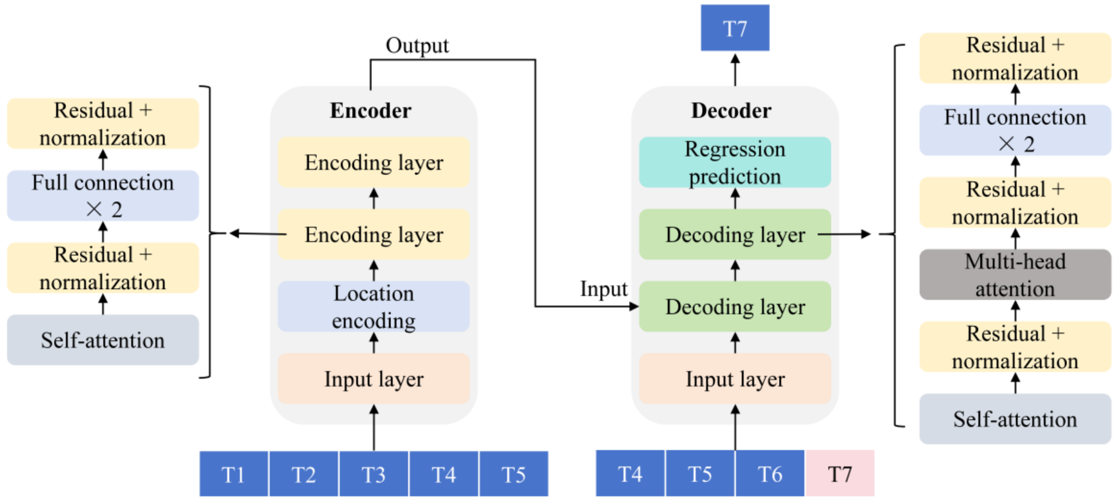


**Figure 2. Transformer model structure.**

### 2.3 Model Operation Plan

Actual production data is needed at the prediction modelling stage. The indicator prediction model becomes operational upon completion. The predicted indicators will substitute the values manually measured and assessed during modelling and are then entered in the prediction model, making the rolling prediction possible. The model operation process diagram is shown in Figure 3.

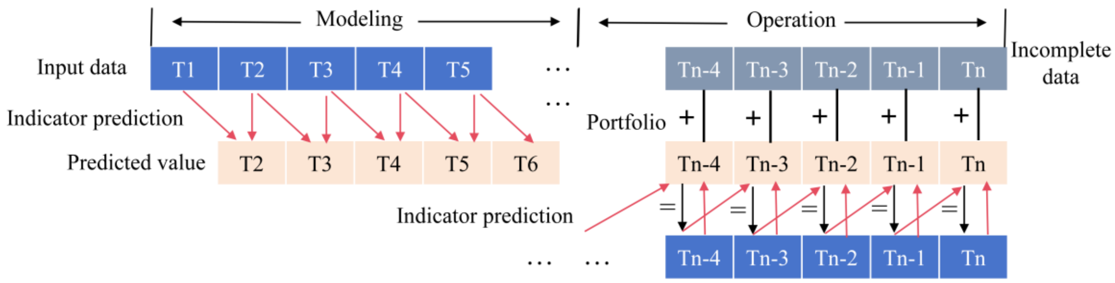


**Figure 3. Model operation plan.**

## 3. Experiment

### 3.1 Experimental Environment

The setup for all data processing and modelling is shown in Table 2.

### 3.2 Data Preprocessing

In this research, a time series data set is established for the predictions of molecular ratio and aluminium output during aluminium electrolysis, and method adopted is also applicable to the prediction of other indicators. A seven-day window is considered as the time series period at the time of establishment. That is, the data collected from seven consecutive days is regarded as a data sample, the molecular ratio and aluminium output on the 7$^{th}$ day will be predicted, and the data collected from the first six days serve as historical data. A completed data sample is composed of 118 data items: 1 cell number, 17 per day for the first six days, and 15 on the 7$^{th}$ day (excluding molecular ratio and aluminium output). The sample also contains two label values: molecular ratio and aluminium output. The data set diagram is shown in Figure 4.

