TATIANE GERCINA DE VILAS

IDENTIFICAÇÃO DE COMPONENTES DE FRAÇÕES DE PETRÓLEO ATRAVÉS DE MASSAS MOLECULARES PARA FORMAÇÃO DE MODELOS MOLECULARES

São Paulo
2012

TATIANE GERCINA DE VILAS

IDENTIFICAÇÃO DE COMPONENTES DE FRAÇÕES DE PETRÓLEO ATRAVÉS DE MASSAS MOLECULARES PARA FORMAÇÃO DE MODELOS MOLECULARES

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Engenharia.

Área de Concentração:
Engenharia Química

Orientador:
Prof. Dr. Cláudio Oller do Nascimento

DEDICATÓRIA

À minha família, que acredita em mim e tem me apoiado ao longo da minha vida.

AGRADECIMENTOS

Agradeço a Deus, por me suster com sua destra fiel e me dar forças para suplantar as dificuldades.

Ao professor Claudio Oller por me permitir desenvolver este projeto e pela orientação concedida.

Aos Professores do Departamento de Engenharia Química da USP.

Aos integrantes do Laboratório de Simulação e Controle de Processos da USP. Em especial para Maria do Carmo, Fabrícia e a Silvia; além de Patrícia, Omar e Willian que tiveram importante papel na minha integração ao laboratório.

Às bibliotecárias Maria Lúcia, e Fátima pela atenção e ajuda com os mais diversos arquivos da biblioteca.

A Jadson, que com sua clareza e seriedade muito me ajudou.

“Com efeito, grandes coisas
fez o SENHOR por nós; por
isso, estamos alegres.”
Salmos
126:3

“Não é o fim.
Tampouco o início do fim.
Apenas o fim do início.”
Winston Churchill

RESUMO

Os modelos moleculares são construídos com base na caracterização da composição do petróleo, que consiste em conhecer informações moleculares detalhadas com o objetivo de identificar os compostos presentes. Este trabalho tem como objetivo construir um banco de dados para a classificação de componentes constituintes de asfaltenos em grupos moleculares. A classificação é feita através da semelhança entre suas massas moleculares em grupos de séries homólogas. Tal classificação é feita por meio da utilização dos dados provenientes da análise de espectrometria de massa e de um algoritmo de busca em linguagem VBA (Visual Basic for Applications), com o objetivo de identificar uma molécula específica em um conjunto de moléculas conhecidas (banco de dados). O método proposto nesse trabalho é baseado na classificação dos compostos através de suas massas moleculares, isto é, os compostos presentes em uma mistura são separados em classes por uma diferença de massa tão pequena quanto se queira. Pode-se, assim, observar a que grupo de moléculas um composto de petróleo pertence por meio de sua massa molecular. Como o banco de dados foi gerado por compostos cuja composição é conhecida, sabe-se que composição um dado composto terá entre as possíveis composições. Para gerar moléculas baseado no conhecimento da composição de petróleo, usam-se as fórmulas gerais dos compostos orgânicos que são formados por séries homólogas. Séries homólogas são séries de compostos orgânicos com propriedades químicas similares cujos componentes diferem entre si por uma massa molecular relativa constante (CH2). Tais séries são geradas e seus pesos moleculares são calculados e armazenados juntamente com suas fórmulas moleculares, em ordem crescente de massa, podendo assim, gerar os grupos moleculares com o auxílio de uma rotina computacional em linguagem VBA.

ABSTRACT

Molecular models are built based on the characterization of the composition of oil, which consists of detailed molecular information to know in order to identify the compounds present. This work aims to build a database for classifying constituent components of asphaltenes in molecular groups. The classification is done by the similarity of their molecular weights with the use of groups of homologous series. This classification is made by using data from the analysis of mass spectrometry and a search algorithm in language VBA (Visual Basic for Applications) with the aim of identifying a molecule in a specific set of molecules known (database data). The method proposed in that work is based on the classification of the compounds through their molecular weights, i.e., the compounds in a mixture are separated into classes by a mass difference as small as one likes. One can therefore see that the group of molecules from petroleum belongs through its molecular weight. Because the database is generated from compounds whose composition is known, it is known that a given compound composition will have all the possible compositions. To generate molecules based on knowledge of the oil composition, it is used the general formulas of organic compounds which are formed by homologous series. Homologous series are series of organic compounds with similar chemical properties whose components differ by a constant relative molecular mass (CH_2). Such series are generated and their molecular weights are calculated and stored together with their molecular formulas, in ascending order of mass, being able to generate the molecular groups with the help of a computing routine in VBA language.

LISTA DE FLUXOGRAMAS

Fluxograma 1 - Fluxograma da rotina em VBA. 40

LISTA DE FIGURAS

Figura 1 - Alguns hidrocarbonetos parafínicos. 17
Figura 2 - Alguns hidrocarbonetos isoparafínicos. 18
Figura 3 - Vetor de estrutura de hidrocarbonetos e matriz estequiométrica S . 31
Figura 4 - Representação em Matriz para uma Mistura de Petróleo 34
Figura 5 - Séries Homólogas de Compostos Aromáticos com um anel. 34
Figura 6 - Fluxograma para ilustrar cálculo das propriedades físico-químicas. 35
Figura 7 - Esquema ilustrativo para busca de grupos de moléculas. 39
Figura 8 - Pseudocódigo da rotina em VBA. 41
Figura 9 - Gráfico de erro para grupo C12H17N3 da amostra 1 48
Figura 10 - Gráfico de erro para grupo C14H17NO da amostra 1..... 49
Figura 11 - Gráfico de erro para grupo C21H13N da amostra 2 50
Figura 12 - Gráfico de erro para grupo C20H19NO da amostra 2..... 51
Figura 13 - Gráfico de erro do grupo C20H16S2 da amostra 3..... 52
Figura 14 - Gráfico de erro para grupo C18H19NO da amostra 3..... 53
Figura 15 - Gráfico de erro para grupo C23H17 da amostra 4. 54
Figura 16 - Gráfico de erro para grupo C27H19NO da amostra 4..... 55

