

Statistical Reconstruction of Gas Oil Cuts

D. Hudebine*, J. Verstraete and T. Chapus

Institut français du pétrole, IFP-Lyon, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize - France
e-mail: damien.hudebine@ifp.fr - jan.verstraete@ifp.fr - thierry.chapus@ifp.fr

* Corresponding author

Résumé — Reconstruction statistique de coupes gazoles — Les coupes gazoles sont des mélanges extrêmement complexes de plusieurs milliers de composés chimiques différents. De ce fait, les analyses pétrolières conventionnelles ne permettent pas d'obtenir un détail moléculaire qui serait pourtant nécessaire aux développements de modèles cinétiques robustes et prédictifs. Récemment, les techniques de chromatographie bidimensionnelle (GC2D) ont entraîné un saut qualitatif important dans le domaine de la caractérisation des gazoles mais celles-ci restent des outils de R&D encore peu généralisés dans l'industrie pétrolière. Par rapport à cette problématique, le but de la reconstruction statistique de gazoles consiste donc à fournir un substitut à l'analyse GC2D en proposant de caractériser les gazoles sous la forme de matrices de fractions molaires de pseudo-composés décrits par famille chimique et nombre d'atomes de carbone. Les analyses utilisées en entrée sont la spectrométrie de masse Fitzgerald, la spéciation soufre (chromatographie monodimensionnelle couplée à un détecteur du soufre par chimiluminescence), les teneurs en azote total et azote basique qui permettent de quantifier les proportions des différentes familles chimiques représentées dans la matrice. La distillation simulée est utilisée quant à elle pour avoir une information sur la volatilité de la coupe gazole. La méthode de reconstruction proposée dans cet article se base principalement sur une distribution statistique de référence du nombre d'atomes de carbone des chaînes alkyles sur les noyaux naphthéno-aromatiques. Pour chaque famille chimique, la connaissance du nombre potentiel de chaînes alkyles et l'estimation de la longueur maximale de ces chaînes permettent alors de déterminer la distribution par nombre d'atomes de carbone en dilatant la distribution de référence. Au final, la matrice de fractions molaires obtenue possède des propriétés très proches des analyses utilisées pour la reconstruction. Elle permet aussi de prédire, avec une grande précision, des analyses complémentaires comme la teneur en hydrogène, la teneur en carbone aromatique ou la densité à 15 °C. Enfin, elle peut être employée très efficacement dans des modèles cinétiques comme ceux utilisés à l'IFP pour prédire les performances d'un procédé d'hydrotraitement de gazoles.

Abstract — Statistical Reconstruction of Gas Oil Cuts — Gas oil cuts are extremely complex mixtures of several thousands of different chemical species. Consequently, conventional petroleum analyses do not allow to obtain the molecular detail that is required for the development of robust and predictive kinetic models. Recently, two-dimensional Gas Chromatographic techniques (GC2D) have greatly improved the knowledge in the field of characterization of gas oils. However, they remain R&D tools and are hardly utilized in the refining industry. Hence, the goal of the statistical reconstruction of gas oils is to provide a surrogate for this GC2D analysis. To this aim, the gas oil cuts are characterized by means of matrices of molar fractions of pseudo-compounds, which are classified by chemical family and by carbon atom number. The input analyses are the Fitzgerald mass spectrometry, the sulfur speciation (one-dimensional gas chromatography coupled to a specific sulfur chemiluminescence detector) and the total nitrogen and basic nitrogen contents, and allow to

quantify the proportions of all the chemical families present in the matrix. The simulated distillation is also used in order to introduce information on the volatility of the gas oil cut. The reconstruction method proposed in this paper is mainly based on a reference statistical distribution of the number of carbon atoms for the side chains connected to the naphtho-aromatic cores. For each chemical family, the knowledge of the number of potential side chains and the estimation of the maximum length of these alkyl chains allow to determine the carbon number distribution by adjusting of the reference distribution. After reconstruction, the properties of the resulting molar fractions matrix are very close to the analyses used for the reconstruction. Moreover, the method allows to predict, with a high precision, complementary analyses such as the hydrogen content, the aromatic carbon content and the density at 15°C. Finally, the matrix can be efficiently used to develop kinetic models like those employed at IFP to predict the performances of gas oil hydrotreating units.

INTRODUCTION

The use of kinetic models in the field of refining is very often limited by the molecular complexity of petroleum fractions. Indeed, oil mixtures (gasolines, atmospheric gas oils, vacuum distillates, atmospheric residues, etc.) consist of thousands of chemical compounds, which are difficult to separate and detect analytically. It is even more difficult to quantify all of the potential species and represent the petroleum mixtures on a molecular level. Hence, the most widely applied kinetic models are based on a lumped representation of the petroleum fractions in order to follow the chemical transformations. The quantitative evolution of these lumps is predicted using simplified kinetic models that describe the reactions between different families together with their kinetic rates, as illustrated in several classic publications concerning catalytic reforming [1-3], catalytic cracking [4-6], hydrocracking [7, 8], (hydro)pyrolysis [9, 10], and hydrotreating [11-14]. However, these kinetic models lack molecular information, and their parameters are very often feed dependent. To develop more accurate kinetic models, it is necessary to build more detailed reaction networks, and hence to get access to a more detailed description of the feeds.

In the particular case of gas oil hydrotreating, the current chromatographic analyses are not powerful enough to identify and quantify all the chemical compounds. While two-dimensional Gas Chromatography (GC2D) is starting to provide very detailed analyses of gas oil fractions, this method is still very difficult to apply [15-17]. The only analytical methods widely available at present are distillation, mass spectrometry and sulfur speciation [18], but they only allow to describe gas oils in terms of large physico-chemical families. To overcome the analytical limitations, numerous efforts have been made to develop molecular reconstruction methods to represent various petroleum fractions by means of mixtures of chemical structures [10, 19-35].

In a first approach, the petroleum fraction can be represented by a fixed set of predefined molecules. Their molar fractions are adjusted in order to obtain a mixture whose properties are

close to the analytical characteristics of the petroleum cut. Allen and Liguras [19-21] used such a method in which they minimized a divergence criterion under 190 linear equality constraints that were derived from the Gas Chromatography, ^{13}C Nuclear Magnetic Resonance (NMR) and ^1H NMR analyses they performed on the actual sample. Khorasheh *et al.* [22] start from the elemental analysis, ^1H NMR and ^{13}C NMR analysis and determine the concentrations of a selected number of structural groups by minimizing an objective function subject to the linear constraints given by the analytical data. Hudebine *et al.* [23, 24] developed an entropy maximization technique to calculate the molar fractions of a predefined set of molecules. With respect to the previous methods, which directly modify the molar fractions of all the molecules present, the entropy criterion was advantageously selected to reduce the size of the optimization problem. Indeed, instead of estimating several hundreds or several thousands of mole fractions, restructuring the optimization problem allows to limit the number of parameters to one Lagrange parameter per available analytical property [24]. A second advantage of the entropy criterion consists in the flexibility of the method, since it adjust the number of Lagrange parameters to the number of analytical data available, irrespectively of the number of molecules in the basis set. The entropy maximization technique has been successfully applied to various types of naphtha fractions [25, 26] and gas oil fractions [24]. Quann and Jaffe [27-29] suggested a similar method but the molecules are replaced by vectors of structural blocks, called "Structure Oriented Lumping" (SOL) vectors. To determine the molecular abundance of each molecule, these authors also minimize the differences between the measured and the predicted properties by means of a non-negative Lagrange multiplier method [29]. Finally, a variant of this method is proposed by Zhang [30], in which the various molecules and their properties are replaced by various gas oil fractions that were characterized in great detail by coupling Gas Chromatography with Mass Spectrometry to result in a composition matrix by carbon number and chemical family. All the above methods have similar drawbacks: for a given petroleum cut, the calculated mole fractions are

strongly dependent on the set of molecules (or vectors or matrices) that was initially chosen by the user. Hence, it is very important to start from a mixture that already represents the petroleum fraction relatively well.

Alternatively, Neurock *et al.* [31, 32] developed a method called “stochastic reconstruction”, in which petroleum cuts are described by means of a set of distributions of molecular structural attributes. These distributions are then sampled by a Monte-Carlo method in order to generate an equimolar mixture of molecules. After coupling this reconstruction step to an optimization loop on the parameters of the distributions of molecular attributes, the method allows to generate mixtures that correctly represent the analytical properties of heavy petroleum fractions [10, 33]. Based on the same concept, Hudebine [23, 24, 34] applied a similar method to the reconstruction of various Light Cycle Oil (LCO) gas oil fractions [24, 34] and petroleum vacuum residues [23]. The main drawback of these methods resides in the computational burden that is linked to the Monte-Carlo approach. Moreover, the random sampling of the distributions of structural attributes induces random noise on the objective function, which impedes the use of derivative-based descent optimizer algorithms. Hence, the optimization loop is generally based on a simulated annealing or a genetic algorithm optimizer, which also adds to the overall calculation time.

To compensate for the drawback of the two above-listed classes of reconstruction methods, Hudebine and Verstraete [34] advantageously coupled the stochastic reconstruction and entropy maximization methods in a two-step molecular reconstruction algorithm. This approach was developed for and validated on LCO gas oils, and was later extended to vacuum gas oils [35]. However, when relumping the stochastically generated set of gas oil molecules into a composition matrix by carbon number and chemical family, some of the elements of the matrix may be absent because none of its molecules has been generated. Moreover, it is preferable to adapt this method for different types of gas oils, such as straight run gas oils, hydrotreated gas oils, LCO gas oils, coker gas oils, etc.

