

Challenges of Modeling Steam Cracking of Heavy Feedstocks

K.M. Van Geem, M.F. Reyniers and G.B. Marin

Laboratorium voor Petrochemische Techniek, UGent, Krijgslaan 281 (S5), 9000 Gent - Belgium
e-mail: kevin.vangeem@ugent.be

Résumé — Les défis de la modélisation du vapocraquage des hydrocarbures lourds — Actuellement les modèles cinétiques du vapocraquage des hydrocarbures, par événements constitutifs, permettent de simuler la conversion de fractions lourdes. Le principal défi pour modéliser le craquage de ces matières lourdes est leur reconstruction moléculaire. Celle-ci dépend du niveau de précision moléculaire requis du réseau réactionnel et de la caractérisation/méthode de reconstruction de la charge, comme cela est illustré pour les condensats du gaz naturel par exemple. La comparaison entre les prédictions et les résultats obtenus dans une unité pilote montre comment les incertitudes au niveau de la reconstruction de la charge se propagent au niveau des résultats de la simulation. La combinaison du modèle cinétique par événements constitutifs et de la méthode de reconstruction basée sur la maximisation de l'entropie de Shannon, permet d'obtenir des résultats précis si la densité, la repartition en poids ou en volume des fractions PIONA, et les points d'ébullition initiaux, à 50 % et finals sont connus. Moins il y a d'indices spécifiés, moins il y a de correspondance entre les résultats simulés et ceux obtenus par expérimentation. La méthodologie développée peut être très simplement élargie à d'autres coupes pétrolières.

Abstract — Challenges of Modeling Steam Cracking of Heavy Feedstocks — Today single event microkinetic (SEMK) models for steam cracking of hydrocarbons allow simulating the conversion of heavy fractions. The key challenge to model the cracking behavior of these heavy feedstocks is related to feedstock reconstruction. The latter depends on the required level of molecular detail of the reaction network and of the feedstock characterization/ reconstruction model. This is illustrated for gas condensate feedstocks. Comparison of yield predictions with yields obtained in a pilot plant illustrate how uncertainties in the feedstock characterization propagate to the simulation results. The combination of a SEMK model and the feedstock reconstruction method based on maximization of the Shannon entropy allows to obtain accurate simulation results, provided that the specific density, the global PIONA weight or volume fractions, and the initial, 50% and final boiling point are known. Specifying less commercial indices results in a decrease of the agreement between simulated and experimentally obtained product yields. The developed methodology can be extended in a straight forward way to any heavy feedstock.

NOMENCLATURE

List of symbols

BP	Boiling point	K
C_j	Concentration of species j	mol m^{-3}
CIP	Coil Inlet Pressure	MPa
CIT	Coil Inlet Temperature	K
COP	Coil Outlet Pressure	MPa
COT	Coil Outlet Temperature	K
c_p	Heat capacity	$\text{kJ mol}^{-1} \text{K}^{-1}$
d	Specific density	kg m^{-3}
d_t	Reactor diameter	m
f	Fanning friction factor	-
F_j	Molar flow rate of component j	mol s^{-1}
f_j	Value of constraint j	variable
$f_{i,j}$	Stoichiometric coefficient of molecule i for constraint j	variable
FBP	Final Boiling Point	K
$\Delta_f H^0$	Standard enthalpy of formation	kJ mol^{-1}
IBP	Initial Boiling Point	K
k	Reaction rate coefficient	s^{-1} or $\text{mol m}^{-3} \text{s}^{-1}$
\tilde{k}	Single event reaction rate coefficient	s^{-1} or $\text{mol m}^{-3} \text{s}^{-1}$
M	Molar mass	kg mol^{-1}
N	Number of molecules in the library	-
n_e	Number of single events	-
p_t	Pressure	MPa
q	Heat flux	$\text{kJ m}^{-2} \text{s}^{-1}$
r_b	Radius of the bend	m
r_V	Reaction rate	$\text{mol m}^{-3} \text{s}^{-1}$
R_V	Net production rate	$\text{mol m}^{-3} \text{s}^{-1}$
S	Shannon entropy	-
T	Temperature	K
w_k	Weighing factor	-
v	Velocity	m s^{-1}
x_i	Mole fraction of component i	-
z	Axial position	m

Greek symbols

α	Conversion factor	-
λ	Lagrange multiplier	-
κ	Lagrange multiplier	-
ν	Lagrange multiplier	-
$\nu_{k,j}$	Stoichiometric coefficient of the library molecule j	-
ξ	Objective function	-
π	Probability	-

ρ	Density of the process gas mixture	kg m^{-3}
ζ	Nekrasov factor for bends	-
φ	Coefficient	-
ξ	Coefficient	-
ω	Circumference	m
Ω	Cross sectional surface area	m^2

Subscript

i	Molecule i
0	Initial

Superscript

lump Lumped

INTRODUCTION

Steam cracking of hydrocarbons is one of the most important processes of the petrochemical industry. In this process hydrocarbons are cracked into commercially more important products such as light olefins and aromatics. Feedstocks ranging from light alkanes such as ethane and propane up to complex mixtures such as naphthas and gas oils are converted at temperatures ranging from 900-1200 K in tubular reactors suspended in large gas-fired furnaces (Froment, 1992). The last four decades much effort has been devoted to the development of detailed mechanistic computer models for use in furnace design and predicting product yields from various types of feedstock over a broad range of cracking conditions. Mathematical modeling has the important advantage that once the model is developed, results can be easily gathered and computer simulations take only a limited time (Dente and Ranzi, 1979). For an accurate description of chemical kinetics applicable over a wide range of process conditions and feedstocks, a detailed microkinetic model is required (Froment, 1992). These microkinetic models (Baltanas and Froment, 1985; Chinnick *et al.*, 1988; Hillewaert *et al.*, 1988; Chevalier *et al.*, 1990; DiMaio and Lignola, 1992; Broadbelt *et al.*, 1994; Ranzi *et al.*, 1995; Prickett and Mavrovouniotis, 1997; Warth *et al.*, 2000; Guillaume *et al.*, 2003; Matheu *et al.*, 2003; Buda *et al.*, 2006; Lozano-Blanco *et al.*, 2006) capture the essential chemistry while a manageable size of the reaction network is maintained.

Due to the depleting reserves of sweet crude oils and the decreasing demand of heavy fractions as fuel, there is an increasing tendency to use low cost fuels such as heavy gas oils, vacuum gas oils (VGO) or gas condensates as feedstocks in steam cracking (Singh *et al.*, 2005). The main challenge from a modeling point of view is that the cracking behavior of the heavy cuts differs significantly

from that of lighter fractions, and developing an accurate kinetic model for these heavy fractions is not straight forward. The difference in cracking behavior between light and heavy fractions is obviously related to typical differences in chemical constituents, *e.g.* heavy fractions contain significant amounts of di-, tri- and poly-aromatic compounds that are not present in light fractions. The most advanced models allow modeling not only of steam cracking of light feedstocks such as naphthas but also of the cracking behavior of heavier feedstocks, such as gas oils, gas condensates and vacuum gas oils (VGO's). In this paper the general structure and assumptions made to develop a single event microkinetic (SEMK) model for both light and heavy feedstocks are addressed. Special attention will be given to the selection of species considered in the SEMK model and the criteria used for lumping of components will be discussed.

Although microkinetic models give accurate simulation results for various industrial processes, *e.g.* hydrocracking (Martens *et al.*, 2001), catalytic cracking (Feng *et al.*, 1993), a major problem associated with their use is that they require a detailed molecular feedstock composition. For the lightest feeds (gas, liquefied gas, naphtha), this molecular characterization can be obtained by gas chromatography. For heavy fractions, containing molecules with more than 25 carbon atoms, no analytical technique is yet powerful enough to detect and quantify the thousands of different compounds that compose an oil fraction and only average structural information can be obtained (Merdrignac and Espinat, 2007). Consequently, the use of detailed microkinetic models requires numerical "reconstruction" of a detailed molecular composition of a complex hydrocarbon mixture from partial analytical data, *i.e.* the so-called commercial indices (*e.g.* the average molecular weight, specific density...), by formulating model hypotheses and using expert knowledge (Hudebine and Verstraete, 2004). Most of the methods used to reconstruct the molecular composition of complex mixtures implement a two-step algorithm (Liguras and Allen, 1989a,b; Neurock *et al.*, 1994; Quann and Jaffe, 1996; Dente *et al.*, 2001; Joo *et al.*, 2001; Hudebine and Verstraete, 2004; Van Geem *et al.*, 2007). In the first step a large set of representative molecules, the so-called library of molecules, is created. In the second step of the algorithm, the mole fractions of the molecules contained in the library are adjusted in such a way that the mixture of library molecules has the same global characteristics as those imposed by the analytically determined commercial indices. A judicious selection of the library of molecules is therefore of the utmost importance for these approaches. For naphtha fractions and gas oils the library can be determined experimentally based on the detailed molecular composition of a huge number of reference fractions with widely varying characteristics (Neurock *et al.*, 1994; Van Geem *et al.*, 2007). For very

heavy fractions, such as heavy asphaltene feedstocks, experimental determination of the library becomes very difficult and stochastic reconstruction has been proven able to yield mixtures that closely mimic the original properties of heavy feedstocks (Trauth, 1993; Trauth *et al.*, 1994). In this work the experimental route is chosen to define the library of molecules for gas condensates because the detailed molecular composition of these fractions can be determined quantitatively using a combination of gas chromatography (GC) and gas chromatography-mass spectrometry (GC-MS). Several methods can be used for adjusting the mole fractions: optimizing an objective function such as the information entropy (Hudebine and Verstraete, 2004; Van Geem *et al.*, 2007) or the enthalpy of the mixture (Liguras and Allen, 1989), using a neural network model (Joo *et al.*, 2000) or using correlations obtained via multiple regression (Dente *et al.*, 2001). A disadvantage of the last two methods is that they are rather limited in their application range and not easily extendable. Moreover, these methods have traditionally a rather limited flexibility as they are based on a limited number of well defined commercial indices. Hence, these methods can only be applied if all the required indices are available and additional analytical information about the mixture cannot improve the predicted composition because this extra information cannot be taken into account. Methods that optimize a specific objective function do not show these disadvantages. These methods seem also more suited for reconstruction of the heavy hydrocarbon fractions. This is because if a neural network method or a multiple regression method is selected an extensive database of reference mixtures is needed to train the model. In particular, the method that optimizes the information entropy seems potentially very powerful. In what follows the advantages and disadvantages of this method for feedstock reconstruction are discussed. The approach is validated *via* comparison with analytical results obtained for a number of gas condensates. The combination of the feedstock reconstruction method with the SEMK model for steam cracking allows to evaluate the required molecular detail of the kinetic model and of the feedstock characterization/reconstruction model to obtain accurate model predictions.

