MOLECULAR RECONSTRUCTION OF VACUUM RESIDUES

Jan J. Verstraete, Hugues Dulot and Damien Hudebine

IFP Lyon - Process Modeling and Design Division
PO Box 3, 69390 Vernaison, France

Introduction

In order to gain insight in the composition of petroleum and for modeling purposes, it is vital to have a representation of the molecular structure of all petroleum fractions. Unfortunately, for petroleum cuts boiling above the naphtha range (gas oils, residues, …), a molecular description can no longer be obtained directly, even by means of cutting-edge analytical techniques.

To get around this lack of molecular detail, a detailed description needs to be "reconstructed" from more global analyses. This process of estimating the composition in terms of individual compounds of a feedstock from overall characterization data of the mixture is referred to as "molecular reconstruction". The basic idea behind these molecular reconstruction algorithms is to generate a discrete set of molecules whose properties mimic the properties of the petroleum cut to be represented.

Several molecular reconstruction techniques have been described in the literature. Liguras and Allen[1] used a predefined set of key components and numerically adjusted their molar fractions in order to obtain a mixture that closely resembles the input analytical data. Unfortunately, this approach is based on an *a priori* definition of the components, while the authors adjust the molecular abundance of each compound based on a large amount of analytical data, which are not necessarily available to refiners. To get around these limitations, Neurock et al.[2] developed a method termed "stochastic reconstruction". To get to a description on a molecular level, a set of distributions of structural attributes was first defined and then sampled by a Monte Carlo algorithm to produce an equimolar mixture of molecules. When coupled to an optimization loop, the method has been proven able to yield synthetic mixtures that closely reproduce the properties of heavy asphaltene fractions[3].

During previous work[4-8], two different algorithms were developed to generate a complex mixture of molecules from standard petroleum analyses: a stochastic reconstruction technique and a reconstruction method based on information entropy maximization. These algorithms and their variations were applied to naphthas[5,8], to Light Cycle Oils[4], and to Vacuum Gas Oils[7]. In the present work, a two-step reconstruction algorithm combining both approaches will be applied to the molecular reconstruction of vacuum residues.

General description of the two-step reconstruction algorithm

The proposed algorithm consists of two distinct steps. After defining a molecule construction scheme tailored to the specific petroleum cut, a large set of molecules is generated via a stochastic reconstruction method similar to the one described by Neurock et al.[2]. This set of molecules is then iteratively improved upon until it yields an equimolar mixture whose properties are close to that of the petroleum fraction. The representativeness of this mixture is improved in the second step of the algorithm by modifying the molar fractions of the various molecules in the set by means of an information entropy maximization method.

Generation of the initial set of molecules. The first step aims at creating a set of molecules that are typical of the petroleum cut to be represented. In our algorithm, a stochastic reconstruction method is used to generate a large set of molecules that are assembled from structural attributes (polycyclic cores, rings, substituents, chains, …). To create a molecule, the type and number of these attributes are selected by randomly sampling a set of parametric distributions of these building blocks via a Monte Carlo approach. The sequence in which the various attributes are sampled is defined by a decision tree, termed "construction scheme". The actual creation of a molecule from the selected structural attributes is based on "building rules" that define how to correctly assemble the various building blocks on a constrained hexagonal map, avoiding the creation of unfeasible molecules. At the same time, the pure component properties of each molecule are calculated from their structure, either directly by inspection (e.g. chemical formula, molecular weight, NMR signature, ...) or numerically by group contribution methods (e.g. density, boiling point, polarity, ...). To obtain a representative equimolar mixture of N molecules, the molecule construction procedure is repeated N times. The properties of this synthetic mixture are subsequently obtained via mixing rules and compared to the available analytical data of the petroleum fraction (elemental analysis, density, average molecular weight, ^{1}H NMR, …) by means of an objective function that quantifies the deviation between the experimental and simulated data. An elitist genetic algorithm finally modifies the parameters of the distributions for the structural attributes to minimize this objective function. A more detailed description of the algorithm can be found elsewhere[5,6]. This stochastic reconstruction step allows to successively modify the generated set of molecules until a synthetic mixture is obtained whose properties approach those of the petroleum cut to be represented.

Determination of molecular abundance in the molecule set. The set of molecules that has been generated in the stochastic reconstruction step only yields the initial representation of the petroleum fraction. In order to improve the representativeness of the set of molecules, the molecular abundance of these molecules is modified by adjusting their molar fractions based on the available analytical data. In our algorithm, this is performed by maximizing the information entropy. This criterion ensures that, in absence of information, the distribution of the set of molecules will remain uniform. The introduction of constraints, i.e. analytical data for the petroleum cut to be represented, will distort the uniform distribution of the set in order to match this information, while imposing non-negative mole fractions, similarly to the approach proposed by Jaffe et al.[9]. As for the stochastic reconstruction step, the deviation between the experimental and simulated data is assessed through an objective function, which is minimized by means of a conjugate gradient optimization technique. For a more detailed description of the algorithm, the reader is referred to the literature[4,5,6,8].

