



Identification and Observability of Lumped Kinetic Models for Vacuum Gas Oil Hydrocracking

Zoltán Till, Tamás Varga, Laura Szabó, Tibor Chován

► To cite this version:

Zoltán Till, Tamás Varga, Laura Szabó, Tibor Chován. Identification and Observability of Lumped Kinetic Models for Vacuum Gas Oil Hydrocracking. *Energy & Fuels*, 2017, 31 (11), pp.12654-12664. 10.1021/acs.energyfuels.7b02040 . hal-02017440

HAL Id: hal-02017440

<https://hal.science/hal-02017440>

Submitted on 13 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

This is the authors' version of the paper titled

Identification and Observability of Lumped Kinetic Models for Vacuum Gas Oil Hydrocracking

To get the formatted version, please use the following DOI: [10.1021/acs.energyfuels.7b02040](https://doi.org/10.1021/acs.energyfuels.7b02040).

Citation:

- MLA Till, Zoltán, et al. "Identification and Observability of Lumped Kinetic Models for Vacuum Gas Oil Hydrocracking." *Energy & Fuels* 31.11 (2017): 12654-12664.
- APA Till, Z., Varga, T., Szabó, L., & Chován, T. (2017). Identification and Observability of Lumped Kinetic Models for Vacuum Gas Oil Hydrocracking. *Energy & Fuels*, 31(11), 12654-12664.
- ISO 690 TILL, Zoltán, et al. Identification and Observability of Lumped Kinetic Models for Vacuum Gas Oil Hydrocracking. *Energy & Fuels*, 2017, 31.11: 12654-12664.



[University of Pannonia](#)



[Faculty of Engineering](#)



[Department of Process Engineering](#)

Identification and Observability of Lumped Kinetic Models for Vacuum Gas Oil Hydrocracking

Zoltán Till^{,†}, Tamás Varga[†], Laura Szabó[‡], Tibor Chován[†]*

[†] Department of Process Engineering, University of Pannonia, 10 Egyetem Street, Veszprém H-8200, Hungary

[‡] MOL Plc., Október Huszonharmadika Str. 18., Budapest, Hungary

* Email: tillz@fmt.uni-pannon.hu

[Supporting information](#)

The so-called “lumping approach” is widely used to study complex processes such as hydrocracking of vacuum residue. In order to describe the composition changes in such systems, not only must the kinetic parameters be determined, but the lumped reactions that occur should also be identified. In this study, the modeling of catalytic hydrocracking of vacuum gas oil has been carried out using six component lumps. Three different identification strategies have been developed to determine the reaction subnetwork containing a given number of reaction pathways that provides the data fit. The strategies were compared according to their tendency to provide increasingly better results, as a function of the number of reactions present. Although, in this way, 40% of the original reaction superstructure was eliminated from the system, the kinetic parameters of the remaining reactions still could not be identified with complete certainty. Hence, the linearized state-space model representations of the reaction networks have been further analyzed with the objective of identifying observable subsystems. The results show that there exists a five-reaction network that is observable and can be determined using the finally proposed identification strategy, while its curve fitting is also satisfactory. It has also been shown that it gives the best results from all possible five-reaction subsystems.

Introduction

The hydrocracking process, which involves catalytic cracking of long-chained hydrocarbon molecules under a hydrogen atmosphere, has a growing importance in the oil industry. Important nonconventional feedstocks such as vacuum gas oil (VGO), various vegetable oils, or even waste cooking oil can be utilized to produce fuels, mainly biogas oil with high quality.¹ Another advantage is that these can be co-processed and might be used in blends in different refinery technologies.^{2,3} The importance of fuels from alternative feedstocks have significantly increased with the oil prices rising in the 2000s, and, although the market situation has since normalized, the so-called biofuels still have great importance. Beyond that, there are environmental as well as legal reasons, e.g., there is a mandatory target of a 20% share of energy from renewable sources in overall energy consumption in the European Union by 2020.⁴ When calculating this share, plant-derived motor fuels count as a renewable source, and most of that is consumed in conventional and renewable source blends.

Chemical kinetic modeling becomes increasingly difficult for complex processes such as the hydrocracking of VGO, where several thousands of individual species can be present and, between them, an even larger number of reactions can occur. While concentration measurement for individual components is a routine task nowadays and one can generate full reaction networks automatically, the identification of all kinetic parameters and the subsequent model reduction is usually not a viable method to find a solution. Opposed to that, there are two common methods for treating these systems:⁵ a priori lumping, which is carried out based on empirical rules such as constraining the total number of species and/or reactions, and a posteriori lumping, where the reaction network is generated first (although its parameters are not identified) and the component grouping is carried out based on the properties of the reaction network.

The lumping approach is a widely used method for largescale reaction networks. A great deal of reported applications comes from the oil industry. A five-lump kinetic model for the hydrocracking of heavy oils under moderate conditions was proposed by Sánchez et al., which was capable of predicting component concentrations with an average absolute error of <5% at temperatures of 380–420 °C and liquid hourly space velocity (LHSV) values of 0.33–1.5 h⁻¹.⁶ The effect of pressure on the kinetics of hydrocracking can also be studied with the lumping approach.⁷ In addition, Nguyen et al. developed a structure-based kinetic model consisting of 16 lumps for the hydrodesulfurization of light gas oil, which can predict the composition of the reactor outlet and provides information about chemical structures as well.⁸ An exhaustive review of heavy petroleum fraction hydrocracking, lumped reaction schemes, and kinetic data has been reported by Ancheyta et al.⁹ The hydrocracking of VGO is usually carried out in a fixed-bed tubular reactor, but other configurations, i.e., ebullated bed reactors (EBRs), can be applied as well.¹⁰

The process of catalytic upgrading of fuels, such as gasoline olefin content removal, can also be modeled with the lumping approach.¹¹ Wang et al. applied lumped kinetic simulation to optimize catalyst grading in shale oil hydrotreatment.¹² While the majority of publications involve quasi-homogeneous phase models, the lumped kinetic modeling approach is applicable for describing multiphase systems in detail as well.^{13,14} A detailed, two-dimensional, nonisothermal, heterogeneous model was established by Forghani et al., by applying two different reaction kinetic networks between four lumps that is also applicable for scale-up design of green diesel production.¹⁵

Lumped kinetic models can also be used in the case of treating various oils from renewable sources, such as biomass tar cracking,¹⁶ catalytic cracking of vegetable oils,^{17,18} or waste cooking oil.¹⁹ Moreover, this approach is not limited to modeling reactions in hydrocarbon mixtures; lumping

techniques can be used in biochemistry as well, for example, for predicting metabolism or modeling of fermentation processes.^{20,21} It can be assumed that the suggested strategies for lumped reaction network analysis and reduction reported here may be used in any of the previous fields with little to no modification.

In the case of chemical reaction networks, the objective of model reduction is to identify a reaction subsystem as sparse as possible for which the calculated concentration values still show reasonable agreement, compared to the full kinetic network. If the reaction system is dense, i.e., a large number of reactions are supposed to occur, there is an emerging possibility that the system becomes overspecified, which means that practically that two or more identified kinetic parameter sets are equivalent and lead to the same composition profiles, as will later be shown.

In order to investigate how reliable the results based on a specific dataset are, determination of the observability of the system can be useful. The concept of observability, i.e., whether the states of a system can be observed (and, consequently, estimated), comes from control theory, where Bayesian state estimators, most notably the Kalman filter and its variants, are widely used to predict the states of the system from model output and measurements using state space models.²² The system is called “observable” if the values of state variables can be determined within a finite time from the values of outputs.^{23,24} State estimators can be effectively used to determine kinetic parameters from measured variables (such as concentrations or temperature) in an observable system.^{25,26} Despite this, the theoretical observability and identifiability of the suggested lumped reaction kinetic networks are almost never studied; hence, the results obtained using lumped kinetic models are often not applicable, for example, for process intensification.

Another way to overcome this problem is using the continuous lumping method to model the chemical system. In that case, the reaction mixture is represented by a continuous function (such as a function of true boiling point (TBP)) that is then discretized in order to recover the concentration of the sought pseudo-components (defined by TBP range).^{27,28} The advantage of this approach is that any number of lumps can be defined and the reaction rate coefficient can be correlated to the normalized TBP, thus reducing the number parameters to be identified.

In the following subsections, different reaction network reduction strategies are presented and compared for a priori selected lumps for the hydrocracking of VGO. The reduced networks are further investigated regarding system observability to determine whether the model parameters can be identified without any uncertainties. The proposed procedure is an effective tool to determine realizable models for the hydrocracking of VGO that are useable in plant modeling or process development.