LISTA DE TABELAS

Tabela 1- Número de Isômeros de Hidrocarbonetos Isoparafínicos. 18
Tabela 2- Séries de Compostos presentes em Frações de Petróleo 37
Tabela 3- Séries Homólogas de Compostos Orgânicos..... 38
Tabela 4- Resultados da classificação da amostra 1 comparada com o banco de dados. 43
Tabela 5- Resultado da classificação da amostra 2 comparada com o banco de dados. 44
Tabela 6- Resultado da classificação da amostra 3 comparada com o banco de dados. 45
Tabela 7- Resultado da classificação da amostra 4 comparada com o banco de dados. 46
Tabela 8- Grupo do C12H17N3 para a amostra 1..... 47
Tabela 9- Grupo do C14H17NO para a amostra 1..... 48
Tabela 10 - Grupo do C21H13N para a amostra 2..... 50
Tabela 11 - Grupo do C20H19NO para a amostra 2..... 51
Tabela 12 - Grupo do C20H16S2 para a amostra 3..... 52
Tabela 13- Grupo do C18H19NO para a amostra 3..... 53
Tabela 14- Grupo do C23H17 para a amostra 4 54
Tabela 15- Grupo do C27H19NO para a amostra 4..... 55
Tabela 16- Intervalos de Peso Molecular das Amostras. 56
Tabela 17- Percentagens de Moléculas Classificadas das Amostras 57

LISTA DE ABREVIATURAS E SIGLAS

ASTM – American Society of Testing and Materials

Da – Dalton

DEQ – Departamento de Engenharia Química

FCC – Fluid Catalytic Cracking

GC – Gas Chromatography

GPC – Gel Permeation Chromatography

IUPAC – International Union of Pure and Applied Chemistry

IR - Infrared

MALDI – Matrix-Assisted Laser Desorption / Ionization

MON – Motor Octane Number

MS – Mass Spectrometry

MTHS – Molecular Type Homologous Series

NMR – Nuclear Magnetic Resonance

PIONA – Paraffins, Isoparaffins, Olefins, Naphthenes, Aromatics

QSRR – Quantitative Structure-Reactivity Relationship

RON – Research Octane Number

RVP – Reid Vapor Pressure

SI – Sistema Internacional

SOL – Structure Oriented Lumping

TBP – True Boiling point

TOF – Time of Flight

uma – Unidade de Massa Atômica

USP – Universidade de São Paulo

VBA – Visual Basic for application

LISTA DE SÍMBOLOS

T_b – Temperatura ou Ponto de Ebulição

d – Densidade Relativa

M – Peso Molecular ou Massa Molar

M_i – Peso Molecular de um Composto i

v – Viscosidade Cinemática

SUMÁRIO

1.Introdução 16
2. Revisão da Literatura 17
2.1.Composição de Frações de Petróleo 17
2.1.1. Parafinas..... 17
2.1.2. Olefinas..... 19
2.1.3. Naftenos 19
2.1.4. Aromáticos..... 19
2.2. Importância da Caracterização das Frações de Petróleo 20
2.2.1. Oportunidades e Limitações 20
2.2.2. Análise Composicional de Frações de Petróleo 21
2.2.3. Espectrometria de Massa 22
2.3. Peso Molecular.....22
2.4. Estimando o Peso Molecular 24
2.4.1. Método de Riazi - Daubert..... 25
2.4.2. Método ASTM..... 25
2.4.3. Método de Lee - Kesler..... 26
2.4.4. Correlação de Goossens 26
2.5. Modelagem em Processos de Refino.....26
2.6. Práticas Atuais em Modelagem de Processos na Refinaria.....27
2.6.1. Modelos de Agrupamento..... 27
2.6.2. Modelos Mecanísticos 29
2.6.3. Método de Agrupamento de Estrutura Orientada 31
2.6.4. Matriz MTHS..... 33
3. Materiais e Métodos.....36
3.1. Espectrometria de Massa.....36
3.2. Séries Homólogas dos Compostos.....36
3.3. Desenvolvimento do Sistema de Busca.....39
4. Resultados e Discussão.....42
5.Conclusões.....59
REFERÊNCIAS..... 60

1. Introdução

Petróleo é uma mistura complexa de hidrocarbonetos com milhares de espécies moleculares diferentes que, de acordo com a teoria mais aceita, é proveniente da decomposição de matéria orgânica em condições específicas de pressão e temperatura em reservatórios subterrâneos. Determinar a composição molecular de componentes de petróleo e seus derivados não é uma tarefa fácil com as técnicas analíticas atuais, devido ao grande número de componentes. Modelar a composição e o comportamento dessas misturas complexas nos processos de refino de petróleo tem, ao longo dos anos, evoluído para um conceito chamado agrupamento (Quann, 1992). O processo de refino de petróleo a cada dia é otimizado por diversas técnicas para atender tanto às demandas de qualidade do mercado quanto às leis ambientais que se tornam cada vez mais rigorosas. Com isso, a indústria do refino de petróleo tem produzido derivados não só livres de contaminantes mas também com propriedades particulares e composição molecular especial. Deste modo, faz-se necessário o desenvolvimento de processos mais específicos e confiáveis para se obter modelos mais significativos e melhor otimizar e gerenciar todo o processo de refino.

Para desenvolver os modelos dos processos de refino, atualmente são utilizados os modelos de agrupamento e de pseudocomponentes, que não fornecem as informações necessárias para uma modelagem molecular detalhada. Com isso, tem se utilizado conhecimentos da termodinâmica, cinética e fundamentos químicos para auxiliar na construção de modelos e no detalhamento dos compostos. Os modelos de agrupamento são rápidos e fáceis de formular e resolver, tornando-os ainda dominantes no processo de controle, design e otimização. (Quann, 1996).

Modernas técnicas analíticas têm sido utilizadas para a caracterização dos compostos de petróleo. Com informações obtidas a partir dessas caracterizações obtêm-se modelos moleculares de refino. Modelos moleculares auxiliam na construção de modelos mais detalhados para unidades de reação e separação para predição de propriedades e composição molecular de produtos. Utilizando caracterização molecular e modelos mais detalhados na

otimização global é possível monitorar, controlar e localizar importantes espécies.

Neste trabalho é desenvolvido um método de descrição da composição molecular de uma fração de petróleo. Utilizando dados obtidos a partir de técnicas analíticas para obtenção de propriedades de moléculas, gerando um banco de dados para sistematizar a identificação de componentes do petróleo.

