Stabilization time modeling for hydroprocessing: 1 identification of the dominant factors

Ngoc-Yen-Phuong Cao, Benoît Celse, Denis Guillaume, Isabelle Guibard, Joris W Thybaut

Stabilization time modeling for hydroprocessing: identification of the dominant factors

Ngoc-Yen-Phuong Cao[1], *Benoit Celse*[1*], *Denis Guillaume*[1], *Isabelle Guibard*[1], *Joris W. Thybaut*[2]

[1] IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France

[2] Ghent University, Laboratory for Chemical Technology, Technologiepark 125, Ghent, B-9052, Belgium

Email addresses of authors:

1. Ngoc-Yen-Phuong Cao: ngoc-yen-phuong.cao@ifp.fr

2. Benoit Celse: benoit.celse@ifpen.fr

3. Denis Guillaume: denis.guillaume@ifpen.fr

4. Isabelle Guibard: isabelle.guibard@ifpen.fr

5. Joris W. Thybaut: Joris.Thybaut@UGent.be

* Corresponding author: Benoit Celse

Email: benoit.celse@ifpen.fr

Address: IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France

Abstract

Hydroprocessing stabilization has been assessed using experimental data acquired at transient conditions. These data were obtained from a hydrotreating pilot plant in a wide range of operating conditions. It has been found that the stabilization evolution follows a first-order response and could be characterized by a stabilization time τ. A linear model was developed to relate τ to its most influential parameters. The model can then be combined with online transient data to predict the steady-state performance. By testing against new data with other feedstocks, the model has been found to provide a good prediction of the stabilization evolution and the ultimate steady-state hydrotreating performance. It is, hence, possible to "online" calculate the steady-state performance without the need to reach this steady state.

Keywords: hydrocarbon, hydrotreating, stabilization, transient data.

1 Introduction

Hydrotreating is a catalytic conversion process in petroleum refining, among others for removing impurities such as nitrogen and sulfur from hydrocarbon streams. It eliminates such heteroatoms in oil fractions from compounds by mixing them with hydrogen and treating this mixture in a fixed bed catalytic reactor at high temperature and pressure. The catalyst used in hydrotreating usually is molybdenum (group VI metal) based and promoted with cobalt or nickel (group VIII metal) and supported on alumina or silica-alumina. The most common catalysts used in hydrotreating are $CoMo/Al_2O_3$ and $NiMo/Al_2O_3$. CoMo favors the sulfur removal and olefin saturation while NiMo is used for removing nitrogen and saturating aromatic compounds (Treese et al., 2015). Hydrotreating/hydrocracking catalysts require several days before achieving stable operation while the residence time in the reactor is of the order of maximum a few hours. The observed transient phenomena can, hence, at most partly be attributed to hydrodynamics

phenomena in the equipment used and mainly to chemical phenomena occurring at the catalyst level. These are distinct from and, hence, are not to be confused with the long-term catalyst deactivation. There is very few literature on this subject. Sau et al. (2005) investigated the effects of organic nitrogen on hydroprocessing reactions. The authors did some hydrocracking experiments using a zeolite-based catalyst and determined the evolution of the conversion with time. When the reactor temperature was increased from 390 to 405 °C while processing a hydrotreated vacuum gas oil feed containing 320 ppmw organic nitrogen, a slow stabilization of the conversion was observed. The phenomenon was explained by the slow rate of nitrogen desorption from the catalyst with increasing temperature. Thanks to this nitrogen desorption, the concentration of active acid sites increases and, hence does the conversion. As the nitrogen desorption is slow, the observed conversion takes long time to reach the steady state as well. Similarly, these authors also investigated the effect of increasing nitrogen content in the feed from 5 to 125 ppmw. The observed time to reach the stabilization of conversion is also around eight days and is attributed to the adsorption of nitrogen on the active sites of the catalyst. According to Elizalde et al. (2016), the Liquid Hourly Space Velocities (LHSV) has an impact on the dynamic behavior of hydrocracking process. LHSV is the ratio of liquid volumetric flowrate and the catalyst volume. It is the inverse of the reactor space time. A lower LHSV is equivalent with a higher space time, so the time to reach the steady state is longer because the product concentration needs more time to reach the steady state. However, the authors reported that the temperature in the range of 380-400 °C does not seem to significantly affect the dynamic behavior. It means that no matter what the exact reactor temperature is, the time to reach the steady state remains similar.

A kinetic model is a significant asset, not to say essential, for the adequate design and simulation of processes, especially of large-scale operations with narrow profit margins. Since crude oil contains a lot of compounds with difficult structures, kinetic modeling of hydroprocessing is a challenging and time-consuming task (Jarullah et al., 2011). There are various approaches for hydroprocessing modeling, such as the detailed kinetic (Schweitzer et al., 1999; Oyekunle and Edafe, 2009; Charon-Revellin et al., 2011; Nguyen et al., 2015; Raghuveer et al., 2016; Nguyen et al., 2017; Doukeh et al., 2018), lumping modeling (Bonnardot, 1998; López García et al., 2010; Lababidi and AlHumaidan, 2011; Tang et al., 2013; Becker et al., 2015; Esmaeel et al., 2016) and black-box approach (Elkamel et al., 1999; Bahmani et al., 2010; Sadighi and Reza Zahedi, 2013; Sadighi and Mohaddecy, 2018). The most common and widely used is the lumping approach which consists of regrouping chemical compounds with similar properties (Oliveira et al., 2016). In these approaches, the model parameters are generally estimated by fitting the model to steady-state experimental data. Stabilization phenomena lead to long experimentation times in order to obtain sufficient steady-state data for kinetic modeling. However, in the transient regime towards this steady state, the analyses of liquid effluent are already carried out at regular time intervals to verify whether the steady state has effectively been reached and to ensure that the reaction is under control. These transient data are currently not used for kinetic modeling because the stabilization behavior is not well understood. It leads to a true challenge, meaning that experimental data acquisition to ‘calibrate’ the model is a time and money consuming task.

The aim of this work is, first, to acquire a better understanding of the stabilization behavior during these transient conditions and secondly, to use these data in order to predict, from the first transient points, the steady-state performance. If this value is far from a target, the operators can

change the operating condition without waiting for stabilization and without the use of a complex model.

2 Materials and methods

2.1 Pilot plant

The experimental data are acquired using a pilot plant located at IFP Energies Nouvelles, Solaize, France. It consists of four parallel fixed beds, operated in down-flow mode, which are used for hydrotreatment of feeds varying from gas oil to vacuum gas oil. The total catalyst volume in each reactor amounts to 50 cm^3. The pilot plant operates at isothermal conditions with the temperature being controlled along the reactor. Once the unit is in continuous operation, the density, nitrogen content and refractive index of the liquid effluent are analyzed every 24 hours. The steady state is ensured by having observed the stabilization of these characteristics. Operating conditions are adjusted after having reached the steady state corresponding with the previous operating conditions.

2.2 Data representation

The acquired data cover 11 Vacuum Gas Oil (VGO) feeds over two similar catalysts. Liquid Hourly Space Velocities (LHSV) vary between 0.5 and 4 h^{-1}, the temperature from 350 to 410 °C and the total pressure between 50 and 140 bar. Performance data are measured in terms of different properties of the liquid effluent such as density, refractive index and nitrogen contents over time on stream, totaling 920 measurements. A series of consecutive points corresponding to one experimental run is denoted as an *'episode'*. The latter characterizes the intended steady state kinetic analyses rather than the systematic, c.q., repeated, analysis of transient behavior. Figure 1 shows 7 episodes corresponding with 42 data points acquired in about 45 days. The first from these 7 episodes corresponds to the start of the test at specific operating conditions, while later

episodes are initiated by a change of operating conditions such as LHSV or temperature. Other variables that may also change between consecutive episodes are pressure and feedstock composition. It is noted that the deactivation phenomena in the pilot plant are negligible for the test with duration less than 45 days.

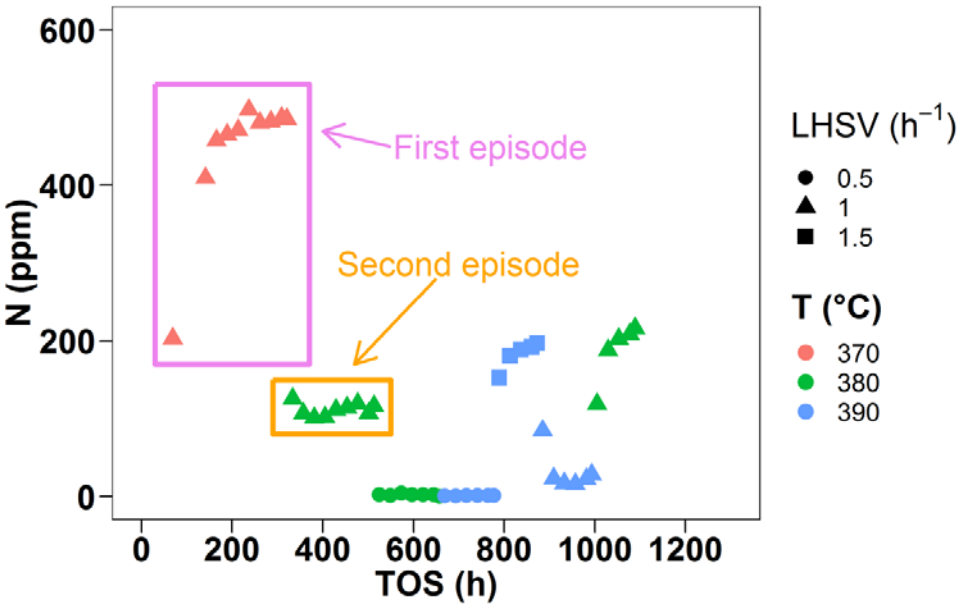


Figure 1. Data illustration - Liquid effluent nitrogen as a function of time on stream (TOS).

2.3 Stabilization modeling

A computational study on the pilot plant has been carried out using a tracer technique to estimate the stabilization time purely due to hydrodynamic phenomena. The objective is to identify whether the experimentally observed stabilization time is determined by hydrodynamic or chemical phenomena. The corresponding dynamic simulation of the pilot plant (without chemical reaction) was carried out using Matlab & Simulink R2014b in order to follow the flow of an unreactive tracer component through the pilot plant.

Exploratory data analysis (Tukey, 1977) (see also Figure 1) shows that the stabilization phenomena follow a first-order transfer function as shown in Equation (1).

$$y_{(TOS)} = y_{init} + \left(y_{final} - y_{init}\right)\left(1 - exp\left(\frac{-(TOS - TOS_{init})}{\tau}\right)\right) \quad (1)$$

Where y is the liquid effluent property at a specific time on stream (such as nitrogen content, density, refractive index); y_{init} is the liquid effluent property corresponding to the first experimental point of episode (ppm); y_{final} is the last experimental point of episode (ppm); TOS is the time on stream (h); TOS_{init} is the time on stream corresponding to y_{init} (h) and τ is the characteristic time of the episode (h). The characteristic time τ of each episode presented in the equation is estimated via nonlinear least-squares technique using Gauss-Newton algorithm (Bates and Watts, 1988).

The estimated characteristic time τ can logically be supposed to depend on some variables such as feed properties and operating conditions. A linear function accounting for possible interaction among the variables is defined to describe the characteristic time τ (Equation (2)).

$$\tau = a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + \cdots + b \quad (2)$$

Where x_1, x_2 are the input variables, which can be LHSV, temperature, pressure or feed properties such as organic nitrogen content, organic sulfur content, resin content in feed, etc.; x_1x_2 is the interaction term between x_1 and x_2; a_1, a_2, a_{12},… and b are the coefficients.

The most influential input variables of the model are determined via variable selection technique called ‘leaps’ (Furnival and Wilson, 1974). The main idea of this technique is based on an efficient branch-and-bound algorithm which was first proposed by Land and Doig (1960). It searches for all possible subset solutions and stores them in the branches of a tree with the full set at the root. The algorithm runs an exhaustive search and shows the best variable subsets. It reveals the linear dependence of stabilization on the input variables. The quality of each subset is evaluated by a metric called adjusted R^2. In terms of statistics, higher adjusted R^2 means a better

quality. However, the number of variables in the model should not to be excessive. The selected model is a compromise between adjusted R^2 and the number of variables in the subset.

The adjusted R^2, see Equation (3), is a measure of model fitting quality taking into account the number of variables. It is usually employed to compare multiple linear models.

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1} \quad (3)$$

Where R^2 is the 'coefficient of determination' or multiple R^2; n is the number of observations in data sample; p is the number of predictors, c.q., variables and/or parameters, in the model. Multiple R^2 is calculated via Equation (4).

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (4)$$

Where $y_i$are observed values; $\hat{y}_i$ are predicted values; $\bar{y}$ are mean of sample.

2.4 Stabilization prediction

One of the properties of Equation (1) is that 95% of the value at steady state is reached after 3τ and 99% of the steady-state value after 5τ. It is a characteristic of first-order function which can be found in the literature (Smith and Campbell, 2016). The time required to reach steady state can be estimated thanks to the τ predicted by the model.

The model for τ can then be used for stabilization performance prediction. For each episode, τ is calculated by the model. Knowing τ and the first two experimental points of the episode suffices to predict the episode evolution and, ultimately, the behavior at steady state using Equation (5), exemplified for the nitrogen content. If the predicted steady-state value is far from the target, the operating condition can be adjusted without waiting for stabilization. The prediction can be updated daily based on the new measurements.

$$N_{final} = \frac{N_2 - N_1}{1 - exp\left(-\frac{TOS_2 - TOS_1}{\tau}\right)} + N_1 \quad (5)$$

Where N_{final} is the liquid product nitrogen at steady state (ppm); N_2 is the liquid product nitrogen corresponding to the second point of episode (ppm); N_1 is the liquid product nitrogen corresponding to the first point of episode (ppm); TOS_2 is the time on stream corresponding to the second point of episode (h); TOS_1 is the time on stream corresponding to the first point of episode (h) and τ is the characteristic time predicted by model (h).

