Universidade de Lisboa

Instituto Superior Técnico

Numerical reconstruction of residue hydrotreating effluents

Sofia Miguel Assunção da Silva Martins

Thesis to obtain The Master of Science Degree in

Chemical Engineering

Supervisors:

Prof. Dr. Filipe José da Cunha Monteiro Gama Freire

Dr. Tiago de Sousa Hipólito Sozinho

Examination Committee

Chairperson: Prof. Dr. Francisco Manuel da Silva Lemos

Supervisor: Prof. Dr. Filipe José da Cunha Monteiro Gama Freire

Members of the Committee: Prof. Dr. Carla Isabel Costa Pinheiro

November 2018

This page was left intentionally in blank.

Acknowledgements

Firstly, I would like to start to express my gratitude to my supervisors, Professor Filipe Gama Freire for his support and for the motivational and kind words given to me along this journey, and to Dr. Tiago Sozinho for his guidance, valuable advices, and all his patience and dedication provided throughout my internship. This thesis would have remained a dream had it not been for the opportunity given by him. They will always have my most sincere gratitude.

I would also like to thank Dr. Jan Verstraete for all the help that he granted me and for the ideas and suggestions without him this thesis would not have been possible, for his immense knowledge and enthusiasm in the subject I express my deepest gratitude.

My sincere thanks also goes to all my friends that have supported me during this internship principally to Marta Oliveira, as well as Daniela Santos for her friendship during this years at *IST*. I wish to thank to my *IFPen* buddies for making my adaptation fast and easy and to my Portuguese colleagues that supported me and made me feel like home.

Finally, I would like to thank to my family beginning to my parents, Luís Martins and Ana Assunção, for their support and for believing me, even when I do not, to my brothers for their unconditional support and for being my source of comfort and energy and to my uncle, Jorge Martins for all his efforts throughout his life to assure our growth. I owe to them my deepest gratitude.

Moreover, I dedicate this work to my grandparents for all the sacrifices they had to make to provide me the best conditions to accomplish my dreams, in particular to Maria de Lourdes Martins, my role model.

This page was left intentionally in blank.

Resumo

O hidrotratamento de resíduos permite a conversão de frações petrolíferas pesadas em produtos de maior valor removendo os heteroátomos contidos nestas frações como S, N e metais. Os resíduos são compostos por diversas moléculas complexas normalmente caracterizadas pelas suas propriedades gerais. De modo a superar esta desvantagem, um complexo processo de análise foi realizado para melhorar a modelização cinética do hidrotratamento de resíduos através dum sistema reacional complexo. Ainda assim, esta análise é cara e demorada.

No presente trabalho, um algoritmo numérico a duas etapas foi testado utilizando os dados das análises globais para criar um set de moléculas no qual as propriedades são consistentes com os dados experimentais, e foi proposta uma nova metodologia com o objetivo de reduzir as análises experimentais necessárias para obter a caracterização detalhada da mistura. Na primeira etapa, um grupo equimolar de moléculas é criado de maneira estocástica usando um diagrama de construção formado por um conjunto de blocos estruturais. Neste trabalho, três diagramas de construção diferentes são discutidos. Na segunda etapa, as frações molares das moléculas são ajustadas maximizando um critério de entropia. Na nova metodologia, uma abordagem designada de discretização molecular foi incluída de modo a que as moléculas de base fossem replicadas de acordo com a sua fração molar para obter uma nova mistura equimolar mais precisa. Ambas as metodologias são aplicadas a diferentes pontos de um efluente de hidrotratamento do mesmo teste para avaliar a predição dos dados analíticos detalhados a partir dos dados analíticos globais. Com a nova metodologia, para o melhor cenário de otimização de dados de input, o erro relativo médio do complexo processo de análise diminuiu de 26% para 18% ou 12% considerando, respetivamente, o diagrama de construção B ou C.

Palavras-Chave: Efluentes de HDT, Reconstrução Molecular, Resíduos de Vácuo, Maximização de Entropia, Algoritmo Estocástico, Discretização Molecular

Abstract

Residue hydrotreating enables the conversion of heavy petroleum fractions into more valuable products by removing the impurities, such as S, N and metals, found in this fraction. Residues are composed of very complex molecules and are generally only characterized by their bulk properties. These bulk properties are not detailed enough to develop a reliable kinetic model. To overcome this drawback a complex analytical procedure was employed on the feeds and effluents to obtain a more detailed characterization of residues, with the objective to improve the kinetic modeling of a residue hydrotreating fixed bed reactor with a complex reaction system. Yet, this analytical procedure is very costly and time consuming.

In the present work, a two-step numerical reconstruction algorithm using global analytical data was tested to create a set of molecules whose properties are consistent with the experimental data. A new methodology was proposed with the aim of reducing the experimental analytical procedures to obtain a detailed mixture characterization. In the first step, an equimolar set of molecules is created using a building diagram by assembling structural blocks in a stochastic manner. Three different building diagrams are discussed in this work. In the second step, the mole fractions of the molecules are adjusted by maximizing an information entropy criterion. In the new methodology, an approach called molecular discretization was included to replicate the molecules based in their molar fraction to obtain a new equimolar mixture that more closely resembles the effluent composition. Both methodologies were applied to different points of the same hydrotreating test to evaluate the prediction of the detailed analytical data from global analytical data. With the new methodology, for the best scenario of optimal input data (2^{nd} REM), the average relative error for the complex analytical procedure decreased from 26% to 18% or 12% considering the building diagram B or C.

Keywords: HDT, Molecular Reconstruction, Vacuum Residues, Entropy Maximization, Stochastic Algorithm, Molecular Discretization

This page was left intentionally in blank.

Table of contents

List of Tables

List of Figures

Nomenclature

AEBP – Atmospheric Equivalent Boiling Point

AR – Atmospheric Residue

BD – Building Diagram

BP – Boiling Point

CCR – Conradson Carbon

CDF – Cumulative Distribution Function

CME – Composition Modeling Editor

^{13}C NMR – Carbon-13 Nuclear Magnetic Resonance

CPU – Central Processing Unit

EA – Elemental Analysis

EOR – End of Run

FBP – Final Boiling Point

FCC – Fluid Catalytic Cracking

HDM – Hydrodemetallization

HDS – Hydrodesulfurization

HDT – Hydrotreating

^{1}H NMR – Proton Nuclear Magnetic Resonance

HPLC - High Performance Liquid Chromatography

IBP – Initial Boiling Point

IEA – International Energy Agency

IFPEN –IFP Energies Nouvelles

LCO – Light Cycle Oil

LHSV – Liquid Hourly Space Velocity

LSFO – Low Sulfur Fuel Oil

MD – Molecular Discretization

MDt – Middle Distillates

MTHS – Molecular Type and Homologous Series

MW – Molecular Weight

PDF – Probability Distribution Function

REM – Reconstruction by Entropy Maximization

SOL – Structure Oriented Lumping

SR – Stochastic Reconstruction

THERMIDOR – Thermal Monitoring for Isoperformance Desulfurization of Oil Residues

TLA – Thermidor Lumps Analysis

VGO – Vacuum Gas Oil

VR – Vacuum Residue

1 Introduction

Petroleum is by far the most commonly used source of energy, however, the debate between using crude oil or alternative energy sources is an important global issue. In 2016, the world's consumption of crude oil reached almost 97 million of barrels per day [1]. Currently crude oil supplies about 40% of the European-28's energy consumption and 94% of its transportation energy [2]. Forecasts made by International Energy Agency (IEA) [3] indicate that the oil demand is expected to continue to grow with the main developing economies, such as China and India.

In the petroleum industry, the production of light conventional crude oil is declining, and projected to decline further in the future. At the same time, residual fuel oil demand continues to decrease, while, the demand for motor fuels continues to increase [4]. Consequently, the necessity of processing heavy oils and residue to obtain more valuable cuts like gasoline and other liquid fuels increases.

The complexity of petroleum increases rapidly with the boiling point as a result of the increasing number of atoms in the molecules and the immense number of their possible structural arrangements [5]. Hence, to characterize a vacuum residue (VR) mixture, due to their high number of atoms per molecule, the exact molecular structures present in heavy petroleum fractions can currently not be unveiled, even when using cutting-edge analytical techniques. The lack of detail about VR impedes the use of traditional kinetic models in which the full characterization of molecules and reactions is possible. A different methodology must therefore be followed in order to be able to develop reliable and accurate kinetic models for these fractions.

To characterize VR fractions, a complex analytical procedure was employed on the feeds and effluents to obtain a more detailed characterization of residues. Yet, this analytical procedure combines many different analyses, is very costly and time consuming. To reduce the analytical burden, it was proposed to limit the analyses on some samples and to estimate more detailed information from a number of reference samples by using data reconciliation based on the overall analyses. In such a procedure, it is necessary to numerically reconstruct a mixture from partial analytical data, model hypotheses and expert knowledge.

In this work, a two-step reconstruction algorithm, SR-REM method, will be used. The first step, called **stochastic reconstruction** (SR), assumes that oil mixtures can be described by distributions of structural blocks. The choice of the blocks and the distributions are based on expert knowledge and partial analytical data. The transformation from a set of distributions into a mixture of molecules is obtained by a Monte-Carlo sampling, while a genetic algorithm adjusts the parameters of distributions. The second term, **reconstruction by entropy maximization** (REM), improves the representativeness of the set of reconstructed molecules by adjusting their molar fractions using Shannon entropy maximization criterion.

The generated set of molecules, accomplished by the SR-REM method, can then be used as input for a molecule-based kinetic models by accessing the molecular information not provided by current analytical techniques [6] overcoming the lack of detail information and create an accurate kinetic model to simulate and optimize the atmospheric residue hydrosulfurization (ARDS) process.

2 Literature Survey

2.1 Crude Oil

2.1.1 Petroleum Generalities

In general, petroleum is a mixture of hydrocarbons containing variable amounts of heteroatoms, such as sulfur, oxygen, nitrogen, and metals as vanadium or nickel, depending on the origin of the crude oil. Crude oils can include three phases: gaseous (natural gas), liquid (crude oil) and solid or semisolid (bitumens and asphalt).

The major components of petroleum are hydrocarbons that display great variation in their molecular structure. The simplest hydrocarbons are a large group of saturated hydrocarbons with chain-shaped molecules known as paraffins. The naphthenes are a series of ring-shaped saturated hydrocarbons range from volatile liquids, like naphtha, to high molecular weight substances as the asphaltene fraction. Paraffins and naphthenes constitute the family group designed by the name saturates. The olefin group contains unsaturated chain-shaped hydrocarbons but, usually, olefins are not found in crude oils due to their thermodynamic instability. The group of aromatics is known as the group of ring hydrocarbons containing one or more aromatic nuclei (unsaturated).

The non-hydrocarbon fraction of petroleum, considered as impurities, includes organic derivatives of sulfur, nitrogen, oxygen and metals such as vanadium or nickel. One of the refining objectives is to remove these impurities because, even its presence in trace quantities, brings a lot of problems during refining, such as discoloration, lack of stability during storage, catalyst poisoning or corrosion, and during the use of petroleum due to environmental issues. Besides the complexity of petroleum itself there are many types of crude oils around the world, some examples are: Brent oil, Iran Light or Heavy Oil, Kuwait oil, etc. [7]. The market value of an individual crude oil depends of its quality characteristics. Two main quality characteristics are density and sulfur content. In Table 2.1 can be found the range of elemental composition for crude oils, the metals are present in trace amounts up to 1000 ppm.

Table 2.1: Range of elemental composition of crude oil.

Element	Composition (%wt)
Carbon	84-87
Hydrogen	11-14
Sulfur	<0.1-8
Oxygen	<0.1-1.8
Nitrogen	<0.1-1.6

Conventional petroleum is often referred to as the crude material to be recovered by pumping operations as a free-flowing dark to light colored liquid. However, crude oils used today are quite different in composition and properties compared to those available 60 years ago mainly because they have higher proportions of nonvolatile

constituents. Heavy oil is a crude oil that has a viscosity typically greater than 10cP and high specific gravity. This type of crude oil also includes low hydrogen-carbon ratio and high asphaltene, sulfur, nitrogen and heavy-metal content. The generic term heavy oil, defined by Speight [8], is often applied to a crude oil that has less than 20°API and usually, but not always, a sulfur content higher than 2% w/w. Extra heavy oil is a material that is in the solid or near solid form at ambient condition, but remains mobile under reservoir conditions (high temperature). The term bitumen includes a wide variety of reddish brown to black materials of semisolid, viscous to brittle character, often without any material boiling point below 350°C, that are immobile under reservoir conditions [8].

Nowadays, the declining reserves of light crude oil have resulted in an increasing need to develop options to upgrade the abundant supply of known heavy oil reserves [9]. Indeed, taken together heavy oil, extra heavy oil and bitumen account for about 70% of all reserves of oil, while heavy oil alone accounts for 15% [10,11]. In addition, there is considerable focus and renewed efforts to handle these heavier crude oils to obtain lighter, more valuable fractions. Crude oils that are light (lower density or high degree of API) and sweet (low sulfur content) are usually priced higher than heavy crude oils. On the other hand, heavy crude oil needs more upgrading, by fractionation or chemical treatment, before refining operations to convert them to light products.

2.1.2 Petroleum Fractions

The fractionation methods available to the petroleum industry allow a reasonable effective degree of separation of hydrocarbon mixtures. Petroleum can be separated into a variety of fractions that are usually characterized by a distillation interval expressed in normal boiling temperature or atmospheric equivalent boiling point (AEBP) temperature. These boiling point temperatures can also be converted into numbers of carbon atoms equivalent to paraffins. Figure 2.1 shows the main fractions obtained in refineries, the most relevant fractions for this work will be described in the follow subsections (2.1.2.2-2.1.2.4).

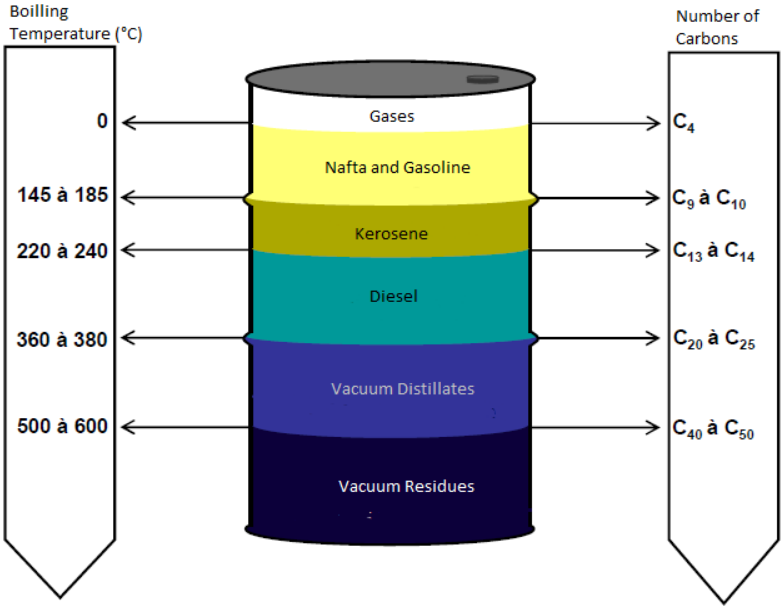


Figure 2.1: Main petroleum fractions with boiling point temperature intervals and number of carbon atoms equivalent to paraffins.(Adapted[12])

The lighter distillates are composed by gases, naphtha and gasoline. Gases are the lightest fraction of petroleum and contains hydrocarbons between C_1 and C_4, mainly methane, and low quantities of C_4^+. Other gases such as hydrogen, carbon dioxide, hydrogen sulfide and carbonyl sulfide can, also, be present. The maximum boiling point (BP) considered is 0°C.

Naphtha and gasoline correspond to the fraction with BP between 0°C and 145°C-185°C, although, the initial BP depends on the amount of C_4 dissolved in the oil. Most of the naphtha fractions are processed further in order to formulate commercial gasoline, the remainder is used to produce the main petrochemical building blocks: ethylene, propylene, benzene, toluene and xylenes.

Kerosene and diesel are considered to be middle distillates with a carbon number varying between C_9 to C_{25} and BP ranges between 145°C-185°C and 360°C-380°C. The fractions are also processed further to produce transportation fuels and heating oils.

The heavier molecules constitute the atmospheric residues, material that boils above 345°C [13]. Such an atmospheric residue (AR) is further distilled under vacuum to avoid thermal decomposition turning into vacuum distillates (vacuum gas oil) and vacuum residues to produce two fractions: vacuum distillates (vacuum gas oil, ot VGO) and vacuum residues (VR). Vacuum residue fraction is the most complex petroleum fraction since VR contains the majority of heteroatoms originally in the crude oil and the heaviest molecules, typically containing more than 40 carbons and high molecular weight whereas their BP is higher than 500°C. As the BP increases, the molecular weight also increases, leading to a huge number of possible isomers, that grow exponentially with the number of carbons as shown in Table 2.2.

Table 2.2: Carbon number and number of posible isomers.

Carbon number	Number of isomers
5	3
10	75
15	4347
20	$36.6*10^4$
30	$41.1*10^8$
40	$62.4*10^{12}$

Crude distillations yield different quantities of each fraction depending on their origin. Thus, more naphtha is distilled from light crude oils and more vacuum residues is obtained from heavy crude oils [13]. Therefore, it is not surprising that considerable attention must be given to the products obtained after vacuum distillation, not only because, as said before, the reserves of light oil petroleum are in decline, but also because refiners need to take the most income, they can from crude oil supplying the demand of the market.

2.1.2.1 Naphthas

The naphtha fraction from crude oils is dominated by normal-paraffins (typically between 25 and 45 wt%), isoparaffins (15 to 40 wt%) and naphthenes (15 to 35 wt%). Only low amounts of monoaromatics are present, typically between 5 and 15 wt% [13]. No diaromatics are present. The carbon number varies between C_5 and C_{10} with the most of the raw naphtha molecules having a low octane number. The amount of heteroatoms is in general very low.

2.1.2.2 Middle distillates

In 2016, middle distillates reached 36% of the world's consumption of crude oil [2], being the most consumed petroleum fraction. The major components in the mid-distillate fraction of petroleum are saturated species, but aromatics, which include compounds having up to three aromatic rings, and heterocycles are present and represent a larger portion of the total [9,13].

Within the saturated constituents, the concentration of n-paraffins decreases regularly with the increase of carbon number. In this cut, the naphthenes or cycloparaffins are mostly constitute by one or two rings having five or six carbons per ring. With the increase of carbon number, the concentration of naphthenes decreases while more complex molecules take place. Alkylated naphthenes, in most cases, may have a single long side chain and one or more methyl or ethyl groups. Substituted three rings naphthenes have also been detected by gas chromatography [9,13].

Within the aromatics of the middle distillates, there are substituted mono-, di-, and tri-aromatics species, but di-aromatics are the most frequent. As for the naphthenes, in most cases, the substituted aromatics typically have a single long side chain and one or more methyl or ethyl groups. The most abundant aromatics are mono-, di- and tri-methyl naphthalenes [13], other substituted two-ring aromatics can be found in smaller quantities [9].

Heteroatoms, such as S and N, are present, although, nitrogen only is found in much lower amounts. The S-heterocycles commonly found are thiacyclanes, benzothiophenes and dibenzothiophenes. There are lesser amounts of dialkyl-, diaryl- and aryl-alkyl sulfides and alkylthiophenes are scarce [13]. Nitrogen can be found as neutral and basic compounds. Pyrrole and indole derivatives, neutral species, account for about two thirds of the nitrogen while the remainder is found in basic alkylated pyridine and alkylated quinoline compounds [9].

2.1.2.3 Vacuum gas oils

Vacuum gas oil (VGO) is commonly processed in refinery to produce more valuable products, such as gasoline or diesel. In general, 60% of the saturate constituents are naphthenes in a conventional crude oil, although may be vary between 20% and 80% due to the high dependence of the paraffins and naphthenes distributions on the origin of the petroleum.

The bulk of saturated constituents consists of iso-paraffins and especially naphthenes. Analytical techniques have shown that the naphthenes contains from one to more than six fused rings with alkyl substitution.

About 50% of the aromatics are mono- and diaromatics, in which alkyl substitution typically involves one long side chain and several short methyl and ethyl substituents. In addition, naphthene rings can also be present in the molecules.

Heterocyclic species are significant contributors to the VGO fraction. Indeed, in terms of sulfur compounds, thiophene and thiacyclane sulfur predominate. Some molecules can even contain more than one sulfur atom. The prevalent forms of thiophene are benzothiophenes and dibenzothiophenes accounting for 30-40% of the total sulfur present [13].

The nitrogen contained in VGO fraction includes higher molecular weight pyridines, quinolines, benzoquinolines, amines, indoles, carbazoles and molecules with two nitrogen atoms, like diazo compounds, with three or four aromatic rings. Typically, one third of the nitrogen compounds are basic species, such as pyridines, the rest are neutral species like amines.

Oxygen compounds also are present in the VGO approximately in even lower quantities than nitrogen compounds. The most identified compounds are the carboxylic acids and phenols [9].

2.1.2.4 Vacuum residues

The vacuum residue (VR) is the most complex fraction of petroleum containing several thousands to one million different species [14]. This fraction contains all the high boiling molecules and the majority of the heteroatoms initially present in the crude oil, creating a very large possible variety of chemical structures. Impurities include sulfur (up to 6%wt), nitrogen (0.1-2%wt), oxygen (0.005-1.5%wt) and metals in trace levels (0 to 1000 wt ppm), such as nickel and vanadium [15].

The VR complexity makes the characterization of individual species impossible. The current analytical techniques do not allow to characterize this fraction. Instead, many methods were proposed to obtain partial information on these mixtures: elemental analysis, average molecular weight, NMR and other methods based on average bulk properties. Separation techniques as simulated distillation, high performance liquid chromatography (HPLC) and SARA analysis have also been applied.

The SARA analysis allows to split fractions according to their solubility and polarity in the follow families: saturates, aromatics, resins and asphaltenes. To achieve the separation a paraffinic solvent, normally n-pentane or n-heptane, is used to precipitate asphaltenes that contain high concentrations of heteroatoms [14]. Asphaltenes are the heaviest and most polar fraction of VR, and can be as much as 50%wt of VR [9]. Furthermore, asphaltenes are known to be coke precursors in acid catalysis and catalyst inhibitors [4]. The soluble fraction, called maltenes, is then fractionated and characterized. The saturated fraction is the lightest and least polar fraction, mainly constituted of paraffinic and naphthenic compounds. Aromatics are intermediate fraction, while resins constitute the most polar fraction of the maltenes. Aromatics are less polar than resins because resins contain higher concentrations of heteroatoms.

The sulfur is the most abundant heteroatom found in VR and is mainly present as sulfide and thiophenic structures [15]. The sulfur compounds are mainly present in aromatics, resins and asphaltenes. With the increase of the complexity of the molecule more difficult is to desulfurize. Neutral (pyrroles, indoles, carbazoles…) and basic nitrogen compounds (amines, pyridines, quinolines, acridines, …) are also present. Although oxygen is present in small concentration in crude oils, the majority of oxygenated compounds are found in the VR. Oxygen-containing compounds may be subdivided in two groups: neutral compounds (furans, aldehydes, ketones, esters, alcohols…) and acid oxygen compounds (phenols, carboxylic acids…). Finally, nickel and vanadium are present in trace levels in the VR.

For the 565°C$^+$ fraction, the elemental data showed that up to 80% of the molecules can contain sulfur. Also, the levels of nitrogen and oxygen start to approach the concentration of sulfur for this fraction. These elements concentrate in the most polar fractions to the extent that every molecule has on average one heteroatom [13]. Besides that, paraffins contribute only with less than 2% of the 10-20% of saturates, the remaining saturates are naphthenes [13].

2.2 Oil Refining

2.2.1 Oil Refining Generalities: from crude oil to vacuum residue

Briefly, oil refining is the separation of petroleum into fractions and the subsequent treating of these ones to yield marketable products. Generally, refinery processes are divided into three categories:

1. Separation processes consisting in the division of crude oil into various fractions depending on the nature of the material, for example, distillation.
2. Conversion processes, such as coking or catalytic cracking, intend to produce valuable materials from crude oil, usually by alteration of the skeletal structure or chemical type of the molecule through cracking.
3. Finishing processes are purification operations like hydrotreating (Subsection 2.2.2) that essentially remove impurities from the product. In addition, processes that accomplish molecular alteration, such as catalytical naphtha reforming, are included in this category.

Normally, four major forces affect the refining industry:

1. The demand for products, such as gasoline, diesel, fuel oil and jet fuel;
2. The feedstock supply dependent of the quality of crude oil and geopolitics;
3. The environmental regulations that include more stringent regulations in relation, for example, to sulfur;
4. The available technology, such as new catalysts and processes.

Since the beginning of oil refining, the demand, feedstock, regulations and technology development have undergone huge alterations [9]. Nowadays, the need for lower-boiling products are continuously increasing, while the feedstocks are becoming more and more heavy and environmental regulations tighter. Technological advances

are therefore crucial for refineries in order to produce greater quantities of lighter products from higher boiling fractions while respecting the environmental regulations.

To convert crude oil into desired products in an economically and environmentally acceptable manner, the configuration of crude oil refinery processes needs to be constantly adapted and upgraded to remain viable and responsive to ever-changing patterns of feedstock supply and market demands. As a result, refineries have been introducing increasingly complex and expensive processes to gain higher yields of lower-boiling products from the higher-boiling fractions and residua, such as residue hydrotreating processes described in the next subsection (2.2.2.1).

Nowadays, several paths may be followed by refiners to increase flexibility of existing facilities. Figure 2.2 shows a typical layout of a modern refinery. The VR is obtained from crude oil after vacuum distillation section. Indeed, the way a modern refinery operates depends not only on the nature of the petroleum feedstock but also on its configuration.

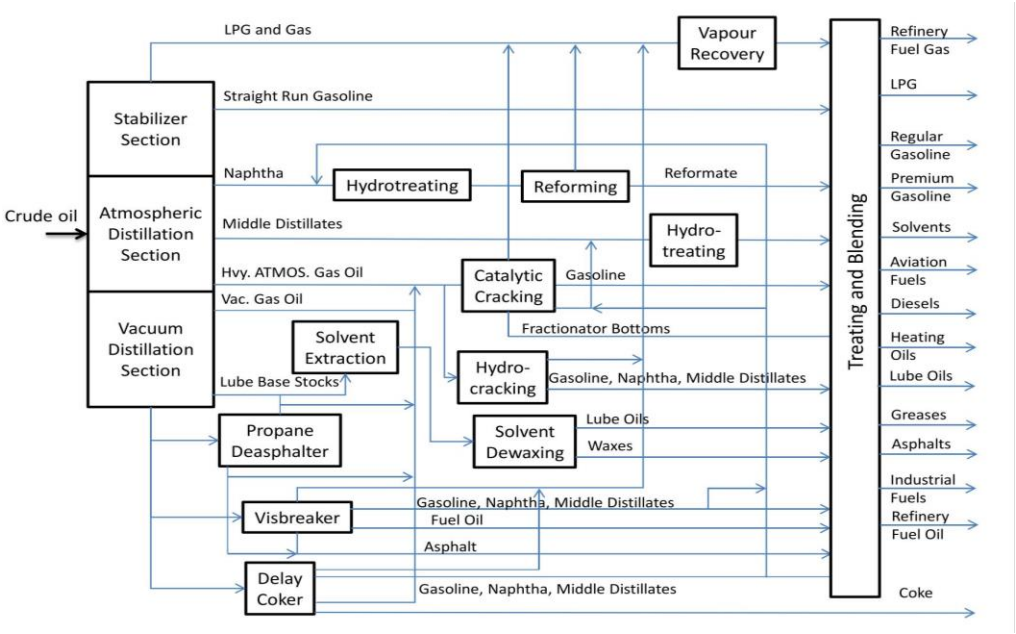


Figure 2.2: Typical layout of a modern refinery [16].

2.2.2 Hydrotreating

The growing demand for light and middle distillates and the increasing production of heavy crude oils have placed hydrotreating as one of the most important secondary petroleum refinery processes. Hydrotreating processes were initially installed in refineries to reduce the sulfur content of fuel oils [17]. Nowadays, hydrotreating processes allow the removal of various impurities, such as metals and heteroatoms (S, N, O), the conversion of higher molecular weight constituents in a feedstock to lower-boiling products, and the improvement of product quality [9].

The use of hydrogen in catalytic processes operates essentially the same manner. First the feedstock is heated and passed with hydrogen gas through a reactor filled with catalyst pellets. The reactor is maintained at a temperature

of 350°C to 460°C and at pressures from 110 to 200 bar, depending on the particular process, the nature of feedstock and the degree of hydrogenation required. After leaving the reactor, excess hydrogen is separated from the treated product and recycled through the reactor after hydrogen sulfide removal. The liquid product passes through a stripping tower in which steam removes dissolved hydrogen and hydrogen sulfide and the liquid product is storage or pumped to next unit [9].

AR and VR are the most difficult feeds to desulfurize and to convert catalytically because they concentrate the majority of the impurities inside large molecules. It is obvious that vacuum residues in which resins, asphaltenes and metals are concentrated require very severe operating conditions to achieve the specified characteristics of the hydrotreated products.

The hydrotreating catalysts, working under very severe operating conditions with very contaminated feedstocks, are deactivated by coke, metals deposition (pore plugging) and, metals poisoning. These catalysts cannot be regenerated [17] and therefore special care has to be taken to extent their life to the maximum possible limits.

Various technologies are available for upgrading heavy oils. Commercial processes mainly use two types of reactors: fixed bed reactors or moving bed reactors [18]. The advantages of using fixed bed reactors are the relative simplicity of construction and operation. The major drawback in these reactors is that a full unit shutdown is required when the catalyst is deactivated due to the accumulation of metals and coke in the catalyst pore mouth, blocking the access of reactants to the internal surface. The moving bed reactors eliminate this need for shutdown by continuously removing some of the spent catalyst and adding fresh catalyst to keep the same catalyst activity level. Of course, moving bed reactors are more complex to operate mostly because of the catalyst handling and attrition.

2.2.2.1 Residue Hydrotreating Processes

Figure 2.3 shows a residue hydrotreating process which includes permutable fixed bed guard reactors that can be switched in operation (Reactor 1a and 1b) allowing the possibility to isolate temporarily one reactor for change out of the guard HDM catalyst.

Figure 2.3: Residue Hydrotreating Process flow diagram with permutable reactor system [17].

In AR and VR hydroprocessing, high amounts of metal sulfides and coke are deposited in the catalyst. Usually, the rate of deposition is highest at the top of the reactor and lead to the end of run (EOR). The EOR conditions can

arise due to the two major causes: pressure drop or catalyst deactivation. A significant increase in pressure drop, usually occurring in the front bed and due to the presence of very fine particles that pass through industrial filters, can lead to shutdown, even though the metal and coke uptake is far from limits. The second cause, catalyst deactivation, is due to coke deposits and metal poisoning that cannot be avoided.

The Permutable Reactor System (PRS) technology allows the replacement of the catalyst of the first reactor without the need of a complete shutdown the unit. When the catalyst reaches its maximum uptake capacity, it is possible to switch reactors, using the second one and continue the process during unloading and reloading of the first reactor.

Considering the large deactivation of catalysts by coke and metal sulfides deposition, the association of several types of catalysts is mandatory in order to maximize the HDM and HDS performance. A new generation of catalysts has been developed for both the HDM and HDS sections, with a higher activity than the previous generation so as to obtain a very light sulfur fuel oil (LSFO), usually with sulfur content lower than 0.3% wt [17]. One typical association comprises three types catalysts containing Ni, Co and Mo supported on alumina [20,21]:

1. an HDM catalyst for metal removal, in the first section of the unit;
2. an HDM/HDS catalyst with balanced HDS and HDM activities;
3. an HDS catalyst designed for sulfur removal, placed in the back end of the unit.

The major function of the first catalyst is metals removal because HDS performance is really limited by metal poisoning. To achieve this, the catalyst must be capable to convert as much asphaltenes and resins as possible, since in these fractions are found most of the metals. A special porous structure is designed for this propose: a bimodal catalyst with macropores and mesopores. Macropores allow the retention of metals and mesopores are responsible for the conversion of asphaltenes and resins [22]. In each catalyst type, the porosity must be adjusted to allow heavy molecules to access the active phase and to obtain a high porous volume to accommodate metals and coke deposits. If the activity is too high, there is a risk of a less homogeneous radial distribution of metals deposited in the particle and hence, less retention. It is therefore better to reduce the catalyst activity rather than prematurely plug up the pores by an uncontrolled coke deposit. The HDM catalyst has a metal retention capacity between 60 and 80% wt of fresh catalyst.

The HDM/HDS catalyst aims at demetallizing but also desulfurizing the VR. It is in general, a mesoporous monomodal catalyst in which a compromise in terms of mean mesopore diameter has to be found between high activity and high diffusional limitations. In order to favor the HDM, an increase of the mean average mesopore diameter allows the access of the larger molecules to the active sites leading to a loss of specific surface and consequently to a decrease of the HDS activity.

The HDS catalysts, located in the third and last section of the process, removes sulfur in the feedstock. These catalysts have been optimized to maximize the HDS activity but keeping an excellent stability [22]. They are designed to develop a large specific surface, but their pores are easily blocked leading to a low metal retention capacity. Their pore size distribution is, generally, monomodal.

The severity of operation and the hydrotreating performances are dictated by the amount of contaminants in the feed and the required limits of the contaminants in hydrotreated effluents. Nevertheless, the overall system can achieve conversions for demetallization higher than 95% wt and desulfurization conversion higher than 90% wt on most residues if the design of the catalyst packages are optimized according the process objectives [17].

2.3 Refinery Modeling (Chemical Kinetic Modeling)

Refinery modeling is vital for process simulation, advanced control and optimization of a modern refinery. Generally, a chemical plant will process its feedstock under well-established conditions with the intent of making the same product that satisfies all legal constraints called product quality specification. However, the refinery's feedstock crude oil, and therefore, its molecular composition, may change on a weekly basis depending on its source, as might the most economical slate of product options due to market conditions.

A refinery process might, therefore, require frequent composition dependent optimizations and process simulation to understand how it is possible to make the most value of the feedstock. Furthermore, environmental regulations of the 21st century place restrictions on the molecular composition of petroleum products in addition to their traditional quality specifications. Hence, the aim of refinery modeling is to provide solutions to minimize the facilities investment and costs and to maximize the profit obeying the regulations.

To establish a reliable kinetic model, the composition of the feedstock is crucial. In a typical chemical process, in which only a few molecular species are present, a sufficiently detailed understanding of the chemistry, thermodynamics and kinetics of the reaction system is known. The kinetic model can then easily be derived. However, in petroleum refining, the number of molecules is enormous, and the number of chemical reactions associated is an order of magnitude greater. For such application, three different approaches of kinetic models have therefore been developed.

The first strategy, called the lumping approach, was developed in the 50's in which the chemical complexity is reduced by grouping chemical compounds into families called lumps. The kinetics are developed for each lump taking into account the reaction path of the lump. Indeed, the kinetics are not associated to a molecular theory, but directly derived from experimental data so, long and expensive experimental work has to be done. One of the other drawbacks of grouping chemical compounds is relative to the fact that not all the molecules included in the lump will have the same reactivity resulting in a bad estimation of the chemical properties of the effluent, due to the correlations used, turning this method into a very limited approach. Although, this approach is simple to apply, requiring few analytical details and a reduced calculation time.

However, over the last decades, the refining industry has been subjected to more stringent regulations and the lumping approach cannot ensure the right prediction of the process performance due to its limitations. Therefore, a detailed kinetic approach was developed. It derives a complex kinetic model based on a detailed explicit description of the reaction pathways, including its reactants, elementary steps and reactions intermediates, overcoming the disadvantages of the approach by lumping. This type of approach needs high calculation times

impeding industrial real-time process monitoring. Besides that, it is mandatory to know the molecules present in the petroleum cut and the time and money spend in characterize the petroleum cuts is huge. Since for heavy fractions conversion processes, and despite significant advances in recent years, it is almost unfeasible to measure the complete molecular composition to obtain all molecular properties and/or study the kinetics of all reactions, another approach needs to be applied.

An intermediate approach, capable to overcome the lack of molecular detail in the lumping approach, and able to reduce the calculation time of the detailed kinetic approach, called molecular approach, was developed. This strategy can retain a molecular description by the simplifications made to reconstruct a feed so that reaction simulation with a molecular-based description is possible. The existing molecular approaches will be reviewed in the next sub-chapter.