**Table 2. Setup.**

| Component | Configuration |
|---|---|
| Operating system | Ubuntu 22.04 |
| Programming language | Python |
| Development framework | PyTorch |
| Processor | Intel i9-13900K |
| Graphics card | NVIDIA GeForce RTX 4090 D (24 GB) |
| Memory | 64 GB |

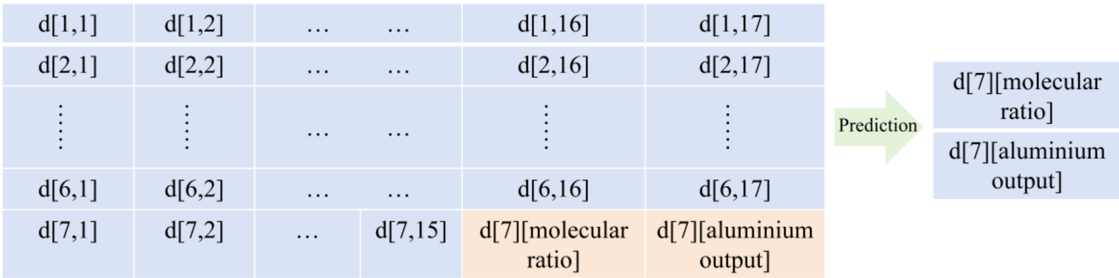


**Figure 4. Data set.**

As a large volume of data was collected from the experiment, some missing data items were identified at the time of data set establishment. Original data has been cleaned based on data information and prediction targets. Subsequently, two data sets were established: one for aluminium output prediction, containing 401 data samples, and another for molecular ratio prediction, containing 344 data samples. The completed data set is divided into training set, validation set and test set in the proportion of 3:1:1, among which the training set and validation set are collectively referred to as the modelling set. Data to be predicted include the molecular ratio and aluminium output at the next time node.

### 3.3 Modelling

In the process of building the key process indicator prediction model for aluminium electrolysis, the tolerance and the maximum number of iterations were respectively set as $10^{-6}$ and 5000 for the PLSR model. In addition, the number of principal components was optimized by means of grid search, with the optimization space of [2–50]. Hyper-parameters were set for the CNN and Transformer models in the following ways: optimizer: adaptive moment estimation (Adam); loss function: root-mean-square error; learning rate: $10^{-4}$; segment size: 16; and number of training rounds: 300. The predictions of molecular ratio and aluminium output were respectively modelled, and the prediction models were evaluated by average relative error (MRE) and mean absolute error (MAE). RE and AE were obtained from equations (1) and (2) separately.

$$MRE(d,d') = \frac{\sum_{i=1}^{n}\left|\frac{d_i - d'_i}{d_i}\right| \times 100\ \%}{n} \tag{1}$$

$$MAE(d,d') = \frac{\sum_{i=1}^{n}\left|d_i - d'_i\right|}{n} \tag{2}$$

where:

$d$ Set of true indicators

$d'$ Set of predicted indicators

$n$ Quantity of data items

## 4. Results and Discussion

### 4.1 Data Preprocessing Results

After the data set was established, in time sequence, the molecular ratio and aluminium output data were analysed every 7 days for the 34 electrolytic cells over three months, as shown in Figure 5. The averages of molecular ratios and aluminium outputs in three months are 2.42 and 1581.64 kg respectively, and both indicators are constantly fluctuating.

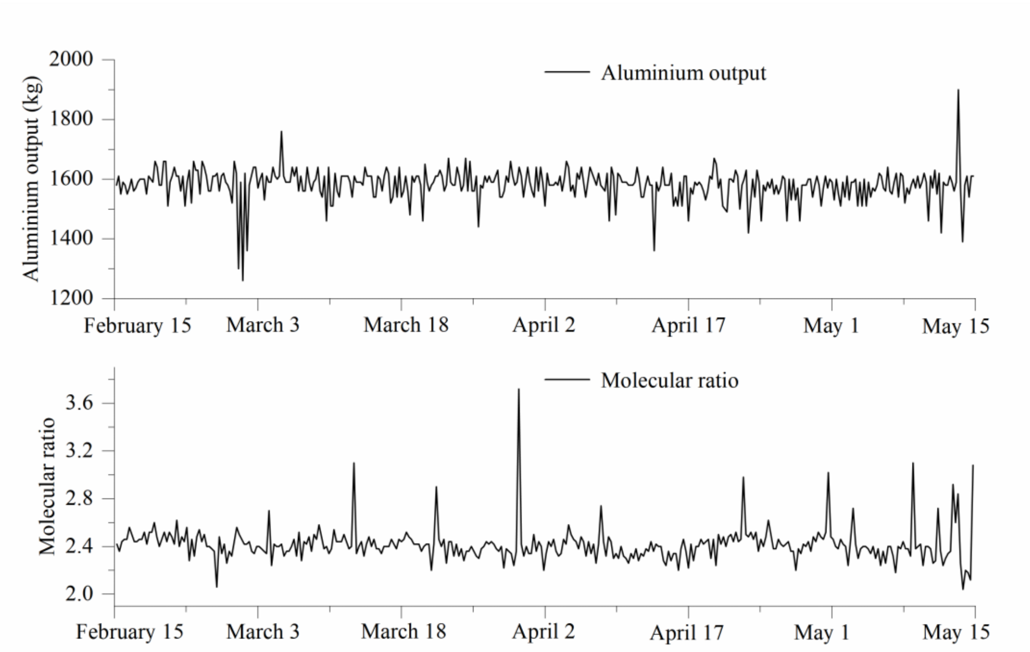


**Figure 5. Data distribution.**

### 4.2 Prediction Model Results

Prediction targets include the molecular ratio and aluminium output, each of which was separately modelled in the process. Table 3 and Table 4 respectively indicate the MRE and MAE for the test sets and modelling sets of the PLSR, CNN and Transformer models. The molecular ratios and aluminium output were predicted accurately through three models. The PLSR model had larger prediction error than the multi-layer network model, and the CNN and Transformer models had approximate prediction errors. However, the CNN model exhibited less stability when processing the long-term time series data, with greater error for prediction of aluminium output.