Reactor Model

A quasi-homogeneous phase plug flow model of a pilot-scale hydrocracking reactor operating under isothermal conditions and its parameters were reported earlier by Sadighi et al. It applied six lumped components, i.e., unconverted vacuum gas oil or residue (VGO), diesel (D), kerosene (K), heavy naphtha (HN), light naphtha (LN), and gas (G).^{29,30} The objective of the present work is to analyze the transitions between the lumped components; therefore, the reported experimental data under different conditions are applied here without any modifications. Assuming a steady-state operation and constant catalytic activity, the component mass balance equations of the system can be formalized as follows:

$$\frac{d(\underline{c} \cdot v)}{d\ell} = \eta \cdot \varepsilon' \cdot \underline{L} \cdot \underline{c} \quad (1)$$

where ℓ is the dimensionless length of the reactor, η is the effectiveness factor of the heterocatalytic reactions, and ε' the catalyst volume fraction. The values of η and ε' are 0.8 and 0.264, respectively; each of these values was taken from the data supplied in the original article. Since the catalyst suffers from deactivation during the hydrocracking of heavy distillates, this assumption is only valid for a limited time operation. c is the component concentration column vector:

$$\underline{c} = [c_{VGO} \ c_D \ c_K \ c_{HN} \ c_{LN} \ c_G]^T \quad (2)$$

The actual space velocity (ν) is calculated from the LHSV, considering the density change along the reactor length (see eqs 3–5). LHSV values reported by Sadighi et al. were 0.5, 1, 1.5, and 2 h⁻¹.²⁹

$$\nu = LHSV \cdot \frac{\rho_{in}}{\rho_{act}} \quad (3)$$

$$\rho_{act} = \sum_{j=1}^6 (Y_j \cdot \rho_j) \quad (4)$$

$$Y_i = \frac{c_i \cdot \nu}{\sum_{j=1}^6 (c_j \cdot \nu)} \quad (5)$$

Here, $\underline{L}$ is the Kirchhoff matrix of the reaction system.³¹ For six components, L is a 6-by-6 square matrix, the diagonal elements of which represent the reaction rate coefficients, where the i^{th} component is consumed (hence, the negative sums in the main diagonal), while the off-diagonal elements contains the rate coefficients of the reactions where the i^{th} product is formed from the j^{th} reactant. Because of the mass conservation law, $\underline{L}$ is a column conservation matrix (i.e., the element summary for each column is zero).

$$\underline{L} = a_{ij} = \begin{cases} -\sum_l k_{il} & \text{if } i = j \\ k_{ji} & \text{if } i \neq j \end{cases} \quad (6)$$

The initial or complete reaction network (also reported by Sadighi et al.^{29,30}) consists of 15 reactions (Figure 1), including all pathways that are encountered during that chemical decomposition. Each reaction is hypothesized to be a pseudo-first-order reaction with Arrhenius-type temperature dependency of the rate coefficient.

$$k_i = k_{0,i} \cdot \exp\left(-\frac{E_{a,i}}{RT}\right) \quad (7)$$

where $k_{0,i}$ is the pre-exponential factor and $E_{a,i}$ is the activation energy of the i^{th} reaction (15 in total), R is the gas constant, and T is the absolute temperature of the reactor. The steady-state reactor model was implemented and solved in MATLAB R2011b using the single-step, second-order solver based on a modified Rosenbrock formula called ode23s.³²

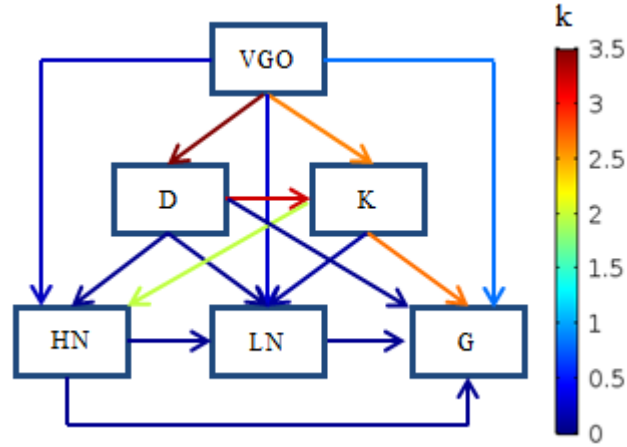


Figure 1. Initial (complete) reaction network consisting 6 component lumps and 15 lumped reactions.

The reactor component balance equations expressed in eq 1 are similar to a state-space model in structure that has a general form as described in eqs 8 and 9. Given that a plug-flow reactor model can be rewritten as a cascade of continuous stirred-tank reactors, the concentration changes along the axial coordinate of the reactor are mathematically equivalent to time derivatives. Therefore, in our case, residence time was defined as the ratio of reactor length and actual space velocity, eliminating the necessity of model conversion.

$$\frac{dx}{dt} = \underline{A} \cdot \underline{x} + \underline{B} \cdot \underline{u} \quad (8)$$

$$\underline{y} = \underline{C} \cdot \underline{x} \quad (9)$$

In the state-space model:

- $\underline{A}$ is the system matrix $\underline{A} = \eta \cdot \varepsilon' \cdot \underline{L}$
- $\underline{x}$ is the state vector $\underline{x} = \underline{c}$,
- $\underline{y}$ is the output vector ($\underline{y} = \underline{c}$; the concentrations of all lumps are measured),
- $\underline{C}$ is the output matrix $\underline{C} = \underline{I}$, and
- the term $\underline{B} \cdot \underline{u}$ is negligible (there are no inputs).

This state space model might be extended to include the reaction rate coefficients as states as well. In this case, the length of the state vector increases with the number of reactions present in the system (a maximum of 15, albeit that gives space to model reduction), shown in eq 10. The output vector remains the same, because the rate coefficients are not measured, while the output matrix is expanded with zero vectors to accommodate to the increased number of states.

$$\underline{x}^e = [\underline{c}^T \quad \underline{k}^T]^T \quad (10)$$

One of the main consequences of the extension of the state-space model that it becomes nonlinear, because $\underline{L}$ contains state variables in this interpretation. Hence, it should be linearized in order to study the observability and identifiability of the system. This step was carried out using complex step differentiation, which has been described elsewhere,³³ that is well-known for its accuracy, even for small step sizes.

Identification Strategies

The main objective is to minimize the difference between the measured and calculated compositions, which is defined as

$$f(\underline{x}^n) = \sum_T \sum_{comp} \sum_{LHSV} \left(\frac{y_m - y_c}{y_m^{max}} \right)^2 \quad (11)$$

where y_m and y_c are the measured and calculated mass concentration values, respectively, while y_m^{max} represents the maximum measured mass concentration value for a specific component (*comp*) including all temperature (T) and $LHSV$ values. Measured mass concentration values originate from the previously mentioned work of Sadighi et al.²⁹ To solve the optimization problem, the NOMAD software package was used, which implements the Mesh Adaptive Direct Search (MADS) algorithm and is a well-suited solver for derivative-free optimization.³⁴ It also has a MATLAB interface available that can be called directly from the OPTI Toolbox.³⁵ The objective function variables are the reaction kinetic parameters; however, one must address the fact that not all reactions are necessarily present in the system. Considering the full and all possible reduced networks, there are 32 777 possible variants that cannot be evaluated on a reasonable time horizon. The solution presents itself as to assign existence variables to each reaction and solve the obtained MINLP problem; however, since the objective function for a lumped kinetic reaction network is highly nonlinear and no a priori information is available for the reactions present, finding the best solution as the global minimum is at least time-consuming and involves various trial-and-error searches, regarding the algorithm parameters, such as the initial function value.

