GAP ANALYSIS AND OPTIMIZATION OF HYDROPROCESSING PREDICTION MODEL

Nedyalko Petrov and Ivan Jordanov
School of Computing, University of Portsmouth, PO1 3HE, UK
Nedyalko.Petrov@port.ac.uk, Ivan.Jordanov@port.ac.uk

ABSTRACT
This paper discusses the process of gap identification, analysis and optimization of an existing hydrotreating process prediction model used in petroleum refineries. The performance of the model is investigated for a set of 16 specially selected crude oil feeds. Global optimization with genetic algorithm is conducted for a number of the model’s parameters. The simulation, testing and validation of the investigated model show improved prediction accuracy and efficiency. MATLAB® is used as a main working environment for this investigation. Most of the tasks are automated and included in a graphical user interface tool that can assist the company for further model analysis, optimization, testing and validation of currently used models.

KEY WORDS
Gap analysis, global optimization, hydrotreating, hydroprocessing, genetic algorithms.

1. Introduction

Petroleum companies undergo continuous research in the way they process different crudes, in order to meet the increased demand for cleaner products. This demand is dictated by more stringent environmental regulations that aim to reduce the pollution caused mainly by the sulphur and nitrogen components, present in these products [1, 2].

To meet the desired specifications, a number of technologies have been developed to refine and upgrade petroleum residues. Catalyst hydrotreating, as one of these technologies, is a key method for modern refineries [1, 3]. It has the capacity to increase the yield of distillates and to reduce the crude impurities, such as sulphur, nitrogen, metals, etc. [4]. Commonly, hydrotreating is carried out in reactors loaded with CoMo or NiMo catalysts [1, 5, 6].

The major issues with this technology are related to the catalyst deactivation, due to the presence of metals in the feed and also the severity of the reactions. The control of the reactor’s operation conditions (such as temperature, pressure, or space velocity, etc.) could have a drastic impact on the catalyst cycle length [7, 8, 9].

A good control of hydrotreating reactions environment is beneficial for keeping the balance between the product quality parameters and the rate of the catalyst deactivation (i.e. its cycle length). It is very important for the economics of the process – the investment made for the unit, hydrogen consumption, heating, cooling, pumping, and compression [1, 8, and 10].

Therefore, the optimisation of the hydrotreating efficiency is of a major interest for refineries [1]. To minimize the costs, time and effort required for experimental work, a mathematical prediction model could be of a great assistance. However, this is not a trivial task, due to the complex physical and chemical transformations that the feeds undergo, as well as the mass-transfer phenomena and the catalyst deactivation mechanisms in the reaction system [11].

Different approaches have been reported, whereas most of them deal with tickle-bed reactor models, applied to hydrotreating of middle distillates [12-15] and vacuum gas oils [16-18]. Only few models (developed by major petroleum companies) have been presented for heavy feeds. A broader summary of available models could be found in [2, 11, and 19].

In order to safely monitor and optimize the operation of their hydrotreating units, petroleum companies develop computer simulation models that allow them to better predict the effects that different feeds have on an apparent catalyst activity. These prediction models play an essential role in monitoring, optimizing, revamping and designing of new hydrotreating units.

As environmental regulations and market demands for clean fuel become more stringent, these companies face new challenges in operating their units. For example, the observation and critical analysis of many short catalyst life cycles of BP diesel hydrotreaters, while producing ultra low sulphur diesel (ULSD), have shown that variations in the feed quality can have a significant impact on the catalyst life. In this context, running distillates from troublesome crudes can shorten the catalyst life by 30%, causing significant financial losses.

To address these problems, BP has commissioned a comprehensive pilot plant study to investigate and simulate the impact of different feedstocks under the ULSD operation conditions.

This paper focuses on investigation, identification and analysis of the gaps between the measured pilot plant data and the predictions made by the current company’s hydprocessing mathematical model and tries to optimize and improve its accuracy and efficiency.
2. Problem Approach

As many repetitive tasks were to be performed on different hydroprocessing model versions using available plant data sets, a way of automating the model execution and data manipulation was needed. MATLAB® was chosen as a main working environment for this investigation because of its advanced analysis, optimization and visualization tools as well as for its rich high-level programming language.

In this context, the first tasks were to investigate how to link the existing hydroprocessing model to MATLAB, in order to import and export data, and how to set up various model parameters and run the model with some basic options within MATLAB.