In this paper, we propose a novel method of characterizing gas oil fractions in terms of a more detailed description than that used in classic lumping methods, but without going all the way to the molecular level. The objective of this work is to describe gas oils at an intermediate level in the form of a matrix of pseudo-compounds, in which the chemical families present in gas oils are subdivided according to their number of carbon atoms. This article outlines the new reconstruction algorithm and illustrates its application to gas oil fractions. The reconstruction technique assumes that the length of all alkyl chains can be described by one, and only one, reference distribution. As the method is based on a statistical representation of this alkyl chain length, it is called “statistical reconstruction” to distinguish it from its related method, “stochastic reconstruction” [10, 23, 31-35].

1 CHARACTERIZATION OF GAS OILS BY MEANS OF MATRICES OF PSEUDO-COMPOUNDS

1.1 The Matrix Description of Gas Oils

To characterize gas oils by means of matrices of pseudo-compounds, these pseudo-compounds have to be defined first. They must be representative of all gas oil cuts and sufficiently accurate to limit the disadvantages of the lumping approaches [36, 37]. Our approach consists in defining different chemical families and discretizing them according to the number of carbon atoms. In this way, each pseudo-compound is characterized by its chemical family and its carbon number, which correspond to the two dimensions of the matrix of pseudo-compounds. This type of description has already been used, for example, for the analytical development of gas chromatography in combination with a mass spectrometer [38] or for representing gas oil fractions [30].

1.1.1 Choice of the Chemical Families

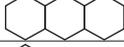
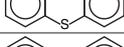
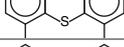
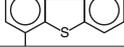
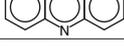
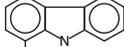
Gas oils are mainly constituted of hydrocarbon, sulfur and nitrogen species. Consequently, the choice of the chemical families has been based on the results obtained from a Fitzgerald mass spectrometry analysis (derived from ASTM D2425) [39-41], gas chromatography analysis with sulfur chemiluminescence detection [18], and nitrogen analyses. These different analyses have been used to define 28 different chemical families that describe all the gas oil cuts, which can be processed in a modern refinery. These families are listed in Table 1. However, it should be noted that their selection is based on the following assumptions:

- all the naphtheno-aromatic cores are described by cycles with 6 carbon atoms, excluding 5-carbon atom cycles;
- all the naphtheno-aromatic cores have a maximum of 3 cycles, excluding any tetracyclic families;
- olefins are not included as their abundance is relatively limited in most cases and since it is impossible to obtain their concentration by means of the above analysis methods;
- the sulfide/mercaptan family is assumed to be an acyclic aliphatic structure;
- the dibenzothiophene family has been divided into 3 subfamilies to take into account their differences in hydrodesulfurization activity [13].

1.1.2 Choice of the Number of Carbon Atoms

The choice of the number of carbon atoms was initially based on classic gas oil cut points. However, in order to allow for heavy gas oils, the maximum number of carbon atoms has been fixed at 30, which corresponds to a boiling point of 450°C for *n*-triacontane. Similarly, in order to be able to manage the cracking phenomena that occur in hydrotreating units, the minimum number of carbon atoms has been set to 1.

TABLE 1
Listing of the chemical families defined in the matrix

Index	Homologous family	Head of the family (name)	Head of the family (structure)
1	Paraffins	Methane	CH ₄
2	Monocycloparaffins	Cyclohexane	
3	Dicycloparaffins	Perhydronaphthalene	
4	Tricycloparaffins	Perhydroanthracene	
5	Benzenes	Benzene	
6	Tetralins	1,2,3,4-tetrahydronaphthalene	
7	Dinaphthenobenzenes	1,2,3,4,5,6,7,8-octahydroanthracene	
8	Naphthalenes	Naphthalene	
9	Tetrahydroanthracenes	1,2,3,4-tetrahydroanthracene	
10	Dihydroanthracenes	7,8-dihydroanthracene	
11	Anthracenes	Anthracene	
12	Pyrenes	Pyrene	
13	Bicyclohexyls	Bicyclohexyl	
14	Cyclohexylbenzenes	Cyclohexylbenzene	
15	Biphenyls	Biphenyl	
16	Mercaptans/sulfides	Methanethiol	CH ₃ -SH
17	Thiophenes	Thiophene	
18	Benzothiophenes	Benzothiophene	
19	Dibenzothiophenes	Dibenzothiophene	
20	4-dibenzothiophenes	4-methylidibenzothiophene	
21	4,6-dibenzothiophenes	4,6-dimethyldibenzothiophene	
22	Anilines	Aniline	
23	Pyridines	Pyridine	
24	Quinolines	Quinoline	
25	Acridines	Acridine	
26	Pyrroles	Pyrrole	
27	Indoles	Indole	
28	Carbazoles	Carbazole	

1.1.3 The Pseudo-Compounds Matrix

The matrix of pseudo-compounds is constructed by discretization of the 28 chemical families between C_1 and C_{30} . Hence, a matrix is obtained of $28 \times 30 = 840$ pseudo-compounds to which 3 more compounds (H_2 , H_2S and NH_3) have to be added, resulting in a total of 843 pseudo-

compounds. However, the actual number of pseudo-compounds is smaller, as some compounds cannot exist in the matrix, *e.g.* alkyl benzenes in C_4 . Hence, the total number of pseudo-compounds drops down to 612. Table 2 shows the current representation of the pseudo-compounds matrix for the gas oils.

TABLE 2
Description of the matrix of pseudo-compounds

H_2	H_2S										NH_3											
Carbon number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	...	30	
Paraffins																						
Monocycloparaffins																						
Dicycloparaffins																						
Tricycloparaffins																						
Benzenes																						
Tetalins																						
Dinaphtenobenzenes																						
Naphthalenes																						
Tetrahydroanthracenes																						
Dihydroanthracenes																						
Anthracenes																						
Pyrenes																						
Bicyclohexyls																						
Cyclohexylbenzenes																						
Biphenyls																						
Sulfides																						
Thiophenes																						
Benzothiophenes																						
Dibenzothiophenes																						
4-Dibenzothiophenes																						
46-Dibenzothiophenes																						
Anilines																						
Pyridines																						
Quinolines																						
Acridines																						
Pyrroles																						
Indoles																						
Carbazoles																						

■ Unfeasible pseudo-compound.

1.2 Properties of the Pseudo-Compounds in the Matrix

1.2.1 Theoretical Aspects

In order to be able to calculate the properties of a mixture of chemical compounds, the composition of the mixture has to be known as well as the individual pure component properties. In this work, the mixture is represented by pseudo-compounds and not by pure components, which poses a difficulty with respect to lumping. Indeed, the pseudo-compound “C₈ alkyl benzene” includes ethyl benzene (NBP = 136.2°C), ortho-xylene (NBP = 143.9°C), meta-xylene (NBP = 139.2°C) and para-xylene (NBP = 138.3°C), which have slightly different properties. The “C₈ alkyl benzene” must therefore have the average properties of the pure components it contains. Moreover, in the literature, thermodynamic databases only contain pure component properties for species with less than 12 to 14 carbon atoms, except for a few very specific chemical families, such as the normal paraffins. Hence, it is not possible to determine the properties of pure components containing more than 14 carbon atoms by using the standard thermodynamic tables.

Consequently, in order to calculate the properties of the pseudo-compounds, it has been decided to use group contribution methods whenever necessary. They are based on the principle that a molecule can be broken down into a certain number of elementary chemical groups which may be either single atoms (aromatic carbon, naphthenic carbon), pairs of atoms (olefinic, cyanide), or larger functional groups (carboxyl, amide). Each group, whatever its place in the molecule, has its own contribution that adds to the properties of the molecule. The general equation used to calculate a property P using a group contribution method is therefore as follows:

$$P = f\left(\sum_i n_i \cdot C_i\right) \quad (1)$$

where: P Property to be determined

$f()$ Relation between the property and the group contributions

n_i Number of groups of type i

C_i Contribution of a group of type i

Group contribution methods are available to predict many different properties. Benson [42-45] devised a group contribution method that predicts the standard enthalpy of formation, the absolute entropy and the heat capacity of various temperatures. The Joback method [46] is particularly useful to predict the critical properties, normal boiling temperature and melting temperature of a molecule from 41 structural groups. More recently, Constantinou and Gani proposed a group contribution method for calculating the normal melting temperature, critical properties, standard free enthalpy, standard enthalpy of formation, standard

vaporization enthalpy and standard melting enthalpy from 251 increasingly complex structural groups [47].

1.2.2 Application to the Statistical Reconstruction Method

For our statistical reconstruction method, the following properties were determined or calculated for each pseudo-compound:

- chemical formula (carbon, hydrogen, sulfur and nitrogen);
- molecular weight;
- aromatic, naphthenic and paraffinic carbon content;
- normal boiling point;
- critical pressure;
- critical temperature;
- critical volume;
- molar volume at 15°C and 20°C;
- acentric factor ω ;
- solubility factor δ ;
- Rackett compressibility factor Z_{RA} .