1 SINGLE EVENT MICROKINETIC MODEL FOR STEAM CRACKING

The SEMK model contains two parts: 1. the single event reaction network and 2. the reactor model equations and the solver of the resulting set of differential algebraic equations. A schematic overview of the SEMK model for steam cracking is presented in Figure 1. In this section the structure of the reaction network for steam cracking and the species that are considered are summarized.

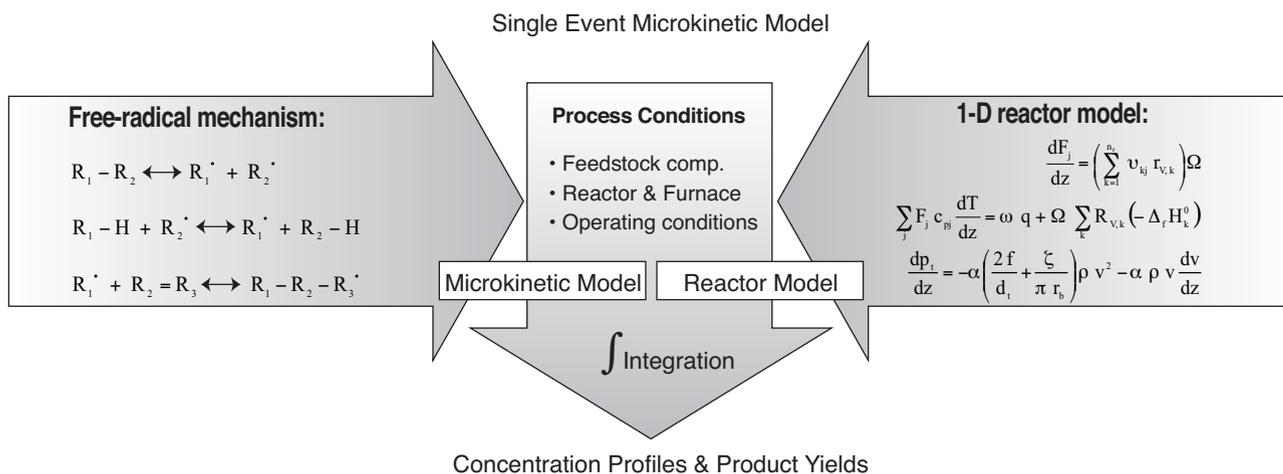


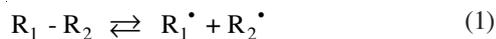
Figure 1

Structure of the single event microkinetic model for steam cracking of hydrocarbons.

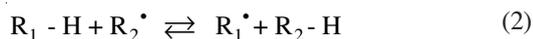
1.1 Structure of the Single Event Reaction Network

Steam cracking of hydrocarbons proceeds through a free radical mechanism and three important reaction families can be distinguished (Rice *et al.*, 1931, 1934):

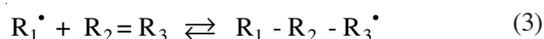
- Carbon-carbon and carbon-hydrogen bond scissions of molecules and the reverse radical-radical recombinations:



- Hydrogen abstraction reactions, both intra- and intermolecular:

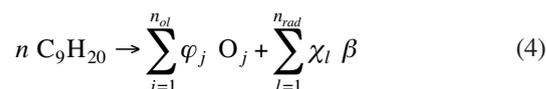


- Radical addition to olefins and the reverse β scission of radicals, both intra- and intermolecular:



Developing a detailed reaction network is a major challenge. On the one hand the size of the reaction network can become huge as the number of reactions and species increases exponentially with the average carbon number of the feedstock (Broadbelt *et al.*, 1994). On the other hand, developing these reaction networks implies that both the thermo-chemistry and kinetic parameters are known. Fortunately the μ radical hypothesis holds for steam cracking of hydrocarbons (Ranzi *et al.*, 1983) which allows drastically reducing the reaction scheme. The μ radical hypothesis implies that for the so-called μ radicals the monomolecular reactions are much faster than the bimolecular ones, and hence, that the latter can be neglected without loss of

accuracy. These μ radicals are for example long chain aliphatic radicals containing more than 5 carbon atoms. This allows distinguishing between two networks: the monomolecular μ network and the β network, that contains both uni- and bimolecular reactions. The kinetics for the former can be described by analytical expressions based on the pseudo steady state assumption (PSSA) for the radical reaction intermediates (Hillewaert *et al.*, 1988). For example the complete reaction scheme shown in Figure 2 starting from the C-C scission of n-nonane can be replaced by one equivalent reaction (Ranzi *et al.*, 2001) of the following form:



Note that the coefficients φ_j and χ_l in the equivalent reaction presented by Equation 4 are not the stoichiometric coefficients (Ranzi *et al.*, 1983). These coefficients are weakly temperature dependent as they are function of the reaction rate coefficients of the elementary reaction steps in the considered reaction scheme. The assumptions made for constructing the reaction network have been verified using a new rate based network generator called RMG (Van Geem *et al.*, 2006). Under typical steam cracking conditions (T : 600–850°C, p_i : 0.15–0.25 MPa) the μ radical hypothesis is indeed valid. Also, the error introduced by applying the quasi steady assumption state for the group of μ radicals is negligible. For each elementary reaction considered in the β network the reaction rate coefficient is calculated based on the Arrhenius expression, using the corresponding activation energy and pre-exponential factor.

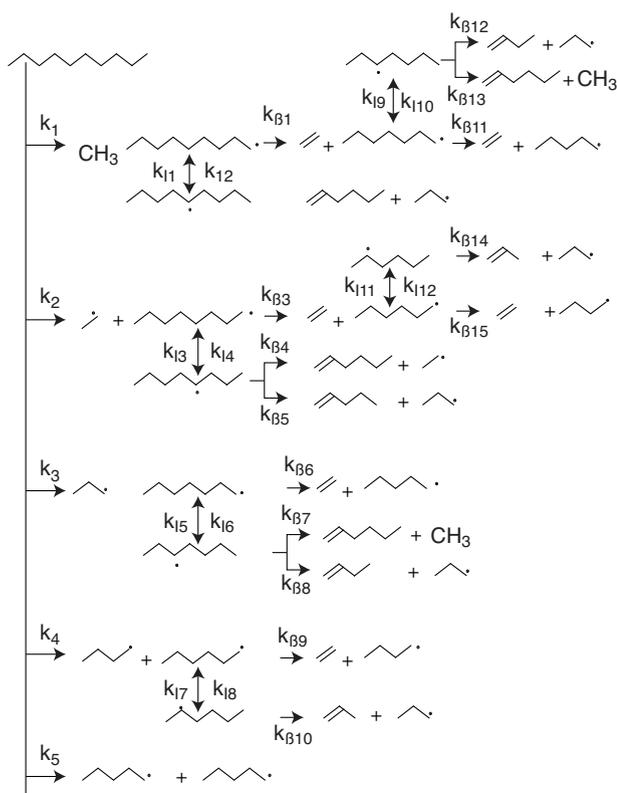


Figure 2

Reaction scheme for *n*-nonane cracking starting from a C-C scission reaction.

For a light naphtha feedstock the number of elementary reactions in the complete network amounts to 119 000. Introducing the single event concept allows calculating all these reaction rate coefficients in a straightforward way. The reaction rate coefficient of an elementary reaction step, k , is then equal to the product of a single event rate coefficient, \tilde{k} , and the number of single events, n_e (Willems and Froment, 1988; Baltanas *et al.*, 1989). The single event rate coefficient \tilde{k} , depends on the reaction family and on the nature of the reactant and product involved in the elementary step and is independent of the number of carbon atoms. These assumptions drastically reduce the number of elementary step rate coefficients. Moreover, using a group additive method and incorporating thermodynamic consistency allows further reducing the number of parameters in the model. In the present model the group additive method of Saeys *et al.* (2003, 2004, 2006) is used. Saeys' method is a consistent extension of Benson's group additivity concept (1976) to transition state theory and maximizes the benefits of using the reaction family concept and thermodynamic consistency.

