Numerical reconstruction of Residue hydrotreating effluents: optimization of stochastic building diagrams

Miguel Miranda[1], Tiago Sozinho[2], Jan Verstraete[2], Filipe Gama Freire[1]

1 Instituto Superior Técnico
2 IFP Energies Nouvelles

The hydrotreating process allows to convert petroleum heavy fractions into lighter products, valuing the residue, while impurities such as the heteroatoms, nitrogen, sulphur and oxygen are removed. The residue, and thus the hydrotreating effluents are very complex mixtures composed by a huge quantity and variety of chemical species, usually characterized by global analysis. However, this information is not enough to develop a kinetic model, therefore analysis at thermidor lumps level are required but, in order to obtain it, a long and costly experimental process is required.

To overcome the difficulties mentioned before, a two-step algorithm, SR-REM, was applied to reconstruct hydrotreating effluents using numerical methods, by means of global analytical properties generating a set of molecules that are consistent with the experimental data and is further applied to other mass balance points of the hydrotreating unit, avoiding extra laboratorial work.

Thus, in the first step of the algorithm an equimolar set of molecules is obtained by sampling molecular structural blocks in a stochastic way following certain rules, the building diagrams. Once the molecules are created a reconstruction by entropy maximization is applied in order to adjust the molar fractions. The methodology described was tested in previous work, showing promising results. However, the first step of the algorithm, the SR step, had some flaws when it aimed to predict some of the thermidor lumps heteroatoms. In the present work, an optimization to the SR step was made, tested to an hydrotreating unit and the results compared to the ones obtained before.

Keywords: Hydrotreating effluents, Molecular Reconstruction, Vacuum Residues, Entropy Maximization, Stochastic Algorithm, Monte Carlo

Introduction

Petroleum , also known as crude oil, is a complex mixture of gaseous, liquid and solid compounds, comprising millions of naturally formed molecules, essentially composed by hydrogen and carbon, as well as significant presence of hetero atoms such as nitrogen, sulphur, oxygen and trace amounts of nickel and vanadium (1).

Nowadays, the world demand for high-quality low boiling products such as gasoline and diesel has been increasing, at the same time the availability of high-quality crude reserves (low sulfur content) is decreasing, meaning that the quantity of natural resources with characteristics that would allow a higher production of lighter distillates is scarce. In consequence, refining these types of crude will yield a higher fraction of heavier distillates with low market value (2).

To convert this low market value products in high value ones, refiners have to be equipped with processes like hydrotreating, that need continuous optimization. Thus, a key factor to improve conversion processes performance is the development of reliable and accurate kinetic models, however, to create such tools the

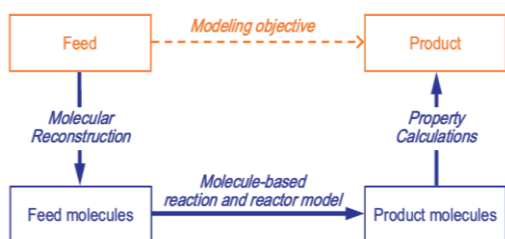feedstock's composition should be known in detail.

Figure 1 Process modelling technology adapted from (1).

At the present time, analytical tools are not able to provide such detailed molecular information. To overcome it, the proposed solution is the numeric reconstruction of heavier cuts using partial analytical data, chemical knowledge and model assumptions.

Several authors suggested new methods like Quann and Jaffe, with a called structure-oriented lumping (SOL) in which hydrocarbon carbon molecules are represented as vectors of 22 structure blocks (3). A method proposed by Zhang and Peng tried an approach named molecular type and homologous series (MTHS) similar to SOL using structural groupings to characterize petroleum mixtures, that will be conceived as matrix (4). Pyl proposed a model which is composed by a matrix generated from a composition modelling editor based on a library of molecules and a range of ebullition points. The author suggests using probability functions that describe chemical families as well as the distribution of carbon in each family (5). All the methods mentioned before are based in a library of molecules, in other words, a database that requires extensive laboratorial work, contradictory from what is pretended.