The chemical formula, the molecular weight and the aromatic, naphthenic and paraffinic carbon contents are calculated directly from the structure of each pseudo-compound. For the other properties, group contribution methods were developed using a property value for the head of a homologous family and adding a contribution for the total number of carbon atoms in the alkyl chains. The property values of the heads of each homologous series have been taken from the Thermodynamic Research Center (TRC) data base [48]. The various properties are therefore determined as follows:

Normal Boiling Point

To determine the normal boiling point, the developed group contribution method consists in taking the normal boiling point of the head of the homologous family and adding a contribution for the each aliphatic carbon atom in the alkyl chains [23]:

$$\exp\left(\frac{T_b}{307.63}\right) = \exp\left(\frac{T_{b,head}}{307.63}\right) + 0.31012 \cdot nC_{ali} \quad (2)$$

where: T_b Boiling point of the pseudo-compound (K)

$T_{b,head}$ Boiling point of the head of the homologous family (K)

nC_{ali} Number of carbon atoms in the alkyl chain (adim)

The normal boiling points of the heads of each homologous family are available in the literature [48], and the values used are shown in Table 3.

Critical Properties

The calculation of the critical properties is based on the same principle, but using group contribution methods derived from

TABLE 3

Properties of the head compounds of the homologous families used to calculate the pseudo-compounds properties

Homologous family	Name of the head of the homologous family	$T_{b,head}$ (K)	$\Delta T_{c,head}$ (adim)	$\Delta V_{c,head}$ (cm ³ /mol)	$\Delta P_{c,head}$ (adim)	$\Delta V_{c,head}^{20}$ (cm ³ /mol)	ω_{head} (adim)	$\Delta H_{v,head}$ (J/m ³) ^{0.5}	ΔT_{polar} (K)
Paraffins	Methane	111	-0.0103	60.6	2.2959	47.9	0.011	8900	0
Monocycloparaffins	Cyclohexane	354	-0.0552	283.8	8.2608	113.3	0.210	30600	5
Dicycloparaffins	Perhydronaphthalene	463	-0.1018	463.0	12.8394	155.6	0.254	46300	-15
Tricycloparaffins	Perhydroanthracene	489	-0.1484	642.2	17.4180	197.1	0.298	62000	-25
Benzenes	Benzene	353	-0.0528	244.2	5.5080	92.8	0.210	31400	5
Tetalins	1,2,3,4-tetrahydronaphthalene	480	-0.0926	420.0	10.0924	137.2	0.328	51200	-15
Dinaphtenobenzenes	1,2,3,4,5,6,7,8-octahydroanthracene	567	-0.1392	599.2	14.6710	178.8	0.360	60500	-25
Naphthalenes	Naphthalene	491	-0.0910	393.6	8.2572	126.4	0.302	49600	-15
Tetrahydroanthracenes	1,2,3,4-tetrahydroanthracene	583	-0.1308	569.4	12.8416	165.8	0.392	59800	-25
Dihydroanthracenes	7,8-dihydroanthracene	578	-0.1300	556.2	11.9240	160.4	0.392	59800	-25
Anthracenes	Anthracene	613	-0.1292	543.0	11.0064	155.0	0.486	57600	-25
Pyrenes	Pyrene	668	-0.1498	611.0	11.9196	159.3	0.507	71500	-45
Bicyclohexyls	Bicyclohexyl	512	-0.1202	557.6	15.5930	189.1	0.428	54400	-15
Cyclohexylbenzenes	Cyclohexylbenzene	519	-0.1144	516.3	12.8431	170.5	0.378	59900	-15
Biphenyls	Biphenyl	527	-0.1086	475.0	10.0932	152.0	0.365	57600	-15
Sulfides	Methanethiol	279	-0.0096	121.4	2.7658	60.8	0.158	22300	-5
Thiophenes	Thiophene	357	-0.0344	192.9	4.1372	82.8	0.193	32200	0
Benzothiophenes	Benzothiophene	494	-0.0726	342.3	6.8864	116.4	0.296	51400	-20
Dibenzothiophenes	Dibenzothiophene	605	-0.1108	491.7	9.6356	145.0	0.398	65600	-30
4-Dibenzothiophenes	4-methylidibenzothiophene	618	-0.1267	543.6	11.0087	159.8	0.437	69900	-30
46-Dibenzothiophenes	4,6-dimethylidibenzothiophene	631	-0.1452	600.2	12.3818	176.2	0.476	74200	-30
Anilines	Aniline	457	-0.0753	268.6	6.4336	91.0	0.378	53300	-5
Pyridines	Pyridine	388	-0.0548	179.4	5.0592	80.4	0.239	37600	-5
Quinolines	Quinoline	510	-0.0930	328.8	7.8084	114.0	0.289	57500	-25
Acridines	Acridine	619	-0.1312	478.2	10.5576	142.6	0.438	67900	-35
Pyrroles	Pyrrole	403	-0.0506	191.4	4.6025	69.1	0.298	42900	-5
Indoles	Indole	526	-0.0888	340.8	7.3517	102.7	0.376	58400	-25
Carbazoles	Carbazole	628	-0.1270	490.2	10.1009	131.3	0.494	61600	-35

that of Marrero and Pardillo [49]. The critical property increments of the head of the homologous families are given in Table 3. The corresponding equations are as follows:

$$\Delta T_c = \Delta T_{c,head} - 0.0144 \cdot nC_{ali} \quad (3)$$

$$T_c = T_b \cdot (0.5881 - 0.9305 \cdot \Delta T_c - \Delta T_c^2)^{-1} \quad (4)$$

$$\Delta P_c = \Delta P_{c,head} + 1.3731 \cdot nC_{ali} \quad (5)$$

$$P_c = (0.1218 + 0.4609 \cdot n_A - \Delta P_c)^{-2} \quad (6)$$

$$\Delta V_c = \Delta V_{c,head} + 56.6 \cdot nC_{ali} \quad (7)$$

$$V_c = 26.3 + \Delta V_c \quad (8)$$

where:

- T_c Critical temperature of the pseudo-compound (K)
- ΔT_c Increment for the critical temperature of the pseudo-compound (adim)
- $\Delta T_{c,head}$ Increment for the critical temperature of the family head (adim)
- P_c Critical pressure of the pseudo-compound (bar)

ΔP_c	Increment for the critical pressure of the pseudo-compound (adim)
$\Delta P_{c,head}$	Increment for the critical pressure of the family head (adim)
n_A	Number of atoms in the pseudo-compound (adim)
V_c	Critical volume of the pseudo-compound (cm ³ /mol)
ΔV_c	Increment for the critical volume of the pseudo-compound (cm ³ /mol)
$\Delta V_{c,head}$	Increment for the critical volume of the family head (cm ³ /mol)

Molar Volume at 15°C and 20°C

The molar volume at 20°C of the pseudo-compound is determined using a group contribution method based on the molar volume at 20°C for the head of a homologous family (Tab. 3) and adding a contribution for the number of carbon atoms in the alkyl chains. To develop this group contribution method, the data were taken from the TRC data base [48]. The group contribution equation is given by:

$$V_m^{20} = V_{m,head}^{20} + 16.38 \cdot nC_{ali} \quad (9)$$

where: V_m^{20}	Molar volume of the pseudo-compound at 20°C (cm ³ /mol)
$V_{m,head}^{20}$	Molar volume at 20°C of the family head (cm ³ /mol)

With the above group contribution method, the density at 20°C can be calculated for the various pseudo-compounds. However, the standardized petroleum analyses measure the density at 15°C. Hence, it is necessary to apply an internally developed temperature correction to the above group contribution method. The molar volume at 15°C is determined from:

$$V_m^{15} = 0.99648 \cdot V_m^{20} - 0.26 \quad (10)$$

where: V_m^{15}	Molar volume of the pseudo-compound at 15°C (cm ³ /mol)
-------------------	--

Acentric Factor ω

For this property, a similar group contribution method was developed based on the acentric factors from [50]. As for the other properties, the values of the acentric factors of the heads of the homologous families are given in Table 3. The corresponding equation is given by:

$$\omega = \omega_{head} + 0.0388 \cdot nC_{ali} \quad (11)$$

where: ω	Acentric factor of the pseudo-compound (adim)
ω_{head}	Acentric factor of the head of the homologous family (adim)

Rackett Compressibility Factor Z_{RA}

The Rackett compressibility factor Z_{RA} is deduced from the acentric factor by the following approximation formula [51]:

$$Z_{RA} = 0.29056 - 0.08775 \cdot \omega \quad (12)$$

where: Z_{RA}	Rackett compressibility factor of the pseudo-compound (adim)
-----------------	--

Solubility Factor δ

The solubility factor is defined as the square root of the ratio of the latent heat of vaporization to the molar volume at 20°C.

$$\delta = \sqrt{\frac{\Delta H_v}{V_m^{20}} \cdot 10^6} \quad (13)$$

where: ΔH_v	Latent heat of vaporization of the pseudo-compound (cm ³ /mol)
δ	Hildebrandt solubility factor of the pseudo-compound ((J/m ³) ^{0.5})

Hence, determining the solubility factor means determining the latent heat of vaporization. For the latter property, the following group contribution method derived from [47] is used:

$$\Delta H_v = \Delta H_{v,head} + 4306 \cdot nC_{ali} \quad (14)$$

where: $\Delta H_{v,head}$	Latent heat of vaporization of the family head (J/mol)
----------------------------	--

1.3 Determination of the Mixture Properties

Petroleum mixtures, like most complex chemical mixtures, are far from having an ideal thermodynamic behavior. While some properties of the mixture can easily be deduced from the pseudo-compounds matrix by mass balance (elemental analysis, average molecular weight, etc.), this is not the case with more complex properties such as the average density or the “True Boiling Point” distillation. For these, mixing rules or simplifying assumptions have to be introduced.