3 Results and discussion

This section exhibits the results of each step in the methodology shown in Figure 2. First part represents the hydrodynamic study of the pilot plant where the experimental data were acquired. The analysis of stabilization behavior and its dependence on operating conditions, feedstock are displayed. Stabilization model is then constructed and the prediction against new data is performed.

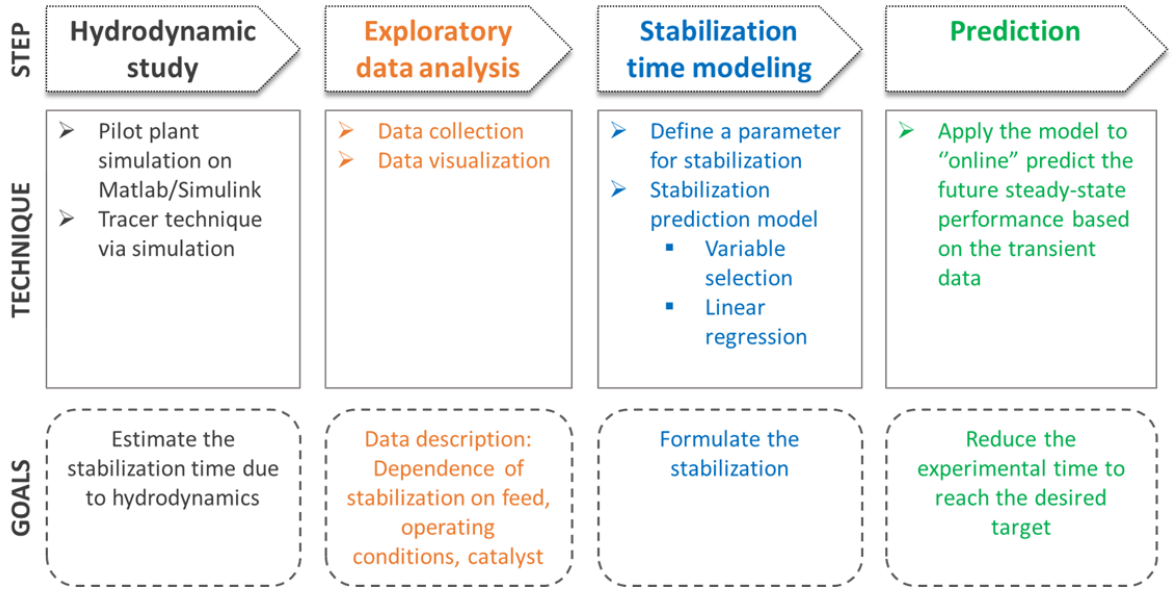


Figure 2. Methodology adopted throughout this work.

3.1 Pilot plant hydrodynamics

A computational analysis of the pilot plant hydrodynamics has been carried out using Matlab & Simulink via the simulation of the behavior of an unreactive tracer component running through the pilot plant. The outlet concentration as a function of time provides insight into the stabilization behavior purely due to hydrodynamics. The pilot plant was simulated as several blocks representing each part of the pilot plant such as a heater, a reactor, a separator to remove hydrogen-rich gas from liquid stream and a stripping column to remove gas dissolved in the liquid stream such as H_2S, NH_3. Each element was modeled by considering it as a plug flow reactor (PFR) or as a continuous stirred-tank reactor (CSTR). Pipeline, heater and reactor were modeled as PFR while separator and stripping column were simulated as CSTR. Mass balance was then written for each block. In the simulation, the input stream is a tracer solution with a concentration set randomly at 1 mol/m^3. The volumetric flowrate of the input stream was simulated at 50 cm^3/h, which is similar to the one in the experimental data that would be used to compare. Temperature and pressure in the reactor were set at 370 °C and 140 bar respectively. There was no simulation of the chemical reaction expected.

Figure 3 compares the transient behavior of tracer concentration in pure hydrodynamics mode with the effluent properties in reactive mode. The figure on the left represents the tracer outlet concentration. The figure on the right shows the transient behavior of liquid product nitrogen in reactive mode. As can be seen, the hydrodynamic stabilization of the pilot plants occurs within 25 hours while the observed stabilization of the pilot plant in reactive mode requires around 200 hours. The hydrodynamic response to a tracer step reached steady state significantly faster than the stabilization of the hydrodenitrogenation behavior. Hence, stabilization is mainly driven by chemical phenomena.

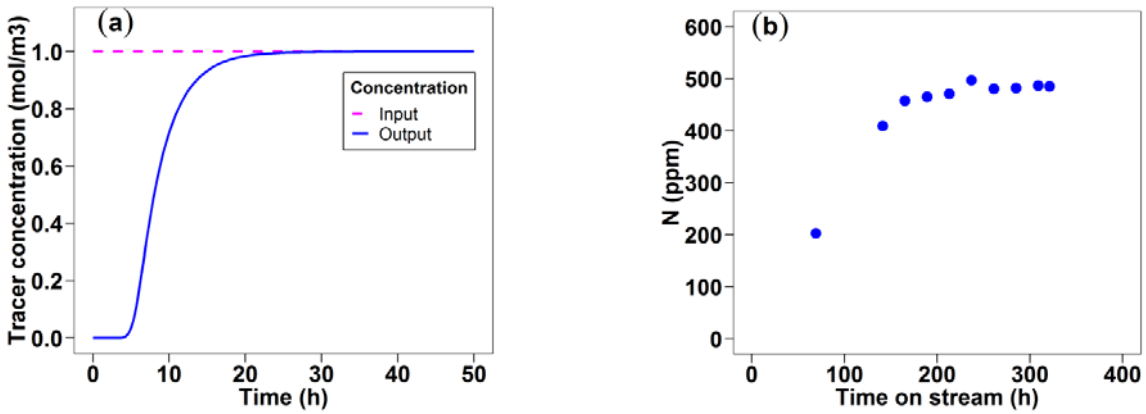


Figure 3. Calculated effluent concentration profiles of tracer (a) and experimentally observed liquid effluent nitrogen content as a function of time on stream in the pilot plant (b) (LHSV = 1 h^{-1}, T = 370 °C, P = 140 bar).

3.2 Exploratory data analysis

Stabilization was found to exhibit a first-order behavior, which is similar to the observations by Sau et al. (2005). The first-order characteristic time τ can be calculated from the evolution of effluent characteristics such as the nitrogen content, density and/or refractive index. Figure 4 compares the transient behavior of these properties, normalized between [0, 1], for the same experimental run. The stabilization behavior of these properties is similar and the values of τ are essentially the same. Following the evolution of stabilization by density and refractive index are more suitable than by the nitrogen content for the following reasons: (1) for episodes with low nitrogen level (< 10 ppm), it becomes difficult to observe the behavior in the nitrogen curve, while it remains feasible for density and refractive index; (2) density and refractive index measurements are typically more precise than for nitrogen and (3) density and refractive index measurements are easier (and less expensive) to carry out than nitrogen measurements.

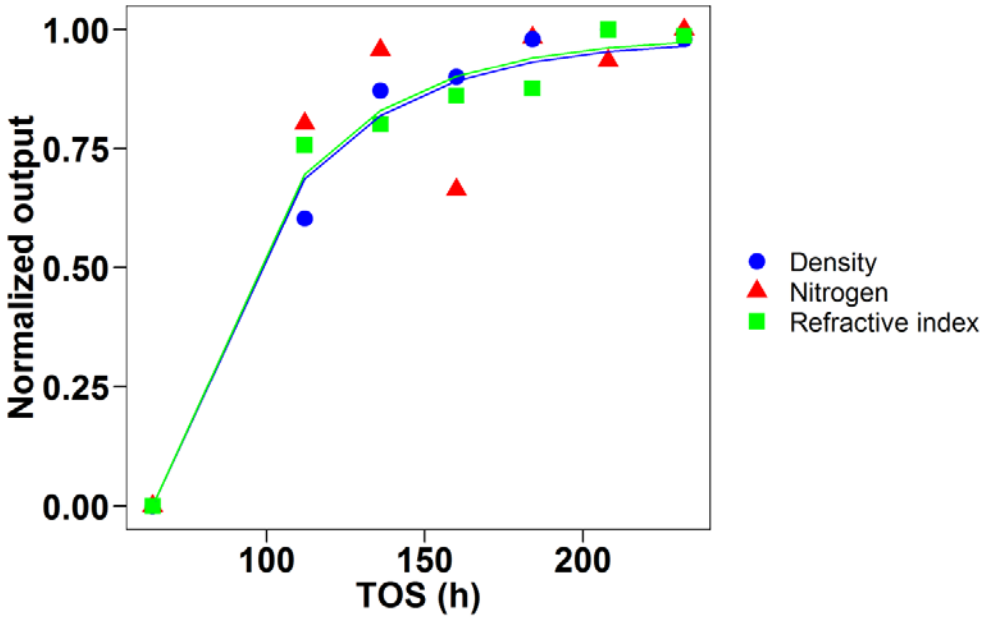
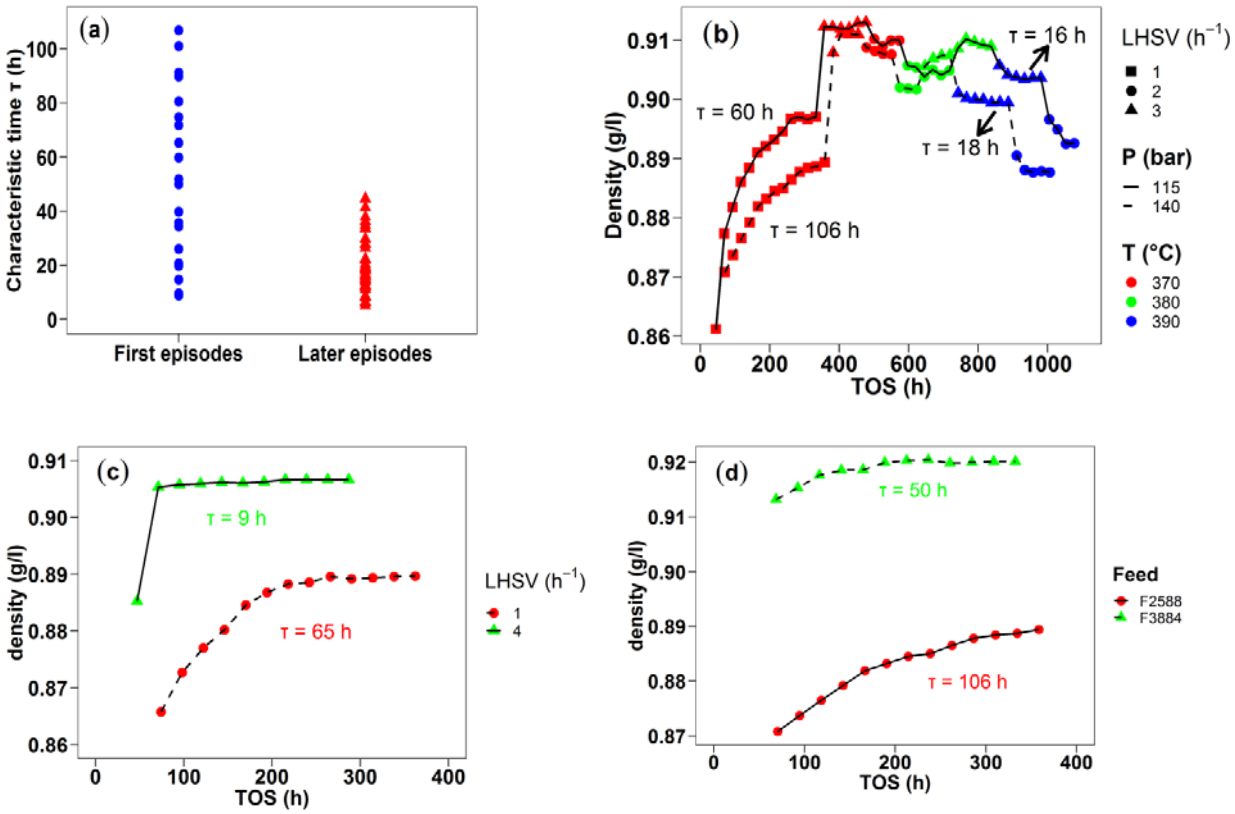


Figure 4. Comparison of transient behavior of nitrogen/density/refractive index (outputs are normalized between [0, 1]). Values of τ are 39.8, 36.3, 39.2 h calculated respectively from density, nitrogen content and refractive index evolution.

There were 54 values of τ which have been classified into 21 first episodes and 33 later episodes, which are shown in Figure 5(a). It can be seen, indeed, that first episodes, on average, require more time to reach steady state as compared to later episodes. It was, hence, decided not to describe the τ of first and later episodes with a single model, but to develop dedicated models, i.e., one for the τ from first episodes and one for the τ of later episodes.



225 Figure 5. (a) Classification of τ regarding the type of episode; (b, c, d) Impact of pressure (solid
226 line: 115 bar, dashed: 140 bar, feed F2588); LHSV (feed F4384, P = 140 bar, T = 370 °C) and
227 feedstock (LHSV = 1 h^{-1}, P = 140 bar, T = 370 °C) respectively on transient behavior of liquid
228 effluent density.

229 Figure 5(b), (c) and (d) show the transient behavior of liquid effluent density as a function of
230 time on stream. Figure 5(b) displays the evolution of stabilization at two different pressures, i.e.,
231 115 and 140 bar. It can be observed that for the first episodes, the evolution of the curve of 115
232 bar (τ = 60 h) is more rapid than the curve of 140 bar (τ = 106 h). For the later episodes, the
233 difference in stabilization behavior between 115 and 140 bar seems to be less pronounced.

234 The impact of LHSV was analyzed using data from two tests with the same feed and identical
235 pressure and temperature (P = 140 bar, T = 370 °C) with LHSVs amounting to 1 and 4 h^{-1}. As

can be seen in Figure 5(c), a higher LHSV leads to a much quicker stabilization. The value of τ in the case of LHSV of 4 h^{-1} was 9 h which is lower than 65 h in the case of 1 h^{-1}.