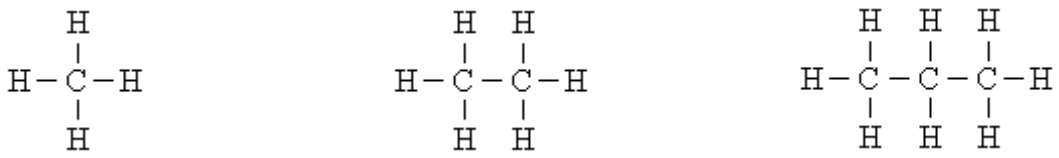
2. Revisão da Literatura

Petróleo é uma mistura complexa de hidrocarbonetos que ocorre em rochas sedimentares na forma de gases (gás natural), líquidos (óleo cru), semi-sólidos (betume) ou sólidos (cera ou asfalto). Os hidrocarbonetos são compostos orgânicos que contém apenas elementos de carbono e hidrogênio e constituem o maior grupo de compostos orgânicos. Geralmente, os hidrocarbonetos são divididos em quatro grupos: parafinas, olefinas, naftenos e aromáticos.

2.1. Composição de Frações de Petróleo

2.1.1. Parafinas

São também chamadas de alcanos e apresentam fórmula geral C_nH_{2n+2}, onde n é o número de átomos de carbono. Parafinas são divididas em normal e isoparafinas. Normal parafinas ou n-parafinas apresentam cadeia aberta, linear e saturada. Abaixo são apresentados exemplos de hidrocarbonetos parafínicos de cadeia normal:

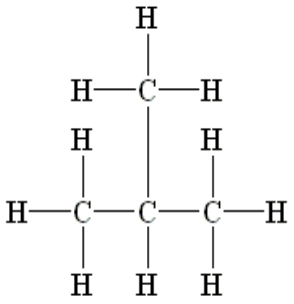


Metano Etano Propano

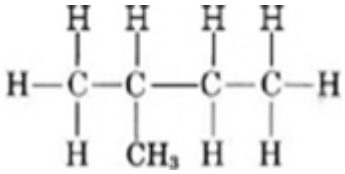
Figura 1 – Alguns hidrocarbonetos parafínicos.

Isoparafinas são hidrocarbonetos parafínicos ramificados e começam com o isobutano (metilpropano), que tem a mesma fórmula molecular do n-butano (C_4H_{10}). Compostos que apresentam estruturas diferentes, porém

mesma fórmula molecular, são chamados de isômeros. Abaixo são apresentados exemplos de hidrocarbonetos isoparafínicos:



Isobutane



Isopentano

Figura 2 – Alguns hidrocarbonetos isoparafínicos.

O número de isômeros cresce rapidamente com o número de átomos de carbono em uma molécula devido rápido aumento de possíveis arranjos estruturais. Para as parafinas no intervalo C_5-C_{12}, o número de isômeros é maior que 600 embora apenas cerca de 200 a 400 deles tenham sido identificados em misturas de petróleo (Altgelt, 1994).

Tabela 1 - Números de Isômeros de Hidrocarbonetos Isoparafínicos (Farah, 2006).

Números de Átomos de Carbono	Número de Isômeros Alcanos
C_1	-
C_2	-
C_3	-
C_4	2
C_5	3
C_6	5
C_7	9
C_8	18
C_9	35
C_{12}	355
C_{15}	4347
C_{25}	36.797.588
C_{40}	62.491.178.805.831

Sob condições padrão de 20ºC e 1 atm, os primeiros quatro membros dos alcanos se apresentam na forma gasosa, enquanto do pentano (C_5H_{12}) ao n-hexadecano ($C_{16}H_{34}$) os alcanos se apresentam como líquidos, e a partir do n-heptadecano ($C_{17}H_{36}$) os hidrocarbonetos se apresentam sólidos. Parafinas do C_1 até C_{40} normalmente são encontradas no óleo cru e representam cerca de 20% em volume. Como as parafinas são completamente saturadas (contendo apenas ligações simples), elas são estáveis e permanecem intactas por longos períodos geológicos.

2.1.2. Olefinas

Olefinas são outro grupo de hidrocarbonetos acíclicos, porém são insaturados (contém ligações duplas e/ou triplas). Hidrocarbonetos insaturados são mais reativos do que os saturados. As olefinas são incomuns em óleo cru devido a sua reatividade com hidrogênio, o que faz com que se tornem saturadas. Entretanto, podem ser produzidas em refinarias por meio de reações de craqueamento. Olefinas são produtos de elevado valor em refinarias e são usados como matéria-prima em plantas petroquímicas para produzir polímeros tais como polietileno.

2.1.3. Naftenos

Naftenos ou cicloalcanos são hidrocarbonetos cíclicos saturados de fórmula geral C_nH_{2n}. Estudos termodinâmicos mostram que aneis naftênicos com cinco e seis átomos de carbono são os mais estáveis hidrocarbonetos naftênicos. O teor de cicloparafinas no petróleo pode chegar a mais de 60%. Geralmente, qualquer mistura de petróleo que contenha hidrocarbonetos com cinco átomos de carbono também contém compostos naftênicos.

2.1.4. Aromáticos

Aromáticos são um importante grupo de hidrocarbonetos encontrados em quase toda mistura de petróleo de qualquer parte do mundo. Aromáticos são hidrocarbonetos cíclicos saturados contendo ligações duplas carbono-

carbono alternadas e seu representante mais simples é o benzeno (C_6H_6). O nome deste grupo é proveniente do fato de que tais hidrocarbonetos apresentam odor característico.

2.2. Importância da Caracterização das Frações do Petróleo

2.2.1. Oportunidades e Limitações

Se a composição das frações do petróleo fosse completamente conhecida, desde as matérias primas até o produto final, certamente seria possível fazer melhor uso do óleo cru e do processo em si. Melhores modelos de processo poderiam ser concebidos, problemas com envenenamento de catalisadores seriam mais facilmente percebidos, entendidos e solucionados. Novos catalisadores seriam desenvolvidos rapidamente e refinarias mais eficientes poderiam ser construídas com a otimização das operações *downstream* (Altgelt,1994).