For the molecular reconstruction techniques, a set of mixing rules has been specifically selected [23], and the calculation methods for distillation data and average density are detailed below. In the case of distillation, the hypothesis consists in assuming a TBP distillation, *i.e.* the pseudo-compounds are ideally separated by increasing normal boiling points. For a simulated distillation, however, the boiling temperatures of the pseudo-compound have to be corrected to account for the residual polarity of the capillary column according to the following equation:

$$T_{b,corrected} = T_b + \Delta T_{polar} \quad (15)$$

where: $T_{b,corrected}$	Boiling point used for calculating simulated distillation (K)
ΔT_{polar}	Increment for the polarity of each homologous family (K)

The ΔT_{polar} values are given in Table 3.

For the average density, the mixture was assumed to be ideal with no excess molar volume calculation. However, this assumption is far from true and causes a bias on the calculation of the average density. After investigating this bias on 17 different gas oils, the average density of a gas oil mixture had to be corrected by means of the following internally developed equation:

$$d = \frac{M_w}{0.9404 \cdot \sum_{i=1}^N x_i \cdot V_{m,i}^{15} + 12.8} \quad (16)$$

where: d Average density of the mixture at 15°C (g/cm³)

M_w Average molecular weight of the mixture (g/mol)

$V_{m,i}^{15}$ Molar volume of the pseudo-compound i at 15°C (cm³/mol)

x_i Molar fraction of the pseudo-compound i (adim)

Now that the pseudo-component property calculations and the mixing rules have been defined and validated, only the matrix composition needs to be determined for each gas oil that has to be represented. This somewhat laborious task is performed by statistical reconstruction, which is described in the next section of this article.

2 THE STATISTICAL RECONSTRUCTION METHOD

2.1 Overview of the Theoretical Aspects

The principle of statistical reconstruction is based on two assumptions:

- the length of the alkyl chains grafted on the naphtho-aromatic cores statistically follows one, and only one, reference distribution that is common to all side chains. For this distribution, a gamma law is chosen, and its standard representation is shown in Figure 1;
- for the chemical families without a core (paraffins and sulfides/mercaptans), another reference statistical distribution is used. Again, a gamma-type distribution is selected (Fig. 1).

Using these assumptions, the molar distribution with respect to carbon number of the pseudo-compounds of a given family can easily be calculated. The sum of the molar fractions of the pseudo-compounds of each chemical family is then equal to 1.0. These normalized molar distributions allow to directly calculate the average molecular weight of each chemical family. Subsequently, the mass spectrometry, sulfur speciation, and nitrogen analyses are used to determine the mass fractions, and then the molar fractions of each chemical family. Finally, combining the molar fractions of

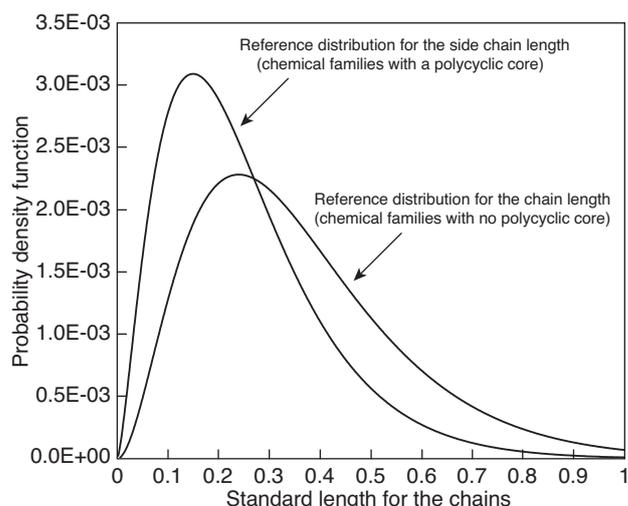


Figure 1

Reference distributions for the length of the side chains and for the length of the paraffins and sulfides/thiols.

each chemical family with its normalized molar carbon number distribution allows to calculate the molar composition of each individual pseudo-compound, thus establishing the entire pseudo-compounds matrix.

Each of these steps of the reconstruction algorithm will now be detailed in the following sections.

2.2 Standard Molar Distributions for Chemical Families with a Polycyclic Core

Establishing a standardized molar distribution for each chemical family with a polycyclic core requires three important data:

- the maximum number of alkyl chains, N_{Chains} , that can be added to the representative core of the chemical family. To calculate this value, the assumption has been made that only the aromatic CH carbon atoms and the naphthenic CH₂ carbon atoms can accept one and only one side chain;
- the reference distribution of the length of the side chains (see Fig. 1);
- the maximum length of a side chain, noted L . L may be a non-integer value, since the reconstruction technique is a statistical method.

By default, the smallest pseudo-compound in a chemical family consists of the naphtho-aromatic core without any side chain. The largest pseudo-compound in this chemical family consists of the same core with N_{Chains} side chains of L carbon atoms each. All the pseudo-compounds within these two limits are then present and their quantity is fixed by the normalized molar distribution.

For example, for the alkyl benzene family, a maximum of 6 side chains can be added, one per aromatic CH carbon

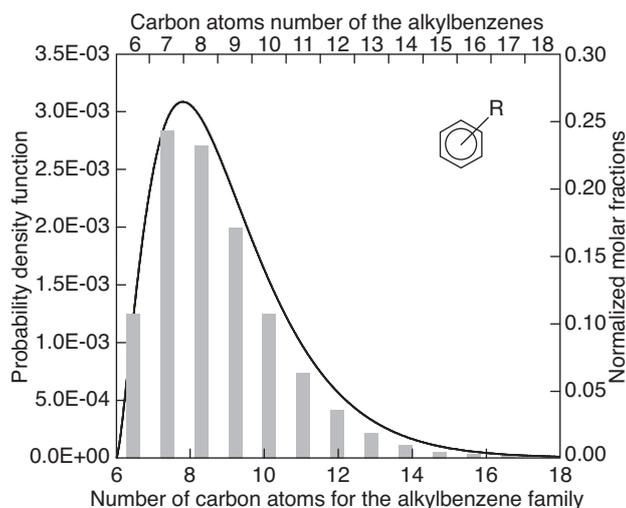


Figure 2

Example of the continuous distribution for the family of alkyl benzenes with $L = 2.0$ (line, left and bottom axes), and of the discrete distribution for the family of alkyl benzenes with $L = 2.0$ (histogram, top and right axes).

atom. Assuming a maximum side chain length L of 2.0, the smallest pseudo-compound in the alkyl benzene family is benzene (alkyl benzene in C_6) and the largest is alkyl benzene in C_{18} (6 carbon atoms for the core + 6 side chains of 2.0 carbon atoms = 18). The reference distribution shown in Figure 1 will first be homothetically distorted to account for these lower and upper limits (Fig. 2). This continuous distribution lies between C_6 and C_{18} in this example and will then be discretized to obtain the normalized distribution of the alkyl benzene family. After discretizing this continuous distribution with respect to the number of carbon atoms of each pseudo-compound, the probability of existence is transformed into the normalized molar fraction of the pseudo-compounds within their chemical family (Fig. 2).

However, the determination of the molar fraction distribution for each chemical family in the matrix is not yet completed, because the distillation cut points of the gas oil have to be taken into account. Indeed, if the gas oil cut to be reconstructed is a 160–400°C fraction, pseudo-compounds whose normal boiling points are below 160°C or above 400°C can not be present and have to be removed from the distribution. In a first approach, a perfect separation can be simulated, using the following rules:

- if the pseudo-compound has a corrected normal boiling point lower (or higher) than the initial (or final) boiling point of the simulated distillation, its molar fraction is set to 0.0;
- if the pseudo-compound has a corrected normal boiling point between the initial point and the final point, its molar fraction is retained.

Finally, the remaining part of the distribution of molar fractions is again renormalized within each chemical family.

However, the cut point adjustment based on a perfect separation leads to a reconstructed pseudo-compounds matrix that does not correctly represent the gas oil, particularly in terms of its calculated simulated distillation (Fig. 3) near the initial and final boiling points. Indeed, by applying the above rules, the calculated initial boiling point is too high, and the calculated simulated distillation curve is quasi-linear between its initial boiling point and its 90% boiling point, lacking the characteristic inflections near the initial boiling points that are due to the distillation overlaps of actual petroleum fractions. Hence, an improved approach was devised to imitate the separation of the molecules near the initial and final boiling points by introducing a separation efficiency factor η . In this second approach, the following rules apply:

- if the pseudo-compound has a corrected normal boiling point lower (or higher) than the initial (or final) boiling point of the simulated distribution, the efficiency factor η is fixed at 0.0;
- if the pseudo-compound has a corrected normal boiling point between the 10 wt% and 90 wt% points, the molar fraction is retained with an efficiency factor η of 100%;
- if the pseudo-compound has a corrected normal boiling point between the initial point and the 10 wt% point, the molar fraction is modified using a separation efficiency factor that is linearly interpolated between these limits;
- if the pseudo-compound has a corrected normal boiling point between 90 wt% point and the final point, the molar fraction is modified using a separation efficiency factor that is linearly interpolated between these limits.