1.2 Species Selection and Lumping Procedures

Microkinetic models require a detailed molecular feedstock composition. For the lightest feeds (gas, liquefied gas, naphtha), this molecular characterization can be obtained by gas chromatography. For heavy fractions, containing molecules with more than 25 carbon atoms, only average structural information can be obtained (Merdrignac and Espinat, 2007). Currently the analysis of the main petroleum fractions remains based on 1-dimensional gas chromatography with its known limitations. For example the peak capacity of the GC separation results in peak overlap starting from C9 carbon molecules. Hence, 1-dimensional gas chromatography does not allow obtaining the detailed analysis of heavy naphtha, kerosene (*i.e.* C8-C15 as a carbon atom range) and middle distillate samples (C15-C30) (Vendeuvre *et al.*, 2007). However, recently major advances were made in the field of detailed characterization of petroleum fractions, *e.g.* comprehensive GC \times GC (Phillips and Xu, 1995; Dalluge *et al.*, 2003) and Fourier transform ion cyclotron resonance mass spectrometry FT-ICR MS (Hughes *et al.*, 2002; Milluns *et al.*, 2006). These new and innovative techniques will allow a more detailed characterization of petroleum fractions in the coming years. This improved characterization methods will surely benefit the model predictions obtained with microkinetic models as these models are able to take full advantage of a more detailed characterization. In the mean time, the current state of the art microkinetic models have to be aware of these analytical shortcomings, but should be able to take maximal advantage of the improved characterization of petroleum fractions using innovative techniques. In the current microkinetic model these considerations were taken into account in the selection of the chemical species considered in the kinetic model. For example in Figure 3 the GC \times GC FID chromatogram of a gas condensate is given. The latter shows that these fractions contain thousands of components, but gathering this complete detailed composition in a reasonable time is very time consuming. First a GC \times GC MS analysis is necessary for the qualitative analysis. For the current gas condensate the first column was a 15 m long DB-5 (0.25 mm \times 0.25 μ m) and the second column was a 1 m long DB-1701 column (0.1 mm \times 0.1 μ m). A quadrupole MS is used for the qualitative analysis, while an FID is used for the quantitative analysis. Peak identification and integration took several days because thousands of different components needed to be identified. Figure 3 shows the different classes of compounds on the FID chromatogram identified using the GC \times GC MS chromatogram. This time consuming operation implies that in industrial practice the complete detailed composition of these fractions is not readily available yet. This shows that, although it is possible to construct very complex SEMK models, one has to be aware of certain

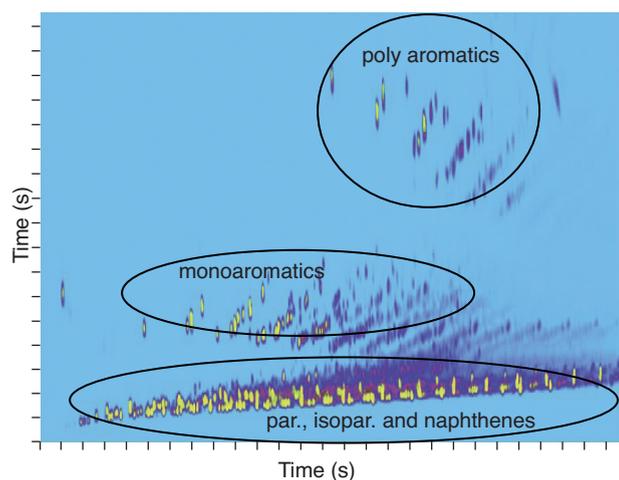


Figure 3

GC x GC FID chromatogram of a gas condensate feedstock used during one of the pilot plant experiments given in Table 5.

restrictions that do not allow to take full advantage of the level of molecular detail considered in the microkinetic model. The microkinetic model obviously contains all the important products and intermediates experimentally observed during the cracking of different hydrocarbon feedstocks. These products are mainly lower olefins and light aromatics. However, the main part of the molecular species considered in the reaction network are molecules that are also present in the feedstocks. These are for example n-paraffinic and iso-paraffinic compounds containing less than 34 carbon atoms.

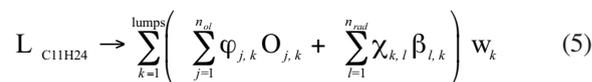
Nonetheless, even for single event microkinetic models, it is not only convenient but also necessary to adopt several simplifications and lumping procedures in order to avoid an excessive number of chemical species and reactions (Ranzi *et al.*, 2007). Increasing the number of molecules in the reaction scheme increases the number of differential equations that have to be integrated, and hence, in a rapid increase of the simulation time. Therefore, a compromise has to be made between computation efforts and prediction accuracy. Table 1 gives an overview of the SEMK components considered in the reaction network. With the current set of SEMK components in the microkinetic model the cracking behavior of the most important types of steam cracking feeds can be described. The maximum carbon number of the molecules is 33 and also a large number of heavy aromatic compounds are considered. The introduction of di-, tri-, poly- and naphthenoaromatic compounds is on the one hand necessary to be able to simulate VGO fractions. On the other hand these molecules form also an important part of the pyrolysis fuel oil

(PFO) fraction. The PFO fraction is the heavy fraction (b.p. > 473 K) formed during steam cracking of liquid feedstocks.

TABLE 1
Overview of the components considered in the SEMK model for steam cracking

Hydrocarbon Type	Lowest Carbon number	Highest Carbon number	Number of Components
hydrogen	-	-	1
n-paraffins	1	33	33
iso-paraffins	4	33	79
naphthenes	5	20	80
olefins	2	32	199
aromatics	6	24	86
TOTAL	0	33	478

Several SEMK components in Table 1 are so-called pseudo-components. For example all iso-paraffinic compounds with more than 11 carbon atoms are lumped into one single pseudo-component per carbon number. Consider for instance a mixture of iso-paraffinic structural isomers I_k . The equivalent reactions of the different pseudo-components are obtained by averaging the equivalent reactions of the structural isomers, taking all the elementary reactions of these isomers into account. The weights w_k are related to an appropriate statistical internal composition of the pseudo-component (Ranzi *et al.*, 2001). An interesting example of the combination of the μ -radical hypothesis and the pseudo-component simplifying rules is constituted by the pseudo-component “ L_{C11} ”, which is a lumped species grouping the different iso-paraffinic compounds with 11 carbon atoms. The final lumped reaction for the standard L_{C11} pseudo-component can be simply obtained by averaging the equivalent reactions of the considered structural isomers. The resulting lumped hydrogen abstraction reaction presents the following stoichiometry (Ranzi *et al.*, 2001):



with $O_{j,k}$ an olefin, $\varphi_{j,k}$ and $\chi_{l,k}$ stoichiometric coefficients, and $\beta_{l,k}$ a β radical. For the pseudo-component a continuity equation needs to be considered only. Note that the differences in the reaction rates for the disappearance of the different structural isomers *via* the equivalent reactions such as Equation 5 should remain as small as possible because only then it is allowed to replace the different structural isomers by a single pseudo-component (Wei and Kuo, 1969; Kuo and Wei, 1969). That is why one pseudo-component is defined per carbon number and per class of components. The composition of each pseudo-component consists then solely of structural isomers.

In principle, the weighing factors w_k should be determined from the detailed analytical composition of the feedstock (Ranzi *et al.*, 2001) implying that for each feedstock new pseudo-components would need to be defined. In practice, fixed weighing factors can be used without loss of accuracy. Based on the analysis of a large number of naphtha and kerosene fractions, Dente *et al.* (2001) found that the distribution of structural isomers is largely independent of the origin of the feedstock (*see Table 2*). Of course, such a procedure implies that the range of applications of the model is restricted. For instance, the decomposition of a given particular structural isomer that has been taken up in a lump cannot

TABLE 2

Relative amounts of branched isomers of C_8H_{18} in naphtha cuts from different origins (Dente *et al.*, 2007)

Origin Isomers	Ponca (wt.%)	Occid. (wt.%)	Texas (wt.%)	Internal weights (wt.%)
2-methylheptane	46.3	36.9	42.1	45.8
3-methylheptane	15.4	28.5	23.4	22.9
4-methylheptane	10.3	10.2	9.3	11.5
2,3-dimethylhexane	3.6	5.4	6.3	3.4
2,4-dimethylhexane	3.1	5.5	4.2	3.4
2,5-dimethylhexane	3.1	5.7	4.0	3.4
3,4-dimethylhexane	6.7	2.6	3.7	3.4
2,3,4-trimethylpentane	0.3-	-	1.1	1.2
3-ethylhexane	4.6	3.5	3.1	3.8
2-methyl-3-ethylpentane	3.1	-	1.5	1.2

be studied using the lumped model. It should however be mentioned that the current computer code allows relaxing the lumping hypotheses and enlarging the kinetic scheme to explicitly include that component, enabling to study the cracking behavior of a specific structural isomer. This lumping flexibility is one of the main advantages of our microkinetic model and allows to easily adapt the model to future improvements in characterization of petroleum cuts.