Neurock suggested a stochastic methodology to transform global available analytical data into detailed molecular descriptions of petroleum feedstocks (6). Neurock described petroleum cuts as sets of distributions of molecular structural attributes (recurring to Boduszynski's work (7)) and through the method of Monte-Carlo generated a large amount of molecules. Integrating the obtained results in an optimization loop in which the distribution parameters are changed towards closing the gap between the generated mixture and overall proprieties of asphaltenes. Neurock proposed an interesting method to recreate petroleum cuts while representing the molecular polydiversity, however it didn't take into account the presence of heteroatoms, which are quite problematic and present in considerable amounts in asphaltenic fractions to be ignored.

Hudebine and Verstraete introduced heteroatoms, improved the optimizer and applied it to different petroleum fractions, nonetheless they noticed that a stochastic reconstruction is not capable to create a mixture of molecules where all the proprieties fitted to the cut to analyse because of the numerous constraints to be respected (8) (9).

To face the limitations of SR, Hudebine proposed to combine the stochastic reconstruction with a matrix method called reconstruction by entropy maximization (REM), developed by Hudebine (10), based on the work of Allen and Liguras, which the main differences are in the way molar fractions are adjusted using the Shannon's information theory. The SR will assume the role of creating an equimolar database of molecules relatively close to the experimental data, and REM adjust the molar fractions to refine results from the first step. The two-step algorithm SR-REM will be the one applied in this work, so a

detailed look will be given to each step of the model.

Stochastic Reconstruction

The method of Stochastic reconstruction aims the creation of an equimolar mixture that has the same proprieties of a certain petroleum fraction. To build this mixture of molecules the following principles are essential:

1. A petroleum molecule can be described as an assembly of molecular attributes and structural blocks (e.g. molecule type, number of rings, number of side chains, length of the chains, etc.).

2. A petroleum fraction can be characterized by means of a set of probability distribution functions (PDF's) for these attributes and structural blocks.

It is important to define properly the molecular attributes; those can vary with the petroleum cuts in which SR will be applied and are chosen by expert knowledge and information prevenient from analytical data. The structural blocks are the elemental constituents of the molecule, however it not enough to fully describe it, thus it is required to apply molecular attributes to describe the main characteristics of the molecule like the number of cycles, the length of the side chains or the number of rings.. From the principles mentioned before, an equimolar mixture of N molecules will be created by sampling the probability distribution function (PDF) via a Monte-Carlo procedure. The PDF's attributes will be processed in a certain order, following certain rules based on chemical knowledge, building diagram B (figure 3), to avoid the formation of impossible or improbable molecules.
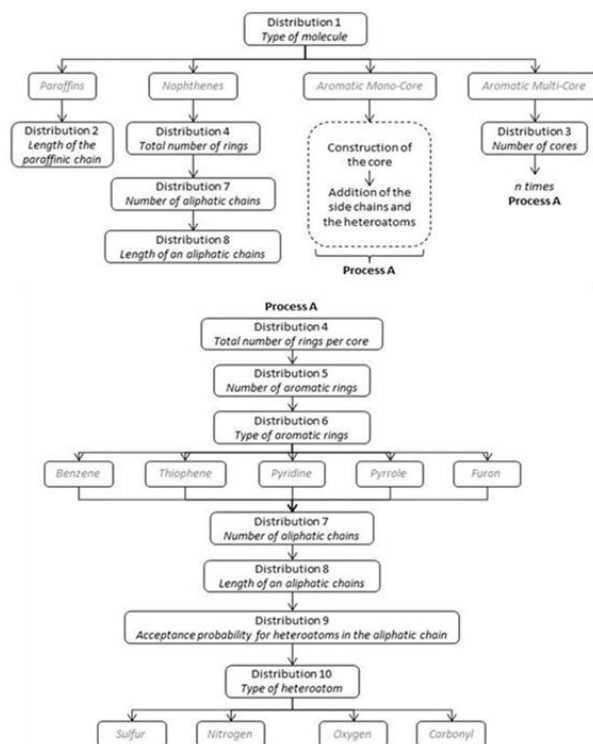

Figure 2 Building diagram B *(13)*.

