



Application of Monte Carlo techniques to LCO gas oil hydrotreating: Molecular reconstruction and kinetic modelling

Maria Lopez Abelairas, Luis Pereira de Oliveira, Jan Verstraete

► To cite this version:

Maria Lopez Abelairas, Luis Pereira de Oliveira, Jan Verstraete. Application of Monte Carlo techniques to LCO gas oil hydrotreating: Molecular reconstruction and kinetic modelling. *Catalysis Today*, 2016, 271, pp.188-198. 10.1016/j.cattod.2016.02.041 . hal-01338408

HAL Id: hal-01338408

<https://hal.science/hal-01338408>

Submitted on 28 Jul 2016

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.

Application of Monte Carlo techniques to LCO gas oil hydrotreating: molecular reconstruction and kinetic modelling.

Maria Lopez Abelairas, Luis Pereira de Oliveira, Jan J. Verstraete*

***Jan J. Verstraete: corresponding author**

E-mail address: jan.verstraete@ifpen.fr

Tel: +33 4 78 02 20 53, ext. 16014; Fax: +33 4 78 02 20 08

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

Abstract

The increasing pressure to satisfy the environmental criteria concerning oil products leads to the need for developing accurate models to predict the performances of the refining processes. In the current study, a stochastic two-step procedure using Monte Carlo techniques is applied to and validated on the hydrotreating of Light Cycle Oil (LCO) gas oils. In the first step, a mixture of molecules representative of the LCO gas oils is generated using a molecular reconstruction method termed SR-REM. Subsequently, the Stochastic Simulation Algorithm (SSA) is applied to simulate the evolution of the mixture composition during hydrotreating. The results show that an accurate representation of eleven different LCO gas oils was obtained by the application of the molecular reconstruction method. The hydrotreating simulations of three LCO gas oils at different operating conditions showed a good agreement with the experimental data obtained at laboratory scale. The current stochastic procedure has demonstrated to be a valid tool for the reconstruction of the composition of LCO gas oils and the simulation of the hydrotreating process.

Keywords

Composition modelling; molecular reconstruction; kinetic modelling; kinetic Monte Carlo; stochastic simulation algorithm; LCO gas oil hydrotreating.

1. Introduction

The pressure on the quality and the maximum impurities content of refinery products has been increased in the last decades due to the environmental concerns associated to their use. The suitable performance of the processes aimed to remove these impurities is essential to satisfy the specifications fixed for oil products. Among these processes, hydrotreating (HDT) is one

of the most mature technologies. The use of hydrogen at high temperature over a catalyst bed allows to remove sulphur, nitrogen and other undesirable elements from the petroleum distillates such as naphthas or gas oils.

The correctness in the prediction of HDT process performance directly depends on the reliability of the used kinetic model. A lumping strategy, in which molecular components are grouped according to their global properties, has usually been applied in the kinetics models of complex hydrocarbons¹⁻³. However, this approach is no longer manageable due to the elevated number of generated groups and reaction pathways. The limited applicability of lumped models has resulted in the development of molecular-based kinetic models^{2,4-9}. In this latter approach, reactant species and products are described by a selection of molecules or small groups of them which react following networks formed by reactions at the molecular level and distinguished by a global kinetic constant. The elemental steps are regrouped within the global reactions scheme, thereby reducing the number of reactive species and the size of reaction network. Following this approach, Lopez Garcia *et al.*⁸ have simulated the hydrotreating of LCO gas oils and, more recently, Pereira *et al.*⁹⁻¹¹ have proposed a kinetic methodology for its application on conversion of petroleum cuts based on the Stochastic Simulation Algorithm (SSA) developed by Gillespie¹².

These kinetics models required a molecular description of the feedstock. However, the high complexity of the petroleum cuts hinders their molecular characterisation even by using the most advanced analytical techniques. To surpass this drawback, a molecular representation of the feedstock can be determined by the application of a molecular reconstruction algorithm. This kind of algorithms provides a set of molecules from petroleum analyses and chemical knowledge. Several approaches have been performed in this sense, from the first model proposed by Liguras and Allen¹³ to the stochastic reconstruction (SR) developed by Neurock *et al.*¹⁴ in which a Monte Carlo method is used to sample the objective function. Hudebine^{15,16} has characterised different petroleum cuts by the application of a two-steps algorithm based on SR algorithm and on reconstruction entropy maximisation (REM). The combination of both methods (SR-REM) overcomes the drawbacks of its separated application, such as noise of the objective function and computational effort. Posteriorly, Pereira *et al.*^{17,18} have modified Hudebine's method to introduce a genetic algorithm which performs the minimisation of the objective function.

The current study is focused on the validation of the SR-SEM and SSA algorithms developed by Pereira *et al.* on the hydrotreating of Light Cycle Oil (LCO) gas oils. A brief description of the stochastic methodology applied to both molecular reconstruction and reaction simulation has been developed. Then, the molecular reconstruction of eleven LCO gas oil has been carried out and the hydrotreating of three of them has been simulated at different operational conditions.

2. Methodology

The stochastic methodology applied in the current work consists of two steps. In the first one, a molecular reconstruction algorithm is used to model the feedstock by means of a set of molecules which mixture properties are similar to those of the feedstock. This algorithm comprises two coupled methods: stochastic reconstruction (SR) and reconstruction by entropy maximization (REM). The application of the SR method enables to obtain a set of molecules typical for a given type of oil fraction. Then, the molar fraction of molecules is adjusted using the REM method in order to improve the predicted mixture properties.

Finally, the generated set of molecules is introduced as input in the second step concerning the molecule-based kinetic modelling of the conversion process. A kinetic Monte Carlo (kMC) method is applied to model the chemical reaction system at a molecular level. This approach is an alternative to traditional deterministic kinetic methods in which a serial of differential equations (ODE) are involved to obtain a temporal evolution of the species in the system.

2.1. Molecular reconstruction

The SR method, from which a representative set of molecules is generated, uses probability distributions functions (PDF's) of structural attributes, such as number of rings or number of side chains, to characterise the oil fraction to be reconstructed^{14,19}. Feedstock analyses, which provide information about atomic abundance (elemental analysis) or structural classes (GC-MS), and the previous knowledge about petroleum cuts should be taken into account to select the proper structural attributes. The selected PDF's are sampled N times using a Monte Carlo method in order to generate an initial set of N molecules (Figure 1). To carry out the transition from PDF's to the set of molecules, the application of a building diagram and a serial of chemical rules is required to avoid the creation of impossible and unlikely molecules. The building diagram is a decision tree that defines the sequence in which PDF's should be sampled to proceed to the posterior assembly of the different attributes, which provides the

structure of the molecule. During the assembly process, the application of chemical rules discards the molecules which do not satisfy the thermodynamic or likelihood criteria.

The application of the Monte Carlo sampling procedure led to the creation of a mixture of N molecules, each one with a molar fraction equal to $1/N$. The average properties of this mixture are calculated and compared with the experimental data by means of an objective function. An optimisation method is used to minimise this objective function through the modification of PDF's parameters. The simulated annealing^{15,19-21} and the genetic algorithms²²⁻²⁴ are both usual methods applied to the optimisation of the objective function in the field of the stochastic reconstruction. However, Schnongs²⁴ has demonstrated that the genetic algorithms are more adapted to the stochastic reconstruction than the simulated annealing since the genetic algorithms are less sensitive to the initial value of the distribution parameters and it has also showed a lower oscillation of the value of the objective function.