1.3 Reactor Modeling

Steam cracking is a non-isothermal, non-adiabatic and non-isobaric process. Hence, the 1 dimensional model equations consist of the transport equations for mass, momentum and energy. The steady state continuity equation for a SEMK component j in the process gas mixture over an infinitesimal volume element with cross sectional surface area Ω , circumference ω and length dz is:

$$\frac{dF_j}{dz} = \left(\sum_{k=1}^{n_r} \nu_{kj} r_{V,k} \right) \Omega \quad (6)$$

with F_j the molar flow rate of SEMK component j , $r_{V,k}$ the reaction rate of reaction k , and ν_{kj} the stoichiometric coefficient of SEMK component j . The energy equation is given by:

$$\sum_j F_j c_{pj} \frac{dT}{dz} = \omega q + \Omega \sum_k R_{V,k} (-\Delta_f H_k^0) \quad (7)$$

with q the heat flux to the process gas, c_{pj} the heat capacity of component j at temperature T , $\Delta_f H_k$ the standard enthalpy of formation of species k , $R_{V,k}$ the net production rate for species k . The momentum equation accounting for friction and changes in momentum is given by:

$$\frac{dp_t}{dz} = -\alpha \left(\frac{2f}{d_t} + \frac{\xi}{\pi r_b} \right) \rho v^2 - \alpha \rho v \frac{dv}{dz} \quad (8)$$

with p_t the total pressure, α a conversion factor, f the Fanning friction factor, ρ the density of the gas mixture, r_b the radius of the bend, d_t the diameter and v the velocity. The boundary conditions of the model equations are:

$$C_j = C_{j0} \quad T = T_0 \quad p = p_0 \quad (z = 0)$$

Integration of the last two model equations is only required when the temperature and/or pressure profile are not imposed. Based on the reactions, and reaction rate coefficients, the production rate $r_{V,k}$ of each SEMK component j by reaction k , can be expressed as a function of the concentration of the involved species. The resulting set of continuity equations forms a set of stiff non-linear first order differential equations. The stiffness of the set of equations originates from the large difference (several orders of magnitude) of the eigen values related to the molecular species on the one hand and the radical species on the other hand. The stiff solver DASSL (Li and Petzold, 1999) is used to integrate the stiff set of differential equations

The developed fundamental simulation model is validated using pilot plant experiments. The latter allows measurement of the kinetics of the cracking reactions (Van Geem *et al.*, 2005) and of the coke deposition in both the radiant coil (Reyniers and Froment, 1995) and the transfer line exchanger [TLE] (Dhuyvetter *et al.*, 2001). The pilot plant set-up consists of 3 parts: a feed section, the furnace containing the suspended reactor coil and the analysis section. The tubular reactor used in this set of experiments has a length of 12.4 m and has an internal diameter of 9.0 mm. These dimensions are chosen to achieve turbulent flow conditions in the coil. The temperature and pressure profile along the reactor can be measured and regulated. Excellent agreement is obtained between the simulated and experimental product yields for a wide range of feedstocks and experimental conditions. The simulation results (Sim1 in Table 3) obtained for a gas condensate experiment illustrate that not only for the ethylene

and propylene yield a good agreement is observed, also for other products such as the C4 fraction and the BTX fraction the agreement is satisfactory. Note that these simulations are performed using the analytically obtained gas condensate composition.

TABLE 3

Simulated and experimentally determined product yields obtained for Condensate 1 in the LPT pilot plant reactor. Conditions: CIT = 823 K; COT = 1093 K; CIP = 0.22 MPa; COP = 0.17 MPa; $F = 1.1 \times 10^{-3} \text{ kg s}^{-1}$; $\delta = 0.5 \text{ kg /kg}$.

Sim1: Simulation using analytically determined composition;
Sim2: Simulation using reconstructed feed composition based on the specific density and 7 ASTM D86 boiling points (see Table 4),
Sim3: Simulation using reconstructed feed composition using all the commercial indices mentioned in Table 4

Product Yields (wt.%)	Sim1	Sim2	Sim3	Experimental
CH ₄	12.3	11.9	12.6	13.2
C ₂ H ₂	0.3	0.3	0.3	0.3
C ₂ H ₄	25.2	25.1	25.6	25.4
C ₃ H ₆	16.1	15.4	15.7	16.0
C ₄ H ₆	5.5	5.8	5.8	5.0
1-C ₄ H ₈	2.0	2.0	1.9	1.7
2-C ₄ H ₈	0.7	0.7	0.7	0.6
iso-C ₄ H ₈	3.1	2.8	2.6	3.2
CPD	1.7	2.2	2.1	1.4
C ₆ H ₆	5.6	5.9	5.6	5.3
C ₇ H ₈	2.8	3.2	3.1	3.0
P/E-ratio (kg/kg)	0.64	0.62	0.62	0.63

2 FEEDSTOCK RECONSTRUCTION BY SHANNON ENTROPY MAXIMIZATION

2.1 Structure of the Reconstruction Method

Shannon's Entropy theory is widely applied in all sorts of engineering fields, ranging from quantum chemistry over civil engineering to hydrodynamics. Shannon's entropy is defined as:

$$S(\pi_i) = -\sum_{i=1}^N \pi_i \cdot \ln \pi_i \quad \text{with} \quad \sum_{i=1}^N \pi_i = 1 \quad (9)$$

in which S represents Shannon's entropy and π_i is the probability of a certain state. Shannon's entropy is a measure of the homogeneity of a probability distribution. The principle of maximum Shannon entropy states that if only partial information concerning possible outcomes is available, the probabilities are to be chosen so as to maximize the uncertainty on the missing information (Shannon, 1948). This implies that the entropy has to be maximized subject to constraints representing the available information. By applying the principle of maximum entropy, the most probable distribution complying

to the given constraints can be obtained. Applying this theory to the composition of petroleum fractions implies that the probabilities in Equation 9 are replaced by the mole fractions x_i of the library molecule i . This method, as applied to feedstock reconstruction developed at IFP (Hudebine *et al.*, 2002; Hudebine, 2003; Hudebine and Verstraete, 2004; Van Geem *et al.*, 2007), uses the analytically determined indices as boundary conditions. We consider the global solution of the following nonconvex problem:

$$\text{MAX} \quad S(x_i) = -\sum_{i=1}^N x_i \cdot \ln x_i \quad (10)$$

$$\text{Subject to:} \quad f_j - \sum_{i=1}^N f_{i,j} \cdot x_i = 0 \quad j = 1, \dots, J$$

with f_j the value of constraint j , *i.e.* a commercial index, $f_{i,j}$ the coefficient of molecule i for constraint j , N the number of molecules in the library and J the number of exact linear constraints. For example, if the average molecular weight $M_{w,av}$ of the mixture is known, the following exact constraint is imposed:

$$M_{w,av} - \sum_{i=1}^N x_i \cdot M_{w,i} = 0 \quad (11)$$

In Figure 4, a schematic overview of the method developed by IFP is given. First, a molecular library is selected. Next, the mole fractions of the molecules contained in the library are adjusted in order to obtain a mixture with the

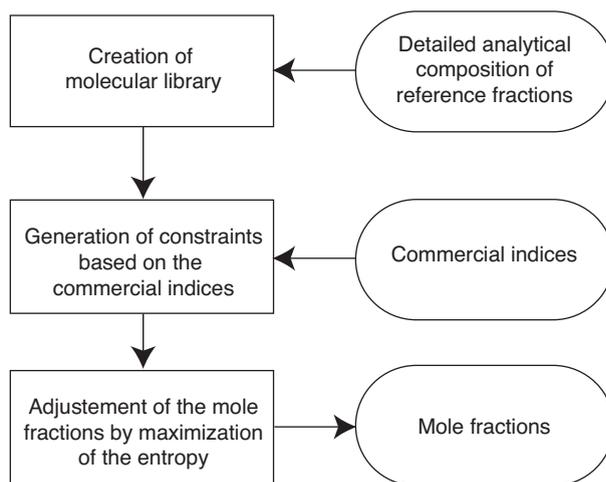


Figure 4

Overview of the feedstock reconstruction method (Van Geem *et al.*, 2007).

desired characteristics, *i.e.* the characteristics imposed by the commercial indices of the feedstock, *e.g.* the average molecular weight or the specific density. Evidently, the mole fractions should also meet the maximization criterion defined in Equation 10. Finding the global optimum of nonconvex problems such as the one defined in Equation 10 is not straightforward and can be very time consuming (Floudas and Pardalos, 1996). Generally, maximum entropy problems are solved using the Lagrange multiplier method (Guiasu and Shenitzer, 1985; Kapur and Kesavan, 1992). The Lagrange multiplier method is used to find the optima of a function $f(x)$ under the constraints $g_j(x)$ that equal zero. A new function which incorporates the function $f(x)$ and all its constraints is introduced:

$$\xi(x) = f(x) + \sum_j \lambda_j \cdot g_j(x) \quad (12)$$

with λ_j a constant variable called the Lagrange multiplier. The optimization problem is then reduced to finding the optima of $\xi(x)$ in x_i and λ_j (Guiasu and Shenitzer, 1985; Kapur and Kesavan, 1992). Solving the optimization problem can be drastically simplified when all the constraints are linear in the variables x_i (Guiasu and Shenitzer, 1985; Kapur and Kesavan, 1992). In this case, the optimization function can be transformed from a nonlinear equation in the N mole fractions x_i into a nonlinear equation in J parameters λ_j . Because N is in the order of 10^2 - 10^6 and J is maximum 13, the gain in the optimization level can be considerable, and the solution of the optimization problem requires only a limited time. However, the linearity requirement imposes an important restriction on the commercial indices and affects strongly which commercial indices can and cannot be used. Information such as the Bureau of Mines Correlation Index, the Watson characterization factor or the vapor pressure cannot be used as commercial indices. The following commercial indices lead to constraints which are linear in the mole fractions: the average molecular weight, the specific gravity, the H/C-ratio, the PIONA weight fractions, boiling points from the true boiling point (TBP) curve, NMR spectra and a detailed distribution of the hydrocarbons per carbon atom. All these commercial indices can be obtained via standardized methods (ASTM methods).