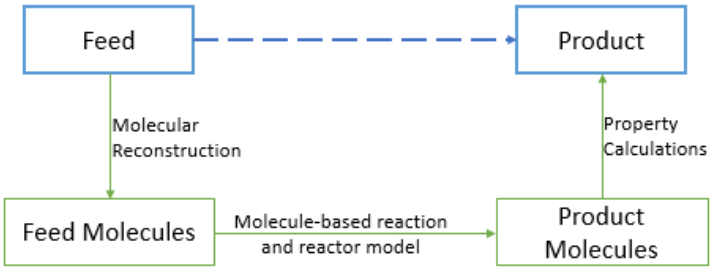


Figure 2.4:Molecular-based approach scheme.

2.4 Molecular Reconstruction of Petroleum Fractions

2.4.1 Introduction

As discussed before in the subsection 2.1.2.4, the complete and quantitative molecular detail of petroleum fractions remains unknown, especially for heavy fractions. Yet, the need for process models to predict accurate and detailed product yields increases in the petroleum refining industry.

Many authors proposed techniques to molecularly describe the feedstocks, some of them utilizing molecules to perform kinetic modelling, so being closer to the detailed approach, others using groups of molecules, called lumps, are closer to the lumping kinetic approach. The objective is to try, from partial analytical data, to create synthetic mixtures of molecules that correctly represent the characteristics of the real mixture by molecular reconstruction.

Molecular reconstruction techniques can be divided in two groups: the first one linked to the representation of the light and middle petroleum fractions and the second one linked to the representation of the heavy fractions.

2.4.1.1 Molecular Reconstruction of Light and Middle Fractions

For the first group, several methods have been developed, mainly according to a deterministic molecular reconstruction approach since the analytical techniques provide a sufficient molecular detail. This approach aims at modeling a set of molecules whose mixture properties are close to those of the real mixture using a predefined

database of compounds. Subsequently, with an objective function, they modify the mole fractions of the compounds. In such technique, the database is established taking into account a large set of analytical data and few assumptions are needed.

Allen and Liguras [23] proposed a database composed by a set of molecules. This set of molecules is defined like key-molecules and must be identified from extensive and powerful analytical work. In this model, the molar fractions are adjusted, by minimizing the average of the formation enthalpies, obtaining a mixture of molecules with properties similar to the real mixture. The reconstruction model by entropy maximization, developed in *IFPEN* [24], is based on the previous one and, adjust the mole fractions of the database applying the principle of maximum Shannon entropy [25]. Hubedine successfully applied this model to LCO gas oils [26] and FCC gasolines [24].

Others models are more close to the lumping approach. Examples of these are the method developed by Quann and Jaffe [27–29], Peng and Zhang [30,31]and Pyl et al. [32], that reconstruct the molecules based on structure blocks. Quann and Jaffe [27–29] developed a method in which molecules are built based on vectors of structure blocks named SOL (*Structure Oriented Lumping*). Peng and Zhang [30,31], based on SOL, created, the MTHS method (*Molecular Type and Homologous Series*) [16,33,34]. This method uses a matrix to represent the molar or mass composition of a petroleum fraction in terms of chemical families (matrix columns) and number of carbons (matrix lines). Theirs method resorts to the interpolation of a well-characterized database to estimate the mole fractions of the lumps in a mixture. More approaches using variations of matrix representations have been proposed in the literature, like the Statistical Reconstruction method developed in *IFPEN* [35,36] and the Pyl method [32].

The Statistical Reconstruction method developed by Hubedine [36] uses a matrix in which each element of the matrix contains the molar fractions of the pseudo-component. Although this seems similar to the MTHS method, the molar fraction of the pseudo-components is not obtained in the same way, because Hubedine suggest to rely on statistical distributions to adjust molar fractions. The Pyl method [32] constructs a matrix using a CME (*Composition Modeling Editor*), where the structure blocks are based on a library of molecules and intervals of boiling points. The molar fractions are obtained considering parametrized probability functions. By optimizing the parameters of these probability functions, it is possible to reconstruct one mixture with approximated properties of the petroleum fraction to be represented.

For low boiling fractions, deterministic molecular reconstruction approaches are the most efficient of the reconstruction methods and they are quite fast in determining the molecular composition of the mixture, when a predefined database already exists. However, to obtain accurate results a good database with compounds that are representative of the mixture is essential. In addition, to develop the database, a lot of advanced analytical tools, expensive and time-consuming, are needed. Those limitations make it impossible to apply the deterministic molecular reconstruction approaches to very complex mixtures like petroleum residues where the chemical analytical data is scarce. Besides that, deterministic molecular reconstruction approaches based on very sophisticated analytical tools such as mass spectrometry or high-resolution chromatography have little use in industry for real-time simulation models of industrial processes.

2.4.1.2 Molecular Reconstruction of Heavy Fractions

For the representation of heavy fractions for which only global analytical data is available, the first approaches considered are based on a one molecule-model. The objective was to generate a single molecule that contains the average characteristics of the mixture to be represented. This approach has been utilized mostly for representing asphaltenes.

Speight [37] created a 2D molecule that represent the asphaltene fraction from its elemental analysis, average molecular weight and proton nuclear magnetic resonance (^{1}H NMR) and developed equations to determine the fractions of seven types of carbon atoms. Hirsch and Altgelt [38] used the same approach as Speight but they suggested to add the average density and replaced some simplifying hypotheses by five floating parameters that had to be estimated by the modeler. Over time, these methods have been improved by adding analytical techniques like ^{13}C NMR [39], functional group analysis and pyrolysis analysis [40], or by splitting the cut to be analyzed by SARA analysis [41] in order to reduce the number of assumptions input for the mixture. The main drawback of this approach is that, a model molecule is far from being representative of all molecules in a mixture and a molecule does not provide the effects of polydispersity of complex mixtures in terms of physical properties and reactivities, which are essential for process modeling.

The other approach often used in molecular reconstruction of heavy fractions is a stochastic molecular reconstruction (SR) approach in which the molecular composition of a mixture is modeled from a set of probability distribution functions (PDF) for molecular structural attributes. Each attribute is represented by a PDF which provides the probability of finding this attribute in a molecule of the mixture. Thus, any molecule can be considered as an assembly of molecular attributes.

The concept of using a PDF to describe complex mixtures was initially developed in the field of polymers for characterizing the molecular size distribution of polymers [42]. After, based on detailed analysis of heavy fractions, Boduszynski [5]; showed that the structural properties of a petroleum fraction follow statistical distributions when plotted against the Atmospheric Equivalent Boiling Point (AEBP), the molecular weight or the number of carbon atoms. Klein's group [43] extended the application of the PDF to the petroleum field by relying on the characterization of petroleum fractions by means of a set of PDFs for molecular attributes. In addition, Klein's group included an optimization loop to estimate the PDF parameters from the overall properties. To do this, each PDF is described by a parameterized mathematical law. Hudebine and Verstraete [26] extended the model by modifying the PDF parameters with an objective function that minimize the difference between the real properties of the mixture and the calculated properties by the algorithm.

The SR method is an ingenious technique to overcome the lack of molecular information, but it also has disadvantages. The selection of the molecular attributes, the PDFs and the sequence of the PDF sampling steps are not trivial tasks to be performed by the modeler. Secondly, this method is computationally very demanding since the optimization loop needs to generate a large number of molecules before reaching an accurate molecular representation. Finally, SR is not always capable of fitting all mixture properties simultaneously due to the numerous constraints imposed by the algorithm.

Given the limitations of SR, *IFPEN* has developed the SR-REM coupling. Hudebine [44] proposed to combine the stochastic reconstruction with the reconstruction by entropy maximization in one method. The SR is responsible for the creation of an equimolar set of typically several thousands of molecules whose average mixture properties approximate the properties of the petroleum fraction to be represented. Subsequently, the REM adjust the molar fractions to improve the characteristics of the calculated mixture. The present work applies the SR-REM method, which will be described in detail below.

2.4.2 Stochastic Reconstruction Method

This part is focused on the first step of the SR-REM method used in this work: stochastic reconstruction. In this first step, a mixture of molecules is built to be representative of the petroleum fraction.

2.4.2.1 General description of the method

The SR step aims, through a random sampling process of parametric distributions of structural blocks, to reconstruct molecules generating an equimolar mixture of N molecules that are representative of a selected petroleum fraction based on two main principles:

1. A petroleum fraction can be characterized by means of a set of probability distribution functions (PDF) for molecular attributes, which are defined as molecules characteristics (e.g. molecule type, length of the paraffinic chains, number of rings etc.);
2. Any petroleum molecule can be viewed as a combination of structure blocks that corresponds to the constituent elements of the molecule (e.g. aromatic cycles, sulfur functions, etc.), which are characterized by molecular attributes.

The choice of chemical attributes and their PDFs are based on expert knowledge of the chemical nature of the petroleum cut. In practice, each distribution is sampled via a Monte-Carlo procedure to identify the structural attributes of one molecule. The sequence of the PDF sampling steps is oriented by a building diagram, which is defined by the modeler. The building diagram defines the hierarchic relationship between the PDFs and the order of the sampling steps. During the assembly of structure blocks chemical rules must be applied in order to avoid the creation of impossible or improbable molecules [45].

A Monte-Carlo procedure is a probability simulation technique used to obtain numerical results relying on repeated random sampling from a probability distribution. Monte-Carlo techniques are valuable tools in numerical integration, simulation, first principles modeling, and optimization. To generate draws from a probability distribution, in a Monte-Carlo simulation, a random value is selected, and the corresponding parameter value is searched for, thereby inverting the probability distribution function. In Monte-Carlo methods, the model is then calculated based on this random value. The result of the model is recorded, and the process is repeated, each time using different randomly-selected values. When the simulation is complete, the large number of results obtained describes not only the average model results, but also the corresponding standard deviation and the entire PDF of the results, by analyzing how many times the model returned a certain result in the simulation.

Once a molecule has been constructed, its properties are calculated by inspection of the structure (e.g., chemical formula, molecular weight, NMR), by group contribution methods (e.g., density, boiling point) or by correlations

(e.g., Conradson carbon). To obtain a mixture, the Monte-Carlo sampling technique must be repeated N times. Thus, a mixture of N molecules can be created, in which each molecule has a molar fraction of 1/N. The average properties of the mixture are calculated and compared with available analytical data through an objective function. The aim is to minimize the value of the objective function by modifying the parameters of the PDFs.

In the Figure *2.5* is represented the block diagram of the SR step. For an illustration of the SR procedure it is recommended to see Annex A exemplifying the construction of a benzothiophene molecule.

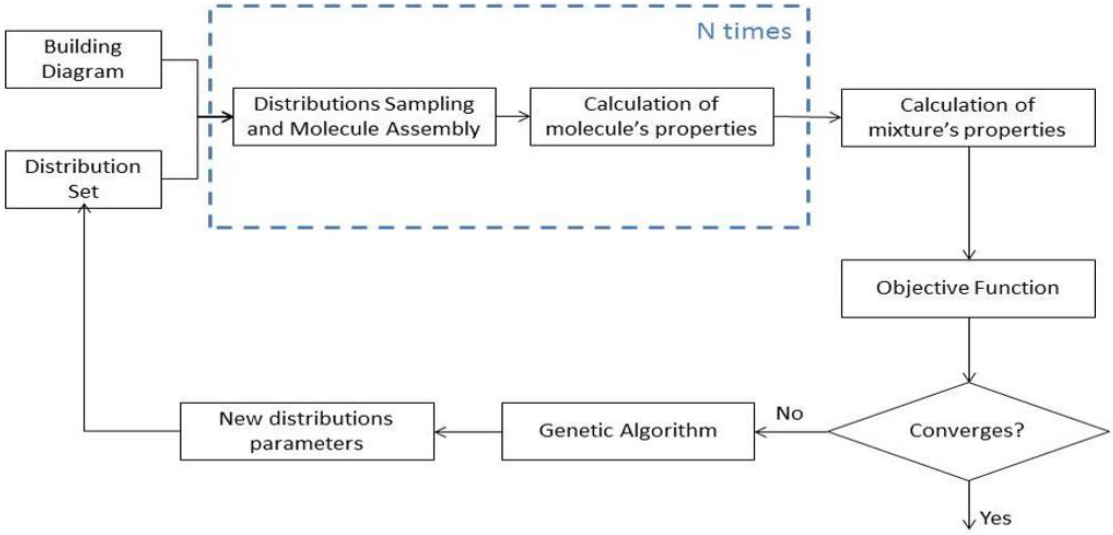


Figure 2.5: Block diagram of the SR step.

2.4.2.2 Molecular Attributes and Structure blocks

The molecular attributes must be chosen according to the information obtained by analytical techniques and by the chemical expertise of the petroleum cut. The structure blocks are the elemental constituents of the molecule. The assembly of these blocks will generate the molecule itself. Indeed, the structure of the blocks is not enough to characterize one molecule. It is necessary therefore to apply molecular attributes to define the main characteristics of the molecule, such as, the number of cycles, the length of the side chains or the number of rings. For example, the structure block "benzene ring" can form different species as naphthalene (2 benzene rings) or phenanthrene (3 benzene rings) or anthracene (3 benzene rings), depending on the number of these blocks added to the molecule and their position.

Moreover, the chosen attributes depend on the type of petroleum fraction reconstructed. Obviously, the light fractions do not contain the same compounds that in the heavy fractions. Also, the type of attributes may depend on the building diagram constructed.

2.4.2.3 Probability distribution functions (PDFs)

Each molecular attribute is associated with a PDF which returns the probability of finding a value for the corresponding attribute. The PDFs may be discrete, such as histograms, or continuous, such as gamma or exponential distribution functions.

The selection of the type of the distribution is based in two factors: flexibility and number of parameters. A flexible PDF is desirable to easily fit the correct form of the experimental observations, as well as to not damage the stochastic nature of the process. A small number of parameters to characterize the PDF is desirable in order to nor overparameterize the model.

Therefore, when the attributes have a narrow range of possible values, a discrete distribution is applied. In this case, the number of parameters of the distribution is given by the number of possible values for the attribute minus one. On the contrary, when a wider range of possible attributes is needed, histograms introduce a large number of parameters and continuous distributions are more suitable. Gamma distributions are used for attributes with large values, this function has two parameters: shape parameter (α) and a scale parameter (β). In some cases, the shape of the PDF resembles a normal distribution, and to reduce the number of parameters, a gamma distribution with β equal to two times α is used. To avoid large values, an exponential distribution is more appropriated. Exponential distributions have one single parameter.

In high detail, Hubedine [44] showed that some molecular attributes, like the number of benzene rings or the number of hexane cycles, do not follow a gamma distribution. Bailleux [46] observed that in vacuum distillates the number of benzene rings have a distribution in which the small values predominate following an exponential distribution.

Finally, the continuous the PDFs are discretized to match the possible values of the attributes, which are integer values. To this aim, these distributions are transformed into cumulative distribution functions (CDF). For this, a minimum and a maximum value are given. Usually, the minimum is equal to zero, except for attributes in which zero would mean the non-existence of a molecule, for example, the number of cores. The maximum value is truncated at 0.99 probability and after renormalized to assure a probability of one. This procedure avoids the continuation until infinite for a probability of 1, where the results are negligible.

2.4.2.4 Building Diagram

Since the petroleum fraction is not only characterized by one PDF, a building diagram must be constructed to describe the sampling order of the different distributions. The main advantage of the building diagram is to prevent the creation of impossible or improbable molecules. Therefore, it is proposed a building diagram which establishes the order between sampling the PDFs when a molecule is constructed.

2.4.2.5 Construction of a molecule

After the appropriate PDFs are found for each molecule attribute and the building diagram is created it is possible to construct one molecule. This is accomplished by two steps:

1. The PDFs from molecular attributes are sampling using the Monte-Carlo procedure to obtain the structure blocks of the molecule;

2. The structure blocks are assembled considering the connections between the different blocks. First, the core of the molecule is built, second the alkyl chains are formed to connect cores, if there are more than one.

The term "core" defines a polycyclic hydrocarbon structure which can be composed of aromatic and heterocyclic rings. After the construction of one molecule, its structure is stored in a molecular representation. In the assembly of structures two approaches can be utilized: one based on a pre-established library of cores from which one of the possible combinations chosen, while the other approach is based on the random construction on a hexagonal map.

In the first approach, a library of cores is created, with all the possible acceptable chemical configurations. After sampling the building blocks, the approach checks in the library all possible combinations and samples a core for the molecule, based on the probability of each core to be present. For heavy fractions, the number of combinations increases exponentially with the number of cycles per core. Hence, in this case, the size of the library would be huge, and the application of this approach is very unpractical. In the work of Hubedine [44] this approach is only used when the molecule does not have more than four cycles per core.

On the other hand, the construction in a hexagonal map consists of assembling the structure blocks in a hexagonal form. This approach is not limited by the number of cycles per core and the molecule construction time scales linearly increase with the number of structure blocks. However, as a disadvantage, it can create some improbable molecules because of its degrees of freedom. To avoid the assembly of improbable or even impossible molecules, a complex set of construction laws and prohibition of certain positions on the hexagonal amp had to be defined.

After the construction of the core, the alkyl chains have to be added. To do that, first the carbons that can accept alkyl chains are identified and a PDF is sampled to determine if the alkyl chain will be connected or not. If a chain has to be connected, a second PDF is generated to decide on the length of the chain. If the chain has heteroatoms or if it is a ramified chain, an analogous procedure will be used.

2.4.2.6 Construction of a mixture of molecules

To create a mixture of N molecules, the procedure to obtain one molecule must be repeated N times. The fraction of a molecule is, then 1/N. For that reason, the mixture is designated as an equimolar mixture.

The size of the mixture is an important factor because the larger the number of molecules the more representative the mixture will be of the PDFs and less random the mixture will become. However, the larger the size of the mixture, the higher will be the time to construct the mixture. Therefore, a tradeoff between the number of molecules and the computation time will be required to get a sufficiently accurate feedstock representation in a time-efficient manner.

2.4.2.7 Calculation of the properties of one molecule and the mixture

The properties of one molecule are calculated after its construction. Table 2.3 provides a list of the properties divided according their calculation procedure. When all the pure components properties are calculated it is possible to determine the properties of the mixture. The mixture properties present in Table 2.4 are calculated by applying linear mixture rules and considering an ideal mixture. A detailed description of the calculation of the molecular

and mixture properties can be found in the following references: the PhD thesis of Hubedine [44] and the PhD thesis of Costa da Cruz [47].

Table 2.3: Molecule properties of the SR and REM methods (adapted from [47]).

Molecular properties calculation methods		
Direct Inspection of the Structure	Group Contributions/ Correlations	
Chemical Formula	Normal Boiling Point	Dynamic Viscosity
Molecular Weight	Melting Point	Kinematic Viscosity
^{13}C NMR	Critical Temperature	Cohesion Energy
^{1}H NMR	Critical Pressure	Standard Enthalpy of Formation
Mass Spectrometry	Critical Volume	Standard Gibbs Free Energy of Formation
PIONA Families	SARA Families	Standart Entropy of Formation
Basic Nitrogen	Specific Gravity	Octanol/Water Partition Coefficient
	Conradson Carbon	

Table 2.4:Mixture properties of the SR and REM methods (adapted from [47]).

Mixture properties calculation methods		
Elemental Analysis	Normal Distillation	Average Molecular Weight
^{13}C NMR	Simulated Distillation	Average Specific Gravity
^{1}H NMR	SARA Families	Average Dynamic Viscosity
Mass Spectrometry	PIONA Families	Average Kinematic Viscosity
Conradson Carbon	Thermidor Lumps*	

*Coded during this work

2.4.2.8 Objective function

Once the properties are calculated they can be compared to the available analyses through an objective function. In this work the objective function (Equation 2.1) is the weighted sum of the deviations between the calculated properties and the experimental data:

$$OF = \frac{1}{N_P}\sum_{i=1}^{N_P} w_i \delta_i \qquad \textit{Equation 2.1}$$

Where, N_P represents the number of properties, w_i is the weight associated to the property i and δ_i represents the difference between the calculated and experimental values of property i. This last factor is determined by Equation 2.2.

$$\delta_i = \frac{1}{N_{M,i}}\sum_{i=1}^{N_{M,i}} \frac{\left|x_{j,i}^{exp} - x_{j,i}^{calc}\right|}{x_{j,i}^{exp}} \qquad \textit{Equation 2.2}$$

In which, $x_{j,i}^{exp}$ is the experimental value of the measure j of the property i, $x_{j,i}^{calc}$, is the calculated value of the measure j of the property i and $N_{M,i}$ is the number of measures of the property j.

2.4.2.9 Optimization Loop

As said, the initially proposed PDFs do not necessarily lead to the best representation of the fraction. In this case, the agreement with the experimental data is not satisfactory, and the parameters must be changed in order to generate another mixture with properties that are closer to that of the sample. The objective function will therefore be minimized by modifying the parameters of the PDF.

Since the molecular reconstruction uses a Monte-Carlo approach to randomly sample the distributions, it is possible, with the same PDFs with constants parameters, to obtain two different mixtures of molecules that do not have exactly the same properties. The Monte-Carlo will therefore, cause noise in the value of the objective function. The noise can be reduced by adding more molecules to the mixture, but this is not desirable due to the increase in the computing time. To minimize a noisy objective function, it is required to use an optimization technique that does not need derivative information to find the global minimum of the function. Several algorithms exist, the goal is to use the approach least sensitive to noise. Two approaches are summarized below:

1. The simulated annealing algorithm [48] bases on the principle of metallurgy that allows the fraction to find a lower energy state through a decreasing or increasing of the temperature slowly. For each defined temperature the parameters of the PDFs are changed. If these new parameters allow a lower energy state, this is a lower value for the objective function, the parameters are kept, otherwise, the values are only kept with a certain probability that will decrease with the decreasing of the simulated temperature.
2. The genetic algorithm, [12], is based on Darwin's theory of evolution that states the most adapted individuals survive and reproduce in their environment. In the case of parameters estimation, the individuals are the ensemble of parameters and the genes are the values of the PDFs parameters. In every generation (iteration), the less adapted individuals are rejected, i.e. the individuals with the higher value for the objective function. The most adapted ones are kept, and a selection of the surviving individuals is reproduced to obtain new individuals for the next iteration. The reproduction is made by randomly passing genes from parents (surviving individuals) to children (new individuals). Some mutations can be performed to expand the optimal search area and avoid local minima.

Comparing this two approaches, Schnongs [49] observed higher oscillations in the objective function for the simulated annealing algorithm when the parameters are not close to the global minimum of the function. Indeed, the genetic algorithm is less sensitive to the starting point because it has many individuals in the population. This reveals that the genetic algorithm is more adapted to the SR problem than the simulated annealing algorithm. For a further understanding of the genetic algorithm, the reference [12] is recommended.

2.4.3 Reconstruction by Entropy Maximization Method

The second step of the SR-REM approach is the reconstruction by entropy maximization. This method is applied in order to improve the properties of the previously generated set of molecules. The REM method can be applied in two ways after the SR step:

1. Direct way: in this case for each petroleum fraction to be reconstructed, a mixture of molecules is generated by the SR step and optimized by the REM adjusting the molar fractions;
2. Indirect way: applying the SR step to one feedstock, a set of molecules is first generated. These molecules have to be diversified and in sufficient number to be used as a library for the REM step for other feedstocks. This method significantly decreases the computing time.

2.4.3.1 General description of the method

In the REM, the molar fractions of a pre-existing mixture are modified in order to achieve the best representation of the petroleum fraction. The REM method is CPU efficient and provides a very good agreement between simulated and experimental properties. However, the algorithm can easily diverge if the initial set of molecules is not correct, inconsistent, or very different from the feedstock type.

The first part is to construct the library of molecules with mixture properties similar to those of the feedstock. The library can be obtained by several approaches, as seen in Section 2.4.1. In our case, the SR method will be used. Once the library of molecules is defined, the maximum entropy criterion is minimized subject to the mass balance and the analytical constrains. A scheme of the REM step can be found in Figure 2.6.

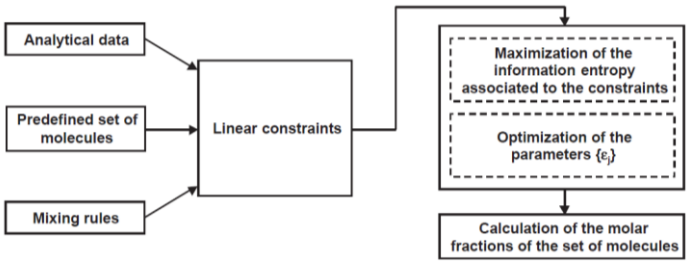


Figure 2.6: Diagram of the REM step ([50]).

The adjustment of the molar fractions to optimize the properties of the mixture is accomplished by maximizing Shannon entropy (Equation 2.3):

$$H = -\sum_{i=1}^{N} x_i \, ln \, x_i \qquad \textit{Equation 2.3}$$

Where, *H* represents the entropy criterion, x_i represents the molar fraction of the molecule *i*, and *N* is the number of molecules in the library. It was demonstrated by Hubedine [24] that the maximization of this criterion ensures:

- In the absence data, no molecule is favored over another, so the molar fraction is 1/N for all molecules presents in the library.

- When data is added to the system, the molar fractions are adjusted to obtain the best correspondence between the properties of the mixture and the added information.

This input that will provide the constrains of the system, is supplied by the analytical data and by the mass balance knowledge. Furthermore, the properties of the mixture depend on the mole fractions and on the pure molecule properties. If they follow a linear law, the average property can be calculated by Equation 2.4.

$$f_j = \sum_{i=1}^{N} f_{i,j} \, x_i \qquad with \; j \in J \qquad \textit{Equation 2.4}$$

Where, f_j is the value of the constrain j, $f_{i,j}$ represents the coefficient of the molecule for the constraint j and J represents the number of constraints. The molar balance given by Equation 2.5 is an additional constraint that is always present.

$$\sum_{i=1}^{N} x_i = 1 \qquad \textit{Equation 2.5}$$

Every constraint is associated with a Lagrange multiplier and added to the entropy criterion. This allows to obtain Equation 2.6.

$$H = -\sum_{i=1}^{N} x_i \, ln\, x_i + \mu(1 - \sum_{i=1}^{N} x_i) + \sum_{j=1}^{J} \lambda_j \left(f_j - \sum_{i=1}^{N} f_{i,j}\, x_i\right) \qquad \textit{Equation 2.6}$$

In which, μ and λ_j are the Lagrange multiplier for the constraints associated with the molar balance and associated with the constraint j, respectively.

After deriving Equation 2.6, with respect to x_i, the following relations are found:

$$x_i = \frac{e^{\left(-\sum_{j=1}^{J} \lambda_j f_{i,j}\right)}}{Z} \quad ,\forall i \in N \qquad \textit{Equation 2.7}$$

$$Z = \sum_{i=1}^{N} \left(e^{\left(-\sum_{j=1}^{J} \lambda_j f_{i,j}\right)}\right) \qquad \textit{Equation 2.8}$$

Introducing Equation 2.7 in Equation 2.6 leads to the final equation:

$$H(\lambda_j) = ln\, Z + \sum_{j=1}^{J} \lambda_j \; f_{i,j} \qquad \textit{Equation 2.9}$$

The value of the entropy criterion no longer depends on the value of x_i, but only on the value of λ_j and the pure molecule properties, thereby effectively reducing the optimization problem from dimension *N* to dimension *J*. Maximization of the entropy criterion supplies the optimal values for the Lagrange parameters, which allow to directly calculate the molar fractions by means of Equation 2.7. It is important to notice that the exponential form of the molar fraction's calculation tends to deform the molar fraction distribution in order to favor the appearance of a limited number of dominating molecules. It is known that the distributions of most characteristics in petroleum fractions are not exponential but rather with a majority of compounds with average properties and few extreme components. To be able to achieve this with the REM method, a library with well approximated properties is essential.

It is also known that the analytical constraints can be non-linear and can contain uncertainties so Equation 2.6 must be transformed. In the present work, only linear constraints will be applied. The developments for other constraints and the corresponding deductions can be found in the work of Hubedine [44].

2.4.4 Conclusions

The SR-REM method allows to obtain a representative mixture of molecules of the petroleum cut from a pre-established library of molecules constructed by the SR. By adjusting the molar fractions with the REM step, a better fit of the mixture properties can be achieved.

Summing up, the algorithm for SR method follows these steps:

1. Identification of the molecular attributes that characterize the structure blocks from feedstock analyses and from expert knowledge;
2. Association of an appropriated PDF for each attribute;
3. Definition of a building diagram that dictates the sequence of the reconstruction process;
4. Construction of a mixture with N molecules by repeating the following algorithm:
   a) Sampling of PDFs via a Monte-Carlo procedure;
   b) Assembling of the structure blocks according to the building diagram and chemical rules;
5. Calculation of pure component properties and mixture properties;
6. Comparison of the calculated mixture against the analytical data with an objective function;
7. Optimization of the PDF parameters by minimizing the objective function.

For the REM method the steps to be considered are:

1. Construction of one *a priori* library of molecules, made by the SR method;
2. Implementation of the constrains associated with analytical data and mixing rules;
3. Adjustment of the molar fractions by maximizing the entropy criterion.

The SR and the REM method combined can deal with the disadvantages of each other. The SR step is CPU consuming and does not provide the best agreement between simulated and experimental data but can provide a consistent mixture to be utilized in the REM step that, otherwise, will have many difficulties trying to optimize the mixture. This REM step will give the best fit for the properties while being more time efficient. For these reasons, coupling both methods seems an excellent solution to eliminate their respective drawbacks and take advantages of their benefits.

In conclusion, the idea of the SR-REM method consists in using a SR step to build a reference mixture. The set of molecules thus obtained can be used in a second step either to improve its representativeness or to rebuild similar petroleum fractions via the REM method. Consequently, this two-step molecular reconstruction algorithm can be used to predict missing analytical data or to generate molecular mixtures that can serve as input to detailed kinetic models.

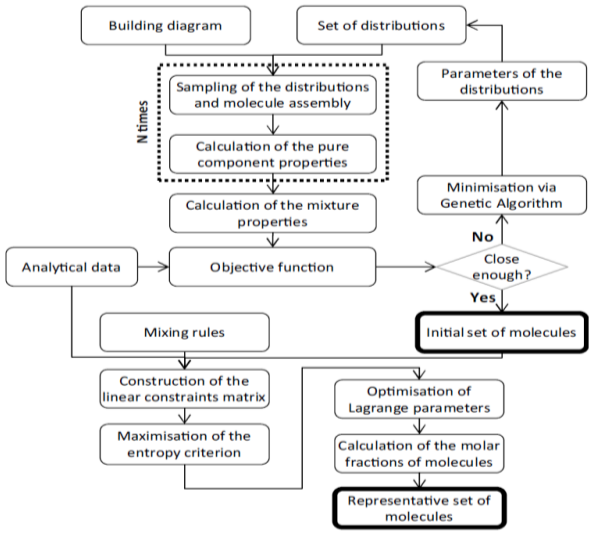


Figure 2.7: Block diagram of the SR-REM method.

3 Objectives

At *IFPEN*, a simulator of the residue hydrotreating process was developed, called *Thermidor*. This simulator deals with a complex three phasic system with 116 reactions that requires as input a detailed description of the feedstock.

Distillation curve, elemental and SARA analysis are partially performed in refineries. On the other hand, the Thermidor Lumps Analysis (TLA) provides a deeper detail about the mixture by providing the elemental analysis for each SARA group in a range of boiling points (see Annex D). This analysis is however very costly and time consuming. Indeed, to preform it, the liquid needs to be distilled in three different cuts: middle distillates (MDt), vacuum gasoil (VGO) and vacuum residue (VR). After that, for each cut, a SARA analysis has to be performed. Finally, for each SARA family, an elemental analysis is done.

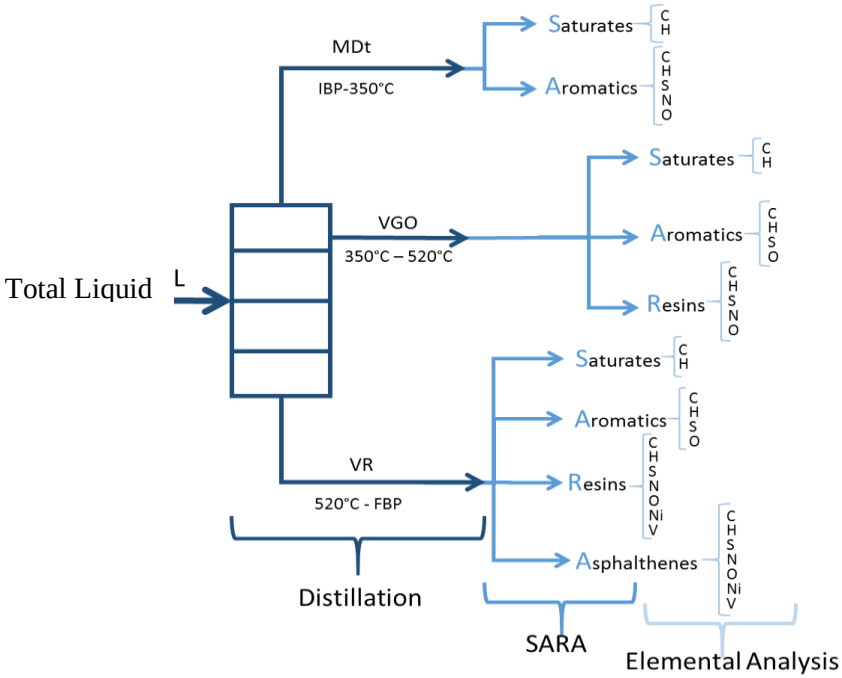


Figure 3.1:Thermidor Lumps Analysis (TLA) scheme.

The prediction of the detailed effluent description by numerical reconstruction may be able to avoid exhaustive and costly experimental work to obtain a full TLA description, thereby saving time and money. It may also allow the utilization of existing pilot plant points for kinetic regression, where only partial analysis is available.

The objective of this work is therefore the numerical study of molecular reconstruction by SR-REM method, analyzing the performance of such models over hydrotreated effluents using macroscopic analytical data to predict the detailed description of the effluent. Further, this work aims at defining the optimal input analytical data that must be used to assure a correct reproduction of the hydrotreated fractions, while limiting the amount of demanded analysis.

4 Methodology

To meet the proposed objectives, the previously described methodology was used in this work to reconstruct hydrotreated effluents coming from a fixed-bed reactor. To apply the methodology the analytical data was provided by *IFP Energies Nouvelles* through the *Proceed Data Base*.

To minimize the calculation time, the proposed methodology consists in reconstructing one hydrotreating pilot plant point assuring the correct reproduction of the molecular mixture by the SR step. This step is computationally intensive when compared to entropy maximization, due to its random character. Therefore, the remaining points of the test were reconstructed using the pre-created library of molecules by applying the REM methodology to assure a reasonable calculation time and reproduction.

The REM needs one library of molecules and experimental data to maximize the entropy criterion adjusting the molar fractions. Although all fractions under study are different, a single reference library with enough molecules is necessary. This reference library will then be manipulated to predict the various mixtures. Taking into account the study made by P. de Oliveira [12] about the impact of the number of molecules on the noise of the objective function and the performance of the optimization algorithm, it was concluded that a mixture of 5000 molecules would assure a good trade-off between calculation time and correct reproduction.

A simple diagram describing the methodology can be found in Figure 4.1. If the obtained library of molecules that is reached as reference has good results, the calculation time can be considerably reduced by applying only the REM, thereby maintaining the efficiency of the molecular reconstruction. Different ways of obtaining the library of molecules were tested in this work.

The library of molecules was obtained by applying two approaches for three different building diagrams. The first approach intends to stochastically reconstruct a VR using the equimolar output mixture as a library of molecules. In Figure 4.1, this approach is denominated SR followed by the letter of the respective building diagram. The second approach uses the SR step to obtain a mixture of molecules, which is optimized by changing its molar fractions through a REM step. The resulting SR-REM mixture will be the new reference database. However, the resulting SR-REM mixture will need to be transformed into an equimolar mixture, since the REM step applied in an indirect way to the other vacuum residues needs as input an equimolar mixture. To perform this transformation, another tool was utilized, molecular discretization (MD), to achieve the new library of molecules. More information concerning the MD procedure can be found in sub-section 5.2.1 In Figure 4.1, this approach is called SR+REM+MD followed by the letter of the respective building diagram. Figure 4.2 and Figure 4.3 present the diagrams of the first and second approach, respectively.