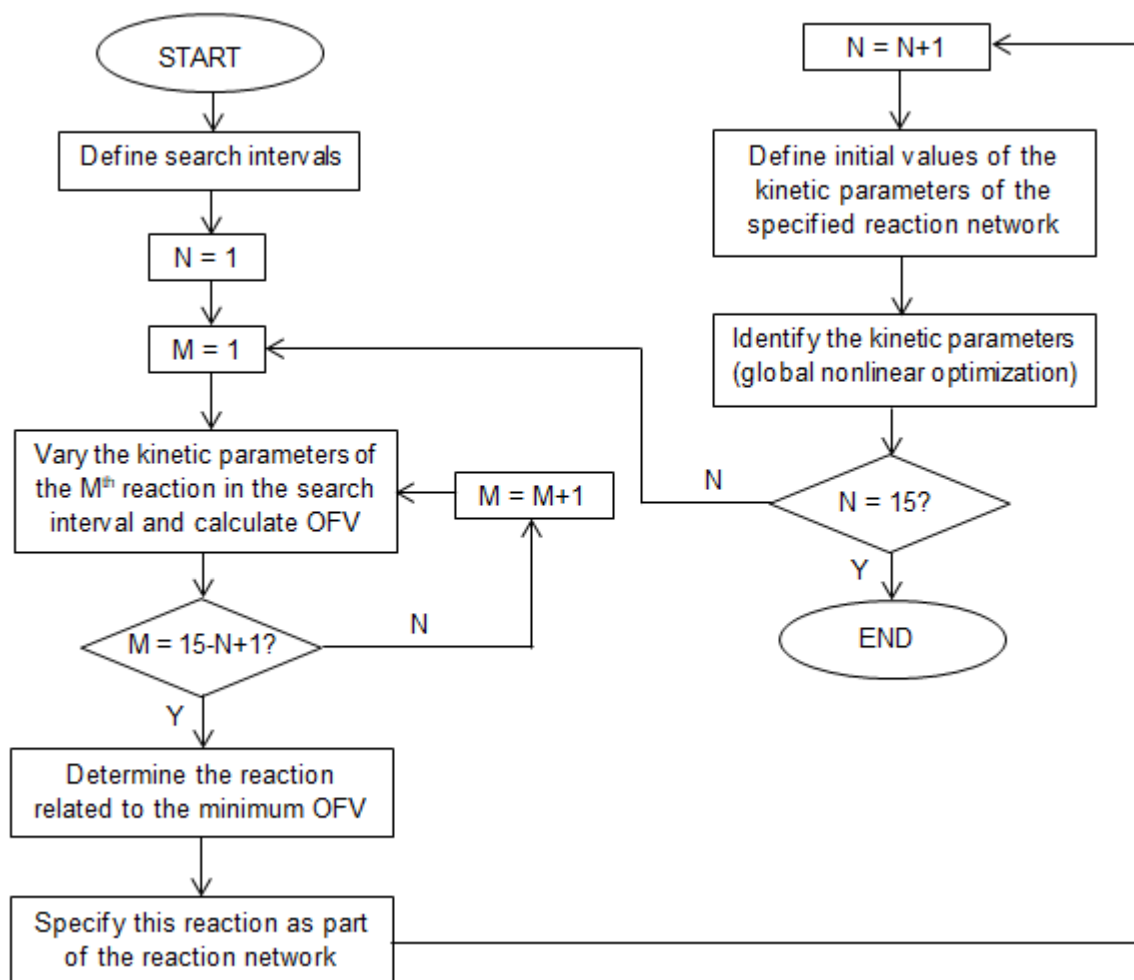


Figure 2. First reaction network identification strategy.

In order to overcome that obstacle, three different approaches were evaluated. The brief concept of the first, sensitivity-based selection strategy (shown in Figure 2) is described as follows:

- (1) For each reaction, the objective function value was calculated in a sensitivity study by varying the kinetic parameters of that reaction only from 1% to 200% of the middle of the related search intervals.
- (2) The reaction associated with the minimum objective function value was specified as part of the reaction network.
- (3) The kinetic parameters of the reaction network were identified by minimizing the function value from eq 11. That means the dimension of the search space was increased by 2 in each step (from 2 to 30 in total). The results from the previous run were used as initial values for the algorithm in the next iteration.
- (4) The specified reaction was marked so that it does not participate in further sensitivity studies.
- (5) The second, third, ..., 15th reaction was identified by repeating steps 1–4.

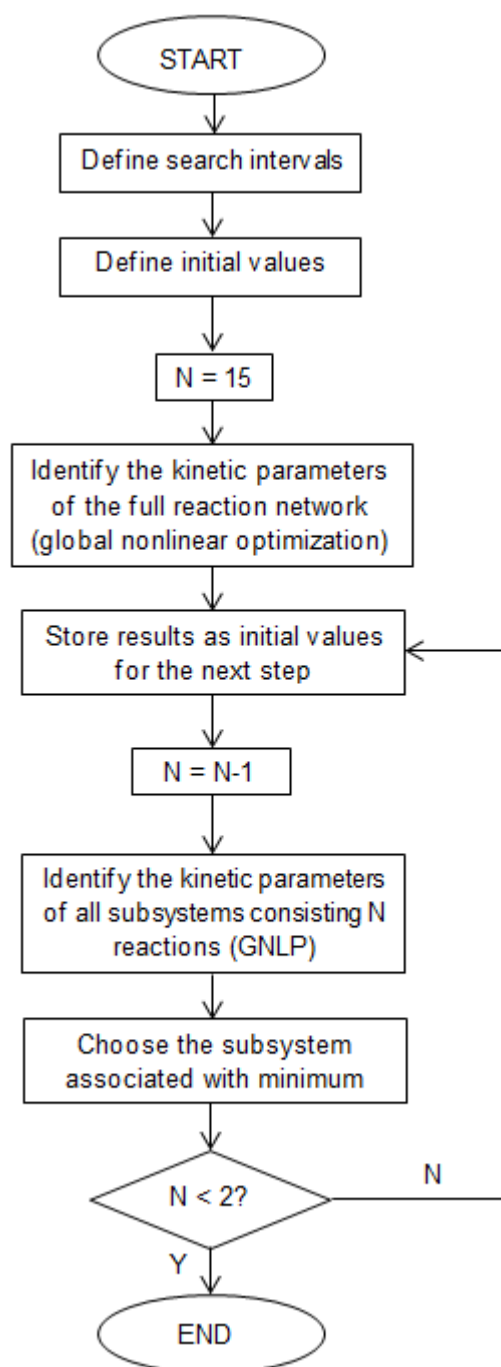


Figure 3. Second reaction network identification strategy.

The second strategy (top-down elimination) consists of the following steps (shown in Figure 3):

- (1) Identification of the kinetic parameters for the full reaction network.
- (2) Identification of the kinetic parameters for each subsystems consisting of one less reaction.
- (3) The elimination of that one reaction where the identified subsystem leads to the best results (i.e., minimum objective function value).
- (4) Further reactions were eliminated stepwise by repeating steps 2 and 3. The parameter values identified in the previous step were used as initial values in the next step.

This strategy is significantly more computation-intensive, as more than 100 nonlinear edge search problems must be solved in total. However, it has better convergence when applying the results of an n -reaction subsystem as the initial values of the next in a given step.

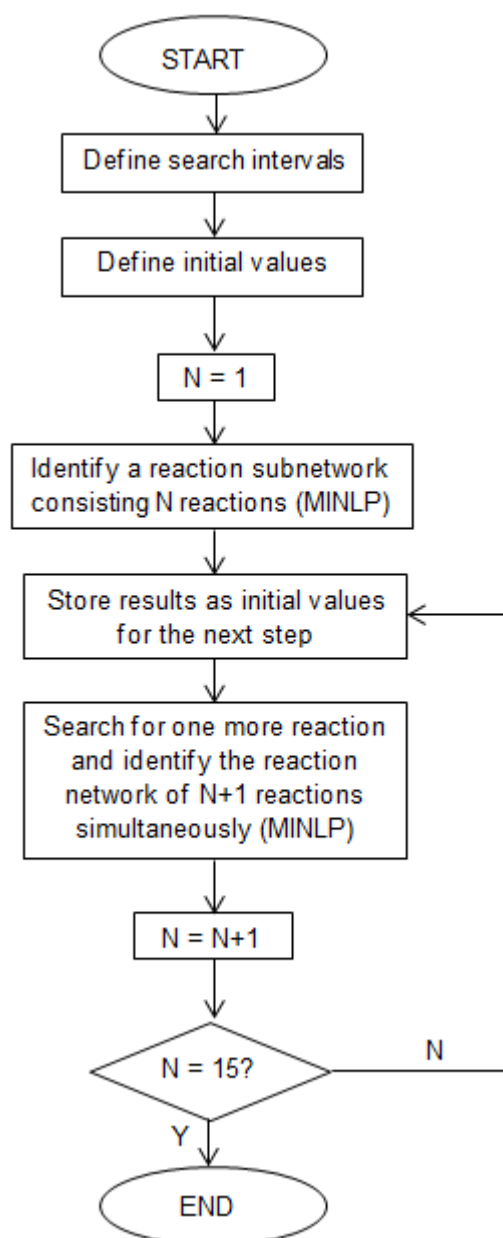


Figure 4. Third reaction network identification strategy.