Applying the Lagrange multipliers method leads to the following optimization criterion (Van Geem *et al.*, 2007):

$$\xi(x) = - \sum_{i=1}^N x_i \cdot \ln x_i + \kappa \cdot \left(1 - \sum_{i=1}^N x_i \right) + \sum_{j=1}^J \lambda_j \cdot \left(f_1 - \sum_{i=1}^N x_i \cdot f_{i,j} \right) \quad (13)$$

The optimum of this function is found by setting the derivatives of this equation with respect to the mole fractions x_i equal to 0. This allows finding the following expressions for the mole fractions (Van Geem *et al.*, 2007):

$$x_i = \frac{\exp \left[- \sum_{j=1}^J \lambda_j f_{i,j} \right]}{\sum_{i=1}^N \exp \left[- \sum_{j=1}^J \lambda_j f_{i,j} \right]} \quad \forall i \in N \quad (14)$$

Substitution of Equation 6 in Equation 9 leads to the following expression for the entropy criterion:

$$E(\lambda) = \ln \left(\sum_{i=1}^N \exp \left[- \sum_{j=1}^J \lambda_j f_{i,j} \right] \right) + \sum_{j=1}^J \lambda_j \cdot f_j \quad (15)$$

The values of λ_j at the optimum permit to calculate the most probable mole fractions of the molecules. In this work, the Rosenbrock method (Rosenbrock, 1960) is used to locate the optimum of the function given in Equation 15 (Van Geem *et al.*, 2007). The Rosenbrock method is a 0th order search algorithm that approximates a gradient search thus combining advantages of 0th order and 1st order strategies (Rosenbrock, 1960). It has been reported that this simple approach is more stable than many sophisticated algorithms and it requires much less calculations of the objective function than higher order strategies (Schwefel, 1981). As only linear constraints are considered, the optimization function can be transformed from a non-linear equation in the N mole fractions x_i into a non-linear equation in J parameters λ_j . The simulation time for obtaining a detailed composition is less 0.1 s on an Intel Pentium III processor of 1.0 GHz (Van Geem *et al.*, 2007).

Note that to use the previously defined method requires the knowledge of the physical properties of the molecules in the library. For some commercial indices such as the average molecular weight, the value for a library molecule can be easily obtained. The density at 25°C of every library molecule is calculated by the group contribution method of Fedors (1974). To calculate the boiling point of a library molecule, the group interaction contribution (GIC) method from Marrero *et al.* (1999) is used. This method is specifically developed for hydrocarbons and is known for its high statistical accuracy.

2.2 Library Construction

As stated previously there are several approaches available for constructing a molecular library (Hudebine, 2003) ranging from experimental methods (Van Geem *et al.*, 2007), over group contribution methods (Hudebine *et al.*, 2002) and even stochastic methods (Neurock *et al.*, 1992; Hudebine and Verstraete, 2004). In this work, the experimental route is chosen to define the library of molecules for gas condensates and gas oils because the detailed molecular composition of these fractions can be determined quantitatively in a reasonable time, using a combination of gas chromatography (GC) and gas chromatography-mass spectrometry (GC-MS). The

molecular library is constructed in the same way as described earlier for naphtha fractions (Van Geem *et al.*, 2007). The detailed molecular compositions of a large number of reference gas condensate allows to select the most important components typically present in this petroleum fraction as possible library molecules. For each type of feedstock a new library has to be selected as every type of petroleum fraction has its own molecular characteristics, independent of its origin.

A list of the selected library molecules for gas condensates is taken up in Appendix. These library molecules are selected based on the detailed molecular composition of a large number of reference gas condensates with widely varying characteristics, *e.g.* their average molecular weight varies between 90 and 125 g mol⁻¹. Only those molecules with a weight fraction higher than 0.3 wt.% in at least one of the reference mixtures are selected. This leads to a total of 107 library molecules that cover on average more than 90 wt.% of the entire feedstock composition. The library includes 17 n-paraffins, 14 aromatics, 18 naphthenes and 57 iso-paraffins. The GC × GC TOF-MS chromatogram of one of the reference feedstocks in Figure 3 shows that a typical gas condensate contains much more components than those selected in the molecular library. If, instead of applying the proposed selection criteria, all the components observed in one of the GC × GC TOF-MS chromatograms of the reference gas condensates would be included in the molecular library, the latter would contain thousands of different library molecules. However, this extension involves mainly structural isomers of one of the library molecules mentioned in Appendix. As was illustrated for naphtha fractions (Van Geem *et al.*, 2007), including more components in the feedstock library does not necessarily lead to more accurate simulation results.

2.3 Feedstock Reconstruction

The proposed method for feedstock reconstruction is able to generate in a minimum of time a detailed molecular composition corresponding to the commercial indices. It should be stressed that the molecular composition with maximum Shannon entropy will only rarely coincide with the analytically determined molecular composition. It can thus be expected that including more information about the mixture, *i.e.* increasing the number of commercial indices considered, leads to an improved agreement between the reconstructed composition and the analytically determined one. For naphtha fractions (Van Geem *et al.*, 2007) this method is able to generate a molecular composition that corresponds reasonably well with the analytically determined one.

The detailed molecular composition of the gas condensate fractions is determined using a combination of gas chromatography (GC) and gas chromatography-mass spectrometry (GC-MS). The calibration factors used for the quantitative analysis are those proposed by Dietz (1967). Calibration

factors for components not explicitly mentioned in the article of Dietz (1967) are calculated using the group contribution method of Dierickx *et al.* (1986). The molecular composition of the gas condensate fraction serves as a basis to determine the detailed PINA (paraffin, isoparaffin, naphthenes, aromatics) weight fractions used for validation purposes. In Figure 5 an overview of the agreement between the simulated and analytically obtained detailed PINA weight fractions per carbon number is shown for Condensate 1. The available commercial indices of Condensate 1 are given in Table 4. Two different simulations are performed. In the first case only the specific density and the PINA weight fractions are considered. In the second simulation all the commercial indices specified in Table 4 are used to reconstruct the feedstock. There are two main advantages of our method over neural network methods or multiple regression methods. The first one is that even for a limited number of available commercial indices the method is still able to reconstruct a composition that meets the conditions set by the boundary conditions. As will be shown further, limiting the number of considered commercial indices has some effects on the yield predictions obtained with the reconstructed composition. The second one is that additional analytical information on the feedstock can easily be taken into account and be used to increase the accuracy of the model predictions. In contrast, neural networks (Joo *et al.*, 2001) or methods based on correlations (Riazi, 2005) have rather limited flexibility as they are based on a limited number of well defined commercial indices and can only be applied if all the indices required by the reconstruction model are available. Moreover, additional analytical information on the feedstock cannot improve the model predictions as this additional information cannot be taken into account.

TABLE 4
Commercial indices of Condensate 1

Commercial Index	Range
Specific density, 15/4°	0.71
PIONA-analysis (wt.%)	
Paraffins	32.3
Iso-paraffins	40.2
Olefins	0.1
Naphthenes	17.8
Aromatics	9.6
Average Molecular Mass (g mol ⁻¹)	100.6
ASTM boiling points (K)	
IBP	302.3
10% BP	324.6
30% BP	343.6
50% BP	374.2
70% BP	414.8
90% BP	496.7
FBP	548.4