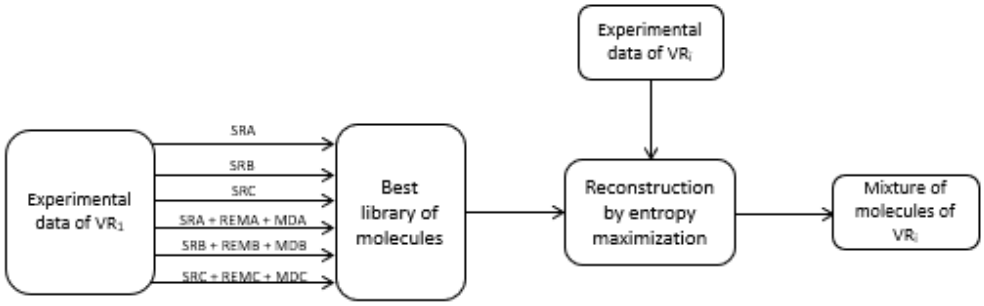


Figure 4.1: Methodology diagram.

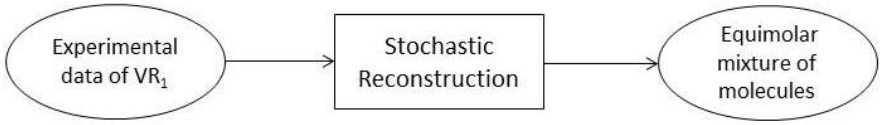


Figure 4.2: SR approach.

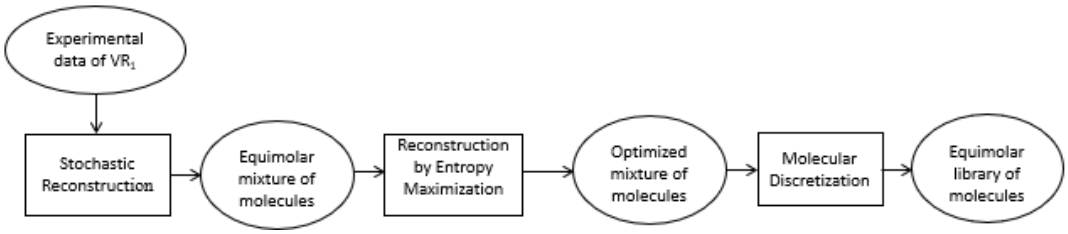


Figure 4.3: SR+REM+MD approach.

In the present work, three building diagrams for VR were used with different quantities of distributions and parameters to evaluate their performance. A brief description of each is given below.

The "Residue A" building diagram is the building diagram implemented by Luís Pereira de Oliveira in his thesis [12] in *IFPEN*. This diagram has 16 distributions and a total of 24 parameters. The details about the chosen distributions, as well as the building diagram can be found in Annex C (Figure C1 and Table C1)

The "Residue B" building diagram (see Figure C2 in Annex C) was developed by Costa da Cruz [47] during her thesis in *IFPEN*. The aim of this building diagram is to reduce the number of attributes as well as the number of parameters compared to the "Residue A" version. This modification allows taking advantage of stochastic nature of the procedure giving it more flexibility. The "Residue B" version is described by 10 distributions with a total of 17 associated parameters. A summary of the distribution set can be found in Table C2 in Annex C.

A more suitable building diagram could be developed by accounting for the TLA property introduced in this work and allowing a deeper research about the molecular reconstruction. This new building diagram was called the "Residue C" building diagram and will be presented in the next chapter.

Finally, the parameters of the genetic algorithm for the optimization loop must be chosen to explore the entire space of solutions to obtain the best set of parameters. In this work, the same parameters proposed by Luís Pereira

de Oliveira [12] are applied. In his thesis is demonstrated that this configuration allows to correctly explore the optimization space with an acceptable calculation time.

Table 4.1:Parameters of the genetic algorithm.

Initial number of individuals	2048
Minimum number of individuals	128
Number of molecules per individual	5000
Maximum percentage of refused molecules	50%
Population percentage kept for the next generation	50%
Average mutation percentage per individual	25%
Number of iterations of the algorithm	100

In addition, for the reconstruction of VR based on each library of molecules created the optimal input analytical data that assures a correct reproduction of the fractions was explored by progressively reducing the comparison criteria used.

5 Results and Discussion

The results and discussion are divided in four different steps:

- Creation of the library of molecules by the SR approach;
- Improvement of the library of molecules through the REM+MD approach;
- Application of the library of molecules for a different effluent (E_i);
- Development of a new building diagram to construct the library of molecules.

As mentioned before, the experimental data was taken from the *Proceed Data Base*, that compiles all pilot scale residue HDT experiments. Four points of the same test were selected: point 0 (ref), point 1, point 2 and point 3. This pilot plant is composed of two reactors that allows the testing of several catalysts activity. In Table 5.1 is presented the main properties of the studied effluents (see Figure 2.3 – reactors).

Table 5.1: Main characteristics of the studied effluents.

	Point 0 (ref)	Point 1	Point 2	Point 3
Density$_{(15°C)}$	0.968 g/cm^3	0.974 g/cm^3	0.961 g/cm^3	0.968 g/cm^3
CCR*	10.87%	11.61%	9.67%	11.08%
Asphalthenes*	3.25%	3.85%	1.88%	3.39%
Sulfur*	2.48%	2.85%	1.97%	2.41%
Nitrogen*	0.26%	0.28%	0.25%	0.27%
Nickel*	13.7 ppm	16.0 ppm	9.7 ppm	8.2 ppm
Vanadium*	29.0 ppm	37.2 ppm	16.6 ppm	15.7 ppm
BP $_{5\% wt}$	337.5°C	341.1°C	317.7°C	334.1°C
BP $_{50\% wt}$	561.4°C	568.3°C	542.8°C	561.2°C
BP $_{95\% wt}$	774.7°C	776.4°C	741.2°C	774.7°C

*%wt

The main difference between the points lies in the operating conditions of the HDT. To obtain a higher conversion of the feed, meaning a more hydrotreated effluent with fewer impurities (S, N and, metals), higher reactor temperature, lower catalyst age and lower LHSV are needed. Since the point 0 has intermediate effluent properties, it could be a more appropriate sample to create suitable molecules that can adjust to the analytical data of the other points. Therefore, this sample may be more propitious to create the library of molecules. After the library of molecules has been created, the three other points were reconstructed by REM. The application case of the library for the point 1 will be explained in detail in section 5.3.

Although metals are found in VR fractions, their concentrations have been neglected. Indeed, since the metals are found in the order of tens or hundreds of ppm, with a mixture of 5000 molecules the probability of creating one molecule is negligible. To have one molecule in which metals are present, a larger mixture would have to be created, up to 100 000 molecules. To create such a mixture, a calculation time of around 4 days would be needed

(16 times more than with 5000 molecules). Moreover, due to the low metals abundance, the overall results would not significantly improve, setting back the relation between calculation time and correct reproduction.

5.1 STEP 1: Stochastic Reconstruction

To apply the SR step to the point 0, a set of structural attributes needs to be chosen as well as the building diagram (BD). Several diagrams can be constructed, leading to different results. In this work, two pre-existing diagrams were compared: building diagram A [12] and building diagram B [47]. Further, a third one was proposed, called building diagram C and discussed in sub-section 5.4.

5.1.1 Comparison results between building diagram A and B

To select the best library of molecules, a comparison criterion was selected. This criterion takes into account all the detailed analytical data necessary to the *Thermidor Simulator* (distillation profile, Elemental Analysis (EA), SARA analysis and Thermidor Lumps Analysis). Using all the detailed analyses the objective function was minimized to fit the experimental data.

The results for the comparison criterion for each analytical information are shown from Figure 5.1 to Figure 5.5. The errors evaluated are the absolute quadratic error (AQ) and the average of the relative error (AR) calculated by the following equations, respectively:

$$AQ\ Error = \sum_N \left(x_{experimental} - x_{calculated}\right)^2 \qquad \textit{Equation 5.1}$$

$$AR\ Error = \frac{1}{N}\left[\sum_N \left|\frac{x_{experimental} - x_{calculated}}{x_{experimental}}\right|\right] \qquad \textit{Equation 5.2}$$

In which x represents the value and N the number of values in each analysis. A summary of the errors will be present in Table 5.2 and Table 5.3. In the top of each column of the graphs, the value of the relative error is given for each BD. When the experimental value is equal to zero, no relative error is presented.

It is also important to consider that all experimental values have an associated error. For the elemental analysis (EA), the uncertainty of the measurement will depend on the type and the abundance of the element. For vacuum residues, typical values are 1.5% wt for carbon (approximately 2% relative), 0.4% wt for hydrogen (approximately 4% relative), 0.2% wt for sulfur (approximately 8% relative), 0.03% for nitrogen (approximately 10% relative), and 0.05% wt for oxygen (approximately 40% relative). Concerning the SARA analysis, the measurement uncertainty is approximately equal for all the families and its around 1%wt. The distillation curve has an uncertainty of 5°C for all points except for the 5% wt point and the end points that have an uncertainty of 20°C. The TLA is affected by the propagation of the errors given by the distillation, SARA and EA to achieve the detailed description of the lumps, thereby increasing the uncertainty of those experimental values.

Taking in consideration Figure 5.1, the fit of the EA is good for both diagrams. All the relative errors are below 5%, except for the oxygen calculated in BD A. These errors are reasonable given the experimental errors. It was expected to find the biggest discrepancies in the heteroatoms, mainly in nitrogen and oxygen since they are present

in lower quantities and, the experimental uncertainty in these elements is larger (around 40% relative to the experimental value). Hence, the values calculated by the algorithm are accepted as a good prediction, even with an oxygen relative error of 7.2% for BD A.

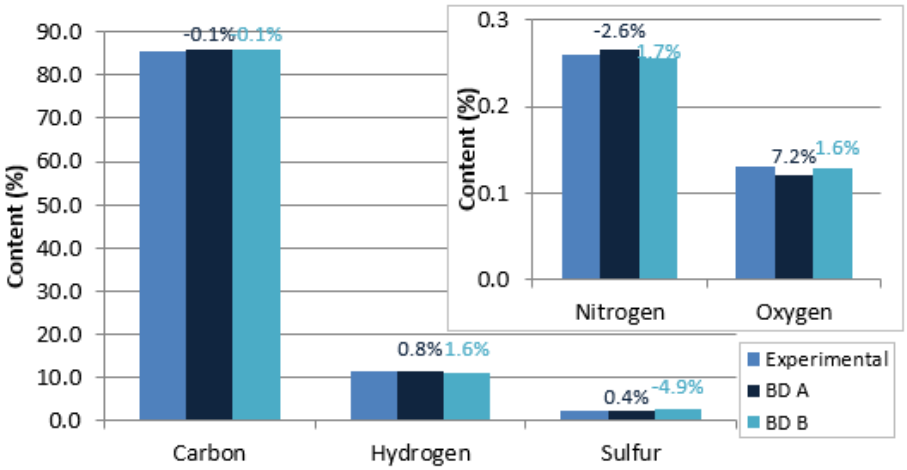


Figure 5.1: Elemental analysis results SR.

For the SARA analysis (Figure 5.2), the fit of BD B is almost perfect, even though BD B has less distributions and less parameters than BD A, and has therefore less degrees of freedom. However, the BD A also adjusts very well the experimental data even with a relative error of 11% for the asphaltenes, as this correspond of an absolute deviation of 0.38%, lower than the experimental uncertainty of 1% wt.

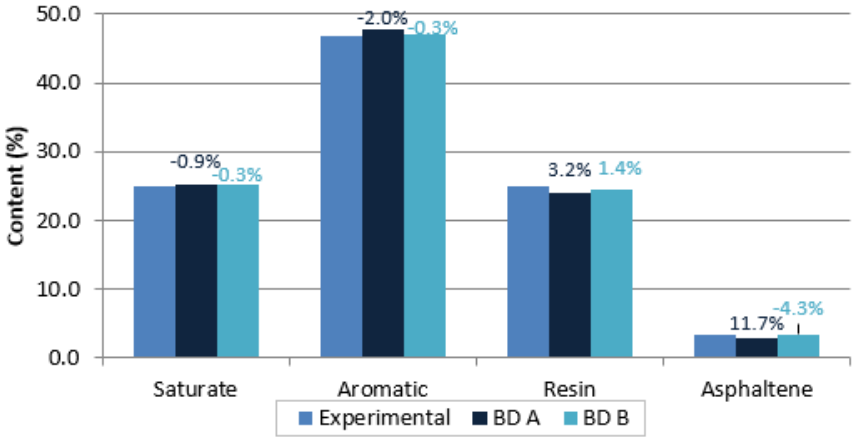


Figure 5.2: SARA analysis results SR.

Concerning the distillation curves, it is possible to affirm that both BDs adjust well to the experimental distillation curve, except for the initial and the final boiling points that are not so well predicted. This can be attributed to both the experimental uncertainty and the group contribution method used for the calculation of the boiling points.

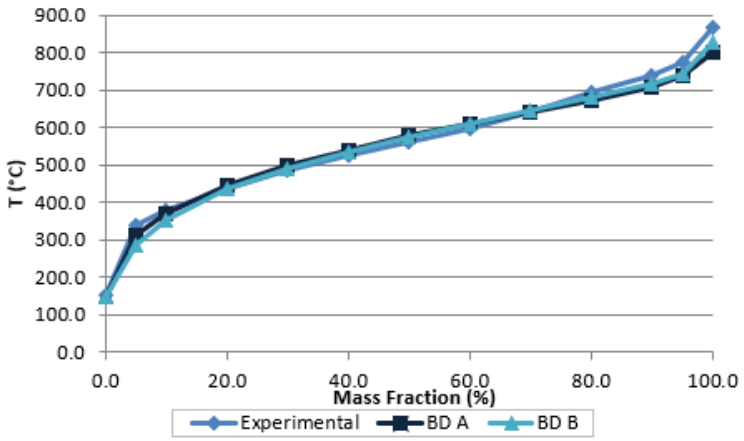


Figure 5.3: Distillation curve results SR.

Regarding Figure 5.4 relative to the TLA a perfect adjustment of the carbon content is accomplished with both BDs. Very low relative errors are observed, always below 5% and most of them below 1%. For hydrogen, the same behavior is observed for both BDs: the hydrogen content for the high-boiling polar products is systematically too high, revealing an unsuitable hydrogen distribution as function of the molecular weight (MW). Overall, the lighter lumps have less hydrogen than experimental observed, and, the heavier lumps have more hydrogen than their corresponding experimental value.

Considering the TLA for the heteroatoms (Figure 5.5), sulfur has an increasing experimental profile in function of the MW. However, this is not predicted by any of the BDs and an unsuitable distribution is revealed. Similarly, the fit for nitrogen and oxygen is not good, even considering the experimental uncertainty. Fortunately, for the HDT process, both these elements are less relevant.

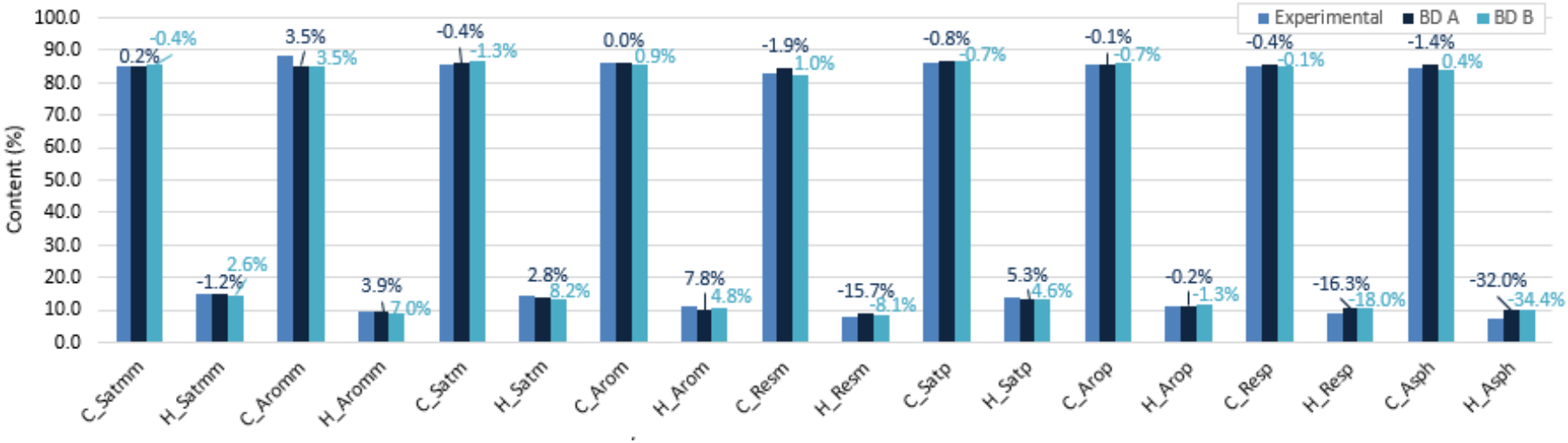


Figure 5.4: Thermidor lumps results SR for C and H.

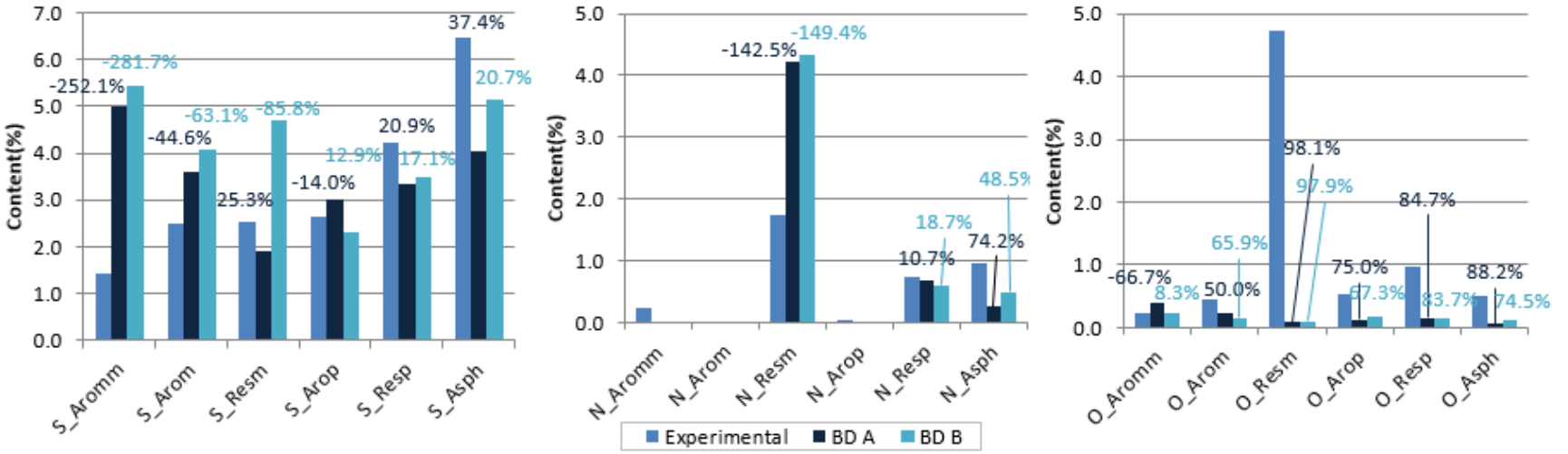


Figure 5.5: Thermidor lumps results SR for heteroatoms

Analyzing Table 5.2 in which the errors for the input analytical data are presented it is possible to conclude that the adjustment of both BDs is very similar, nevertheless BD B is slightly better. Considering the relative error of the TLA, the detailed composition that we want to predict, a better fitting needs to be reached.

Table 5.2: Comparison criteria errors.

	EA		SARA		Distillation		TLA	
BD	AQ Error (%wt)^2	AR Error	AQ Error (%wt) ^{2}	AR Error	AQ Error (°C) ^{2}	AR Error	AQ Error (%wt) ^{2}	AR Error
SRA	1.9E-02	2.2%	1.7	4.5%	8401.2	3.3%	75.7	27.4%
SRB	5.6E-02	2.0%	0.2	1.6%	7038.3	3.5%	80.1	27.8%

In Table 5.3, other properties not given to the algorithm as input analytical data but available in *Proceed Data Base* were calculated and compared. These two analyses, CCR and density, allow to verify the results of the reconstruction. Comparing the calculated values with the experimental values (Table 5.1) and considering the experimental uncertainty for the CCR analysis is 0.85% wt (corresponding to a relative error of approximately 8%) and for the density is 1.06 kg/m^3, it is possible to affirm that the CCR prediction is not good but the prediction of the density is very reasonable. The main reasons for the incorrect prediction of the CCR content is that the correlation used to calculate the CCR value is not very precise since it takes only into account the H content.

Table 5.3: CCR and density errors.

	CCR			Density		
BD	Calculated Value	AQ Error (%wt) ^{2}	AR Error	Calculated Value	AQ Error (g/cm^3) ^{2}	AR Error
SRA	5.1%	33.3	53.1%	0.978	1.0E-04	1.05%
SRB	4.7%	38.1	56.8%	0.977	8.6E-05	0.96%

It is important to note that CCR and density are not used in the simulator as input analytical data to predict the HDT performance, so the values were only calculated to see how well the algorithm fits them, even if they are not used as comparison criteria. However, those analyses are available in the *Proceed Data Base* and a sensitivity analysis of the algorithm was done to see the behavior of the mixture obtained using more input analytical data.

5.1.2 Sensitivity analysis to the input analytical data

To create the library of molecules, different input analytical data can be used. Hence, a sensitivity analysis was performed.

The available analytical data coded is given in Table 2.4. The available analytical data of the effluent fractions extracted from *Proceed Data Base* is shown in Table 5.4. These analyses are available for the global hydrotreated effluent and for the three different cuts (MDt, VGO and VR).

Table 5.4: Available properties on Proceed Data Base.

Mixture properties of *Proceed Data Base*		
Elemental Analysis Density	Conradson Carbon Molecular Weight Thermidor Lumps	Simulated Distillation SARA Analysis

In Table 5.5 are presented the results obtained from this sensitivity analysis for BD B. The first case, considering four analyses was presented in detail in the previous sub-chapter. For the case with five analyses, CCR was added to try to better fit this property since it was the analysis with the biggest error. Finally, for the case with six analyses density was also added.

Table 5.5: Sensitivity analysis AQ Error BD B.

Analyses	**EA (%wt) [2]**	**SARA (%wt) [2]**	**Distillation (°C) [2]**	**TLA (%wt) [2]**	**CCR (%wt) [2]**	**Density (g/cm^3) [2]**
4	5.6E-02	1.8E-01	7.0E+03	80.1	38.1	8.7E-05
5	3.1	19.7	2.4E+04	140.4	9.4E-01	1.4E-03
6	4.2	20.4	4.4E+04	176,1	7.6E-01	1.7E-03

Table 5.6: Sensitivity analysis AR Error BD B.

Analyses	**EA**	**SARA**	**Distillation**	**TLA**	**CCR**	**Density**
4	2.0%	1.6%	3.5%	27.8%	56.8%	1.0%
5	20.0%	14.5%	5.7%	31.6%	8.9%	3.9%
6	20.9%	7.6%	10.1%	33.8%	8.0%	4.2%

Considering the results in Table 5.5 and Table 5.6, one can conclude that when more analyses are furnished to the algorithm , less accurate is the prediction of the TLA, the analytical input necessary to the simulator. However, the algorithm should be able to fit all the analytical data without damaging the other analyses, unless the reconstruction algorithm receives contradicting information from the various analyses. Hence, these results suggest that some inconsistencies may be present in the experimental data. For this reason, a global mass balance of the tested points was performed to verify the inconsistencies between the global (total liquid analysis) and the detailed (TLA) data. The mass balance sheets can be found in 0.

In the mass balance sheets, inconsistencies were found in the elemental analysis, mainly for oxygen, as the detailed analysis show more oxygen than the global liquid analysis. This can be explained by different reasons, among then the oxidation of the effluent by air along the different analysis (distillation, SARA analysis and elemental analysis). To understand the impact of these inconsistencies, the same sensitivity analysis was performed with a perfect mass balance, in which the global analysis was made equal to the sum of the detailed analysis.

Table 5.7: Consistent sensitivity analysis AQ Error BD B.

Analyses	EA (%wt)^2	SARA (%wt)^2	Distillation (°C)^2	TLA (%wt)^2	CCR (%wt)^2	Density (g/cm^3)^2
4	6.9E-02	6.5	7.6E+03	81.5	32.1	1.3E-04
5	9.9E-02	250.5	1.3E+04	42.4	4.9E-03	4.0E-04
6	2.6E-01	138.8	4.9E+04	46.3	9.0E-04	6.8E-06

Table 5.8: Consistent sensitivity analysis AR Error BD B.

Analyses	EA	SARA	Distillation	TLA	CCR	Density
4	1.8%	6.6%	3.8%	29.3%	52.2%	1.2%
5	2.9%	28.7%	4.0%	25.9%	0.6%	2.1%
6	2.2%	15.7%	6.9%	23.3%	0.3%	0.3%

Comparing the tables Table 5.7 and Table 5.8 with consistent mass balance with the previous ones with inconsistencies, the EA instead of increasing the error by one order of magnitude remains in the same order of magnitude with a deviation ranging between 2% and 3% when adding more analyses. For the SARA analysis, the values are higher but also this global analysis accounts with the biggest experimental uncertainty. For the distillation curve, the fit is better without inconsistencies. As expected, when adding CCR and density as input data, the error decreases significantly and the impact in the fitting of the other properties is much lower. In conclusion, the algorithm works quite well when given more analyses as input data without inconsistencies.

5.1.3 Partial Conclusions

As seen in this step, SR algorithm fits almost perfectly the global analytical data, but some incorrect distributions, mainly for H and S, were found when using TLA. In other words, the algorithm adjusts very well the partial analytical data (EA, SARA and distillation curve), but for the more detailed analytical description, the adjustment could not reach the expected performance level. Therefore, in an attempt to find a better library of molecules the REM should be applied to fine tune the properties of the mixture, the results for this approach will be explained in the next sub-chapter.

Additionally, a comparison between BD A and BD B was done. For the SR step, BD B demonstrated to be more suitable to the creation of the library of molecules since the errors are lower, even though BD B has a lower number of distributions and parameters, and therefore a lower degree of freedom to adjust the data and create more suited molecules.

Concerning the sensitivity analysis made on BD B, it was demonstrated that the algorithm works independently of the amount of analytical data given when the analyses are consistent between each other. Otherwise, if inconsistence analyses are used the algorithm is pushed in different directions when trying to fit all the input data simultaneously, resulting in high deviations relative to the experimental values. Regarding this fact, the analyses utilized as comparison criteria must be well picked when inconsistent analyses are used because they will have a

big influence in the quality of the final reconstruction. Moreover, to obtain a correct reconstruction of the petroleum fractions precise information needs to be provided.

Since the goal of this SR step is to generate the best library of molecules to be used to characterize the other points of the test achieving, the best results are those obtained by using four analyses, in which CCR and density are not considered.

5.2 STEP 2: REM + MD

The objective of this second step is to try to obtain a better library of molecules, compared to the previous one obtained by the SR procedure. In this procedure, the REM will generate an optimal mixture of molecules by adjusting the molar fractions of the equimolar mixture given by SR. However, to use this library of molecules for the other points of the test, an equimolar mixture is required. To overcome this, a procedure called molecular discretization (MD) was applied. The next sub-chapter will focus on the description of this procedure.

5.2.1 MD general description

After the REM of an effluent, each molecule has its own molar fraction, leading to a non-equimolar mixture. The molecular discretization algorithm replicates each molecule n times in order to reobtain an equimolar mixture usable by the REM algorithm to reconstruct other reactor effluents. This procedure was developed by Luís Pereira de Oliveira in his PhD thesis. For deeper understanding of this procedure, the reader is referred to reference [12].

For the replication of molecules, the concept of repeat frequency (fr) is applied. The replication of molecules is done knowing that the repeat frequency is directly proportional to the molar fractions, x_i. The repeat frequency is calculated by rounded the value to the nearest integer considering a replication factor, *K*.

$$fr = ROUND\ (x_i.K) \qquad \textit{Equation 5.3}$$

The concept of the MD algorithm will be explained by means of an example. For example, in a mixture after REM, molecule A has a molar fraction of 0.8 and molecule B a molar fraction equal to 0.2. If the replication factor, *K*, is equal to 10, the algorithm will create a mixture with 10 molecules. In this case, this algorithm will create the new mixture by replicating the molecule A 8 times and the molecule B 2 times, none of them is eliminated, as illustrated in Figure 5.6. This also means that the molecules in the mixture with a molar fraction lower than 0.05 will be eliminated.

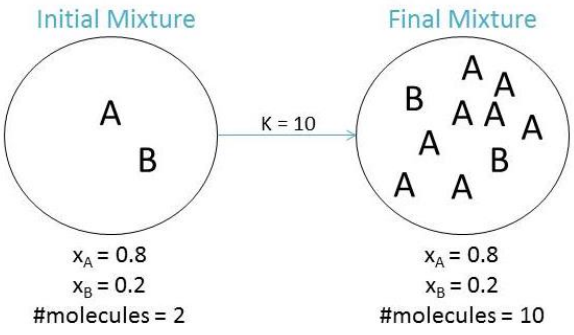


Figure 5.6: Illustration of a MD example

The replication factor, *K*, has two functions: the first one is to impose a maximum number of molecules in the discretized mixture. The second one enables to eliminate the molecules with a molar fraction lower than $1/2K$ since the procedure will round those values to nearest integer that is zero. For that reason, the value of *K* should be as high as possible, otherwise, the MD will eliminate a large number of molecules, consequently degrading the representativeness of the mixture. Meanwhile, a high value for *K* means a larger number of molecules, increasing the CPU effort and the calculation time.

With the current implementation and on a 32 bits computer with 8Mb memory, a replication factor value higher than 30 000 cannot be used. In this work, a value of 10 000 was applied to assure a good trade-off between the representativeness of the obtained mixture, CPU effort and calculation time (see Annex B).

5.2.2 Comparison of the REM results between building diagrams A and B

To compare to the results from step 1, the SR+REM was performed on the same input analytical data, i.e. the EA, SARA distillation, and TLA before the MB corrections.

In Figure 5.7 is represented the REM results for the EA. The general tendency is a perfect fit for C, H, and S independently of the BD used. An almost perfect fit is achieved for N. For O, the experimental value with the biggest inconsistence in the MB, the algorithm has problems to fit this value because it is trying to adjust at the same time EA and TLA with different oxygen content values.

Figure 5.8 and Figure 5.9 show the comparison between the relative errors for the library of molecules obtained applying only the SR (step 1) and applying the SR+REM (step 2) for BD A and BD B respectively. In both figures, a strong decrease of the errors is observed after the REM procedure, except for O. Due to the MB inconsistences, an error of around 80% is found for BD A (absolute error equal to 0.11%) and around 140% for BD B (absolute error equal to 0.18%). Those errors have a great impact in the AR error: without inconsistencies the AR error would be around 2% for both schemes and with inconsistencies is equal to 17% for the BD A and 29% for BD B (Table 5.10).

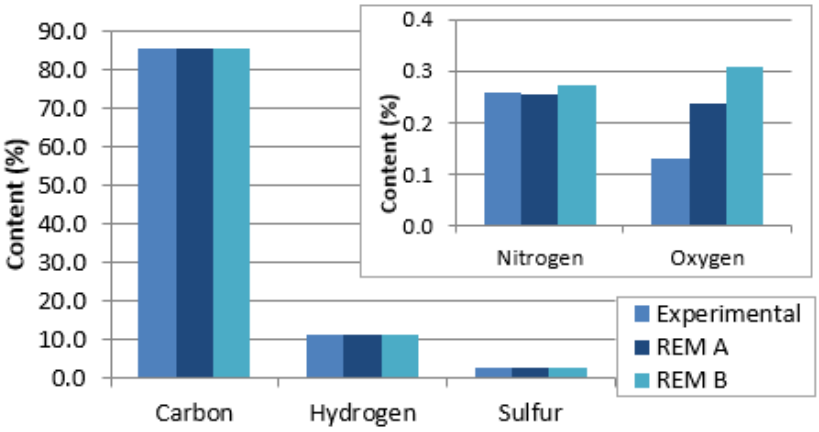


Figure 5.7: Elemental analysis results REM.

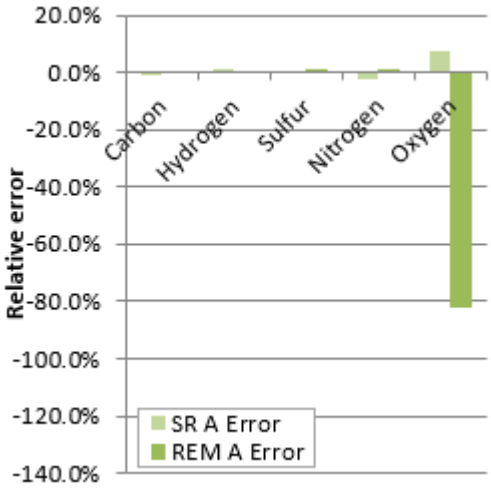


Figure 5.8: EA Errors comparison BD A.

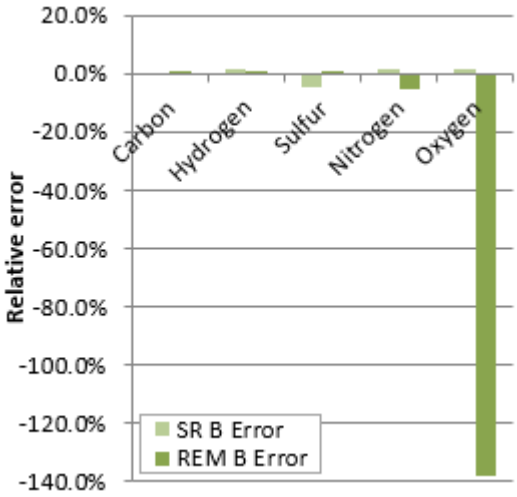


Figure 5.9: EA Errors comparison BD B.

As illustrated in Figure 5.10, a perfect fit is reached by the algorithm for the SARA analysis by adjusting the molar fractions of the molecules obtained in the SR step. Furthermore, in Figure 5.11 and Figure 5.12, is observed the reduction of the error after REM procedure for both BDs, mainly for the BD A because the SR error was bigger. For the BD A, all errors are very low, while for BD B only the absolute deviation for the asphaltenes remains large.

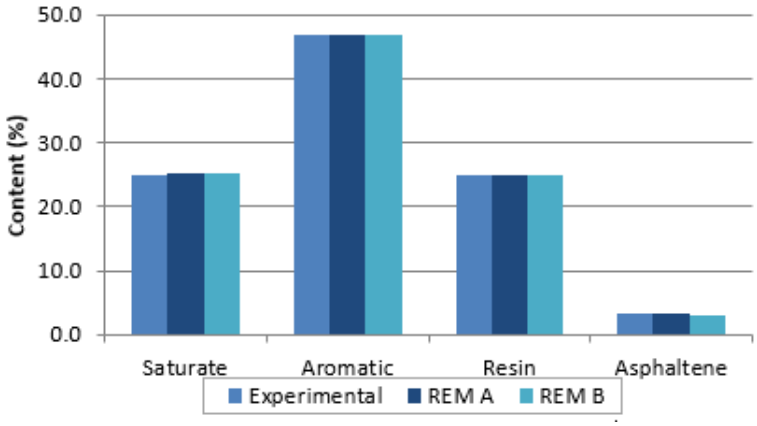


Figure 5.10: SARA analysis results REM.

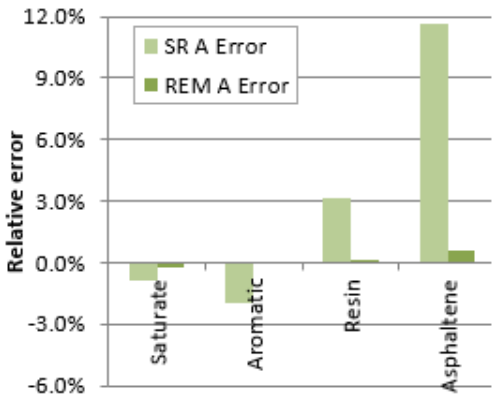


Figure 5.11: SARA Errors comparison BD A.