Devido ao grande número de componentes de uma fração de petróleo, conhecer a composição completamente é difícil devido a dois principais obstáculos: a enorme complexidade das frações de petróleo e as limitações das técnicas de medição. Conhecimento de todos os compostos presentes no petróleo é desnecessário devido a grande similaridade entre os membros de um dado grupo ou subgrupo de compostos se estes são divididos, selecionados e separados apropriadamente. A nomenclatura desses compostos deveria se basear em termos estritamente químicos. Entretanto, o grande número de espécies moleculares no petróleo torna certos agrupamentos inevitáveis. Anteriormente foi mencionada a terminologia básica destes grupos tais como parafinas, olefinas, naftenos e aromáticos. Tradicionalmente é aceita a definição de grupos mais amplos chamados grupos de compostos, tais como saturados, aromáticos e polares. Estes grupos podem ser isolados com certa limitação por métodos cromatográficos simples. Com o auxílio de métodos cromatográficos mais refinados frações mais estreitas, chamadas classes de compostos, podem ser separadas. Exemplos de tais classes de compostos são parafinas, naftenos e aromáticos. A partir destes pode-se definir grupos mais restritos chamados tipos de compostos, tais como mononaftenos, dinaftenos, benzotiofenos, sulfitos, pirrois, entre outros. Estes

não podem ser mais separados, entretanto podem ser identificados por métodos espectroscópicos (MS, IR, NMR). O passo seguinte em direção ao aumento dessa diferenciação conduz ao conceito de séries homólogas que são séries de compostos químicos orgânicos com propriedades químicas similares cujos componentes diferem entre si por uma massa molecular relativa constante (CH_2).

2.2.2. Análise composicional de frações de petróleo

Quanto maior o ponto de ebulição de uma fração de petróleo mais difícil se torna esta análise. A composição molecular de naftas leves cujo intervalo de ebulição começa em 130℃, pode ser totalmente deter minada por uma análise tão simples quanto cromatografia gasosa. A composição de naftas pesadas, com intervalo de ebulição de 130°-220℃, é mais dif ícil de determinar. A análise de cromatografia gasosa chamada GC PIONA (parafinas, isoparafinas, olefinas, naftenos, aromáticos) reflete estas dificuldades quando se utiliza do agrupamento de componentes em cinco classes de compostos ao invés de determinar compostos individuais. A análise composicional detalhada de naftas pesadas ainda é acessível via análises simples tais como GC-MS(Altgelt, 1994).

Destilados médios, cujo intervalo de ebulição é de 220°-345℃, já representam um desafio à análise GC-MS. A determinação incompleta de compostos individuais torna a interpretação e quantificação dos dados muito difíceis, senão impossível. Neste intervalo, a cromatografia gasosa não consegue mais separar todos os componentes e a espectrometria de massa pode ver e analisar misturas de espécies pertencentes a tipos de compostos diferentes ao invés de compostos simples. Técnicas de espectrometria de massa são preferíveis nestes casos (Farah, 2006).

Frações de petróleo com ponto de ebulição acima de 345℃ são muito complexas e geralmente requerem o uso extensivo de esquemas de separação e caracterização, associadas a técnicas diversas de análise tais como espectrometria de massa.

2.2.3. Espectrometria de massa

É um tipo de análise de fração de petróleo que visa a identificação de grupos moleculares com base na ionização e fragmentação por meio da colisão com elétrons. Tal identificação se dá pela medição do peso molecular dos compostos. Enquanto compostos de petróleo voláteis e leves podem ser analisados por cromatografia gasosa, compostos mais pesados e não voláteis podem ser analisados e identificados por espectrometria de massa (Speight, 1998). Com a análise de espectrometria de massa, hidrocarbonetos que apresentam pontos de ebulição similares podem ser identificados. Com esta análise, peso molecular, fórmula molecular dos hidrocarbonetos e suas quantidades podem ser determinados (Riazi, 2005).

Os resultados provenientes do espectrômetro de massa são armazenados em um computador e processados por programas fornecidos pelo fabricante do equipamento. Programas comerciais de processamentos de dados estão disponíveis no mercado e avaliam os dados de vários modos: alguns em termos de unidades de massa atômica exata e outros pela aplicação de correções aos dados. Entretanto, devido ao elevado custo das licenças de tais programas, alguns usuários optam pelo desenvolvimento de seu próprio programa de processamento de dados.

2.3. Massa Molecular

Massa molecular ou massa molar de uma substância, representado por M, é a massa de uma molécula dessa substância relativa à unidade de massa atômica u.m.a ou Dalton (Da). No SI a unidade de massa molecular é dada em kg/kmol. A massa molecular é representada pelo mesmo número em todos os sistemas de unidades. A massa molecular de um componente puro é determinada a partir da fórmula química e do peso atômico dos elementos. O peso atômico dos elementos encontrados no petróleo C = 12.011; H= 1.008; S=32.065, O=16.0 e N=14.01, como estabelecido pelo padrão IUPAC (*International Union of Pure and Applied Chemistry*). Massa molecular é um dos parâmetros de caracterização de hidrocarbonetos.

Massa molecular é uma propriedade indicativa do tamanho e da estrutura molecular. É uma propriedade importante que normalmente os laboratórios não medem e falham em relatar quando analisam várias propriedades das frações de petróleo. Isto ocorre talvez devido à baixa precisão das medidas de massa molecular das frações de petróleo, principalmente para frações pesadas. Entretanto, deve-se atentar para o fato de que a incerteza experimental em valores de massa molecular reportados é menor do que os erros associados com os métodos de estimação para este importante parâmetro. Como frações de petróleo são misturas de hidrocarbonetos, a massa molecular de uma mistura é definida como um valor médio chamado massa molecular médio ou, de modo mais simples, massa molecular de uma mistura e é calculado pela seguinte expressão:

$$M = \sum x_i M_i$$

Onde x_i e M_i são respectivamente a fração molar e a massa molecular do componente i. A massa molecular da mistura, M, representa a razão entre a massa total e o número de moles total da mistura. Para se conhecer exatamente o peso molecular da mistura faz-se necessário conhecer a composição exata de todos os compostos na mistura. Para frações de petróleo este conhecimento exato não é viável devido ao elevado número de componentes presentes na mistura. Portanto, medidas experimentais de massa molecular de uma mistura são mais importantes do que a exata composição de todos os compostos da mistura.

Existem três métodos que são amplamente usados para medir a massa molecular de várias frações de petróleo. São eles crioscopia, método da pressão de vapor e o método de cromatografia por exclusão de tamanho. Para frações pesadas de petróleo e compostos asfaltênicos o método de cromatografia por exclusão de tamanho é normalmente usado para medir a distribuição de massa molecular na fração. Este método é principalmente usado para determinar pesos moleculares de polímeros no intervalo de 2000 a 2000000. A cromatografia por exclusão de tamanho também é chamada de cromatografia de permeação em gel (GPC) e é descrita no método ASTM D5296. No método GPC, através da comparação do tempo de eluição de uma amostra com o tempo de eluição da solução padrão a massa molecular da

amostra pode ser determinado. Este método é normalmente aplicado a resíduos e asfaltenos pesados em óleo cru e determina o peso percentual de vários constituintes versus o peso molecular (Riazi, 2005).

