XLIII CILAMCE
21–25, November, 2022
FOZ DO IGUAÇU – PR – BRAZIL

MACHINE LEARNING APPLICATION TO ASSESS A PROCESS CONTROL OF A CATALYTIC CRACKING UNIT

Guilherme Lopes de Campos[1], Troner Assenheimer de Souza[1], Víctor Rolando Ruiz Ahón[1], Ninoska Isabel Bojorge-Ramirez[1]

[1]*Graduate Program in Chemical Engineering, Universidade Federal Fluminesce*
R. Passo da Pátria, 152-470,24210-240, Niterói - RJ, Brasil
guilherme_lcid.uff.br

Abstract.

Catalytic cracking is extensively applied in the downstream oil and gas industry to process an oil characterized by a high value of API degree, in other words, petroleum with a high percentage of heavy hydrocarbons, resulting in many products with smaller molecular weight. The biggest challenge of this process is the evaluation and control of catalyst deactivation, a phenomenon characterized by inducing the inactivity of disponible porous regions of the catalysts that take place the chemical species transfer or reagent to inside the porous, after the chemical reaction discharge to external medium products with a less molecular chain. One alternative workable in this industry is a continuous regular substitute of deactivated catalyst, using the chemical or pyrolysis reactivation and returning to the chemical process. To ensure the efficiency is crucial to determine the substitute frequency and the amount of reactivated catalyst to maintain the maximum yields of the process. To analyze aspects of catalyst deactivation, a system control project that aims to ensure the conversion obtained at the reactor and the flow of the reactivated catalyst is required. Since both variables are explicitly important to the problem, it is possible to define the optimal set-point for system control by monitoring these variables. So, a strategy based on ratio control, that uses mathematical modeling to obtain the set-point to the conversion and flow of the reactivated catalyst, classification as control and manipulate variables, respectively. Thus, get the suitable ratio in accord at the set-point of the process. As a way of evaluating and optimizing the flow of catalyst employed, after the projected control system, it will apply a model of machine learning, Support Vector Machine (SVM), a supervised method that with a data set prescribe a hyperplane and assess each point of distance on relation this plane for determining which class best representing the data. This work aims to study which variables influence the deactivation catalyst and suggest ways to mitigate and control them.

Keywords: Dynamic models, Process control, Machine Learning, FCC,SVM.

1 Introduction

Chemical process is extensively used nowadays, appling the natural and artificial resources to transform an inputs into product with high value aggregate. During these process many waste was generated for many reasons: Technical problems, bad inputs about cleanness, optimization at operation plant and knowledge about the phenomenal at each stage. Applied at a catalytic cracking unit the main challenge is to connect the conversion obtained at the reactor and the flow rate of fresh catalytic. These values have essential to optimization and control for getting the maximum yields and minimum amount of waste, including coke [1–3]. To the necessity to control and optimizetion two variables, conversion and flow rate of fresh catalyst, as appropriate dynamics models to have a preview understand temporally how to work this mechanism and obtained transfer functions to project a system control [4] . Simultaneously in this project the control process optimization parameters could be obtained at machine learning techniques, even as support vector machines (SVM) using a data set, It is possible to get the principal factors and parameters to analyse this process.

2 Dynamics and machine learning methods

The first step to modeling a dynamic system obtains the variables, constants and parameters included in phenomenon. Basically, the following steps to get a dynamic model:

- Define problem;
- Understand the phenomenon that would like study;
- List model simplifying assumptions;
- Apply mass and energy balance;
- Obtain the values of constants and parameters;
- Selection appropriate initial and boundary conditions;
- Represent using block diagram;
- Solve the models using numerical or analytical methods;
- Interpret the results and validated if experimental data are available.

Follow these steps to get a mathematical model of phenomena, the next step is to analyse if they correctly representing this conditions and assumptions listed before [2]. The machine learning steps change a little because the input is data and not relations between variables, briefly described by [5]:

- Extract the data;
- Applied the pre-processing data analysing: outliers, type of data, quality, amount, time available;
- Using method of data engineering, select the best variables to employ at this model;
- Select the model more appropriate to data type select at step data engineering;
- Solve the numerical methods;
- Interpret and validated model.