Once the initial set of molecules is obtained, the second step based on the reconstruction by entropy maximisation (REM) is applied. During this step, the molar fraction of the molecules in the mixture is modified to obtain a mixture with properties closer to those experimentally determined¹⁵. To carry out the adjustment of the parameters, the REM method, based on Shannon's information theory²⁵, uses the maximisation of an information entropy criterion ($E(x_i)$) given by the Eq. (1):

$$E(x_i) = - \sum_{i=1}^N x_i \cdot \ln(x_i) \quad (1)$$

with

$$\sum_{i=1}^N x_i = 1 \quad (2)$$

where x_i corresponds to the molar fraction of molecule i and N is the number of molecules present in the initial mixture determined by applying the SR method. This criterion guarantees that, without constraints or available information, a given molecule cannot be preferred to others, so that the distribution of the predefined set of molecules remains uniform. The introduction of a matrix of constraints from analytical data, mixing rules and the initial set of molecules distorts this uniformity in order to match this information. These constraints are introduced in the entropy maximisation criterion by means of the Lagrange multiplier method. In the case of linear constraints were used, the problem is reduced to an optimisation of J parameters of a no-linear equation that can be performed by the conjugate gradient method.

Once the Lagrange parameters are determined, the molar fraction of the molecules is calculated in order to obtain the representative mixture of molecules.

2.2. Reaction modelling

The goal of the kMC method is to describe the chemical reaction system at the molecular level by following the transformations of a discrete population of molecules. Contrary to deterministic methods, the temporal evolution of the system is not described by a set of differential equations, but by a single probabilistic function termed the chemical master equation (CME)^{26,27}. This equation determines the different status of the system in a sequence of discrete-time steps. The evolution of the system is a result from chemical reactions, those have associated a probability function, which depends on the reactant molecule and the rate constant.

However, CME cannot usually be solved by analytical or numerical methods due to its high complexity. In this case, a Monte Carlo procedure developed by Gillespie¹² and termed Stochastic Simulated Algorithm (SSA) is used to solve CME. This method is based on an event-space approach, i.e., the evolution of the system is followed event by event, resulting in disparate time intervals between two consecutive events or, in this particular case, reactions. The scheme followed in the application of the kMC method has been illustrated in Figure 2.

The initial step of the procedure consists in the “molecular discretisation” of the set of molecules generated by applying the SR-REM method. During this step, the replication of the molecules in the mixture is performed according to their molar fraction and it is based on the same principle as the procedure developed by Hudebine and Verstraete²¹ for the creation of a representative set of molecules for gasoline fractions. A replication factor (F_{rep}), defined as the maximum number of molecules in the discrete mixture, is multiplied by the mole fraction of each molecule in order to calculate the number of replications. F_{rep} also defines the lowest mole fraction to be retained since the molecules with a mole fraction lower than $0.5/F_{rep}$ are discarded. The factor 0.5 arises from the rounding off to the closest integer.

In the next step of the kMC procedure, all potential reaction events should be identified and listed taking into account a set of reaction rules that generates the reactions by inspecting the structure of the molecules. Once all potential reaction events are identified, the normalised probability for each reaction event is calculated as the ratio of its stochastic rate (R_v) to the sum of the rates of all reaction events (M), as shown in Eq. (3):

$$P_v = \frac{R_v}{\sum_i^M R_i} \quad (3)$$

with

$$R_v = h_v \cdot c_v \quad (4)$$

where h_v is the number of distinguishable combinations of the reactant molecule and c_v is the stochastic rate parameter. For monomolecular reactions, the number of distinguishable combinations of the reactant molecules h_v equals 1, while for bimolecular reactions, it is equal to the number of available molecules of the second reactant. The stochastic rate parameter c_v is closely linked to the deterministic rate parameter (k_v) of each reaction. In the case of monomolecular reactions, k_v and c_v are equal, while for bimolecular reactions, c_v is equal to the ratio of k_v by reaction volume V .¹²

The effect of the reaction temperature on the stochastic rate is taking into account by means of the Arrhenius expression reflected in Eq. (5):

$$c_v = c_v(T_{ref}) \cdot \exp\left(-\frac{E_a}{R} \cdot \left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right) \quad (5)$$

where $c_v(T_{ref})$ is the stochastic rate constant of the reaction type v at the reference temperature T_{ref} , E_a is the activation energy, R is the ideal gas constant, T is the system temperature and T_{ref} is the reference temperature.

The subsequent step of the algorithm is aimed to the determination of the cumulative reaction probability distribution (D_R), which grouped all M reactions that can take place in the mixture at a given time t .

A first random number (RN_1) is obtained to determine the reaction time step (Δt) according to the Eq. (6), as proposed Gillespie¹²:

$$\Delta t = -\frac{\ln(RN_1)}{\sum_i^M R_i} \quad (6)$$

The distribution D_R is randomly sampled in order to select the next reaction (μ) to be executed in the next reaction time step. For this purpose, a second random number between 0 and 1 is drawn (RN_2) and used to select the next reaction from the cumulative distribution D_R , as shown in Eq. (7):

$$D_R(\mu - 1) < RN_2 \leq D_R(\mu) \quad (7)$$

Once the reaction and the time step have been defined, the system is updated by the execution of the selected reaction, which implies the replace of the reactant(s) by the product(s) and the increase of simulation time in Δt . The simulation continues reaction by reaction until the final simulation time is reached and a set of product molecules is obtained.

Due to its stochastic nature, the algorithm should be executed several times in order to provide several sets of molecules which average can be a proper representation of the molecules in the reaction system.

3. Application to the molecular reconstruction of LCO gas oils

LCO gas oils are gas oil fractions obtained from the catalytic cracking process and composed by a mixture of saturates (paraffins and naphthenes), olefins and aromatic hydrocarbons⁸.

These mixtures also show a marginal presence of nitrogen but an important content in sulphur, which is present in heterocyclic structures, such as thiophenes and benzothiophenes²⁸. Their boiling point is in the range of 120°C to 450°C, which corresponds to the molecules with a number of carbons between C₈ and C₃₀²⁹. According to this data and the mass spectrum, LCO gas oils are formed by molecules with none or one poly-aromatic structure with a number of rings which usually ranges from 0 to 4³⁰⁻³².

In the current work, it has been performed the molecular reconstruction of eleven LCO gas oils using the SR and REM algorithms. The PDF's of molecular attributes in Table 1 is used to generated the set of molecules that would represent the LCO gas oils. Two kind of PDF's (histograms and gamma functions) were used to defined molecular attributes. The use of histograms is limited to the attributes with a narrow range of possible values (less than three) while the gamma functions describe the attributes with a higher number of possible values. These PDF's have been defined taking into account the previous knowledge about LCO gas oils composition. Firstly, it is considered that the molecules in mixture are only constituted by carbon, hydrogen and sulphur due to the negligible presence of nitrogen species. Then, nitrogen content is fully assigned to carbon fraction, the most abundant element in mixture. Another hypothesis is that paraffins, which would also include the marginal olefins fraction, are formed by alkyl chains with at least 8 carbons. It is also considered that, according to experimental data, each molecule is formed by a maximum of 4 cycles, in which linear alkyl chains can be the unique substituents, and that the sulphur containing compounds present in the mixture only belong to the thiophenes family and its derivatives.

217 In order to define the relationships between distributions and also identify the sampling steps,
218 a building diagram is needed for the reconstruction of the molecules in LCO gas oils (Figure
219 3). First, the type of molecule to be constructed (saturate or aromatic) is determined by
220 distribution 1. In the case of saturate molecules, distribution 3 allows to determine the number
221 of cyclohexane rings (ranged from 0 to 4). When the number of rings present is null, the
222 molecule to be constructed is a paraffin, which chain length is determined by the application
223 of distribution 8. Contrary, in the case of a naphthenic units with between 1 and 4
224 cyclohexane rings, it is necessary to determine the number (distribution 6) and type
225 (distribution 7) of the side chains. In the case of aromatic molecules, the number of benzenes,
226 cyclohexanes and thiophenes is determined by distributions 2, 3 and 4, respectively. These
227 three distributions are sampled until the total number of rings is inferior or equal to 4. The
228 number of benzenes is always 1 or higher to avoid the creation of naphthenic molecules in
229 this route. Besides, according to experimental data, most of the thiophene rings do not contain
230 cyclohexane rings, so that distribution 3 is not sampled when thiophene rings are present.
231 Once the number and type of cycles is defined, the distributions 5 and 6 are used to find the
232 number of side chains in naphthenic and aromatic rings, respectively, and the distribution 7 is
233 applied to determine the type of chain.