Afterwards, gaps identification and analysis were performed on the investigated hydroprocessing model. Comparison between the predicted and desired outputs used was made. The available data was considered feed by feed, reactor by reactor and as a whole.

Next, several more advanced manipulation functions were implemented (such as run plans, feed data extraction, reactor data extraction, results visualization, exportation, etc.) in order to facilitate the model optimization process. Most of them were included in a graphical user interface (GUI) tool for easier user interaction.

Finally, model optimization incorporating genetic algorithms was performed for improving the model prediction capabilities.

3. Existing Hydroprocessing Model Characteristics

The investigated hydroprocessing model has been implemented in FORTRAN as a dynamic-link library (dll). The model input includes more than 250 parameters and produces more than 60 outputs. Some of those inputs are specific to the individual crude feeds, while others are related to specific processes in the refinery, they are about to be processed in, and the kinetics of the chemical reactions.

The current interface used for interaction with the model has been implemented in VBA (Visual Basic for Applications) and uses Microsoft Excel for receiving the model inputs and presenting the predicted results.

For the purposes of this paper and for preserving the confidentiality, the hydroprocessing model will be further discussed as a ‘black box’, without giving too much detail about the way it works.

This model has played an important role in monitoring, optimizing and designing new hydroprocessing units. However, its performance features some gaps in its prediction capabilities that need further investigation and analysis.

4. Model Gap Analysis

The characteristics of the products delivered by refineries should meet specific regulation criteria [1,2]. Although producing low sulphur and nitrogen species is highly desirable from environmental point of view, it has to trade off with very high manufacturing costs. Therefore, a good hydroprocessing prediction model is of great importance for the proper monitoring and control of the company’s hydroprocessing units.

4.1 Model Analysis Data

A pilot plant study was carried out in 4 parallel small reactors under three different conditions. It included 16 diverse feeds from different BP refineries across the globe. The selected feeds were covering most of the refinery diesel sources, such as straight run (SR) diesel feeds, light cycle gas oil (LCO), light coker gas oil (LCGO) and distillates form residue hydrocracking units (RHU). Four feeds were processed in each reactor and a reference feed was run between them, in order to monitor the catalyst activity decrease.

Observations for each condition were made and a number of relevant parameters were recorded. In this way, a data set containing information for 64 different experimental conditions was generated (its entries will be further referred to as condition points). Each condition point consisted of a vector of refinery measured data (such as reactor bed temperature, pressure, etc.), taken at a specific moment of time.

4.2 Model Analysis Approach

The above discussed data set was used for the creation of 64 vectors (describing the 64 recorded condition points), that were to be fed as inputs to the model. Some of the input vector’s elements were directly extracted from the measurements, while others were derived after an interpolation of the 3 condition points available for each feed. The MATLAB’s polynomial interpolation implementation was used to calculate these values.

After deriving the input vectors for the model, it was run and the predicted results were recorded. Analysis on the model performance was conducted for two of the most desired output parameters – the product sulphur and nitrogen components. The work was focused mainly on those two parameters because of their importance for the products’ quality.

The analysis was performed using MATLAB and its Global Optimization and Statistics Toolboxes. Graphics showing the relationship between the measured and the predicted results were created and critically analysed and compared (the comparison results are discussed in more detail in the next two subsections).
4.3 Sulphur Gap Analysis

The model’s sulphur prediction performance is shown in Figure 1. The x-axis represents the measured product sulphur values in parts per million (ppm) and the y-axis – the predicted sulphur values in ppm. The diagonal line shows where the desired results should be, i.e. where the condition points should lie if the model predicted sulphur values coincide with the measured ones. Each condition point is labelled according to its entrance in the database. However, no further details behind these points could be given considering the confidentiality of the data.

The results show that in general the model underestimates the product sulphur rates. The mean gap between the predicted and the measured product sulphur was 8.78 ppm, whereas the peak gaps could reach up to 50 ppm (or more than 50% of the measured values, see conditions 37 and 57 on Figure 1) for some of the troublesome feeds.

4.4 Nitrogen Gap Analysis

The nitrogen prediction performance results are shown in Figure 2. Analogically to the sulphur analysis case, the measured total product nitrogen is given on the x-axis and the predicted one – on the y-axis. The diagonal line represents a reference line where the condition points should lie in the ideal case.

