Prediction of exhaust gas content of PTA oxidation units based on the improved parameter adaptive LSTM

Yongming Han,[1] Jiajun Huang,[1] Bowen Xu,[2] Tianxu Hao,[2] Min Liu,[3,4,*] Xingxing Zhang,[5] and Zhiqiang Geng[1,*]

[1]College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

[2]China Kunlun Contracting & Engineering CO., LTD, Beijing 100037, China

[3]Beijing Advanced Innovation Centerfor Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing 100029, China

[4]Department of Automation, Tsinghua University, Beijing 100084, China

[5]Sustainable energy research Center (SERC), Dalarna University, Falun 79188, Sweden

*Correspondence: lm@buct.edu.cn (M. L.); gengzhiqiang@mail.buct.edu.cn (Z. G.)

Received: October 12, 2024; Accepted: November 11, 2024; Published Online: November 18, 2024; https://doi.org/10.59717/j.xinn-energy.2024.100052

Citation: Han Y., Huang J., Xu B., et al., (2024). Prediction of exhaust gas content of PTA oxidation units based on the improved parameter adaptive LSTM. The Innovation Energy 1(4): 100052.

Dear Editor,

This letter presents an improved LSTM(Long Short-Term Memory) method based on parameter adaptation technology, which focuses on predicting the waste gas content in the oxidation unit of PTA(Purified Terephthalic Acid) to optimize energy consumption. The tail gas components that need continuous monitoring in the PTA oxidation unit mainly include key gases such as oxygen, carbon monoxide, and carbon dioxide. Particularly, the oxidation reaction of PX(P-Xylene) in the PTA oxidation unit consumes a significant amount of energy. Abnormal concentrations of these gases may lead to excessive energy consumption or decreased process efficiency. However, by accurately predicting abnormal changes in gas content ten minutes ahead of time, factory operators can dynamically adjust energy consumption in a timely manner, ensuring efficient energy utilization.

Traditionally, prediction models rely on manually determined autoregressive coefficients for prediction. However, in actual industrial environments, data characteristics vary significantly across different time periods, locations and operating conditions. If manual determination of autoregressive coefficients is required for each prediction task, it undoubtedly increases the complexity and cumbersome nature of the application. To address this challenge, this paper innovates on the basis of the original LSTM model by incorporating parameter constraint control mechanisms of the AIC(Akaike Information Criterion) and BIC(Bayesian Information Criterion),[1] achieving adaptive adjustment of autoregressive coefficients. This improvement enables the proposed model to more effectively adapt to datasets from different locations and operating conditions, significantly enhancing the model applicability and prediction accuracy.

The early warning system applying the method proposed in this paper has demonstrated remarkable effectiveness in optimizing energy use, reducing unnecessary energy consumption, and improving overall operational efficiency. Tests conducted on multiple gas-related datasets show that the prediction accuracy of this proposed method surpasses other methods, making it an effective tool for reducing energy consumption in PTA production.

The PTA oxidation device tail gas content data set is a univariate time series data set. For time series prediction problems, as the accumulation of time series data increases and the data dimension grows, its research methods are also continuously improved. There are currently four major categories of methods: traditional time series modeling methods, analysis methods based on time series decomposition, machine learning-based methods, and deep learning-based methods.

The traditional time series forecasting method is mainly based on determining the time series parameter model, solving the model parameters, and using the solved model to complete future prediction work. The ARIMA (Auto-Regressive Integrated Moving Averages) is a widely used statistical method for time series forecasting. There are many improved algorithms based on the ARIMA, such as the SARIMA(Seasonal Autoregressive Integrated Moving Average).[2]

The time series decomposition method has always been a very useful method in time series analysis, which believes that a time series is often the superposition or coupling of long-term trends, seasonal changes, cyclic fluctuations and irregular fluctuations. A typical time series decomposition method was Facebook's open source Prophet model, which decomposed the time series into the sum of trend items, seasonal items (weekly or monthly), holiday items, and noise items.