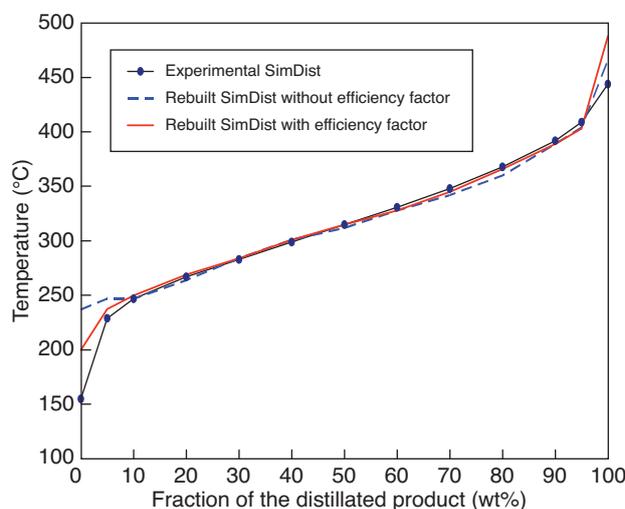


Figure 3

Experimental and reconstructed Simulated Distillation curves of a typical SRGO without and with the use of the separation efficiency factor η .

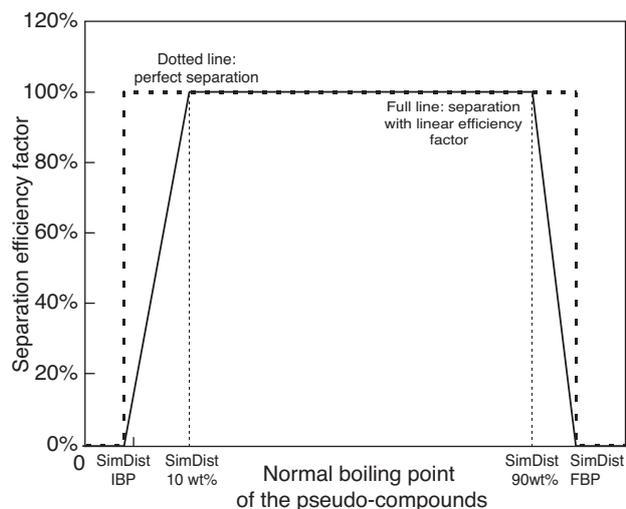


Figure 4

Separation efficiency methods used for pseudo-compounds distribution near the extremes of the simulated distillation boiling point range of a gas oil.

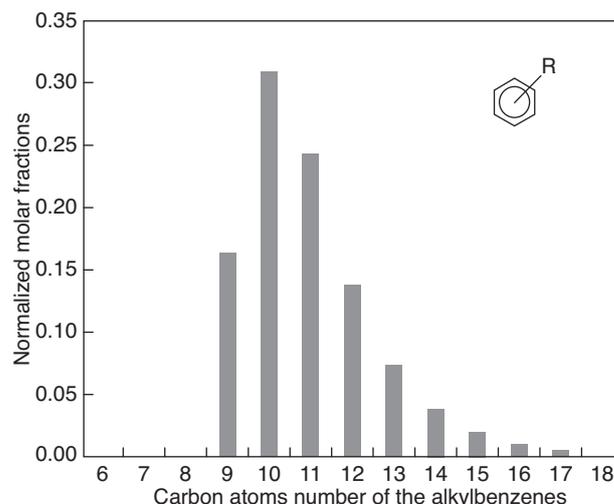


Figure 5

Example of a discrete distribution for the family of alkyl benzenes with $L = 2.0$ and the distillation cuts 160–400°C (after the use of the separation efficiency factor).

Applying this linearly interpolated separation efficiency factor significantly improved the results (Fig. 3), especially near the initial boiling point. This is particularly important for the flash calculations, as those needed for gas oil hydrotreating models. Figure 4 summarizes the two approaches. The current separation method is still being perfected by using non-linearly interpolated separation efficiency factors, which can be calculated by means of parameterized sigmoid functions.

As for the example with the normalized distribution of the alkyl benzene family, and assuming the gas oil fraction to be reconstructed is a 160–400°C cut, only alkyl benzenes with 9 to 24 carbon atoms can be encountered. Hence, the distribution shown in Figure 2 has to be modified to take this additional constraint into account. Consequently, the alkyl benzenes in C_6 , C_7 and C_8 are removed, while the mole fractions of the other alkyl benzenes have been corrected by means of their respective separation efficiency factor. Figure 5 shows the new renormalized molar distribution of the alkyl benzene family after including the cut point information.

2.3 Standard Molar Distributions for Paraffins and Sulfides/Mercaptans

For paraffins and sulfides/mercaptans, there is no naphtho-aromatic core and the above procedure based on the addition of side chains with a limited length distribution can no longer be used. The distillation cut points can, however, be used to determine the number of carbon atoms of the smallest and the largest pseudo-compound in the chemical family. A second reference distribution will therefore be used (Fig. 1)

and homothetically modified to lie between these two extremes.

Using the same example, for a 160–400°C cut, the smallest paraffin is the paraffin with 10 carbon atoms, while the largest one has 24 carbon atoms. The second reference distribution shown in Figure 1 therefore undergoes a translation to carbon number 10, is homothetically expanded to reach 24 carbon atoms, and is discretized. Again, separation effects close to the initial and final boiling points will arise, and are resolved in the same way as for the families with a naphtho-aromatic core by applying a separation efficiency factor. Figure 6 gives an example of the renormalized molar distribution of the paraffin family for a 160–400°C gas oil cut.

2.4 Determination of the Abundance of the Different Chemical Families in the Matrix

Once the renormalized molar distributions within each chemical family have been determined, their overall absolute abundance has to be quantified. This absolute abundance is determined from four physicochemical analyses: the mass spectrometry, the sulfur speciation, and the analyses of total and basic nitrogen.

The mass spectrometry determines the mass fractions of the hydrocarbon families directly in weight fractions. The sulfur speciation gives the mass fractions of the sulfur families expressed in weight ppm of sulfur, while the nitrogen analyses establish the mass fractions of the nitrogen in weight ppm of nitrogen. Since the sulfur and nitrogen

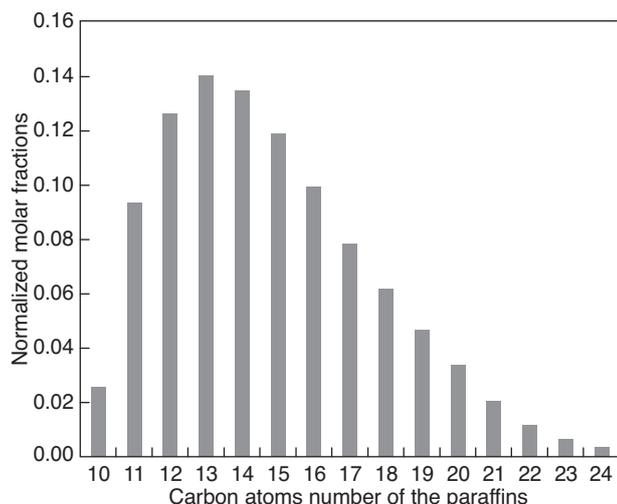


Figure 6

Example of a discrete distribution for the family of paraffins when the gas oil is a 160-400°C distillation cut (*i.e.*, after using of the separation efficiency factor).

analyses are given as the amount of the respective atoms, the mass fractions of the corresponding families in the mixture can only be determined by using the average molecular weight of the family, which is obtained from the previously calculated renormalized molar distributions.

Working with actual data showed that the four analyses are not necessarily consistent with one another due to their analytical uncertainties. Hence, they have to be renormalized so that the sum of the mass fractions of all the chemical families is equal to 1. As the analytical uncertainty of the mass spectrometry is the largest, the renormalization is only applied to the hydrocarbon families.

The overall absolute abundance of the different chemical families is now defined on a mass basis. To convert the mass fractions into molar fractions, each of the mass fractions of the chemical families is divided by their average molecular weight and then renormalized to 1.

2.5 Molar Fractions Matrix

The final step consists in distributing the overall molar fraction of each family over the pseudo-compounds they comprise according to their renormalized molar distribution. Consequently, the gas oil cut is therefore fully characterized by its matrix of molar fractions, which can now be used to determine the properties of the mixture.

By construction, the amounts of sulfur and nitrogen in the families of the matrix are strictly equal to the values obtained by the sulfur and nitrogen analyses. The mass spectrometry of the matrix is very slightly modified due to the renormalization. The average molecular weight and distillation curve

may, however, be completely false. Indeed, these properties depend strongly on the choice of the maximum length L of the side chains, which therefore needs to be determined through the minimization of an objective function.

2.6 Optimization of the Parameters of the Model

Mass spectrometry, sulfur speciation and nitrogen analyses provide chemical detail of gas oil fractions and are used to reconstruct the first dimension of the pseudo-compounds matrix. However, it is also necessary to have information concerning the second dimension of the matrix, *i.e.* the carbon atom distribution. Simulated distillation appears to be the most appropriate and widely available analysis for this purpose. As previously mentioned, the initial and final boiling points can be used as cut-off points for the carbon number distributions, while the 50 wt% point allows to determine the average molecular weight of the gas oil through API correlations, and hence reveals the length distribution of the alkyl chains.