234 Elemental analysis (carbon, hydrogen, sulphur and nitrogen), average molecular weight,
235 specific gravity, simulated distillation, mass spectrometry (MS), ^1H NMR and ^{13}C NMR were
236 used to characterise the LCO gas oils analytically. Elemental analysis is used to remove the
237 molecules which are composed by an element with a null experimental value. For example, if
238 sulphur presence has not been detected, all sulphureted molecules (benzothiophenes and
239 dibenzothiophenes) are eliminated. In a similar way, all molecules belonging to a
240 spectrometric family which has not been found in mass spectrometry analyses are suppressed.
241 Spectrometric families are determined according to the number of aromatic and thiophene
242 rings in compounds. Thus the following families can be distinguished: saturates (paraffins and
243 cycloparaffins), monoaromatics (alkylbenzenes, tetralins...), diaromatics (naphthalenes,
244 diphenyls...), triaromatics and compounds with more than three aromatic rings (anthracenes,
245 phenanthrenes...), benzothiophenes and dibenzothiophenes. Meanwhile, simulated distillation
246 enables to discard the molecules with a boiling temperature 5°C under the initial or 5°C over
247 the final boiling point. This correction is introduced due to the bias between the real boiling
248 point and the value obtained from simulated distillation curve³³. The “experimental”

molecular weight was calculated using an API correlation based on specific gravity and simulated distillation ³⁴.

3.1. Sensitivity of the stochastic method

The average properties of the simulated mixture of molecules are calculated by the application of a number of linear mixing rules. An objective function (*OF*) is used to perform the comparison between calculated and measured properties. This function is the sum of the relative deviations between calculated and experimental data and it is given by Eq. (8):

$$OF = \frac{1}{N_p} \cdot \sum_{i=1}^{N_p} \delta_i \quad (8)$$

where N_p is the number of properties present in objective function and δ_i is the relative deviation between the calculated and experimental values of property i which is calculated following Eq. (9):

$$\delta_i = \frac{1}{N_{M,i}} \cdot \sum_{j=1}^{N_{M,j}} \frac{|X_{j,i}^{exp} - X_{j,i}^{calc}|}{X_{j,i}^{exp}} \quad (9)$$

where $N_{M,i}$ is the number of measurements of property i (for example, the number of points in the simulated distillation) and $X_{j,i}^{exp}$ and $X_{j,i}^{calc}$ are, respectively, the experimental and calculated values of measurement j of property i .

An elitist genetic algorithm developed by Pereira *et al.* ¹⁸ was applied in order to minimise this objective function by the modification of the parameters of the PDF's for the structural attributes. Genetic algorithms ^{22–24}, as well as simulated annealing ^{15,19–21} and particle swarm optimization ³⁵, are robust global optimizers, a characteristic required for their application in stochastic models. Besides, according to its elitist nature, the applied methodology allows to a certain percentage of individuals to move to the following generation (in this case, a percentage of 50% was considered), with which the best individuals (i.e. those with the lowest *OF* values) remain in the next generation.

The number of molecules is an important parameter in molecular reconstruction since an excessively low number would lead to an inadequate representation of the attributes of the represented mixture. Otherwise, the use of a large number of molecules would imply a significant increase in calculation time. In order to determine the optimal value for this parameter, the sensibility of the stochastic method has been tested by the generation of 100 mixtures of N molecules from a set of parameter distributions determined for a LCO gas oil.

N value was varied in a range between 10 and 20000 molecules and the average, minimum, maximum and standard deviation values were determined for the objective function in each point (Figure 4). In addition, the time necessary to perform the calculation was also represented.

It was observed that the increase of the size of the mixture leads to a drop of the objective function until attaining an asymptotic value at $N = 5000$ (Figure 4.a). The increase of the number of molecules also implies a rapprochement of the minimum and maximum values and a reduction in the standard deviation of the objective function. However, the drop of the standard deviation is very mild at $N > 5000$, while the calculation time continues to rise linearly with the size of the mixture (Figure 4.b). Then, a set of 5000 molecules was considered as optimal to perform the stochastic reconstruction of LCO gas oils.

3.2. Construction of the molecular representation of LCO gas oils

The analyses used in this case for the molecular reconstruction were the elemental analysis, MS and simulated distillation. The elemental analysis provides information on the number of thiophenes, while the information on the chemical structure is provided by MS and by simulated distillation in the case of the average molecule size and size distribution. The experimental and simulated properties were compared in Table 2 for a LCO gas oil. Properties that were not used in reconstruction step (specific gravity, molecular weight, ^1H NMR and ^{13}C NMR) were calculated from the predicted molecular mixture and also compared with the corresponding experimental values.

The properties of the mixture obtained after the application of the SR step are mostly in good agreement with the analytical results. Some deviations are nevertheless observed in the simulated distillation curve. These differences may be explained by the fact that the distillation curve is calculated under the assumption that all molecules are perfectly separated by the use of increasing boiling points. However, this is not strictly true (especially in the region near the initial and final distillation points), since the chromatographic columns are not completely apolar and separation is influenced by the polarity of the molecules²⁹.

On the other hand, the coupled application of SR and SEM schemes led to a significant improvement in the agreement between predicted and experimental data for most of the properties (Table 2). The most meaningful effects were observed on more specific and detailed analyses, such as NMR and MS, while only small improvements could be detected in the more global measurements, such as in the case of elemental analysis.

On the following step, a total of eleven LCO gas oils have been reconstructed, and the properties of the predicted mixtures have been compared with the corresponding experimental data in order to perform the validation of the current methodology (Figure 5). It was found a good agreement with the analytical results for most of the evaluated properties. However, a deviation was detected for the predicted average molecular weight, which was mostly placed below the value calculated from experimental data. This may be explained by the alterations in the determination of the distillation curve, which imply a reduction in the boiling point of the molecules with respect to their normal value. It leads to the prediction of molecules smaller than their counterpart and, consequently, to reduce the average molecular weight of the mixture.

These results show that the predicted set of molecules are proper molecular representations of the corresponding LCO gas oils. Some of these simulated mixtures have been used as inputs for the kinetic model of the hydrotreating of LCO gas oil.

4. Application to the hydrotreating of LCO gas oils

Once the representative set of molecules has been generated for each LCO gas oil, the hydrotreating reactions are simulated by applying the kMC method. The experimental data used in this step has been collected from previous studies developed at IFP Energies nouvelles about the hydrotreating of these oil fractions ³⁶.

Three main type of reactions are involved in the hydrotreating process: hydrodesulphurisation (HDS), hydrodenitrogenation (HDN) and hydrogenation of benzene and cyclohexane rings and of olefins. However, the reactions of hydrodenitrogenation and hydrogenation of olefins will be neglected due to the lack of nitrogen and olefins in the mixture. Similarly, the HDS reactions are limited to heterocyclic structures since it has been considered that sulphur is only present in the mixture as thiophenes and their derivatives.

Hydrogenation of aromatic rings (HDA), an exothermic and reversible reaction, is thermodynamically limited at typical hydrotreating conditions, so that the reverse reaction of dehydrogenation should also be considered. Korre *et al.* ³⁷ have indicated that no partially hydrogenated ring compounds, such as dihydronaphthalene and hexahydrophenanthrene, were found in the product stream, which shows that hydrogenation proceeds in a ring-by-ring manner. Besides it was detected that the higher number of aromatic rings, the higher reactivity of the molecule. This fact implies that polyaromatic species, such as pyrene, would be the most reactive whereas that the HDA of monoaromatics, such as alkyl benzenes and tetralines,

would be the slowest. Other qualitative trend observed during the process was that, for the molecules with more than two aromatic rings, the hydrogenation of the rings in the middle of the molecule goes faster than in the case of the external rings^{38,39}. To this, it should be added the fact that the present of alkyl substituents and of cyclohexane rings increases the hydrogenation rate³⁷.