The present article it will apply both steps to understand FCC model.

2.1 Dynamic model to FCC

To model a process that involve chemical reaction, it will be important to know chemical species and stoichiometry coefficient to correct quantification. The reaction simply at the catalyst process as presented in equation 1. Where A is the heavy oil, B light oil and Coke the product that disables the porous of catalyst. The coke composition's high ratio between carbon and hydrogen [6–8]

$$A \rightarrow B + Coke \tag{1}$$

From the chemical reaction that represents coke formation, a mole balance was applied to heavy oil in the conversion function, considering the hypotheses [9–11] :

- Stoichiometry coefficients are one;
- Kinetic reaction of first order;
- Deactivation by coke formation, express to equation 2.The function is multiplied at the rate of reaction and related to the coke concentration and time, with parameters α_2 and A, explained in table 1, [1]

$$a(t) = \frac{1}{1 + \alpha_2 A t^n} \tag{2}$$

- Isothermal model;
- Physical properties constants;
- There is not feed of light oil and coke at chemical reactor;
- The final model is classified as a differential equation of first order non linear, parameters constant and non homogeneous.

Through mole balance apply at light oil and express at conversion function considering the hypotheses list at paragraph 2.1, it was obtained the dynamics model of FCC , equation 3:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -\frac{1}{\tau}(1 + X) + k_{reac} a(t)(1 - X), X(0) = 0 \tag{3}$$

To project a ratio control it was obtained the transfer function relate the conversion and catalytic activity using Laplace transform at deviation variables present at equation 4, classified as first order system control [12]. The expression of gain and characteristic time are also available with values list at table 1.

$$\frac{X(s)}{A(s)} = \frac{K_1}{\tau_1 s + 1}, K_1 = \frac{k\tau}{1 + k\tau}, \tau_1 = \frac{\tau}{1 + k\tau}, \frac{X(s)}{A(s)} = \frac{0.94}{120s + 1} \tag{4}$$

Table 1 shows parameters and variables used in the dynamic model to get a numerical solution using the medium order method to solve differential equations.

Table 1. Parameters dynamic model

Variable	Meaning	Value
X	Conversion [%]	Dependent variable
t	Time [s]	Independent variable
K_1	Gain first order dynamic system [-]	0.94
τ	Residence time [s]	2000
k	Kinetic constant $[s^{-1}]$	7.85×10^{-3}
a	Deactivation catalyst function	Function relate by time
α_2	Parameter of % coke in petroleum (varies between 0.42 - 0.88) [-]	0.44
A	Express the kinetic coke formation	0.05
n	Constant that related time and coke formation	0.54

2.2 Machine Learning - SVM

Support vector machine is a supervised learning method to classification and regression analysis used at many areas of knowledge. The principle is to employ a data set and define a hyperplane with a maximum margin by weighting each variable present in the set. After defining this references, calculate the distance between each point relative to the hyperplane these difference is used to classify the data [13]. The article it will apply the Lagrange polynomial to classify the variables that must be important to the deactivation catalyst.

3 Results

3.1 Dynamic model results

The solution of dynamic model was obtained solve numerically (Bogacki-Shampine method) equations 3 at Simulink™ (R2022a) with 1800 seconds. Figure 1 was the model project. It was divided into two parties: Dynamic model and ratio control model:

- Dynamic model: Solve the nonlinear differential equation numerically at mole balance to light oil considering deactivation by coke formation. To project initially selected the block for mathematical function and order to transfer signal, obtained the maximum conversion approximately 93%, this values agree with thermodynamics principles and deactivation. To project a ratio control was used the value of conversion calculated by model and flow rate of fresh catalyst. These flow was stipulate using correlation cited by [7] predict by the follow physical properties: Superficial velocity, Transversal area and specific mass of catalyst, resulting at $150kg/s$, thus, divided the maximum conversion to flow rate, obtained the ratio to control equal 0.007722.
- Ratio control model: Connect the information of dynamic model and export to a feedback control at the value of set point the ratio optimized (7.73×10^{-3}). Using the PI tuning by automatic model Simulink get at

