OPERATING OPTIMIZATION OF REFINING UNITS USING MACHINE LEARNING AND HEURISTIC ALGORITHM

FEI LUO[1], LIANG SUN[1], BEI ZHUAGN[1], NAN WANG[1], YAN-MING CAO[1], JUN XU[1]

[1] Industrial Chain Optimization Research Center, CNPC PetroChina Planning and Engineering Institute, Beijing 100083, China
E-MAIL: luofei5@petrochina.com.cn, sunliang_cppei@petrochina.com.cn, zhuangbei@petrochina.com.cn, wangnan0301@petrochina.com.cn, caoyanming@petrochina.com.cn, xujunme@petrochina.com.cn

Abstract:

The production and operation level of refining units significantly impacts the economic benefits of enterprises. However, due to their complex processes and numerous parameters, it is difficult to efficiently regulate them using traditional methods. This study combines production processes and data analysis to build regression models and optimization algorithms for optimizing the production and operation of these units. This technical route is applied to improve optimization objectives of continuous catalytic reforming unit (e.g. the octane barrel is increased by 0.44%). It can also recommend more optimal key parameters of production operation in real-time, which has guiding significance for actual production.

Keywords:

Refining units, Key parameters, Machine learning, Heuristic algorithm, Operating optimization

1. Introduction

In the field of energy and chemical engineering, the operational level of refining units is of utmost importance as it is related to corporate profits, as well as the supply and demand of products [1]. The production and operation level of these units depends on numerous parameters of production processes, and a large amount of data is generated during the operation. Moreover, data has become an important part of production factors and is driving refining enterprises forward at a rapid pace [2].

At present, some units still rely on experience or simulation calculations to adjust these parameters [3], [4]. However, problems such as low production and operation levels, high energy consumption, and unstable product quality still exist. Therefore, achieving production and operation optimization through data-driven methods and recommending better key production parameters in real-time is a necessary means to improve the production and operation level of these units [5].

2. Methodology

The technical route of this study is shown in Figure 1, mainly consisting of three parts: identification and analysis of key parameters, data acquisition and processing, and regression and optimization models.

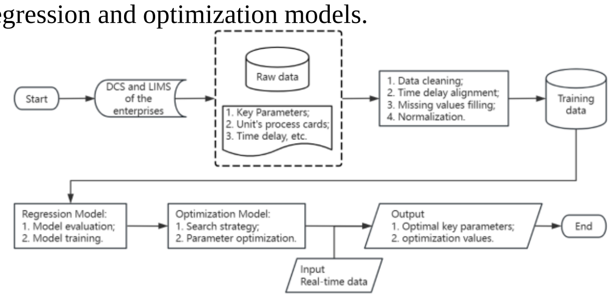

FIGURE 1. Flow chart of the methodology

2.1. Identification and Analysis of Key Parameters

Firstly, it is necessary to identify and analyze the key parameters that are most relevant to optimization objectives from complex processes and numerous parameters [6].

1) Parameter Selection: In this study, through on-site investigations of the units and discussions with experts, the optimization objectives for the production and operation of the units were determined. For example, the octane barrel is one of the most reasonable, scientific, and operationally optimizable major comprehensive technical indicators for the continuous catalytic reforming units. The key parameters affecting the optimization objectives were sorted out, such as feedstock properties, temperature, pressure, light oil ratio, and so on [7]. Among them, the selected key parameters have different meanings. Some are adjustable process parameters during production operation, some are used to calculate production operation objectives, and some are operating conditions of the unit and cannot be adjusted arbitrarily.

2) Correlation Analysis: For the selected key parameters, this study employed Pearson Correlation Coefficient and Maximum Information Coefficient to analyze the linear and non-linear correlations between the key parameters and the optimization objectives, so as to identify the parameters that are strongly correlated with the optimization objectives.