The feedstock is also found to have an impact on the stabilization. Figure 5(d) compares the evolution of liquid effluent density as a function of time on stream between two feedstocks. The evolution with F3884 having a higher resin and nitrogen content stabilizes quicker than the one with F2588.

Other factors, such as temperature, catalyst were found to exert a less pronounced impact on the stabilization behavior.

3.3 Model for τ prediction

τ is modeled using linear regression accounting for interactions. As indicated before, two linear models for τ were built; one for first episodes and another for other episodes.

3.3.1 First episodes model

There are in total 21 values of characteristic time τ corresponding to first episodes. The possible input variables are operating conditions (LHSV, temperature, total pressure) and feed properties (organic nitrogen content N_{feed}, organic sulfur content S_{feed}, resin of feed res_{feed}, density d_{feed} and the weighted average temperature of simulated distillation TMP_{feed}). All input variables and the interaction terms which are products of operating conditions and feed properties are used in variable selection procedure. Figure 6 shows the evolution of the adjusted R^2 of the 5 best subsets of each size as a function of the size of subsets. It seems that 4 variables represents a good compromise between model accuracy and overfitting (bias/variance compromise), i.e., although the adjusted R^2 still slightly increases when including more than 4 variables, the increase seems so marginal that it is preferred to limit the number of variables in the model to 4.

Table 1 presents an overview of the selected variables in the best subsets with the number of variables in subset from 1 to 7. The 5 best subsets with 4 variables are detailed as well. Among these 5 best subsets, the subset (LHSV, P, res_{feed}, LHSV*res_{feed}) giving the highest adjusted R^2 is selected. This model contains only 1 interaction term LHSV*res_{feed}, which also appears in the 4 left. It is a robust and less complicated model.

The model presents an adjusted R^2 of 0.79 and a multiple R^2 of 0.83. The corresponding parity diagram is given in Figure 7. The results are quite good for all the feedstocks.

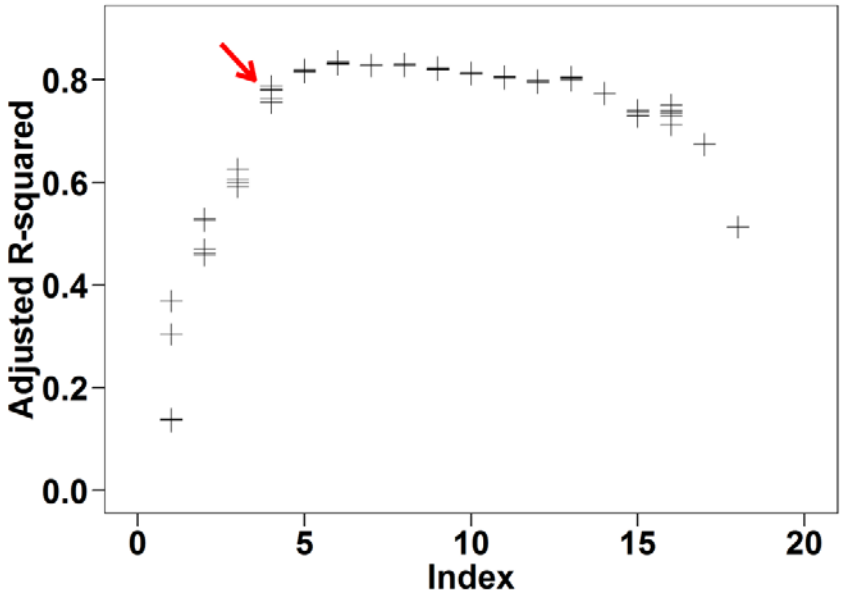


Figure 6. Variable selection using leaps for linear model with interaction (first episodes).

269 Table 1. Selected variables corresponding to each subset obtained from variable selection in
270 Figure 6 (only show the best selected subsets of 'from 1 to 7' variables and 5 best subsets of 4
271 variables)

Index	Selected variables	Adjusted R^2
1	P	0.369
2	P, LHSV*S_{feed}	0.528
3	T, P, LHSV*S_{feed}	0.625
4a	**LHSV, P, res_{feed}, LHSV*res_{feed}**	**0.788**
4b	LHSV, res_{feed}, LHSV*res_{feed}, P*TMP_{feed}	0.782
4c	res_{feed}, LHSV*res_{feed}, LHSV*TMP_{feed}, P*TMP_{feed}	0.780
4d	LHSV, P, LHSV*res_{feed}, T*res_{feed}	0.763
4e	LHSV, LHSV*res_{feed}, T*res_{feed}, P*TMP_{feed}	0.756
5	LHSV, T, P, res_{feed}, LHSV*res_{feed}	0.819
6	P, res_{feed}, LHSV*res_{feed}, LHSV*TMP_{feed}, T*res_{feed}, P*N_{feed}	0.835
7	LHSV, P, N_{feed}, LHSV*Res_{feed}, T*TMP_{feed}, T*N_{feed}, P*N_{feed}	0.828

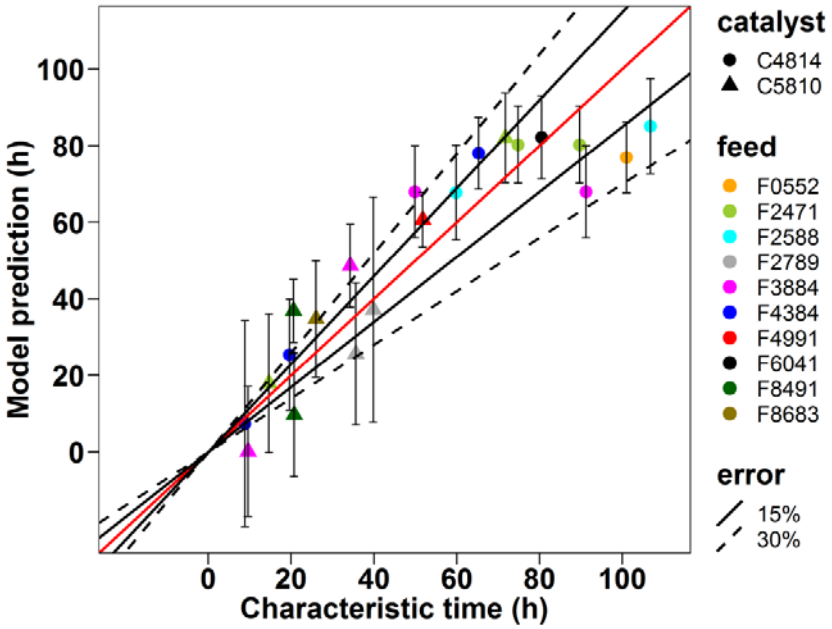


Figure 7. Parity plot with 95% confidence interval ($R^2 = 0.83$) (first episodes model).

The coefficients of the model as well as the statistical values are presented in Table 2. The global p value amounts to 5.3×10^{-6} and the F value for the global significance of the regression exceeds the 95% quantile by far, which both indicate that the model is statistically significant. All variables are individually significant as well as their p value is sufficiently low. As can be seen, a negative correlation was determined between LHSV and τ. The modeling confirmed that temperature is not a dominant factor. These two results are coherent with those acquired by Elizalde et al. (2016). A positive proportionality between pressure and τ was observed. However, the impact of pressure is less pronounced than that of other variables. The interaction term shows that the impact of LHSV on τ depends on the value of feed resin, i.e., the polar components with high molecular weight.

Table 2. Coefficient and statistical values for first episodes model

	Intercept	LHSV	P	res_{feed}	LHSV*res_{feed}
Coefficient	63.339	-69.055	0.693	-7.081	5.481
p-value	3.33×10^{-3}	3.15×10^{-6}	7.1×10^{-6}	5.31×10^{-5}	1.9×10^{-4}

Global p-value	F	$F_{95\%}$
5.3×10^{-6}	19.55	3.01

An overview of the diagnostics for the multiple linear regression with interaction yielding the combination (LHSV, P, res_{feed}, LHSV* res_{feed}) as best performing one, is displayed in Figure 8. The analysis is as below:

- Residuals vs. fitted shows that there is linear trend.
- The normal Q-Q indicates that the residuals are almost normally distributed.
- The assumption of equal variance is also valid by checking the scale-location.
- Residuals vs. Leverage show that there is no influential observation. Even the observation #13 is on the Cook's distance line (red dashed line on the Residual vs Leverage plot), it is not an influential observation. Because when this point is excluded from the regression, multiple R^2 changes from 0.8301 to 0.8335, which means a small impact.

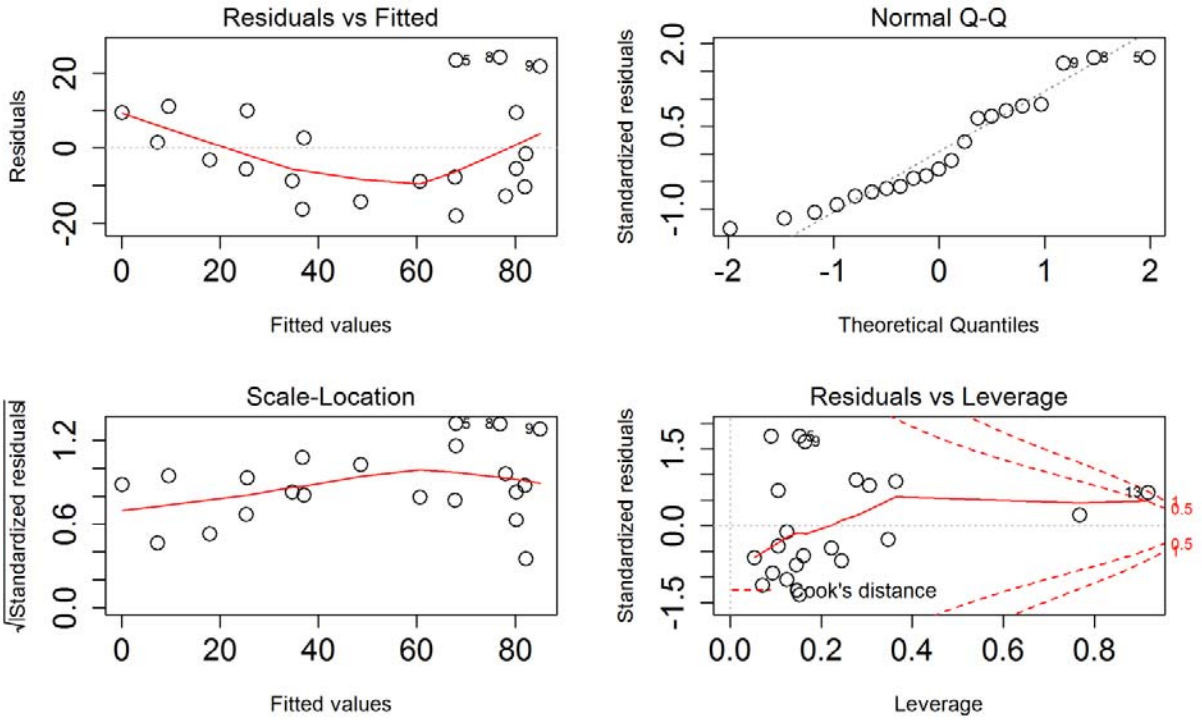


Figure 8. Diagnostic plots for linear model with interaction (first episodes model).

3.3.2 Later episodes model

There are 33 observations of τ in the whole data set for later episodes. The same procedure for variable selection was applied. The input variables are the feed properties, operating conditions of the current episode as well as of the previous episode. Variables relating to the previous episode are, e.g., temperature and LHSV of this previous episode: T_{pre}, $LHSV_{pre}$; and ratio of temperature and LHSV of the episode and the previous one. Only a selection of interaction terms was allowed based on the inspiration acquired by modeling the first episodes, to keep the number of considered variables and, hence, adjustable parameters, within reasonable constraints. Table 3 shows the 17 input variables taken into account in variable selection procedure.

Table 3. Variables used in variable selection technique for later episodes regression

Feed	Operating conditions	Condition switching	Interaction terms
N_{feed}	LHSV	$\frac{T}{T_{pre}}$	LHSV*res_{feed}
S_{feed}	T	$\frac{LHSV}{LHSV_{pre}}$	LHSV*N_{feed}
res_{feed}	P	$\frac{T_{pre}}{T}$	LHSV*S_{feed}
d_{feed}	$LHSV_{pre}$	$\frac{LHSV_{pre}}{LHSV}$	
TMP_{feed}	T_{pre}		

Figure 9 show the evolution of the adjusted R^2 of the best subsets as a function of the number of variables. On similar grounds as for the first episodes modeling (Section 3.3.1), for the later episodes regression 7 variables is determined to be a good compromise between accuracy and over fitting for these later episodes.

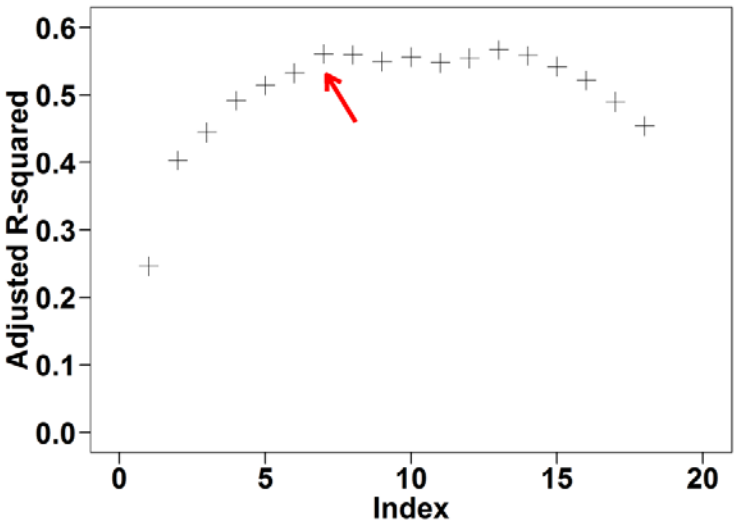


Figure 9. Variable selection using leaps for later episodes model.