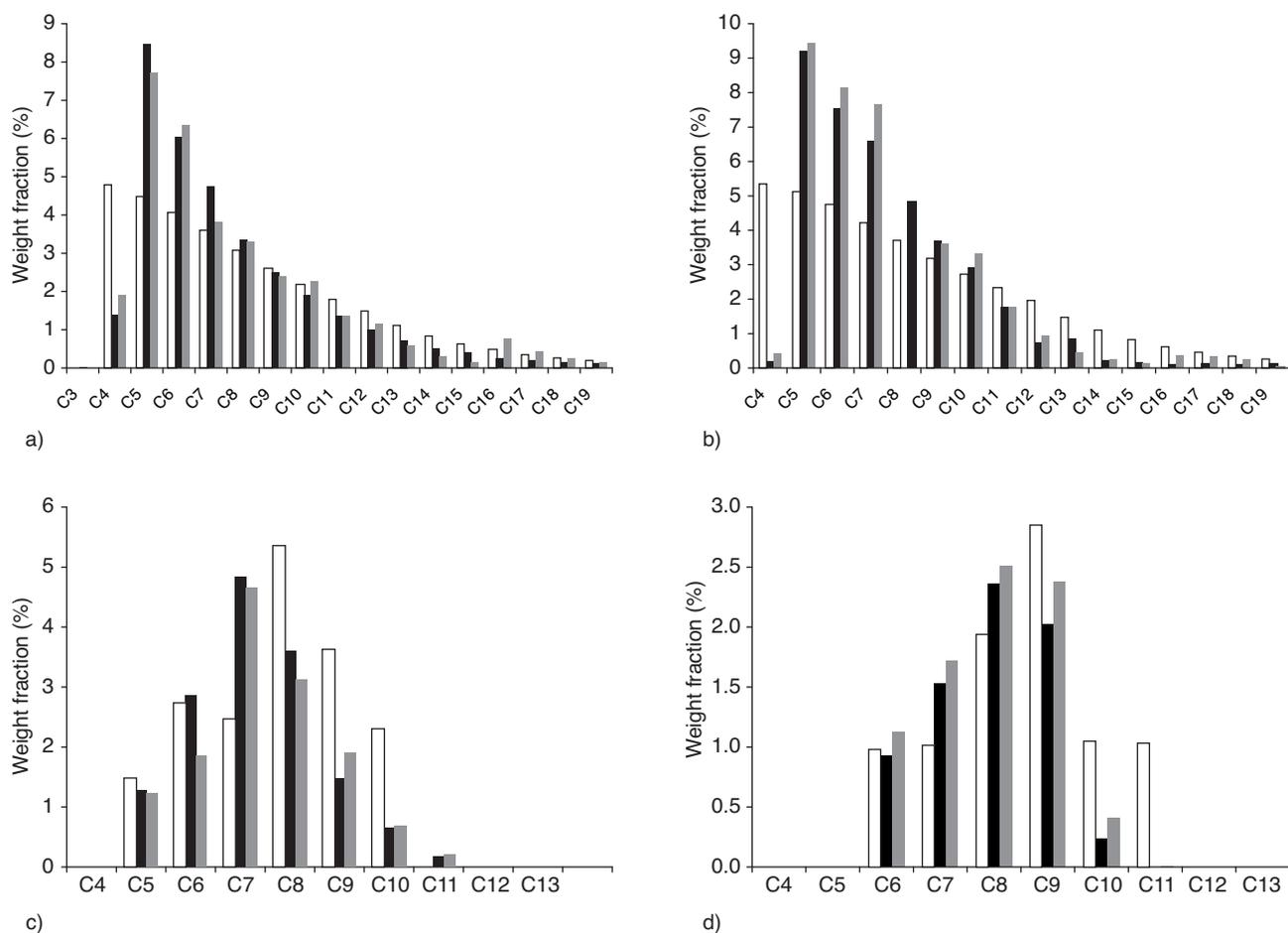


Figure 5

Comparison of the simulated and analytically determined detailed PIONA weight fractions. (a) Weight fractions of paraffinic components, (b) Weight fractions of iso-paraffinic components, (c) Weight fractions of naphthenic components, (d) Weight fractions of aromatic components. \square Simulations using the commercial indices mentioned in Table 4, \blacksquare Experimental results, \blacksquare Simulated with commercial indices from Table 4 + 7 ASTM D86 boiling points (IBP = 302.3 K; 10% BP = 324.6 K; 30% BP = 70.6 K; 50% BP = 374.2 K; 70% BP = 414.8 K; 90% BP = 537.0 K; FBP = 593.0 K).

From Figure 5, it is clear that for the first simulation, considering only the specific density and the global PINA weight fractions, the agreement between the simulated and analytically determined PINA weight fractions per *C* number is rather poor. On the other hand, including the ASTM-D86 boiling points in the second simulation leads to a good agreement. In the first simulation, no information related to the molecular carbon number distribution is specified. In the second simulation, the extra boundary conditions set by the 9 points of the ASTM D86 boiling point curve provide a good indication of the carbon number distribution. Obviously, specifying these extra commercial indices has also an effect on the accuracy of the simulation results obtained with the SEMK model for steam cracking, as will be shown in the next section.

Note that even upon using all the commercial indices mentioned in Table 4 the agreement between simulated and analytical composition is not perfect. The differences illustrated in Figure 5 can be expected for several reasons. Firstly, as stated previously, the method is statistical, implying that it selects a single composition with maximum Shannon entropy out of a number of possible compositions that meet all the boundary conditions. Hence, some differences are inherent to the statistical nature of method. Secondly, deviations from the analytically determined compositions can also be caused by the inaccuracy of some of the commercial indices. For example, the reconstruction method applied in the present work uses boiling points from the true boiling point curve. However, the latter is only rarely known as in practice ASTM D86, ASTM

D1160 or ASTM D2887 boiling point curves are determined. Therefore, a transformation of ASTM boiling point curves into the true boiling point curve is required and the empirical methods used to that purpose, *e.g.* the Riazi-Daubert method (Riazi, 2005), undeniable introduces some inaccuracies. As shown by Van Geem *et al.* (2006), this problem cannot be overcome by imposing uncertainties on the boiling points as these imposed uncertainties introduce at the same the possibility to make the mole fraction distribution more uniform.

3 VALIDATION USING PILOT PLANT EXPERIMENTS

The combination of the feedstock reconstruction program and the single event microkinetic model is our ultimate goal. Therefore a set of pilot plant experiments are used to test the capabilities to simulate the cracking behavior of gas condensates. In a first step, the molecular composition of these condensates is reconstructed. For each feedstock the following set of commercial indices is known: the global PINA weight fractions, the specific density, and 7 ASTM D86 boiling points (IBP, 10% BP, 30% BP, 50% BP, 70% BP, 90% BP, FBP). All this information is used as input and allows to determine the weight fractions of each of the 107 molecules considered in the molecular library. In a second step, the reconstructed composition is translated to a SEMK-input composition. As discussed previously, in the reaction network a certain degree of lumping is applied. For example all iso-paraffinic compounds containing more than 11 carbon atoms are lumped into one single pseudo-component per carbon number. The gas condensate is then represented by 62 SEMK components, with some of them so-called pseudo-components. In total, 15 pilot plant experiments are simulated for 8 different gas condensates, in which the experimental conditions vary over a range of the experimental conditions presented in Table 5. The flow rate of the hydrocarbon feedstock is varied between 8.3×10^{-4} and 1.3×10^{-3} kg s⁻¹, while the coil outlet temperature varies from 1073 K to 1113 K. The dilution varies from 0.3 kg_{steam}/kg_{naphtha} to 1.0 kg_{steam}/kg_{butane}. The coil outlet pressure varies from 0.16 MPa to 0.18 MPa. These conditions correspond with a P/E-range (propylene to ethylene ratio) from 0.50 to 0.65 kg/kg.

The parity plots (*see Fig. 6*) for the yields of the main cracking products methane, ethylene, propylene and benzene show that the combination of the feedstock module with the 1-dimensional reactor model is able to accurately simulate the product yields over a wide range of process conditions. A crucial element for success is the minimal number of commercial indices specified. If for example only the boiling point curve and the specific density are used for the feed reconstruction, the accuracy of the simulated product yields can become rather poor. As illustrated in Table 3 (Sim1), the SEMK-model accurately predicts the yields obtained for a

pilot plant experiment of Condensate 1 using the analytically determined feedstock composition. Table 3 also present simulated yields obtained using reconstructed feed compositions by considering a different number of commercial indices. The results Sim2, *i.e.* using the ASTM D86 boiling point curve and the specific density as input to reconstruct the feed composition, differ significantly from the experimental values and from the simulated yields obtained with the detailed analytical composition. On the other hand, if also the global PINA weight fractions are used as input for the feedstock reconstruction, the agreement between the experimental and simulated data (Sim3 in Table 3) is quite satisfactory. This example clearly illustrates that the accuracy of the predicted yields can be improved by including more commercial indices. Note that the simulation results obtained for the yields of iso-C4H8 and cyclopentadiene using the combination of the feedstock reconstruction method and the SEMK model are not as good as those obtained with the combination of the analytically obtained composition and the SEMK model. This has two reasons. On the one hand it should be a priori clear that the composition with maximum Shannon entropy will differ from the analytically determined one. Our results show that this method is able to generate a detailed molecular composition that corresponds reasonably well to the analytically determined one, but some differences are inevitable and these are partly responsible for the differences shown in Table 3. Another source of differences is caused by difficulties to distinguish between different isomers because

TABLE 5

Commercial indices to reconstruct the molecular composition of the gas condensates and the range of experimental conditions used in the pilot plant cracking tests.

COMMERCIAL INDICES	
Specific density, 15/4°	0.69-0.74
PIONA-analysis (wt.%)	
Paraffins	32.7-40.5
Iso-paraffins	24.9-44.3
Olefins	0-0.6
Naphthenes	12.2-22.1
Aromatics	7.0-19.7
ASTM D86 boiling points (K)	
IBP	300-308
50% BP	368-385
FBP	537-593
PROCESS CONDITIONS	
COP (MPa)	1.7-1.8
COT (K)	1073-1113
Dilution (kg _{steam} /kg _{Hydrocarbon})	0.3-1.0
Hydrocarbon Flow Rate (kg/s)	8.3×10^{-4} - 1.3×10^{-3}
P/E-ratio (wt.%/wt.%)	0.50-0.64

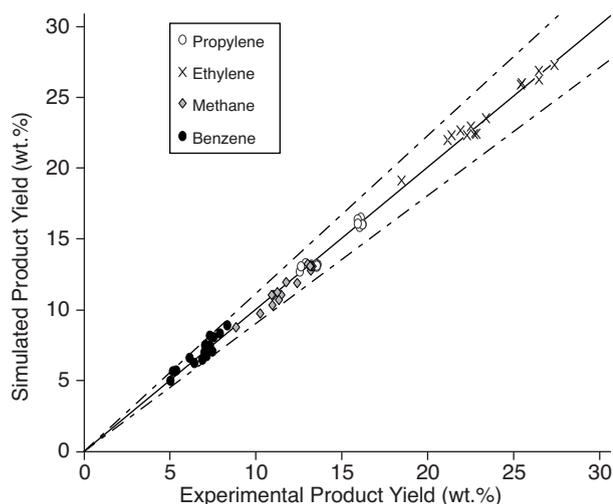


Figure 6

Parity plot for the yields of methane, ethylene, propylene and benzene obtained with 8 different gas condensates for 15 pilot plant experiments. Product yields were simulated using the feed reconstruction module combined with the SEMK model. All the commercial indices specified in Table 5 were used for the feed reconstruction. An overview of the range of experimental conditions is also specified in Table 5.

these components have similar physical properties (boiling point, density). The differences between the reconstructed composition and the analytically determined composition can propagate to the results obtained with the SEMK model. To obtain reasonable accurate simulation results with the SEMK model in combination with the feedstock reconstruction method, the specific density, the global PIONA weight or volume fractions, and the initial, 50% and final boiling point are required.