Finally, a third strategy (shown in Figure 4) was defined that covers the conventional MINLP approach with a stepwise extension of the search space:

- (1) For the first step, there are 15 binary existence variables and 30 kinetic parameters. The reaction subnetwork to be identified consists of one reaction (e.g., the sum of existence variables is 1). The program searches for the reaction where the value of the objective function is at minimum and identifies the parameters of that reaction while keeping the values of all other kinetic parameters at zero at the same time.
- (2) For each subsequent step, the number of existence variables is decreased by one as the program searches for an additional reaction to expand the reaction network to minimize the value of the objective function further. The kinetic parameters of that additional reaction, as well as those of the previously added reactions, are identified.

The vector of search variables, $\underline{x}$, from eq 11 has a general form that is expressed in eqs 12 and 13.

$$\underline{x}^n = \left[\underline{exs}^T \quad \underline{k}_0^{nT} \quad \underline{E}_a^{nT} \right]^T \quad (12)$$

$$\underline{x}^{act} = \underline{x}^n \circ (\underline{u}_b - \underline{l}_b) + \underline{l}_b \quad (13)$$

$\underline{x}^n$ is a column vector that consists of the normalized values of existence variables, pre-exponential factors, and activation energies searched for in a given step of a given strategy (hence, the length of $\underline{exs}$ is zero for the first two strategies). All variables were normalized between 0 and 1 for faster and better convergence. The values of upper and lower bounds (summarized in Table 1) were determined from a set of preliminary studies regarding the complete reaction network as no a priori data are available for the lumped reactions.

Table 1. Lower and Upper Bounds of Kinetic Parameters of the Specified Reactions Used in Every Identification Strategy

Reaction	$k_0 [\text{m}^3 \text{h}^{-1} \text{m}_{\text{cat}}^{-3}]$		$E_a [\text{J mol}^{-1}]$	
	l_b	u_b	l_b	u_b
VGO $\rightarrow$ D	$3.6 \cdot 10^9$	$3.6 \cdot 10^{10}$	0	$4 \cdot 10^5$
VGO $\rightarrow$ K	$3.6 \cdot 10^{16}$	$3.6 \cdot 10^{17}$		
VGO $\rightarrow$ HN	360	3600		
VGO $\rightarrow$ LN	$3.6 \cdot 10^{13}$	$3.6 \cdot 10^{14}$		
VGO $\rightarrow$ G	360	3600		
D $\rightarrow$ K	$3.6 \cdot 10^{15}$	$3.6 \cdot 10^{16}$		
D $\rightarrow$ HN	36	360		
D $\rightarrow$ LN	3.6	36		
D $\rightarrow$ G	3.6	36		
K $\rightarrow$ HN	$3.6 \cdot 10^9$	$3.6 \cdot 10^{10}$		
K $\rightarrow$ LN	360	3600		
K $\rightarrow$ G	360	3600		
HN $\rightarrow$ LN	$3.6 \cdot 10^8$	$3.6 \cdot 10^9$		
HN $\rightarrow$ G	0.36	3.6		
LN $\rightarrow$ G	$3.6 \cdot 10^7$	$3.6 \cdot 10^8$		

Results and Discussion

The values of the objective function, which can be interpreted as the sum of normalized error between measured and calculated yields for all six lumps, are shown in Figure 5. The objective function values decrease exponentially, depending on the number of reactions considered in the model, depicted by a purple dashed line that is a result of the regression of an exponential decay function to the averages of the three objective function values against each set of reactions present. The values from each of the strategies decrease monotonously, which that suggests that the results are mostly free from numerical errors. The results from the sensitivity-based elimination strategy are the closest to the average, likely because of its relatively simple structure. For the second case, small objective function values were obtained, even for a smaller number of reactions than in case of the other two strategies as the search space in this case included a higher number of reaction network

subsystems, leading to more accurate results (at the cost of being more resource intensive, however). The results from the MINLP-based stepwise extension are the least accurate; for example, there is a plateau between five and seven reactions where the objective function value remains the same, which leads to the conclusion that, while it seems to be a somewhat evident choice, this is the least suitable for model reduction in our case.

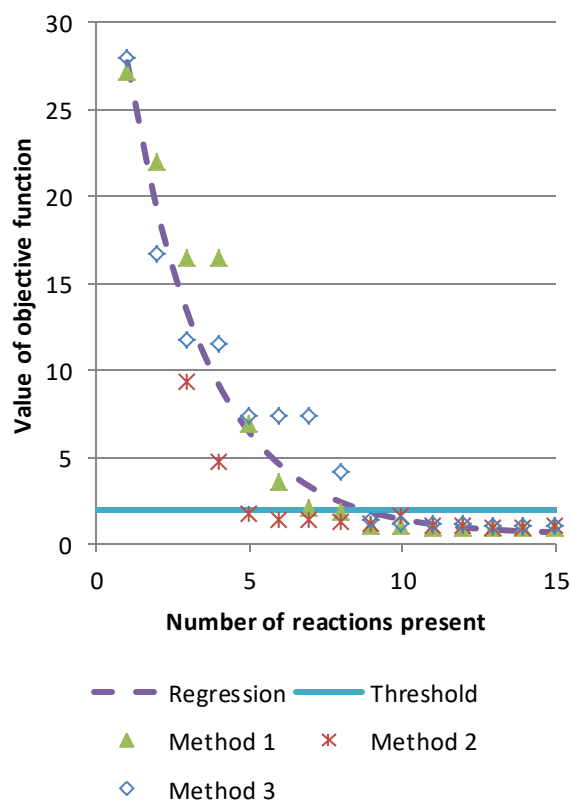


Figure 5. Objective function values for reaction subnetworks obtained by different model reduction methods.

There can be a threshold assigned for the objective function value above that the identified system poorly agrees with the measured data (or even some of the lumps do not appear in the model results). The value of this limit is ~ 2.0 (marked with a solid blue line in Figure 5), two times higher than the lowest obtained value of the objective function. This threshold value is an inherent part of the problem specified and cannot be considered as a general result; nevertheless, it can be used to evaluate the results. Hence, it can be assumed that a minimum of five reactions are necessary to describe the concentration changes of each defined lump in the system. It is apparent from the shape of the exponential function in Figure 5 that nine reactions are sufficient to minimize the objective function value. If there are more reactions present in the identified subsystem, the difference between the measured and calculated data series remains essentially the same. The square of Pearson correlation coefficients between measured and calculated data for each lump and the overall data series are listed in Table 2. The results are compared to the original work by Sadighi et al.³⁰ and show a considerable improvement, even for a five-reaction subsystem, probably due to the use of more-complex optimization algorithms. With the correlation factors for individual components taken into consideration, the results for the nine reaction subsystem obtained from the third strategy can be assumed to be the best, predicting the yields of all lumps except Diesel (D)

adequately, and the latter may well indicate some underlying errors in the measured data, although there is no information available regarding this issue.

Table 2. Pearson Correlations between Measured and Calculated Data

	Original work	5 reactions	9 reactions		
		Method 2	Method 1	Method 2	Method 3
VGO	0.852	0.906	0.912	0.912	0.909
Diesel	0.018	0.409	0.346	0.318	0.401
Kerosene	0.731	0.920	0.921	0.913	0.916
Heavy Naphtha	0.814	0.893	0.890	0.892	0.897
Light Naphtha	0.652	0.937	0.945	0.947	0.827
Gas	0.816	0.753	0.796	0.789	0.841
Overall	0.589	0.928	0.932	0.931	0.931

Consequently, the nine reaction subsystems identified with each strategy have been investigated more thoroughly. The reduced reaction networks for all strategies are shown in Figure 6. It can be seen that there are only five of the nine reactions ($VGO \rightarrow D$, $VGO \rightarrow LN$, $VGO \rightarrow G$, $D \rightarrow K$, and $K \rightarrow HN$, highlighted with red directed lines) are present in all three subsystems; in other words, these reaction networks differ significantly from each other, even though the resulting pseudocomponent yields are actually the same. Some of the differences can be explained with underlying consecutive reactions, i.e., a $VGO \rightarrow K$ lumped reaction can be substituted with a $VGO \rightarrow D \rightarrow K$ consecutive pathway by choosing the values of the rate coefficients appropriately, whereas, in the case of other reactions (i.e., $LN \rightarrow G$), these substitutions cannot be carried out.