These 7 variables selected are: LHSV, S_{feed}, $LHSV_{pre}$, T_{pre}, LHSV*res_{feed}, $\frac{LHSV}{LHSV_{pre}}$, $\frac{LHSV_{pre}}{LHSV}$. The model gives an adjusted R^2 amounting to 0.55 and a multiple R^2 to 0.66. Figure 10 shows the parity diagram of the model. The quality of the model is lower than that of first episode model, which is, at least partly, explained by the fact that the results are closer to each other, i.e., there is less spread for a higher number of variables.

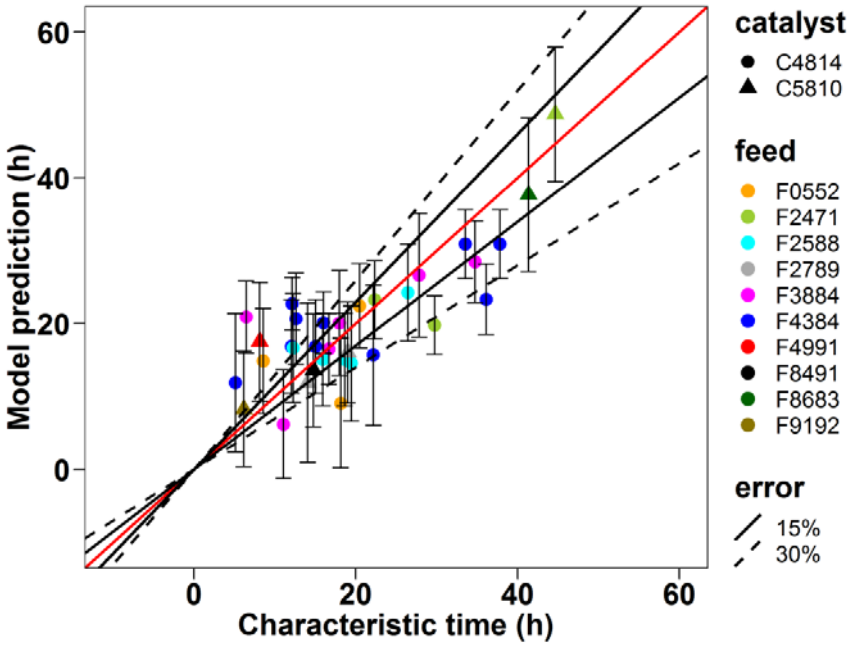


Figure 10. Parity plot with 95% confidence interval (R^2 = 0.66) (later episodes model).

The regression is globally statistically significant since the global p-value is sufficiently small and the F value exceeds the critical value.

Table 4 displays the parameter values obtained for the model and their individual statistical significance. The p values for the individual variables indicate that S_{feed} and $LHSV_{pre}$ are at the margin of statistical significance (P = 0.11 and P = 0.07, respectively) and that the other variables are significant. LHSV, sulfur and resin content in the feed represent the operating conditions and feed properties. Other variables represent the impact of previous conditions on the

τ. The τ of later episodes is more complicated to model than the first episodes and depends on the conditions of the previous episode.

τ fitted by data has an uncertainty due to the measurement error of the output such as liquid effluent nitrogen or density. In another database, one episode was repeated several times. The values of τ fitted by the obtained data are 45, 45, 50, 64, 54 and 59 (hours). The obtained mean is 53 hours and the standard deviation is about 8 hours. Besides that, the residual standard deviation of the model is 7. It is shown that τ fitted by data varies from one experiment to another experiment under the same conditions. The latter can lead to the inaccuracy of model since the values of fitted τ were used to train the model. The proposed model can be sufficient since it simulates well the order of magnitude of fitted τ and the model is easy to establish and understand.

Table 4. Coefficient and statistical values for later episodes model

	Intercept	LHSV	S_{feed}	$LHSV_{pre}$
Coefficient	177.349	-10.943	2.589	8.851
p-value	8.32×10^{-5}	0.036	0.114	0.071

	T_{pre}	$LHSV*res_{feed}$	$\frac{LHSV}{LHSV_{pre}}$	$\frac{LHSV_{pre}}{LHSV}$
Coefficient	-0.379	-0.828	8.832	-11.713
p-value	8.34×10^{-4}	3.85×10^{-3}	5.78×10^{-3}	0.020

Global p-value	F	$F_{95\%}$
1.4×10^{-4}	6.83	2.40

Diagnostics of model for later episodes is discussed as below (see also Figure 11).

- Residuals vs. fitted shows that there is a linear trend.
- The normal Q-Q indicates a normal distribution.
- Scale-location shows that the assumption of equal variance is valid.
- Residuals vs. Leverage show no influential observation.

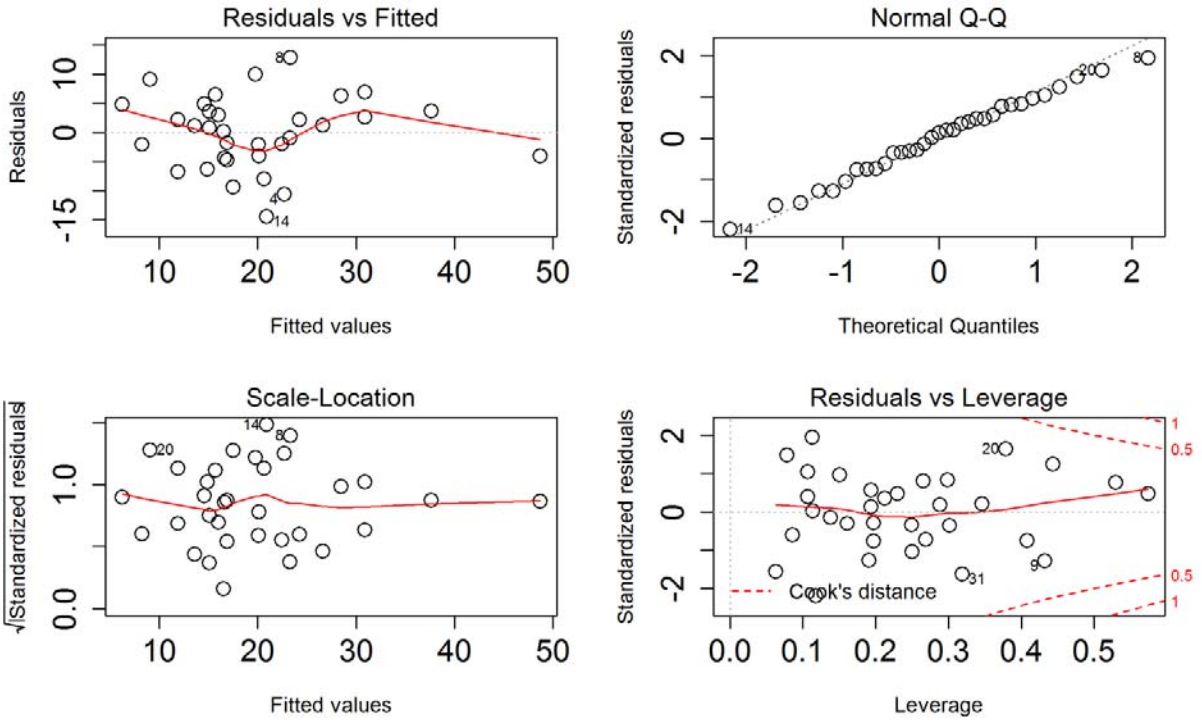


Figure 11. Diagnostic plots for linear model with interaction (later episodes model).

3.4 Prediction

The model was used to predict τ and the steady-state performance was calculated via Equation (5) assuming that the first two points of the episode have already been measured. The prediction was carried out on a new experimentation with another feedstock operating on another pilot plant but the same catalyst.

3.4.1 First episodes

First episodes model is tested using new data from other feedstocks. Figure 12 shows two examples of such predictions by the model. The first two points of the episode which are used to predict the steady-state value are shown as circles. The remaining data in the episode are represented by triangles. The predicted nitrogen values are very close to the measured ones. The model has been found to provide a good prediction of the stabilization evolution and the steady-state value.

Using the model for τ prediction enables the calculation of the steady-state performance without the need to wait for reaching steady state. The advantage is that if the steady-state value could be far from target, operators can change the operating condition without waiting for stabilization.

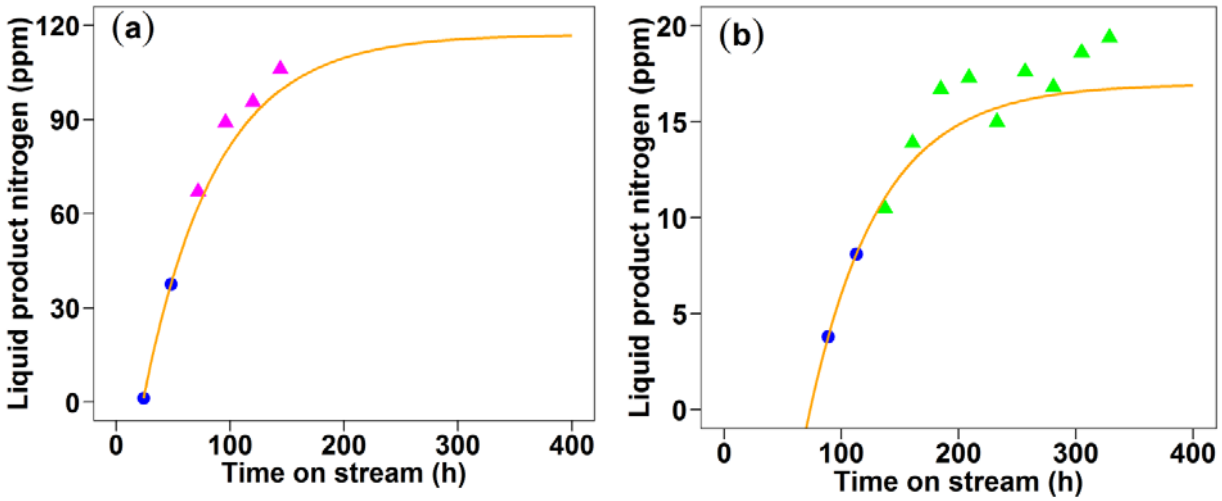


Figure 12. Two test predictions using the first two points of episode (points: experimental data, solid line: model prediction; Test (a): LHSV = 1.34 h^{-1}, P =132 bar, T = 370 °C; Test (b): LHSV = 1.71 h^{-1}, P =140 bar, T = 386 °C).

3.4.2 Later episodes

Later episodes model is tested using new data from other feedstocks as well. Figure 13 compares the prediction (solid line) and the experimental data. Again, the first two points of episode which are used to estimate the steady-state value are shown as circles. The remaining data in the episode are represented by triangles. The model predicts well the evolution of episodes 2, 3, 5, 7 and 8. Since nitrogen values are low in episode 7 and 8, the prediction is considered not far from the experimental value, even if the evolution is not perfectly in line with the experimentally observed one. Moreover, during the first point of an episode, hydrodynamic effects still may have a certain impact, indicating that we should be careful with the reliability of this first point.

Regarding episode 4 and 6, the stabilization behavior did not follow the first-order response since the reactor temperature was slightly adjusted to reach an intended nitrogen level. The reactor temperature profile is shown in Figure 14. A target of 30 ppm was fixed for the value of liquid effluent nitrogen in episode 4. The reactor temperature was firstly set at 378.7 °C. The prediction using the first two points was 25 ppm (see solid line in episode 4, Figure 13). The reactor temperature was, hence, gradually decreased to 377.1 °C. The temperature profile was considered constant from the fourth point. The steady-state value prediction using the fourth and fifth point was 32.8 ppm, which is not far from the target (see dashed line in Figure 13). The third point was not used for prediction since the temperature was adjusted during that time. Similar to episode 4, the liquid effluent nitrogen target of episode 6 was 50 ppm. The temperature was firstly set at 374.7 °C for episode 6 and the steady-state forecast using the first two points was 80 ppm. The temperature was gradually modified to 376.7 °C. The steady-state value prediction using the fourth and fifth point was 52 ppm, which was close to the target (see also dashed line in episode 6, Figure 13). As the liquid product nitrogen values seem to be

sensitive to the reactor temperature value, it is needed to carefully control the temperature profile. As can be seen, the steady-state performance prediction depends not only on the predicted value of τ but also on the first two "online" measurements. This 'online' prediction is, hence, flexible and the operators should consider the certainty of experimental data and adapt to what really happen (modification of operating conditions, measurement error and equipment problem).

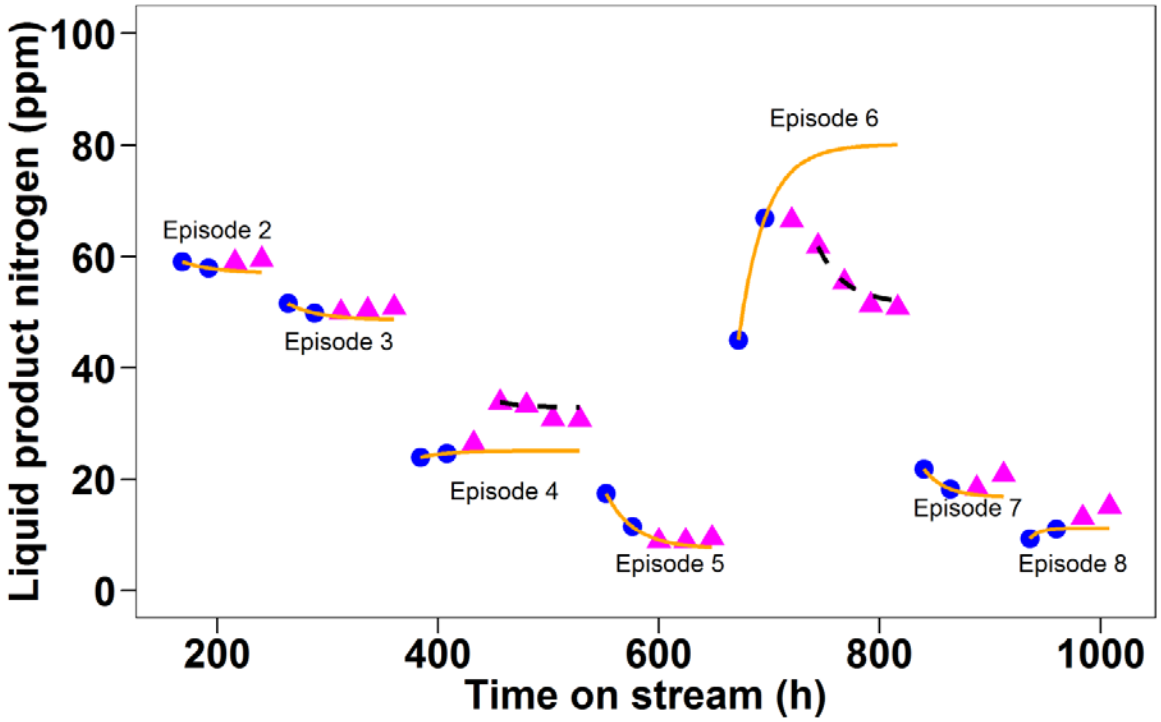


Figure 13. Prediction with the first two points of other episodes (points: experimental data, solid line: model prediction using the first two points, dashed line: prediction while discarding the first three points).