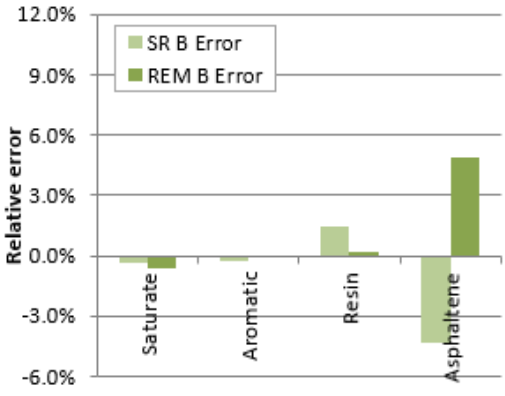


Figure 5.12: SARA Errors comparison BD B.

Figure 5.13 shows the behavior of the distillation curve when applying SR-REM to BD A and compares these results with those for the SR. It is verified that the adjustment is better after the REM procedure. Additionally, the last point of the distillation curve cannot be improved because the library of molecules created in the SR does not have molecules with a higher BP so the same value is found for the SR mixture and for the SR+REM mixture.

The same trend is found for BD B in Figure 5.14. However, a worst fit is observed for the second point of the distillation curve. This means that, after applying the REM, a higher molar fraction was attributed to the light molecules leading to a lower BP for the 5%wt point. This adjustment remains suitable, because the simulator uses the lump, in this case the MDt lump that covers the molecules between the IBP and 350°C.

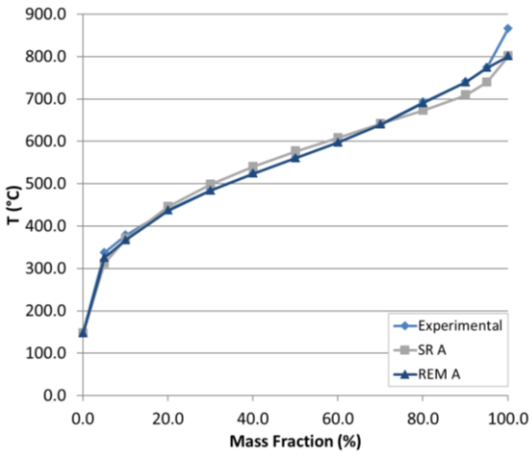


Figure 5.13:Distillation curve results REM A.

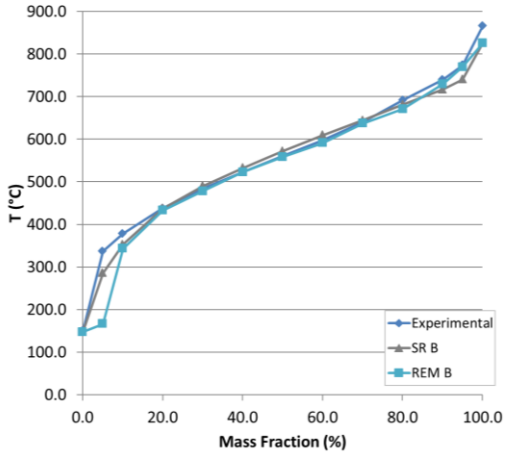


Figure 5.14: Distillation curve results REM B.

The TLA results for carbon are illustrated in

Figure *5.15*. One can observe a perfect fit, similar to those created by the SR procedure: all the relative errors are lower than 5%, which is a good value for the application of the algorithm due to the propagation of errors in this particular analysis.

Figure 5.16 presents the comparison for BD A of the fit for the C content applying SR or SR+REM, while the same comparison is present for BD B in Figure 5.17. As shown in these two figures, the C distribution has lower errors in the majority of the data, accumulating all deviations in the C content of the VGO resins (for BD A) or in the C content of the VGO resins and the asphaltenes (for BD B).

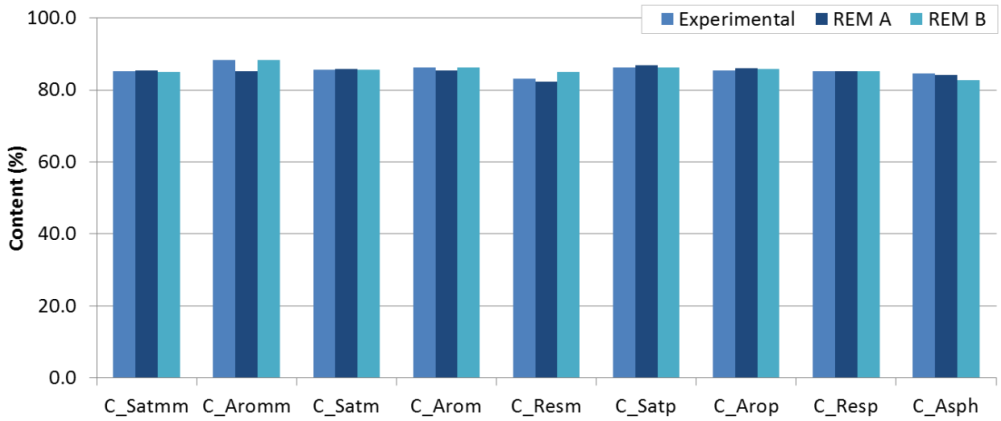


Figure 5.15: Thermidor lumps results REM for C.

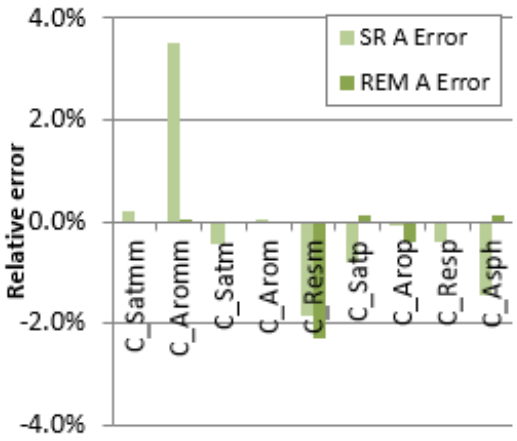


Figure 5.16: C Errors comparison BD A.

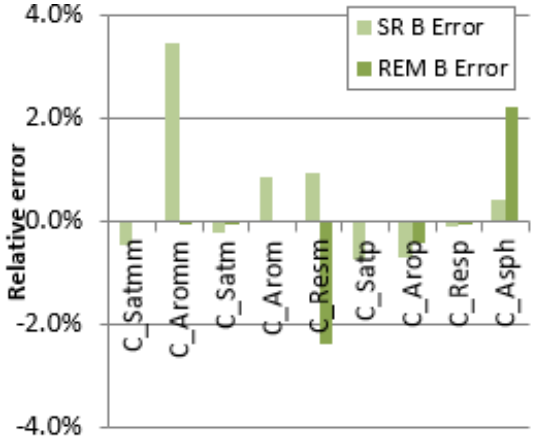


Figure 5.17: C Errors comparison BD B.

For hydrogen, as observed in Figure 5.18 the fit for the lighter lumps is perfect, however, for the heavier lumps the distribution was not so accurate, these molecules remain less aromatic than they should.

As shown in Figure 5.19 and Figure 5.20, the result found for carbon is also obtained for the hydrogen distribution: the errors in the hydrogen content of the various lumps were almost reduced to zero by adjusting the molar fractions of the molecules, except the resins of the VGO, which now accounts for the biggest deviation, significantly above 5%.

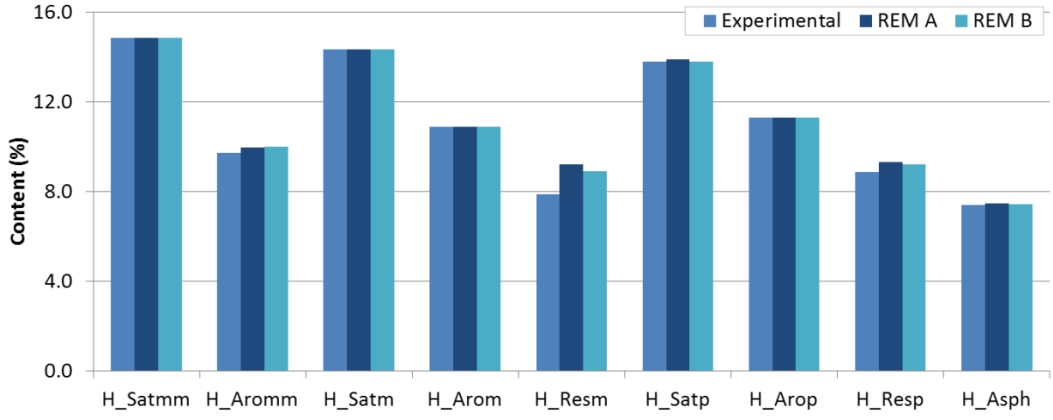


Figure 5.18: Thermidor lumps results REM for H.

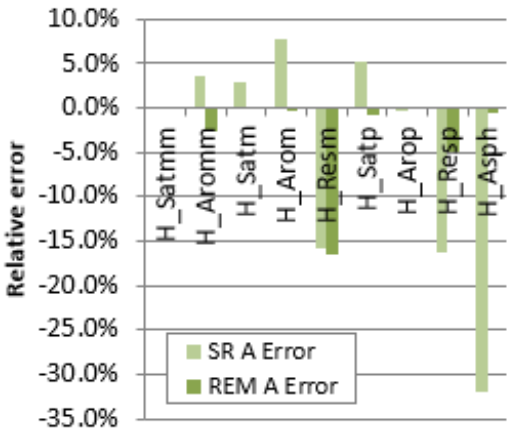


Figure 5.19: H Errors comparison BD A.

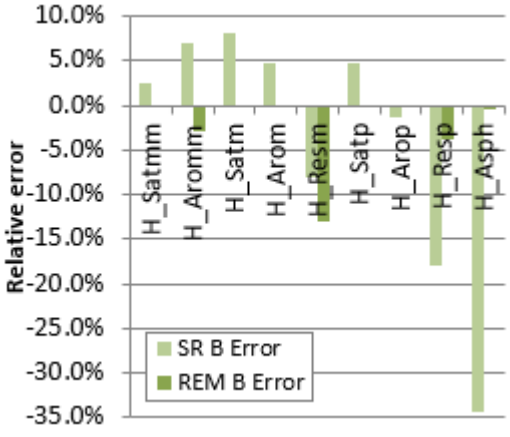


Figure 5.20: H Errors comparison BD B.

Comparing the sulfur distribution shown in Figure 5.21 (SR-REM) to the sulfur distribution in with Figure 5.5 (SR only) shows that, when REM is applied, the experimental sulfur profile can be achieved independently of the BD used, thereby significantly improving the fit (Figure 5.22 and Figure 5.23). After the optimization by entropy maximization, both diagrams result in a sulfur distribution with almost zero errors, except for the sulfur content on VGO resins and of the asphaltenes. BD A performs somewhat better, especially for the asphaltenes.

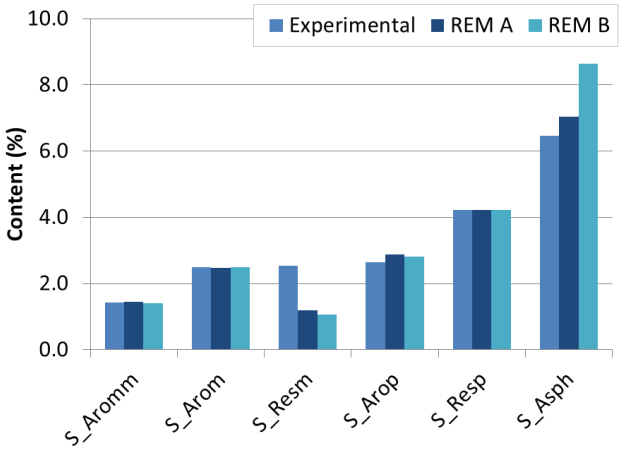


Figure 5.21: Thermidor lumps results REM for S.

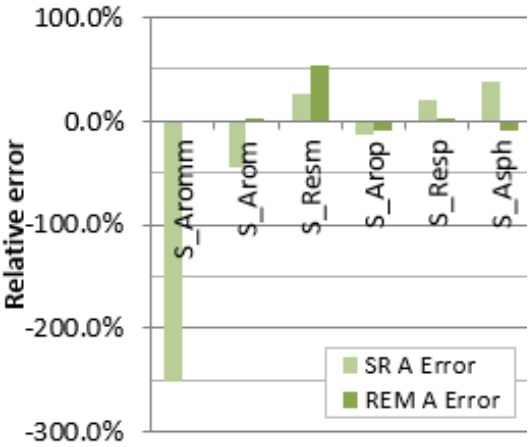


Figure 5.22: S Errors comparison BD A.

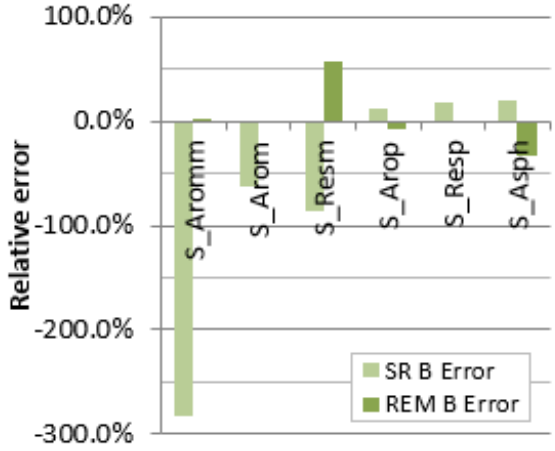


Figure 5.23: S Errors comparison BD B.

For nitrogen and oxygen, when comparing both BDs after REM (Figure 5.24 and Figure 5.25), very similar results are observed. Concerning the comparison between the SR and the SR+REM methods (Figure 5.26 to Figure 5.28), both BDs reduce the error for the nitrogen content in the asphaltenes to zero, negatively affecting the nitrogen content of the two resin fractions. For oxygen, BD A successfully reduced all errors, achieving an almost zero error for the lighter aromatic lumps and the heavier resins through REM, while the remaining lumps still present a somewhat larger deviation. However, BD B still exhibits very big deviations for three out of the six lumps.

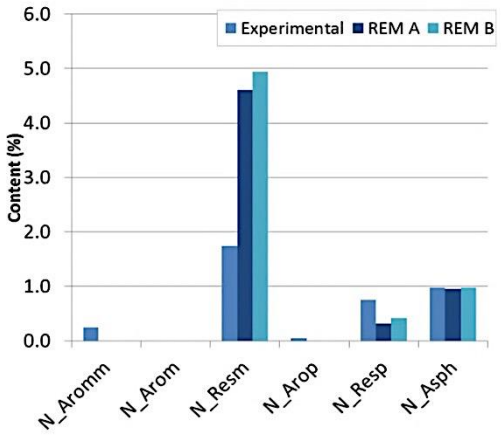


Figure 5.24: Thermidor lumps results REM for N.

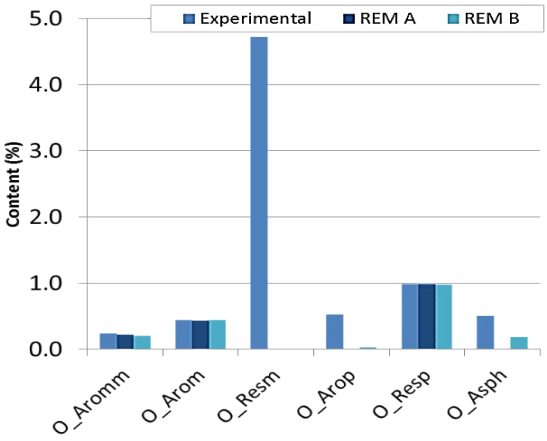


Figure 5.25: Thermidor lumps results REM for O.

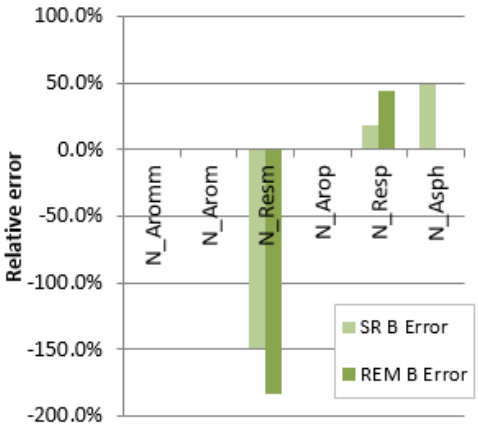


Figure 5.26: N Errors comparison BD A.

Figure 5.27: N Errors comparison BD B.

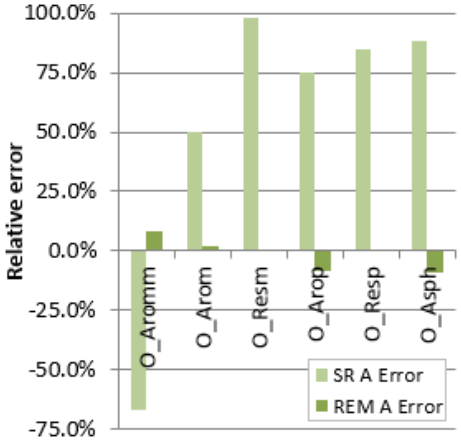


Figure 5.28: O Errors comparison BD A.

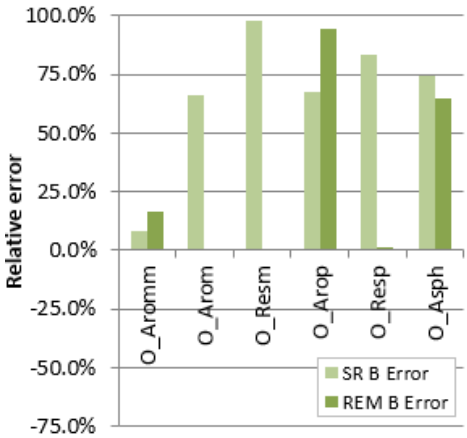


Figure 5.29: O Errors comparison BD

Analyzing Table 5.9 and Table 5.10, which provide the global comparison, the REM is validated as an optimization procedure to create a more accurate library of molecules. It is also important to underline that the mixture of molecules obtained by the BD A had worst results for the SR only. After applying the REM, the results are better for BD A.

Table 5.9: Comparison SR and SR+REM AQ Errors.

Analyses	EA (%wt) [2]	SARA (%wt) [2]	Distillation (°C) [2]	TLA (%wt) [2]	CCR (%wt) [2]	Density (g/cm^3) [2]
SR A	1.9E-02	1.7	8.4E+03	75.7	33.3	1.0E-04
REM A	1.5E-02	3.9E-03	4.5E+03	39.2	12.7	1.3E-04
SR B	5.6E-02	1.8E-01	7.0E+03	80.1	38.1	8.7E-05
REM B	5.2E-02	5.6E-02	3.3E+04	48.8	14.2	6.6E-05

Table 5.10: Comparison SR and SR+REM AR Errors.

Analyses	EA	SARA	Distillation	TLA	CCR	Density
SR A	2.2%	4.5%	3.3%	27.4%	53.1%	1.1%
REM A	17.0%	0.2%	1.3%	7.9%	32.8%	1.2%
SR B	2.0%	1.6%	3.5%	27.8%	56.8%	1.0%
REM B	29.0%	1.5%	5.8%	12.3%	34.7%	0.8%

5.2.3 Comparison between REM and MD

As mentioned above, the MD procedure was applied to the REM results. The obtained results will not be shown in detail, but as expected, the lower the replication factor lower is the accuracy of the results. A replication factor *K* equal to 10 000 was shown to give results closer to the REM (Annex B).

This comparison is critical to assure a good representativeness of the mixture. Indeed, as previously explained, this methodology deletes all the molecules with a molar fraction inferior to a threshold that is a function of the replication factor. For example, with a replication factor of 10 000 all the molecules with molar fraction inferior to $5*10^{-5}$ will be deleted on MD procedure (Figure 5.30). Hence, some characteristics can disappear by eliminating molecules that are present in low concentration. The number of eliminated molecules will therefore have an impact on the MD results by decreasing the representativeness of the mixture.

Figure 5.30 shows the distributions in molar fraction for the molecules after REM for each BD. BD B has a more homogeneous distribution, which leads to deleting 3207 of 5000 molecules, or 64% of the mixture, when applying MD with a replication factor of 10 000. BD A has a molar fraction distribution in which molar fractions above $1*10^{-5}$ predominated. Hence, fewer molecules are deleted during MD, only 1754 molecules, representing 39% of the mixture, after MD with a replication factor of 10 000.

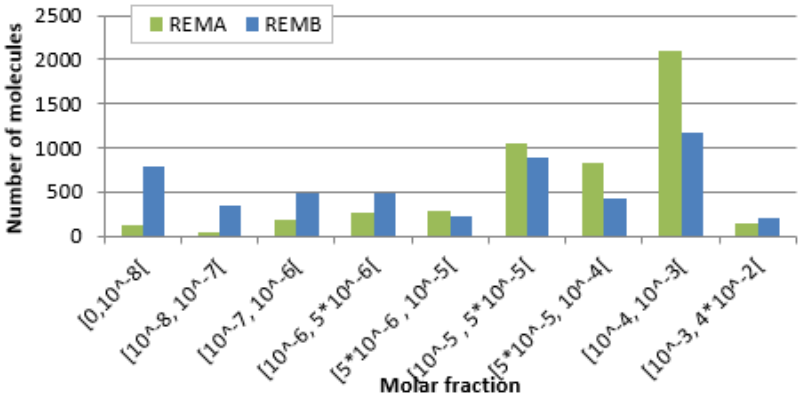


Figure 5.30: Molar fraction distributions of the REM mixtures.

Table 5.11: Comparison REM and MD AQ Error.

Analyses	EA (%wt) [2]	SARA (%wt) [2]	Distillation (°C) [2]	TLA (%wt) [2]	CCR (%wt) [2]	Density (g/cm^3) [2]
REM A	1.5E-02	3.9E-03	4.5E+03	39.2	12.7	1.3E-04
REM + MD A	1.1E-02	3.6E-01	4.4E+03	44.2	4.9E-03	1.1E-04
REM B	5.2E-02	5.6E-02	3.3E+04	48.8	14.2	6.6E-05
REM + MD B	4.6E-02	2.4E-01	4.1E+04	60.1	9.0E-04	5.5E-05

Table 5.12: Comparison REM and MD AR Error.

Analyses	EA	SARA	Distillation	TLA	CCR	Density
REM A	17.0%	0.2%	1.3%	7.9%	32.8%	1.2%
REM + MD A	14.8%	1.1%	1.1%	13.1%	0.6%	1.1%
REM B	29.0%	1.5%	5.8%	12.3%	34.7%	0.8%
REM + MD B	28.2%	1.2%	7.0%	17.6%	0.3%	0.8%

As a conclusion, only slight differences are observed REM and REM+MD. Hence, the MD procedure used to return an equimolar mixture is validated using the replication factor of 10 000 and it was shown that the small loss in representativeness will not affect the results greatly.

5.2.4 Partial Conclusions

It was seen in this sub-chapter that applying the REM to the mixture globally improves the results, as was expected, because this procedure optimizes the mixture by adjusting the molar fractions of the molecules in order to be closer to the experimental data. However, issues are found on the oxygen content due to the mass balance inconsistencies.

According to the results in Section 5.2.2, BD A has the best results after applying the REM procedure. Therefore, this diagram builds better molecules. This is verified in Figure 5.30, in which the molar fraction distribution has a low dispersion, indicating that the REM procedure only introduces a strong variation in the mole fraction for only

a small number of molecules. For BD A, the molecules are mostly in the high concentration range, so, the majority of the molecules are important to achieve properties similar to the experimental data values. On other hand, for BD B, the molar fraction distribution has a higher dispersion, meaning that some of the created molecules are highly representative, receiving a higher molar fraction in the REM method, to less suitable ones to which a lower molar fraction is attributed.

The MD method with the chosen replication factor of 10 000 deletes 39% of the molecules obtained after REM A and 64% of the molecules of REM B. Nevertheless, considering the results of Section 5.2.3, the mixture obtained after MD ensures a good representativeness, because in both cases the results obtained after MD only differ slightly from the REM results.

The application of this new procedure (SR+REM+MD) has as main advantages: the achievement of a proper sulfur profile and a better fitting of the hydrogen distribution, therefore decreasing the TLA error and ensuring a better reproduction of the experimental data.

5.3 STEP 3: Effluent (E_i) Application Cases

The procedure illustrated in Figure 5.31 is proposed to reconstruct a detailed composition of the effluents (E_i) based on the pre-created library of molecules, while exploring the effect of the number of input analytical data on the correct reproduction of the effluent.

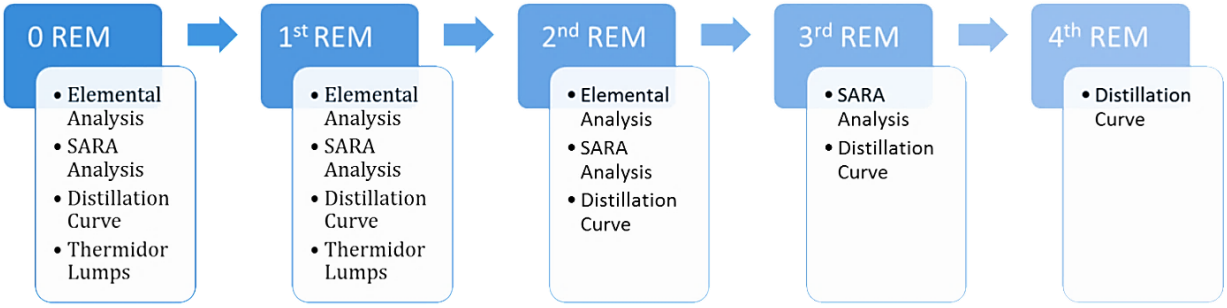


Figure 5.31: E_i reconstruction procedure scheme.

Case zero and the first case use all the analytical data that must be given to the simulator. In those two cases, there is no reducing of the analytical work. However, both cases enable to understand the procedure and allows to compare the results from the maximum input analytical data to the minimum input analytical data.

The difference between them lies in the mass balance, for the 1st REM the mass balance is inconsistent while for the 0 REM case, the mass balance was adjusted to obtain perfectly consistent data (inputs equal to outputs – see Annex E). Thereby, an overall assessment of the effect of inconsistencies is obtained. The most relevant case is the second one (2nd REM) in which the REM is applied taking as input the global analytical data necessary to the simulator to predict the TLA. The third and fourth cases are limiting cases in which a strongly reduced amount of information is given.

Step 3 has been implemented for three different points: point 1, point 2, point 3. To simplify the discussion of the results, and given the similarity of the results and procedures, only the first point will be present in detail. Because of the existence of four library of molecules (step 1 or step 1+2 for BD A or BD B), a series of results will be shown and discussed.

5.3.1 Point 1: REM Results

5.3.1.1 0 REM Case

The 0 REM case is the case in which the perfect mass balance for point 1 is explored (Annex E). This case allows to verify the possible problems of the BD used to construct the molecules that could be hidden by the inconsistencies presents in the other cases.

For the EA, as shown in the two subsequent images, the step 1 library, or SR library, (Figure 5.32) presents better results than the step 1+2 library, or SR+REM+MD library, (Figure 5.33). Since the REM procedure fine-tunes the mixture so as to reduce the inconsistencies in the library used, the atoms present in lower quantities, N and O, may be deleted during the MD procedure, limits the possibilities to adjust both atoms due to the reduction in number of species. Analyzing each figure separately, for step 1, a better adjustment is accomplished by BD A. However, for step 1+2, BD B presents the best adjustment after the MD procedure. Overall, the library obtained by BD A after the SR step (step 1) is the best library of all.

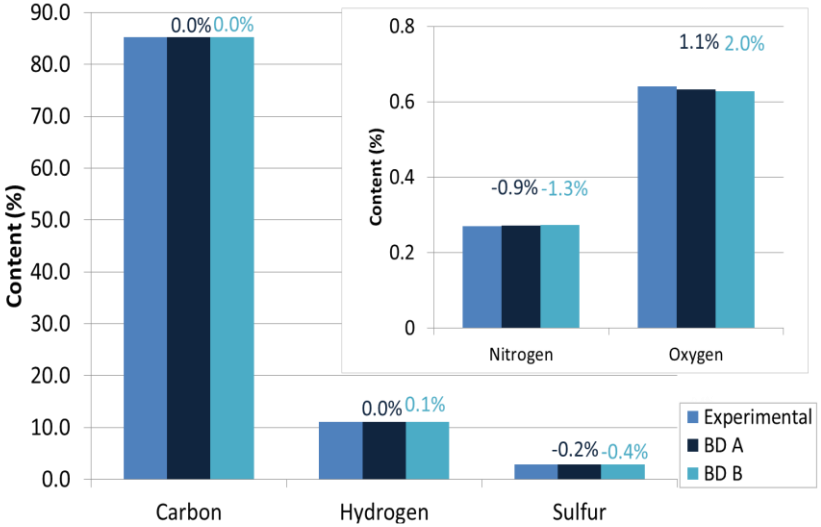


Figure 5.32: EA results 0 REM (step 1).

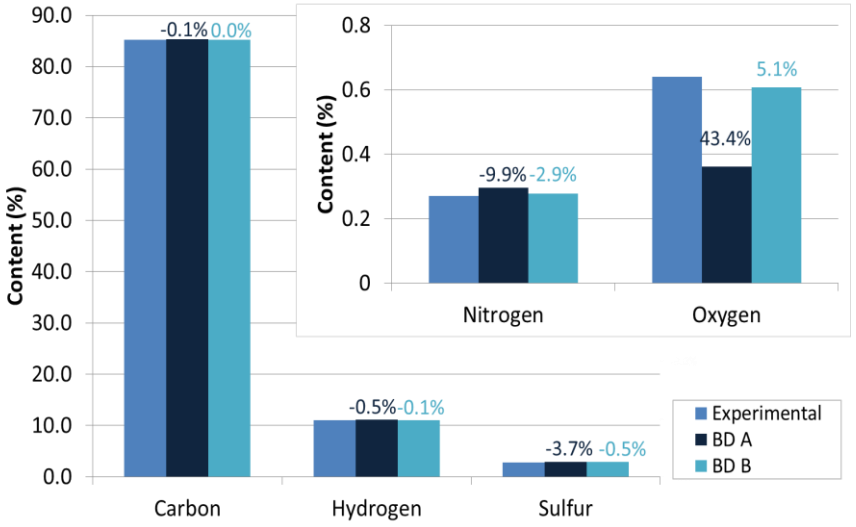


Figure 5.33: EA results 0 REM (step 1+2).

On the following page, the results for the SARA analysis and the distillation curve are illustrated. Comparing the SARA analysis for both approaches (Figure 5.34 and Figure 5.36), a perfect reproduction of SARA families is observed. The main difference observed is on the prediction of the asphaltenes, since they are presented in lower quantities (< 4%). Furthermore, the improvement when applying the step 2 to BD A, in which the MD procedure allows to fine tune the library of molecules, must be underlined. Also, for this analysis, the best adjustment was found for BD A using the second approach (step 1+2).

The agreement for the distillation curve, is illustrated in Figure 5.35 and Figure 5.37. BD A undergoes large deviations when the optimized library of molecules is used (step 1+2), because in this library, the adjusted molar fractions favors saturates with a lower BP. The best results are given by the BD B in both approaches (step 1 and step 1+2), since the second approach does not substantially affect the result. Furthermore, since more molecules are deleted less saturates are presented in the mixture leading to a higher BP for the molecules and ensuring a good distillation curve distribution. Also, since the molecules in the library have a BP varying between 148.1°C and 801.4°C for all the cases, the initial BP of this point (151°C) can be reached, but the final BP of 854.8°C cannot be reached due the maximum BP of the library.

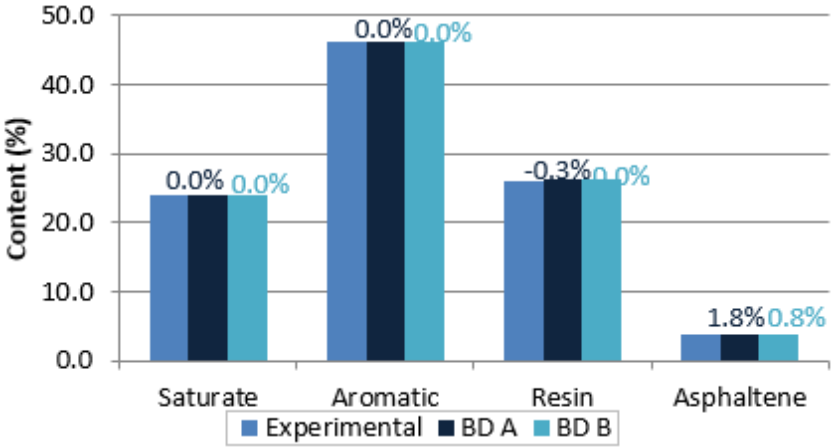


Figure 5.34: SARA analysis 0 REM (step 1).

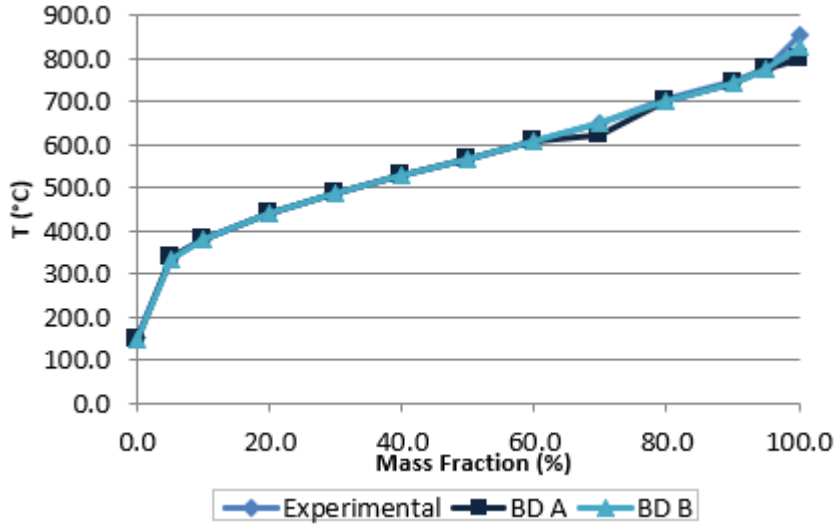


Figure 5.35: Distillation curve 0 REM (step 1).

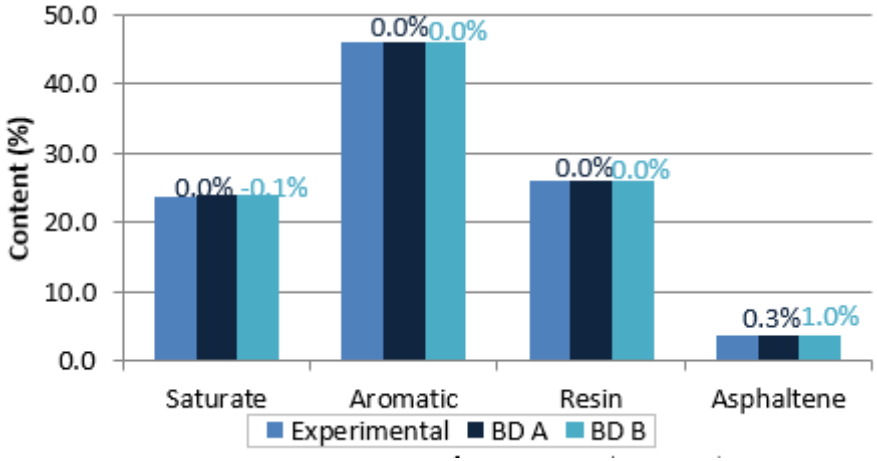


Figure 5.36: SARA analysis 0 REM (step 1+2).