2.2. Data acquisition and processing

The data in this paper are collected from the Distributed Control System (DCS) and Laboratory Information Management System (LIMS) of the refining enterprise. The data types include real-time instrument data used to describe the operation status of the unit, as well as laboratory data of material properties and compositions. The data period ranges from 2022 to the present. In addition, it also includes unstructured data such as the time-delay relationships between each unit and optimization objectives, the calculation relationships between optimization objectives and parameters, and the process cards of the unit.

To ensure the accuracy of the algorithm model, data pre-processing is necessary [8]. Firstly, data cleaning is carried out. The raw data is analyzed in combination with process requirements. Outliers in the historical data are removed, and the data format is verified to eliminate non-numerical data. Then, according to the time-delay relationships between each unit in the unit's process flow and the optimization objectives, time-delay alignment is performed when merging the data, and then the missing values are filled. Finally, min-max normalization is carried out to eliminate the influence of the dimensions of different parameter units.

2.3. Regression and optimization model

In this study, the regression and optimization models are combined to form a method integrating machine learning and heuristic algorithm.

1) Regression Model: The regression model takes key parameters that affect optimization objectives as inputs and outputs production operation optimization objectives. The models include Linear Regression, K-Nearest Neighbors Regression, and Support Vector Regression. Considering the existence of certain non-linear relationships among the data, it also includes tree models such as Boosting model (GBRT, Adaboost), Bagging model (Random Forest), etc.

2) Optimization Model: The input of optimization model includes key parameters that affect the optimization objectives, adjustable parameters in the key parameters, upper and lower bounds of adjustable parameters, search step size of adjustable parameters, ideal values of optimization objectives, and the output are the values of recommended adjustable parameters and optimized objective. The search strategy of the model can be selected as Simulated Annealing Algorithm, Sequential Least Squares Programming, etc.

3. Results

3.1. Regression model training and results

In the training phase, the preprocessed dataset is divided into a training set and a test set at a ratio of 7:3. A 5-fold cross validation is performed on each regression model. The model with the highest average R2 score is selected as the regression model for this optimization objective. Then, all historical data is used for training, and the model parameters are saved for use in the optimization phase. The data pre-processing method used in this study can reduce the mean square error of the regression model by an average of 11.5%.

3.2. Optimization results of key parameters

In the optimization phase, the trained regression model is used as the objective function, the input key parameter values are used as the initial solution, and the upper and lower bounds and step size of the adjustable variables are used as constraint to optimize the adjustable variables. Among them, the upper and lower bounds of adjustable parameters and the search step size need to conform to the actual business situation. For example, if a temperature parameter is recommended to be adjusted from the current 528.08 ℃ to 528.09 ℃ after optimization, this change of 0.01 is unreasonable. Not only can the instrument not be so precise, but it is also smaller than the change caused by unit fluctuations.

The model of this study has been applied to refining enterprise and has achieved remarkable results in a continuous industrial validation test for 30 days on the continuous catalytic reforming unit of D Petrochemical Company.

As shown in Table 1, D Petrochemical Company has made 4 adjustment records of the continuous catalytic reforming unit after discussions, by referring to the real-time recommended optimal key parameters and the improved values of the optimization objective. Among them, "Rec. Time" is the time of the system recommends the key parameters, "Adj. Time" is the actual adjustment time of the unit by the enterprise, "Key Parm." is the key parameter to be adjusted, "Cur. Val." is the current actual value of the key parameter, "Rec. Val." is the recommended value of the key parameter, "Rec. Adj." is the adjustment value of the key parameter ("↑" and "↓" respectively represent the increase and decrease), "Act. Adj." is the actual adjustment value of the key parameter made by the enterprise, and "Obj. Pred." is the predicted value of the optimization objective under the recommendation of the key parameters.