Table 2. PI Controller

Parameter	Value
K_c	0.0056
τ_i	0.004

Figure 2 put forward the solution of dynamic model at 900 seconds. The solution has a exponential behavior,

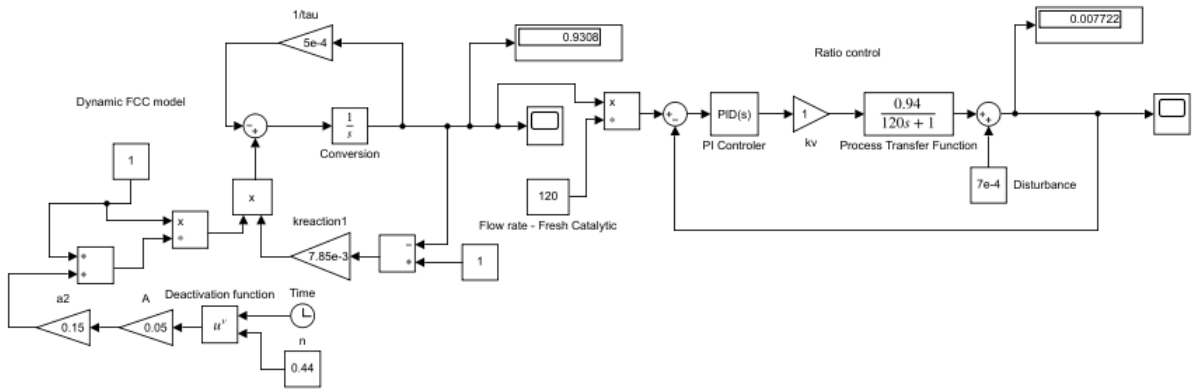

Figure 1. FCC - Solution dynamic model with ratio control

initially by zero because there is not feed of light oil at chemical reactor until 93% at maximum value achieved. Analysing the figure after 600 seconds conversion doesn't change, where represent the maximum value obtained. Thus, could be concluded that in approximately each period of 600 seconds is necessary change the catalyst present at reactor.

Figure 3 present the system control at a value of disturbance 7×10^{-4}, it is important to note that this value count effect of conversion and flow rate catalyst, because the transfer function input at the model related the both variables. Through control system the process it was stabilised after a time of 1200 seconds, characterized by a high sensible operation by a long time to stabilized and return at control.