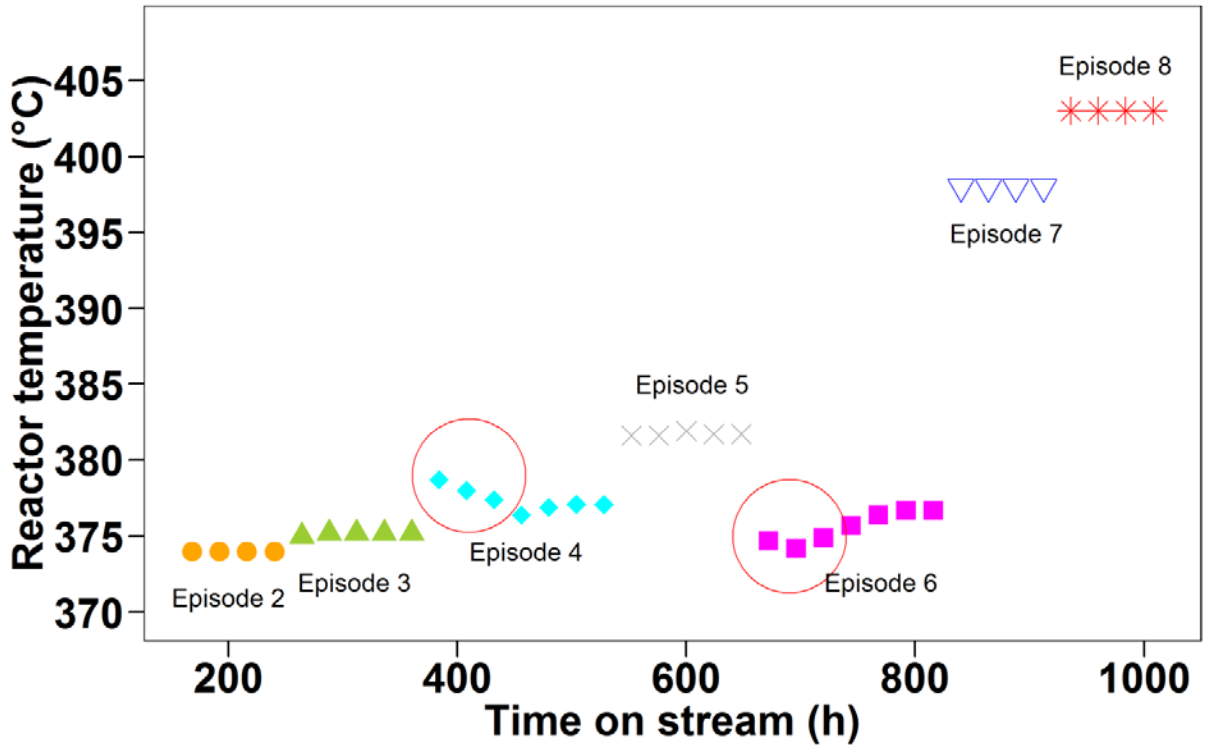


Figure 14. Reactor temperature profile corresponding to each episode

Even if the quality of the model seems limited, the model reproduces well the order of magnitude of τ. The impact of τ predicted by model on the stabilization evolution prediction was carried out for one episode. Figure 15 shows the steady-state prediction based on the first two points and different values of predicted τ. In fact, a variation of ±15% and ±30% was added on τ predicted by model in order to study the impact of predicted τ on the steady-state prediction. As can be seen, the nitrogen value comes to its steady state at about 14 – 20 ppm for ±30% variation of predicted τ. The steady-state value of nitrogen is 15.5 – 18.5 ppm for ±15% variation. It is shown that a small variation of τ does not impact significantly on the prediction. Moreover, the model can be employed at the experimentation scheduling step. The predicted τ (in hours) can be used to estimate the order of magnitude of the time required to reach steady-state (in days). As explained in Section 2.4, the steady state can be considered to reach after about 3τ – 5τ.

The proposed methodology can be applied to any reaction where the chemical stabilization requires significantly more time than the hydrodynamic stabilization and analysis time to forecast the evolution towards the steady state. The latter can help the operator to follow the trend of the stabilization behavior. It enhances the decision making in process operation to reach an intended target.

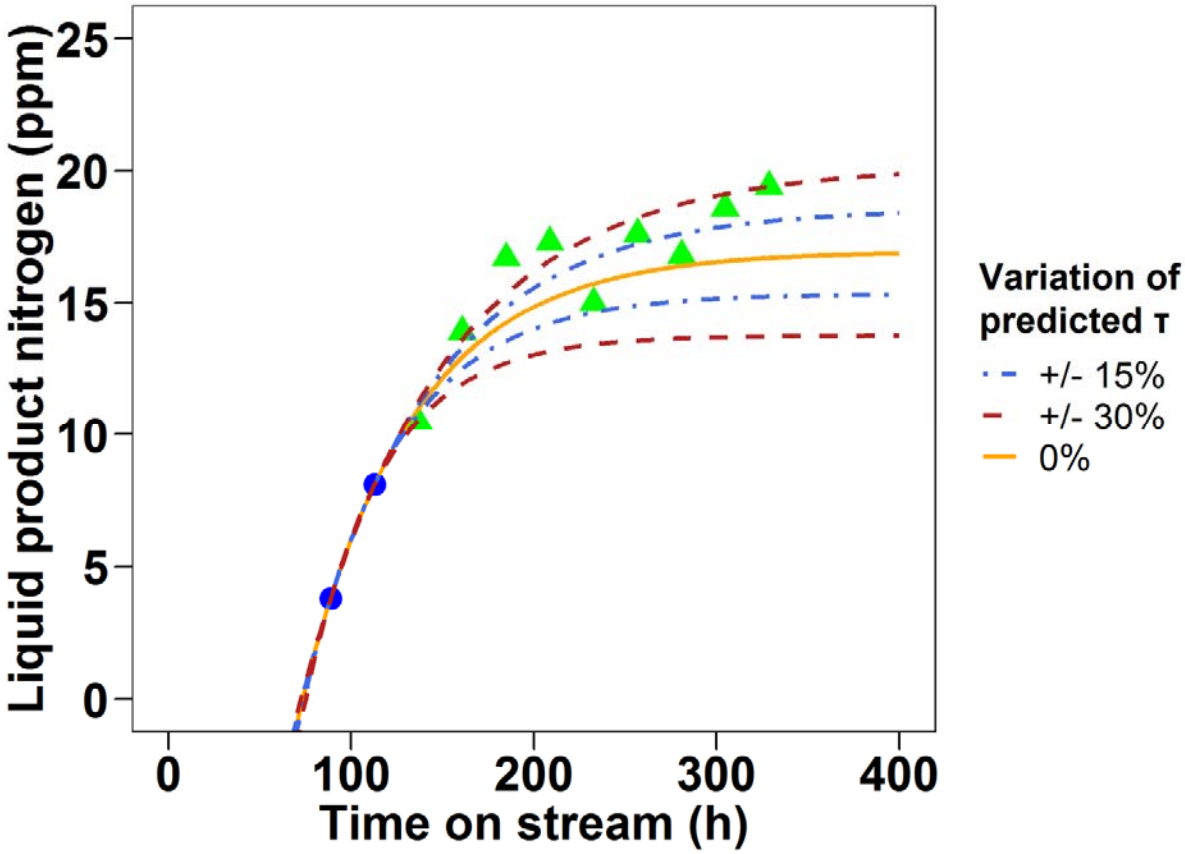


Figure 15. Sensitivity of predicted τ by model on the steady-state prediction (points: experimental data; line: prediction based on the first two points and different values of τ)

4 Conclusions

Stabilization is crucial for hydroprocessing. In this work, transient hydrotreating data were used to probe the details of this stabilization, which was found to follow a first-order response. Two models, i.e., one for the first episodes, another for the later episodes, for stabilization time 'τ' prediction were constructed. It was found that a higher LHSV leads to a quicker stabilization. The extent of the impact of LHSV on τ depends on the feed resin content. A proportional relationship between pressure and stabilization was found. Temperature is not a dominant factor. Stabilization of later episodes depends not only on the feedstock and operating conditions but also on the operating conditions of the previous episode.

The availability of two initial measurements during stabilization typically suffices to predict the steady-state hydrotreating performance. Good prediction results were obtained, particularly for the first episodes. The stabilization of later episodes is more complicated to predict. The prediction results indicate that transient data can be used to optimize the steering of pilot plants to reach quickly some targets.

Corresponding author: Benoit Celse

Email: benoit.celse@ifpen.fr

Notes: The authors declare no competing financial interest.

Nomenclature

CSTR: continuous stirred-tank reactor

LHSV: liquid hourly space velocity (h^{-1})

P: pressure (bar)

PFR: plug flow reactor

T: temperature (°C)

TOS: time on stream (h)

ppm: parts-per-million

VGO: vacuum gas oil

References

Bahmani, M., Sharifi, K., Shirvani, M., 2010. Product Yields Prediction of Tehran Refinery Hydrocracking Unit Using Artificial Neural Networks. Iranian Journal of Chemical Engineering 7.

Bates, D.M., Watts, D.G., 1988. Nonlinear Regression Analysis and Its Applications. Wiley, New York.

Becker, P.J., Celse, B., Guillaume, D., Dulot, H., Costa, V., 2015. Hydrotreatment modeling for a variety of VGO feedstocks: A continuous lumping approach. Fuel 139, 133–143. 10.1016/j.fuel.2014.08.032.

Bonnardot, J., 1998. Kinetic modelling of hydro-treatment reactions by study of different chemical groups, Lyon (France).

Charon-Revellin, N., Dulot, H., López-García, C., Jose, J., 2011. Kinetic modeling of vacuum gas oil hydrotreatment using a molecular reconstruction approach. Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles 66 (3), 479–490. 10.2516/ogst/2010005.

Doukeh, R., Bombos, M., Trifoi, A., Mihai, O., Popovici, D., Bolocan, I., Bombos, D., 2018. Kinetics of thiophene hydrodesulfurization over a supported Mo–Co–Ni catalyst. Comptes Rendus Chimie 21 (3-4), 277–287. 10.1016/j.crci.2017.07.001.

Elizalde, I., Trejo, F., Muñoz, J.A.D., Torres, P., Ancheyta, J., 2016. Dynamic modeling and simulation of a bench-scale reactor for the hydrocracking of heavy oil by using the

474 continuous kinetic lumping approach. Reac Kinet Mech Cat 118 (1), 299–311.
475 10.1007/s11144-016-0995-8.
476 Elkamel, A., Al-Ajmi, A., Fahim, M., 1999. Modeling the hydrocracking process using artificial
477 neural networks. Petroleum Science and Technology 17 (9-10), 931–954.
478 10.1080/10916469908949757.
479 Esmaeel, S.A., Gheni, S.A., Jarullah, A.T., 2016. 5-Lumps kinetic modeling, simulation and
480 optimization for hydrotreating of atmospheric crude oil residue. Appl Petrochem Res 6 (2),
481 117–133. 10.1007/s13203-015-0142-x.
482 Furnival, G.M., Wilson, R.W., 1974. Regressions by Leaps and Bounds. Technometrics 16 (4),
483 499–511.
484 Jarullah, A.T., Mujtaba, I.M., Wood, A.S., 2011. Kinetic model development and simulation of
485 simultaneous hydrodenitrogenation and hydrodemetallization of crude oil in trickle bed
486 reactor. Fuel 90 (6), 2165–2181. 10.1016/j.fuel.2011.01.025.
487 Lababidi, H.M.S., AlHumaidan, F.S., 2011. Modeling the hydrocracking kinetics of atmospheric
488 residue in hydrotreating processes by the continuous lumping approach. Energy Fuels 25 (5),
489 1939–1949. 10.1021/ef200153p.
490 Land, A.H., Doig, A.G., 1960. An Automatic Method of Solving Discrete Programming
491 Problems 28, 497–520.
492 López García, C., Hudebine, D., Schweitzer, J.-M., Verstraete, J.J., Ferré, D., 2010. In-depth
493 modeling of gas oil hydrotreating: From feedstock reconstruction to reactor stability analysis.
494 Catalysis Today 150 (3-4), 279–299. 10.1016/j.cattod.2009.08.002.
495 Nguyen, M.-T., Pirngruber, G.D., Chainet, F., Tayakout-Fayolle, M., Geantet, C., 2017. Indole
496 hydrodenitrogenation over alumina and silica–alumina-supported sulfide catalysts—

497 Comparison with quinoline. Ind. Eng. Chem. Res. 56 (39), 11088–11099.
498 10.1021/acs.iecr.7b02993.
499 Nguyen, M.-T., Tayakout-Fayolle, M., Pirngruber, G.D., Chainet, F., Geantet, C., 2015. Kinetic
500 modeling of quinoline hydrodenitrogenation over a NiMo(P)/Al2O3 catalyst in a batch
501 reactor. Ind. Eng. Chem. Res. 54 (38), 9278–9288. 10.1021/acs.iecr.5b02175.
502 Oliveira, L.P. de, Hudebine, D., Guillaume, D., Verstraete, J.J., Joly, J.F., 2016. A Review of
503 Kinetic Modeling Methodologies for Complex Processes. Oil Gas Sci. Technol. – Rev. IFP
504 Energies nouvelles 71 (3), 45. 10.2516/ogst/2016011.
505 Oyekunle, L.O., Edafe, O.A., 2009. Kinetic modeling of hydrodenitrogenation of pyridine.
506 Petroleum Science and Technology 27 (6), 557–567. 10.1080/10916460802104156.
507 Raghuveer, C.S., Thybaut, J.W., Marin, G.B., 2016. Pyridine hydrodenitrogenation kinetics over
508 a sulphided NiMo/γ-Al2O3 catalyst. Fuel 171, 253–262. 10.1016/j.fuel.2015.12.042.
509 Sadighi, S., Mohaddecy, S.R.S., 2018. Evaluating the ability of R for modeling a commercial
510 scale VGO hydrocracking plant using Artificial Neural Network (ANN). Petroleum and Coal
511 60 (3), 358–364.
512 Sadighi, S., Reza Zahedi, G., 2013. Comparison of kinetic-based and artificial neural network
513 modeling methods for a pilot scale vacuum gas oil hydrocracking reactor. Bull. Chem. React.
514 Eng. Catal. 8 (2). 10.9767/bcrec.8.2.4722.125-136.
515 Sau, M., Basak, K., Manna, U., Santra, M., Verma, R.P., 2005. Effects of organic nitrogen
516 compounds on hydrotreating and hydrocracking reactions. Catalysis Today 109 (1-4), 112–
517 119. 10.1016/j.cattod.2005.08.007.