However, as given in ASTM method D2887, the initial and final boiling points contain large analytical uncertainties. For this reason, it was decided to use the average molecular weight and the 10 wt% and 90 wt% points of the simulated distillation as the target analytical data. The parameters to be modified for this purpose are as follows:

- the maximum length L of the side chains (which strongly influences the average molecular weight of the mixture);
- the initial boiling point of the simulated distillation;
- the final boiling point of the simulated distillation.

At each iteration, the pseudo-compounds matrix is recalculated together with its mixture properties. By construction, the mass spectrometry, sulfur speciation and nitrogen speciation are always very close to the analytical data. The average molecular weight and distillation data (SimDist 10 wt% and 90 wt% points) converge towards their experimental values by modifying the parameters of the model. Upon convergence, the resulting pseudo-compounds matrix has properties that are very close to those given by the mass spectrometry, the sulfur speciation, the nitrogen analyses and the simulated distillation. This pseudo-compounds matrix is representative of the gas oil to be reconstructed. Figure 7 summarizes the overall statistical reconstruction method.

3 APPLICATION TO THE RECONSTRUCTION OF GAS OILS

3.1 Reconstruction of a Straight-Run Gas Oil

3.1.1 SR Gas Oil Analysis

The following analyses are required for the statistical reconstruction of a SR gas oil: density, simulated distillation,

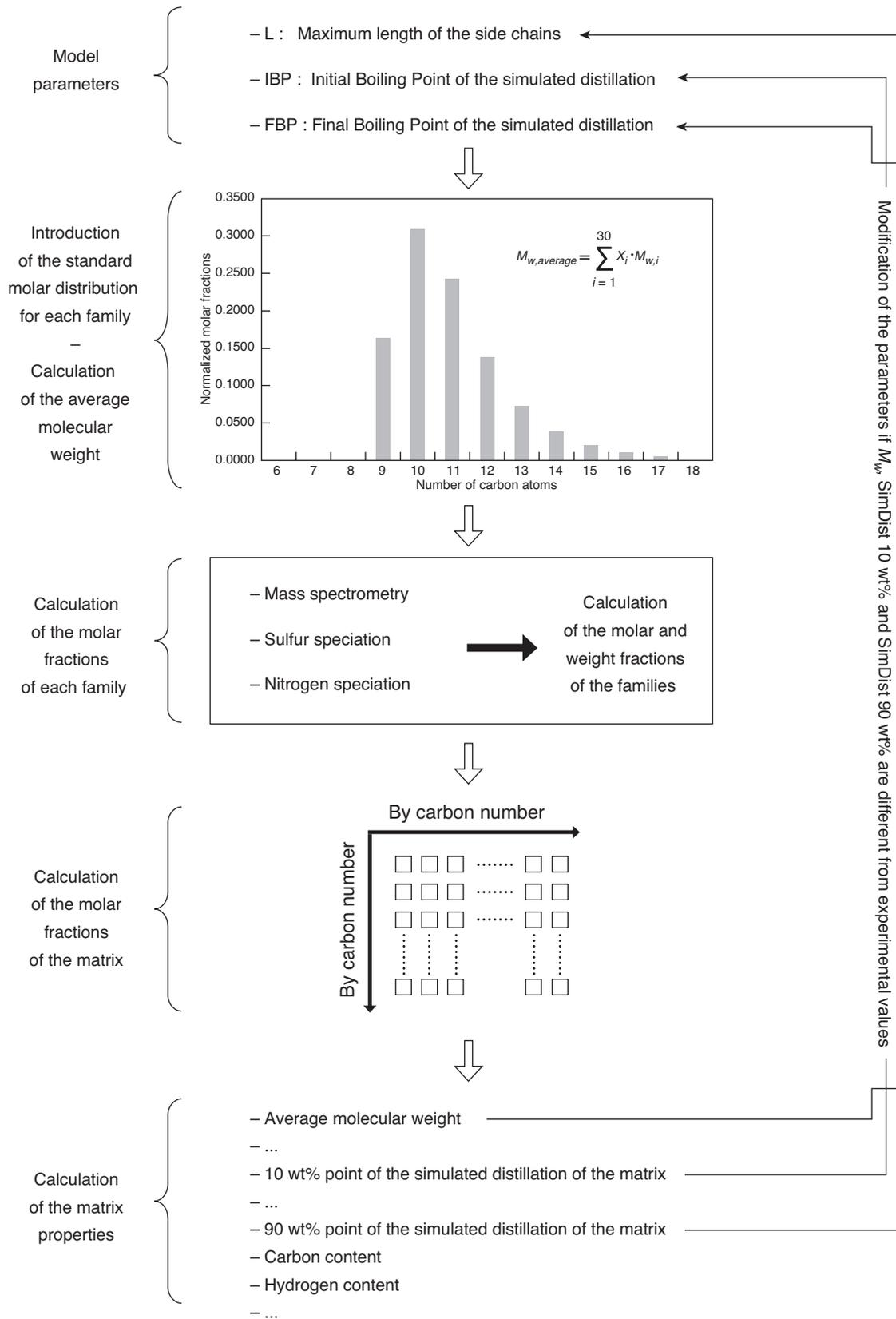


Figure 7

Diagram for the statistical reconstruction of gas oils.

TABLE 4
Analyses of a typical SRGO
and associated pseudo-compounds mixture properties

		Analytical data of the SRGO	Properties of the associated matrix
Density at 15°C	g/cm ³	0.8537	0.8548
Hydrogen content	wt%	13.25	13.32
Sulfur content	wt%	1.35	1.35
Nitrogen content	wt%	0.0127	0.0127
Basic nitrogen content	wt%	0.0047	0.0047
Aromatic carbon content	%	15.0	13.24
0 wt% (SimDist)	°C	160	216
5 wt% (SimDist)	°C	219	245
10 wt% (SimDist)	°C	245	250
20 wt% (SimDist)	°C	272	269
30 wt% (SimDist)	°C	290	284
40 wt% (SimDist)	°C	304	301
50 wt% (SimDist)	°C	317	315
60 wt% (SimDist)	°C	330	328
70 wt% (SimDist)	°C	344	342
80 wt% (SimDist)	°C	359	360
90 wt% (SimDist)	°C	376	378
95 wt% (SimDist)	°C	386	394
100 wt% (SimDist)	°C	404	447
Paraffins	wt%	41.9	40.63
Monocycloparaffins	wt%	17.3	16.77
Dicycloparaffins	wt%	8.5	8.24
Tricycloparaffins	wt%	2.1	2.04
Alkylbenzenes	wt%	7.9	7.66
Indanes/Tetralins	wt%	3.5	3.39
Indenes/Dinaphtenobenzenes	wt%	1.4	1.36
Naphthalenes	wt%	4.8	4.65
Acenaphtenes	wt%	2.0	1.94
Acenaphtylenes	wt%	1.3	1.26
Phenanthrenes/Anthracenes	wt%	1.3	1.26
Polyaromatics	wt%	0.0	0.00
Sulfides/Mercaptans	wt%		5.74
Thiophenes	wt%		0.00
Benzothiophenes	wt%	5.0	1.42
Dibenzothiophenes	wt%	3.0	3.43
Anilines/Pyridines	wt%		0.00
Quinoleines	wt%		0.07
Acridines	wt%		0.00
Pyrroles	wt%		0.00
Indoles	wt%		0.06
Carbazoles	wt%		0.07
Sulfures/Mercaptans	wt ppm S	7226	7226
Thiophenes	wt ppm S	0	0
Benzothiophenes	wt ppm S	2081	2081
Dibenzothiophenes	wt ppm S	937	937
4-Dibenzothiophenes	wt ppm S	1821	1821
4,6-Dibenzothiophenes	wt ppm S	1435	1435
Anilines	wt ppm N		0
Pyridines	wt ppm N		0
Quinolines	wt ppm N		47
Acridines	wt ppm N		0
Pyrroles	wt ppm N		0
Indoles	wt ppm N		40
Carbazoles	wt ppm N		40

mass spectrometry, sulfur speciation and nitrogen analysis. These data are listed in Table 4.

3.1.2 Optimization of Parameters

Using the simulated distillation and d_4^{15} density, API correlations are employed to estimate the average molecular weight of the gas oil fuel: a value of 242.5 g/mol was obtained in this case. The reconstruction then consists in defining the following 3 parameters:

- maximum length of the side chains L ;
 - initial boiling point of the simulated distillation;
 - final boiling point of the simulated distillation;
- so as to obtain a pseudo-compound mixture with the following properties:
- average molecular weight: 242.5 ± 5.0 g/mol;
 - 10 wt% boiling point of the simulated distillation: $245 \pm 5.0^\circ\text{C}$;
 - 90 wt% boiling point of the simulated distillation: $376 \pm 5.0^\circ\text{C}$.

Upon convergence, the reconstructed matrix has a molecular weight of 240.5 g/mol, a 10 wt% SimDist point of 250°C and a 90 wt% SimDist point of 378°C , which are all within the required precision. The following parameter values were obtained for this matrix: the maximum length L of the alkyl chains is 4.0, the initial boiling point of the SimDist is 215.6°C and the final boiling point 447.4°C .