TABLE 1. Adjustment records of the unit

Rec. Time	Adj. Time	Key Parm.	Cur. Val.	Rec. Val.	Rec. Adj.	Act. Adj.	Obj. Pred.
2025/2/19 10:00	2025/2/19 11:00	T01	526.8	526.1	↓ 0.7	↓ 1.0	91.43
		T03	526.9	526.2	↓ 0.7	↓ 1.0	
		T05	528.9	528.1	↓ 0.8	/	
		T07	529.0	528.0	↓ 1.0	/	
2025/2/26 8:00	2025/2/26 9:00	T01	526.6	526.0	↓ 0.6	/	91.49
		T03	525.9	525.3	↓ 0.6	/	
		T05	528.8	528.0	↓ 0.8	↓ 1.0	
		T07	528.9	528.1	↓ 0.8	↓ 1.0	
2025/3/11 14:00	2025/3/11 16:00	T01	525.2	524.5	↓ 0.7	↓ 1.0	90.93
		T03	525.1	524.4	↓ 0.7	↓ 1.0	
		T05	527.1	526.5	↓ 0.6	/	
		T07	527.0	526.4	↓ 0.6	/	
2025/3/14 8:00	2025/3/14 10:00	T01	524.0	523.3	↓ 0.7	↓ 1.0	91.29
		T03	524.0	523.3	↓ 0.7	↓ 1.0	
		T05	527.0	526.3	↓ 0.7	↓ 1.0	
		T07	527.0	526.3	↓ 0.7	↓ 1.0	

The algorithm optimized the octane barrel. At 2025/2/19 10:00, it recommended a set of adjustment values for the adjustable key parameters such as T01, T03, T05 and T07. Among them, the current value of T01 is 526.8 ℃, the recommended value is 526.1 ℃, recommending a decrease of 0.7 ℃. The current value of T03 is 526.9 ℃, the recommended value is 526.2 ℃, recommending a decrease of 0.7 ℃, and so on. It is predicted that the value of the octane barrel can be optimized from 91.24 to 91.32.

TABLE 2. The results after adjustment

Rec. Time	Adj. Time	Obj. Ture.	Obj. Pred.	Obj Adj.
2025/2/19 10:00	2025/2/19 11:00	91.24	91.43	91.32
2025/2/26 8:00	2025/2/26 9:00	91.19	91.49	91.32
2025/3/11 14:00	2025/3/11 16:00	90.82	90.93	90.83
2025/3/14 8:00	2025/3/14 10:00	91.15	91.29	91.22

Subsequently, based on the actual production situation, the enterprise adopted the adjustment recommendations for some key parameters of the unit. At 2025/2/19 11:00, the key parameters of T01 and T03 were decreased by 1°C respectively. After the adjustment of this set of operations was completed, the average value of the octane barrel within two hours after the unit stabilized was calculated to be 91.32 (Obj. Adj.), as shown in Table 2.

According to statistics, during the industrial validation period, the average actual value of the octane barrel was 91.17, an increase of 0.44 compared to the baseline period (last month's average) of 90.73.

4. Discussion

4.1. Interpretation of results

From the optimization results, it can be seen that the recommended adjustment scheme of key parameter is in line with the actual business situation. Firstly, the light oil component in the reforming reaction plays a crucial role. Its characteristics will directly affect the carbon deposition value and activity of the catalyst, thus affecting the reaction efficiency [9]. Secondly, the regulation of the reaction temperature is a crucial link in production optimization. Increasing the reaction temperature can lead to cracking reactions, resulting in an increase in hydrogen gas and affecting the yield of C_5+ products (hydrocarbons with 5 or more carbon atoms), causing resource waste and economic benefit losses [10].

Therefore, reasonably reducing the reaction temperature is a measure to optimize the production process, improve product quality, and enhance production efficiency. These aspects verify the reasonable of the results.

4.2. Limitations and uncertainties

Indeed, although this study has achieved positive results in the production and operation optimization, there are still limitations and uncertainties in the research. On the one hand, the accuracy of the model is limited by the distribution and the quality of the data [11]. Factors such as instrument precision, unit fluctuations, and data entry can all lead to incomplete or abnormal data. On the other hand, this study has only been successfully applied to a continuous catalytic reforming unit. There are uncertainties regarding whether it can be successfully applied to other units (such as hydrocracking units), and how to combine the optimization of the overall process flow with the optimization of unit operations in order to pursue the maximization of the overall profit of the refinery.