Application to vacuum residues

As mentioned above, an appropriate molecule construction scheme needs to be defined to reconstruct vacuum residues. This requires specifying the type and distribution of each structural attribute, and defining the construction scheme and building rules.

The choice of the structural attributes and of their distributions has been guided by literature information of the chemical nature of these heavy cuts, but care has been taken to minimize the number of structural attributes. For the residue fractions, the molecules have been divided in 4 types: paraffins, naphthenes, aromatics/resins, or asphaltenes. A first distribution defines the type of molecule to be constructed. For the paraffins, their chain length is determined by a gamma distribution. For non-paraffins, the number of polycyclic cores per molecule was limited to one for naphthenes, aromatics and resins, while a gamma distribution was used for the number of polycyclic cores per asphaltene molecule. For each polycyclic core, the total number of cycles per core is also characterized by a gamma distribution. Possible cycles that are considered in the stochastic

reconstruction step are benzenes, thiophenes, pyridines, pyrroles, furanes and cyclohexanes. The length of the side-chains is defined by an exponential distribution, while histograms are used to define the distributions for the presence of sulfur, nitrogen and oxygen atoms in these side-chains. The complete list of distributions that are used to build a residue molecule is given in Table 1.

Table 1. Description of the Structural Attributes Used

Structural attribute	Distribution	Values	Par.
1. Molecule type *(Paraf., Napht., Aro./Res., Asph.)*	Histogram	0 .. 3	3
2. Number of polycyclic cores in an asphaltene	Gamma	> 1	2
2. Number of cycles per core	Gamma	> 1	2
4. Number of benzenes	Gamma	> 0	2
5. Type of hetero-cycles *(thioph., pyrid., pyrrole, furane)*	Histogram	0 .. 3	3
6. Number of thiophenes	Histogram	1 or 2	1
7. Number of pyridines	Histogram	1 or 2	1
8. Number of pyrroles	Histogram	1 or 2	1
9. Number of furanes	Histogram	1 or 2	1
10. Alkyl chain presence on peripheral carbon atoms	Histogram	0 or 1	1
11. Length of the paraffins	Gamma	> 0	2
12. Length of an alkyl chain *(lateral / inter-core)*	Exponential	> 0	1
13. Aro/Res: Sulfur substitution for aliphatic CH_2	Histogram	0 or 1	1
14. Aro/Res: Other atom substitution for aliphatic CH_3 / CH_2	Histogram	0 or 1	1
15. Asph: Sulfur substitution for aliphatic CH_2	Histogram	0 or 1	1
16. Asph: Other atom substitution for aliphatic CH_3 / CH_2	Histogram	0 or 1	1
17. Type of atom substitution *(oxygen / nitrogen)*	Histogram	0 or 1	1
18. Oxygen structure *(alcohol, ketone, carboxylic acid)*	Histogram	0 .. 2	2
19. Hydroxyl substitution on an aromatic H	Histogram	0 or 1	1

As explained above, the construction scheme defines the sequence in which the distributions are sampled. The distribution defining the type of molecule is sampled first. If the molecule to be constructed is a paraffin, distribution 11 is used to determine its chain length. If the molecule to be constructed is a naphthene, distribution 3 is sampled to define the total number of cyclohexanes in the core. Once the various structure elements are assembled into a polycyclic core, each peripheral carbon atom is tested to check whether an alkyl chain should be inserted or not (distribution 10), while the length of each chain is determined from distribution 12. For the aromatics and resin molecules, the size of its core is sampled from distribution 3, while the number of benzenes, and the type and number of thiophene, pyridine, pyrrole, and furane cycles are defined through distributions 3 to 9. As for the naphthenes, distribution 10 is used for each peripheral carbon atom to check whether a side chain should be present, while distribution 12 is used to determine the length of each chain. The type and position of the hetero-atoms on the side-chains are then defined by distributions 13, 14, 17, and 18, while distribution 19 is used to add hydroxyl substituents to unsubstituted aromatic carbon atoms. For the asphaltenes, an archipelago representation was chosen. To construct an asphaltene molecule, several polycyclic cores are constructed in an analogous way as for the aromatics and resins. The number of polycyclic cores to be constructed is defined by distribution 2, while the length of the alkyl chains between the cores is sampled from distribution 12.