O método da pressão de vapor é baseado na medição da diferença entre a pressão de vapor da amostra e a pressão de vapor do solvente padrão com pressão de vapor maior que a amostra. Uma solução contendo cerca de 1 g da amostra em 25 mL do solvente referência é preparada. Esta solução, que apresenta pressão de vapor menor do que o solvente, tende a condensar os vapores do solvente na amostra estabelecendo, assim, uma diferença de temperatura que é medida por dois sensores. A molaridade da solução é calculada utilizando-se curvas calibradas. Este método é descrito pela norma ASTM D 2503 e é aplicável a óleos com um ponto de ebulição inicial maior que 220ºC (Denis *et al*,1997). Uma incerteza experimental típica para a medição do peso molecular é por volta de 5% (Speight, 1998).

O terceiro e mais amplamente usado método de determinação da massa molecular de uma mistura de petróleo de composição desconhecida é o método crioscópico, que é baseado na depressão do ponto de congelamento. O ponto de congelamento de uma solução é uma medida da concentração da solução. O aumento da concentração do soluto conduz à diminuição do ponto de congelamento da solução. A relação entre a depressão do ponto de congelamento e a concentração é linear. Para hidrocarbonetos orgânicos, benzeno é normalmente usado como solvente. Curvas de calibração podem ser preparadas por meio da medição de pontos de congelamento de diferentes concentrações de soluto com soluto e solvente conhecido (Hsu, 2003).

2.4. Estimando o Peso Molecular

Para hidrocarbonetos puros provenientes de grupos homólogos são apresentados a seguir alguns métodos para obtenção do peso molecular.

2.4.1. Método de Riazi-Daubert

Os métodos clássicos para obtenção do peso molecular se utilizam do ponto de ebulição bem como da massa específica dos compostos. Para frações pesadas de petróleo tais informações podem não estar disponíveis. Por esta razão Riazi e Daubert desenvolveram uma correlação em função da viscosidade cinemática baseada no peso molecular de frações pesadas no intervalo de 200 – 800:

$$M = 223{,}56\ [v_{38}^{(-1{,}2435+1{,}1228d)}\ v_{99}^{(3{,}4758-3{,}038d)}]\ d^{-0{,}6665} \quad (1)$$

Os três parâmetros de entrada são viscosidades cinemáticas (em cSt) a 38ºC e 98,9ºC representados por v_{38} e v_{99}, respectivamente e a densidade relativa, d, a 15,5ºC. Deve-se notar que viscosidades a duas temperaturas diferentes representam dois parâmetros independentes: sendo um o valor da viscosidade e o outro o efeito da temperatura na viscosidade. O uso de um terceiro parâmetro é necessário para caracterizar a complexidade de hidrocarbonetos pesados que seguem uma relação de energia potencial a três parâmetros. Este método só é recomendado quando o ponto de ebulição dos compostos não está disponível.

2.4.2. Método ASTM

O método contido na norma ASTM D 2502 apresenta um gráfico para calcular o peso molecular de óleos viscosos usando a viscosidade cinemática a 38ºC e 98ºC. Este método foi desenvolvido empiricamente por Hirschler em 1946 e é dado pela seguinte equação:

$$M = 180 + K(H_{38} + 60) \quad (2)$$

Onde

$$K = 4{,}145 - 1{,}733 \log_{10}(VSF - 145)$$

$$VSF = H_{38} - H_{99}$$

$$H = 870 \log_{10} [\log_{10} (v + 0{,}6)] + 154$$

Sendo v a viscosidade cinemática em cSt.

2.4.3. Método de Lee – Kesler

Segundo Kesler e colaboradores, a massa molecular é relacionada ao ponto de ebulição (T_b) e à densidade relativa (d) através da seguinte correlação:

$$M = -12272,6+9486,4\ d + (8,3741-5,9917d)T_b + (1-0,77084d-0,02058d^2) \times (0,7465-222,466/T_b)10^7/T_b + (1-0,80882d + 0,02226d^2) \times (0,3228-17,335/Tb)10^{12}/T_b^3$$

Dados de massa molecular elevada também foram usados na obtenção das constantes. Esta correlação é recomendada para pontos de ebulição acima de 477ºC.

2.4.4. Correlação de Goossens

Mais recentemente, Goossens relacionou a massa molecular com o ponto de ebulição (T_b) e densidade relativa (d) a 20ºC usando dados de 40 hidrocarbonetos puros e 23 frações de petróleo da seguinte forma:

$$M = 0,01077\ T_b^{\beta}/d_4^{20}$$

Onde $\beta = 1,52869\ \ln[T_b/(1078- T_b)]$

Entretanto, a aplicação prática desta equação é limitada a frações de petróleo de baixa massa molecular devido à indisponibilidade de dados de destilação concernentes a frações pesadas.

2.5. Modelagem em Processos de Refino

A indústria do refino de petróleo é uma indústria vital, que consiste em converter óleo cru e gás natural em vários óleos, lubrificantes e produtos petroquímicos que são utilizados para melhorar nossa qualidade de vida. Com o intuito de melhorar o processo de refino, muitos autores têm combinado a ideia de modelagem molecular com métodos de programação matemática avançada para alcançar melhor desempenho no projeto, além de otimização

integrada. A ideia principal para este tipo de abordagem é alcançar um melhoramento macroscópico baseado em fundamentos de fenômenos microscópicos. Em outras palavras, buscar informações moleculares para construir modelos moleculares de refino e aumentar o entendimento em vários processos.

Tais modelos são implementados em sistemas de otimização integrados, que são utilizados para selecionar matéria-prima e catalisadores, otimizar as condições de diferentes processos e determinar quais delas são viáveis economicamente. Deste modo, o lucro da refinaria vai ser maximizado aumentando a quantidade de moléculas desejáveis para formar produtos mais valiosos, e minimizar produtos indesejáveis como aromáticos e sulfurados em compostos da gasolina e em emissões gasosas. (Hu, 2001)

Descrever o comportamento de misturas de petróleo é importante para o projeto e operação de processos de refino, mas o principal desafio consiste em identificar os milhares de componentes e estabelecer suas complexas reações catalíticas. A análise de matéria-prima é limitada às medidas de propriedades de massa tais como conteúdo de carbono e hidrocarboneto, tipo de carbono e hidrogênio RMN, limitadas análises de espectroscopia de massa de tipos de componentes predominantes e distribuição de pontos de ebulição, (Quann, 1992).