HDS is an exothermic and basically irreversible reaction at the hydrotreating conditions⁴⁰ through which sulphur is removed from the petroleum cuts in form of hydrogen sulphide. The removal of this element is one of the main aims of hydrotreating process due to its catalyst poisoning and pollution potential¹¹. The HDS reactivity of sulphureted species closely depends on their structure. According with this, non-heterocyclic compounds present a very higher reactivity at hydrotreating conditions than heterocyclic species, such as thiophenes family and derivatives. Different reactivities are also found within this last compounds family, in which thiophenes (Ts) and benzothiophenes (BTs) show a higher reactivity than dibenzothiophenes (DBTs)⁴¹. The hydrotreating rates of the heterocyclic species are also influenced by the presence of substituent groups in the positions adjacent to the sulphur atom⁹. The desulphurisation process in these species goes through two different pathways: (1) the direct removal of the sulphur atom without ring hydrogenation (hydrogenolysis pathway) or (2) the saturation of the thiophene and the fused benzene rings prior to the removal of the sulphur atom from the thiophene ring (hydrogenation pathway). The HDS of Ts, BTs and unsubstituted DBTs mostly follows the hydrogenolysis pathway whereas hydrogenation becomes the preferred pathway for DBTs at the same time that the degree of substitution increases^{40,41}.

Taking into account the above information, the reaction types reflected in Table 3 has been considered during the stochastic modelling of LCO hydrotreating. A set of reaction rules is also established in order to identify the reaction events from the structure of the molecule. Hydrogenation of aromatic rings and dehydrogenation of saturated rings have been included in the current model as ring-by-ring reactions while it have been considered that sulphur is removed from thiophene rings as H₂S. Besides, the rate constants of the HDT reactions must be specified.

In the case of hydrogenation and dehydrogenation reactions, their reaction rates can be calculated from Eq. (10) and (11), respectively:

$$R_{Hydro} = c_{Hydro} \cdot p_{H_2} \quad (10)$$

$$R_{Dehydro} = c_{Dehydro} = \frac{c_{Hydro}}{K_{eq}} \quad (11)$$

where c_{Hydro} and $c_{Dehydro}$ are the stochastic reaction constants for hydrogenation and dehydrogenation (s^{-1}), K_{eq} is the equilibrium constant ($mol \cdot (mol \cdot atm)^{-1}$), and p_{H_2} is the partial pressure of hydrogen (atm).

The expressions for the calculation of hydrogenation rate constants and the equilibrium constants have been developed in previous studies¹⁷, starting from the quantitative structure/reactivity correlations (QS/RCs) proposed by Korre *et al.*³⁷. They are given in Eq. (12) and (13), respectively:

$$\begin{aligned} \ln(c_{Hydro}) = & -4.65 - 7.06 \cdot n_H^{1.18} - 1.08 \cdot \left| \frac{\Delta H_R^0}{R} \right| \\ & + 0.616 \cdot N_{AR} + 0.330 \cdot N_{SR} - 2.80 \cdot N_{TR} \\ & - \frac{E_a}{R} \cdot \left(\frac{1}{T} - \frac{1}{T_{Ref}} \right) \end{aligned} \quad (12)$$

$$\begin{aligned} \ln(K_{eq}) = & 2.952 - 13.215 \cdot n_H + 5.196 \cdot 10^{-3} \cdot \left| \frac{\Delta H_R^0}{R} \right| \\ & - 0.784 \cdot N_{SR} \\ & - \frac{\Delta H_R}{R} \cdot \left(\frac{1}{T} - \frac{1}{T_{Ref}} \right) \end{aligned} \quad (13)$$

where n_H is the stoichiometric coefficient for molecular hydrogen, ΔH_R^0 is the heat of reaction at 25°C ($J \cdot mol^{-1}$), N_{AR} is the number of aromatic rings in the molecule, N_{SR} is the number of saturated rings, N_{TR} is the number of thiophenes fused to the aromatic ring, E_a is the activation energy ($J \cdot mol^{-1}$), R is the ideal gas constant ($J \cdot (K \cdot mol)^{-1}$), T is the temperature of the system (K), and T_{ref} is the reference temperature (K).

The values of the coefficients in the K_{eq} expression proposed by Korre *et al.*³⁷ have been modified in order to obtain a better agreement between predicted and experimental equilibrium data. The experimental data of a set of 21 compounds, including anthracene, phenanthrene and benzene, have been used to determine the new constants. For the hydrogenation rate constant c_{Hydro} , the N_{TR} term was added to reflect the effect of the presence of thiophene rings on the hydrogenation of aromatic rings, as observed in hydrodesulphurisation kinetics⁴⁰. An exponential coefficient was also added to the n_H term to improve the prediction of the kinetics of benzoic compounds. The hydrogen exponential coefficient, the coefficient for the thiophene correction and the independent term in the expression for k_{SR} were determined by minimising the difference between the predicted and

experimental data for the gas oil LCO1 (Table 4) during its hydrotreating at 300 °C (Figure 6.a).

A very good agreement between experimental and predicted data was obtained for monoaromatic compounds. Besides, the fast hydrogenation of the aromatic compounds with more than two aromatic rings was also properly reflected but some differences could still be observed for saturates and diaromatic compounds. Despite they followed the same tendency than the predicted data, the experimental data showed a larger removal of diaromatic compounds and also a higher concentration of saturated compounds. The introduction of hydrogen exponential coefficient may be the cause for this deviation since, although it provides a better prediction of benzoic compounds kinetics, it worsens the results for other molecules, such as naphthalenes.

HDS reactions have been classified in five families according to the structure of the reacting molecule and a rate parameter was assigned to each one (Table 4). The reaction rates are the component directly given by the corresponding rate constant following Eq. (15):

$$R_{HDS} = c_{HDS} \quad (15)$$

The rate constants c_{HDS} were obtained from the fit to the experimental data of the LCO gas oil which was previously used to determine the coefficients in Eq. (15) (Figure 6.b). In this case, the predicted curves perfectly fit to experimental data of both BT and DBT compounds.

In order to validate the current modelling methodology, the hydrotreating of the three LCO gas oils in Table 4 has been simulated at different conditions of temperature and partial H_2 pressure. These LCO gas oils present different composition profiles. LCO1 and LCO2 both present a significant content in sulphur aromatics (mainly DBT in LCO1 and BT in LCO2) while, in the case of LCO3, the content in these compounds is very lower (less than 1.5% wt) and the addition of sulphides such as dimethyl disulphide (DMDS) is required in order to reach an adequate partial pressure of H_2S over the process. Besides, LCO3 is mainly composed by saturated compounds (almost 60% wt) unlike the other two gas oils, in which diaromatics are the most abundant compounds.

The predicted results after 1 h of operation have been compared with analytical data in the parity plots depicted in Figure 7. A good agreement with experimental data was found for saturated and monoaromatic compounds with an average relative error lower than 14%. The prediction of the species with three or more aromatic rings was accurate in most of the

experimental points and a very good fit was also found for sulphur aromatic compounds. Performance variations due to modifications in operational conditions, i.e. temperature and hydrogen partial pressure, were contemplated and well-reflected on the current model. On the other hand, as in the case of LCO1, some differences were also detected for diaromatic compounds in LCO3. High complexity of mixtures, which leads to the use of correlations to determine the reactivity of each molecule in hydrogenation/dehydrogenation processes, implies a greater difficulty to accurately predict the content in all aromatic species, although a good global agreement between predicted and experimental data was reached.