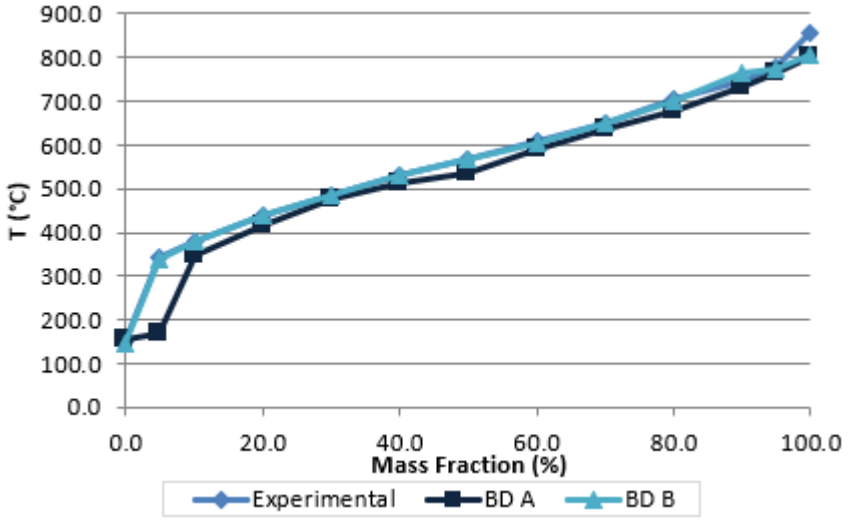


Figure 5.37: Distillation curve 0 REM (step 1+2).

As shown in Figure 5.38 and Figure 5.40, for the carbon and hydrogen distribution of the TLA, both C and H present a good adjustment, independently of the used molecule library (as the deviations are lower than 5%). However, both schemes and approaches fail to provide a satisfactory reconstruction of the H in Res^-. When comparing the results between approaches, the errors of the library constructed by only applying the stochastic reconstruction are smaller due to the higher representativeness of the mixture since the REM procedure disfavors the molar fraction of the Res^- fraction, because of the lower amount present.

Observing the results for the heteroatoms (Figure 5.39 and Figure 5.41), the results between schemes are very close due to the similarity of the distributions to add the atoms: S, N and O. For S, the increasing profile of the content is achieved for the SR library of molecules used and it is maintained for the step1+2 library of molecules. For both approaches the calculated contents of Res^- and asphaltenes are unsuitable. Hence, the S in asphaltenes is over-predicted and S in the Res^- is under-predicted. Moreover, for Res^-, even though molecules with S are present in both libraries, the MD procedure tends to eliminate them due to the low content of this lump, revealing an inappropriate construction for this point.

For N and O, the performance of the results for the point follows the same trend observed for the various libraries of molecules, so the same unsuitable contents are found. For the second approach, due to the loss in representativeness after applying the MD procedure, the errors are larger than those from the first approach, because it is more difficult to make the adjustment when the population of molecules is reduced. Usually, since these atoms are present in lower quantities the molecules in which they were introduced during the construction of molecules have a lower molar fraction after the REM procedure, and can eventually be eliminated of the mixture during MD. In summary, it is assumed as the best approach corresponds to step 1 using BD A due to the lower relative errors.

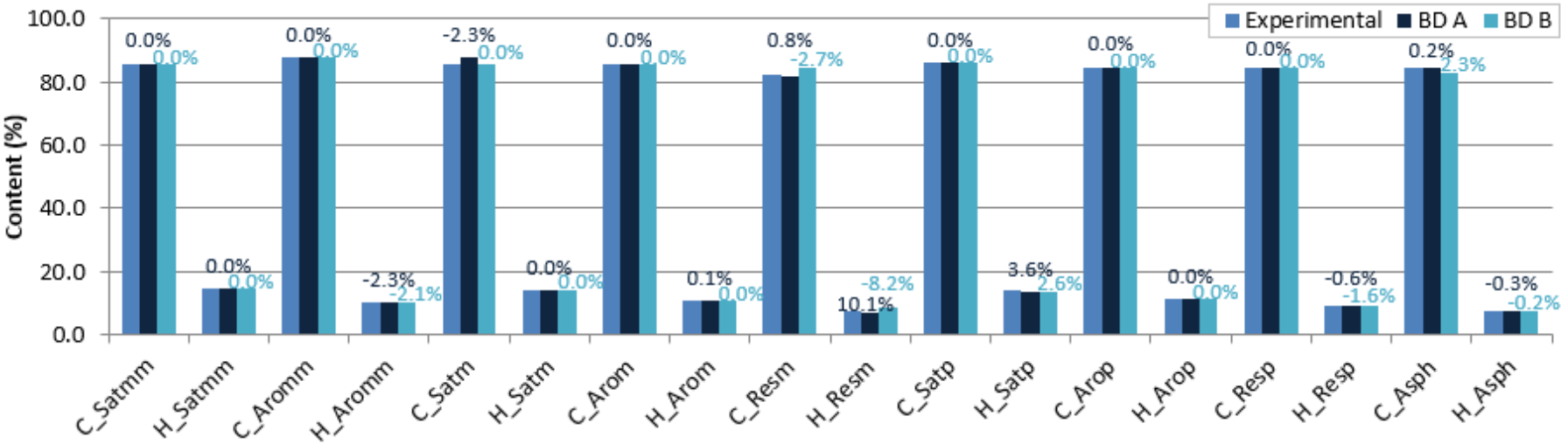


Figure 5.38: Thermidor lumps 0 REM for C and H (step 1).

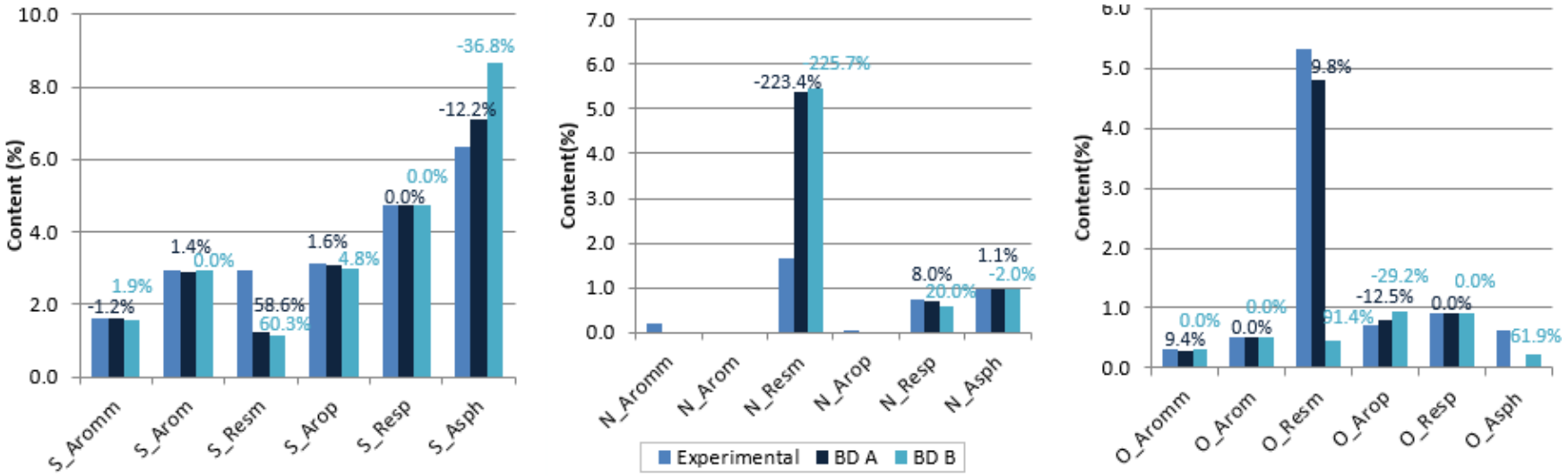


Figure 5.39: Thermidor lumps 0 REM for heteroatoms (step 1).

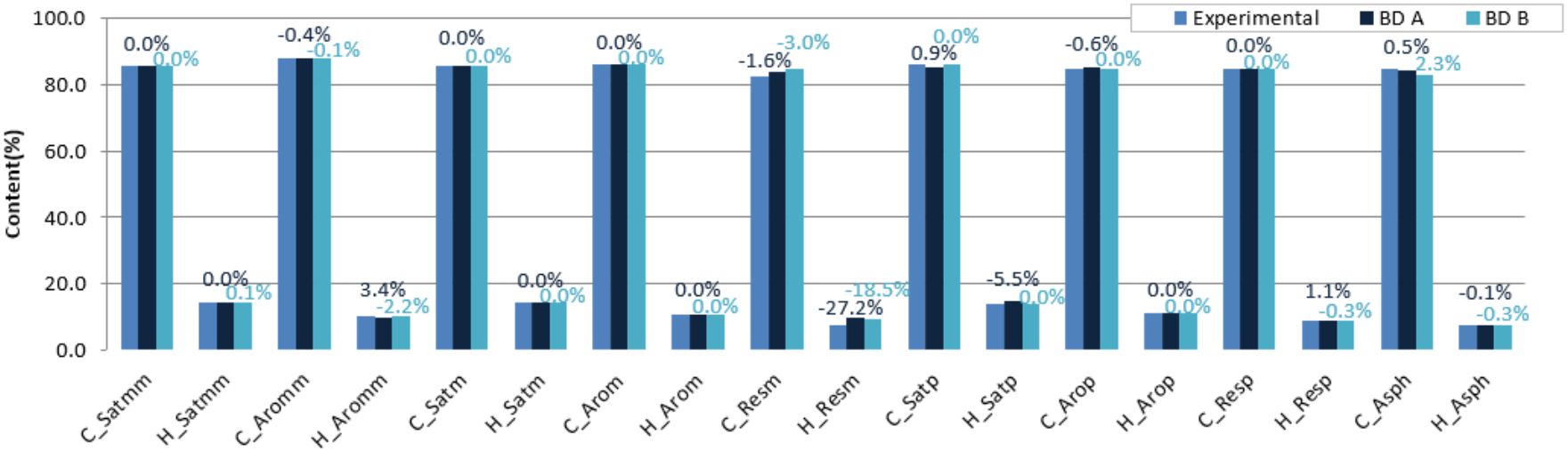


Figure 5.40: Thermidor lumps 0 REM for C and H (step 1+2).

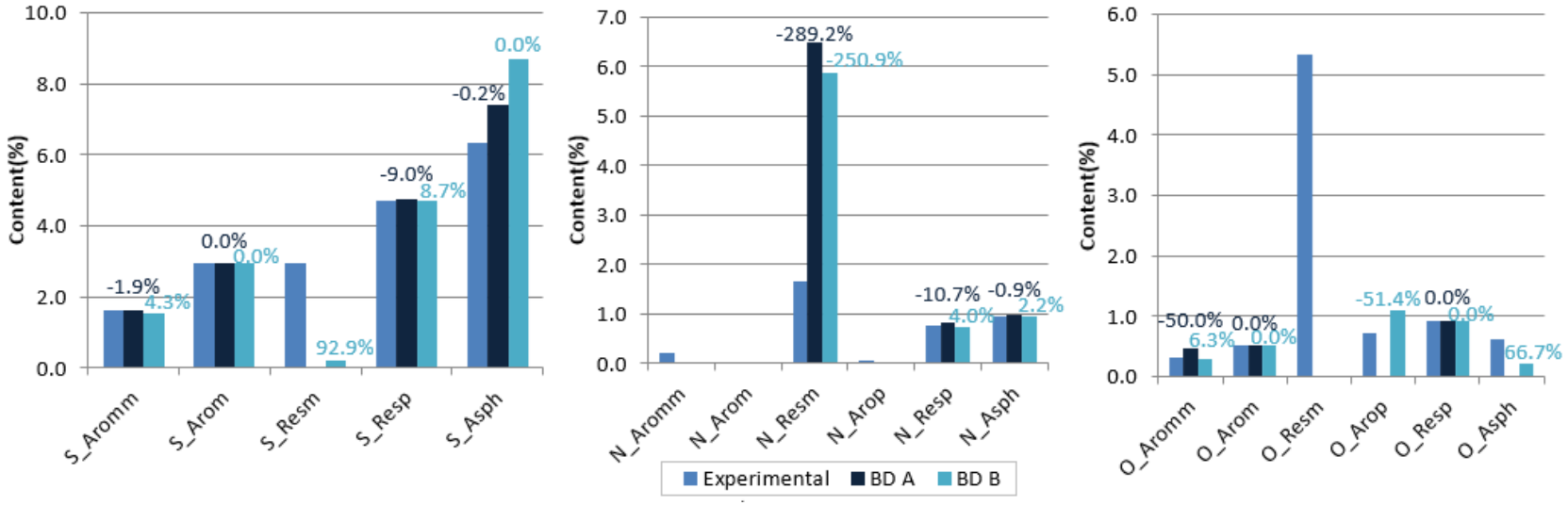


Figure 5.41: Thermidor lumps 0 REM for heteroatoms (step 1+2)

5.3.1.2 1st REM Case

The 1st REM case is the one in which the actual measured data are used as is, even if the introduce an inconsistent mass balance. In this case, the detailed analytical data is used as comparison criteria. Hence, no reduction of analytical data is achieved, but this case is important to understand the behavior of the algorithm. In Figure 5.42 and Figure 5.43 , the EA results are illustrated. Comparing with the previous case (with a consistent mass balance), for O the errors are bigger due to the inconsistencies between the EA and TLA. However, because it is reconstructed from an inconsistent library of molecules, the errors are lower than in the previous case. Additionally, the step 1+2 presents better results due to the pre-optimization performed on the set of molecules utilized.

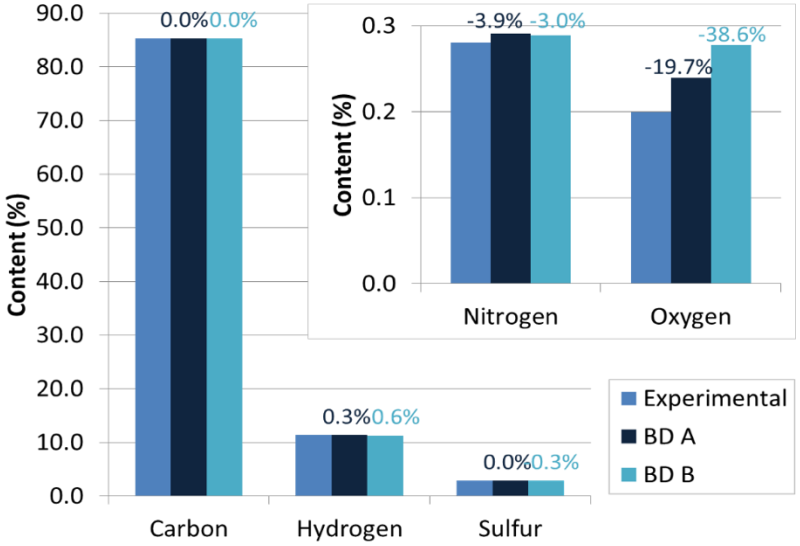


Figure 5.42: Elemental analysis results 1st REM (step 1).

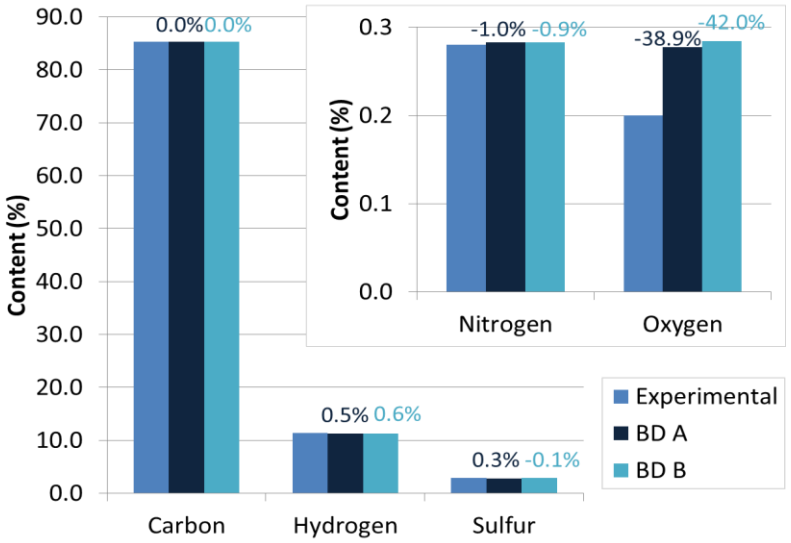


Figure 5.43: Elemental analysis results 1st REM (step 1+2).

For the SARA analysis, the results are almost the same when comparing to the 0 REM case, since there was no inconsistence for SARA families. Hence, the adjustment is almost perfect (Figure 5.44 and Figure 5.45). Considering the distillation curve obtained with BD B, the prediction of the MDt cut for both approaches is less suitable for lower BP. This is consistent with the higher content of saturates found on SARA results. The best fit is reached with BD A, using the non-optimized molecule library.

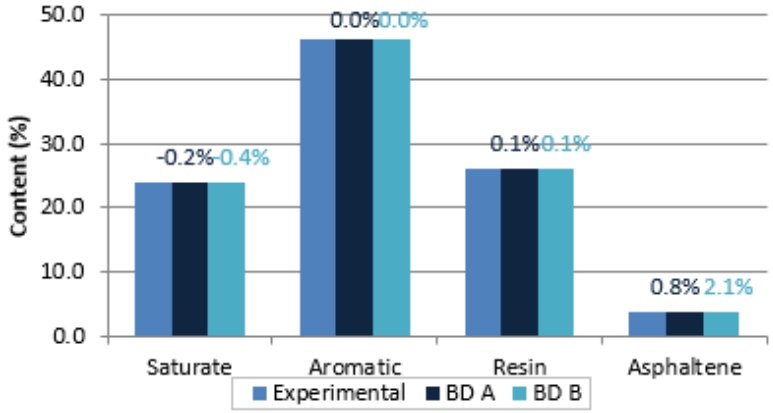


Figure 5.44: SARA analysis results 1st REM (step 1).

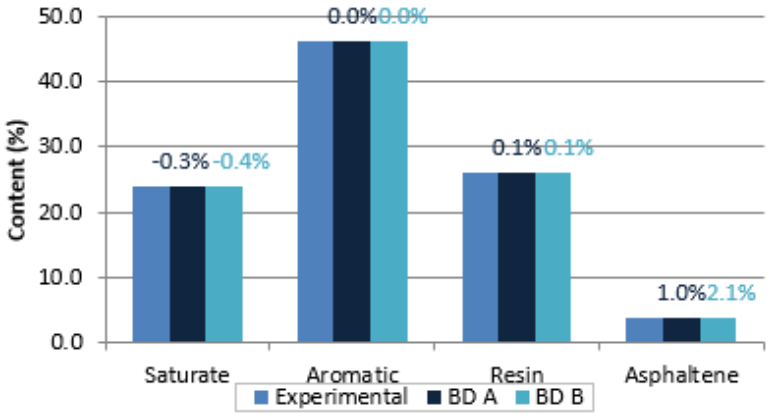


Figure 5.45: SARA analysis results 1st REM (step 1+2).

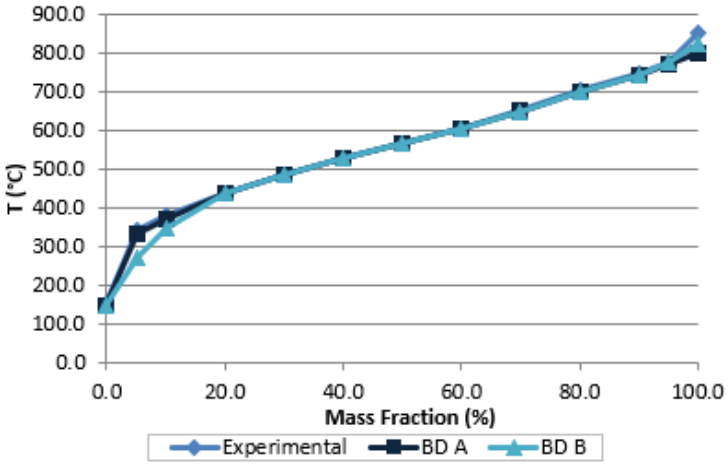


Figure 5.46: Distillation Curve 1st REM (step 1).

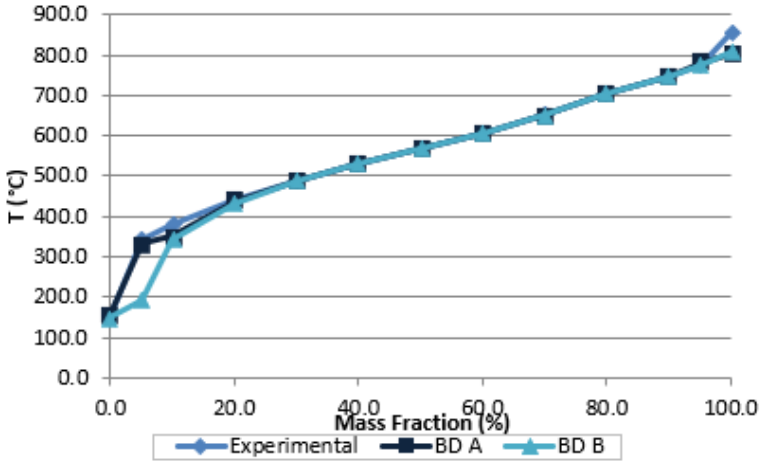


Figure 5.47: Distillation Curve 1st REM (step 1+2).

Again, as for the previous case, the Res$^-$ have a poor adjustment, mainly because this lump is calculated by difference between the total and the sum of the other lumps. As expected, the errors for the O content increased significantly because of the inconsistent data between the TLA and EA, as it is impossible to find a value for the O content that assures a correct reproduction of both experimental values. It is possible to conclude, that the higher the inconsistencies the more difficult it is to assure a correct distribution of the lump contents. In general, it is concluded that the best results were achieved for the first approach (step 1) using the BD A.

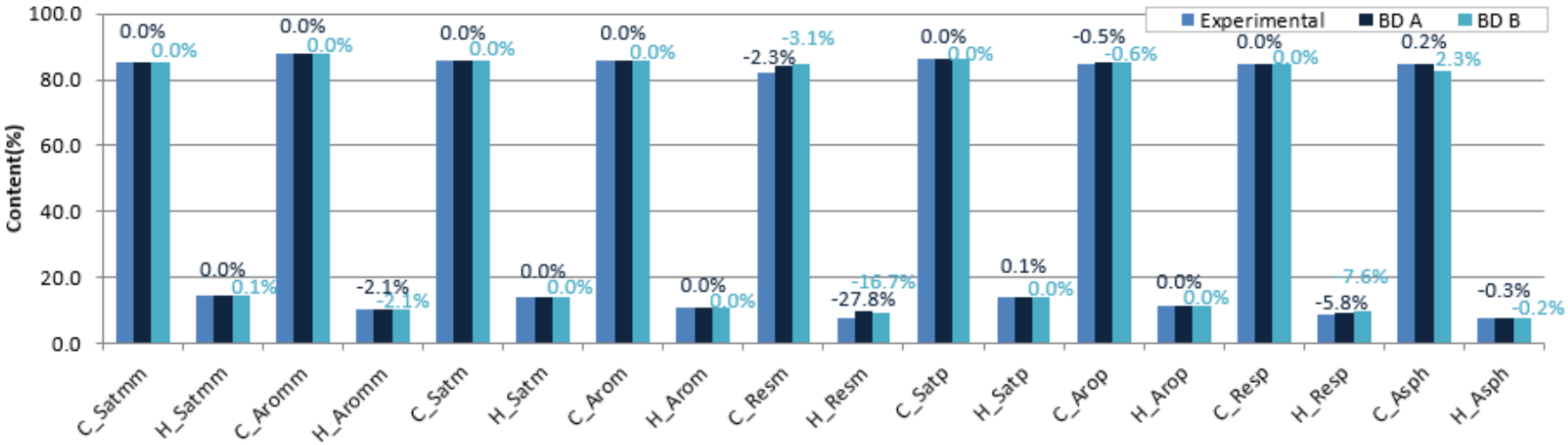


Figure 5.48: Thermidor lumps 1st REM for C and H (step 1).

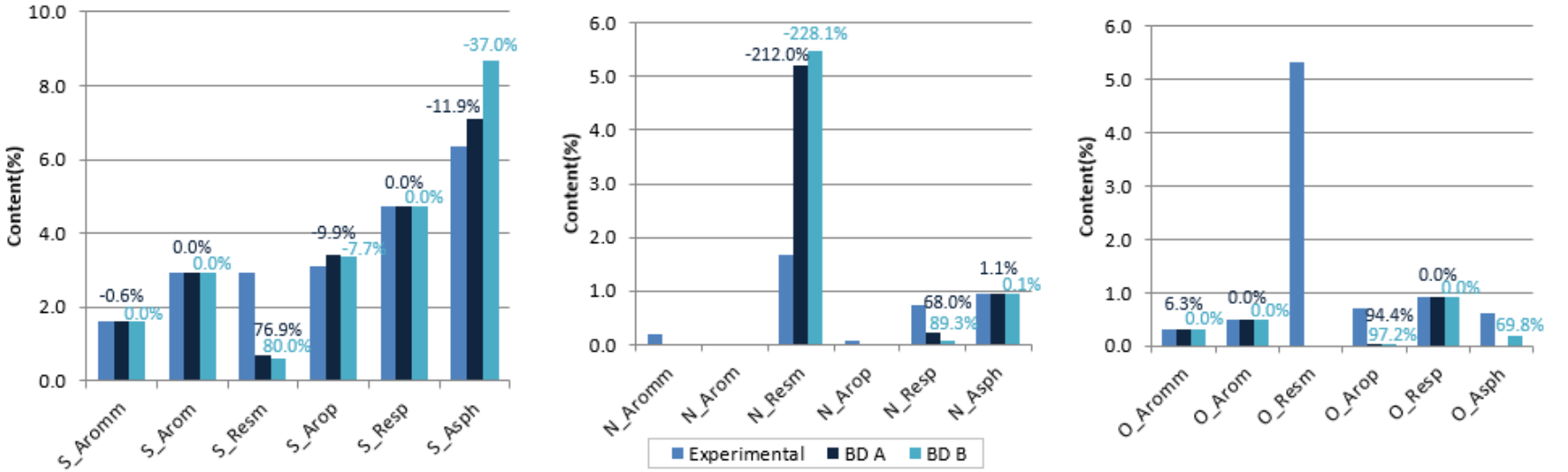


Figure 5.49: Thermidor lumps 1st REM for heteroatoms (step 1).

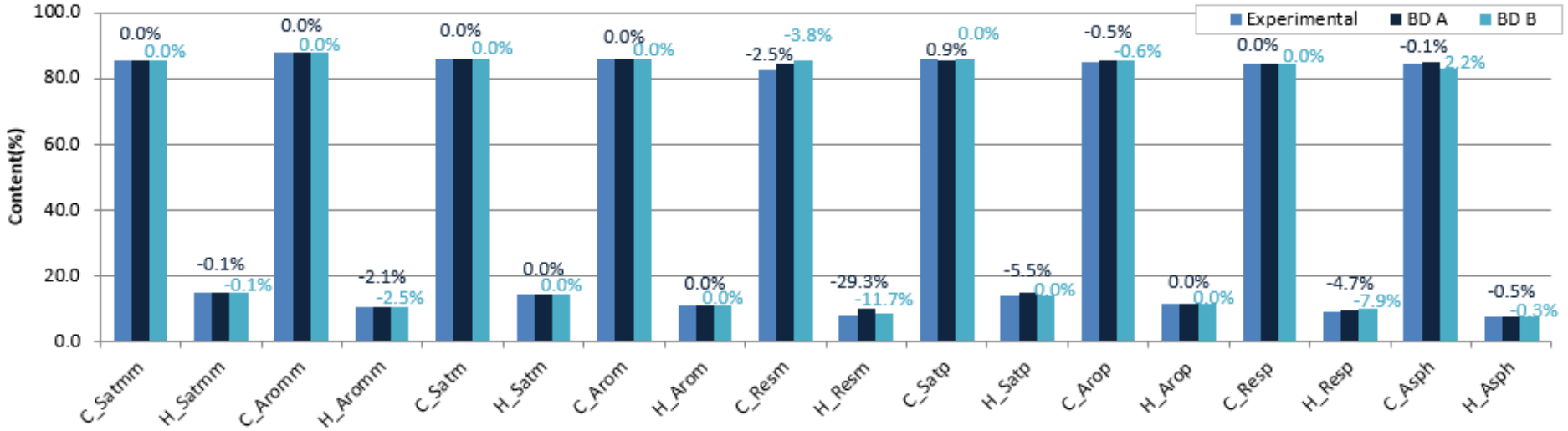


Figure 5.50: Thermidor lumps 1st REM for C and H (step 1+2).

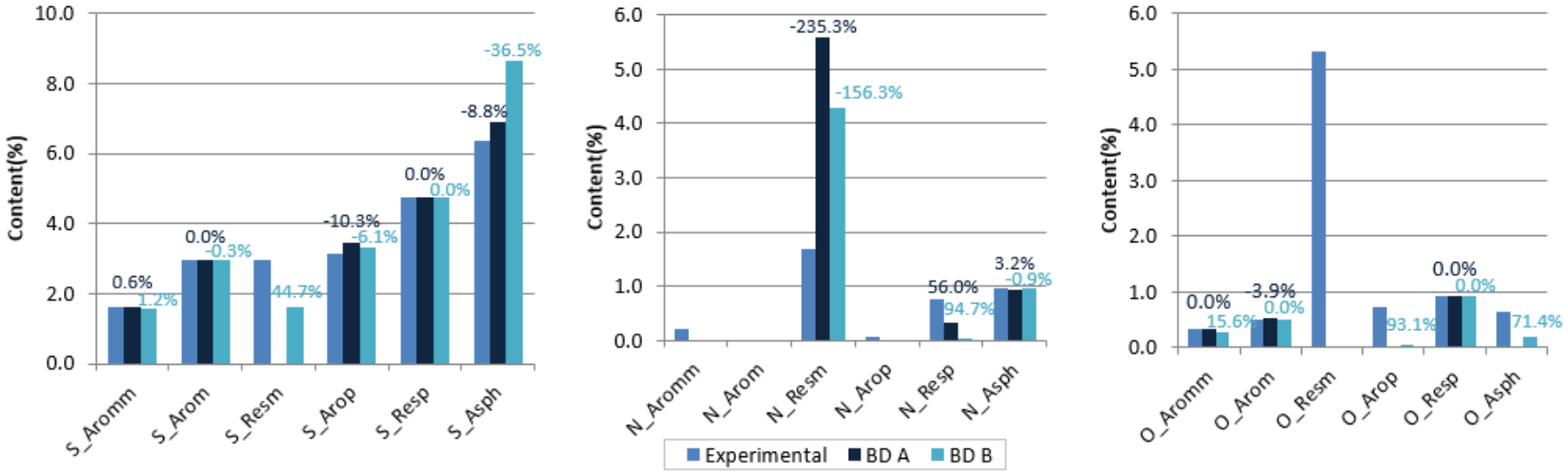


Figure 5.51: Thermidor lumps 1st REM for heteroatoms (step 1+2).

5.3.1.3 2nd REM Case

The 2nd case is identical to the first case, but the analytical data fed to the algorithm is reduced. In this case, only the global analytical data was considered in order to predict the TLA values. Given our overall goal, from an operating point of view, this case is the most important one. It will therefore be described in high detail. For the EA, shown in Figure 5.52 and Figure 5.53, the algorithm is able to achieve a perfect result for both BDs. Indeed, since the TLA was not given as input analytical data the algorithm is not hindered by inconsistencies

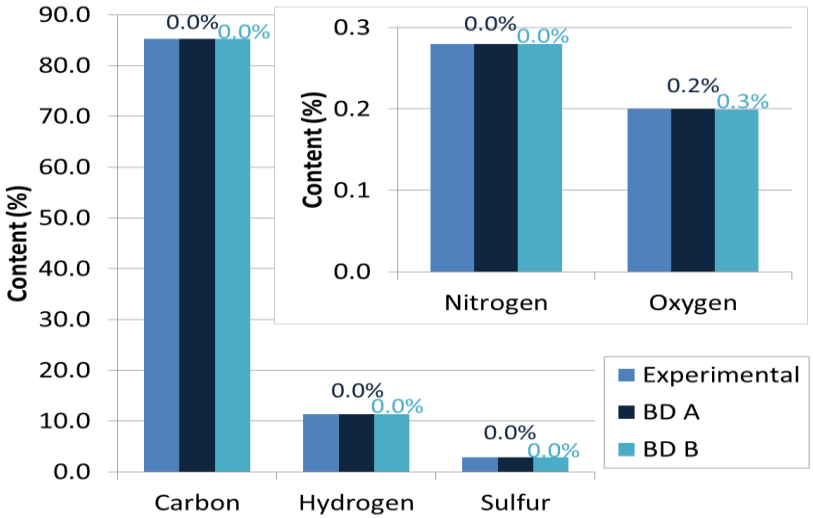


Figure 5.52: Elemental analysis results 2nd REM (step 1).

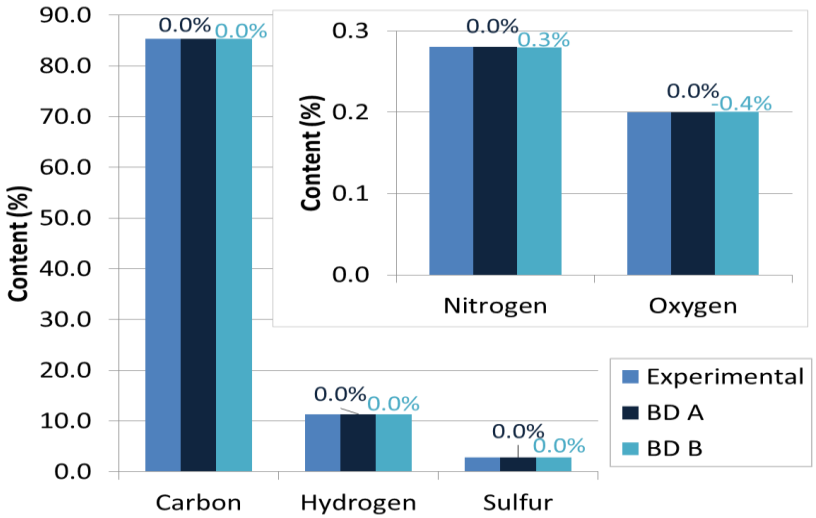


Figure 5.53: Elemental analysis results 2nd REM (step 1+2).

For the SARA analysis, results are perfect for both BDs. This is again due to the elimination of the TLA as input analytical data, thereby hiding the inconsistences of the mass balance. The same happens in the distillation curve that is perfectly adjusted to the experimental values except for the last point. This is due to the limit imposed by the library of molecules which has with a final BP that is lower than in this application point.

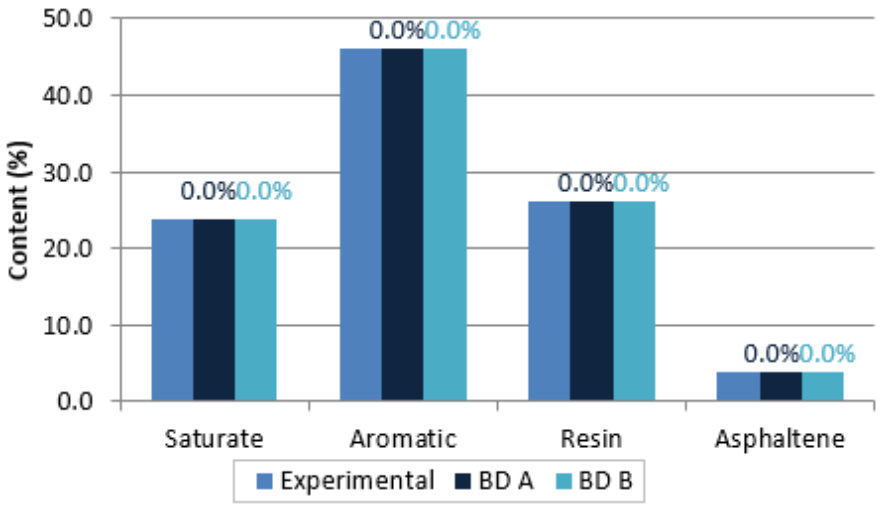


Figure 5.54: SARA analysis results 2nd REM (step 1).