The procedure will be repeated until the number of N molecules is reached, for each molecule the properties are calculated, permitting to calculate the properties of the global mixture, which will allow to compare the results with properties presented by the experimental data using an objective function. In order to minimize the objective function errors, the parameters of the distribution can be changed to find the mixture with characteristics closer to the experimental one.
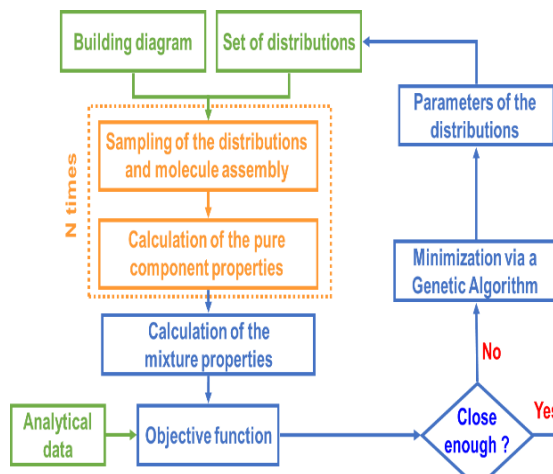

Figure 3 Scheme of Stochastic Reconstruction adapted from *(13)*.

Reconstruction by entropy maximization

Once the equimolar library of molecules is created using the SR, a different approach is made, a reconstruction by entropy maximization consists in adjusting the molar fractions of an existent mixture in order to reach the mixture's properties pretended. The adjustment was developed by Hudebine and Verstraete and is based in the matrix method of Allen et Liguras. However the criteria to refine the molecular fractions is to maximize the entropy term having in account a certain number of constrains. Basically, REM, involves 3 steps:

1.-The first one is the creation of a library of molecules, which in our case is achieved by SR,

2.-The definition of the constrains associated to: the analytical data and the respect of the mass balance among others,

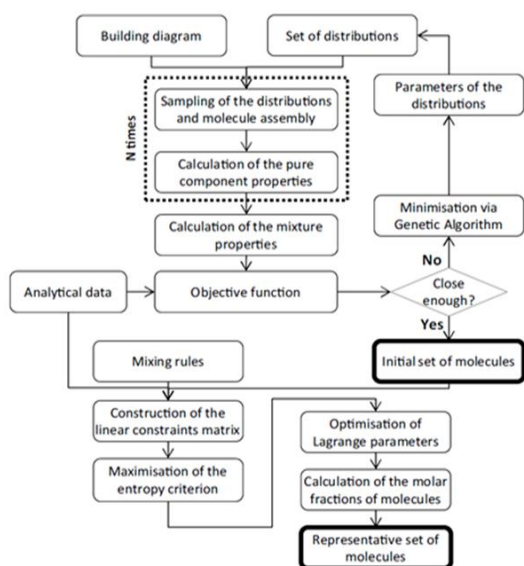**3.-**The adjustment of the molar fractions maximizing entropy.

Figure 4 Scheme of SR-REM method.

The SR-REM method is an efficient way to reproduce petroleum cuts combining two different steps, SR, assuming that oil mixtures can be described by distributions of structural blocks, generates the database of an equimolar mixture of molecules, that will be refined by the adjustment of molar fractions performed by the REM step using the maximization of the entropy criteria under linear constrains.

Methodology

The required analytical data was supplied by IFP and it is respective to a hydrotreating unit, that is composed by four different mass balances performed at different operating conditions. The strategy is to use the first point, balance 1498, as reference to generate the library of molecules by stochastic reconstruction, which a detailed analysis is required. It is important to highlight that the SR will have a huge impact on the performance of the whole method, since the molecules built will influence the REM results, in other words, if molecules are not suitable, REM will have a poor adjustment.

Once the molecules database is built, the REM step will be applied to the other mass balances referring to different operating conditions of the hydrotreating unit.

Optimization

The first runs of the algorithm had as objective the reproduction results obtained by Sofia Martins (11), having them as starting point for the present work (**SR C**). Afterwards, the mass balance discrepancies among the various analyses was fixed (**SR C P**). The remaining effort was trying to optimize the distribution parameters "manually" by adapting the best distribution result from a previous simulation as input to the next one (**SR C OPTM**). The results from these different approaches are displayed in figure 5.
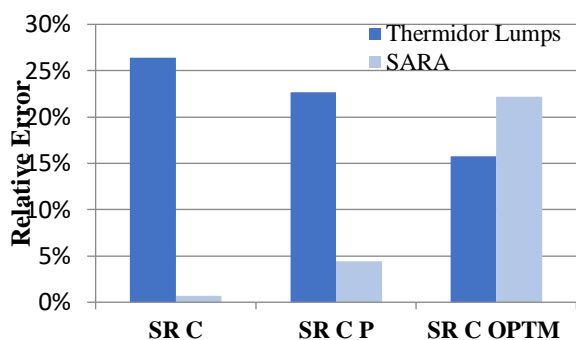
Figure 5 Results from the various SR approaches with diagram C.