5. Conclusions

A stochastic two-steps procedure has been described and validated on the hydrotreating of three LCO gas oils at different operational conditions. In the first step, a molecular reconstruction algorithm is used to represent the feedstock by means of a set of molecules which mixture properties are similar to those of feedstock. This algorithm is formed by the combination of two methods: stochastic reconstruction (SR) and reconstruction by entropy maximization (REM). The SR method leads to obtain a set of molecules typical of a given feedstock while the molar fraction of molecules is adjusted using the REM method in order to improve the predicted mixture properties. Then, the generated set of molecules is used as input in the second step concerning the molecule-based kinetic modelling of the hydrotreating process and in which a kinetic Monte Carlo (kMC) method is applied. The expressions for the calculation of adsorption, equilibrium and hydrogenation/dehydrogenation rate constants have been obtained by applying the quantitative structure/reactivity correlations (QS/RCs).

The current stochastic methodology has demonstrated to be a valid tool for the reconstruction of LCO gas oils by combining their available analytical data. This procedure can be used not only to get the molecular information that is not provided by analytical data but also to obtain an accurate molecular representation which can be used as input in kinetic models. Monte Carlo techniques have also given good results in their application on the kinetic simulation of the hydrotreating of LCO gas oils. Moreover, the simulation of the reactions does not require a pre-defined kinetic network due to it is generated as the reactions proceed. Besides, the use of the corrected QS/RCs led to a good prediction of molecular reactivities in complex mixtures of hydrocarbons such as LCO gas oils.

462

463 References

- 464 1. Quann, R. & Jaffe, S. Structure-Oriented Lumping: Describing the Chemistry of
465 Complex Hydrocarbon Mixtures. *Ind. Eng. Chem. Res.* **31**, 2483–2497 (1992).
- 466 2. Quann, R. & Jaffe, S. Building useful models of complex reaction systems in
467 petroleum refining. *Chem. Eng. Sci.* **5**, 1615–1635 (1996).
- 468 3. Vynckier, E. & Froment, G. in *Kinet. Thermodyn. Lumping Multicomponent Mix.*
469 (Astarita, G. & Sandler, S.) 131–161 (Elsevier B.V., 1991).
- 470 4. Klein, M. T., Hou, G., Bertolacini, R. J., Broadbeit, L. J. & Kumar, A. *Molecular*
471 *modeling in heavy hydrocarbon conversions*. (CRC Press Taylor & Francis Group,
472 2006).
- 473 5. Weekman, V. W. Lumps, models, and kinetics in practice. *AIChE Monogr. Ser.* **75**, 1–
474 29 (1979).
- 475 6. Sanchez, S., Rodriguez, M. A. & Ancheyta, J. Kinetic Model for Moderate
476 Hydrocracking of Heavy Oils. *Ind. Eng. Chem. Res.* **44**, 9409–9413 (2005).
- 477 7. Verstraete, J. J., Le Lannic, K. & Guibard, I. Modeling fixed-bed residue hydrotreating
478 processes. *Chem. Eng. Sci.* **62**, 5402–5408 (2007).
- 479 8. López García, C., Hudebine, D., Schweitzer, J.-M., Verstraete, J. J. & Ferré, D. In-
480 depth modeling of gas oil hydrotreating: From feedstock reconstruction to reactor
481 stability analysis. *Catal. Today* **150**, 279–299 (2010).
- 482 9. Pereira de Oliveira, L., Verstraete, J. J. & Kolb, M. A Monte Carlo modeling
483 methodology for the simulation of hydrotreating processes. *Chem. Eng. J.* **207-208**,
484 94–102 (2012).
- 485 10. De Oliveira, L. P., Verstraete, J. J. & Kolb, M. Molecule-based kinetic modeling by
486 Monte Carlo methods for heavy petroleum conversion. *Sci. China Chem.* **56**, 1608–
487 1622 (2013).

- 488 11. De Oliveira, L. P., Verstraete, J. J. & Kolb, M. Simulating vacuum residue
489 hydroconversion by means of Monte-Carlo techniques. *Catal. Today* **220-222**, 208–220
490 (2014).
- 491 12. Gillespie, D. T. A general method for numerically simulating the stochastic time
492 evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434 (1976).
- 493 13. Liguras, D. K. & Allen, D. T. Structural models for catalytic cracking. 2. Reactions of
494 simulated oil mixtures. *Ind. Eng. Chem. Res.* **28**, 674–683 (1989).
- 495 14. Neurock, M., Nigam, A., Trauth, D. & Klein, M. T. Molecular representation of
496 complex hydrocarbon feedstocks through efficient characterization and stochastic
497 algorithms. *Chem. Eng. Sci.* **49**, 4153–4177 (1994).
- 498 15. Hudebine, D. Reconstruction moléculaire de coupes pétrolières. PhD thesis. (2003).
- 499 16. Hudebine, D. & Verstraete, J. J. Reconstruction of Petroleum Feedstocks by Entropy
500 Maximization. Application to FCC Gasolines. *Oil Gas Sci. Technol. – Rev. d'IFP*
501 *Energies Nouv.* **66**, 437–460 (2011).
- 502 17. Pereira de Oliveira, L. Développement d'une méthodologie de modélisation cinétique
503 de procédés de raffinage traitant des charges lourdes. PhD thesis. (2013).
- 504 18. Oliveira, L. P. De, Vazquez, A. T., Verstraete, J. J. & Kolb, M. Molecular
505 Reconstruction of Petroleum Fractions : Application to Vacuum Residues from
506 Different Origins. *Energy & Fuels* **27**, 3622–3641 (2013).
- 507 19. Trauth, D. M., Stark, S. M., Petti, T. F., Neurock, M. & Klein, M. T. Representation of
508 the Molecular Structure of Petroleum Resid through Characterization and Monte Carlo
509 Modeling. *Energy & Fuels* **8**, 576–580 (1994).
- 510 20. Trauth, D. M. Structure of complex mixtures through characterization, reaction and
511 modelling. PhD thesis. (1993).
- 512 21. Hudebine, D. & Verstraete, J. J. Molecular reconstruction of LCO gasoils from overall
513 petroleum analyses. *Chem. Eng. Sci.* **59**, 4755–4763 (2004).

- 514 22. Verstraete, J. J., Revellin, N. & Dulot, H. Molecular reconstruction of vacuum gasoils.
515 *Prepr. Pap. Chem. Soc. Div. fuel Chem.* **49**, 20–21 (2004).
- 516 23. Verstraete, J. J., Schnongs, P., Dulot, H. & Hudebine, D. Molecular reconstruction of
517 heavy petroleum residue fractions. *Chem. Eng. Sci.* **65**, 304–312 (2010).
- 518 24. Schnongs, P. *Reconstruction moléculaire de coupes pétrolières lourdes*. (2005).
- 519 25. Shannon, C. E. A Mathematical Theory of Communication. *bell Syst. Tech. J.* **27**, 379–
520 423, 623–656 (1948).
- 521 26. Gillespie, D. T. A rigorous derivation of the chemical master equation. *Phys. A Stat.*
522 *Mech. its Appl.* **188**, 404–425 (1992).
- 523 27. Gillespie, D. T. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **58**,
524 35–55 (2007).
- 525 28. López García, C., Becchi, M., Grenier-Loustalot, M. F., Païsse, O. & Szymanski, R.
526 Analysis of aromatic sulfur compounds in gas oils using GC with sulfur
527 chemiluminescence detection and high-resolution MS. *Anal. Chem.* **74**, 3849–57
528 (2002).
- 529 29. Hudebine, D. & Verstraete, J. J. Molecular reconstruction of LCO gas oils from overall
530 petroleum analyses. *Chem. Eng. Sci.* **59**, 4755–4763 (2004).
- 531 30. Cooper, B. H. & Donnis, B. B. L. Aromatic saturation of distillates : an overview. *Appl.*
532 *Catal. A Gen.* **137**, 203–223 (1996).
- 533 31. Castex, H., Boulet, R., Juguin, J. & Lepinasse, A. Analyse des kérosènes et des gas oils
534 moyens par spectrométrie de masse à moyenne résolution. *Oil Gas Sci. Technol.* **38**,
535 523–532 (1983).
- 536 32. Fafet, A., Bonnard, J. & Prigent, F. New Developments in Mass Spectrometry for
537 Group-Type Analysis of Petroleum Cuts. *Oil Gas Sci. Technol.* **54**, 429–452 (1999).
- 538 33. ASTM D2887. in *Annu. B. ASTM Stand.* (2007).
- 539 34. API 282.1. in *API Tech. Handb.* (1987).