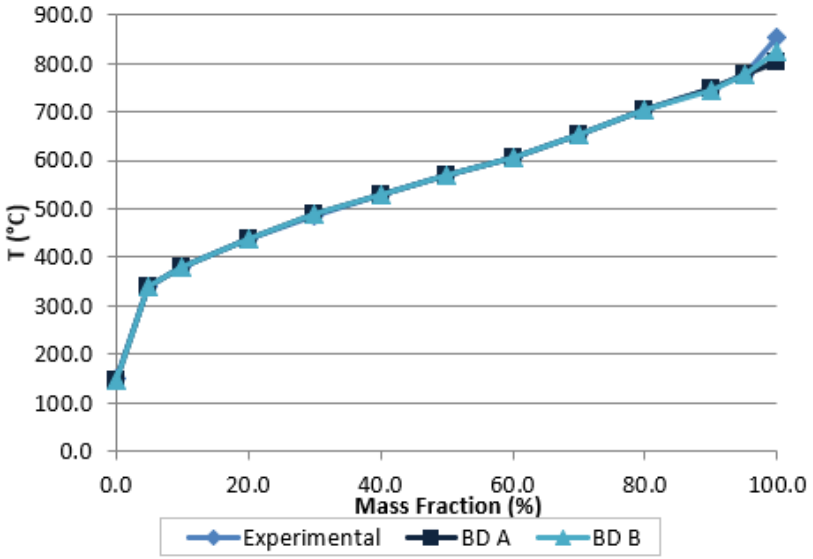


Figure 5.55: Distillation Curve 2nd REM (step 1).

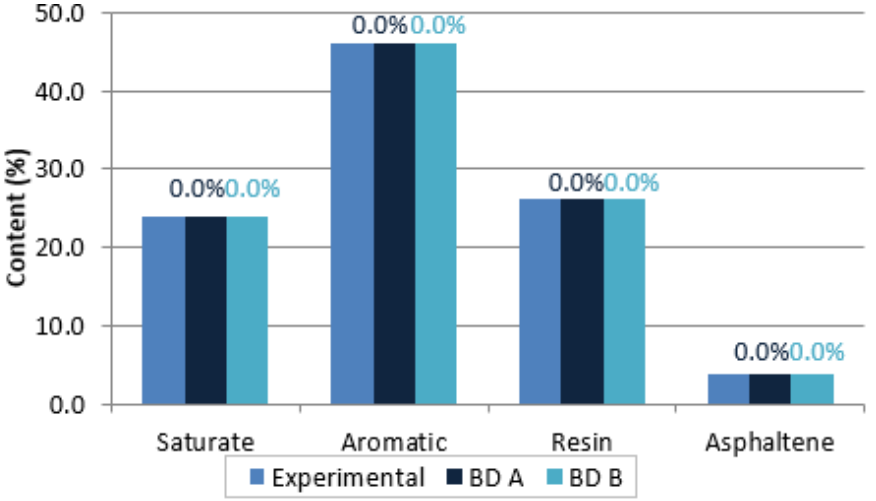


Figure 5.56: SARA analysis results 2nd REM (step 1+2).

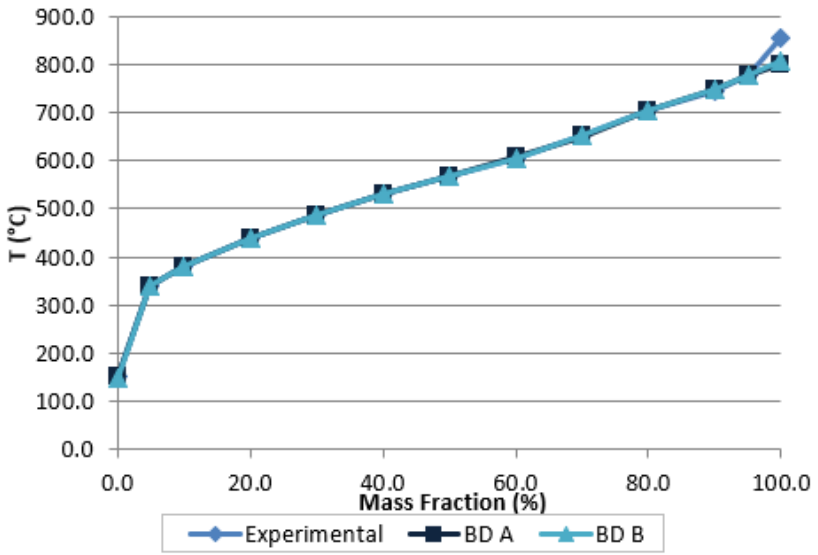


Figure 5.57: Distillation Curve 2nd REM (step 1+2).

Regarding the prediction of the TLA *(*Figure *5.58* to Figure *5.61*), for carbon, a suitable fit was obtained for each case (<5% relative error). However, for hydrogen, considering the observation made during step 1, an incorrect distribution is gain observed in Figure *5.58.* In general, the lower MW lumps have less hydrogen that the experimental value and, on the opposite, for the higher MW lumps the hydrogen content is over-predicted. When analyzing step 1+2, Figure *5.60*, the distribution is more suitable than in step 1 due to the previous adjustment of the library of molecules through the REM approach. However, not all the fits are correct as shown for the resins.

Concerning the heteroatoms, using only step 1, the library of molecules is not previously adjusted, and hence the sulfur profile is still wrong. When using step 1+2, the library was fine-tuned during the REM procedure impose the correct profile of sulfur. Hence, even when the TLA is not used as input analytical data, the sulfur profile is maintained, and the results are greatly improved.

For nitrogen and oxygen, as seen before, the BDs cannot correctly reproduce their distributions. Fortunately, both these atoms are not the major concerns for the simulator. In particular in this case and compared to the cases before, the best option is to combine the approach using the step 1+2 to the BD A.

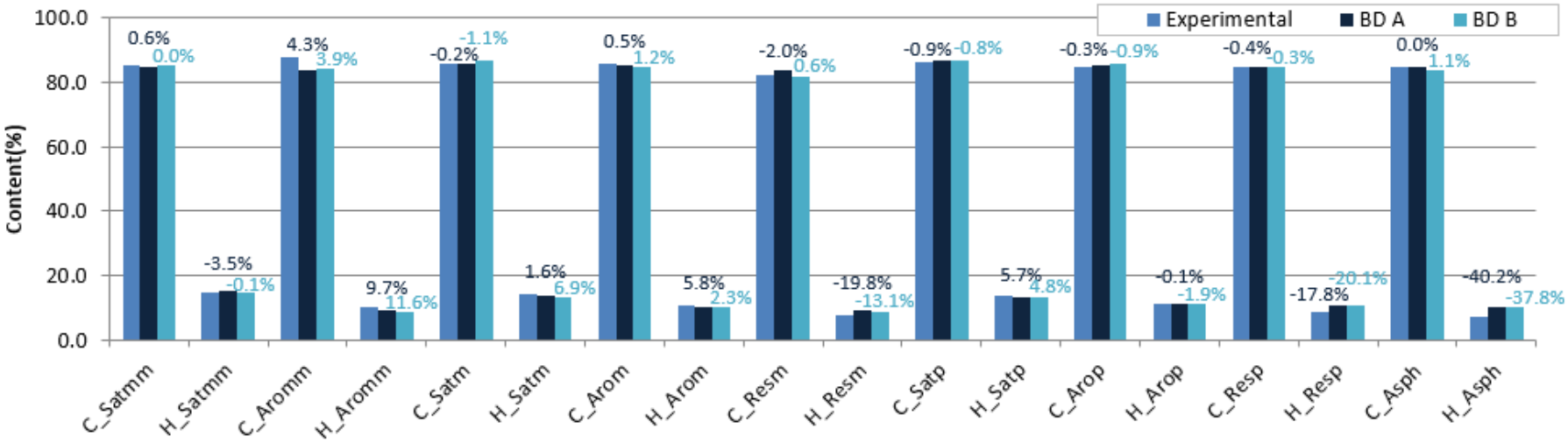


Figure 5.58: Thermidor lumps 2nd REM for C and H (step 1).

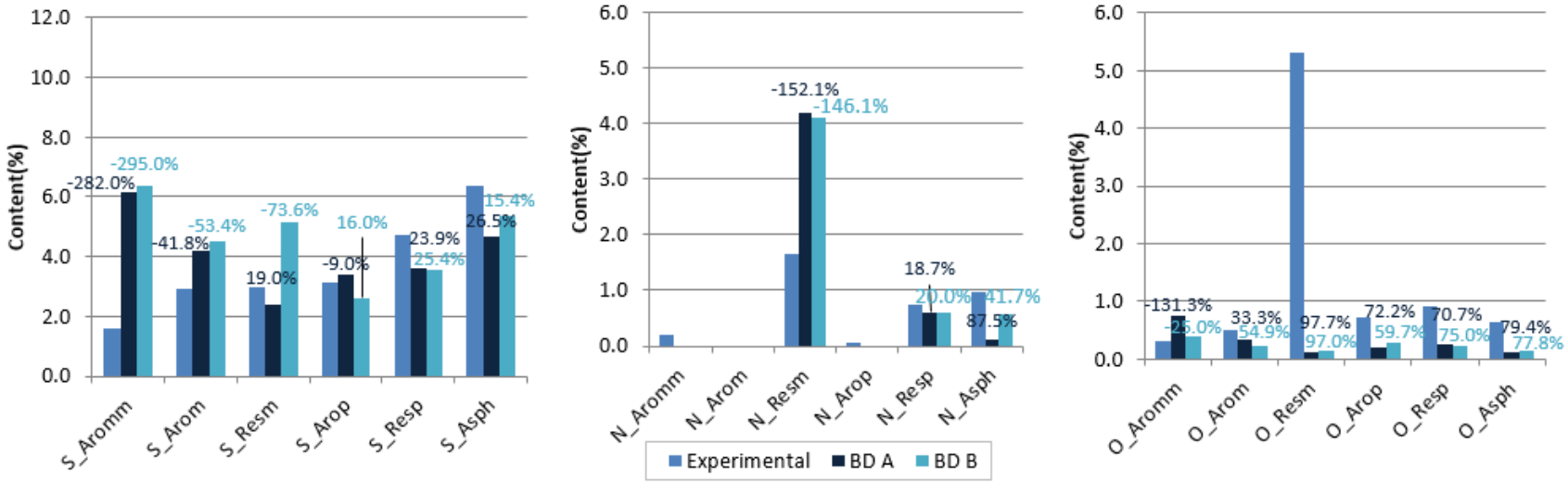


Figure 5.59: Thermidor lumps 2nd REM for heteroatoms (step 1).

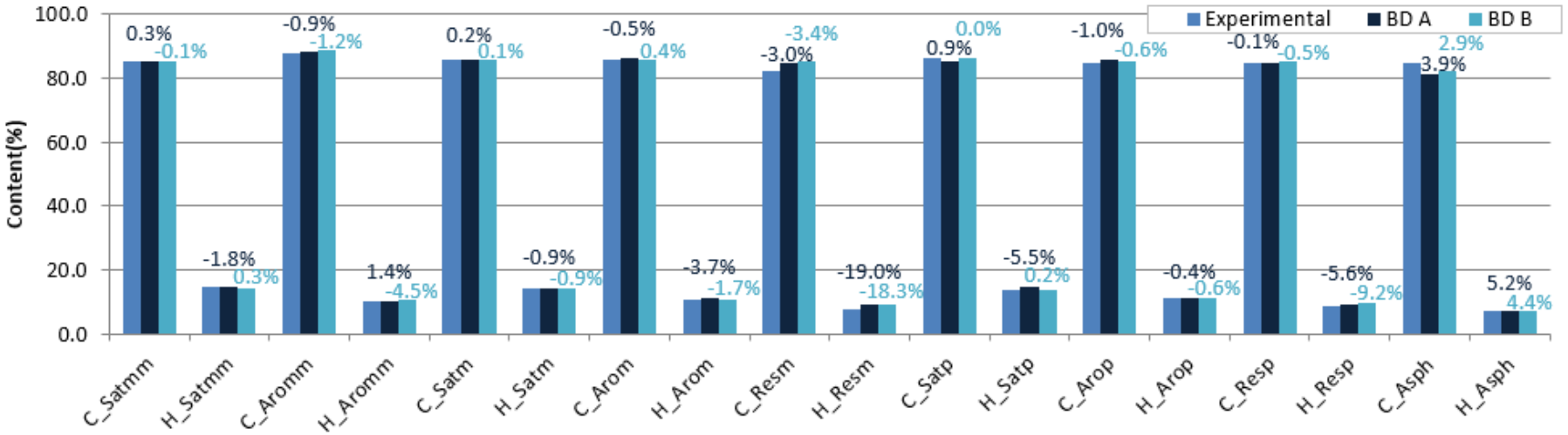


Figure 5.60: Thermidor lumps 2nd REM for C and H (step 1+2).

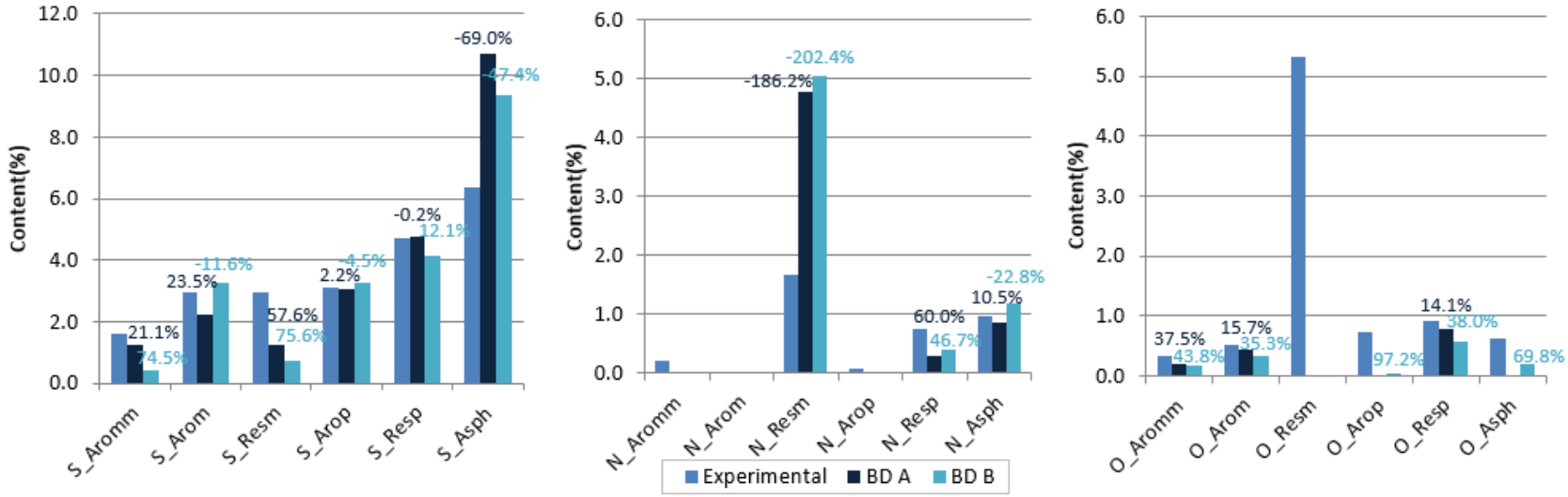


Figure 5.61: Thermidor lumps 2nd REM for heteroatoms (step 1+2).

5.3.1.4 3th REM Case

On the 3rd REM case, the analyses provided to the algorithm were further reduced: only the distillation curve and the SARA analysis were supplied.

Considering the EA (Figure *5.62* and Figure *5.63*), it is possible to notice the deterioration of the results. This is explained by the absence of information on the sample. Therefore, the algorithm will report an EA close to that of the library of molecules.

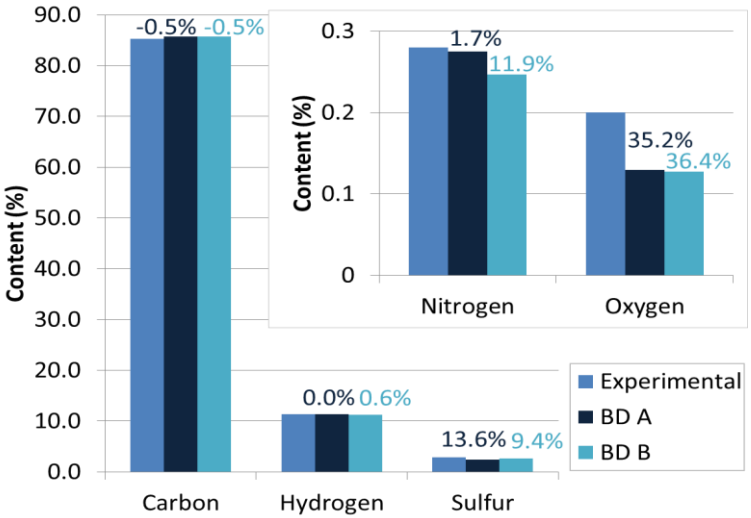


Figure 5.62: Elemental analysis results 3rd REM (step 1).

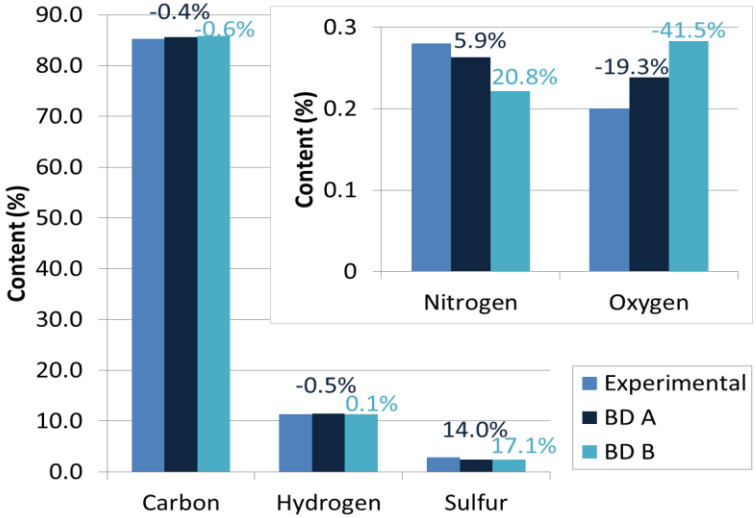


Figure 5.63: Elemental analysis results 3rd REM (step 1+2).

For the SARA analysis and for the distillation curve, the results show a good accuracy (Figure *5.64* to Figure 5.67), since the algorithm fits both analyses without inconsistencies.

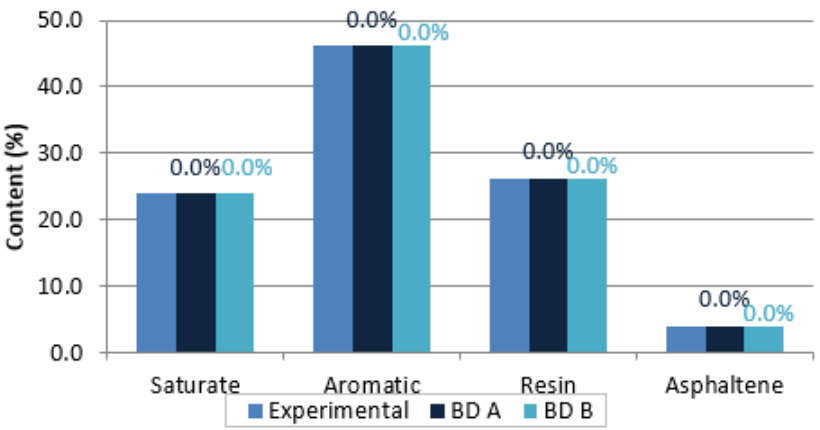


Figure 5.64: SARA analysis results 3rd REM (step 1).

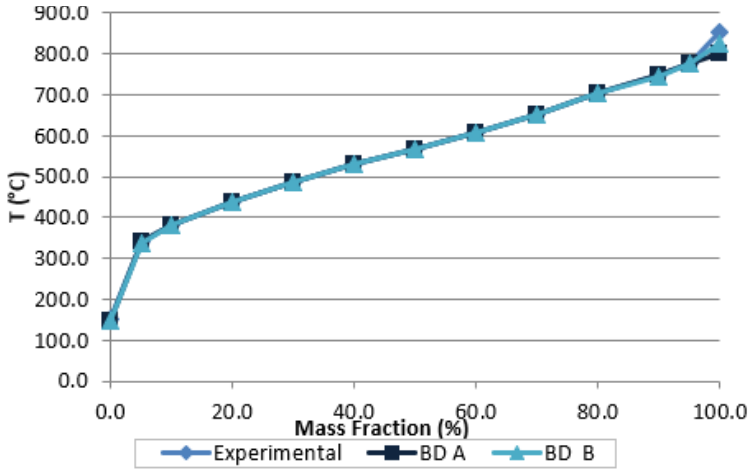


Figure 5.65: Distillation Curve 3rd REM (step 1).

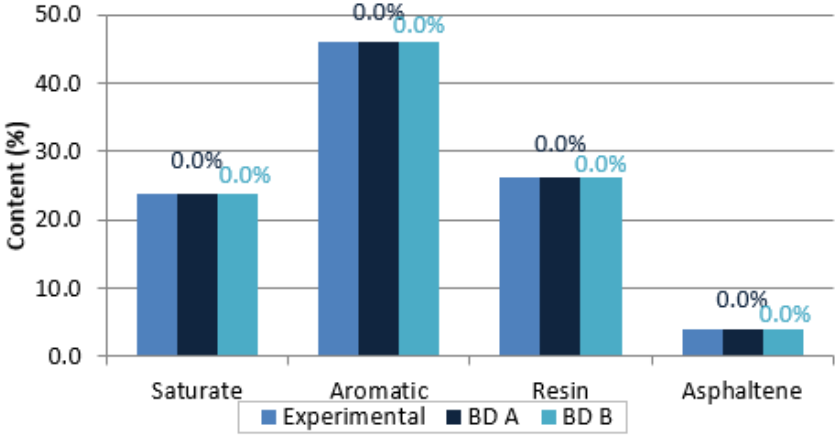


Figure 5.66: SARA analysis results 3rd REM (step 1+2).

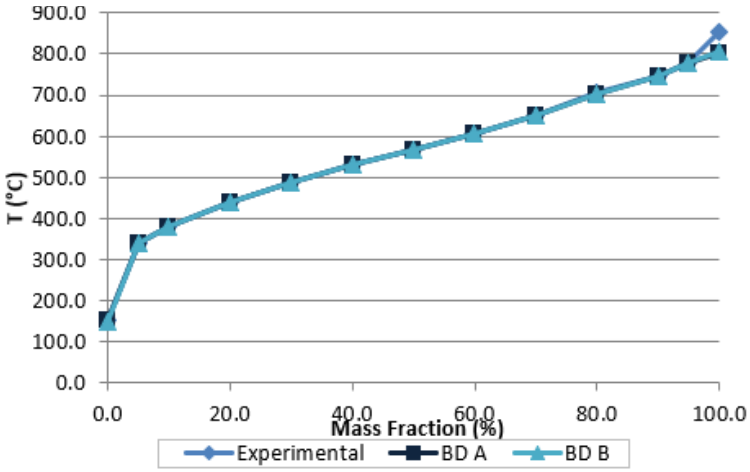


Figure 5.67: Distillation Curve 3rd REM (step 1+2).

Due to the freedom given to the algorithm by the lower number of analytical data, the same phenomenon happens in the TLA as for the EA. Moreover, since the degree of detail is higher than in the EA, the results are even poorer. However, the library of molecules used has characteristics that have been fitted on a reference point. Hence, its characteristics are approximated to the experimental values, and the results still fall in a reasonable range, even though a new mixture has been created using only two analyses.

As before, similar observations can be made, the carbon content is the easiest to predict, because of the higher content therefore a good adjustment is achieved. For the hydrogen content, the lower MW lumps are underpredicted and the higher MW lumps overpredicted in hydrogen. Again, the asphaltenes and resins are not aromatic enough. For the heteroatoms, by applying the second approach, the sulfur content retains the right profile whicg does not happen using only the stochastic reconstruction procedure. For nitrogen and oxygen, the same observation as before can be made.

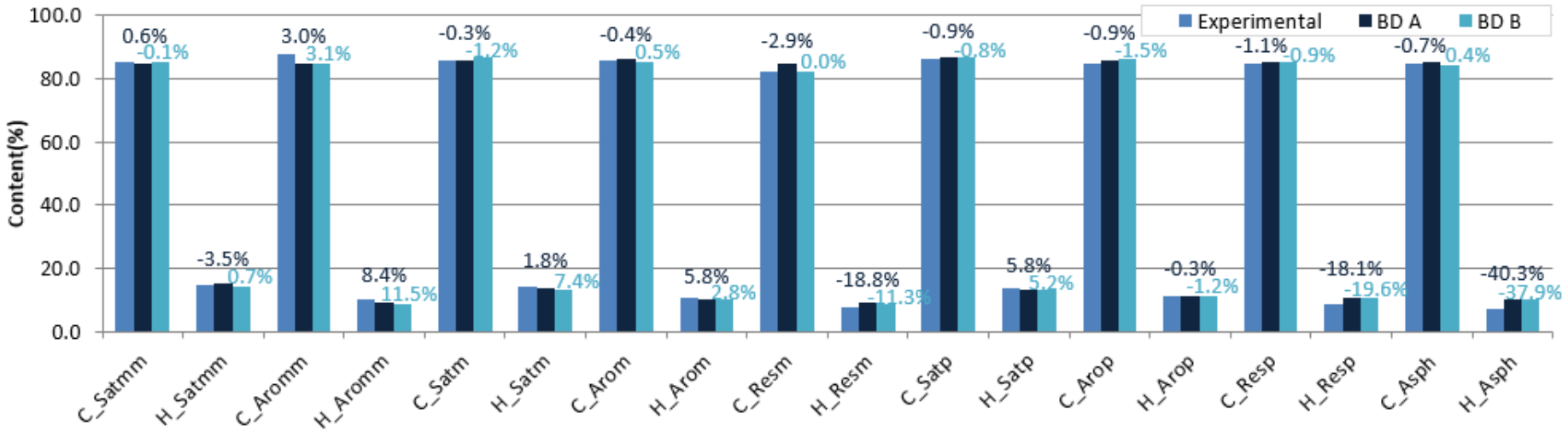


Figure 5.68: Thermidor lumps 3rd REM for C and H (step 1).

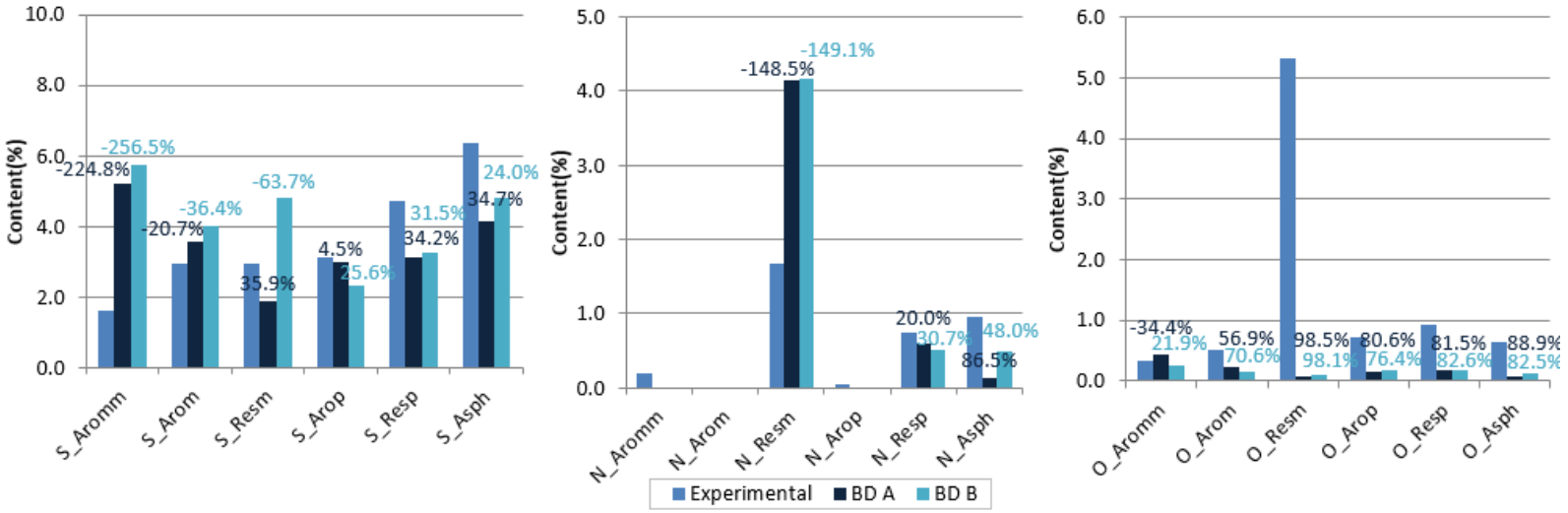


Figure 5.69: Thermidor lumps 3rd REM for heteroatoms (step 1).

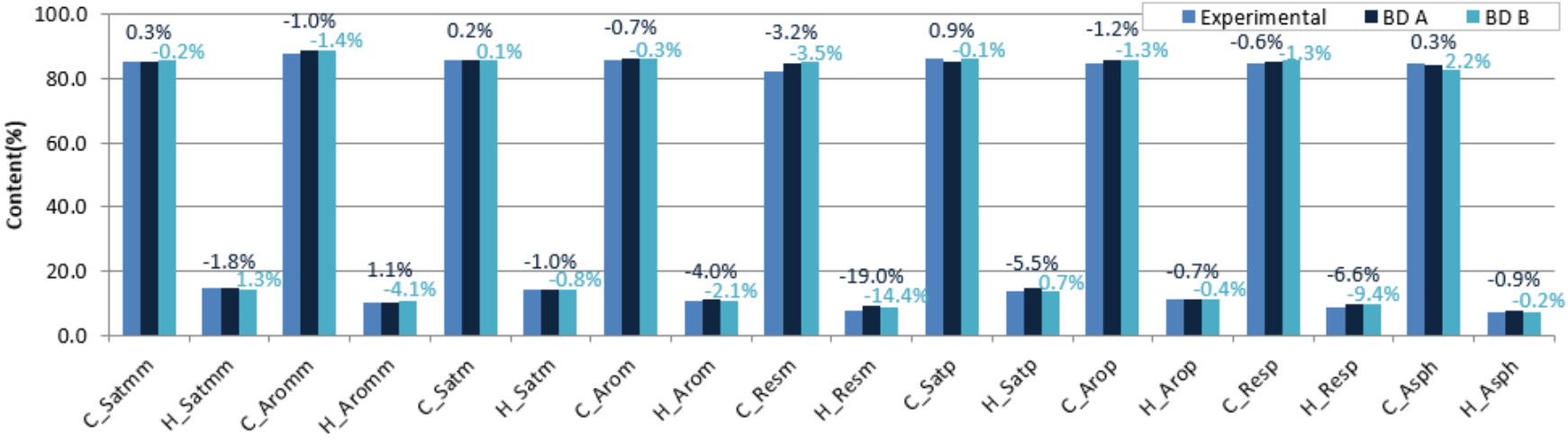


Figure 5.70: Thermidor lumps 3rd REM for C and H (step 1+2).

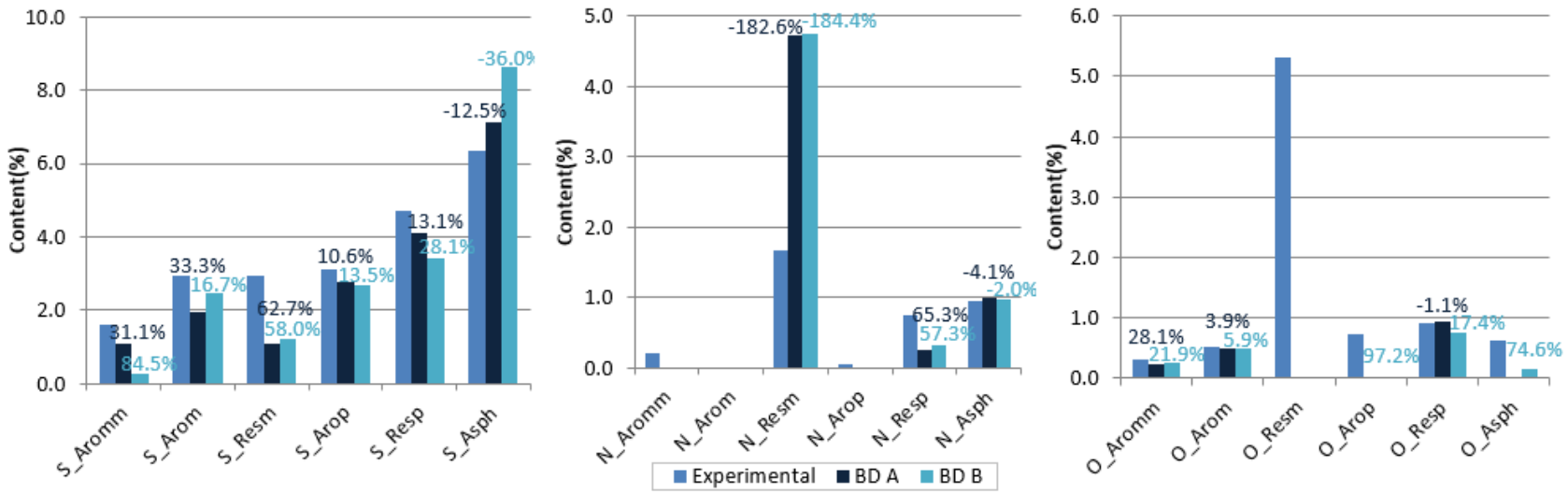


Figure 5.71: Thermidor lumps 3rd REM for heteroatoms (step 1+2).

For the 4th REM case, the SARA analysis was supplied to the algorithm. Due to the similarity of the results and conclusions to the 3rd REM case, the 4th REM cased will not be presented.

5.3.2 Partial Conclusions

To summarize, the global errors of the analyses are shown in Table 5.13 to Table 5.16. Considering the cases without inconsistencies, 0 REM case, the results for TLA have the lower error, since this is the perfect case it is possible to affirm that for the BD A the minimum error that could be achieved is 8% and for the BD B around 12%. When adding the inconsistencies to the MB the algorithm (1st REM) is not able to fit simultaneously the EA and TLA so the errors for both analyses increases until find the calculated values that allows the minimum value for the objective function.

In the 2nd REM case, the inconsistent MB is hidden from the algorithm since TLA is not given as input analytical data and its prediction is done. For this case, the best results are obtained with BD B (lower absolute error value) using the library of molecules previously fine tune, step 1+2. Since the application point is from the same test, the type

of molecules found are similar and the fine tuning done to the library of molecules allows to narrow the calculation of the TLA, presenting best results that the less restricted library (step 1). Furthermore, the global analytical data given as input adapts the values perfectly to the experimental ones.

As mentioned before, the 3rd REM and 4th REM cases are very similar due to the low imposed restrictions, which gives a huge freedom to the algorithm that cannot be compensated by the fine-tuned library of molecules. Moreover, since for the 4th REM case the SARA analysis is also not used as input the algorithm presents a considerable error for this analysis.

In addition, the CCR and density were compared to evaluate the behavior of the algorithm. The CCR is calculated based in a poor correlation that only takes into account the H content of the mixture so the less deviated is the H content the less accurate is the CCR value. In general, the density is well predicted, and better results are given by using the library of molecules optimized due to the alteration of the molar fractions that gives the higher impact to the most suitable ones.

Table 5.13: Comparison error results step 1 and BD A.

Analyses	**EA**		**SARA**		**Distillation**		**TLA**		**CCR**		**Density**	
Cases	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (°C) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (g/cm^3) [2]	AR Error
0 REM	8.0E-05	0.4%	2.0E-04	0.1%	3.8E+03	1.1%	23.9	8.0%	4.0	17.3%	6.6E-04	2.6%
1st REM	9.4E-03	8.1%	6.2E-03	0.4%	3.1E+03	1.2%	56.6	11.6%	13.0	31.1%	1.5E-04	1.3%
2nd REM	3.9E-07	0.0%	2.0E-04	0.0%	2.9E+03	0.7%	96.3	28.0%	49.1	60.4%	2.5E-04	1.6%
3rd REM	3.7E-01	10.2%	0.0E+00	0.0%	2.9E+03	0.7%	89.1	25.9%	52.0	62.1%	2.6E-04	1.7%
4th REM	3.9E-01	10.9%	1.1E+00	2.3%	2.9E+03	0.7%	89.0	25.9%	53.4	63.0%	2.3E-04	1.5%

Table 5.14: Comparison error results step 1 and BD B.