Since it is impossible to reconstruct a crude fraction if the SARA and the Thermidor Lumps (TL) are not coherent. Only the TL were gave as input to the code, knowing that if the reconstruction was able to adjust to it, all the global analysis would be fitted as consequence.

This exercise showed an objective function of zero which either means that all data was perfectly fitted or the TL was not being taken into account, a debug through the algorithm showed that, actually the detailed analysis, TL, was not being considered by the objective function. Thus, it was important to test all three building diagrams (**A, B, C**) with the objective function fixed and considering the TL.
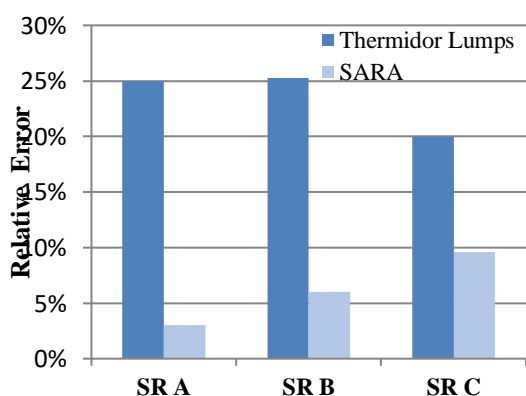

Figure 6 Results for building diagrams (A,B,C) after fixing objective function by including the TL.

As can be seen in Figure 6, the conclusion was the same as before, the best adjustment for TL, which is diagram C, still has the worst result for SARA, showing that the problem of this

incoherence prevails. A further look into the algorithm was made to identify the origin of this problem. This analysis showed that the TL was not considered correctly in the objective function, since there was not a dependency between the TL and the SARA analysis, in other words the sum of the TL did not correspond to 100%.

The solution found was to use the global composition of an element respective to a family by boiling point in SARA in the objective function, instead of the elementary composition inside of each family. Thus, is possible to ensure the concordance between SARA and TL, as it is shown in figure 7.
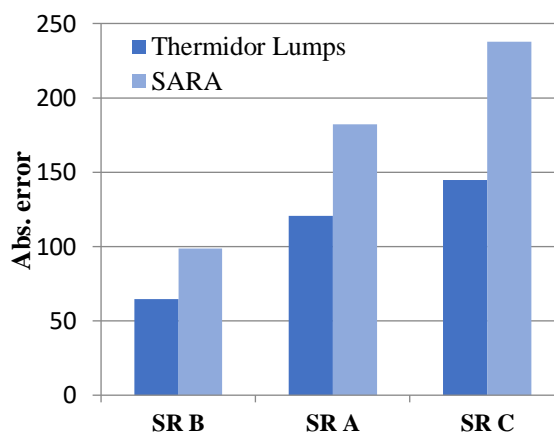

Figure 7 Global approach, concordance between thermidor lumps and SARA.

Algorithm improvement

The obtained molecules have a higher content of hydrogen than expected in Resins and Asphaltenes, meaning that the created molecules in these subfamilies are not aromatic enough. In this same families but respective to sulphur the content needs to be increased, however for Aromm and Arom it needs to be lowered.

Besides the optimizations at the subfamilies level described before, it was necessary to improve the fitting to SARA families as well. For that the method Wiehe (12), which is a diagram

used by the algorithm to classify molecules in each SARA family, was analysed.

Results

For a better understanding of the displayed results it is important to schematize the line of thought followed. The first output from the model analysed was the results to diagram B, in which, the objective was to understand if the changes made were effective and resulted in the improvement of the previous results, mainly in SR step. Having the data from SR B, it is essential to see the impact of those in REM B.

Essentially, understanding if the diagram with the adjustments made was able to build better molecules to fit to the real mixture.