518 Schweitzer, J.-M., Galtier, P., Schweich, D., 1999. A single events kinetic model for the
519 hydrocracking of paraffins in a three-phase reactor. Chemical Engineering Science 54, 2441–
520 2452. 10.1016/S0009-2509(99)00006-8.
521 Smith, C.A., Campbell, S.W., 2016. A First Course in Differential Equations, Modeling, and
522 Simulation. Taylor & Francis Group, Boca Raton.
523 Tang, X., Li, S., Yue, C., He, J., Hou, J., 2013. Lumping kinetics of hydrodesulfurization and
524 hydrodenitrogenation of the middle distillate from chinese shale oil. Oil Shale 30 (4), 517.
525 10.3176/oil.2013.4.05.
526 Treese, S.A., Pujadó, P.R., Jones, D.S.J. (Eds.), 2015. Handbook of Petroleum Processing.
527 Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA.
528

Figures & Tables

Figures

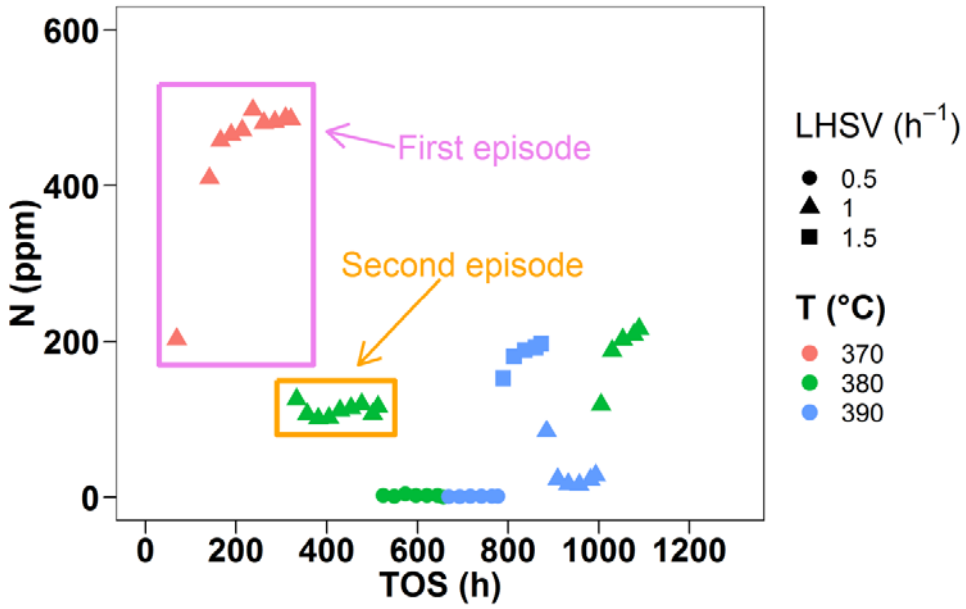


Figure 1. Data illustration - Liquid effluent nitrogen as a function of time on stream (TOS).

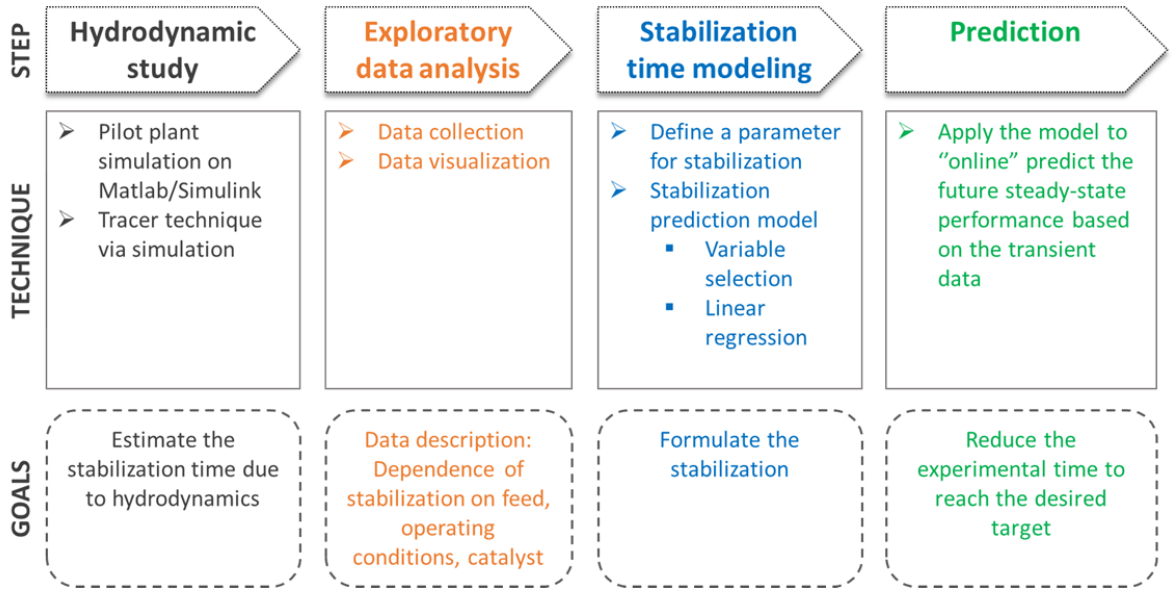
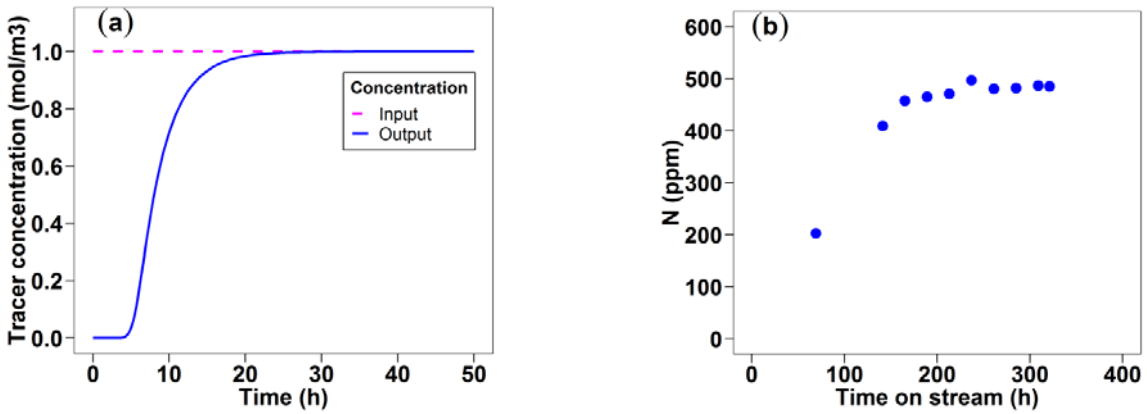
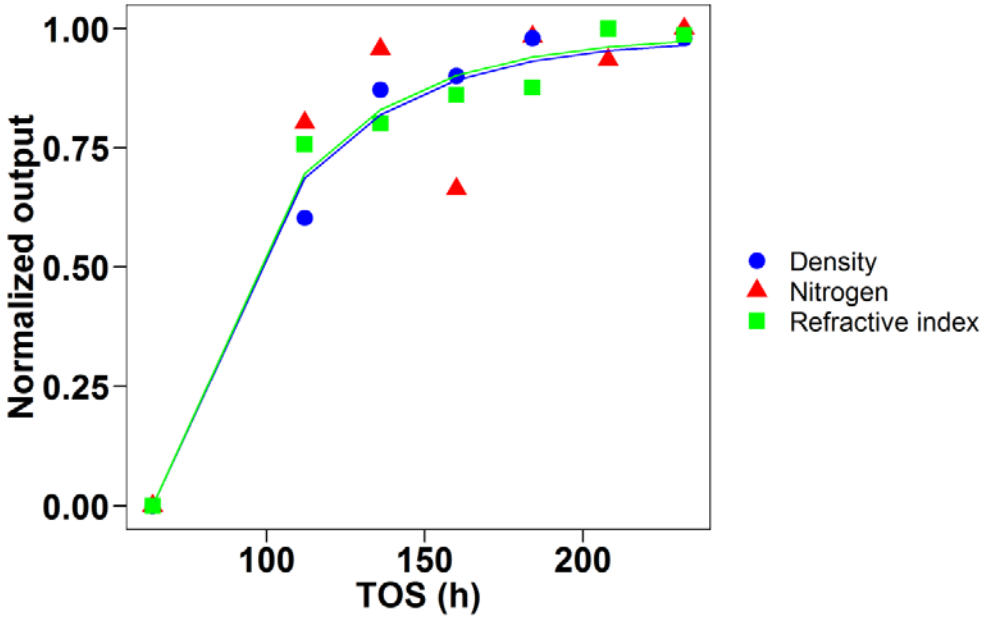


Figure 2. Methodology adopted throughout this work.



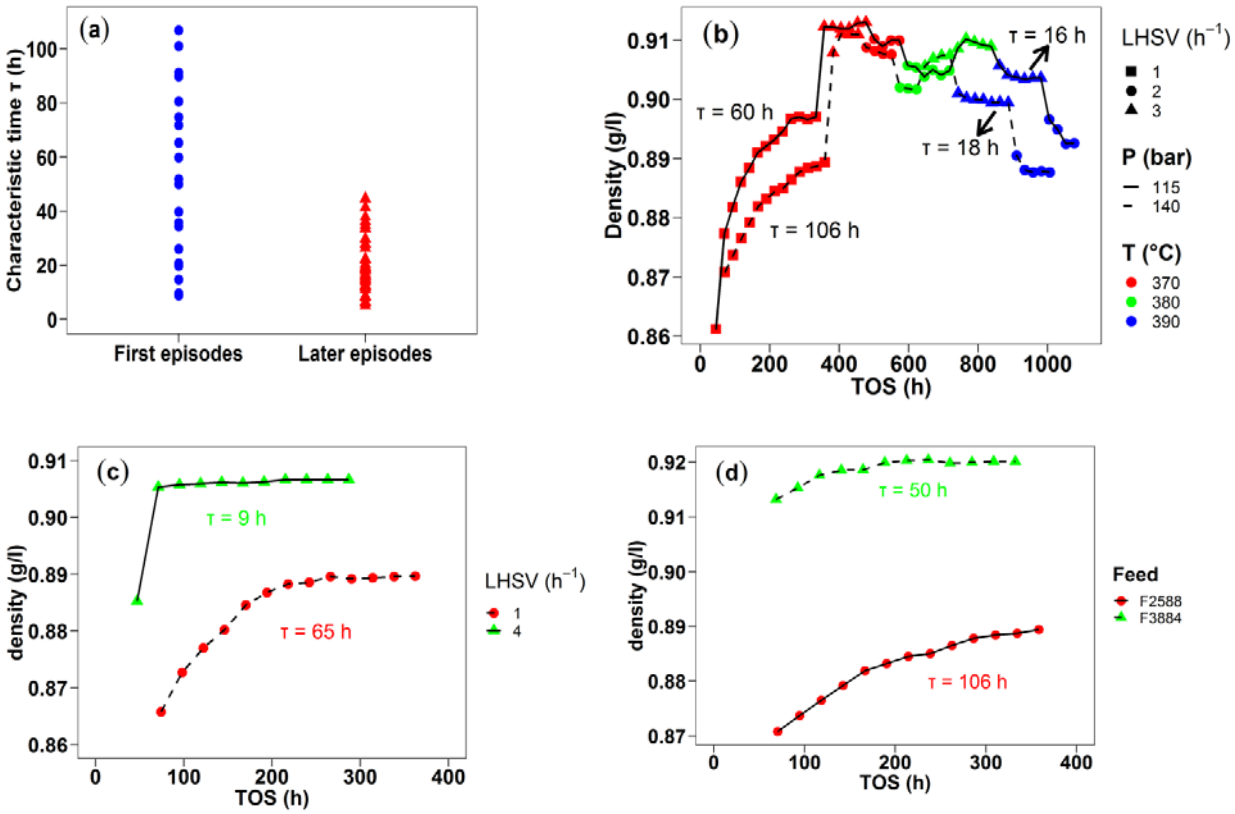
535

536 Figure 3. Calculated effluent concentration profiles of tracer (a) and experimentally observed
537 liquid effluent nitrogen content as a function of time on stream in the pilot plant (b) (LHSV = 1
538 h^{-1}, T = 370 °C, P = 140 bar).