Analyses	**EA**		**SARA**		**Distillation**		**TLA**		**CCR**		**Density**	
Cases	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (°C) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (g/cm^3) [2]	AR Error
0 REM	4.1E-04	0.8%	2.1E-03	0.3%	9.02E+02	0.6%	56.32	12.3%	11.6	29.4%	5.2E-04	2.3%
1st REM	1.3E-02	8.7%	1.6E-02	0.7%	7.22E+03	2.9%	68.0	14.3%	27.1	44.9%	7.1E-05	0.9%
2nd REM	2.0E-06	0.1%	0	0.0%	8.43E+02	0.5%	97.3	26.3%	57.9	65.5%	1.6E-04	1.3%
3rd REM	2.8E-01	11.7%	0	0.0%	8.44E+02	0.5%	88.0	26.7%	54.9	63.8%	2.7E-04	-1.7%
4th REM	2.7E-01	11.8%	3.1	7.7%	8.43E+02	0.5%	88.1	26.7%	53.4	63.0%	2.6E-04	1.6%

Table 5.15: Comparison error results step 1+2 and BD A.

Analyses	**EA**		**SARA**		**Distillation**		**TLA**		**CCR**		**Density**	
Cases	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (°C) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (g/cm^3) [2]	AR Error
0 REM	1.0E-01	11.5%	9.8E-03	0.5%	3.8E+04	7.4%	70.7	9.3%	0.2	-3.4%	1.4E-04	1.2%
1st REM	3.3E-03	4.8%	3.9E-03	0.3%	3.9E+03	1.5%	65.1	8.1%	17.7	36.3%	1.0E-04	1.0%
2nd REM	5.4E-07	0.0%	0.0	0.0%	2.9E+03	0.7%	84.1	12.3%	14.5	32.8%	9.2E-05	1.0%
3rd REM	2.7E-01	8.0%	0.0	0.0%	2.9E+03	0.7%	58.3	11.1%	18.6	37.1%	4.6E-05	0.7%
4th REM	3.1E-01	9.3%	4.8E+00	9.5%	2.9E+03	0.7%	62.5	11.3%	21.3	39.7%	1.4E-05	0.4%

Table 5.16: Comparison error results step 1+2 and BD B.

Analyses	**EA**		**SARA**		**Distillation**		**TLA**		**CCR**		**Density**	
Cases	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (°C) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (g/cm^3) [2]	AR Error
0 REM	1.4E-03	1.7%	1.2E-03	0.2%	2.6E+03	0.9%	71.31	12.3%	18.6	37.1%	3.0E-04	1.8%
1st REM	1.2E-02	8.5%	1.6E-02	0.7%	2.6E+04	4.9%	58.6	12.2%	30.4	47.5%	5.3E-05	0.7%
2nd REM	4.3E-06	0.1%	0.0	0.0%	2.3E+03	0.7%	75.3	18.5%	34.9	50.9%	4.6E-05	0.7%
3rd REM	4.7E-01	16.0%	0.0	0.0%	2.3E+03	0.6%	68.8	16.5%	30.4	47.5%	4.9E-05	-0.7%
4th REM	5.1E-01	17.4%	6.6	12.2%	2.3E+03	0.7%	79.9	17.6%	30.4	47.5%	5.6E-05	0.8%

In conclusion, the molecule library obtained by the second approach is considered more suitable for reduced analytical input data, taking into account the pre-adjustment of by attributing through REM procedure a higher molar fraction to the most suitable molecules. Therefore, the BD A presents better results considering the TLA the BD B was considered more appropriated to construct the library of molecules due to the higher stochastic character provided by a lower number of distributions and parameters. Although, for both diagrams not suitable distributions were found for the H and S in function of the molecular weight, to improve the results two solutions could be applied.

To ameliorate the effluent reconstruction, two ways can be taken. The first one was increasing the replication factor increasing the representativeness of the mixture. This method will maintain the same library of molecules by eliminating fewer molecules and keeping the optimal mixture achieved by the REM procedure. The second one was the development of a new BD to construct more suitable molecules. This BD should provide a higher freedom to create distributions in function of the MW. Since changing the replication factor of the MD procedure to delete less molecules will not correct the problems found for the H and S distributions, it was decided to construct a new BD to solve the root of the problem: the construction of the molecules.

5.4 Development Building Diagram C

The objective of this building diagram is to reduce the error associated to thermidor lumps analysis, only the best BD will be discussed from 2 proposed schemes edified. The details about the chosen distributions, as well as the building diagram can be found in Annex C (Figure C3 and Table C3).

The edification of this scheme is based on the BD B, in which to calculate the cyclohexane rings an equation was used considering the difference between the distribution 4 and 5 (Annex C – Figure C2: BD B) which may raise the question if the correlation is suitable. Then, the distribution 5 considering the quantity of aromatic rings was eliminated and a new parameter specific to the cyclohexane rings was added to the type of rings distribution. Consequently, this new BD called BD C has 9 distributions and the same number of parameters as BD B: 17 parameters. It is expected that this alteration gives more flexibility to the H distribution and afterward the addition of the heteroatoms to the chain can be improved. Since it was already settled that the second approach is more suitable only those results applying the new BD would be discussed.

Table 5.17: Results library of molecules BD C.

			Calculated Values	
Experimental Values			SR	SR+REM
C	wt%	85.67	85.93	85.63
H	wt%	11.46	11.38	11.38
S	wt%	2.48	2.33	2.44
N	wt%	0.26	0.22	0.24
O	wt%	0.13	0.14	0.31
Saturates	wt%	25.03	26.67	25.08
Aromatics	wt%	46.85	44.95	46.85
Resins	wt%	24.87	24.26	24.84
Asphalthenes	wt%	3.25	4.12	3.24
Initial BP	°C	150.5	148.1	148.1
5% BP	°C	337.4	220.5	170.8
10% BP	°C	378.9	272.4	348.1
20% BP	°C	438.8	344.0	421.5
30% BP	°C	484.9	396.1	479.5
40% BP	°C	524.3	447.0	521.9
50% BP	°C	561.4	499.1	554.5
60% BP	°C	598.1	554.9	594.6
70% BP	°C	640.6	615.9	626.6
80% BP	°C	692.5	653.0	673.8
90% BP	°C	740.4	677.7	699.4
95% BP	°C	774.7	693.6	750.1
Final BP	°C	866.6	750.2	750.1

Considering Table 5.17, the construction of molecules accomplished by the new scheme has a good fitting for the global analytical data excepting the distillation curve in which the BP are successively lower than the experimental value, either the higher molecules construct by the BD are not heavy enough to achieve the higher BPs of the mixture, creating only molecules with BP until 750°C.

Applying the REM procedure to the library constructed the properties are fine-tuned and improvements considerer the EA and the SARA analysis enable an almost perfect fit. As before, due to the inconsistencies in the O MB the calculated value is not suitable. Although some enhancements were done in the distillation curve, and the middle points are very close to the experimental values, the weight given to the saturates results in a very low BP for 5%wt. For this case, only the H and S distribution will be seen to evaluate the performance of the BD C since the other elementals have approximately the same characteristics that before (Figure 5.72 and Figure 5.73).

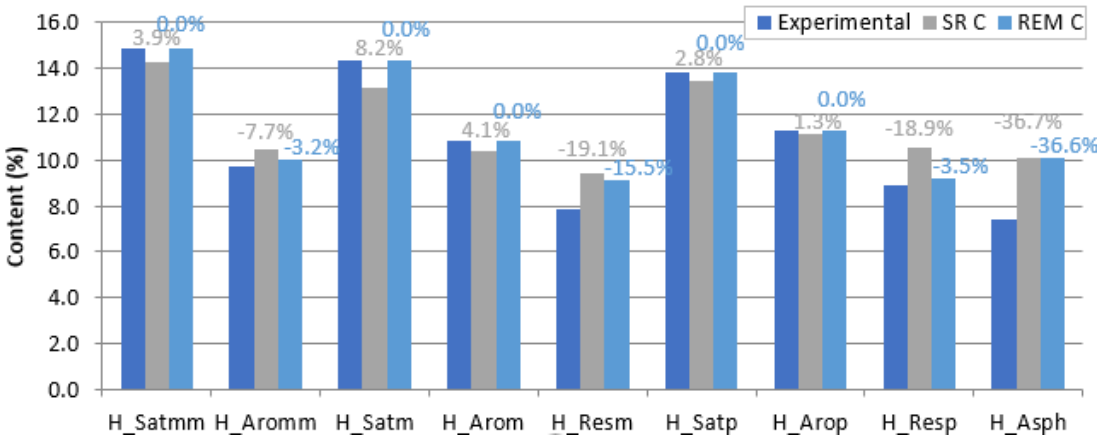


Figure 5.72: H distribution BD C.

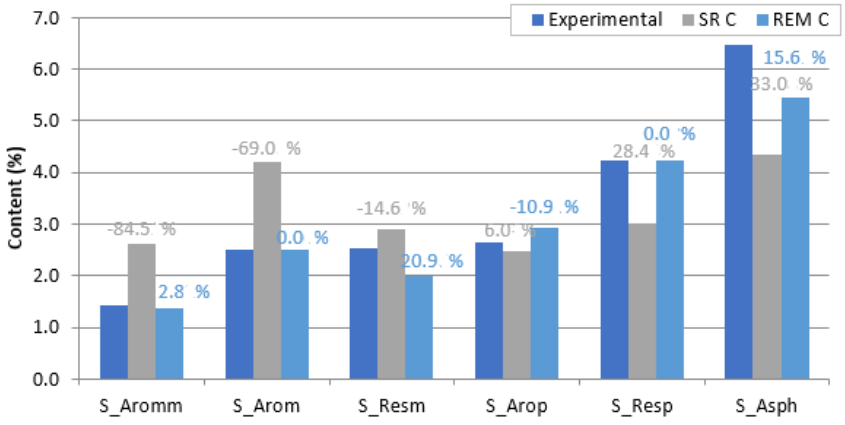


Figure 5.73: S distribution BD C.

The results of both distributions are equivalent to the results found for the BD B given the proximity of both schemes. However, the H in the asphaltenes is not well optimized since this scheme increased the hydrogen content on the EA relatively to the BD B, and the S is better adjust in general and also observing the S in asphaltenes is possible to see instead of an higher value that the experimental one as happened in the BD B the calculated value is lower and a lower error is associated.

After the REM procedure, the MD procedure was applied: On Annex B the impact of the replication factor (K = 10 000) can be seen. Noticing that the BD C is derived from the BD B the results for the type of molar fraction distribution are similar (Figure 5.74), this means a more homogenous distribution throughout the range of molar fractions, so unsuitable molecules have less impact and it is attribute a lower molar fraction and molecules appropriated to the experimental data will receive a higher molar fraction. Accordingly, to these more molecules will be eliminated, and the question of representativeness can be raised again. Be that as it may, the difference between REM and MD results will be study next.

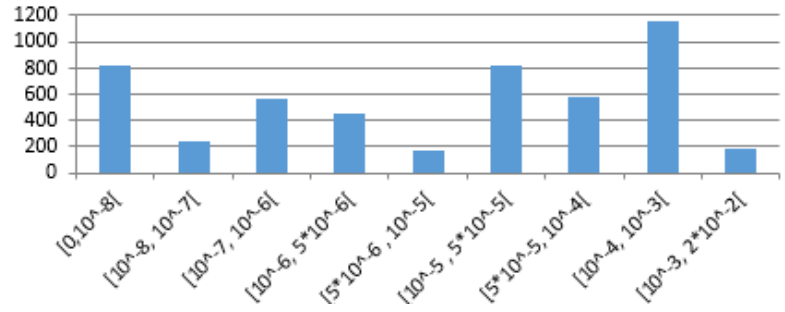


Figure 5.74:Molar fraction distribution REM C.

Analyzing the results between the following tables, since the values of the errors are almost equal is concluded that the MD procedure enables once again to reconstruct an equimolar mixture without damaging greatly the representativeness of the mixture.

Table 5.18: Comparison REM and MD AQ Error BD C.

Analyses	**EA (%wt) [2]**	**SARA (%wt) [2]**	**Distillation (°C) [2]**	**TLA (%wt) [2]**	**CCR (%wt) [2]**	**Density (g/cm^3) [2]**
REM C	4.1E-02	3.5E-03	4.5E+04	45.3	20.9	1.5E-04
REM + MD C	4.0E-02	8.5E-02	4.4E+04	45.2	3.5	1.7E-04

Table 5.19: Comparison REM and MD AR Error BD B.

Analyses	**EA**	**SARA**	**Distillation**	**TLA**	**CCR**	**Density**
REM C	29.5%	0.2%	7.2%	7.7%	42.0%	1.3%
REM + MD C	28.7%	0.5%	7.2%	10.3%	17.2%	1.4%

When employing this mixture to the VR application point the results improved significantly (Table 5.20) although in the library used even with the improvement in the TLA the distillation curve had worst results. Due to the random character of the procedure a deeper research regarding this situation needs to be study in order to find a way of improving the TLA without prejudicing the remaining analytical data while construction the library.

Table 5.20: Comparison criteria errors BD C.

	EA		**SARA**		**Distillation**		**TLA**	
Case	AQ Error (%wt) [2]	AR Error	AQ Error (%wt) [2]	AR Error	AQ Error (°C) [2]	AR Error	AQ Error (%wt) [2]	AR Error
2nd REM	1.9E-06	0.1%	0.0	0.0%	1.3E+04	1.8%	62.4	12.2%

6 Conclusions and Perspectives

In this work, a new methodology to reconstruct residue hydrotreating effluents was developed. This new methodology consists in applying three different approaches:

- Stochastic reconstruction (SR),
- Reconstruction by entropy maximization (REM),
- Molecular discretization (MD).

The best BD is considered to be the one in which the stochastic character of the procedure is higher, this means the ones with less distributions and parameters that allows a good fit of the experimental data, in this case both BD B and BD C are in this situation since both reconstruct well the Ei although the BD B presents best results when edifying an equimolar set of molecules.

Additionally, the thermidor lumps (TL) reconstruction needs to be improved to achieve an average relative error lower than 5%, during this work a minimum of 12% was reached for the application case and 8% for the library of molecules, these values correspond to the 2nd REM case, the most relevant case when reducing the experimental work since the TLA is the most costly and time consuming.

Considering the MD procedure, the replication factor imposed deletes around 40% for the BD A and around 60% for the BD B and BD C, since this factor is closely related with the loss of representativeness of the mixture the lower the elimination of molecules the closer the mixture is of the optimal one given by the REM procedure.

To improve the results, and as perspectives of a future work, is proposed to research deeper the edification of the BD to create more suitable molecules with a molar fraction distribution considerably above of the molar fractions eliminated by the replication factor imposed. With these modifications is expected to decrease the TL error to better fit the experimental data and increasing the representativeness of the mixture by deleting less molecules.

Finally, it is suggested to study the propagation of errors of the TLA to fully understand the uncertainties of the experimental values and the consequences in the calculated ones. Also, to validate the methodology other points need to be tested.

7 Bibliographie

[1] Eurostat. Energy balance sheets 2014 data, *Statistical Books*, 2016.

[2] BP. BP Statistical Review of World Energy 2017, 2017.

[3] Internacional Energy Agency. Oil 2017: Analysis and Forecasts to 2022, *Market Report Series*, 2017.

[4] GAUTHIER T., Danial-Fortain P., MERDRIGNAC I., Guibard I., QUOINEAUD A. Studies on the evolution of asphaltene structure during hydroconversion of petroleum residues, *Catalysis Today*, 2008, **130,** 2-4, 429-438. DOI: 10.1016/j.cattod.2007.10.005.

[5] Boduszynski M.M. Composition of heavy petroleums : 2. Molecular characterization, *American Chemical Society Journal*, 1988, **2,** 5.

[6] Oliveira L.P. de, Verstraete J.J., Marques J., Kolb M., eds. *Two-step Molecular Reconstruction Algorithm of Heavy Petroleum Fractions.* 13th Internacional Conference on Petroleum Phase Behavior and Fouling, 10-14 June. St. Petersburg Beach, Florida, USA, 2012, 1 p.

[7] Energy Internacional Agency. Crude Oils have different quality characteristics. https://www.eia.gov/todayinenergy/detail.php?id=7110, consulté 4 avril 2018.

[8] Speight J. G. *The desulfurization of heavy oils and residua*. Marcel Dekker, New York, 2005.

[9] Speight J. G. *The Chemistry and Technology of Petroleum*. Taylor&Francis CRC Press, 2014.

[10] BP. Heavy Oil vs. Light Oil.

[11] Schlumberger. The Defining Series: Heavy Oil, 2016.

[12] Pereira de Oliveira, Luís Carlos. *Développement d'une méthodologie de modélisation cinétique de procédés de raffinage traitant des charges lourdes*. Université de Lyon, 2013, 474 p.

[13] Barker C., Robbins W.K., Hsu C.S., Drew L.J. Petroleum, 2005.

[14] Oliveira L.P. de, Vazquez A.T., Verstraete J.J., Kolb M. Molecular Reconstruction of Petroleum Fractions: Application to Vacuum Residues from Different Origins, *Energy & Fuels*, 2013, **27,** 7, 3622-3641. DOI: 10.1021/ef300768u.

[15] Verstraete J.J., Schnongs P., Dulot H., Hudebine D. Molecular reconstruction of heavy petroleum residue fractions, *Chemical Engineering Science*, 2010, **65,** 1, 304-312. DOI: 10.1016/j.ces.2009.08.033.

[16] Liu L. *Molecular Characterization and Modelling for Refining Processes-thesis*. University of Manchester, 2015, 175 p.

[17] Billon A., Peries J.P., Espeillac M., Courieres T., eds. *HYVAHL F versus HYVAHL M, Swing Reactor or Moving Bed.* NPRA annual meeting, 17-19. San Antonio, Texas, March 1991, 23 p.

[18] Ancheyta J., Sánchez S., Rodríguez M.A. Kinetic modeling of hydrocracking of heavy oil fractions: A review, *Catalysis Today*, 2005, **109,** 1-4, 76-92. DOI: 10.1016/j.cattod.2005.08.015.

[19] Plantier S., ed. *Hyvahl TM & RFCC Synergy.* Axens' Iran Seminar, 19 July. Tehran, 2017, 12 p.

[20] Kressmann S., Harlé V., Kasztelan S., Guibard I., Tromeur P., Morel F., eds. *Maximizing Cycle Length of Vacuum Residue Hydrodesulfurization Unit.* 218th Am Chem Soc National Meeting, 22 August. New Orleand, 1999.

[21] Ferreira C., Tayakout-Fayolle M., Guibard I., Lemos F. Hydrodesulfurization and hydrodemetallization of different origin vacuum residues : New modeling approach, *Fuel*, 2014, **129**, 267-277. DOI: 10.1016/j.fuel.2014.03.056.

[22] Kressmann S., Guillaume D., Roy M., Plain C. A New Generation of Hydroconversion and Hydrodesulfurization Catalysts, 5-6 December 2004, **14th Annual Symposium - Catalysis and Petroleum Refining & Petrochemicals**.

[23] Allen D.T. STRUCTURAL MODELS OF CATALYTIC CRACKING CHEMISTRY, in *Kinetic and thermodynamic lumping of multicomponent mixtures* : *Proceedings of an ACS Symposium on kinetic and thermodynamic lumping of multicomponent mixtures, Atlanta, GA, April 15, 1991.* Éd. G. Astarita, S. I. Sandler. Elsevier, Amsterdam, Oxford, New York etc., 1991, 163-180.

[24] Hudebine D., Verstraete J.J. Reconstruction of Petroleum Feedstocks by Entropy Maximization. Application to FCC Gasolines, *Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles*, 2011, **66,** 3, 437-460. DOI: 10.2516/ogst/2011110.

[25] Shannon C.E. A Mathematical Theory of Communication, *Bell System Technical Journal*, 1948, **27,** 4, 623-656. DOI: 10.1002/j.1538-7305.1948.tb00917.x.

[26] Hudebine D., Verstraete J.J. Molecular reconstruction of LCO gasoils from overall petroleum analyses, *Chemical Engineering Science*, 2004, **59,** 22-23, 4755-4763. DOI: 10.1016/j.ces.2004.09.019.

[27] Quann R.J., Jaffe S.B. Structure-oriented lumping: describing the chemistry of complex hydrocarbon mixtures, *Industrial & Engineering Chemistry Research*, 1992, **31**, 2483-2497.

[28] Quann R.J., Jaffe S.B. Building Useful Models of Complex Reaction Systems, *Chemical Engineering Science*, 1996, **51,** 10, 1615-1635.

[29] Jaffe S.B., Freund H., Olmstead W.N. Extension of Structure-Oriented Lumping to Vacuum Residua, *Industrial & Engineering Chemistry Research*, 2005, **44,** 26, 9840-9852. DOI: 10.1021/ie058048e.

[30] Zhan Y. *A molecular approach for characterisation and poperty predictions of petroleum mixtures with applications to refinery modelling*. University of Manchester, 1999.

[31] Peng B. *Moelcular modeling of petroleum processes*. University of Manchester, 1999.

[32] Pyl S.P., Hou Z., van Geem K.M., Reyniers M.-F., Marin G.B., Klein M.T. Modeling the Composition of Crude Oil Fractions Using Constrained Homologous Series, *Industrial & Engineering Chemistry Research*, 2011, **50,** 18, 10850-10858. DOI: 10.1021/ie200583t.

[33] Mi Saine Aye M., Zhang N. A novel methodology in transforming bulk properties of refining streams into molecular information, *Chemical Engineering Science*, 2005, **60,** 23, 6702-6717. DOI: 10.1016/j.ces.2005.05.033.

[34] GOMEZPRADO J., ZHANG N., THEODOROPOULOS C. Characterisation of heavy petroleum fractions using modified molecular-type homologous series (MTHS) representation, *Energy*, 2008, **33,** 6, 974-987. DOI: 10.1016/j.energy.2007.11.006.

[35] López García C., Hudebine D., Schweitzer J.-M., Verstraete J.J., Ferré D. In-depth modeling of gas oil hydrotreating : From feedstock reconstruction to reactor stability analysis, *Catalysis Today*, 2010, **150,** 3-4, 279-299. DOI: 10.1016/j.cattod.2009.08.002.

[36] Hudebine D., Verstraete J., Chapus T. Statistical Reconstruction of Gas Oil Cuts, *Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles*, 2011, **66,** 3, 461-477. DOI: 10.2516/ogst/2009047.

[37] Speight J. A structural investigation of the constituents of Athabasca bitumen by proton magnetic resonance spectroscopy, *Fuel*, 1970, **49,** 1, 76-90. DOI: 10.1016/0016-2361(70)90010-4.

[38] Hirsch E., Altgelt K.H. Integrated structural analysis. Method for the determination of average structural parameters of petroleum heavy ends, *Anal. Chem.*, 1970, **42,** 12, 1330-1339. DOI: 10.1021/ac60294a005.

[39] Suzuki T., Itoh M., Takegami Y., Watanabe Y. Chemical structure of tar-sand bitumens by 13C and 1H n.m.r. spectroscopic methods, *Fuel*, 1982, **61,** 5, 402-410. DOI: 10.1016/0016-2361(82)90062-X.

[40] FAULON J.-L. *Prediction, elucidation et modelisation moleculaire : Algorithmes et applications en geochimie*. École nationale superieure des mines, 1991.

[41] Ali F.A., Ghaloum N., Hauser A. Structure Representation of Asphaltene GPC Fractions Derived from Kuwaiti Residual Oils, *Energy & Fuels*, 2006, **20,** 1, 231-238. DOI: 10.1021/ef050130z.

[42] Oliveira L.P. de, Hudebine D., Guillaume D., Verstraete J.J., Joly J.F. A Review of Kinetic Modeling Methodologies for Complex Processes, *Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles*, 2016, **71,** 3, 45. DOI: 10.2516/ogst/2016011.

[43] Neurock M., Nigam A., Trauth D., T. Klein M. Molecular Representation of Complex Hydrocarbon Feedstocks through Efficient Characterization and Stochastic Algorithms, *Chemical Engineering Science*, 1994, **49,** 24A, 4153-4177.

[44] Hudebine D. *Reconstruction moléculaire de coupes pétrolières*. Lyon: École normale supérieure, 2003, 2 vol. (269 p., pagination multiple).

[45] Oliveira L.P. de, Verstraete J.J., Kolb M. Simulating vacuum residue hydroconversion by means of Monte-Carlo techniques, *Catalysis Today*, 2014, **220-222**, 208-220. DOI: 10.1016/j.cattod.2013.08.011.

[46] Bailleux G. *Extension d'un code de reconstruction moleculaire de coupes petrolieres*, 2003.

[47] Costa da Cruz, Ana Rita. *Rapport d' avancement de mi-thêse : Compositional and kinetic modelling of bio-oils from fast pyrolysis of lignocellulosic biomass*. Université de Lyon, 2017, 111 p.

[48] Alvarez A., Castañeda L.C., Ancheyta J. On the application of petroleum feedstock modeling techniques for developing molecule-based models of hydrocarbon conversion processes, *Catalysis Today*, 2014, **220-222**, 198-207. DOI: 10.1016/j.cattod.2013.04.009.

[49] Schnongs P. *Reconstruction moléculaire de coupes pétrolières lourdes*, 2005.

[50] Lopez Abelairas M., Oliveira L.P. de, Verstraete J.J. Application of Monte Carlo techniques to LCO gas oil hydrotreating : Molecular reconstruction and kinetic modelling, *Catalysis Today*, 2016, **271**, 188-198. DOI: 10.1016/j.cattod.2016.02.041.

Annex A: Construction of one molecule

To properly understand the procedure of creating a molecule, one example will be described taking into account the benzothiophene molecule. To create N molecules a similar procedure has to be repeated N times.

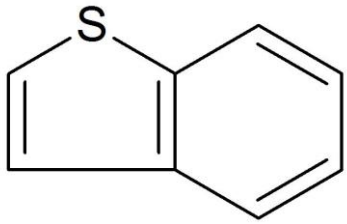


Figure A1: Benzothiophene molecule.

Figure A1 defines some of the molecular attributes presents in the heavy petroleum fractions, as exemplified for the benzothiophene molecule. The current example follows the building diagram B detailed in Annex C (Figure C2).

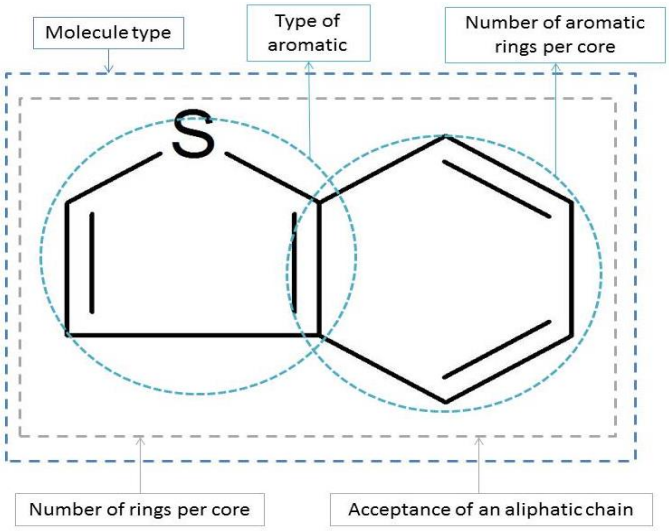


Figure A2: Structure blocks examples for benzothiophene molecule.

As mentioned, each molecular attribute is defined by a PDF and a random number is selected to draw the molecule. Figure A3 highlight this procedure, enabling the generation of the benzothiophene molecule. First, the type of molecule is sampled and a molecule with a single core will be designed. Secondly, since it is an aromatic, it must have at least one ring that is aromatic so both PDFs the number of rings in the core and the number of aromatic rings starts is distribution in one. For the benzothiophene both distributions take the value of two what means that no cyclohexanes are found in the molecule. In addition, it was declared that at least one benzene is present in aromatic cores, and then the PDF corresponding to the type of aromatic is sampled once to discover the other type of aromatic: thiophene. Finally, the acceptance of an aliphatic chain is sampled since it does not allow aliphatic chains the molecule already has is structure blocks: thiophene and benzene.

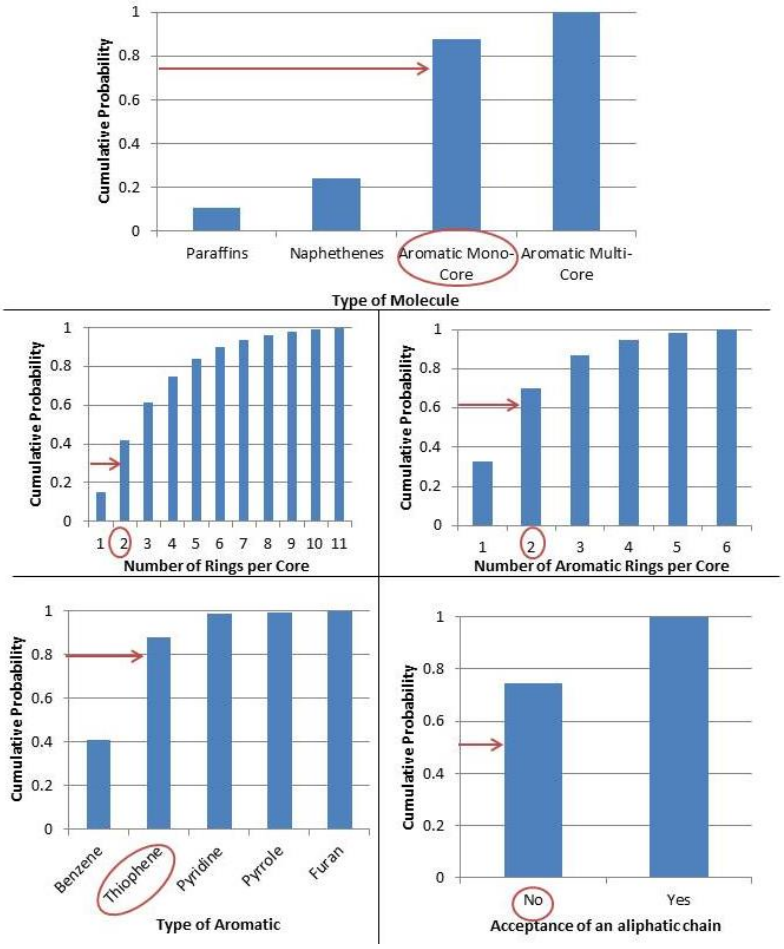


Figure A3: Sampling process of the molecular attributes of benzothiophene molecule.

The core construction of the benzothiophene is illustrated in Figure A4. The first core is placed in the center of the hexagonal map, thiophene. To avoid the creation of impossible molecules certain positions are forbidden or less-favored by given a probability of acceptation for the next structure blocks assembled. For example, to adjust the benzene only two positions with the same probability are allowed.

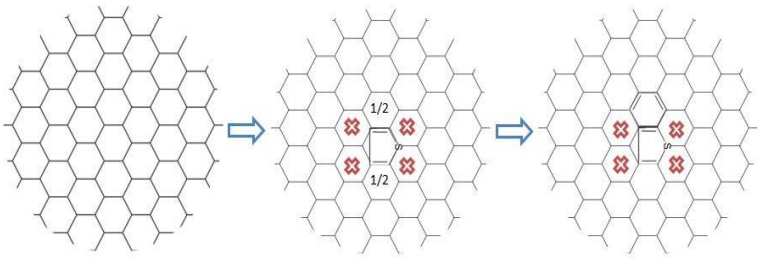


Figure A4: Assembling of the benzothiophene in the hexagonal map.

After the construction of the molecule, its pure properties are calculated, either by inspection of the structure, group contributions methods or correlations. As last remark, after the creation of N molecules the properties of the mixture are calculated and compared with the analytical data through an objective function. If convergence is not achieved the process undergoes in an optimization loop, named genetic algorithm, which will modify the parameters of the distribution. The procedure will be tested again until the properties match the number of iterations imposed.

Annex B: MD: Influence of the *K* factor

As mentioned in sun-section 5.2.1, the replication factor has a big influence in the obtained mixture used as molecule library, and consequently has, an impact in the quality of the results. Therefore, in this section, a study about this parameter is done.

The study was done for each REM mixture file containing 5000 molecules obtained by the application of the three building diagrams (A, B and C) with different *K* values: 1000, 5000, 10 000, 20 000, 30 000, 50 000, 100 000, 500 000, 1 000 000, 10 000 000. *K* values up to 30 000 were study only to understand how the mixture behaves when K tends to infinity, once the memory of the computer used does not allowed it.

For evaluating the representativeness of the mixture two variables were studied: the eliminated molecules in MD procedure and the sum of the molar fractions of the molecules survived in MD, therefore, replicated to obtain an equimolar mixture. These variables are related to the dispersion of the mixture. The goal is to achieve a better adjustment of the mixture under study without losing the representativeness of the mixture.

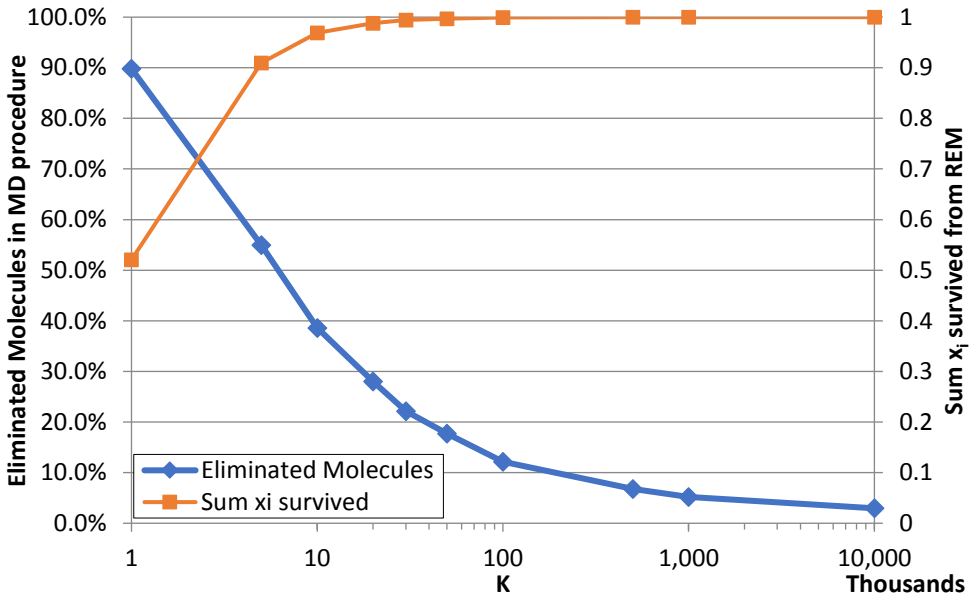


Figure B1: K impact in the mixture file MD A.

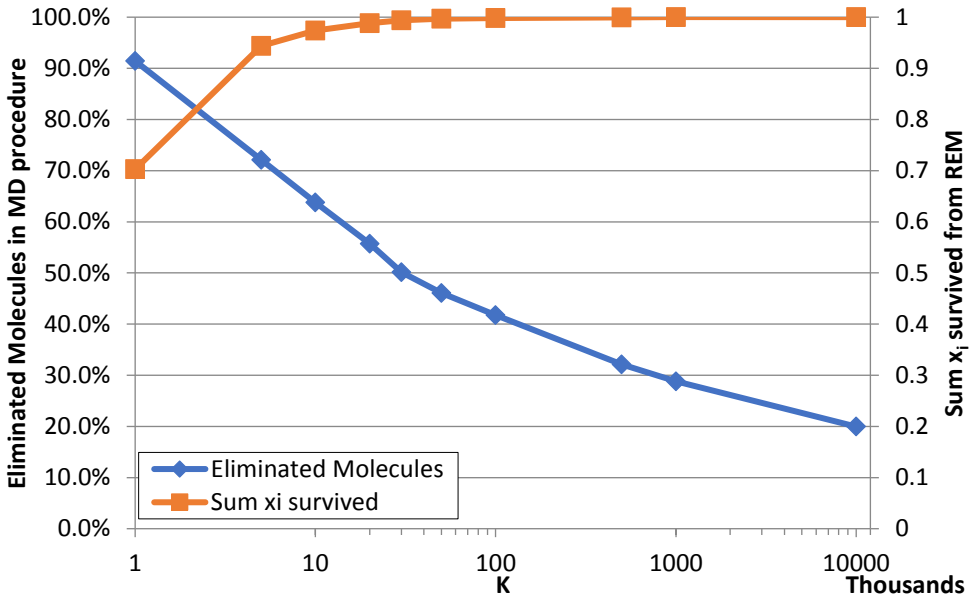


Figure B2: K impact in the mixture file MD B.