SR B

As seen by the results presented in Figure 8, the algorithm can generate a set of molecules that can reproduce accurately the global properties exhibited by the mixture which the analytical data is from. However, all the small discrepancies presented in each analysis will be reflected in model predictions' performance of the thermidor lumps, since it is a result from all the data presented before.
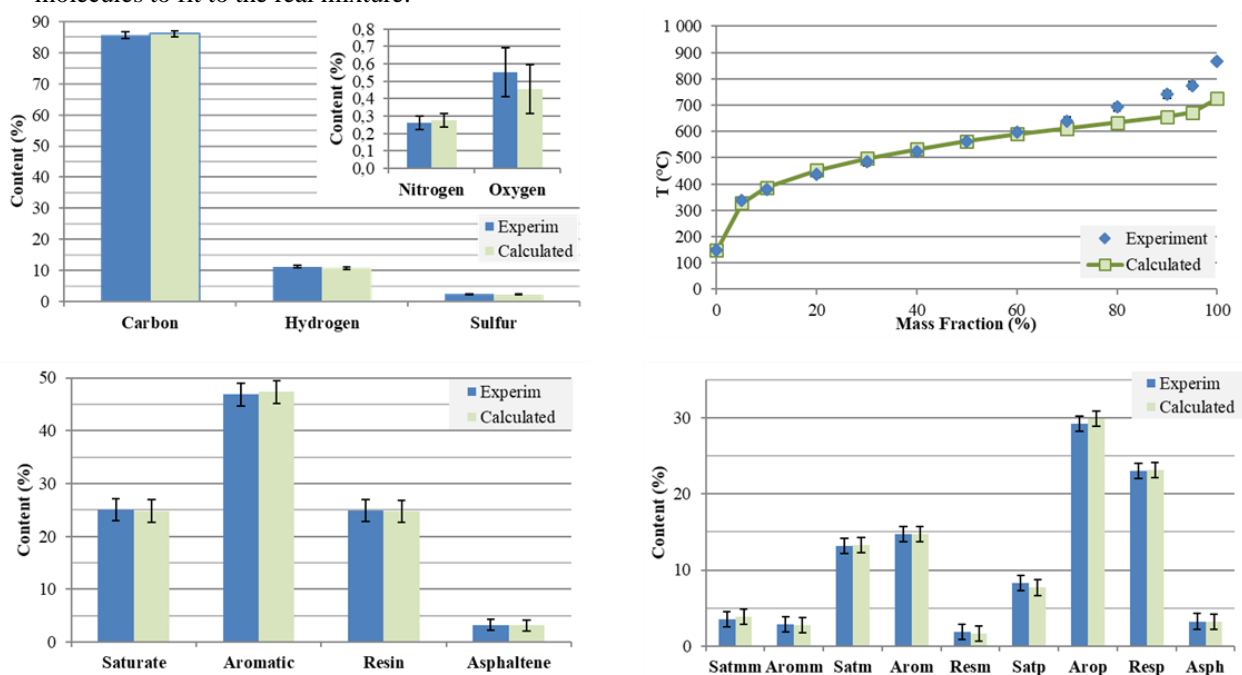

Figure 8 SR B global analyses (elemental analysis, distillation curve, SARA and subSARA results for building diagram B.
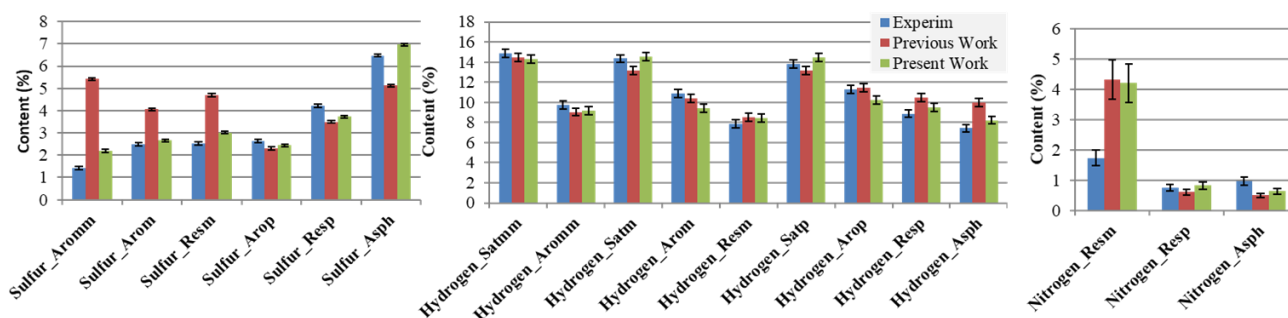

Figure 9 SR B Thermidor Lumps results for building diagram B.

The results obtained for Thermidor lumps are displayed in Figure 9 for H, S and N distributions. An overall good fitting is achieved with the values obtained by the model following the trend of the analytical data, with most of the results within the experimental error.