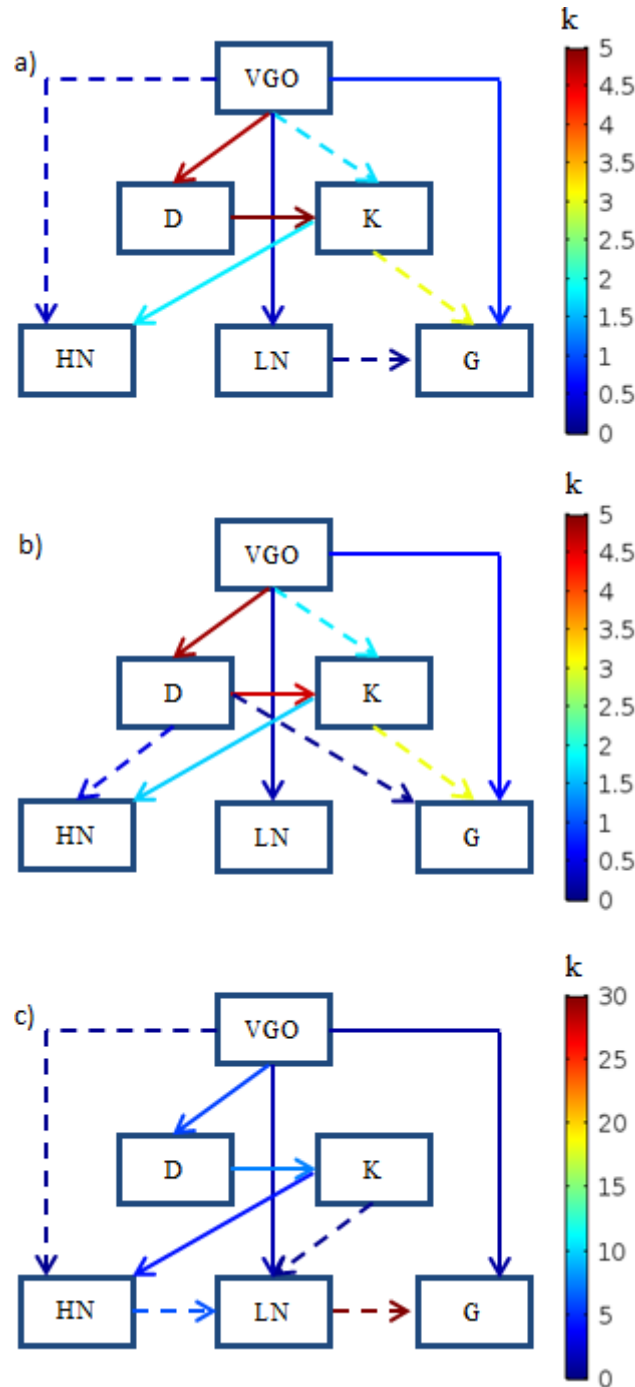


Figure 6. Reaction networks consisting 6 lumps and 9 reactions, obtained by (a) Method 1, (b) Method 2, and (c) Method 3. Reactions that are not present in all three networks are marked with dashed lines.

The existence of different reaction networks leading to almost the same results might explain the numerical uncertainties encountered while solving the optimization problems; however, more importantly, this raises some doubt about whether there exists a proper solution in the case of subnetworks consisting of nine (or more) reactions such that the objective function value has a global minimum.

In order to address this problem systematically and quantitatively, the observability of the extended state-space model was investigated. If the concentrations of six lumps are measured and 21 states are observable, all 15 reaction rate coefficients can be determined from the 6 states measured.

However, if less than 21 states are observable, not all of the rate coefficients can be identified adequately. That has the practical consequence that two or more sets of parameter values in the model lead to the same output values. In other words, more kinetic parameters were identified in the given system than theoretically possible; therefore, that solution cannot be considered to be valid.

The number of observable states was determined by calculating the rank of the observability matrix (eq 14), where A is the system matrix, C is the output matrix, and n is the number of states in the system:

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \quad (14)$$

The state-space model was linearized around the reactor outlet; the results for the component VGO at 380 °C for each LHSV values are shown in Figure 7. For the sake of better understanding, both the original and the linearized state-space model were solved around the reactor outlet; naturally, the values for $\ell > 1$ bear no physical meaning. The figures indicate good agreement between the results from solving the actual and the linearized state-space models; therefore, the system matrix of the latter can be used for the calculation of the observability matrix.

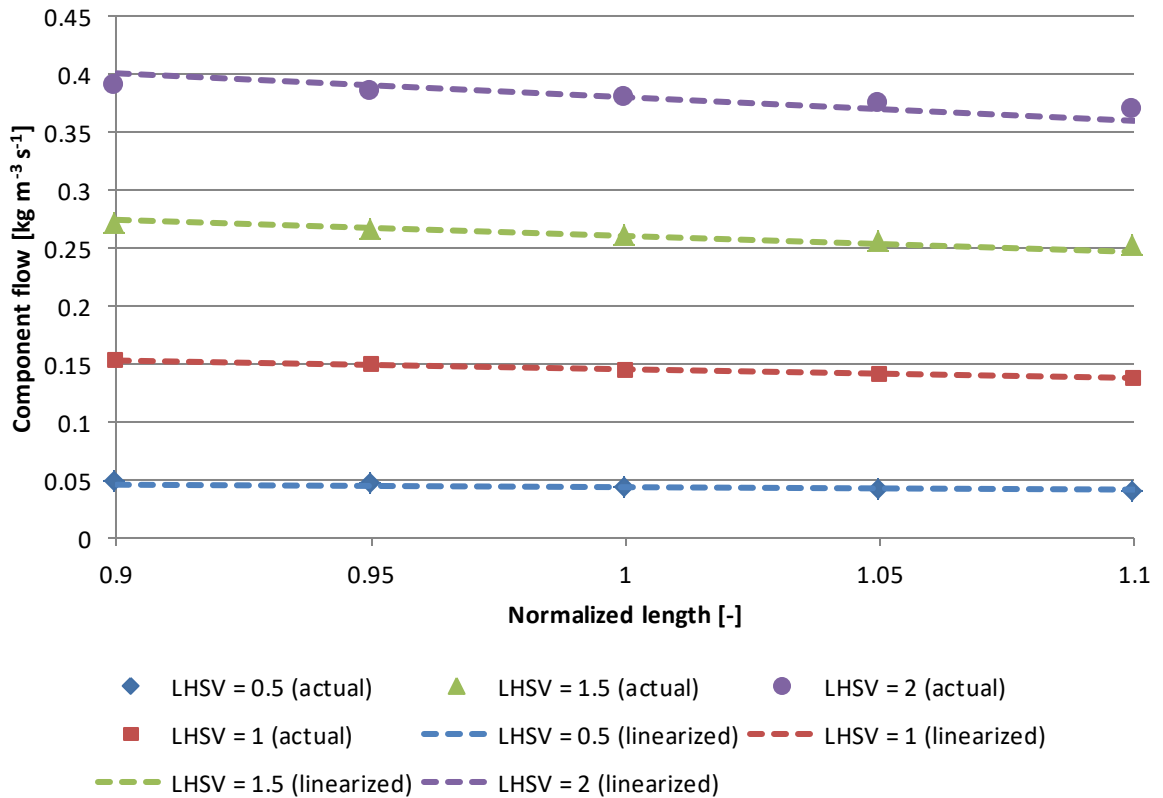


Figure 7. Results of state-space model linearization at the outlet of the reactor for VGO at 380 °C.

In Figure 8, the number of observable states is depicted as a function of states for the three strategies and for all identified reaction subnetworks. The discrete values are connected with solid lines for better readability. The maximum number of observable states equals to the number of actual states in the model (marked with the blue line). It can be seen that the maximum number of

observable states are only 11, meaning only 5 reaction rate coefficients can be identified from the available measurement data, because of the 6 components. It is also clear that only 1 from the 3 identified 5-reaction subsystems has 11 observable states; two others have only 10. The reason behind this is that the remaining 10 reactions were eliminated in such a way that the remaining five are not independent from each other. Such elimination could occur because the strategies do not consider any criteria for observability. The lesser number of observable states might as well explain the high difference between the objective function values that resulted from different strategies shown in Figure 5. In the case of the second top-down elimination strategy, the number of observable states for the reaction subsystems containing 1 and 2 reactions are actually less than the number of measurements; this is because only one lump has other-than-zero concentration at the reactor outlet, which counts as an actual measurement when calculating the rank of the observability matrix.

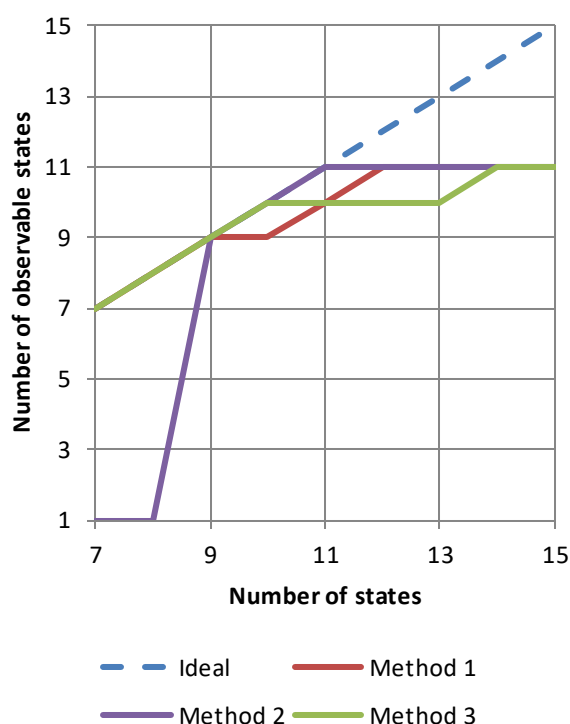


Figure 8. Observability of reaction networks consisting 1–9 reactions, obtained using different methods.