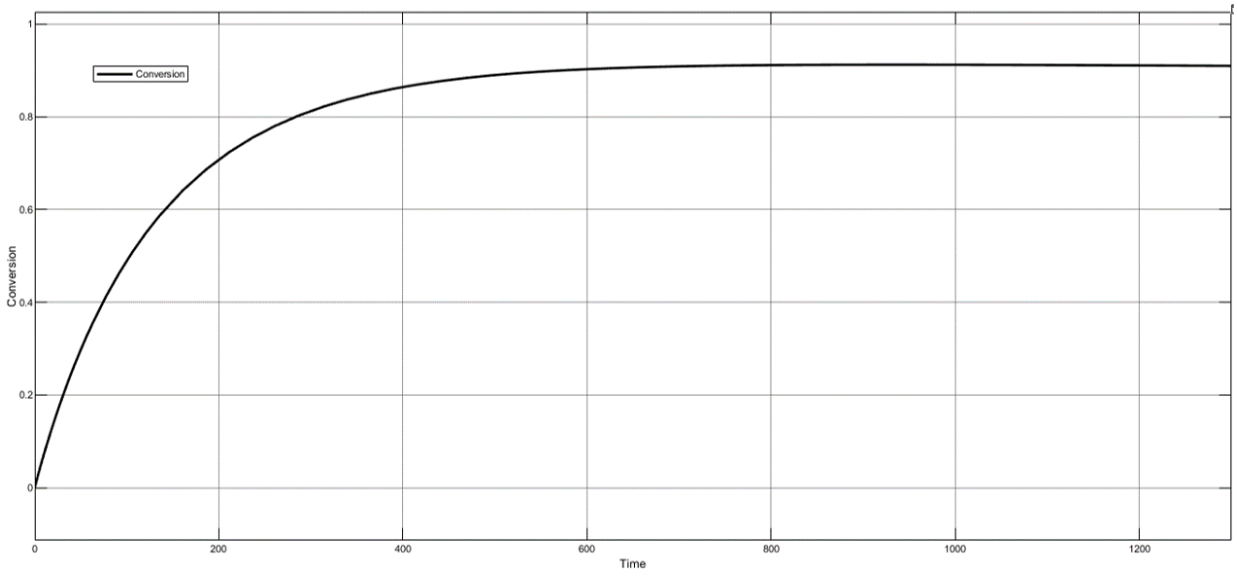

Figure 2. Dynamic model of FCC

3.2 Machine learning model and data engineering

After understanding the dynamic model and its results, used the simulated data to model a supervised model of machine learning method, support vector machines. This purpose is to predict which parameters is most important in catalytic cracking unit that was used at the dynamic model. So, these parameters were evaluated:

- Residence time τ: Express the ratio between reactor volume and volumetric flow to calculate the time necessary to achieve a conversion value. It is also used to study fluid flow and yield at a chemical reactor;
- Kinetic constant: Obtained experimentally, quantify how the reaction occurs;

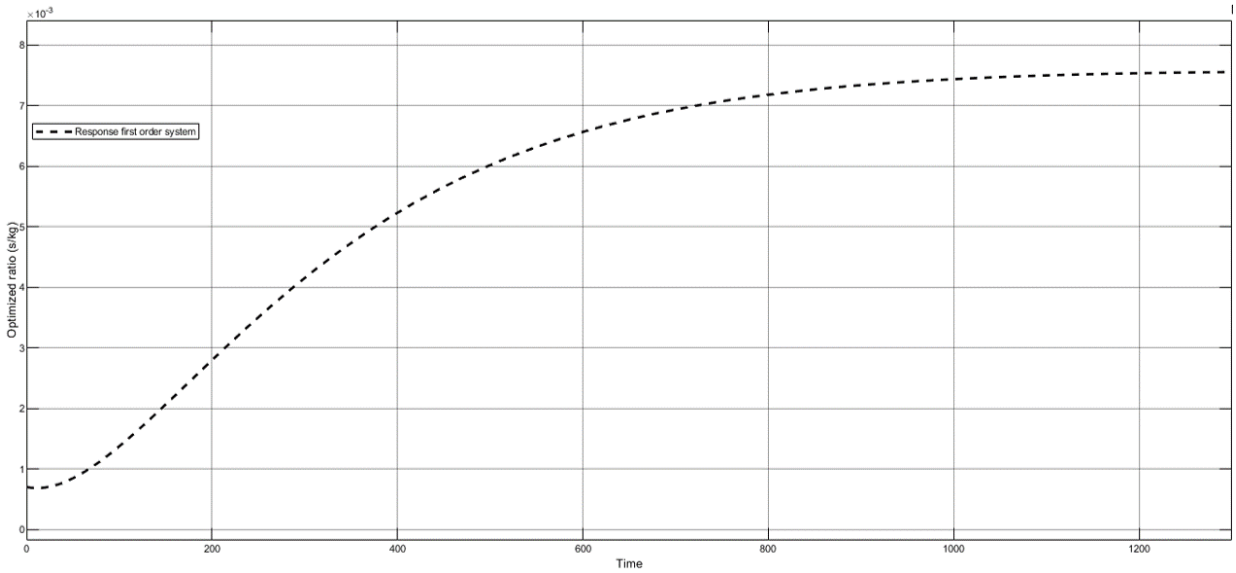

Figure 3. Response of system first order at a ratio control

- n: Present at deactivation function as a power of time. Calculate also experimentally using physical properties of petroleum that was cracking. Could get a value into interval 0.44 (light oil) until 0.88 (heavy oil);
- A: Proportional constant at deactivation function also is calculate experimentally and express the rate of coke formation during FCC unit;
- α_2: Mean the coke concentration at heavy oil at the feed of chemical reactor.