- 540 35. Kennedy, J. & Eberhart, R. C. *Swarm intelligence*. (Morgan Kaufmann Publishers,
541 2001).
- 542 36. Lopez Garcia, C. Analyse de la réactive des composés soufrés dans les coupes
543 pétrolières: cinétique et modélisation de l'hydrotraitement. PhD thesis. (2000).
- 544 37. Korre, S., Neurock, M., Klein, M. T. & Quann, R. J. Hydrogenation of polynuclear
545 aromatic hydrocarbons.2. Quantitative Structure/Reactivity Correlations. *Chem. Eng.*
546 *Sci.* **49**, 4191–4210 (1995).
- 547 38. Stanislaus, A. & Cooper, B. H. Aromatic hydrogenation catalysis: a review. *Catal. Rev.*
548 **36**, 75–123 (1994).
- 549 39. Korret, S. C., Klein, M. T. & Quann, R. J. Polynuclear Aromatic Hydrocarbons
550 Hydrogenation . 1 . Experimental Reaction Pathways and Kinetics. *Ind. Eng. Chem.*
551 *Res.* **34**, 101–117 (1995).
- 552 40. Girgis, M. J. & Gates, B. C. Reactivities, Reaction Networks, and Kinetics in High-
553 Pressure Catalytic Hydroprocessing. *Ind. Eng. Chem. Res.* 2021–2058 (1991).
- 554 41. Whitehurst, D. D., Isoda, T. & Mochida, I. Present State of the Art and Future
555 Challenges in the Hydrodesulfurization of Polyaromatic Sulfur Compounds. *Adv.*
556 *Catal.* **42**, 345–471 (1998).

Figure captions

Figure 1. Flowchart of the SR-REM algorithm

Figure 2. Flowchart of the kMC method

Figure 3. Building diagram for LCO gas oils

Figure 4. Evolution as a function of the number of molecules in the predicted mixture of **(a)** the average value of the objective function (white diamonds) and the maximum (black triangles) and minimum (black squares) values **(b)** the CPU time required (white diamonds) and the standard deviation of the objective function (black crosses)

Figure 5. Parity plots for the predicted properties of eleven gas oils after SR-REM application: **(a)** density (g/cm^3), **(b)** molecular weight (g/mol), **(c)** ^{13}C NMR signatures (% mole): primary saturated CH_3 (black diamonds), secondary saturated CH_2 (white squares), tertiary saturated CH (white rounds), aromatic CH (black crosses), condensed aromatic C (white triangles) and substituted aromatic C (black rounds), and **(d)** ^1H NMR signatures (% mole): diaromatic H (black diamonds), monoaromatic H (white squares), H alpha (white rounds), H beta (black crosses) and H gamma (white triangles)

Figure 6. Comparison between experimental (points) and predicted values (lines) for LCO1: **(a)** for aromatic families: saturates (black diamonds and black continuous line), monoaromatics (black squares and black discontinuous line), diaromatics (white squares and grey discontinuous line) and compounds with more than two aromatic rings (white diamonds and grey continuous line); and **(b)** for sulphur families: benzothiophenes (black squares and continuous line) and dibenzothiophenes (white squares and discontinuous line). All data are expressed as weight fraction (%)

Figure 7. Parity plots for the compounds concentration predicted by means of kMC method after 1 h of hydrotreating: **(a)** saturates, **(b)** monoaromatics, **(c)** diaromatics, **(d)** compounds with more than two aromatic rings and **(e)** sulphur compounds: benzothiophenes (black diamonds) and dibenzothiophenes (white rounds). In the case of benzothiophenes all points overlap each other due to their similar values close to point (0.0, 0.0)

Table 1. Structural attributes of LCO gas oil molecules

	Structural attribute	Values	Distribution	Parameter names
1	Type of molecule	0 or 1	Histogram	0
2	Number of benzene rings per core	1, ..., 4	Gamma	1 and 2
3	Number of naphthenic rings per core	0, ..., 4	Gamma	3 and 4
4	Number of thiophenes per core	0 or 1	Histogram	5
5	Acceptance probability for aromatic carbon CH	0 or 1	Histogram	6
6	Acceptance probability for naphthenic carbon CH ₂	0 or 1	Histogram	7
7	Type of the side chains	0, 1 or 2	Histogram	8 and 9
8	Length of the paraffinic chains	> 0	Gamma	10 and 11

Table 2. Comparison between experimental and calculated properties after the SR and REM steps

		Experimental	Simulated (after SR)	Simulated (after SR- REM)
<i>Elemental analysis (%wt)</i>	Carbon	89.95	89.85	89.99
	Hydrogen	9.21	9.33	9.15
	Sulphur	0.84	0.82	0.86
<i>Simulated distillation (°C)</i>	Initial boiling point	143	143	143
	5%	214	175	211
	10%	231	211	232
	20%	250	250	250
	30%	259	268	267
	40%	273	288	270
	50%	284	314	284
	60%	299	334	299
	70%	314	351	315
	80%	330	369	334
	90%	349	385	355
	95%	363	393	383
	Final boiling point	398	403	403
<i>Mass spectrometry (%wt)</i>	Saturates	11.00	10.49	11.12
	Monoaromatics	12.90	17.07	12.98
	Diaromatics	58.40	49.25	58.36
	Triaromatics	8.50	16.40	8.48
	Tetraaromatics	3.00	1.53	2.97
	Benzothiophenes	4.00	1.39	3.96
	Dibenzothiophenes	2.20	3.86	2.13
<i>C NMR (%wt)</i>	Aromatic CH	39.10	34.23	36.62
	Fused aromatic C	8.00	12.17	12.03
	Substituted aromatic C	15.30	14.69	14.95
	Saturated CH ₃	15.20	15.62	14.96
	Saturated CH ₂	19.60	19.35	18.19
	Saturated CH	2.70	3.93	3.25
<i>H NMR (%wt)</i>	Diaromatic H	25.70	24.26	27.64
	Monoaromatic H	2.30	3.40	2.59
	Saturated alpha H	33.90	27.39	27.61
	Saturated beta H	27.30	29.33	27.80
	Saturated gamma H	10.80	15.61	14.37
<i>Others</i>	Molecular weight (g/mol)	192	194	184
	Density 20°C (g/cm ³)	0.9701	0.9710	0.9729

Table 3. Reaction types and their parameters (*c*)

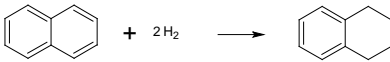
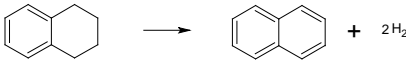
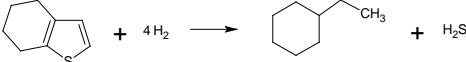


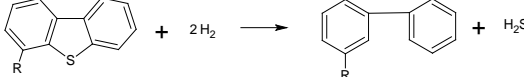

Type of reaction	Reaction	<i>c</i> (s ⁻¹)
Hydrogenation		Eq. (10)
Dehydrogenation		Eq. (11)
Hydrodesulphurisation	Hydrogenolysis of thiophene and their derivatives 	138.90
	Hydrogenolysis of BT and their derivatives 	55.60
	Hydrogenolysis of DBT and their derivatives 	19.40
	Hydrogenolysis of 4-alkyl-DBT and their derivatives 	0.56
	Hydrogenolysis of 4,6-dialkyl-DBT and their derivatives 	0.06

Table 4. Hydrotreating conditions (partial H₂ pressure and temperature) and properties from SR-REM algorithm for the three LCO gas oils used as input in the kMC method

	LCO1	LCO2	LCO3
Hydrotreating conditions			
p_{H_2} (atm)	69.1	95.1	46.1
T (°C)	300/330	300/330	300/310/320
Composition			
C (%wt)	88.01	88.57	87.23
H (%wt)	10.51	9.74	12.57
S (%wt)	1.48	1.69	1.45
Saturates (%wt)	28.58	19.85	59.84
Monoaromatics (%wt)	16.14	22.23	9.55
Diaromatics (%wt)	34.24	38.96	26.20
Triaromatics and more (%wt)	11.49	8.95	3.19
BT (%wt)	2.70	6.39	0.67
DBT (%wt)	6.86	3.61	0.56

Figure 1.