If the subsystem containing 5 reactions identified in the second strategy is observable, that will mean the identified kinetic data represent an exact and unique solution that is the global minimum for 5 reactions, given that the performance of the optimization algorithm was good enough to find it. In order to prove this, all 5-reaction combinations from a set of 15 reactions, a total of 3003 subsystems, must be identified. The task is time-consuming but can be performed within a reasonable time horizon. From the results, the subsets where the objective function value was <10 were taken and plotted on a histogram, using a bin size of 0.5 (Figure 9). The results indicates that the top-down elimination strategy is capable of finding the global minimum, as only 2 from 3003 5-reaction combinations resulted in objective function values of <2 (1.755 and 1.811, to be more accurate), from which the smaller value coincides with the result given by top-down elimination (1.753), and moreover, the identified reaction subset here is exactly the same as the one identified

by the second strategy. Table 3 lists the values of reaction rate coefficients at 410 °C that agree to at least two significant digits; thus, the two solutions are essentially equal.

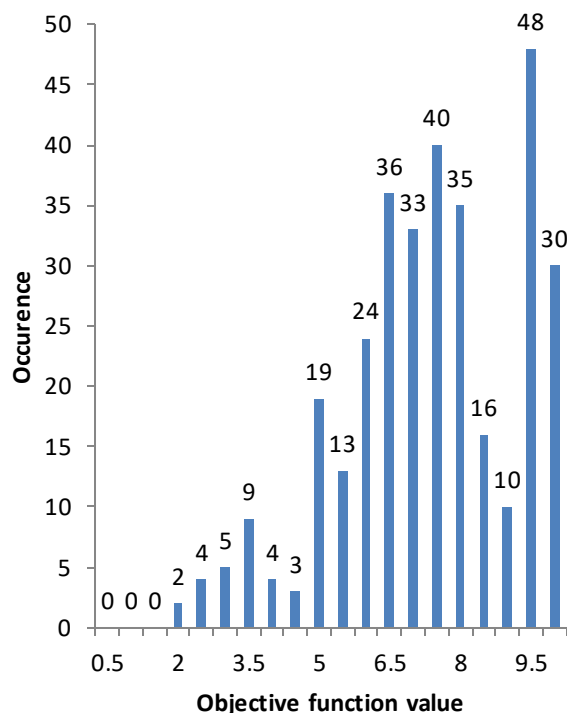


Figure 9. Occurrence of objective function values for various 5-reaction networks.

Table 3. Reaction Rate Coefficients at 410 °C for 5-Reaction Networks

Reactions	VGO → D	VGO → LN	D → K	K → HN	K → G
Best result from the identification of all possible subsystems	6.23	0.204	10.42	2.27	4.81
Result from Method 2	6.29	0.206	10.34	2.27	4.83

Another method to prove that the optimal values have been calculated is to perform a sensitivity analysis on the determined parameters. This was performed by means of $\pm 25\%$ perturbations of the pre-exponential factors and $\pm 5\%$ perturbations of activation energies, on one parameter at a time, following the method described by Alcázar and Ancheyta.³⁶ In each step, the objective function (eq 11) was evaluated; the resulting curves are shown in Figure 10. This analysis again indicates that the estimated parameter values are at the optimum, since the minimum values of all 10 curves are at 0% and any perturbation results in higher objective function values.

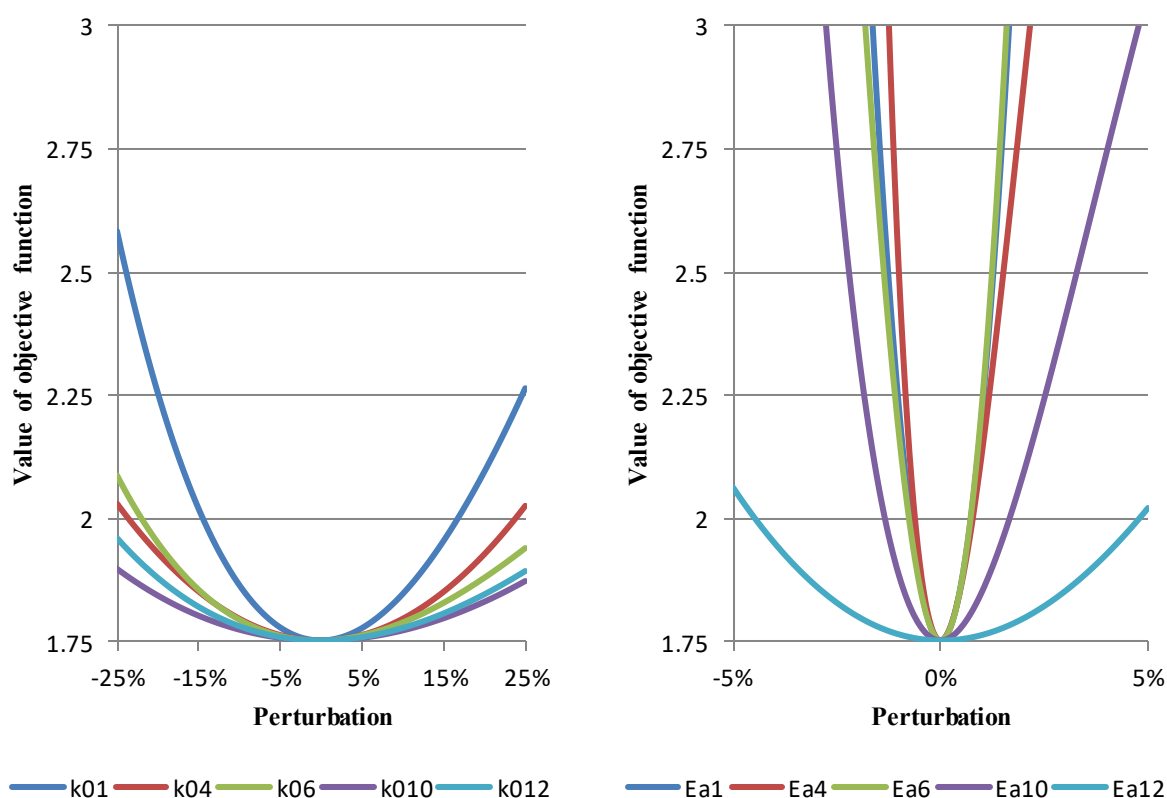


Figure 10. Sensitivity analysis of calculated parameters of the 5-reaction model from the second strategy.

The final solution from the model reduction and identification is depicted in Figure 11. The model presented here somewhat disagrees with the well-known behavior of the hydrocracking reaction; in other words, the reactions do not follow a consecutive pattern (for example, Light Naphtha is produced from VGO instead of Heavy Naphtha). The main reason behind that observation is that the observability criterion strongly limits the number of reactions present in the final model; one pseudo-component can only be produced via one reaction pathway. In other words, there is only one possible 5-reaction subsystem where the consecutive reaction pathway occurs. The objective function value associated with the consecutive pathway is 9.2, which is more than five times higher than the final solution suggested in Figure 11 and Table 3. Hence, there are two solutions: the first includes more reactions, so the consecutive pathway is present and also the value of the objective function is acceptable, and the second excludes some reactions from the consecutive pathway and includes the same number of nonconsecutive reactions, so the resulting model gives a good description of the concentration changes of the pseudo-components while the system remains also observable. While the former approach describes the behavior of hydrocracking better, the latter ensures the certainty regarding the values of the identified kinetic parameters, which is more favorable, in terms of process development and intensification. Moreover, the algorithm presented here utilizes no a priori information about the reaction chemistry, and the purpose of these strategies is to identify a reaction system and its kinetic parameters suitable of describing the concentration changes of pseudo-components based only on measured values of concentration; hence, the only constraints present regarding the structure of the identified network are the reactions that are included in the full network, i.e., the superstructure.