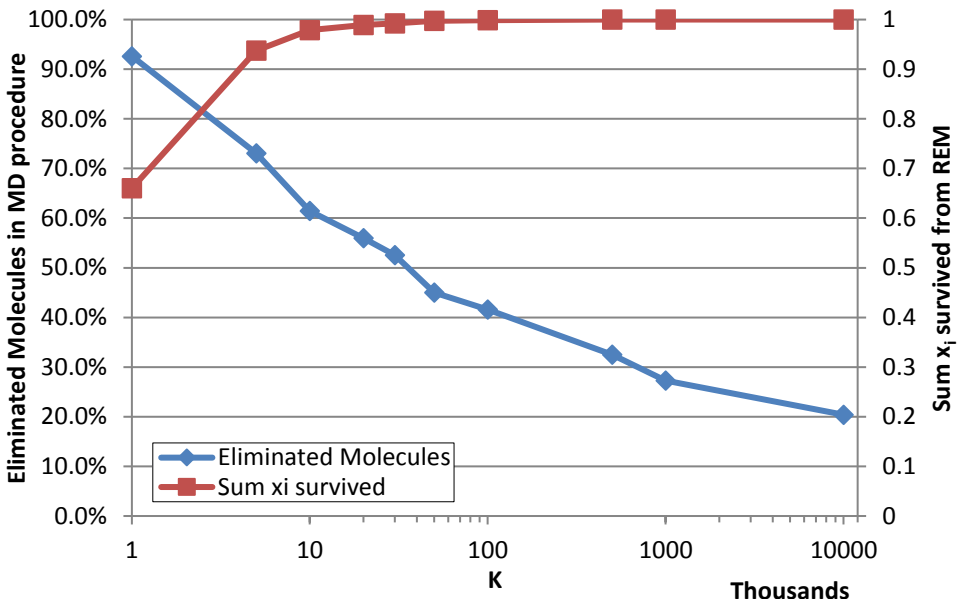


Figure B3: K impact in the mixture file MD C.

From Figure B1 to Figure B3, it is possible to analyze the evolution of the MD mixture in function of the *K* parameter. It is known that the closer the mixture is of the REM mixture the greater representativeness is achieved, this means the closer the sum of the molar fractions survived is close to 1 and the fewer molecules are eliminated. However, this implies a huge value for *K*.

In order to guarantee a good commitment between the representativeness of the mixture with a reasonable *K* value, it was imposed a minimum sum of the molar fractions equal to 0.97 which corresponds to a *K* value of 10 000.

Annex C: Building Diagrams

Table C1: Definition of the structural attributes used in Residue A version.

Number	Structural Attributes	Distribution Type	Values
1	**Type of the molecule***	Histogram	0-1-2-3
2	**Number of cores**	Exponential	>1
3	**Type of heterocycle****	Histogram	0-1-2-3
4	**Number of benzene rings per core**	Exponential	>0
5	**Total number of rings per core**	Gamma	>1
6	**Number of thiophenes per core**	Histogram	0-1-2
7	**Number of pyridines per core**	Histogram	0-1-2
8	**Number of pyrroles per core**	Histogram	0-1-2
9	**Number of furans per core**	Histogram	0-1-2
10	**Acceptance probability for a peripheral carbon**	Histogram	0-1
11	**Length of the paraffinic chain**	Gamma	>1
12	**Length of an alkyl chain (lateral and intercore)**	Exponential	>1
13	**Probability of sulfur substitution for aliphatic CH_3 or CH_2**	Histogram	0-1
14	**Substitution probability of a carbon atom by a heteroatom**	Histogram	0-1
15	**Type of heteroatom substitution*****	Histogram	0-1
16	**Type of oxygen group********	Histogram	0-1
* Type of the molecule: 0-Paraffin, 1-Naphthene, 2-Mono-Core, 3-Multi-Core ** Type of heterocycle: 0-Thiophene, 1-Pyridine, 2-Pyrrole, 3-Furan *** Type of heteroatom substitution: 0-Nitrogen, 1-Oxygen ****Type of oxygen group: 0-Ether function, 1-Carbonyl function			

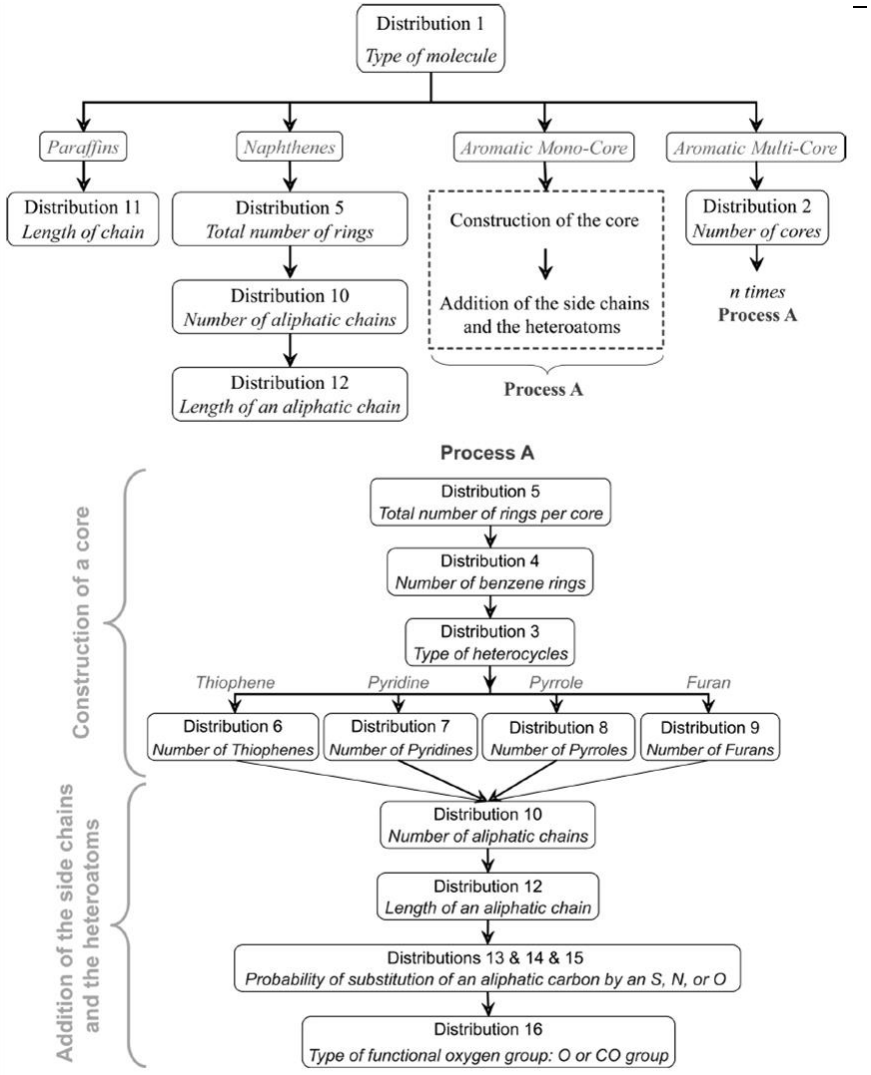


Figure C1: Residue A building diagram version.

Table C2: Definition of the structural attributes used in Residue B version.

Number	Structural Attributes	Distribution Type	Values
1	**Type of the molecule***	Histogram	0-1-2-3
2	**Length of the paraffinic chain**	Gamma	>1
3	**Number of cores**	Exponential	>1
4	**Total number of rings per core**	Gamma	>1
5	**Number of aromatic rings**	Exponential	>0
6	**Type of aromatic ring****	Histogram	0-1-2-3-4
7	**Number of aliphatic chains**	Histogram	0-1
8	**Length of an aliphatic chain**	Exponential	>1
9	**Acceptance probability for heteroatoms*****	Histogram	0-1
10	**Type of heteroatom********	Histogram	0-1-2-3
* Type of the molecule: 0-Paraffin, 1-Naphthene, 2-Mono-Core, 3-Multi-Core ** Type of aromatic ring: 0-Benzene, 1-Thiophene, 2-Pyridine, 3-Pyrrole, 4-Furan *** Acceptance probability for heteroatoms: 0-No, 1-Yes ****Type of heteroatom: 0-Sulfur, 1-Nitrogen, 2-Oxygen, 3-Carbonyle			

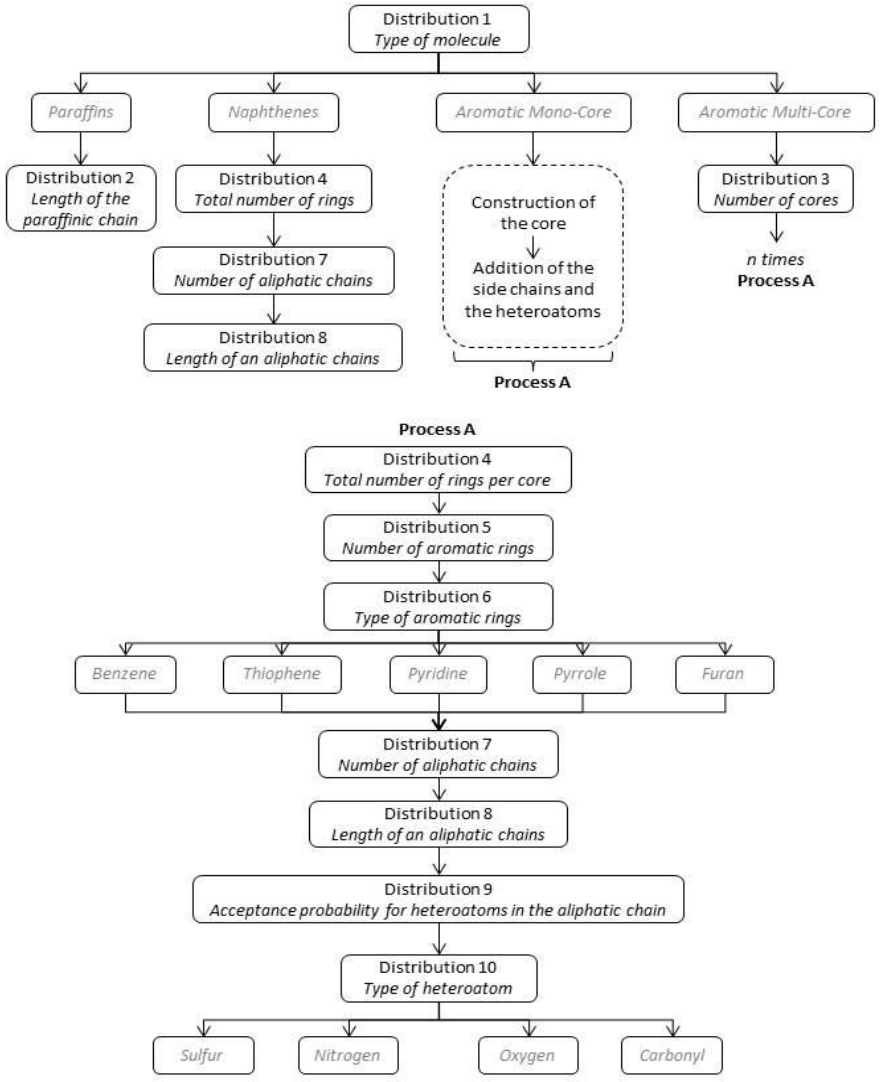


Figure C2: Residue B building diagram version.

Table C3: Definition of the structural attributes used in Residue C version.

Number	Structural Attributes	Distribution Type	Values
1	**Type of the molecule***	Histogram	0-1-2-3
2	**Length of the paraffinic chain**	Gamma	>1
3	**Number of cores**	Exponential	>1
4	**Total number of rings per core**	Gamma	>1
5	**Type of ring****	Histogram	0-1-2-3-4-5
6	**Number of aliphatic chains**	Histogram	0-1
7	**Length of an aliphatic chain**	Exponential	>1
8	**Acceptance probability for heteroatoms*****	Histogram	0-1
9	**Type of heteroatom********	Histogram	0-1-2-3

* Type of the molecule: 0-Paraffin, 1-Naphthene, 2-Mono-Core, 3-Multi-Core
** Type of ring: 0-Cyclohexane, 1-Benzene, 2-Thiophene, 3-Pyridine, 4-Pyrrole, 5-Furan
*** Acceptance probability for heteroatoms: 0-No, 1-Yes
****Type of heteroatom: 0-Sulfur, 1-Nitrogen, 2-Oxygen, 3-Carbonyle

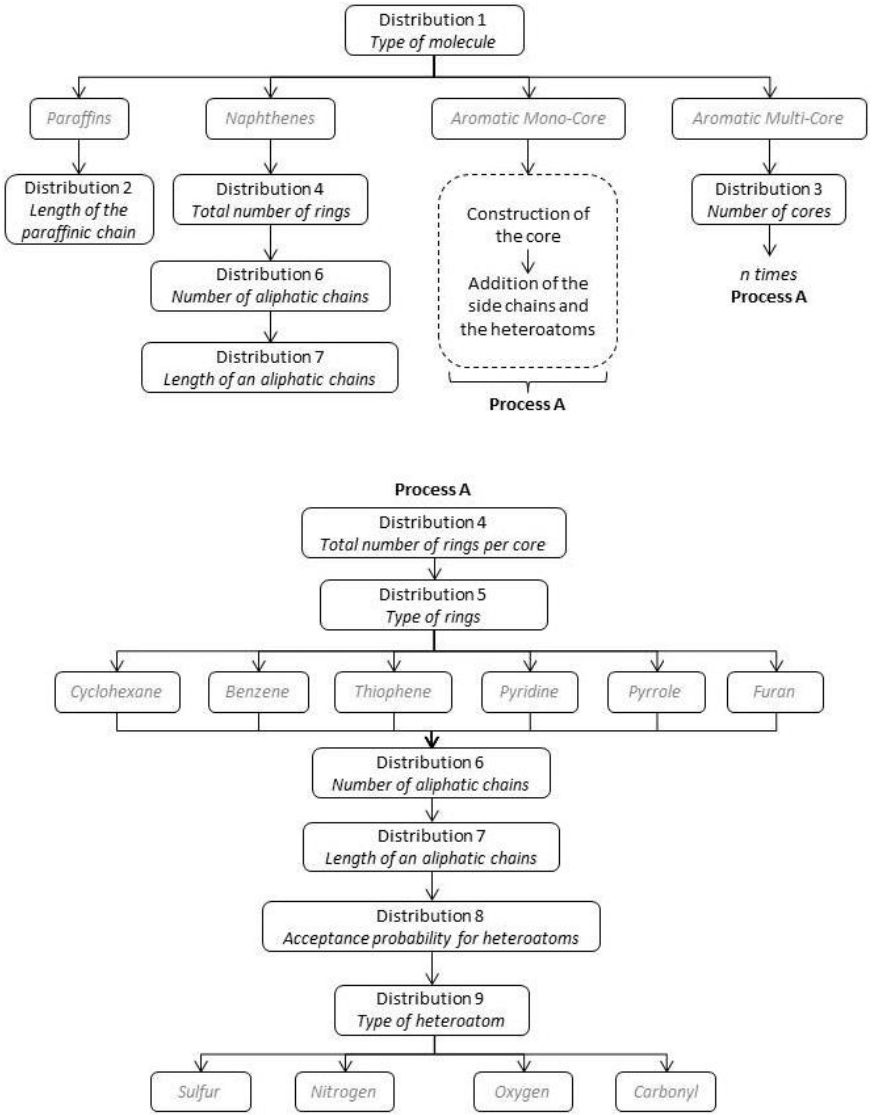


Figure C3: Residue C building diagram version.

Annex D: Thermidor Lumps

Table D1: Details about Thermidor Lumps.

Lump Family Name	Elemental Family	SARA Family	Range of NBP
C_Satmm	Carbon	Saturates	[IBP, 350°C [(MDt)
H_Satmm	Hydrogen		
C_Aromm	Carbon	Aromatics	
H_Aromm	Hydrogen		
S_Aromm	Sulfur		
N_Aromm	Nitrogen		
O_Aromm	Oxygen		
C_Satm	Carbon	Saturates	[350°C, 520°C [(VGO)
H_Satm	Hydrogen		
C_Arom	Carbon	Aromatics	
H_Arom	Hydrogen		
S_Arom	Sulfur		
N_Arom	Nitrogen		
O_Arom	Oxygen		
C_Resm	Carbon	Resins	
H_Resm	Hydrogen		
S_Resm	Sulfur		
N_Resm	Nitrogen		
O_Resm	Oxygen		
C_Satp	Carbon	Saturates	[520°C, FBP [(VR)
H_Satp	Hydrogen		
C_Arop	Carbon	Aromatics	
H_Arop	Hydrogen		
S_Arop	Sulfur		
N_Arop	Nitrogen		
O_Arop	Oxygen		
C_Resp	Carbon	Resins	
H_Resp	Hydrogen		
S_Resp	Sulfur		
N_Resp	Nitrogen		
O_Resp	Oxygen		
C_Asph	Carbon	Asphaltenes	
H_Asph	Hydrogen		
S_Asph	Sulfur		
N_Asph	Nitrogen		
O_Asph	Oxygen		

Annex E: Mass Balance Sheets

As mentioned before, the data was recovered from *Proceed Data Base*. The mass balance was performed for a calculation base of 100 kg regarding the available analytical data, in some cases, a normalization of the experimental data was necessary. Table E1 presents the values for the experimental data.

Table E1: Analytical data for point 0 and 1.

Elemental Analysis (% mass)	Point 0	Point 1	Thermidor Lumps (% mass)	Point 0	Point 1
C_LIQTOTAL	85.66	85.31	C_Satm	85.66	85.78
H_LIQTOTAL	11.46	11.36	H_Satm	14.34	14.22
S_LIQTOTAL	2.48	2.85	Total_Satm	100.00	100.00
N_LIQTOTAL	0.26	0.28	C_Arom	86.19	85.90
O_LIQTOTAL	0.13	0.20	H_Arom	10.88	10.65
Nickel_LIQTOTAL	1.37E-03	1.60E-03	S_Arom	2.49	2.94
Vanadium_LIQTOTAL	2.90E-03	3.72E-03	N_Arom	0.00	0.00
Total_LIQTOTAL	100.00	100.00	O_Arom	0.44	0.51
C_IBP_350	86.59	86.44	Total_Arom	100.00	100.00
H_ IBP _350	12.55	12.62	C_Resm	83.12	82.34
S_ IBP _350	0.64	0.71	H_Resm	7.89	7.72
N_ IBP _350	0.11	0.09	S_Resm	2.53	2.95
O_ IBP _350	0.11	0.14	N_Resm	1.74	1.67
Total_ IBP_350	100.00	100.00	O_Resm	4.72	5.32
C_350_520	85.83	86.08	Total_Resm	100.00	100.00
H_350_520	12.15	12.13	C_Satp	86.19	86.11
S_350_520	1.38	1.67	H_Satp	13.81	13.89
N_350_520	0.13	0.13	Total_Satp	100.00	100.00
O_350_520	0.51	0.00	C_Arop	85.50	84.81
Total_350_520	100.00	100.00	H_Arop	11.29	11.29
C_520+	85.43	85.02	S_Arop	2.64	3.12
H_520+	10.79	10.66	N_Arop	0.05	0.06
S_520+	3.19	3.60	O_Arop	0.52	0.72
N_520+	0.34	0.37	Total_Arop	100.00	100.00
O_520+	0.24	0.35	C_Resp	85.16	84.64
Nickel_520+	2.18E-03	2.59E-03	H_Resp	8.88	8.95
Vanadium_520+	4.59E-03	5.90E-03	S_Resp	4.22	4.73
Total_520+	100.00	100.00	N_Resp	0.75	0.75
SARA Analysis (% mass)	**Point 1**	**Point2**	O_Resp	0.98	0.92
Saturate_ IBP _350	55.0	56.0	Nickel_Resp	2.74E-03	2.99E-03
Aromatic_ IBP _350	45.0	44.0	Vanadium_Resp	4.85E-03	5.71E-03
Total_ IBP _350	100.0	100.0	Total_Resp	100.00	100.00
Saturate_350_520	44.3	44.5	C_Asph	84.58	84.58
Aromatic_350_520	49.5	49.7	H_Asph	7.41	7.42
Resines_350_520	6.2	5.8	S_Asph	6.47	6.34
Total_350_520	100.0	100.0	N_Asph	0.97	0.96
Saturate_520+	13.0	11.6	O_Asph	0.51	0.63
Aromatic_520+	45.8	44.5	Nickel_Asph	1.97E-02	2.00E-02
Resines_520+	36.1	37.9	Vanadium_Asph	4.30E-02	4.67E-02
Asph_520+	5.1	6.0	Total_Asph	100.00	100.00
Total_520+	100.0	100.0			

For the thermidor lumps of the MDt fraction was assumed that the difference of the mass percentages of the hydrogen between VGO and VR in saturates is the same that between MDt and VGO. This assumption allowed to calculate the C_Satmm, since saturates contain only carbon and hydrogen, and the aromatics elemental analysis was calculated by difference between the total elemental analysis and the elemental analysis.

The absolute error for the elemental analysis was calculated by and the absolute error of the SARA analysis through:

$$EA\ error = X_i - \sum_J X_{i,J} \qquad \textit{Equation E1}$$

$$SARA\ error = J_i - \sum_i X_{i,J} \qquad \textit{Equation E2}$$

In which:

X_i – Element *X* in the fraction *i* (*X* = C, H, S, N, O, Ni, V; *i* = LIQTOTAL, MDt, VGO or VR);

$X_{i,j}$ – Element *X* in the fraction *i* belonging to the SARA family *J* (*J* = saturates, aromatics, resins or asphaltenes);

J_i– SARA family *J* (*J* = saturates, aromatics, resins or asphaltenes) in the fraction i (*i* = TOTAL, MDt, VGO or VR).

The relative error is measured dividing the absolute error by X_i or J_i in the case of the elemental analysis or SARA analysis, respectively.

1. Point 0

x_MDt	x_VGO	x_VR
6.48%	29.73%	63.79%

IBP – 350°C

kg	Sat	Aro	Total	EA error	Relative error
Total	3.56	2.91	6.48	0.00E+00	0.00%
C_MDt	3.03	2.58	5.61	0.00E+00	0.00%
H_ MDt	0.53	0.28	0.81	0.00E+00	0.00%
S_ MDt		0.04	0.04	0.00E+00	0.00%
N_ MDt		0.01	0.01	0.00E+00	0.00%
O_ MDt		0.01	0.01	0.00E+00	0.00%
SARA error	0.00	0.00			
Relative error	0.00%	0.00%			

TOTAL

Kg	Sat	Aro	Res	Asph	Total	EA error	Relative error
Total	25.03	46.85	24.87	3.25	100.00	0.00E+00	0.00%
C_TOTAL	21.46	40.24	21.14	2.75	85.66	5.86E-02	0.07%
H_ TOTAL	3.56	5.18	2.19	0.24	11.46	2.82E-01	2.46%
S_ TOTAL		1.18	1.02	0.21	2.48	7.54E-02	3.04%
N_ TOTAL		0.02	0.20	0.03	0.26	6.62E-03	2.50%
O_ TOTAL		0.22	0.31	0.02	0.13	-4.23E-01	-325.45%
Ni_ TOTAL			0.0006	0.0006	0.0014	9.81E-05	7.16%
V_ TOTAL			0.0012	0.0014	0.0029	3.83E-04	13.22%
SARA error	0.00	0.00	0.00	0.00	0.00		
Relative error	0.00	0.00	0.00	0.00	0.00		

350-520°C

kg	Sat	Aro	Res	Total	EA error	Relative error
Total	13.17	14.72	1.84	29.73	0.00E+00	0.00%
C_VGO	11.28	12.68	1.53	25.50	1.87E-02	0.07%
H_VGO	1.89	1.60	0.15	3.64	-2.30E-02	-0.63%
S_VGO		0.37	0.05	0.41	-2.14E-03	-0.52%
N_VGO		0.00	0.032	0.032	6.58E-03	20.50%
O_VGO		0.065	0.087	0.15	-1.31E-04	-0.09%
SARA error	0.00	0.00	0.00	0.00		
Relative error	0.00	0.00	0.00	0.00		

520°C - FBP

kg	Sat	Aro	Res	Asph	Total	EA error	Relative error
Total	8.29	29.22	23.03	3.25	63.79	0.00E+00	0.00%
C_VR	7.15	24.98	19.61	2.75	54.49	7.05E-03	0.01%
H_ VR	1.15	3.30	2.05	0.24	6.73	1.53E-01	2.28%
S_VR		0.77	0.97	0.21	1.95	8.02E-02	4.11%
N_VR		0.01	0.17	0.03	0.22	-1.28E-05	-0.01%
O_VR		0.15	0.23	0.02	0.39	-2.41E-01	-61.16%
Ni_ VR			0.0006	0.0006	0.0012	1.19E-04	9.34%
V_ VR			0.0011	0.0014	2.52E-03	4.12E-04	16.35%
SARA error	0.00	0.00	0.00	0.00			
Relative error	0.00	0.00	0.00	0.00			

Assumption Saturates			
%mass	Res	VGO	MD
C_MDt	86.19%	85.66%	85.13%
H_MDt	13.81%	14.34%	14.87%

2. Point 0: Perfect Case

x_MDt	x_VGO	x_VR
6.48%	29.73%	63.79%

IBP – 350°C

kg	Sat	Aro	Total	EA error	Relative error
Total	3.56	2.91	6.48	0.00E+00	0.00%
C_MDt	3.03	2.58	5.61	0.00E+00	0.00%
H_ MDt	0.53	0.28	0.81	0.00E+00	0.00%
S_ MDt		0.04	0.04	0.00E+00	0.00%
N_ MDt		0.01	0.01	0.00E+00	0.00%
O_ MDt		0.01	0.01	0.00E+00	0.00%
SARA error	0.00	0.00			
Relative error	0.00%	0.00%			

TOTAL

350-520°C

kg	Sat	Aro	Res	Total	EA error	Relative error
Total	13.17	14.72	1.84	29.73	0.00E+00	0.00%
C_VGO	11.28	12.68	1.53	25.50	1.87E-02	0.07%
H_VGO	1.89	1.60	0.15	3.64	-2.30E-02	-0.63%
S_VGO		0.37	0.05	0.41	-2.14E-03	-0.52%
N_VGO		0.00	0.03	0.03	6.58E-03	20.50%
O_VGO		0.06	0.09	0.15	-1.31E-04	-0.09%
SARA error	0.00	0.00	0.00	0.00		
Relative error	0.00	0.00	0.00	0.00		

kg	Sat	Aro	Res	Asph	Total
Total	25.03	46.85	24.87	3.25	100.00
C_TOTAL	21.46	40.24	21.14	2.75	85.60
H_ TOTAL	3.56	5.18	2.19	0.24	11.18
S_ TOTAL		1.18	1.02	0.21	2.41
N_ TOTAL		0.02	0.20	0.03	0.26
O_ TOTAL		0.22	0.31	0.02	0.55
Ni_ TOTAL			6.31E-04	6.41E-04	1.27E-03
V_ TOTAL			1.12E-03	1.40E-03	2.52E-03
SARA error	0.000	0.000	0.000	0.000	0.000
Relative error	0.00%	0.00%	0.00%	0.00%	0.000

520°C - FBP

kg	Sat	Aro	Res	Asph	Total	EA error	Relative error
Total	8.29	29.22	23.03	3.25	63.79	0.00E+00	0.00%
C_VR	7.15	24.98	19.61	2.75	54.49	7.05E-03	0.01%
H_ VR	1.15	3.30	2.05	0.24	6.73	1.53E-01	2.28%
S_VR		0.77	0.97	0.21	1.95	8.02E-02	4.11%
N_VR		0.01	0.17	0.03	0.22	-1.28E-05	-0.01%
O_VR		0.15	0.23	0.02	0.39	-2.41E-01	-61.16%
Ni_ VR			6.31E-04	6.41E-04	1.27E-03	1.19E-04	9.34%
V_ VR			1.12E-03	1.40E-03	2.52E-03	4.12E-04	16.35%
SARA error	0.00	0.00	0.00	0.00			
Relative error	0.00	0.00	0.00	0.00			

Assumption Saturates			
%mass	Res	VGO	MD
C_MDt	86.19%	85.66%	85.13%
H_MDt	13.81%	14.34%	14.87%

3. Point 1

x_MDt	x_VGO	x_VR
4.40%	31.39%	64.20%

IBP – 350°C

kg	Sat	Aro	Total	EA error	Relative error
Total	2.466	1.938	4.403	0.00E+00	0.00%
C_MDt	2.107	1.699	3.806	0.00E+00	0.00%
H_ MDt	0.359	0.197	0.556	0.00E+00	0.00%
S_ MDt		0.031	0.031	0.00E+00	0.00%
N_ MDt		3.96E-03	3.96E-03	0.00E+00	0.00%
O_ MDt		6.16E-03	6.16E-03	0.00E+00	0.00%
SARA error	0.00	0.00			
Relative error	0.00%	0.00%			

TOTAL

kg	Sat	Aro	Res	Asph	Total	EA error	Relative error
Total	23.88	46.11	26.15	3.85	100.00	0.00E+00	0.00%
C_TOTAL	20.50	39.33	22.10	3.26	85.31	1.17E-01	0.14%
H_ TOTAL	3.38	5.08	2.32	0.29	11.36	2.92E-01	2.57%
S_ TOTAL		1.38	1.20	0.24	2.85	1.66E-02	0.58%
N_ TOTAL		0.02	0.21	0.04	0.28	1.04E-02	3.70%
O_ TOTAL		0.29	0.32	0.02	0.20	-4.36E-01	-218.22%
Ni_ TOTAL			7.27E-04	7.70E-04	1.60E-03	1.02E-04	6.40%
V_ TOTAL			1.39E-03	1.80E-03	3.72E-03	5.33E-04	14.32%
SARA error	0.00	0.00	0.00	0.00	0.00		
Relative error	0.00	0.00	0.00	0.00	0.00		

Assumption Saturates			
%mass	Res	VGO	MD
C_MDt	86.11%	85.78%	85.45%
H_MDt	13.89%	14.22%	14.55%

350-520°C

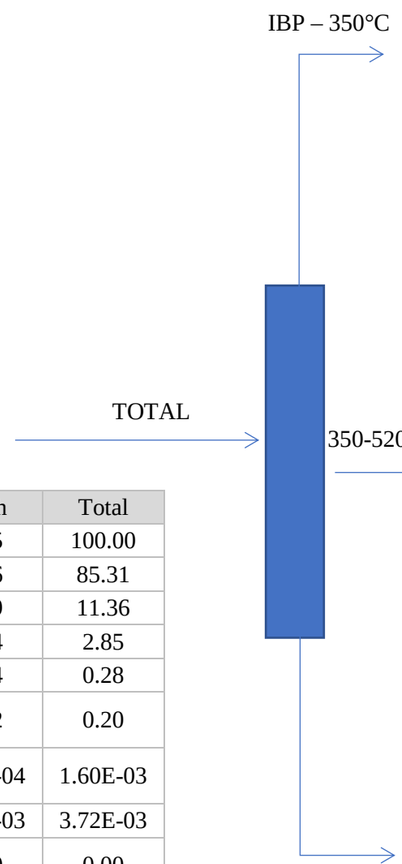
kg	Sat	Aro	Res	Total	EA error	Relative error
Total	13.97	15.60	1.82	31.39	0.00E+00	0.00%
C_VGO	11.98	13.40	1.50	26.88	1.37E-01	0.51%
H_VGO	1.99	1.66	0.14	3.79	1.92E-02	0.51%
S_VGO		0.46	0.05	0.51	1.03E-02	2.00%
N_VGO		0.0E+00	3.04E-02	3.04E-02	1.04E-02	34.21%
O_VGO		7.96E-02	9.69E-02	1.76E-01	-1.76E-01	-100.00%
SARA error	0.00	0.00	0.00	0.00		
Relative error	0.00	0.00	0.00	0.00		

520°C - FBP

kg	Sat	Aro	Res	Asph	Total	EA error	Relative error
Total	7.45	28.57	24.33	3.85	64.20	0.00E+00	0.00%
C_VR	6.41	24.23	20.60	3.26	54.50	8.56E-02	0.16%
H_ VR	1.034	3.226	2.178	0.286	6.724	1.20E-01	1.79%
S_VR		0.89	1.15	0.24	2.29	2.24E-02	0.98%
N_VR		1.71E-02	1.83E-01	3.70E-02	2.37E-01	3.21E-05	0.01%
O_VR		2.06E-01	2.24E-01	2.43E-02	4.54E-01	-2.29E-01	-50.49%
Ni_ VR			7.27E-04	7.70E-04	1.50E-03	1.65E-04	11.04%
V_ VR			1.39E-03	1.80E-03	3.19E-03	6.01E-04	18.85%
SARA error	0.00	0.00	0.00	0.00			
Relative error	0.00	0.00	0.00	0.00			

4. Point 1: Perfect Case

x_MDt	x_VGO	x_VR
4.40%	31.39%	64.20%



kg	Sat	Aro	Total	EA error	Relative error
Total	2.466	1.938	4.403	0.00E+00	0.00%
C_MDt	2.107	1.699	3.806	0.00E+00	0.00%
H_ MDt	0.359	0.197	0.556	0.00E+00	0.00%
S_ MDt		0.031	0.031	0.00E+00	0.00%
N_ MDt		3.96E-03	3.96E-03	0.00E+00	0.00%
O_ MDt		6.16E-03	6.16E-03	0.00E+00	0.00%
SARA error	0.00	0.00			
Relative error	0.00%	0.00%			

kg	Sat	Aro	Res	Asph	Total
Total	23.88	46.11	26.15	3.85	100.00
C_TOTAL	20.50	39.33	22.10	3.26	85.31
H_ TOTAL	3.38	5.08	2.32	0.29	11.36
S_ TOTAL		1.38	1.20	0.24	2.85
N_ TOTAL		0.02	0.21	0.04	0.28
O_ TOTAL		0.29	0.32	0.02	0.20
Ni_ TOTAL			7.27E-04	7.70E-04	1.60E-03
V_ TOTAL			1.39E-03	1.80E-03	3.72E-03
SARA error	0.00	0.00	0.00	0.00	0.00
Relative error	0.00	0.00	0.00	0.00	0.00

kg	Sat	Aro	Res	Total	EA error	Relative error
Total	13.97	15.60	1.82	31.39	0.00E+00	0.00%
C_VGO	11.98	13.40	1.50	26.88	1.37E-01	0.51%
H_VGO	1.99	1.66	0.14	3.79	1.92E-02	0.51%
S_VGO		0.46	0.05	0.51	1.03E-02	2.00%
N_VGO		0.0E+00	3.04E-02	3.04E-02	1.04E-02	34.21%
O_VGO		7.96E-02	9.69E-02	1.76E-01	-1.76E-01	-100.00%
SARA error	0.00	0.00	0.00	0.00		
Relative error	0.00	0.00	0.00	0.00		

kg	Sat	Aro	Res	Asph	Total	EA error	Relative error
Total	7.45	28.57	24.33	3.85	64.20	0.00E+00	0.00%
C_VR	6.41	24.23	20.60	3.26	54.50	8.56E-02	0.16%
H_ VR	1.15	3.30	2.05	1.03	3.23	2.18	0.29
S_VR		0.89	1.15	0.24	2.29	2.24E-02	0.98%
N_VR		1.71E-02	1.83E-01	3.70E-02	2.37E-01	3.21E-05	0.01%
O_VR		2.06E-01	2.24E-01	2.43E-02	4.54E-01	-2.29E-01	-50.49%
Ni_ VR			7.27E-04	7.70E-04	1.50E-03	1.65E-04	11.04%
V_ VR			1.39E-03	1.80E-03	3.19E-03	6.01E-04	18.85%
SARA error	0.00	0.00	0.00	0.00			
Relative error	0.00	0.00	0.00	0.00			

Assumption Saturates			
%mass	Res	VGO	MD
C_MDt	86.11%	85.78%	85.45%
H_MDt	13.89%	14.22%	14.55%