2.6. Práticas atuais em modelagem de processos na refinaria

A modelagem química é a etapa mais importante do projeto da refinaria, controle e otimização. Frequentemente, para modelagem de processos em refinaria constroem-se modelos onde moléculas individuais na matéria-prima do hidrocarboneto são agrupadas em amplas, porém mensuráveis, categorias de classes de compostos ou faixas de pontos de ebulição (Hu, 2001).

2.6.1. Modelos de Agrupamento

Agrupamento é a agregação de grandes conjuntos de espécies dentro de um conjunto menor de pseudo-espécies no contexto de reações químicas. O uso de agrupamento é motivado por, no mínimo, dois amplos interesses.

Primeiro, de significância estatística e de fácil uso nos modelos de reação que possibilita uma formulação de simples descrição de agrupamento cinético. Segundo, e talvez o mais prático, a química analítica frequentemente fornece apenas grupos ou agrupamentos de moléculas como base para a formulação de modelos. A ideia básica fundamental para os modelos de agrupamento é a associação de uma propriedade, por exemplo, reatividade com uma identidade de espécie molecular. Em modelos de agrupamento discretos, isso será frequentemente a associação de uma constante de taxa de reação com ponto de ebulição ou solubilidade de agrupamentos classificados.

Reações em cadeia são construídas refletindo o entendimento de processos químicos e são construídas para esses agrupamentos. Modelos de agrupamento são rápidos e fáceis de formular e resolver. Isto explica o fato de eles permanecerem dominantes em processos de controle, projeto e otimização. Os modelos de agrupamento têm sido aplicados para modelar quase todos os processos de reações na refinaria incluindo craqueamento catalítico, hidrotratamento (Gray, 1990), hidrocraqueamento (Stangeland, 1974), hidropirólise (Parnas, 1988), asfaltenos e resíduos de pirólise. Os modelos de agrupamento são muito úteis quando o principal requerimento são informações amplas como rendimento ou número de octano. No entanto, os modelos de agrupamento não podem fornecer detalhadas informações moleculares tais como conteúdo de parafinas, olefinas, benzeno e enxofre, que são muito importantes para as atuais operações em uma refinaria (Hu, 2001).

As abordagens de modelos de agrupamento são também amplamente utilizadas em processos de separação (Katz, 1933) e em *softwares* comerciais tais como ASPEN, PRO II e HYSIS (Katz, 1933). Mas o método de agrupamento é diferente dos usados pelos modelos dos processos de reação. Por exemplo, na modelagem da destilação, misturas de petróleo são caracterizadas em termos de estreitas faixas de ponto de ebulição como pseudocomponentes, que é diferente dos agrupamentos químicos usados para modelagem de reação. A incompatibilidade nos métodos de agrupamento faz a modelagem de ambos os processos de reação e separação inconsistentes (Hu, 2002).

Modelos de agrupamento falham frequentemente quando utilizados para diferentes matérias-primas devido às diferenças de composição dentro dos

mesmos grupos definidos. A composição atual desses grupos em termos de componentes moleculares pode mudar com a conversão global do sistema e mascarar a verdadeira cinética. Modelos de agrupamento simplificados não podem ser utilizados para interpretar os efeitos das propriedades catalíticas nos aspectos fenomenológicos da catálise química, devido ao fato de que os mecanismos fundamentais da catálise química não são incorporados dentro do esquema cinético. Os modelos de agrupamento possuem insuficientes detalhes para predizer sutis mudanças nas propriedades dos produtos, este fato sendo sua principal limitação (Quann, 1992).

2.6.2. Modelos Mecanísticos

Recentes avanços em química analítica fornecem modeladores com finos agrupamentos de reagentes, de fato, até materiais na faixa do ponto de ebulição do gasóleo. Isso é razoável para visualizar os agrupamentos como isômeros de moléculas. Esta nova informação estrutural revela as limitações dos modelos de agrupamento baseados em ponto de ebulição global. Por exemplo, moléculas com pontos de ebulição muito similares podem sofrer diferentes mecanismos de reação com diferentes cinéticas globais. Estas novas informações disponíveis têm motivado a formulação de uma nova classe de modelos de estruturas moleculares explícitas ou modelos mecanísticos. Esses modelos capturam os detalhes e a complexidade da mistura de hidrocarbonetos multicomponentes. Entre as vantagens desse modelo de estrutura explícita está o fato de que a formulação molecular permite a introdução de fundamentos de físico-química e química orgânica para os modelos de reação química (Nigam, 1993).

O desafio para este tipo de modelagem é que tanto as matérias-primas quanto os produtos contém milhares de componentes que reagem através de caminhos com milhares constantes de taxas de reações associadas. Estudos anteriores utilizando complexas matérias-primas de gasóleo conduziu a modelos empíricos que prediz a inter conversão de agrupamentos de pseudocomponentes tais como parafinas, olefinas, naftênicos e aromáticos ou classes globais de reagentes, como ponto de ebulição. Esses modelos são tipicamente específicos da matéria-prima e as informações de fundamentos é

obscurecida devido à natureza multicomponente de cada agrupamento (Russell, 1994).

Simulações com modelos mecanísticos foram realizadas por Nigam, para a pirólise de hidrocarbonetos pesados. Uma típica mistura complexa pode conter da ordem de 10^8 espécies moleculares distintas. Cada molécula exibe uma estrutura molecular relativamente complexa, elas podem ter acima de 5 aneis aromáticos e um grande número de substituições. A simulação mecanística direta da pirólise das moléculas de hidrocarbonetos pesados e suas misturas são difíceis devido ao grande número de espécies, reações e constantes de taxas associados. A complexidade dos hidrocarbonetos pesados diminui quando as moléculas são visualizadas como sendo pertencentes a uma classe de compostos, tais como parafinas, isoparafinas, alquilbenzenos, alquilnaftenos, alquilaromaticos, entre outras. A complexidade da reação diminui quando os passos das reações são agrupados em clássicas, quebras de ligação, beta-cisão, transferência de hidreto, entre outras. A cinética de cada família de reação foi resumida em termos do relacionamento da quantidade estrutura-reatividade (QSRR). A essência para essa abordagem de agrupamento é a similaridade nas estruturas moleculares e nos passos elementares (Nigam, 1993).