To model a support vector machine algorithm, the variables list at 3.2 was selected as predictor variables and the conversion as response variables. So, the dynamic model change the values of each parameters separately and maintains all other constants, running the model at each change, build the data set, using to input the machine learning method.

The table 3 to gives three information: parameters available and its respects values and RMSE error multiply by 10, for just convenience to compare the data. Each parameters was available at three value to compare and select the best considering as criteria results with smaller values of RMSE error and dynamic behavior when change its at simulation show at black font. Regarding the residence time, it is possible to observe that higher values allow a greater conversion, accounting for the effects of catalytic deactivation, which is also a function of time; High reaction constants positively impacted the conversion of the chemical reaction in a directly proportional relationship, therefore, it assumes the operation of the chemical reactor at high temperatures. The remaining parameters (n, α and A) are characteristic of the petroleum used at catalytic cracking, which are obtained experimentally. However, the model was consistent with the principles of catalysis, organic chemistry and chemical reactions, that is, oils with lower levels of paraffinic or high-density hydrocarbons, require a shorter operating time and amount of inputs to process it.

3.3 Machine learning model results

For a better understanding of the results, each variable will be evaluated individually herewith the operational significance of the catalytic cracking reactor.

- Residence time τ: Increase the residence time show a good operation condition for this reaction compare with others results. For explain the results it is necessary remember that a cracking reaction occurs two phenomenon, mass transfer and chemical reaction, thus with a additional time more reactant can be inside at catalyst, consequently more conversion of products.
- Kinetic constant: Express directly how the reaction happen high values consequently a short period of time to process a heavy oil; k_1 is the best value achieve, however to operate on this condition, only is possible if increase the reactor temperature;
- n: For an evaluation of parameters, agree a logical concepts, where a light oil is more rapidly to process as a heavy oil;

- A: Here the same idea, because also is obtained by the same petroleum data;
- α_2: Relate the coke concentration and catalyst activate. Thus, a petroleum with a high coke concentration tend leave a more time at reactor consequently need more flow rate of fresh catalyst.

Table 3. Results of machine learning regression

Parameters	τ_1	τ_2	τ_3	k_1	k_2	k_3	n_1	n_2	n_3	A_1	A_2	A_3	α_1	α_2	α_3
Values	4×10^{-4}	3×10^{-4}	7×10^{-4}	7.85×10^{-2}	7.85×10^{-1}	7.85×10^{-6}	0.5	0.66	0.88	0.01	0.1	0.25	0.1	0.25	0.5
RMSE $\times 10$	**0.00814**	0.01648	0.01567	**0.14430**	0.52680	0.27243	0.00857	**0.05340**	0.14889	**0.10034**	0.52811	0.56606	**0.01312**	0.05063	0.11195

4 Conclusions

Dynamic and machine learning methods studied the catalytic cracking unit. A mathematical model was added to the deactivation catalyst by coke formation, considering a rate low of activation in the time function with experimentally adjusted parameters. It was found that each period of 600 seconds is necessary for a new feed of catalyst to process the chemical reaction. Solve the nonlinear differential equation by a numerical method Bogachi- Shampine available at Simulink (R2021a) get the maximum conversion of 93%. After discussing the dynamic, the method was getting a transfer function to project an advanced control system in an optimized ratio related to a maximum conversion and the flow rate of fresh catalyst at the value of 7.7×10^{-3}, representing the set point value. Through the set point project, a feedback control with a PI controller to maintain the process at these references with a arbitrary disturbance value of 7.0×10^{-4} leave approximately 1200 seconds to process the return reference values. After understanding the dynamic operation, follow a machine learning study, using the support vector machines to predict the most important variables of chemical operation. It is possible to conclude that the ratio control was successfully advanced techniques to maintain the process under control considering all nonlinearity characterized chemical reaction model, showing a rapid response and mitigating effects of the other phenomena present. The machine learning results could be a sequence of which variables have a major influence on the chemical reaction in ascending order: n, α, A, k, and τ. This reaction strongly depends on resident time and kinetic constant. This work aimed to expand the knowledge about dynamics models, process control and machine learning applied in chemical engineering. We suggest future studies considering other laws of deactivation catalyst that account for more experimentally parameters, available control techniques to mitigate a coke formation and project a forecasting time series to increase catalytic activity.