3.1.3 Properties of the Pseudo-Compounds Matrix

Once the gas oil has been reconstructed in the form of the pseudo-compounds matrix with its molar fractions, the mixture properties can be calculated and compared to the

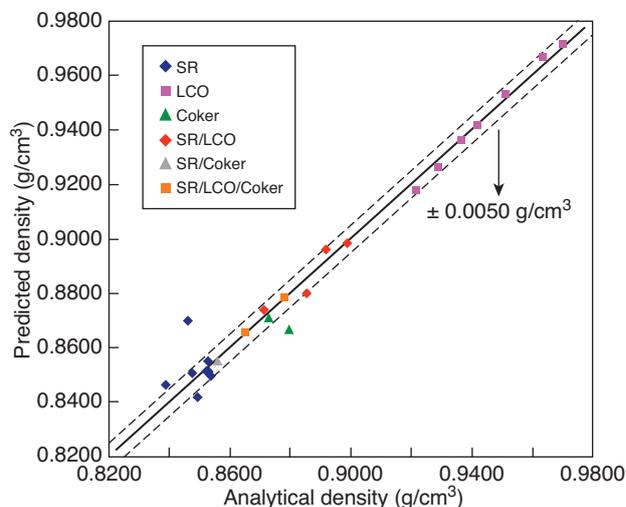


Figure 8

Parity plot for the prediction of the density at 15°C of gas oil cuts (full line: parity line; dotted lines: ± 0.0050 g/cm³).

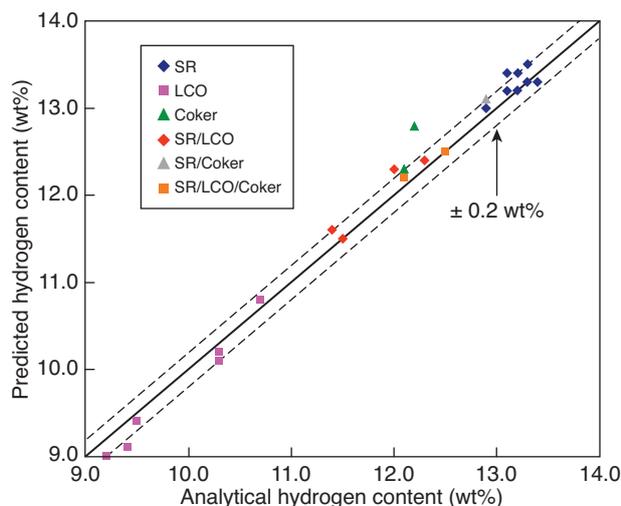


Figure 9

Parity plot for the prediction of the hydrogen content of gas oil cuts (full line: parity line; dotted lines: ± 0.2 wt%).

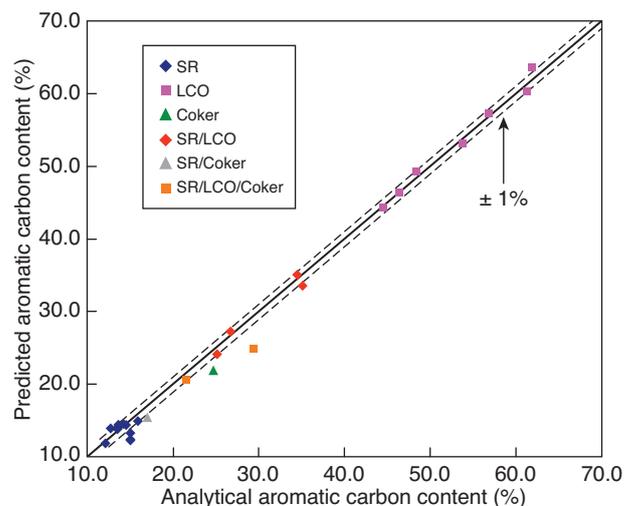


Figure 10

Parity plot for the prediction of the aromatic carbon content of gas oil cuts (full line: parity line; dotted lines: $\pm 1\%$).

analytical data. This comparison is shown in Table 4. The following observations can be made:

- the sulfur speciation of the reconstructed mixture is identical to the analytical data. This is quite normal as these are the input data that are used for the statistical reconstruction;
- the mass spectrometry of the mixture is very close to the analytical mass spectrometry except for the BT and DBT families. With the other families, there are slight differences due to the renormalization;
- the simulated distillation of the mixture is very close to the experimental distillation curve between 10 wt% and 90 wt%. There are, however, more significant differences near the initial boiling point, and even more importantly near the final boiling point, due to the cut point management method, which still needs to be further improved;
- the unused analyses can be predicted from the reconstructed mixture composition, such as the hydrogen content or the aromatic carbon content. As can be noticed, both values are well predicted.

3.2 Application to Other Types of Gas Oils

The statistical reconstruction method illustrated for the SR gas oil can be applied for any type of gas oil. To test the validity of this reconstruction method, the following gas oils were analyzed in detail and reconstructed:

- 8 Straight-Run (SR) gas oils;
- 7 Light Cycle Oil (LCO) gas oils;
- 2 Coker Gas Oils (CGO);
- 7 mixtures of SR, LCO and/or CGO fractions.

After reconstruction, the density, hydrogen content and aromatic carbon content of the associated matrices were calculated and compared to the analytical data. The comparison is shown in the form of parity plots (Fig. 8, 9 and 10). Excellent results were obtained and confirm that the gas oils were successfully represented by their pseudo-compounds matrices.

CONCLUSION

In this work, a novel reconstruction method has been developed that allows to create a detailed characterization of Straight Run gas oils, Light Cycle Oils and Coker Gas Oils starting from the 5 following analyses: mass spectrometry, sulfur speciation, nitrogen analyses, density and simulated distillation. The technique reconstructs a pseudo-compounds matrix with the molar fractions of 612 pseudo-compounds, covering 28 different chemical families over a range of carbon atoms going from 1 to 30. This reconstruction method, termed statistical reconstruction, is based on an assumed statistical distribution for the length of the side chains. Compared to the previously developed reconstruction technique, the statistical reconstruction is more rapid and more precise than the stochastic reconstruction methods [10, 23, 24, 31-33], while being more robust than the entropy maximization algorithm [23-26] or the combined approach [23, 24, 34, 35]. However, the statistical reconstruction is less flexible because it needs necessarily all the information given by the simulated distillation, Fitzgerald mass spectrometry, gas chromatography analysis with sulfur chemiluminescence detection, and nitrogen analyses.

The properties of each pseudo-compound can be calculated either directly or by group contribution methods, enabling the creation of matrices for each property, e.g. the molecular weight or the normal boiling point. The major advantage of this approach is that these property matrices are calculated once and for all and not for each gas oil, which would be the case when one uses a much coarser lumping scheme. Furthermore, mixing rules were derived based on simplifying assumptions in order to determine the mixture properties directly from the properties of the pseudo-compounds and from the matrix composition of each gas oil.

The reliability and representativeness of the reconstruction technique was confirmed by comparing the properties of the reconstructed matrices with existing gas oil analyses. The results also illustrate the potential of the method, as it was shown that unused analyses, e.g. the hydrogen content or the aromatic carbon content, can be predicted from the pseudo-compounds matrix.

Concerning the distillation analyses, there is still some room for improvement, since the reconstruction is not perfect near the initial and final boiling points, which may eventually pose problems when calculating the gas-liquid repartition, for example. The separation efficiency criterion η would need to be improved in order to better use the information provided by simulated distillation. This will be dealt with in a later article. It would also be useful to develop new contribution methods in order to directly estimate other major gas oil properties, such as the cetane index or viscosity.

Finally, this novel reconstruction method to characterize gas oils is also used as input for predictive kinetic and process models for hydrotreating or hydrocracking, which will be presented elsewhere [52].