The analysis of the results shows that the model has serious problems in the prediction of the total product nitrogen. A mean gap of 5.03 ppm was calculated, where the maximal discrepancy could reach up to 60 ppm (condition 37 on Figure 2) for some of the condition points. However, analysis also showed that most of the condition points lie on a line, indicating good optimization opportunities.

5. Model Optimization

The optimization was performed by tuning some of the model setup parameters, so that it could better predict the product sulphur and nitrogen values. As the logic incorporated in the model is complex and the functional relations are not necessary continuous, a genetic algorithm optimization approach was chosen, in order to aim for a global solution.

A group of 20 model setup parameters were chosen for evaluation. The inclusion of these parameters in the group was driven by their significance and influence on the model’s prediction performance, as well as on the chemical reactions expert knowledge incorporated in the model. The investigated parameters were related to the pre-exponential factors and activation energies for some important chemical reactions embedded in the model.

The process of optimization was performed in MATLAB® environment using its Global Optimization Toolbox functionality. It is discussed in more detail in the next four subsections.

5.1 Linking the model to MATLAB®

Two major approaches were considered for linking the hydrotreating model to MATLAB®. First, the possibility of using the already created Excel model interface was investigated. The process of importing data, exporting data and execution of Excel macro in MATLAB® was a relatively simple and easy task. However, the main disadvantage of this approach was the added loss of performance. The model’s execution time was of crucial importance, especially when hundreds of thousands test runs were to be evaluated. Therefore, another approach was necessary and functionality for calling the hydroprocessing model through its FORTRAN dll was
developed. This was not a straight forward process, as a direct calling of FORTRAN dlls is not supported by MATLAB®. To go around this issue, a header file in C++ style was defined for the dll and the C++ import procedure was followed. As a result of this approach, the execution time needed for evaluation of a single test condition was reduced over five times, when compared to the initial approach.

5.2 GUI Tool for Model Analysis and Optimization

For speeding up the model analysis, optimization, testing and validation, a MATLAB® based graphical user interface (GUI) tool was developed and implemented. It included functionality for defining, saving, loading feeds and reactors data; for defining, saving, loading and running of refinery reactor work plans; and for running optimizations and testing and validation of the optimized models.

Module for visualization and comparison of different model prediction capabilities was implemented, which also included option for exportation of the model comparison data and performance graphs for further investigation and analysis.

The developed GUI tool was used for deriving the results presented in this paper. A screenshot of its main window is shown in Figure 3.

Figure 3: Graphical User Interface Tool Screenshot

5.3 Model Optimization Approach

The optimization was performed on several stages, as some of the selected model setup parameters were not fully independent. As a result, running optimization for all the 20 parameters at the same time, was leading to identifying a number of local minima with insignificant differences in values. Also, although some optimization constraints were defined, many of these minima were not resulting in chemically sound combinations of the parameters. Therefore, the optimization was performed in groups of independent variables and considering some other model output parameters and chemical engineers’ expertise.

Global optimization, employing genetic algorithms was applied. The mean square error (MSE) between the measured data (M) and the one predicted by the model (P), over the \( n \) condition points was chosen as a fitness function for the optimization (1), in order to emphasize the reduction of the largest prediction gaps [20].

\[
F = \frac{1}{n} \sum_{i=1}^{n} (P_i - M_i)^2
\] (1)

Constraints for each of the optimized setup parameters were defined, so that they could take chemically reasonable values. Chromosomes with five genes, representing four sulphur and one nitrogen parameters were used. Starting with a population of 20 chromosomes, different selection techniques (stochastic uniform, reminder, roulette, and tournament), and reproductive operators - cross-over (scattered, single point, two points) and mutation (Gaussian, constraint dependent) were used for the genetic algorithm runs. The stopping condition was set to reaching 100 generations or a zero error (whichever first).

The optimization process was divided into 5 steps. First, the model was optimized for 8 setup parameters, related to the nitrogen predictions. This was done in two steps – in groups of four, based on the parameters nature and chemical engineers expertise. The nitrogen’s performance optimization was conducted before the sulphur one, because of the inhibition effects that nitrogen could have on the product sulphur values. Next, a two step optimization was performed for the sulphur related parameters – in groups of five. Again, the grouping of the parameters was decided based on their chemical nature. Finally, an optimization of two, related to both sulphur and nitrogen, parameters was conducted.