De outro modo, modelos mecanísticos moleculares podem fornecer muitas informações sobre a química da molécula. Klein e colaboradores realizaram pesquisas em modelagem mecanística para as reações de craqueamento catalítico (Watson, 1997), hidroprocessamento (Russell, 1994), e pirólise (Broadbelt, 1994). Esses modelos foram desenvolvidos com base em etapas de reações fundamentais tais como beta-cisão, hidreto shift, metil shift, quebras de ligações, entre outros, que envolvem transições elementares de sítios ativos, por exemplo, superfície intermediária em reações catalíticas ou radicais livres em reações térmicas (Walter, 1995).

Froment e colaboradores também desenvolveram detalhes moleculares para modelos mecanísticos, que segue superfícies intermediárias e transformação molecular fundamental para hidrocraqueamento (Baltanas, 1989), craqueamento térmico (Clymans, 1984), e hidrodesulfirização (Froment, 1994). Como também craqueamento catalítico (Svoboda, 1995). Os modelos mecanísticos têm a vantagem de serem ricos em fundamentos químicos. No

entanto eles enfrentam várias dificuldades nas aplicações em problemas de otimização. Isto acontece devido às reações em cadeias serem baseadas em modelos moleculares mecanísticos que podem conduzir a uma quantidade muito grande de espécies moleculares. Deste modo o grande número de espécies e reações envolvidas pode conduzir a um enorme número de equações algébricas diferenciais, que são difíceis de tratar numericamente (Hu, 2002). Além disso, modelos mecanísticos requerem determinações precisas das composições das misturas, que são difíceis de medir.

2.6.3. Método de Agrupamento de Estrutura Orientada – SOL

Na tentativa de superar os problemas acima, a Mobil desenvolveu uma estrutura de agrupamento orientada (SOL), abordagem feita para modelar os processos de hidroprocessamento catalítico (Quann, 1996) e craqueamento catalítico fluidizado (FCC). Uma complexa mistura de petróleo no método SOL é descrita usando um vetor de incrementos estruturais, como mostrado na Figura 3.

		A6	A4	A2	N6	N5	N4	N3	N2	N1	R	br	me	H	AA	NS	RS	AN	NN	RN	NO	RO	KO
2,3,5 Trimetilexano		0	0	0	0	0	0	0	0	0	9	3	0	1	0	0	0	0	0	0	0	0	0
Benzeno		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Naftaleno		1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fenantreno		1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pireno		1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Carbazol		2	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0
Benzofurano		1	0	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	1	0	0
Fluoranona		2	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1
Adamantano		0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Colesterol		0	0	0	1	0	2	1	0	0	10	2	2	-1	0	0	0	0	0	0	0	1	0
Água	H_2O	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0

Figura 3 - Vetor de estrutura de hidrocarboneto a e matriz estequiométrica S (Jaffe ,2005).

No agrupamento de estrutura orientada, moléculas de hidrocarbonetos podem ser descritas como um vetor, com os elementos do vetor representando características estruturais para construir qualquer molécula. Diferentes moléculas com os mesmos grupos estruturais, ou seja, isômeros são representados pelo mesmo vetor (Quann, 1992).

Cromatografia líquida ou cromatografia gasosa acoplada à espectroscopia de massa (GC/MS) pode determinar a distribuição de número de carbono de muitas séries homólogas, rendendo uma análise e caracterização do óleo em termos de milhares de tipos de moléculas. Embora muitas moléculas de hidrocarbonetos possam ser representadas por um conjunto de grupos estruturais, o método apenas indica quais grupos estruturais estão presentes na molécula, não seus arranjos específicos dentro da molécula. Uma estrutura específica de vetor pode representar muitas moléculas tendo os mesmos conjuntos de grupos estruturais, mas com diferentes posicionamentos espaciais. Esses são chamados isômeros estruturais nesse contexto e pode envolver diferentes posicionamentos de aneis naftênicos aromáticos, numerosas configurações de grupos alquil, posicionamento das ramificações e de heteroátomos e ligações olefínicas. O raciocínio que está por trás negligencia o arranjo de grupos estruturais no vetor e a suposição é que as propriedades físicas e químicas de isômeros estruturais são similares. O agrupamento de isômeros estruturais reduz a complexidade do sistema de milhões para cerca de milhares de moléculas.

Técnicas analíticas para diferenciar isômeros ainda não estão disponíveis. Resolução de isômeros pelo método CG é possível apenas em moléculas de baixo peso molecular na faixa da nafta (C4-C9). Técnicas analíticas para destilar hidrocarbonetos de cadeias longas são incapazes de distinguir isômeros estruturais, exceto em alguns casos especiais tais como normal parafina ou anéis aromáticos simples. Apenas massa molecular e número de aneis aromáticos em que incrementos estruturais podem ser atribuídos são medidos (Quann, 1992).

A estratégia para lidar com a complexidade de isômeros é o uso de um conjunto de convenções para selecionar um isômero para representar o agrupamento de isômeros estruturais correspondente a uma particular estrutura de vetor. Essas convenções são baseadas em conhecimento de

estrutura molecular de hidrocarbonetos de óleo cru e argumentos termodinâmicos. Compostos orgânicos derivados de petróleo tendem a ter regularidade na estrutura padrão devido a sua origem biológica (Quann, 1992).

Regras de reação, por exemplo, saturação de aneis e remoção de enxofre, derivadas de reações químicas fundamentais são então aplicadas para determinar as mudanças do conjunto de vetores estruturais e assim construir para as reações em cadeia. Uma vantagem do método SOL é que ele permite o desenvolvimento de uma descrição matemática e algorítmica de processos químicos (Hu, 2002).

2.6.4. Matriz MTHS

Hu recentemente apresentou uma estrutura de modelagem molecular para os processos da refinaria. Ao contrário do método de caracterização de parâmetro estrutural médio, propôs o método de caracterização que é baseado em uma matriz de representação MTHS (tipo molecular série homóloga). Esse método usa grupos homólogos e número de carbono para caracterizar uma fração de petróleo. Este método pode efetivamente representar informações moleculares básicas para óleo cru, que pode incluir 10^8 diferentes moléculas (Hu, 2001).