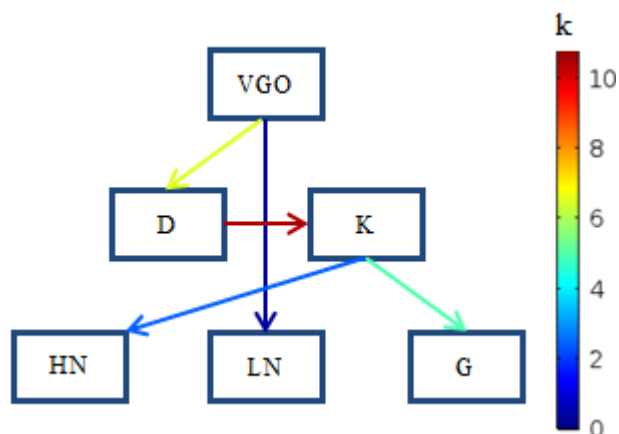


Figure 11. Reaction network consisting of 5 reactions, obtained by Method 2.

The measured and calculated product compositions as a function of LHSV at 410 °C for 15 and 5 reactions present are shown in Figures 12 (VGO, Diesel, Kerosene) and 13 (Heavy Naphtha, Light Naphtha and Gas). The shapes of the curves are mostly the same for the 5-reaction model and the original model consisting of 15 reactions. In some cases, the resulting curve from the former fits worse to the measured data (e.g., Gas); whereas, in other cases, the curve fitting is better (VGO). In some cases, somewhat poor correlation between measured and calculated data can be observed for both models (e.g., Diesel). However, the results of the sensitivity analysis on the kinetic parameters implies that this issue is related, at least in part, to measured data and could be overcome by the application of different lumping strategies. This may well be part of a further investigation.

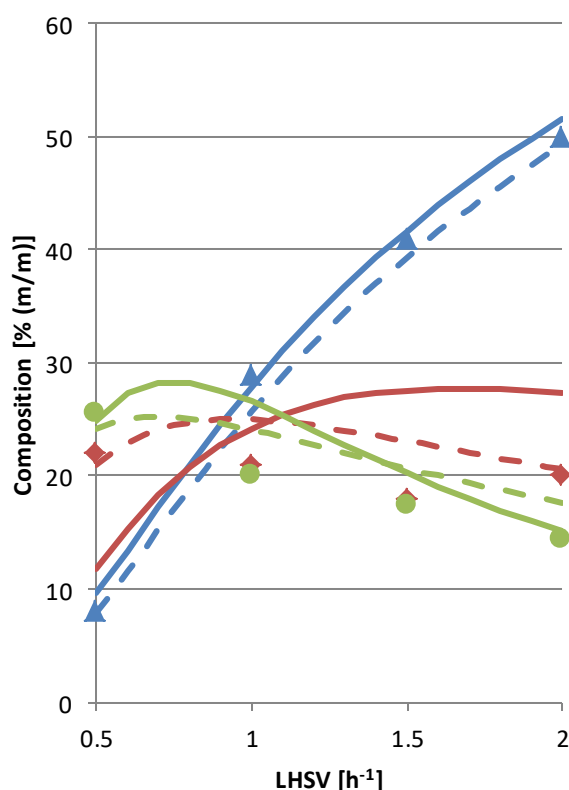


Figure 12. Concentration of VGO (blue), Diesel (red), and Kerosene (green) at the reactor outlet for different LHSV values at 410 °C. Data points represent measured data, dashed lines indicate the 15-reaction system results, and solid lines represent the 5-reaction system results.

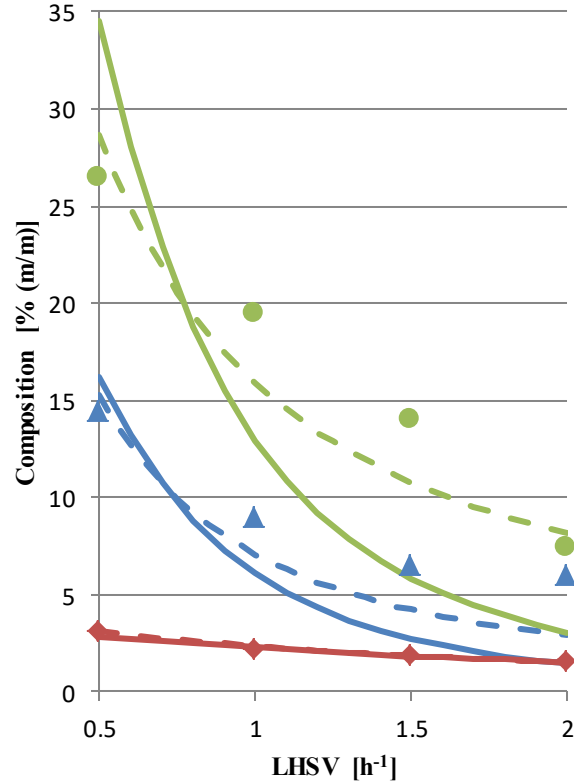


Figure 13. Concentration of Heavy Naphtha (blue), Light Naphtha (red), and Gas (green) at the reactor outlet for different LHSV values at 410 °C. Data points represent measured data, dashed lines indicate the 15-reaction system results, and solid lines represent the 5-reaction system results.

Conclusions and Future Work

In this work, multiple optimization studies for the heterocatalytic oxidation of HCl into Cl₂ (the HCl conversion process) were carried out for a fixed-bed tubular reactor. Three different design methods were considered, the application of graded catalyst beds, multiple cooling zones, and finally, the coupling of the previous two into a generalized reactor zone. The number of zones (N) varied from 1 to 6 in our investigations. The results obtained by using the initial objective function for HCl conversion maximization have shown that while reaching the equilibrium conversion is possible, this method does not properly address the temperature changes occurring in the reactor. Hence three other objective functions have been implemented, and the results were evaluated using the axial temperature gradient of the reactor temperature as a main indicator. With this method, the temperature deviation can be reduced to a reasonably low level using four catalyst beds with different activities. While splitting of the reactor shell is not necessary, a more flexible system can be achieved with it as shell side temperature may be altered in continuous operation of the reactor. Future work may consider the dynamic behavior of the reactor, for example, the effect of depleting activity of the catalyst that could be different in each zone and, along with that, control strategies to maintain the determined temperature profile.

Supporting Information

The Supporting Information is available free of charge on the [ACS Publications](https://pubs.acs.org/doi/10.1021/acs.energyfuels) website at DOI: [10.1021/acs.energyfuels](https://pubs.acs.org/doi/10.1021/acs.energyfuels).

Kinetic data for all reaction networks identified in each strategy ([PDF](#))

Author Information

Corresponding Author

*E-mail: tillz@fmt.uni-pannon.hu.

ORCID

Zoltán Till: [0000-0003-0691-6768](https://orcid.org/0000-0003-0691-6768)

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

Acknowledgements

We acknowledge the financial support of Széchenyi 2020, under the GINOP-2.3.2-15-2016-00053. The proposed optimization methodology can be applied in the development of catalytic production technology of high hydrogen content (in molecular structure) fuels.

Nomenclature

A	system matrix
B	input matrix
C	output matrix
c	mass concentration [kg m^{-3}]
E_a	activation energy [J mol^{-1}]
exs	existence variables
I	identity matrix
k	reaction rate coefficient [$\text{m}^3 \text{h}^{-1} \text{m}_{\text{cat}}^{-3}$]
k_0	pre-exponential factor [$\text{m}^3 \text{h}^{-1} \text{m}_{\text{cat}}^{-3}$]
ℓ	dimensionless length [-]
L	Kirchhoff matrix of the reaction system
l_b	lower bounds of optimization variables
$\mathcal{O}$	observability matrix
R	gas constant [$\text{J mol}^{-1} \text{K}^{-1}$]
T	temperature [K]
v	actual space velocity [h^{-1}]
u	state-space model input
u_b	upper bounds of optimization variables
x	model state (in state-space model)
x^{act}	actual optimization variables
x^n	normalized optimization variables

y	state-space model output
Y	pseudocomponent yield [-]

Abbreviations

D	Diesel lump
G	Gas lump
HN	Heavy Naphtha lump
K	Kerosene lump
LHSV	Liquid Hourly Space Velocity [h^{-1}]
LN	Light Naphtha lump
OFV	Objective function value
VGO	Vacuum Gas Oil lump

Subscripts

b	bounds
cat	catalyst
comp	pseudocomponent, lump
in	inlet

Superscripts

act	actual
n	normalized
T	transpose

Greek letters

ϵ'	catalyst volume fraction [$\text{m}_{\text{cat}}^3 \text{m}_{\text{reactor}}^{-3}$]
η	catalyst effectiveness factor [-]
ρ	density [kg m^{-3}]

References

- (1) Sahu, R.; Song, B. J.; Im, J. S.; Jeon, Y.-P.; Lee, C. W. A review of recent advances in catalytic hydrocracking of heavy residues. *J. Ind. Eng. Chem. (Seoul, Repub. Korea)* **2015**, 27, 12-24.
- (2) Sági, D.; Holló, A.; Varga, G.; Hancsók, J. Fuel Purpose Hydrotreating of Free Fatty Acid By-products and Heavy Straight Run Gas Oil. *Chem. Eng. Trans.* **2016**, 52, 883-888.
- (3) Rana, B. S.; Kumar, R.; Tiwari, R.; Kumar, R.; Joshi, K.; Garg, M. O.; Sinha, A. K. Transportation fuels from co-processing of waste vegetable oil and gas oil mixtures. *Biomass Bioenergy* **2013**, 56, 43-52.
- (4) Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC (Text with EEA relevance). *Off. J. Eur. Communities: Legis.*, **2009**, 140, 16-62.
- (5) de Oliveira, L. P.; Hudebine, D.; Guillaume, D.; Verstraete, J. J. A Review of Kinetic Modeling Methodologies for Complex Processes. *Oil Gas Sci. Technol.* **2016**, 71 (3), 45.