There is a close connection between the work of time series data forecasting and regression analysis in machine learning methods. Machine learning algorithms are mainly divided into the following categories according to their implementation methods: the SVR(Support Vector Regression),[3] the GBRT(Gradient Boosting Regression),[4] and the HMM(Hidden Markov Models).[5]

Deep neural networks construct past time series feature representations through multiple nonlinear layers to learn the internal change patterns of time series. For different application scenarios, the hidden layer in deep learning adopts different implementation forms, and thus uses different neural network architectures for implementation, which are mainly divided into the following categories: methods based on convolutional neural networks,[6] methods based on recurrent neural networks,[7] methods based on self-attention networks,[8] methods based on graph neural networks,[9] and methods based on residual fully connected networks.[10]

Previous related research has shown that forecasting speed and efficiency are evolving from traditional time series forecasting methods to deep learning methods, especially when combined with self-attention networks, thereby improving time series forecasting capabilities. However, actual prediction application scenarios place higher requirements on prediction methods. The proposed method in this letter can automatically determine the autoregressive coefficient based on various types of gas data, and can quickly adapt to changes in the data set caused by changes in point locations and gas types to be monitored. At the same time, the prediction effect of this proposed method has significant advantages on such small data sets.

For the prediction of the exhaust gas content of the PTA oxidation device, the parameter adaptive LSTM model can be used to predict effectively various exhaust gas data. Parameter adaptive LSTM model training includes four stages(Figure 1A): Stages 1-3 perform data preprocessing such as data cleaning, analyze the characteristics of the data set to determine adaptively autoregressive parameters, and input the results into the LSTM model for training. In the fourth stage, we verify the validity of the experiment by performing cross-comparisons with multiple models and datasets. A detailed analysis of these four stages is as follows:

Stage 1: The data preprocessing stage requires cleaning the original data and converting it into a data format suitable for model training. The original data is a text file transferred directly from the factory project. This file data has two characteristics. First, the data set contains dirty data with extremely special values. Second, the time series data of this data set is not regular. Therefore, after reading the text file, data cleaning to remove dirty data and interpolation processing to complete the time series data need to be performed. The final processed dataset with no dirty data and uniform time is the output of the first stage.

Stage 2: The parameter adaptive algorithm is used to analyze the data set output from stage 1 to obtain the regression parameters. The parameter adaptive algorithm proposed in this letter is a method for determining autoregressive parameters that fully utilizes the Akaike information criterion and the