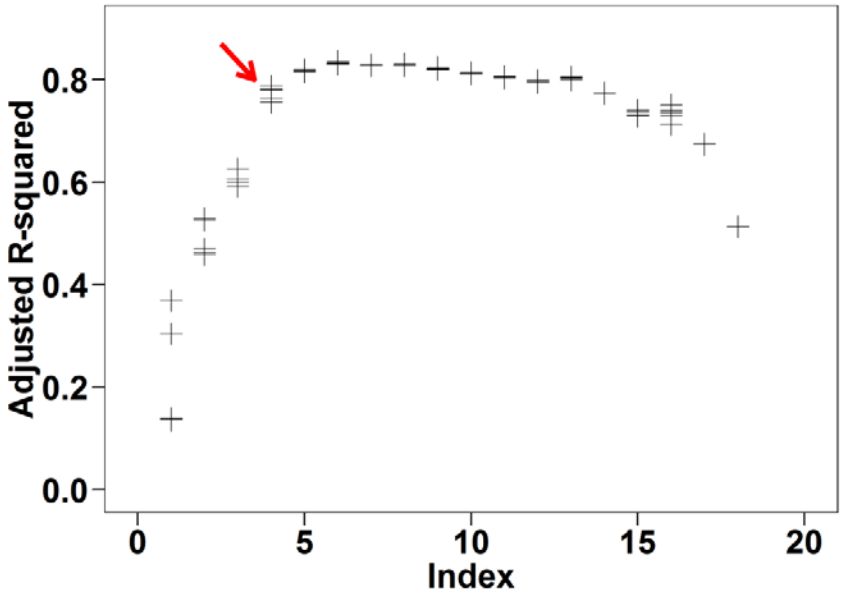


539

540 Figure 4. Comparison of transient behavior of nitrogen/density/refractive index (outputs are
541 normalized between [0, 1]). Values of τ are 39.8, 36.3, 39.2 h calculated respectively from
542 density, nitrogen content and refractive index evolution.

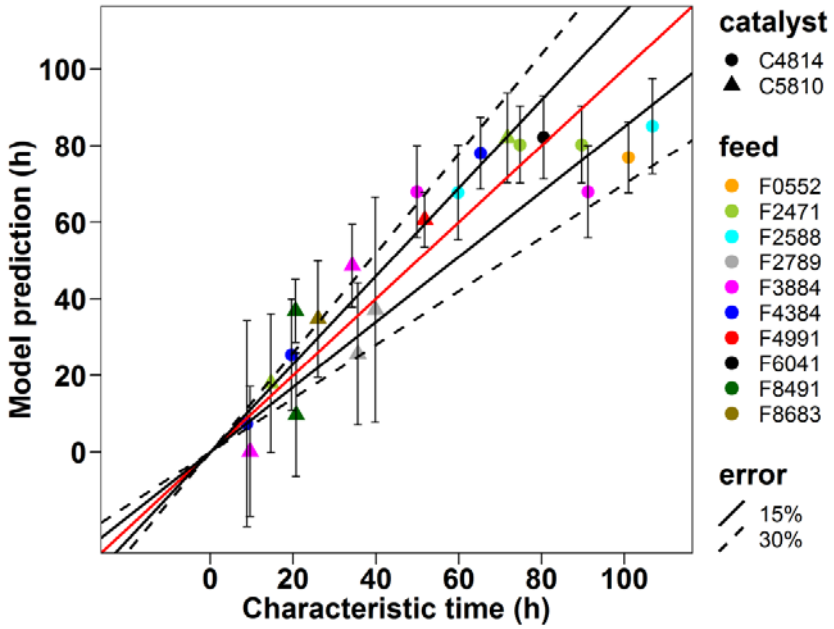


543 Figure 5. (a) Classification of τ regarding the type of episode; (b, c, d) Impact of pressure (solid
544 line: 115 bar, dashed: 140 bar, feed F2588); LHSV (feed F4384, P = 140 bar, T = 370 °C) and
545 feedstock (LHSV = 1 h^{-1}, P = 140 bar, T = 370 °C) respectively on transient behavior of liquid
546 effluent density.