- (6) Sánchez, S.; Rodríguez, M. A.; Ancheyta, J. Kinetic model for moderate hydrocracking of heavy oils. *Ind. Eng. Chem. Res.* **2005**, *44* (25), 9409-9413.
- (7) Sánchez, S.; Ancheyta, J. Effect of pressure on the kinetics of moderate hydrocracking of Maya crude oil. *Energy Fuels* **2007**, *21* (2), 653-661.
- (8) Nguyen, T. T. H.; Teratani, S.; Tanaka, R.; Endo, A.; Hirao, M. Development of a Structure-Based Lumping Kinetic Model for Light Gas Oil Hydrosulfurization. *Energy Fuels* **2017**, *31* (5), 5673-5681.
- (9) Ancheyta, J.; Sánchez, S.; Rodríguez, M. A. Kinetic modeling of hydrocracking of heavy oil fractions: A review. *Catal. Today* **2005**, *109* (1), 76-92.
- (10) Martínez, J.; Sánchez, J. L.; Ancheyta, J.; Ruiz, R. S. A review of process aspects and modeling of ebullated bed reactors for hydrocracking of heavy oils. *Catal. Rev.* **2010**, *52* (1), 60-105.
- (11) Jiang, H.; Huang, S. Eight-Lump Reaction Kinetic Model for the Maximizing Isoparaffin Process for Cleaning Gasoline and Enhancing Propylene Yield. *Energy Fuels* **2016**, *30* (12), 10770-10776.
- (12) Wang, H.; Dai, F.; Yang, Y.; Li, Z.; Li, C.; Zhang, S. Catalyst Grading Optimization and Kinetic Simulation of the Shale Oil Hydrotreating Process. *Energy Fuels* **2017**, *31* (4), 4563-4360.
- (13) Cheng, Z. M.; Huang, Z. B.; Yang, T.; Liu, J. K.; Ge, H. L.; Jiang, L. J.; Fang, X. C. Modeling on scale-up of an ebullated-bed reactor for the hydroprocessing of vacuum residuum. *Catal. Today* **2014**, *220*, 228-236.
- (14) Calderon, C. J.; Ancheyta, J. Modeling of Slurry-Phase Reactors for Hydrocracking of Heavy Oils. *Energy Fuels* **2016**, *30* (4), 2525-2543.
- (15) Forghani, A. A.; Jafarian, M.; Pendleton, P.; Lewis, D. M. Mathematical modelling of a hydrocracking reactor for triglyceride conversion to biofuel: model establishment and validation. *Int. J. Energy Res.* **2014**, *38* (12), 1624-1634.
- (16) Ledesma, E. B.; Mullery, A. A.; Vu, J. V.; Hoang, J. N. Lumped Kinetics for Biomass Tar Cracking Using 4-Propylguaiacol as a Model Compound. *Ind. Eng. Chem. Res.* **2015**, *54* (21), 5613-5623.
- (17) Ong, Y. K.; Bhatia, S. The current status and perspectives of biofuel production via catalytic cracking of edible and non-edible oils. *Energy* **2010**, *35* (1), 111-119.
- (18) Han, Y.; Stankovikj, F.; Garcia-Perez, M. Co-hydrotreatment of tire pyrolysis oil and vegetable oil for the production of transportation fuels. *Fuel Process. Technol.* **2017**, *159*, 328-339.
- (19) Periyasami, B. Reaction pathway analysis in thermal cracking of waste cooking oil to hydrocarbons based on monomolecular lumped kinetics. *Fuel* **2015**, *158*, 479-487.
- (20) Song, H. S.; Ramkrishna, D. Cybernetic models based on lumped elementary modes accurately predict strain-specific metabolic function. *Biotechnol. Bioeng.* **2011**, *108* (1), 127-140.
- (21) Rodríguez-Fernández, D. E.; Rodríguez-León, J. A.; De Carvalho, J. C.; Sturm, W.; Soccol, C. R. The behavior of kinetic parameters in production of pectinase and xylanase by solid-state fermentation. *Bioresour. Technol.* **2011**, *102* (22), 10657-10662.
- (22) Auger, F.; Hilairet, M.; Guerrero, J. M.; Monmasson, E.; Orlowska-Kowalska, T.; Katsura, S. Industrial applications of the Kalman filter: A review. *IEEE Trans. Ind. Electron.* **2013**, *60* (12), 5458-5471.

- (23) Kalman, R. E. Mathematical description of linear dynamical systems. *J. Soc. Ind. Appl. Math., Ser. A* **1963**, 1 (2), 152-192.
- (24) Liu, Y.-Y.; Slotine, J.-J.; Barabási, A.-L. Observability of complex systems. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, 110 (7), 2460-2465.
- (25) Wentzell, P. D.; Karayannis, M. I.; Crouch, S. R. Simultaneous kinetic determinations with the Kalman filter. *Anal. Chim. Acta* **1989**, 224, 263-274.
- (26) Gui, M.; Rutan, S. C. Determination of initial concentration of an analyte by kinetic detection of the intermediate product in consecutive first-order reactions using an extended Kalman filter. *Anal. Chem.* **1994**, 66 (9), 1513-1519.
- (27) Laxminarasimhan, C. S.; Verma, R. P.; Ramachandran, P. A. Continuous lumping model for simulation of hydrocracking. *AIChE J.* **1996**, 42 (9), 2645-2653.
- (28) Sildir, H.; Arkun, Y.; Cakal, B.; Gokce, D.; Kuzu, E. A dynamic non-isothermal model for a hydrocracking reactor: model development by the method of continuous lumping and application to an industrial unit. *J. Process Control* **2012**, 22 (10), 1956-1965.
- (29) Sadighi, S.; Ahmad, A.; Irandoukht, A. Modeling a pilot fixed-bed hydrocracking reactor via a kinetic base and neuro-fuzzy method. *J. Chem. Eng. Jpn.* **2010**, 43 (2), 174-185.
- (30) Sadighi, S.; Zahedi, G. R. Comparison of Kinetic-based and Artificial Neural Network Modeling Methods for a Pilot Scale Vacuum Gas Oil Hydrocracking Reactor. *Bull. Chem. React. Eng. Catal.* **2013**, 8 (2), 125-136.
- (31) Szederkényi, G. Computing sparse and dense realizations of reaction kinetic systems. *J. Math. Chem.* **2010**, 47 (2), 551-568.
- (32) Shampine, L. F.; Reichelt, M. W. The MATLAB ODE Suite. *SIAM J. Sci. Comput.* **1997**, 18 (1), 1-22.
- (33) Martins, J. R. R. A.; Sturdza, P.; Alonso, J. J. The complex-step derivative approximation. *ACM T. Math. Software* **2003**, 29 (3), 245-262.
- (34) Le Digabel, S. Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm. *ACM T. Math. Software* **2011**, 37 (4), 44.
- (35) Currie, J.; Wilson, D. I. OPTI: Lowering the Barrier Between Open Source Optimizers and the Industrial MATLAB User. In *Foundations of Computer-Aided Process Operations*, Savannah, United States, January 8–13, 2012; Sahinidis, N., Pinto, J., Eds.; Elsevier: Toronto, 2012.
- (36) Alcázar, L. A.; Ancheyta, J. Sensitivity analysis based methodology to estimate the best set of parameters for heterogeneous kinetic models. *Chem. Eng. J. (Amsterdam, Neth.)* **2007**, 128 (2), 85-93.