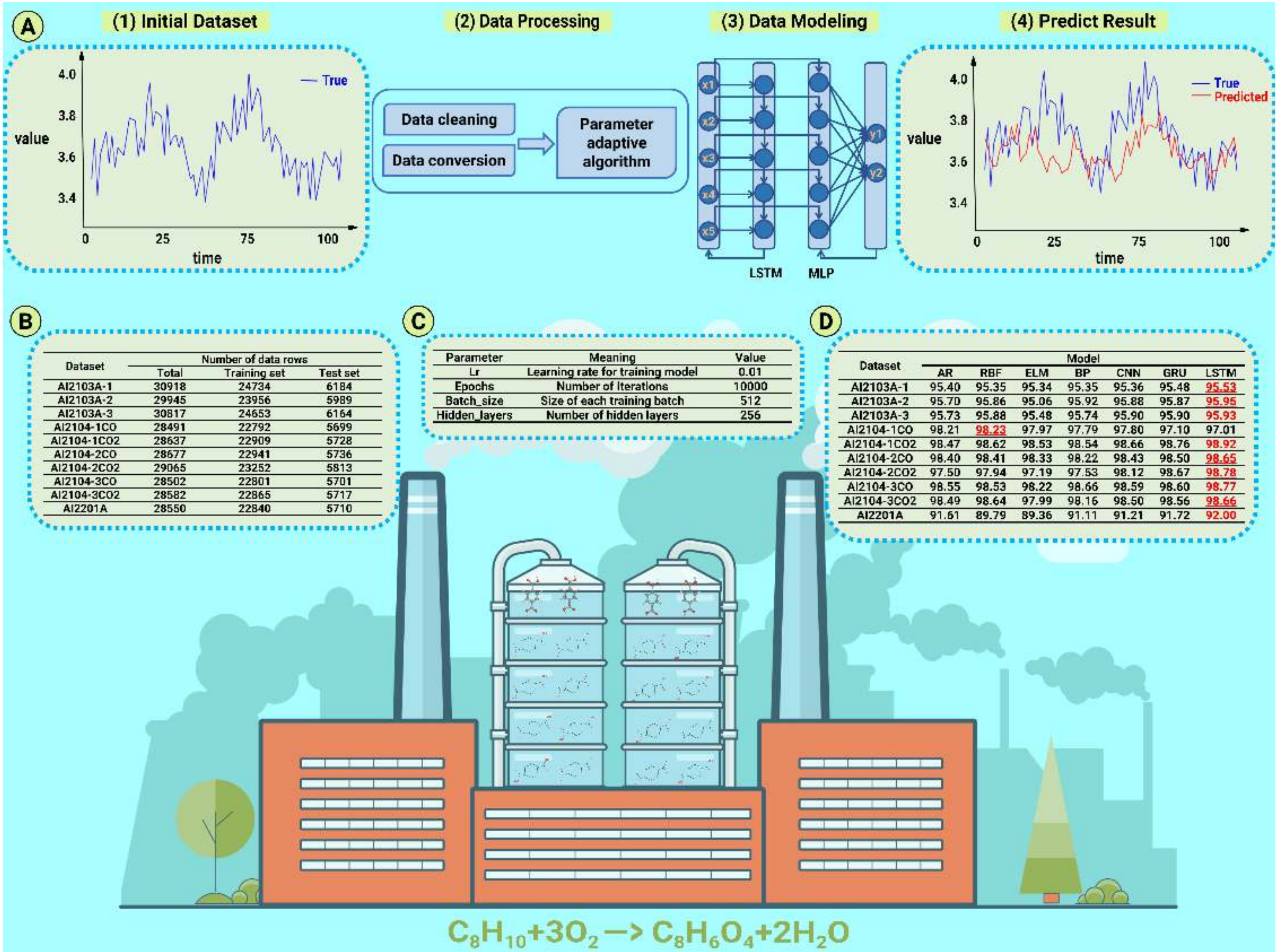

Figure 1. Relevant details of the proposed method (A) Process of the proposed method. (B) Summary of Ten Datasets. (C) LSTM's Experimental Parameters. (D) Accuracy of Each Model on Each Dataset.

Bayesian information criterion. Its principle is: in the selection of a certain range of autoregressive parameters, the values of the AIC and the BIC are calculated for the processed data set, and the CIC is calculated using a weighted method. Finally, the autoregressive parameter with the smallest CIC is selected as the most suitable parameter for the dataset.

$$AIC = -2 \times log(L) + 2 \times k \tag{1}$$

$$BIC = -2 \times log(L) + k \times log(n) \tag{2}$$

$$CIC = 0.5 \times AIC + 0.5 \times BIC \tag{3}$$

where L represents the likelihood function value of the model, indicating how well the model fits the data, k represents the number of parameters of the model (including intercept terms), and n is the sample size. When the CIC value is the smallest, the autoregressive parameter lag is the output of stage two.

Stage 3: The purpose of training the LSTM network is to mine the patterns existing in time series data.First, the LSTM network is built using the PyTorch framework. Then, the preprocessed data set is input into the LSTM network to train the network. In each training iteration, the real values and the feature values used for the prediction are input into the network, calculating the loss function, and updating and optimizing the parameters of the network through backpropagation.

Stage 4: In order to facilitate the quantification of experimental results, we use the accuracy as the evaluation index of the model. Multiple datasets are fed into the LSTM model and compared with the AR(Autoregressive), the BP(Back Propagation), the RBF(Radial Basis Function), the ELM(Extreme Learning Machines), the CNN (Convolutional Neural Network) and the RNN(Recurrent Neural Network) models.

$$ACC = 1 - \frac{|P - T|}{T} \tag{4}$$

where ACC represents the accuracy of the model; P represents the output value predicted by the model; T represents the true value at the time the model wants to predict.

Our experiments use an Intel Core i9-13900HX CPU @ 2.20GHz, NVIDIA GeForce RTX 4060 Laptop GPU, 16G RAM, and Windows 11 64bit. We use Python 3.7 and Pytorch 1.10.2 to implement the algorithmic idea.

The PTA is currently mainly produced by air oxidation of PX. The typical production process mainly includes a PX oxidation unit and a crude terephthalic acid hydrorefining unit. During this process, due to the occurrence of side reactions and incomplete reactions, the exhaust gas of the device will contain gases such as oxygen, carbon monoxide, and carbon dioxide. For the validity of the method proposed in this letter, we obtained 10 data sets from the PTA oxidation plant during its operation by recording the exhaust gas content of the plant. These 10 datasets are for three gases including oxygen, carbon monoxide and carbon dioxide at seven points including AI2103A-1, AI2103A-2, AI2103A-3, AI2104-1, AI2104-2, AI2104-3 and AI2201A(Figure

1B).

In order to illustrate the effectiveness of the proposed method, comparative experiments are conducted. By continuously adjusting parameters for testing, the suitable model parameters are obtained(Figure 1C). The results(Figure 1D) show that the prediction results of the LSTM model is the best among all models. This proposed model successfully solves the challenge of making effective time series predictions for different data sets.