For S distribution sulphur in Aromm, Resp and Asph subfamilies have the poorest fitting, for H distribution Arom, Arop and Asph are the ones with a not so accurate adjustment, and in N distribution a considerable error is present in Resm.

Knowing that one of the biggest issues in SR step was related to hydrogen and sulphur distribution, it is interesting to compare the obtained results with those of the previous work. This exercise is present in figure 9, where is possible to observe an relative error decrease from 28% to around 12%, where hydrogen distribution decreased from 9% to 7% and sulphur error dropped from 80% to an error around 17%.

REM B

The REM consists in a molar fraction adjustment of a pre-existent equimolar mixture, in our case, the one generated by the SR B, to adjust to the analytical analysis. Although REM requires a pre-defined library of molecules to be performed, this methodology allows to recreate a mixture requiring less time and less analytical data than SR step, relying only in global analysis.

The three-analysis observed (figure 10) show an almost perfect fit to the experimental data to this new point of mass balance in the process.

Consequently, the REM step adjusts accurately the molar fractions of the library of molecules generated by SR to meet the global properties of a new mixture.
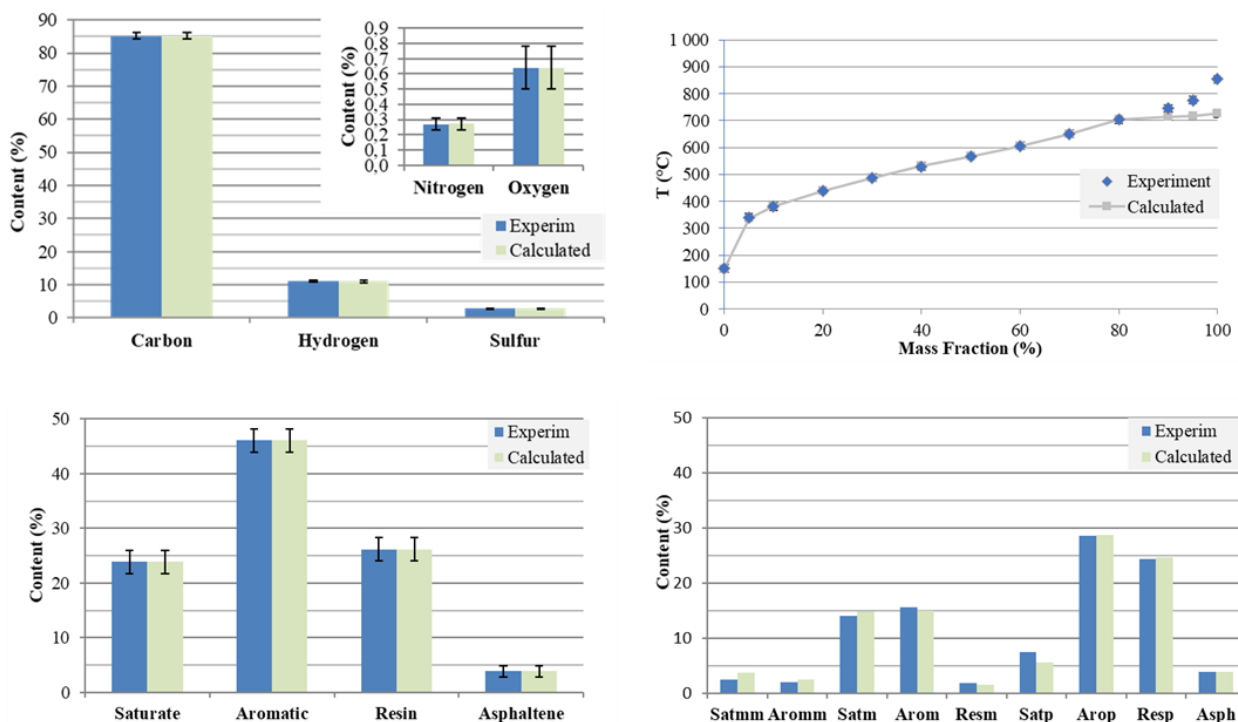