**Table 3. Relative errors of prediction model.**

| Prediction indicators | MRE (%) | | | | | |
|---|---|---|---|---|---|---|
| | PLSR | | CNN | | Transformer | |
| | Modelling set | Test set | Modelling set | Test set | Modelling set | Test set |
| Molecular ratio | 2.20 | 2.89 | 1.95 | 2.56 | 1.79 | 2.50 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Aluminium output | 1.46 | 1.49 | 1.24 | 1.50 | 1.27 | 1.40 |

**Table 4. Absolute errors of prediction model.**

| **Prediction indicators** | **MRE (%)** | | | | | |
|---|---|---|---|---|---|---|
| | **PLSR** | | **CNN** | | **Transformer** | |
| | Modelling set | Test set | Modelling set | Test set | Modelling set | Test set |
| Molecular ratio | 0.053 | 0.073 | 0.046 | 0.064 | 0.043 | 0.062 |
| Aluminium output (kg) | 22.456 | 23.674 | 19.077 | 23.790 | 19.540 | 22.130 |

The Transformer model is effective in exploring the nonlinear relationships across time steps and the correlation between indicators and parameters among the time series data. It has demonstrated the effectiveness of predicting the key indicators and parameters of aluminium electrolysis through the integration of multi-source time series data with Transformer, enabling timely and accurate understanding of electrolytic cell indicators and state. The Transformer-based molecular ratio and aluminium output predictions are shown in Figures 6 and 7.

According to the research results, although the independent samples in the modelling data set are taken from different electrolytic cells, or when the furnace bottom voltage drop, sedimentation, ledge thickness and cell age are different at that time, the key indicator prediction model established based on mass historical data by efficient modelling approaches performs well. The results have proven that multi-source data fusion and deep learning modelling will effectively reduce the influence of interfering factors and accurately extract associated features, making the key aluminium electrolysis indicator prediction possible.