At each stage, the setup parameter values that were to be accepted for the optimized model were decided by refinery chemical experts, based on the model’s prediction performance and the chemical consistency of the combination of parameter values. This was necessary, because solutions with an insignificant difference in the value of the fitness function were sometimes resulting in chemically more sound combination of parameters.

5.4 Optimization Results

The performance of the optimized model was compared with the initial model. Its prediction capabilities were examined feed by feed, reactor by reactor and as a whole. A sample model improvement results for randomly chosen feed are shown in Figure 4. The blue line represents the measured sulphur data and the red and green lines – the predictions made by the initial and the optimized models, respectively. The reactor’s temperature (as one of the important inputs) is shown on the x-axis. The labels on the points denote the condition indices in the data set. It can be seen that the prediction gaps are now reduced for the optimized model and its predictions follow better the measurements’ trend. The mean gap for the feed was almost halved - from 6.38 ppm to 3.68 ppm.
The comparison results of the initial and the optimized models in predicting the product sulphur for all the feeds is shown in Figure 5. The predictions made by the optimized model are given with green markers. It could be seen that they are now shifted closer to the desired diagonal line. As a result, the mean difference for the sulphur predictions was reduced from 8.78 ppm of the initial model, to 4.60 ppm. Also, the model reliability was improved and the peak sulphur prediction gaps were closed from 50 ppm to up to 20 ppm.

Figure 5: Product Sulphur Optimization Results for all Feeds

Analogically, Figure 6 presents the model comparison results for the same randomly selected feed given in Figure 4, but this time for the product nitrogen predictions. It can be seen that the optimized model not only gives results closer to the measured values, but also follows better the trend of the product nitrogen change when varying the reactor inlet temperature. The improvement is more explicit for the lower temperature range, where the gap was closed from about 12 to 5 ppm. As a result, the mean difference was reduced from 4.96 ppm to 2.87 ppm.

Figure 6: Product Nitrogen Optimization Results for a Sample Feed

The optimization results of the product nitrogen predictions for all the feeds are given in Figure 7. It shows that the model predictions have moved closer to the reference line. As a result of the optimization, the mean discrepancy of the nitrogen predictions was reduced from 5.03 ppm to 2.66 and the peak gaps were closed from 60 ppm to up to 10 ppm.

Figure 7: Product Nitrogen Optimization Results for all Feeds
only on expert knowledge (without parameter optimization).

Although a straightforward comparison with other hydrotreating mathematical models is not possible, due to the different nature of the underlying computational apparatus, model input parameters, setup parameters, error measurements and test data (size and feed’s diversity), it has to be noted that the optimized model prediction results are in good agreement or even superior to those reported by other authors.

For example, in [20], the authors propose a catalyst evaluation methodology that allows apparent reactions kinetics and catalyst life to be predicted. Although no error metrics are provided, the reported results are very satisfactory from both pilot plant and commercial plant tests. Again, good agreement of the predicted values in the studied range, when compared to the experimental data, is achieved by the model, presented in [1]. In it, the reactor setup parameters are optimised following Levenberg-Marquardt’s algorithm for minimizing the sum of square of the residuals between the experimental data and model predictions. The same algorithm for obtaining the optimal set of kinetic parameters is employed with similar success in [11]. In [21] a hydrotreating model taking into account more reactor configuration properties is investigated and the reported predictions’ accuracy falls into a ±10% interval. Good modelling results are also achieved in [8, 22 and 23].

6. Conclusion

A model optimization approach for improving the accuracy and efficiency of an existing hydrotreating model is studied. Its prediction gaps are investigated and a set of model setup parameters are chosen for consideration. Global optimization for the selected parameters is performed using genetic algorithm on a specially composed diverse feedstock data set.

As a result, the model performance become more reliable as a whole and the peak differences between the predicted and the experimentally measured values are reduced up to about 20 ppm for the sulphur and up to 10 for the nitrogen. This would result in providing better opportunities for planning the operation of the company’s hydrotreating units and improving the economics of the process by reducing of the feed processing time, hydrogen consumption, catalyst deactivation, etc.

As a continuation of the work, further analysis and evaluation of the optimized model are to be conducted for a broader pilot and commercial plant data.

Finally, it has to be noted that the same or very similar approach could be considered for improving other prediction models, especially if they are relying only on expert knowledge.

References

[14] M. Bhaskar, G. Valavarasu, B. Sairam, K.S. Balaraman, K. Balu, Three-phase reactor model to simulate the performance of pilot-plant and industrial