CONCLUSIONS

The detailed quantitative characterization of petroleum fractions is the current limit for exploiting the full power of SEMK models. New innovative techniques such as GC \times GC will allow a more detailed characterization in the coming years, however in the mean time some degree of lumping in SEMK models is unavoidable. The current SEMK model, with at least one pseudo-component defined per carbon number and per class of components, allows to simulate adequately the steam cracking of gas condensates. A main feature of the computer code is that the lumping hypotheses can easily be relaxed, enabling to study the cracking behavior of

a particular structural isomers. This lumping flexibility is one of the main advantages of this SEMK model and allows to easily adapt the model to future improvements in characterization of petroleum cuts.

The method for feedstock reconstruction based on Shannon's entropy criterion can be used to reconstruct gas condensate feedstocks with a rather limited library containing 107 molecules. The reconstructed detailed PIONA weight fractions correspond reasonably well with the analytically determined one indicating a reasonable correspondence between the two compositions. The number of analytically determined commercial indices taken into account for the feed reconstruction significantly affects the reconstructed composition. The feed reconstruction method allows easy implementation of additionally acquired analytical data and thus allows to improve the agreement between simulated and analytically determined composition of the feed. Comparison of the simulation yields with pilot plant data obtained from a set of 15 pilot plant experiments performed in the LPT-pilot plant installation shows a good agreement between the simulated and experimentally determined product yields. To obtain accurate simulation results, at least the specific density, the global PIONA weight or volume fractions, and the initial, 50% and final boiling point are required.

ACKNOWLEDGEMENTS

Kevin M. Van Geem holds a Postdoctoral Fellowship of the Fund for Scientific Research, Flanders, Belgium.

REFERENCES

- Baltanas M.A., Froment G.F. (1985) Computer-generation of reaction networks and calculation of product distributions in the hydroisomerization and hydrocracking of paraffins, *Comput. Chem. Eng.* **9**, 1, 71-81.
- Baltanas M.A., Van Raemdock K.K., Froment G.F., Mohedas S.R. (1989) Fundamental kinetic modeling of hydroisomerization and hydrocracking on Noble metal loaded faujasites -1. rate parameters for Hydroisomerization, *Ind. Eng. Chem. Res.* **28**, 7, 899-910.
- Beirnaert H.C., Alleman J.R., Marin G.B. (2001) A fundamental kinetic model for the catalytic cracking of alkanes on a USY zeolite in the presence of coke formation, *Ind. Eng. Chem. Res.* **40**, 5, 1337-1347.
- Broadbelt L.J., Stark S.M., Klein M.T. (1994) Computer-generated Pyrolysis Modeling – on-the-fly Generation of Species, Reactions, and Rates, *Ind. Eng. Chem. Res.* **33**, 790-799.
- Buda F., Heyberger B., Fournet R., Glaude P.A., Warth V., Battin-Leclerc F. (2006) Modeling of the gas-phase oxidation of cyclohexane, *Energ. Fuel.* **20**, 4, 1450-1459.
- Chevalier C., Warnatz J., Melenk H. (1990) Automatic generation of Reaction-mechanism for the Oxidation of higher Hydrocarbons, *Ber. Bunsen-Gesellsch.* **94**, 1362-1367.

- Chinnick S.J., Baulch D.L., Aysough P.B. (1988) An expert System for Hydrocarbon Pyrolysis, *Chemometr. Intell. Lab.* **5**, 1, 39-52.
- Dalluge J., Beens J., Brinkman U.A.T. (2003) Comprehensive two-dimensional gas chromatography: a powerful and versatile analytical tool, *J. Chromatogr. A* **1000**, 1-2, 69-108.
- Dente M., Ranzi E. (1979) Detailed Prediction of Olefin Yields from Hydrocarbon Pyrolysis through a Fundamental Simulation Program SPYRO, *Comput. Chem. Eng.* **3**, 61.
- Dente M., Ranzi E., Bozzano G., Faravelli T., Valkenburg P.J.M. (2001) Heavy Component Description in the Kinetic Modelling of Hydrocarbon Pyrolysis, *AIChE Spring National Meeting*, Houston, TX, USA.
- Dente M., Bozzano G., Faravelli T., Marongiu A., Ranzi E. (2007) Kinetic Modeling of Pyrolysis processes in Gas and Condensed Phase, *Adv. Chem. Eng.* **32**, in press.
- Dhuyvetter I., Reyniers M.F., Froment G.F., Marin G.B. (2001) The Influence of Dimethyl Disulfide on Naphtha Steam Cracking, *Ind. Eng. Chem. Res.* **40**, 4353-4362.
- Dietz W.A. (1967) Response Factors for Gas Chromatographic Analyses, *J. Gas Chromatogr.* **5**, 68.
- Dierickx J.L., Plehiers P.M., Froment G.F. (1986) Online Gas-Chromatographic Analysis of Hydrocarbon Effluents – Calibration Factors and their Correlation, *J. Chromatogr.* **362**, 155.
- DiMaio F.P., Lignola P.G. (1992) KING, a Kinetic Network Generator, *Chem. Eng. Sci.* **47**, 2713.
- Fedors R. (1974) A Method for Estimating Both the Solubility Parameters and Molar Volumes of Liquids, *Polym. Eng. Sci.* **14**, 147.
- Feng W., Vynckier E., Froment G.F. (1993) Single-event kinetics of catalytic cracking, *Ind. Eng. Chem. Res.* **32**, 12, 2997-3005.
- Floudas C.A., Pardalos P.M. (1996) *State of the art in global optimization: Computational methods and Applications*, Kluwer, New York.
- Froment G.F. (1992) Kinetics and Reactor Design in the Thermal Cracking for Olefin Production, *Chem. Eng. Sci.* **47**, 2163.
- Guiasu S., Shenitzer A. (1985) The principle of maximum entropy, *Math. Intell.* **7**, 42-48.
- Guillaume D., Surla K., Galtier P. (2003) From single events theory to molecular kinetics - application to industrial process modeling, *Chem. Eng. Sci.* **21**, 4861-4869.
- Hillewaert L.P., Dierickx J.L., Froment G.F. (1988) Computer-Generation of Reaction Schemes and Rate-Equations for Thermal-Cracking, *AIChE J.* **34**, 17-25.
- Hudebine D. (2003) Reconstruction Moléculaire de Coupes Pétrolières, *PhD Thesis*, École Normale Supérieure de Lyon.
- Hudebine D., Vera C., Wahl F., Verstraete J.J. (2002) Molecular Representation of Hydrocarbon Mixtures from Overall Petroleum Analyses. *AIChE 2002 Spring Meeting*, New Orleans, LA, March, 10-14, Paper 27a.
- Hudebine D., Verstraete J.J. (2004) Molecular Reconstruction of LCO Gasoils from overall Petroleum Analyses, *Chem. Eng. Sci.* **59**, 4755-4763.
- Hughey C.A., Rodgers R.P., Marshall A.G. (2002) Resolution of 11 000 compositionally distinct components in a single Electrospray ionization Fourier transform ion cyclotron resonance mass spectrum of crude oil, *Anal. Chem.* **74**, 16, 4145-4149.
- Joo E., Park S., Lee M. (2001) Pyrolysis Reaction Mechanism for Industrial Naphtha Cracking Furnaces, *Ind. Eng. Chem. Res.* **40**, 2409-2415.
- Kapur J.N., Kesavan H.K. (1992) *Entropy Optimization Principles with Applications*, Academic Press, San Diego.
- Kuo J.C.W., Wei J. (1969) A Lumping Analysis in Monomolecular Reaction Systems – Analysis of Approximately Lumpable System, *Ind. Eng. Chem. Fund.* **8**, 124.
- Li S., Petzold L.R. (1999) Design of New DASPK for Sensitivity Analysis, *UCSB Technical report*.
- Liguras D.K., Allen D.T. (1989a) Structural Models for Catalytic Cracking. 1. Model Compound Reactions, *Ind. Eng. Chem. Res.* **28**, 665.
- Liguras D.K., Allen D.T. (1989b) Structural Models for Catalytic Cracking. 2. Reactions of Simulated Oil Mixtures, *Ind. Eng. Chem. Res.* **28**, 674.
- Lozano-Blanco G., Thybaut J.W., Surla K., Galtier P., Marin G.B. (2006) Fischer-Tropsch synthesis: Development of a microkinetic model for metal catalysis, *Oil Gas Sci. Technol.* **61**, 4, 489-496.
- Marrero M.J., Pardillo F.E., Fernandez B.S. (1999) Estimation of Hydrocarbon Properties from Group-Interaction Contributions, *Chem. Eng. Commun.* **176**, 161.
- Martens G.G., Thybaut J.W., Marin G.B. (2001) Single-event rate parameters for the hydrocracking of cycloalkanes on Pt/US-Y zeolites, *Ind. Eng. Chem. Res.* **40**, 8, 1832-1844.
- Matheu D.M., Dean A.M., Grenda J.M., Green W.H. (2003) Mechanism Generation with Integrated Pressure dependence: A New Model for Methane Pyrolysis, *J. Phys. Chem. A* **107**, 8552-8565.
- Merdrignac I., Espinat D. (2007) Physicochemical characterization of petroleum fractions: the State of the Art, *Oil Gas Sci. Technol.* **62**, 1, 7-32.
- Mullins O.C., Rodgers R.P., Weinheber P., Klein G.C., Venkataraman L., Andrews A.B., Marshall A.G. (2006) Oil reservoir characterization via crude oil analysis by downhole fluid analysis in oil wells with visible-near-infrared spectroscopy and by laboratory analysis with electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Energ. Fuel.* **20**, 6, 2448-2456.
- Neurock M., Nigam A., Trauth D., Klein M.T. (1994) Molecular Representation of Complex Hydrocarbon Feedstocks through Efficient Characterization and Stochastic Algorithms, *Chem. Eng. Sci.* **49**, 4153.
- Phillips J.B., Xu J.Z. (1995) Comprehensive multidimensional gas chromatography, *J. Chromatogr. A* **703**, 1-2, 327-334.
- Prickett S.E., Mavrovouniotis M.L. (1997) Construction of Complex Reaction Systems, *Comput. Chem. Eng.* **21**, 1219-1235.
- Quann R.J., Jaffe S.B. (1996) Building Useful Models for Complex Reaction Systems in Petroleum Refining, *Chem. Eng. Sci.* **51**, 1615.
- Ranzi E., Dente M., Plerucci S., Biardi G. (1983) Initial Product Distribution from Pyrolysis of Normal and Branched Paraffins, *Ind. Eng. Chem. Fund.* **22**, 132-139.
- Ranzi E., Dente M., Goldaniga A., Bozzano G., Faravelli T. (2001) Lumping procedures in Detailed Kinetic Modeling of gasification, pyrolysis, partial oxidation and combustion of hydrocarbon mixtures, *Prog. Energ. Combust.* **27**, 99-139.
- Ranzi E., Faravelli T., Gaffuri P., Sogaro A. (1995) Low Temperature Combustion – Automatic-Generation of Oxidation Reactions and Lumping procedures, *Combust. Flame* **102**, 179-192.
- Reyniers M.F., Froment G.F. (1995) Influence of Metal Surface and Sulfur Addition on Coke Deposition in the Thermal Cracking of Hydrocarbons, *Ind. Eng. Chem. Res.* **34**, 773-785.
- Riazi M.R. (2005) *Characterization and properties of petroleum fractions*, ASTM International, West Conshohocken.
- Rice F.O. (1931) The Thermal Decomposition of Organic Compounds from the Standpoint of Free Radicals, *J. Am. Chem. Soc.* **53**, 1959.