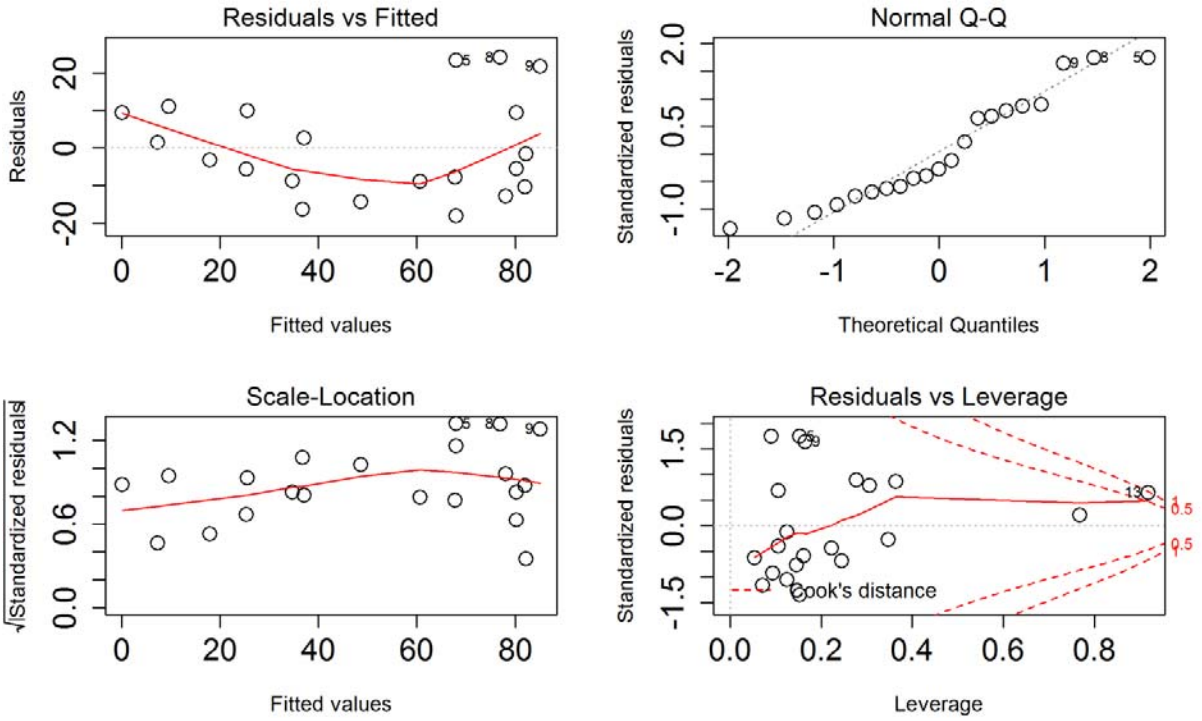
547

548 Figure 6. Variable selection using leaps for linear model with interaction (first episodes).



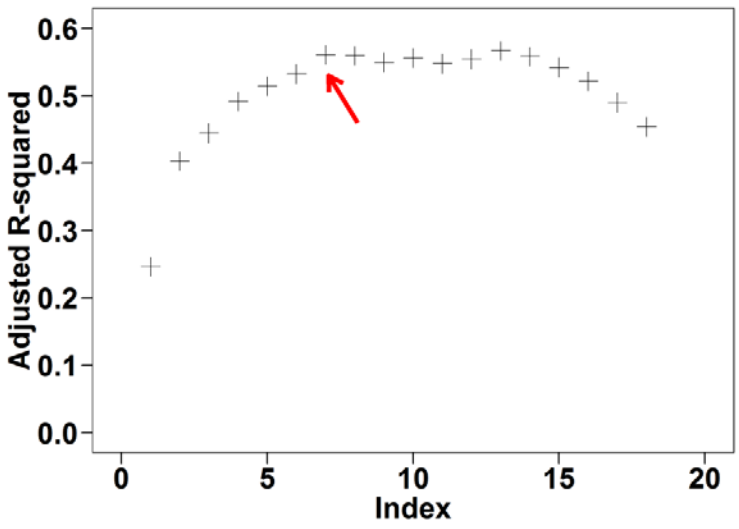
549

550 Figure 7. Parity plot with 95% confidence interval ($R^2 = 0.83$) (first episode model).



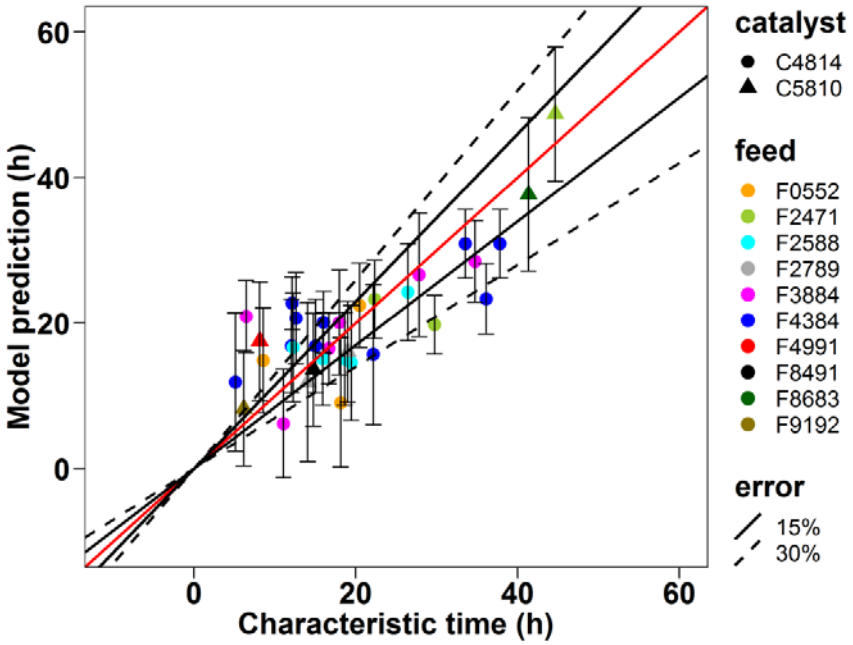
551

552 Figure 8. Diagnostic plots for linear model with interaction (first episodes model).



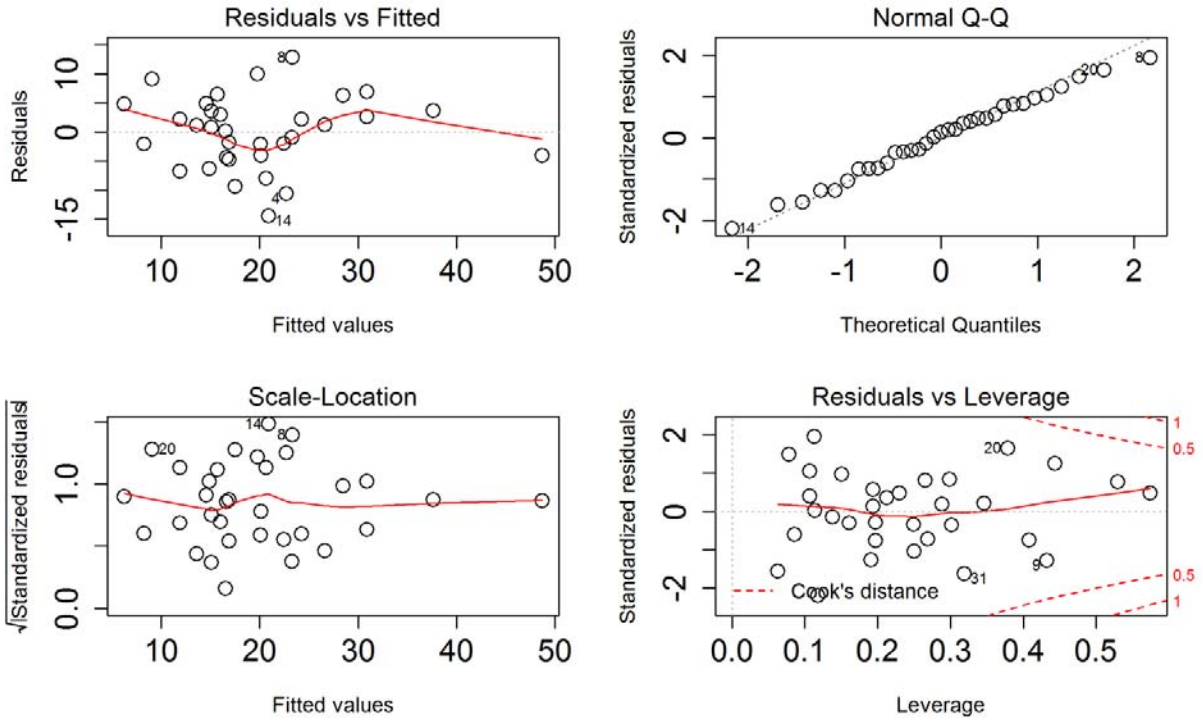
553

554 Figure 9. Variable selection using leaps for later episode model.



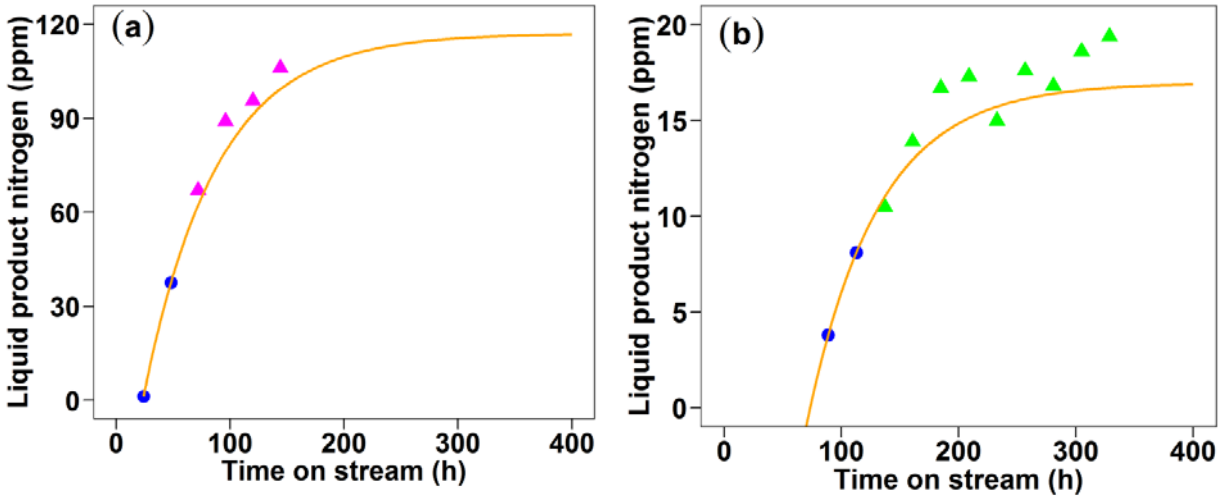
555

556 Figure 10. Parity plot with 95% confidence interval (R^2 = 0.66) (later episodes model).



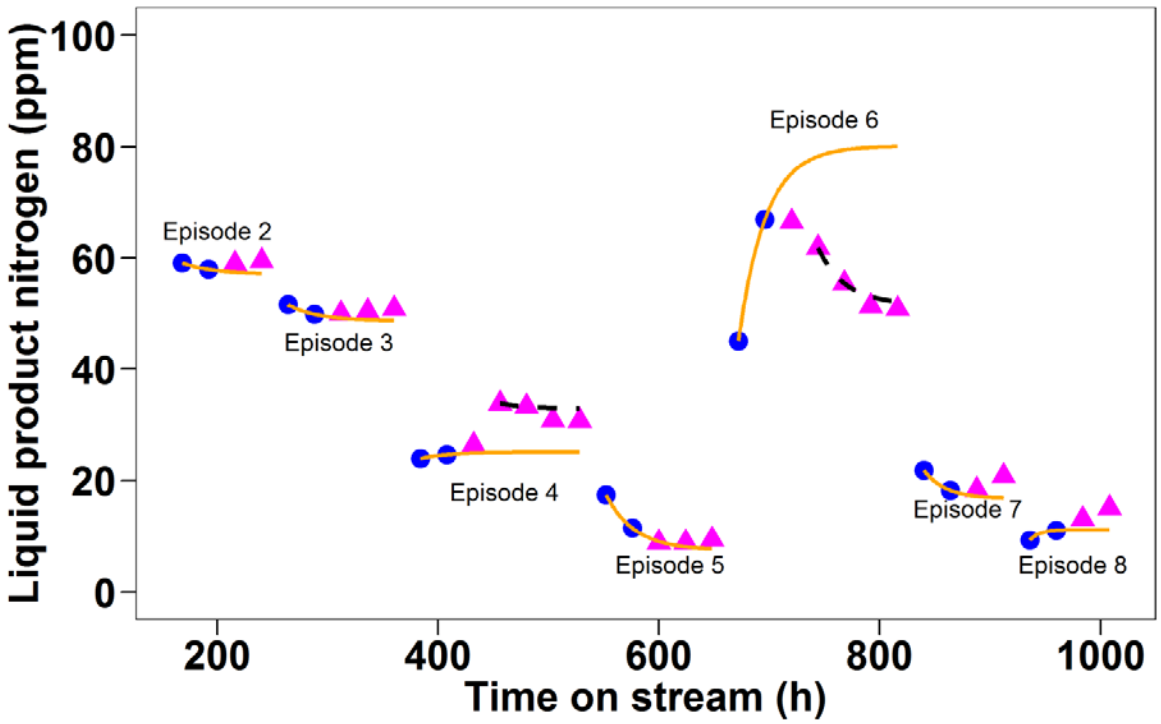
557

558 Figure 11. Diagnostic plots for linear model with interaction (later episodes model).



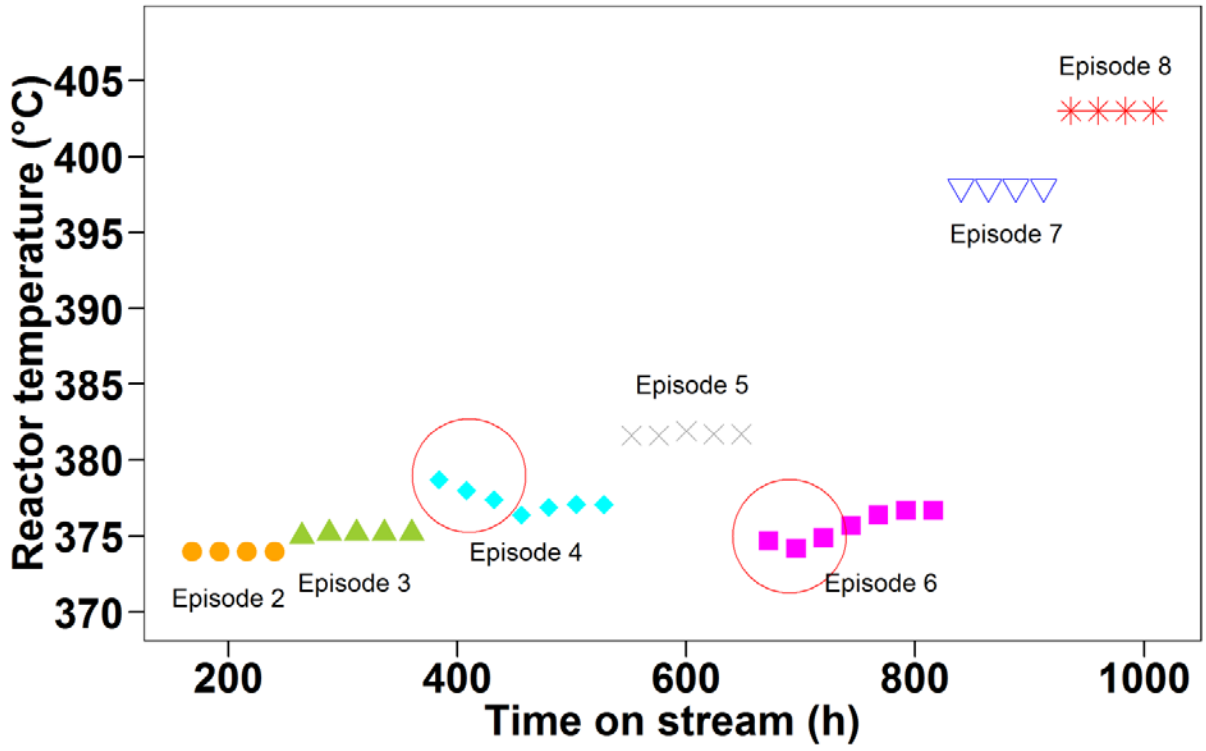
559 Figure 12. Two test predictions using the first two points of episode (points: experimental data,
560 solid line: model prediction; Test (a): LHSV = 1.34 h^{-1}, P =132 bar, T = 370 °C; Test (b): LHSV
561 = 1.71 h^{-1}, P =140 bar, T = 386 °C).

562



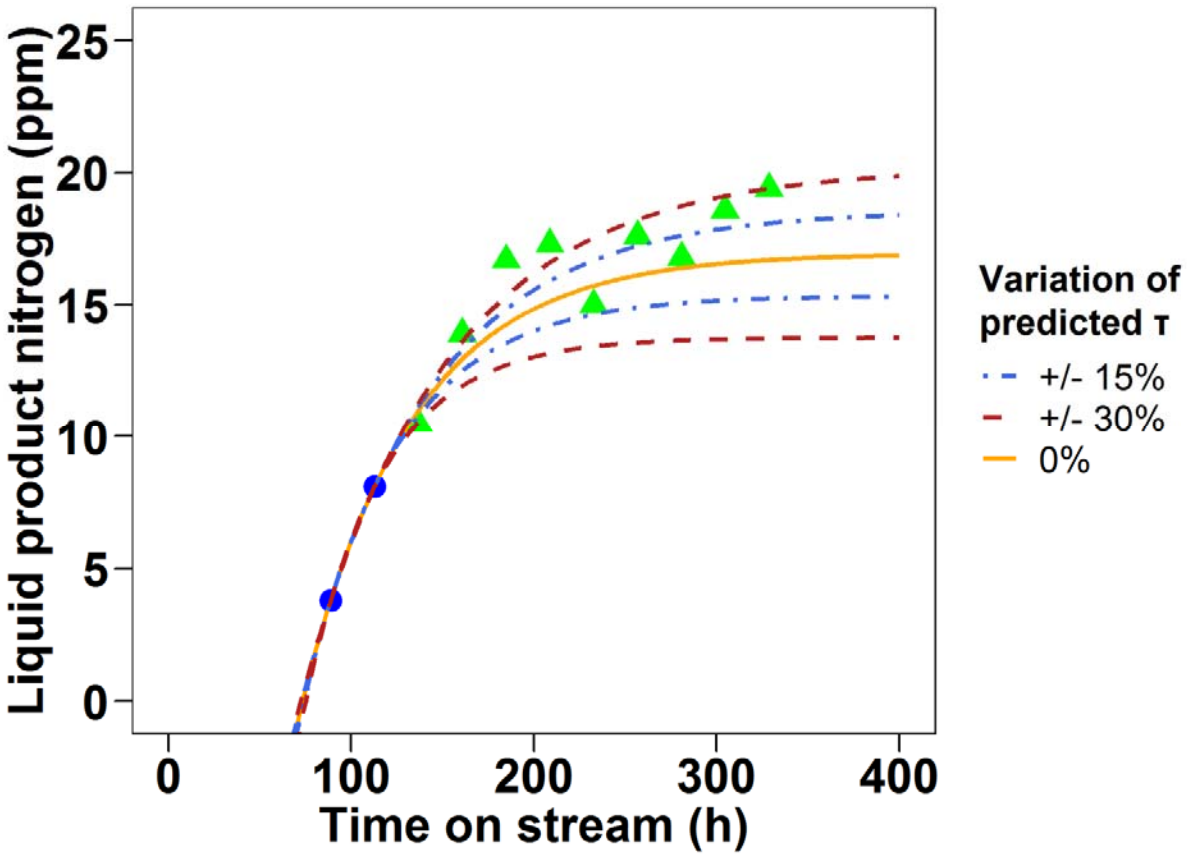
563

564 Figure 13. Prediction with the first two points of other episodes (points: experimental data, solid
565 line: model prediction using the first two points, dashed line: prediction while discarding the first
566 three points).



567

568 Figure 14. Reactor temperature profile corresponding to each episode



569

570 Figure 15. Sensitivity of predicted τ by model on the steady-state prediction (points :

571 experimental data; line: prediction based on the first two points and different values of τ)

572

573 **Tables**

574 Table 1. Selected variables corresponding to each subset obtained from variable selection in
575 Figure 6 (only show the best selected subsets of 'from 1 to 7' variables and 5 best subsets of 4
576 variables)

Index	Selected variables	Adjusted R^2
1	P	0.369
2	P, LHSV*S_{feed}	0.528
3	T, P, LHSV*S_{feed}	0.625
4a	**LHSV, P, res_{feed}, LHSV*res_{feed}**	**0.788**
4b	LHSV, res_{feed}, LHSV*res_{feed}, P*TMP_{feed}	0.782
4c	res_{feed}, LHSV*res_{feed}, LHSV*TMP_{feed}, P*TMP_{feed}	0.780
4d	LHSV, P, LHSV*res_{feed}, T*res_{feed}	0.763
4e	LHSV, LHSV*res_{feed}, T*res_{feed}, P*TMP_{feed}	0.756
5	LHSV, T, P, res_{feed}, LHSV*res_{feed}	0.819
6	P, res_{feed}, LHSV*res_{feed}, LHSV*TMP_{feed}, T*res_{feed}, P*N_{feed}	0.835
7	LHSV, P, N_{feed}, LHSV*Res_{feed}, T*TMP_{feed}, T*N_{feed}, P*N_{feed}	0.828

577

578

579 Table 2. Coefficient and statistical values for first episodes model

	Intercept	LHSV	P	res_{feed}	LHSV*res_{feed}
Coefficient	63.339	-69.055	0.693	-7.081	5.481
p-value	3.33×10^{-3}	3.15×10^{-6}	7.1×10^{-6}	5.31×10^{-5}	1.9×10^{-4}

580

Global p-value	F	$F_{95\%}$
5.3×10^{-6}	19.55	3.01

581

582 Table 3. Variables used in variable selection technique for later episodes regression

Feed	**Operating conditions**	**Condition switching**	**Interaction terms**
N_{feed}	LHSV	$\frac{T}{T_{pre}}$	LHSV*res_{feed}
S_{feed}	T	$\frac{LHSV}{LHSV_{pre}}$	LHSV*N_{feed}
res_{feed}	P	$\frac{T_{pre}}{T}$	LHSV*S_{feed}
d_{feed}	$LHSV_{pre}$	$\frac{LHSV_{pre}}{LHSV}$	
TMP_{feed}	T_{pre}		

583

584

585 Table 4. Coefficient and statistical values for later episodes model

	Intercept	LHSV	S_{feed}	$LHSV_{pre}$
Coefficient	177.349	-10.943	2.589	8.851
p-value	8.32×10^{-5}	0.036	0.114	0.071

586

	T_{pre}	$LHSV*res_{feed}$	$\frac{LHSV}{LHSV_{pre}}$	$\frac{LHSV_{pre}}{LHSV}$
Coefficient	-0.379	-0.828	8.832	-11.713
p-value	8.34×10^{-4}	3.85×10^{-3}	5.78×10^{-3}	0.020

587

Global p-value	F	$F_{95\%}$
1.4×10^{-4}	6.83	2.40

588

589