With this molecular construction scheme, the stochastic reconstruction step was used to generate an initial set of 5000 molecules while adjusting the parameters of the distributions to minimize the objective function. The latter contained the differences between the experimental and calculated values for the elemental analysis, the average molecular weight, the ^{1}H NMR signature, the SARA analysis and the partial Simulated Distillation curve for an Arabian Light vacuum residue. As can be seen in Table 2, the properties of the initial equimolar mixture of molecules obtained at the end of the stochastic reconstruction step are already close to the corresponding experimental values. The elemental analysis was well predicted, although there was a surplus of oxygen. For the ^{1}H NMR analysis, the resulting aromatic hydrogen content was too low and the aliphatic structures on the molecules are probably too large. The molecular weight was overestimated, however. Some differences also exist in the simulated distillation curve. Lastly, the recalculated SARA analysis correlates relatively well with the experimental data. More particularly, good results were obtained in the distribution of the molecules between the aromatic and resin families, for which no adjustable parameter was provided. In conclusion, this set of 5000 molecules obtained by stochastic reconstruction step can be considered as representative of an Arabian Light vacuum residue.

The entropy maximization step of the algorithm adjusted the molar fractions of the various molecules to further minimize the differences between experimental and calculated values in the objective function. The properties of the resulting final mixture are very close to the corresponding experimental values, indicating that this mixture is a good representation of the actual vacuum residue.

Table 2. Comparison of the Properties of the Arabian Light Vacuum Residue and of the Corresponding Molecular Set

	Exp.	Initial	Final
Simulated Distillation (ASTM D2887)			
0% (°C)	439	425	425
5% (°C)	536	504	532
10% (°C)	556	541	554
20% (°C)	583	574	581
Elemental analysis			
Carbon (wt%)	84.7	84.4	84.7
Hydrogen (wt%)	10.2	9.9	10.2
Sulfur (wt%)	4.1	4.2	4.1
Oxygen (wt%)	0.7	1.2	0.7
Nitrogen (wt%)	0.29	0.27	0.29
^{1}H NMR analysis			
Aromatic H (at%)	7.0	4.3	6.9
Alpha H (at%)	11.0	12.6	11.0
Beta H (at%)	60.8	63.1	60.9
Gamma H (at%)	21.2	20.0	21.2
Liquid chromatography SARA analysis			
Saturates (wt%)	16.3	16.8	16.3
Aromatics (wt%)	58.7	57.0	58.7
Resins (wt%)	18.9	20.9	18.9
Asphaltenes (wt%)	6.1	5.4	6.1
Other analysis			
Average molecular weight (g/mol)	842	1017	997

Conclusions

A molecular reconstruction algorithm was developed for vacuum residues. The proposed two-step algorithm first generates an initial equimolar set containing a large number of molecules in a stochastic reconstruction step. The representativeness of this mixture is subsequently refined by modifying the molar fractions of the various molecules in an information entropy maximization step. Applied to an Arabian Light vacuum residue, the properties of the resulting set of molecules are very close to actual vacuum residue to be represented, indicating that this molecular description can now be used to gain insight in the composition of this fraction or as input to a detailed reaction model for hydrocarbon conversion processes.

Acknowledgements. The authors wish to acknowledge Ph. Schnongs for his contribution to this work.

References

(1) Liguras, D.K.; Allen, D.T.; *Ind. Eng. Chem. Res., 28* (6), 674-683; **1989**
(2) Neurock, M.; Nigam, A.; Trauth, D.; Klein, M.T.; *Chem. Eng Sci., 49* (24A), 4153-4177; **1994**
(3) Trauth, D.M.; Stark, S.M.; Petti, T.F.; Neurock, M.; Klein, M.T.; *Energy & Fuels*, 8 (3), 576-580, **1994**
(4) Hudebine, D.; Vera, C.; Wahl, F.; Verstraete, J.J.; 2002 AIChE Spring Meeting, New Orleans, LA, March, 10 – 14, **2002**
(5) Hudebine, D.; Ph.D. thesis, Ecole Normale Supérieure - Lyon, **2003**
(6) Hudebine, D.; Verstraete, J.J.; *Chem. Eng. Sci.*, 59, 4755–4763, **2004**
(7) Verstraete, J.J.; Revellin, N.; Dulot H.; Hudebine, D. ; *Prepr. Pap.-Am. Chem. Soc., Div. Fuel Chem.*, 49 (1), 20-21 , **2004**.

(8) Van Geem, K.M.; Hudebine, D.; Reyniers, M.F.; Wahl, F.; Verstraete J.J.; Marin, G.B.; *Comp. Chem. Eng.*, 31, 1020–1034, **2007**
(9) Jaffe, S.B.; Freund, H.; Olmstead, W.N.; *Ind. Eng. Chem. Res.*, 44, 9840-9852, **2005**