- Rice F.O., Herzfeld K.F. (1934) The Thermal Decomposition of Organic Compounds from the Standpoint of Free Radicals. VI. The Mechanism of Some Chain Reactions, *J. Am. Chem. Soc.* **56**, 284.
- Saeyns M. (2003) Ab initio Modelling as a Tool for the Sustainable Development of Chemical processes, *PhD Dissertation*, UGent.
- Saeyns M., Reyniers M.F., Marin G.B., Van Speybroeck V., Waroquier M. (2004) Ab initio group contribution method for activation energies for β scissions and radical additions, *AIChE J.* **50**, 426-444.
- Saeyns M., Reyniers M.F., Van Speybroeck V., Waroquier M., Marin G.B. (2006) Ab initio group contribution method for activation energies of hydrogen abstraction reactions, *Chem. Phys. Chem.* **7**, 188-199.
- Shannon C.E. (1948) A Mathematical Theory of Communications, *Bell System Tech. J.* **27**, 379.
- Singh J., Kumar M.M., Saxena A.K., Kumar S. (2005) Reaction pathways and product yields in mild thermal cracking of vacuum residues: A multi-lump kinetic model, *Chem. Eng. J.* **108**, 3, 239-248.
- Trauth D.M. (1993) Structure of complex mixtures through characterization, reaction, and modeling, *PhD Thesis*, University of Delaware.
- Trauth D.M., Stark S.M., Petti T.F., Neurock M., Klein M.T. (1994) Representation of the molecular structure of petroleum resid through characterization and Monte Carlo modeling, *Energ. Fuel.* **8**, 3, 576-580.
- Van Geem K.M., Reyniers M.F., Marin G.B. (2005) Two Severity Indices for Scale-Up of Steam Cracking Coils, *Ind. Eng. Chem. Res.* **44**, 3402-3411.
- Van Geem K.M., Reyniers M.F., Marin G.B., Song J., Mattheu D.M., Green W.H. (2006) Automatic Network generation using RMG for Steam Cracking of n-Hexane, *AIChE J.* **52**, 2, 718-730.
- Van Geem K.M., Hudebine B., Reyniers M.F., Whal F., Verstraete J.J., Marin G.B. (2007) Molecular reconstruction of naphtha steam cracking feedstocks based on commercial indices, *Comput. Chem. Eng.* **31**, 1020-1034.
- Vendeuvre C., Ruiz-Guerrero R., Bertoncini F., Duval L., Thiebaut D. (2007) Comprehensive two-dimensional Gas chromatography for Detailed Characterisation of Petroleum products, *Oil Gas Sci. Technol.* **62**, 1, 43-55.
- Verstraete J.J., Revellin N., Dulot H., Hudebine D. (2004) Molecular Reconstruction of Vacuum Gas Oils, *Prepr. Pap.-Am. Chem. Soc., Div. Fuel Chem.* **49**, 1, 20-21.
- Warth V., Battin-Leclerc F., Fournet R., Glaude P.A., Come G.M., Scacchi G. (2000) Computer based generation of reaction mechanisms for gas-phase oxidation, *Comput. Chem.* **24**, 541-560.
- Wei J., Kuo J.C.W. (1969) A Lumping Analysis in Monomolecular Reaction Systems – Analysis of the Exactly Lumpable System, *Ind. Eng. Chem. Fund.* **8**, 114.
- Willems P.A., Froment G.F. (1988) Kinetic modelling of the Thermal Cracking of Hydrocarbons. 1. Calculation of frequency factors, *Ind. Eng. Chem. Res.* **27**, 1959-1966.

Final manuscript received in May 2007

APPENDIX

Molecular library for gas condensate feedstocks (107 components)

C3	propane	C4	isobutane, butane
C5	2-methyl-butane, cyclo-pentane, pentane, 2,2-dimethyl-propane	C6	2-methyl-pentane, 3-methyl pentane, 2,3-dimethyl-butane, 2,2-dimethyl-butane, hexane, methyl-cyclopentane, cyclohexane, benzene
C7	2,2-dimethyl-pentane, 2,3-dimethyl-pentane, 2,4 dimethyl-pentane, 3,3-dimethyl-pentane, 2-methyl-hexane, 3-methyl-hexane, 4-methyl-hexane, heptane, 1,3-dimethyl-cyclopentane, ethyl-cyclopentane, methyl-cyclohexane, 1,2-dimethyl-cyclopentane, 1,1-dimethyl-cyclopentane, toluene	C8	2-methyl-heptane, 3-methyl-heptane, 4-methyl-heptane, 2,4-dimethyl-hexane, 2,5-dimethyl-hexane, 3,3-dimethyl-hexane, 2,3,4-trimethyl-pentane, 3-ethyl-hexane, octane, iso-propyl-cyclopentane, 1,4-dimethyl-cyclohexane, (1-methyletyl)-cyclopentane, 1,3-dimethyl-cyclohexane, ethyl-cyclohexane, ethyl-benzene, o-xylene, p-xylene, m-xylene
C9	2-methyl-octane, 3-methyl-octane, 4-methyl-octane, 2,6-dimethyl-heptane, 3,3-dimethyl-heptane, 3-ethyl-heptane, nonane, n-butyl-cyclopentane, 1,1,3-trimethyl-cyclohexane, (1-methylethyl)-cyclohexane, n-propyl-cyclohexane, 1,2,3-trimethyl-benzene, n-propyl-benzene, 1,3,5-trimethyl-benzene, 2-ethyl-1-methyl-benzene	C10	2-methyl-nonane, 3-methyl-nonane, 4-methyl-nonane, 2,3-dimethyl-octane, 2,5-dimethyl-octane, 2,6-dimethyl-octane, 3,3-dimethyl-octane, 3,4-dimethyl-octane, decane, n-butyl-cyclohexane, (1-methylpropyl)-benzene, naphthalene, 1,2,3,4-tetrahydro-naphthalene
C11	2-undecene, 2-methyl-decane, 3-methyl-decane, 4-methyl-decane, 3,5-dimethyl-nonane, 2,6-dimethyl-nonane, undecane, n-pentyl-cyclohexane, 1-methyl-naphthalene	C12	2,5-dimethyl-decane, 3-methyl-undecane, 4-methyl-undecane, dodecane, 1-ethyl-naphthalene
C13	2,6-dimethyl-undecane, 3-methyl-dodecane, 4-methyl-dodecane, tridecane	C14	2,6-dimethyl-dodecane, 3-methyl-tridecane, 4-methyl-tridecane, tetradecane
C15	3-methyl-tetradecane, 4-methyl-tetradecane, pentadecane	C16	3-methyl-hexadecane, 4-methyl-hexadecane, hexadecane
C17	3-methyl-hexadecane, heptadecane	C18	3-methyl-heptadecane, octadecane
C19	3-methyl-octadecane, nonadecane	C20	