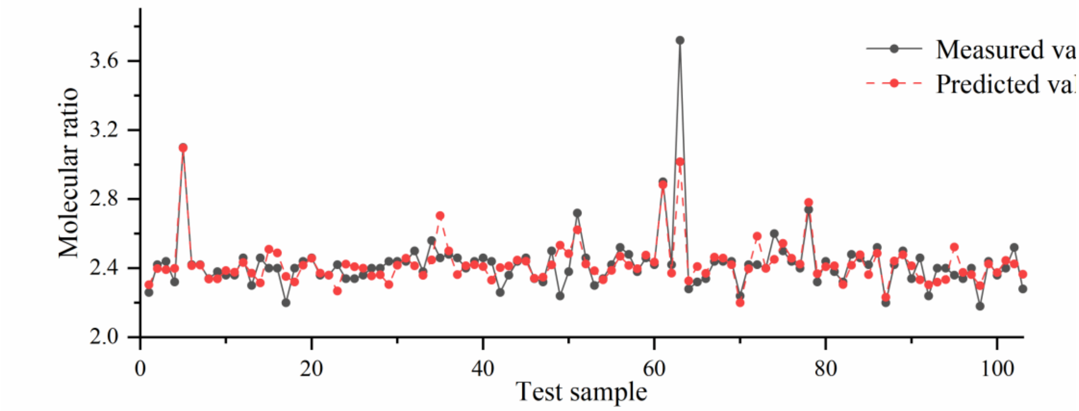


**Figure 6. Molecular ratio prediction results.**

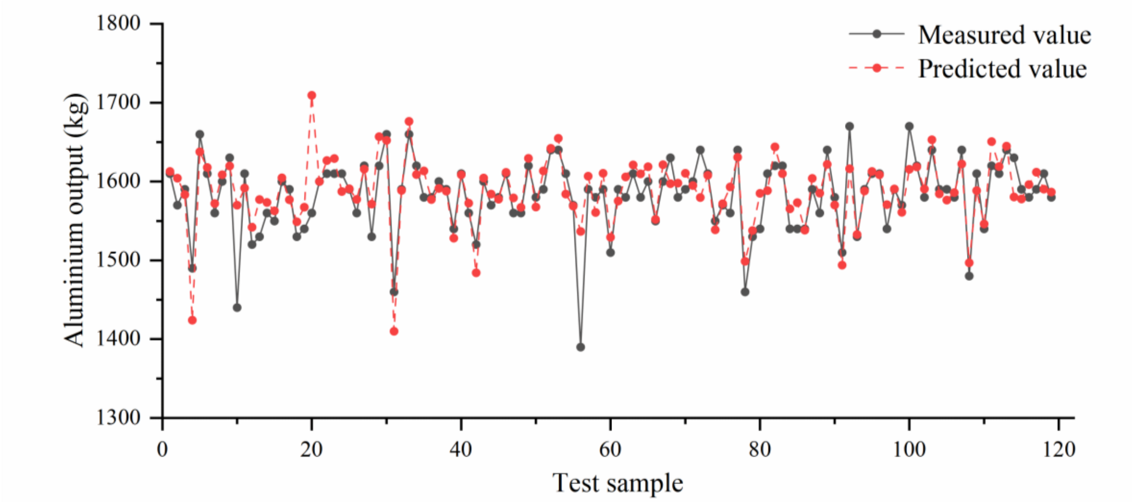


**Figure 7. Aluminium output prediction results.**

## 5. Conclusions

In this research, the multi-source time series data is collected from 34 electrolytic cells in a smelter over three consecutive months. After data pre-processing, a time series data set is designed and established. Prediction models were built with three machine learning frameworks, namely PLSR, CNN and Transformer, for two key indicators of aluminium electrolysis, i.e., molecular ratio and aluminium output. Effective features of sequence data are captured and the synergistic effect of aluminium electrolysis is explored by taking full advantage of the Transformer architecture's strengths in self-attention layer and multi-head attention mechanism, and fully exploring the dependency between multi-source data information and long-term time series data, to build the key indicator prediction model for industrial scenarios. The Transformer model has the best predictions of the two key indicators: for the molecular ratio prediction model, the relative error is 2.50 %, which is optimal, and the absolute error is 0.062; for the aluminium output prediction model, the relative error is 1.40 %, which is optimal, and the absolute error is 22.130 kg. According to the results, the Transformer-based key electrolytic cell indicator prediction model is effective in production and able to predict the aluminium output and molecular ratio accurately and promptly. What's more, it reduces subjective human intervention through the rolling prediction solution, keeping the indicators and parameters in a reasonably optimal level. It supports the process optimization and energy consumption improvement of electrolytic cells with data available, raises the energy consumption efficiency, and offers valuable insights into the intelligent upgrading of the aluminium electrolysis industry.

In the subsequent experimental research, there will be two extended topics:

1) More key indicators are added for timely and accurate prediction, and more fine-grained data is collected for more refined indicator predictions.
2) Visualization analysis is conducted on the prediction model so that effective features are selected and optimized regulation model is built to guide and regulate production "reversely".

## 6. References


1. Ning Kang, Production of Electrolytic Aluminium, Metallurgical Industry Press. 2015 (in Chinese).
2. Gang Yin, et al., Research on PID Search-Optimized CNN-LSTM-Attention Method for Predicting Electrolytic Temperature in Aluminium Reduction Cells, *Chinese Journal of Scientific Instrument*, 2025, 46(01): 324–337 (in Chinese).

3. Yong Chen, et al., Research and Application of Key Indicator Prediction for Aluminium Electrolysis, *Computer Engineering and Applications*, 2019, 55(12): 250–258 (in Chinese).
4. Wang J et al., Development of Data-knowledge-driven Predictive Model and Multi-objective Optimization for Intelligent optimal Control of Aluminium Electrolysis Process, *Engineering Applications of Artificial Intelligence*, 2024, 134108664.
5. Berg T E L et al., A Novel Hybrid Analysis and Modeling Approach Applied to Aluminium Electrolysis Process, *Journal of Process Control*, 2021, 105: 62–77.
6. Ziling Zhao, Jinhong Li, Research and Application of Molecular Ratio Prediction Algorithm in Aluminium Electrolysis, *Light Metals*, 2022, (01): 30–39 (in Chinese).
7. Xiaofeng Ni, Bin Cao, Prediction of Aluminium Output in Aluminium Reduction Cells Based on MLP Neural Network, *Intelligent Computer and Applications*, 2021, 11(08): 139–142 (in Chinese).
8. Vaswani A, et al., Attention Is All You Need, arXiv, 2017.
9. Hui Lu, New Methods for Predicting and Measuring Key Physicochemical Parameters of Complex Aluminium Electrolytes, University of Science and Technology Beijing, 2021 (in Chinese).
10. Yujie Chang, et al., Anode Effect Recognition Method Based on Feature Explanation CNN of Multivariate Time Series, *Light Metals*, 2023, (12): 18–24 (in Chinese).