Figure 10 REM B global analyses (elemental analysis, distillation curve, SARA and subSARA results for building diagram B.
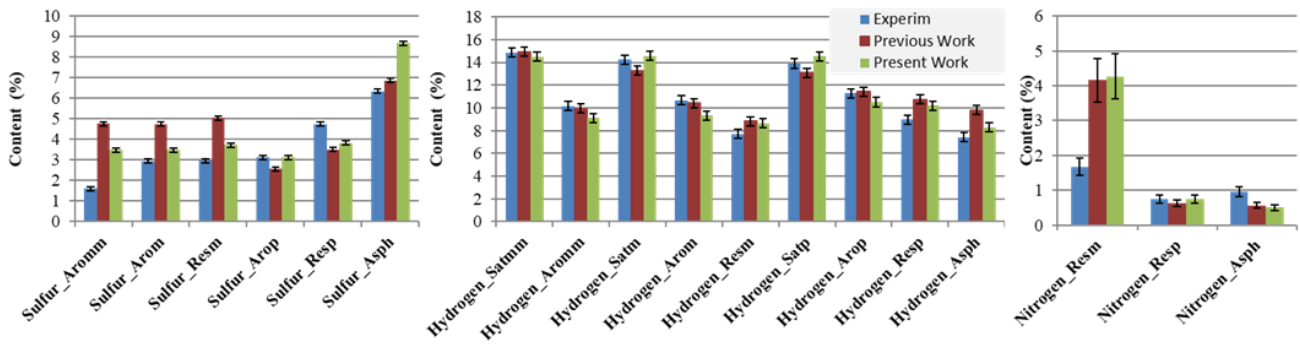
Figure 11 REM B Thermidor Lumps results for building diagram B and comparison with previous work

To understand if the model is capable to meet the experimental data related to the subSARA families, this parameter was added to the REM code, however it was not introduced in the matrix of constrains, in other words it is not take into account to adjust the molar fractions, thus being a consequence of the SARA, elemental analysis and distillation curve adjustments. The results respective to subSARA are displayed in figure 10, a very good fit is obtained for aromatics, resins and asphaltenes, however a poor fitting is exhibited for saturates (Satmm, Satm and Satp).

Finally, the results for Thermidor lumps, Figure 11, where can be seen that the values obtained by REM follow the trend of the analytical data, but there are some considerable differences in sulphur distribution, especially for Aromm and Asph subfamilies. Regarding the Hydrogen distribution the main issues are related to Resp, Arom and Aromm. Nitrogen distribution can only fit for Resp while the remaining values are not well adjusted.

To summarize, in the previous work, the deviations in the SR step were the same as the ones REM. In the present work, after optimizing and decreasing the error respective to the SR step the results shown by REM are improved. However, does not have the similarities with SR's errors like before, exposing bigger errors. This means that the optimization of SR step can only improve REM to a certain level and that can be clearly seen by the presented values in table 1.

In table 1 is presented a summary of the results obtained and, as observed before, all the outcomes from global analysis are improved in the REM step, however for detailed analysis, subSARA and Thermidor Lumps, the REM stage cannot refine the output from SR step and gets even worst results.

A detail that is important to highlight is the fact that although the REM step is able to achieve a perfect SARA adjustment, but when it comes to the subSARA families the outcome is worse than the obtained in SR. So, a better fit in SARA did not make a better adjustment in subSARA which

Table 1 Summary of SR and REM results for diagram B

Building Diagram B	Elemental Analysis		Distillation Curve		SARA		subSARA		Thermidor Lumps	
	Abs. Error	Rel. Error	Abs. Error	Rel. Error	Abs. Error	Rel. Error	Abs. Error	Rel. Error	Abs. Error	Rel. Error
SR	0,36	6,26%	42323,05	5,34%	0,22	1,10%	1,00	4,16%	38,60	12,80%
REM	1,39E-06	0,05%	21189,91	2,34%	0	0,00%	6,65	14,30%	58,88	19,34%

should be expected since one is consequence of the other. This can mean that the problem in REM can be correlated to the incapacity to adjust subSARA, since it is not being considered in the matrix of constrains for REM method. Since, subSARA is not well predicted, therefore Thermidor lumps show a worst outcome than SR, passing from a 12% relative error to 19%.

Conclusions

In this work an improvement of the methodology to reconstruct residue hydrotreating effluents is proposed. The methodology consisted in two different steps. The first one, a stochastic reconstruction (SR) followed by a second named reconstruction by entropy maximization (REM).