REFERENCES

- Kmak W.S. (1971) A Kinetic Simulation Model of the Powerforming Process, *AICHE National Meeting*, Houston, TX, March 3, 1971.
- Ramage M.P., Graziani K.R., Krambeck F.J. (1980) Development of Mobil's Reforming Model, *Chem. Eng. Sci.* **35**, 1, 41-48.
- Marin G.B., Froment G.F., Lerou J.J., De Backer W. (1983) Simulation of a Catalytic Naphtha Reforming Unit, *Proceedings of the Third International Congress on Computers & Chemical Engineering* **27**, 2, C117/1-C117/7.
- Weekman V.W., Nace D.M. (1970) Kinetics of catalytic cracking selectivity in fixed, moving, and fluid bed reactors, *AICHE J.* **16**, 3, 397-404.
- Jacob S.M., Gross B., Voltz S.E., Weekman V.W. (1976) A Lumping and Reaction Scheme for Catalytic Cracking, *AICHE J.* **22**, 4, 701-713.
- Pitault I., Nevicato D., Forissier M., Bernard J.R. (1994) Kinetic Model Based on a Molecular Description for Catalytic Cracking of Vacuum Gas Oil, *Chem. Eng. Sci.* **49**, 24A, 4249-4262.
- Stangeland B.E. (1974) A Kinetic Model for the Prediction of Hydrocracker Yields, *Ind. Eng. Chem. Process Design Dev.* **13**, 1, 71-76.
- Koseoglu R.O., Phillips C.R. (1988) Kinetic Models for the Non-Catalytic Hydrocracking of Athabasca Bitumen, *Fuel* **67**, 6, 906-915.
- Parnas R.S., Allen D.T. (1988) Compound Class Modeling of Hydropyrolysis, *Chem. Eng. Sci.* **43**, 10, 2845-2857.
- Trauth D.M. (1993) Structure of Complex Mixtures through Characterization, Reaction, and Modeling, *PhD Thesis*, University of Delaware.
- Mosby J.F., Buttke R.D., Cox J.A., Nikolaides C. (1986) Process Characterization of Expanded-Bed Reactors in Series, *Chem. Eng. Sci.* **41**, 4, 989-995.
- Gray M.R. (1990) Lumped Kinetics of Structural Groups: Hydrotreating of Heavy Distillate, *Ind. Eng. Chem. Res.* **29**, 4, 505-512.
- López-García C., Roy-Auberger M., Chapus T., Baco F. (2003) Analysis and Kinetic Modeling in ULSD Hydrotreating, *Preprint Paper of the American Chemical Society's Division of Fuel Chemistry* **48**, 2, 681-682.
- Schweitzer J.M., Kressmann S. (2004) Ebullated bed reactor modeling for residue conversion, *Chem. Eng. Sci.* **59**, 22-23, 5637-5645.
- Wang F.C.Y., Robbins W.K., Di Sanzo F.P., McElroy F.C. (2003) Speciation of Sulfur-Containing Compounds in Diesel by Comprehensive Two-Dimensional Gas Chromatography, *J. Chromatogr. Sci.* **41**, 10, 519-523.
- Dartiguelongue C., Hudebine D., Bertoncini F., Lopez-Garcia C., Chapus T. (2006) Comparison of experimental and modelled data for sulfur molecular distribution in diesel feeds from various origins, *Preprint Paper of the American Chemical Society's Division of Petroleum Chemistry* **51**, 2, 299-302.
- Adam F., Bertoncini F., Brodusch N., Durand E., Thiébaud D., Espinat D., Hennion M.C. (2007) New benchmark for basic and neutral nitrogen compounds speciation in middle distillates using comprehensive two-dimensional gas chromatography, *J. Chromatogr. A* **1148**, 1, 55-64.
- López García C., Becchi M., Grenier-Loustalot M.F., Pâisse O., Szymanski R. (2002) Analysis of aromatic sulfur compounds in gas oils using GC with sulfur chemiluminescence detection and high-resolution MS, *Anal. Chem.* **74**, 15, 3849-3857.
- Liguras D.K., Allen, D.T. (1989) Structural Models for Catalytic Cracking 1. Model Compound Reactions, *Ind. Eng. Chem. Res.* **28**, 6, 665-673.
- Liguras D.K., Allen, D.T. (1989) Structural Models for Catalytic Cracking 2. Model Compound Reactions, *Ind. Eng. Chem. Res.* **28**, 6, 674-683.
- Allen D.T., Liguras D.K. (1991) Structural Models of Catalytic Cracking Chemistry: A Case Study of a Group Contribution Approach to Lumped Kinetic Modeling, *Chemical Reactions in Complex Mixtures: The Mobil Workshop*, Sapre A.V., Krambeck F.J. (eds), Van Nostrand Reinhold, New-York.
- Khorasheh F., Khaledi R., Gray M.R. (1998) Computer generation of representative molecules for heavy hydrocarbon mixtures, *Fuel* **77**, 4, 241-253.
- Hudebine D. (2003) Reconstruction moléculaire de coupes pétrolières, *PhD Thesis*, École Normale Supérieure de Lyon.
- Hudebine D., Vera, C., Wahl F., Verstraete J. (2002) Molecular Representation of Hydrocarbon Mixtures from Overall Petroleum Analyses, *2002 AIChE Spring Meeting*, New Orleans, LA, March 10-14, 2002, Paper 27a, 15-24.
- Van Geem K.M., Hudebine D., Reyniers M.F., Wahl F., Verstraete J.J., Marin G.B. (2007) Molecular reconstruction of naphtha steam cracking feedstocks based on commercial indices, *Comput. Chem. Eng.* **31**, 9, 1020-1034.

- 26 Hudebine D., Verstraete J.J. (2009) Reconstruction of Petroleum Feedstocks by Entropy Maximization. Application to the FCC gasolines, *Oil Gas Sci. Technol. – Rev. IFP* (in press – this volume).
- 27 Quann R.J., Jaffe S.B. (1992) Structure-Oriented Lumping: Describing the Chemistry of Complex Hydrocarbon Mixtures, *Ind. Eng. Chem. Res.* **31**, 11, 2483-2497.
- 28 Quann R.J., Jaffe S.B. (1996) Building Useful Models of Complex Reaction Systems in Petroleum Refining, *Chem. Eng. Sci.* **51**, 10, 1615-1635.
- 29 Jaffe S.B., Freund H., Olmstead W.N. (2005) Extension of Structure-Oriented Lumping to Vacuum Residua, *Ind. Eng. Chem. Res.* **44**, 9840-9852.
- 30 Zhang Y. (1999) A Molecular Approach for Characterization and Property Predictions of Petroleum Mixtures with Applications to Refinery Modelling, *PhD Thesis*, University of Manchester.
- 31 Neurock M. (1992) A Computational Chemical Reaction Engineering Analysis of Complex Heavy Hydrocarbon Reaction Systems, *PhD Thesis*, University of Delaware.
- 32 Neurock M., Nigam A., Trauth D.M., Klein M.T. (1994) Molecular Representation of Complex Hydrocarbon Feedstocks through Efficient Characterization and Stochastic Algorithms, *Chem. Eng. Sci.* **49**, 24, 4153-4177.
- 33 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.
- 34 Hudebine D., Verstraete J.J. (2004) Molecular Reconstruction of LCO Gasoils from Overall Petroleum Analyses, *Chem. Eng. Sci.* **59**, 22-23, 4755-4763.
- 35 Verstraete J.J., Revellin N., Dulot H., Hudebine D. (2004) Molecular reconstruction of vacuum gasoils, *Preprint Paper of the American Chemical Society's Division of Fuel Chemistry* **49**, 1, 20-21
- 36 Wei J., Kuo J.C.W (1969) Lumping Analysis in Monomolecular Reaction Systems. Analysis of the Exactly Lumpable System, *Ind. Eng. Chem. Fund.* **8**, 1, 114-123.
- 37 Li G., Rabitz H.A. (1989) A General Analysis of Exact Lumping in Chemical Kinetics, *Chem. Eng. Sci.* **44**, 6, 1413-1430.
- 38 Fafet A., Magné-Drisch J. (1995) Analyse quantitative détaillée des distillats moyens par couplage CG/MS, *Oil Gas Sci. Technol. – Rev. IFP* **50**, 3, 391-404.
- 39 Castex H., Boulet R., Juguin J., Lepinasse A. (1983) Analysis of Kerosenes and Middle Distillates by Medium-Resolution Mass Spectrometry, *Oil Gas Sci. Technol. – Rev. IFP* **38**, 4, 523-532.
- 40 Fafet A., Bonnard J., Prigent F. (1999) New Developments in Mass Spectrometry for Group-Type Analysis of Petroleum Cuts. First Part: Improving Quantification of Sulphured Aromatic Compounds in Middle Distillates, *Oil Gas Sci. Technol. – Rev. IFP* **54**, 4, 439-452.
- 41 Fafet A., Bonnard J., Prigent F. (1999) New Developments in Mass Spectrometry for Group-Type Analysis of Petroleum Cuts. Second Part: Development and Validation of a New Inlet System for Heavy Cuts, *Oil Gas Sci. Technol. – Rev. IFP* **54**, 4, 453-462.
- 42 Benson S.W. (1968) *Thermochemical Kinetics*, J. Wiley and Sons, New York.
- 43 Benson S.W., Buss J.H. (1958) Additivity Rules for the Estimation of Molecular Properties. Thermodynamic Properties, *J. Chem. Phys.* **29**, 546-572.
- 44 Benson S.W., Cruickshank F.R., Golden D.M., Haugen G.R., O'Neal H.E., Rodgers A.S., Shaw R., Walsh R. (1969) Additivity Rules for the Estimation of Thermochemical Properties, *Chem. Rev.* **69**, 279-324.
- 45 Benson S.W. (1976) *Thermochemical Kinetics*, Second edition, J. Wiley and Sons, New York.
- 46 Joback K.G., Reid R.C. (1987) Estimation of Pure-Component Properties from Group-Contributions, *Chem. Eng. Commun.* **57**, 1-6, 233-243.
- 47 Constantinou L., Gani R. (1994) New Group Contribution Method for Estimating Properties of Pure Compounds, *AIChE J.* **40**, 10, 1697-1710.
- 48 Thermodynamic Research Center (1999) TRC Thermodynamic Tables, *Thermodynamics Research Center*, The Texas A&M University System, College Station, TX.
- 49 Marrero-Morejon J., Pardillo-Fontdevila E. (1999) Estimation of Pure Compound Properties Using Group-Interaction Contributions, *AIChE J.* **45**, 3, 615-621.
- 50 Reid R.C., Prausnitz J.M., Poling B.E. (1987) *The Properties of Gases and Liquids*, Fourth Edition, McGraw-Hill Book Co., New York.
- 51 Yamada T., Gunn R.D. (1973) Saturated Liquid Molar Volumes. Rackett Equation, *J. Chem. Eng. Data* **18**, 2, 234-236.
- 52 López García C., Hudebine D., Schweitzer J.M., Verstraete J.J., Ferré D. (2009) In-Depth Modeling of Gas Oil Hydrotreating: From Feedstock Reconstruction to Reactor Stability Analysis, *Catal. Today* (in press).

Final manuscript received in July 2009
Published online in November 2009

Copyright © 2009 Institut français du pétrole

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than IFP must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee: Request permission from Documentation, Institut français du pétrole, fax. +33 1 47 52 70 78, or revueogst@ifp.fr.