In this letter, we propose an innovative improved LSTM method, which innovatively integrates a parameter adaptation algorithm with the LSTM model and utilizes the AIC and the BIC as guidelines to achieve optimal parameter selection for various classification tasks. The core improvement of this proposed method lies in its automated determination of the most suitable autoregressive parameter based on a comprehensive assessment of data characteristics and LSTM model parameters, a step that traditionally required manual adjustment.

Looking ahead, we plan to conduct in-depth comparative analyses of this research method with other advanced algorithms, such as the SOFNN-HPS(Self-Organizing Fuzzy Neural Network with Hierarchical Pruning Scheme), to reveal their respective strengths and limitations. Meanwhile, we are also committed to expanding the application of the proposed improved LSTM method in a wider range of industrial processes, aiming to explore its more profound research prospects and practical application value, thereby promoting technological advancements and application innovations in related fields.

REFERENCES

1. Zhang, Y., Li, R., and Tsai, CL.. (2010). Regularization Parameter Selections via Generalized Information Criterion. J. Am. Stat. Assoc. **105**(489): 312–323. DOI: 10.1198/jasa.2009.tm08013.
2. Vagropoulos, Stylianos I., et al. (2016). Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting. IEEE international energy conference. 1-6. DOI: 10.1109/ENERGYCON.2016.7514029.
3. Chen, C., Lu, N., Jiang, B., et al. (2021). A Risk-Averse Remaining Useful Life Estimation for Predictive Maintenance. IEEE/CAA J. Autom. Sinica. **8**(2): 412–422. DOI: 10.1109/JAS.2021.1003835.
4. Zuo, H., Yan, G., Lu, R., et al (2024) A Multi-Task Learning Based Runoff Forecasting Model for Multi-Scale Chaotic Hydrological Time Series. Water Res. Manag. **38**(2): 481-503. DOI: 10.1007/s11269-023-03681-z.
5. Zahari, A., and Jafreezal J.. (2015). A novel approach of hidden Markov model for time series forecasting. Proceedings of the 9th International Conference on Ubiquitous Information Manag. Commun. **91**: 1–5. DOI: 10.1145/2701126.2701179.
6. Ye, Z., Wang, N., Zhou, J., et al. (2024). Organic crystal structure prediction via coupled generative adversarial networks and graph convolutional networks. The Innovation. **5**(2): 100562. DOI:https://doi.org/10.1016/j.xinn.2023.100562.
7. Han, Y., Li, Z., Hu, X., et al. (2024). Novel Long Short-Term Memory model based on the attention mechanism for the leakage detection of water supply processes. IEEE Trans. Syst. Man. Cybern. **54**(5): 2786–2796. DOI: 10.1109/TSMC.2024.3350200.
8. Cai, H., Liu, X., Sun, L., et al. (2024). Battery internal short circuit diagnosis based on vision transformer without real data. The Innovation Energy **1**(3): 100041. DOI: 10.59717/j.xinn-energy.2024.100041.
9. Wu, Z., Pan, S., Long G., et al. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining. 753–763. DOI:10.1145/3394486.3403118.
10. Oreshkin, B. N., Amini, A., Coyle, L., et al. (2021). FC-GAGA: Fully connected gated graph architecture for spatio-temporal traffic forecasting. Proceedings of the AAAI conference on artificial intelligence. **35**(10): 9233–9241. DOI:arXiv:2007.15531. DOI: 10.1609/aaai.v35i10.17114.

FUNDING AND ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (62422303 and 62373035) and the Fundamental Research Funds for the Central Universities, China (buctrc202337). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Yongming Han: Supervision, Writing – review & editing, Methodology, Resources

Jiajun Huang: Experiment, Data curation, Methodology, Resources, Software, Writing - original draft

Bowen Xu: Data curation, Resources

Tianxu Hao: Data curation, Resources

Min Liu: Supervision, Writing – review & editing, Investigation, Methodology, Software, Writing - original draft

Zhiqiang Geng: Supervision, Funding acquisition, Writing – review & editing.

Xingxing Zhang: Software, Writing – review & editing.

DECLARATION OF INTERESTS

Xingxing Zhang serves as an Editorial Board Member for The Innovation Energy. The authors declare no other competing interests that could appear to have influenced the submitted work reported in this paper. This role did not influence the objectivity, peer review process, or editorial decisions regarding this manuscript.

DATA AND CODE AVAILABILITY

Data are available from the corresponding author upon reasonable request.