Acknowledgements. I would like to thank the master's program at Universidade Federal Fluminesce. Financial support from the agência nacional de petróleo e gás natural (ANP). To my academic advisors, professors and graduate colleagues.

References

[1] H. Fogler. *Essentials of Chemical Reaction Engineering*. International Series in the Physical and Chemical Engineering Sciences. Pearson Education, 2017.
[2] N. Ghasem. *Modeling and Simulation of Chemical Process Systems*. CRC Press, 2018.
[3] N. Ghasem. *Computer Methods in Chemical Engineering*. CRC Press, 2021.
[4] C. Kravaris and I. K. Kookos. *Understanding Process Dynamics and Control*. Cambridge Series in Chemical Engineering. Cambridge University Press, 2021.
[5] I. Pan, L. R. Mason, and O. K. Matar. Data-centric engineering: integrating simulation, machine learning and statistics. challenges and opportunities. *Chemical Engineering Science*, vol. 249, pp. 117271, 2022.
[6] R. McFarlane, R. Reineman, J. Bartee, and C. Georgakis. Dynamic simulator for a model iv fluid catalytic cracking unit. *Computers & Chemical Engineering*, vol. 17, n. 3, pp. 275–300. Industrial challenge problems in process control, 1993.
[7] J. Ren, J.-P. Cao, X.-Y. Zhao, and Y.-L. Liu. Recent progress and perspectives of catalyst design and downstream integration in biomass tar reforming. *Chemical Engineering Journal*, vol. 429, pp. 132316, 2022.
[8] Q. Khanh Tran, H. Vu Ly, T. Anh Vo, H. Tae Hwang, J. Kim, and S.-S. Kim. Highly selective hydrodeoxygenation of wood pallet sawdust pyrolysis oil to methyl phenol derivatives using cobalt and iron on activated carbon supported catalysts. *Energy Conversion and Management: X*, vol. 14, pp. 100184, 2022.

[9] H. Xiang, R. Xin, N. Prasongthum, P. Natewong, T. Sooknoi, J. Wang, P. Reubroycharoen, and X. Fan. Catalytic conversion of bioethanol to value-added chemicals and fuels: A review. *Resources Chemicals and Materials*, vol. 1, n. 1, pp. 47–68, 2022.

[10] Y. Lin, W. Zhang, H. Machida, and K. Norinaga. Cfd simulation of the sabatier process in a shell-and-tube reactor under local thermal non-equilibrium conditions: Parameter sensitivity and reaction mechanism analysis. *International Journal of Hydrogen Energy*, vol. 47, n. 34, pp. 15254–15269. GREEN HYDROGEN FOR GLOBAL DECARBONIZATION, 2022.

[11] N. Remperova, J. Prech, M. Kubu, K. Golabek, J. F. Minambres, M.-F. Hsieh, A. Turrina, and M. Mazur. Gas-phase isomerisation of m-xylene on isoreticular zeolites with tuneable porosity. *Catalysis Today*, vol. 390-391, pp. 78–91. CIS-2021 Young Researchers, 2022.

[12] M. P. Kumar and N. Kaistha. Evaluation of ratio control schemes in a two-temperature control structure for a methyl acetate reactive distillation column. *Chemical Engineering Research and Design*, vol. 87, n. 2, pp. 216–225, 2009.

[13] A. Faul. *A Concise Introduction to Machine Learning*. CRC Press, 2019.