5. Conclusion

5.1. Study summary

This study is based on real data, forming a complete technical route from the analysis and selection of key parameters, data collection and processing, selection and training of regression models, search strategy of optimization model, to ultimately improving production operation objectives. Innovatively combining process and data to form effective method for processing key parameter data, and applying machine learning and heuristic algorithm to optimize the operation of refining units.

In addition, from 2025/2/14 0:00 to 2025/3/15 0:00, the time period during which the actual value of the octane barrel of the unit was higher than 90.73 accounted for 95.2% of the total time. This indicates that this method provides a practical and feasible solution for enhancing the production and operation level of refining units.

5.2. Future study directions

In the future, we will further conduct data analysis and processing by integrating the past operating conditions of the unit, so as to improve the quality of data collection and, in turn, enhance the prediction performance of the regression model [12]. Meanwhile, this method will also be industrially validated on other refining units. If successfully applied, we will also consider combining the overall process optimization with the unit operation optimization, shifting from focusing on local optimality of the unit to achieving the global optimality that maximizes the overall profit the refinery.

References

[1] Pedram Pasandide, and M. Rahmani, “Simulation and optimization of continuous catalytic reforming: Reducing energy cost and coke formation”, International Journal of Hydrogen Energy, Vol. 46, pp. 30005-30018, Aug. 2021.

[2] Angga W. Wibow, “Prediction analysis of greeting gestures based on data augmentation and LSTM”, Proceedings of ICMLC2024, Miyazaki, pp. 377-384, September 2024.

[3] Badiea S. Babaqi, and Mohd S. Takriff, “Yield and energy optimization of the continuous catalytic regeneration reforming process based particle swarm optimization”, Energy, Vol. 206, pp. 1-12, Sep. 2020.

[4] Badiea S. Babaqi, and Mohd S. Takriff, “Modeling and parameter tuning for continuous catalytic reforming of naphtha in an industrial reactor system”, Processes, Vol. 11, pp. 1-18, Sep. 2023.

[5] Nan Wang, “Dynamic personalized graph neural ordinary differential equation network for multivariate time series prediction of chemical processes”, 2024 IEEE 13th DDCLS Conference, Kaifeng, pp. 2044-2049, August 2024.

[6] Hossein Moeinzadeh, and K. Yong, “A critical analysis of parameter choices in water quality assessment”, Water Research, Vol. 258, pp. 1-10, Jul. 2024.

[7] Ahmet C. Serfidan, and M. Türkay, “Octane optimization with a combined machine learning and optimization approach”, Computer Aided Chemical Engineering, Vol. 50, pp. 221-226, Jul 2021.

[8] Badiea S. Babaqi, and Mohd S. Takriff, “Mathematical modeling, simulation, and analysis for predicting improvement opportunities in the continuous catalytic regeneration reforming process”, Chemical Engineering Research and Design, Vol 132, pp. 235-251, Apr. 2018.

[9] Chang. S. Hsu, Springer handbook of petroleum technology, Springer, Berlin, 2017.

[10] Janardhan L. Hodala, and Dong J. Moon, “Catalyst design for maximizing C_5+ yields during fischer-tropsch synthesis”, International Journal of Hydrogen Energy, Vol. 46, pp. 3289-3301, Jan. 2021.

[11] Mehdi Mohammadpoor, and F. Torabi, “Big data analytics in oil and gas industry: an emerging trend”, Petroleum, Vol. 6, pp. 321-328, Dec. 2020.

[12] Vijay Keswani, “Fair classification with partial feedback: an exploration-based data collection approach”, Proceedings of the 41st ICML, Vienna, pp. 23547-23576, Jul. 2024.