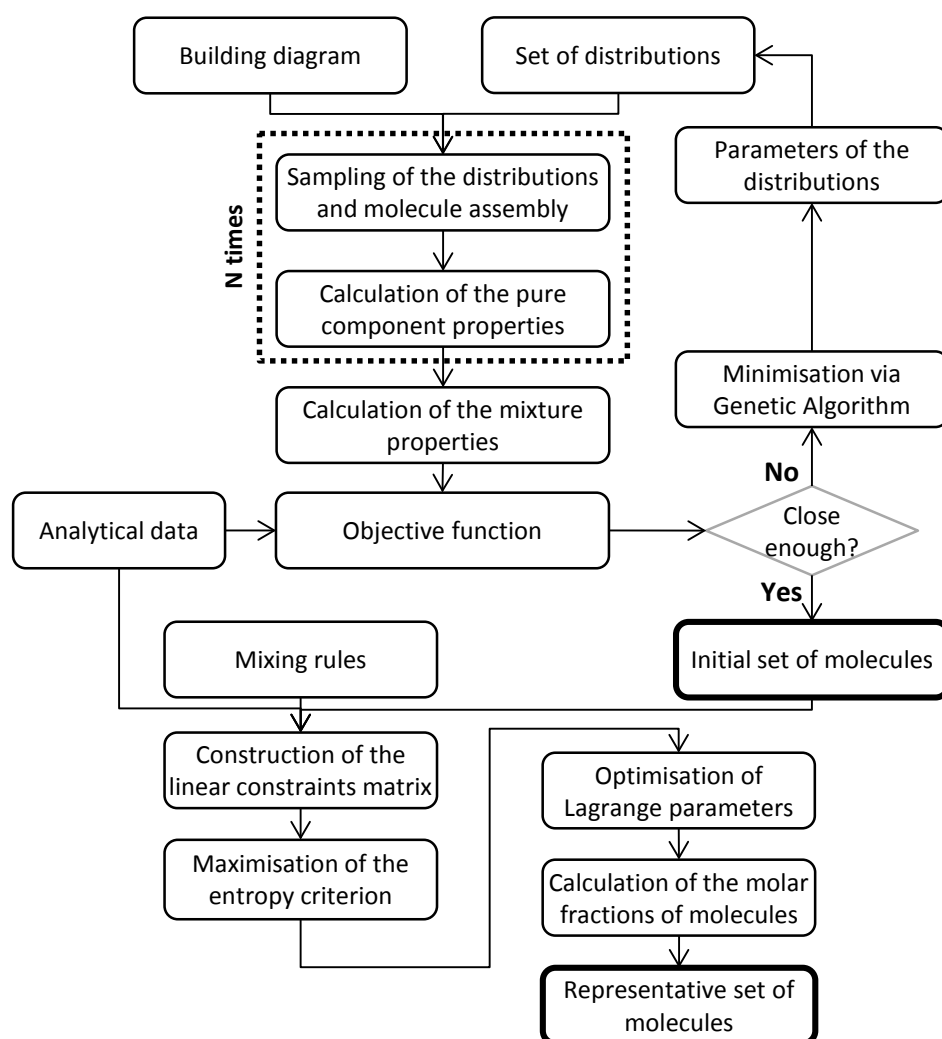


Figure 2.

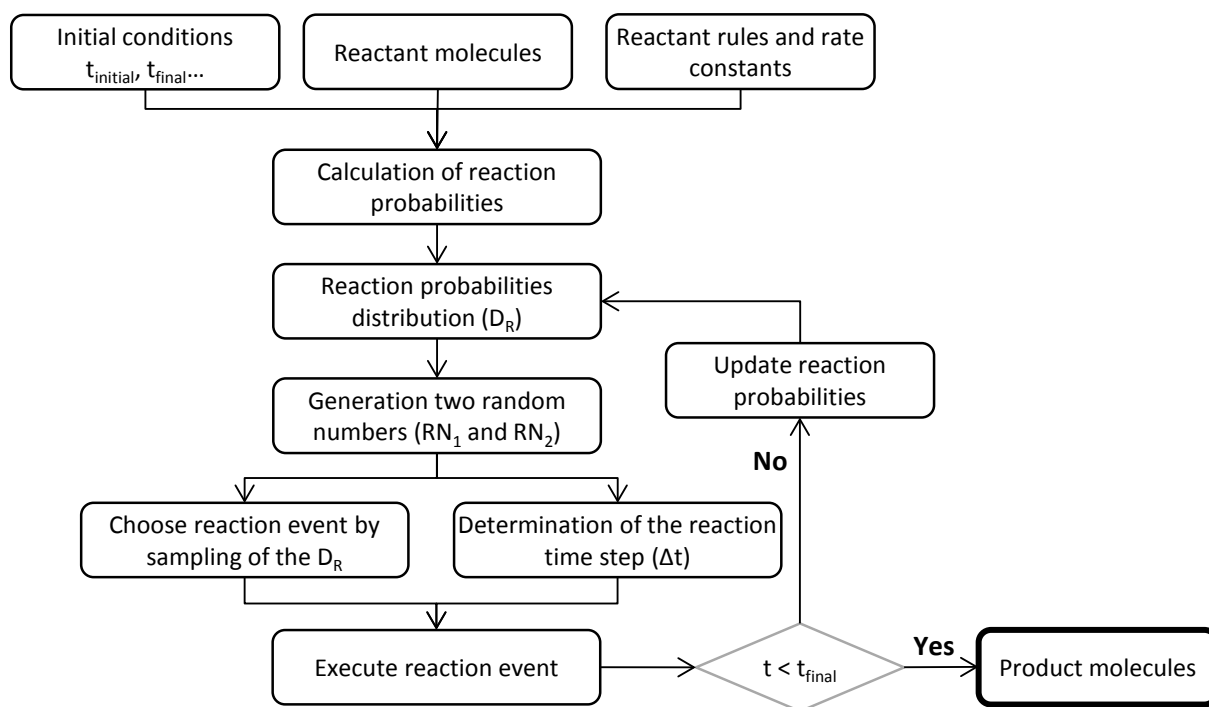


Figure 3.