The first approach intended to get better results in the SR, since it was demonstrated by previous works that this step would greatly influence the REM step. The modifications led to reducing the BD B SR relative error from 27.75% to 12.80%, indicating that a more suitable library of molecules was built to reconstruct the residue hydrotreating effluents by the BD's.

Applying the second step, REM, to the generated set of molecules from BD B to recreate another point of the hydrotreating residue effluents using only global analysis, it was possible to improve the previous REM results obtained decreasing the error of BD B from 24.31% to 19.34%.

An overall improvement for this methodology was achieved with errors decreasing in both SR and REM step considerably.

Another particular detail is that the REM errors are always higher than the ones achieved by the SR in detailed analysis (subSARA and TL), and the opposite was expected since the REM is a refinement method that should get closer to the analytical data by changing the molar factions.

One reasonable explanation can lay on the fact that the model is producing relatively low errors results or at least really close to the experimental ones and so REM is not being able to suffer the same magnitude of error's reduction.

An important conclusion to make is to understand what it is holding back the REM errors to get lowered to the same level to SR, if in one hand, the problem is related with BD's and the way the library of molecules is built or, in the other hand, the problem is related to the REM algorithm.

Having the last conclusion in mind and getting BD's out of the picture, for further developments of the present topic , should be interesting to have a deeper look into the REM step and understand why detailed analysis such subSARA and thermidor lumps are not being capable to adjust molar fractions to get a better prediction. One interesting approach to optimize and study the REM behaviour can be the introduction of subSARA in the constrains matrix of this method.

Another key factor that could have influenced the performance of all the diagrams tests is related to the fact that only three types of distributions were used to describe the structural blocks of molecules, leaving room to try another types of distributions similarly to what was done in previous works at IFP.

Bibliography

1. **Speight, James G.** *The Chemistry and Technology of Petroleum.* 5 th. s.l. : CRC Press, 2014.

2. **International Energy Agency.** *Oil 2019 : Analysis and forecast to 2024.* 2019.

3. *Structure-Oriented Lumping: Describing the Chemistry of Complex Hydrocarbon Mixtures.* **Quann, R.J. , Jaffe, S.B.** s.l. : Industrial & Engineering Chemistry, 1992, Vol. 31. 2483-2497.

4. **Zhang, Y.** *A Molecular Approach for Characterization and Property Predictions of Petroleum Mixtures with Applications to Refinery Modelling.* s.l. : PhD of the University of Manchester, 1999.

5. *Modeling the Composition of Crude Oil Fractions Using Constrained.* **Steven P. Pyl, † Zhen Hou,‡ Kevin M. Van Geem,† Marie-Franc-oise Reyniers,*,† Guy B. Marin,† and.** 2011.

6. *MOLECULAR REPRESENTATION OF COMPLEX.* **MATTHEW NEUROCK, ABHASH NIGAM,DANIEL TRAUTH, MICHAEL T.** 24A, s.l. : Chemical Engineering Science, Vol 49, No. 24A, pp. 4153~177, 1994, 1994, Vol. 49. pp 4153-4177.

7. *Composition of Heavy Petroleums. 2. Molecular.* **Boduszynski, Mieczyslaw M.** 5, s.l. : Energy & Fuels an american chemical society journal , 1987, Vol. 2.

8. *Reconstruction of Petroleum Feedstocks by Entropy Maximization. Application to FCC.* **Hudebine, D., Verstraete, J.J.** s.l. : Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles, 2011, Vol. 66.

9. *Molecular reconstruction of LCO gasoils from overall petroleum analyses.* **Hudebine, D., Verstraete, J.J.,.** s.l. : Chemical Engineering, 2004 , Vols. 59 (22-23).

10. **Hudebine, D.** *Reconstruction moléculaire de coupes pétrolières.* Lyon : thèse de doctorat de l'école normale supérieure, 2003.

11. **Martins, Sofia.** *Numerical Reconstruction of residue hydrotreating effluents.* 2018.

12. **Liang, I. A. Wiehe and K. S.** Asphaltenes, resins, and other petroleum macromolecules . *Fluid Phase Equilibria.* 1996, 117, pages 201-210.

13. *Molecular Reconstruction of Petroleum Fractions: Application to Vacuum Residues from Different Origins.* **L. Pereira de Oliveira, A. Trujillo Vazquez, J. J. Verstraete, and M. Kolb.** s.l. : ACS Publications, 2013.