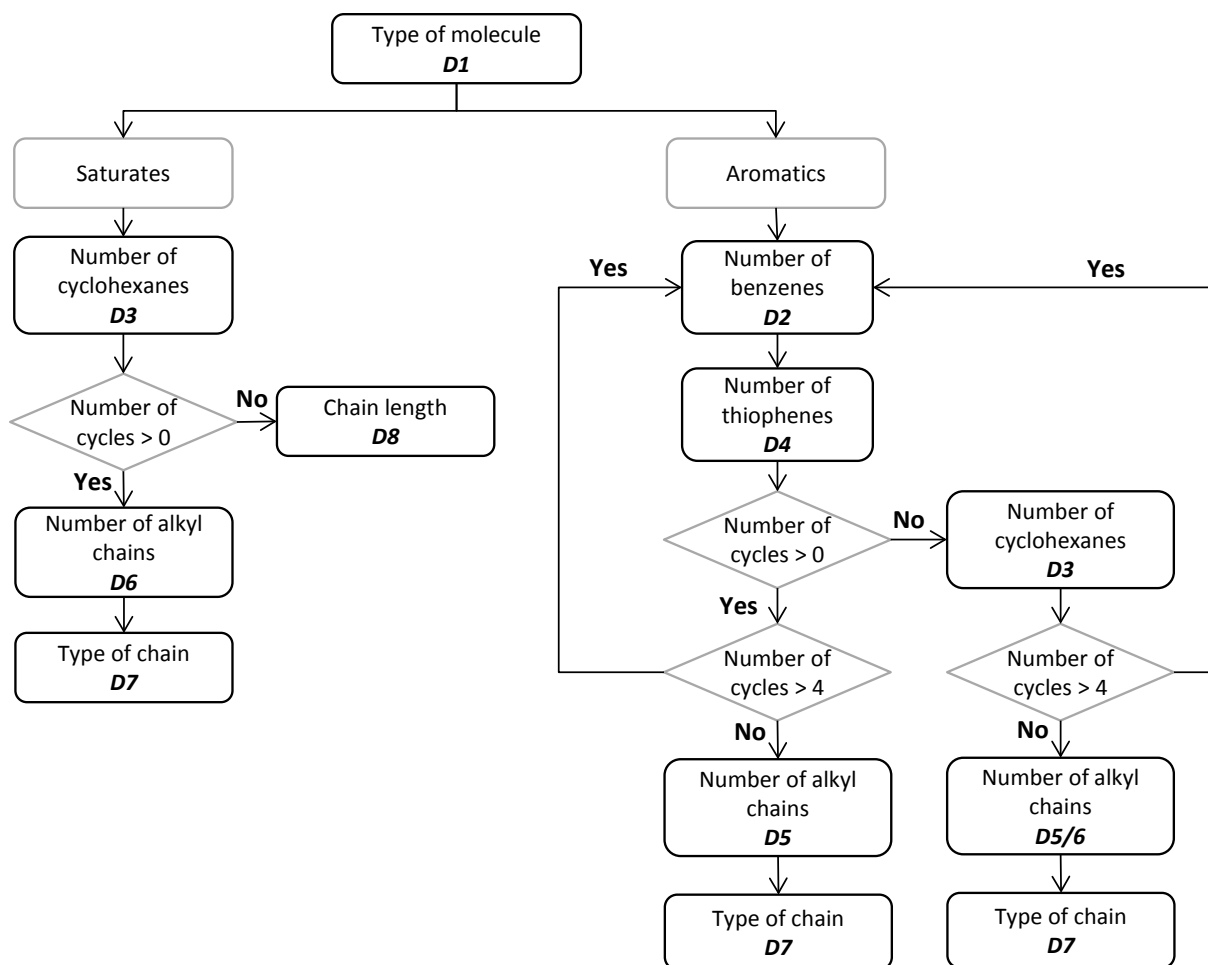


Figure 4.

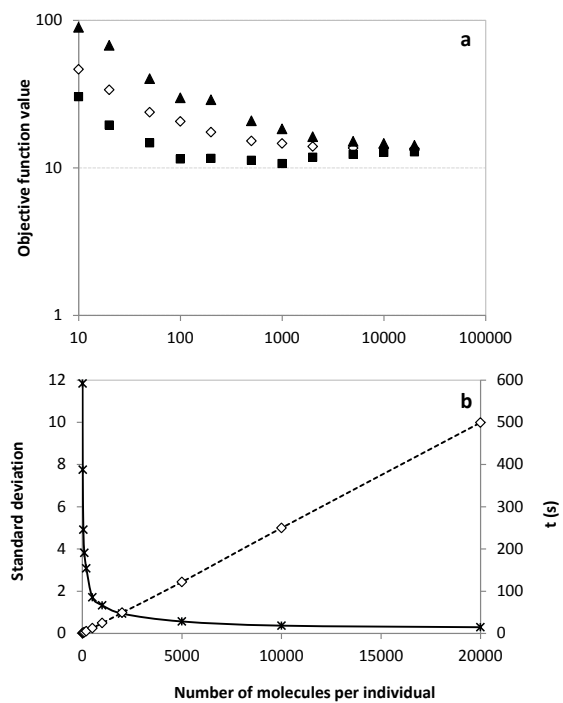


Figure 5.

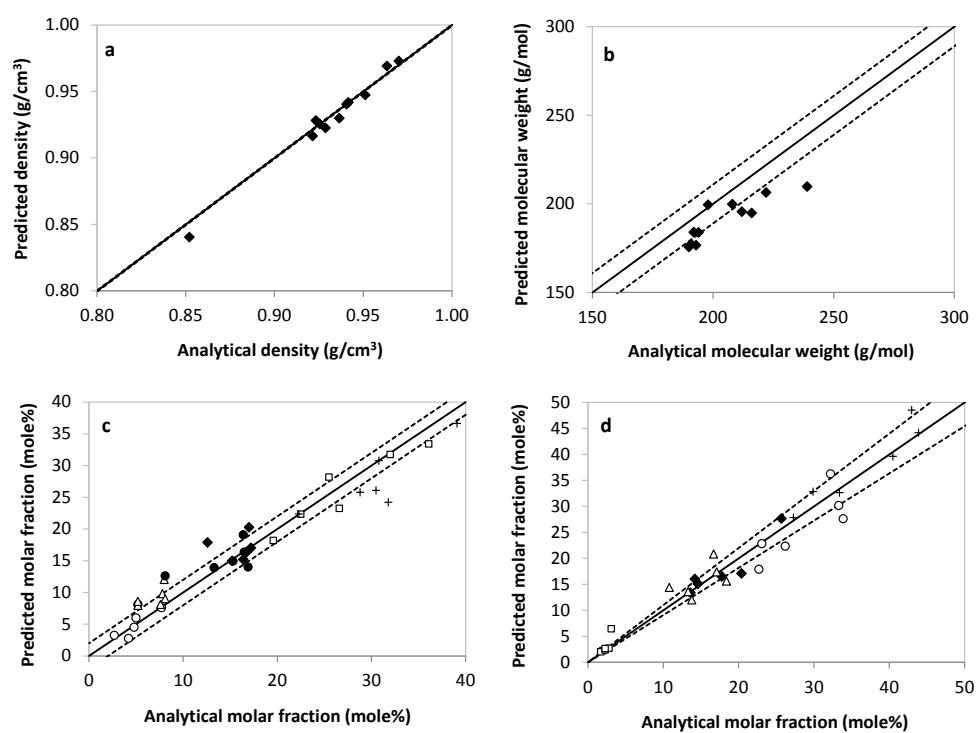


Figure 6.

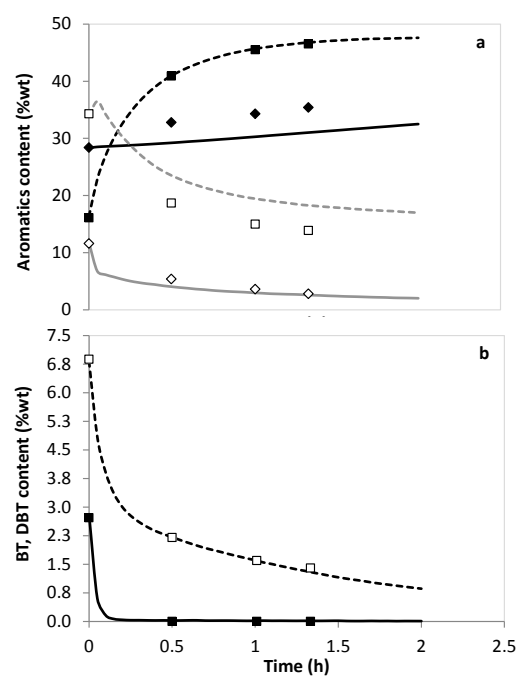


Figure 7.