Para modelar processos usando informações moleculares, é necessário modelar as correntes da refinaria, que são misturas de petróleo consistindo de numerosas moléculas de hidrocarbonetos que pode ser divididas em parafinas, olefinas, naftênicos, aromáticos e compostos com heteroátomos. É impossível descrever cada molécula de uma corrente e também desnecessário porque muitos tipos de moléculas estão presentes em pequenas quantidades (Hu, 2002). Os tradicionais métodos de caracterização para misturas de hidrocarbonetos são baseados em pseudocomponentes e compostos de classes e parâmetro estrutural médio (Oka, 1977). A matriz MTHS na Figura 4 é usada para representar composição molecular de misturas de petróleo, onde nP, iP, O, N e A representam diferentes séries homólogas correspondente a normal parafinas isoparafinas, olefinas, naftenos e aromáticos, respectivamente.

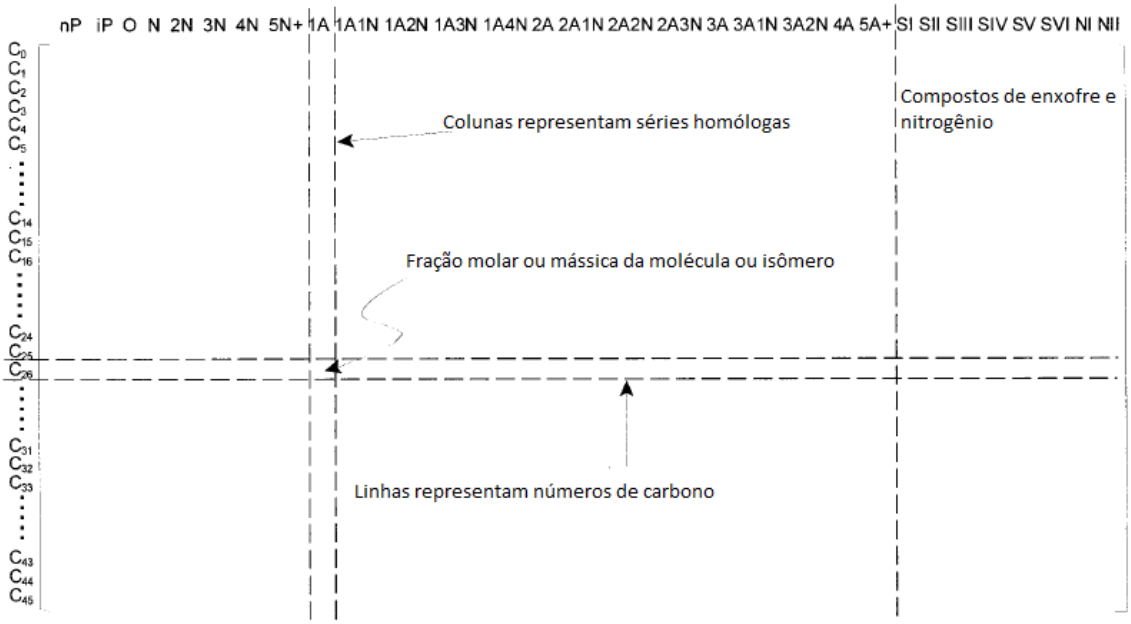


Figura 4 - Representação em Matriz para uma Mistura de Petróleo (Hu, 2002).

As séries homólogas têm as mesmas estruturas-base, mas com variável número de carbonos. A Figura 5 mostra uma série homóloga de aneis aromáticos.

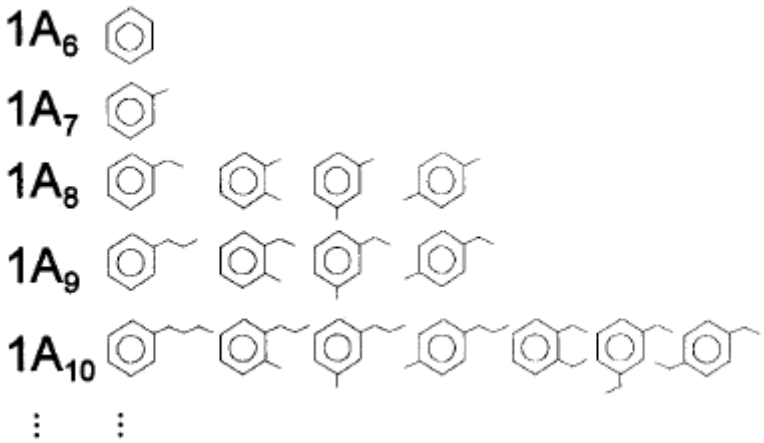


Figura 5 - Séries Homólogas de Compostos Aromáticos com apenas um anel (Hu, 2002).

Os tipos moleculares, tais como aneis benzênicos e aneis naftênicos, capturam as diferentes estruturas bases de cada série. As moléculas que pertencem a uma mesma série homóloga, com o mesmo número de carbono têm propriedades físicas similares e assim eles são agrupados em um único

componente como elemento da matriz. O tipo molecular junto com o número específico de carbono determina a reatividade e propriedades físicas e químicas das espécies. As séries homólogas também formam a base para correlações moleculares de estruturas e propriedades.

Depois de obter a matriz molecular para diversas correntes, são conhecidos os parâmetros de estruturas moleculares tais como, tipo molecular e número de carbono, dos quais é possível determinar correlações moleculares de estrutura e propriedades, (as quais não se encontram na literatura pública), determinam-se propriedades moleculares fundamentais tais como, peso molecular, densidade e ponto de ebulição. Essas moléculas são misturadas de acordo com a composição e acumuladas por volume percentual para gerar a distribuição da curva do verdadeiro ponto de ebulição (TBP) das correntes. A curva TBP é então convertida dentro de um ensaio ASTM (Sociedade Americana para Testes e Materiais), curva de destilação ASTM D86 usando correlações existentes. Seguindo esses passos, propriedades de correntes incluindo densidade, peso molecular, viscosidade, índice de refração, ponto de orvalho, ponto de fluidez, número de cetano e conteúdo de parafinas, naftênicos e aromáticos, podem ser estimadas. Número de octano (RON), número de octano motor (MON), pressão de vapor Reid (RVP) de correntes pode ser calculados por correlações baseados em composições moleculares (Hu, 2002). A sequência para determinação das propriedades está ilustrada na Figura 6, abaixo:

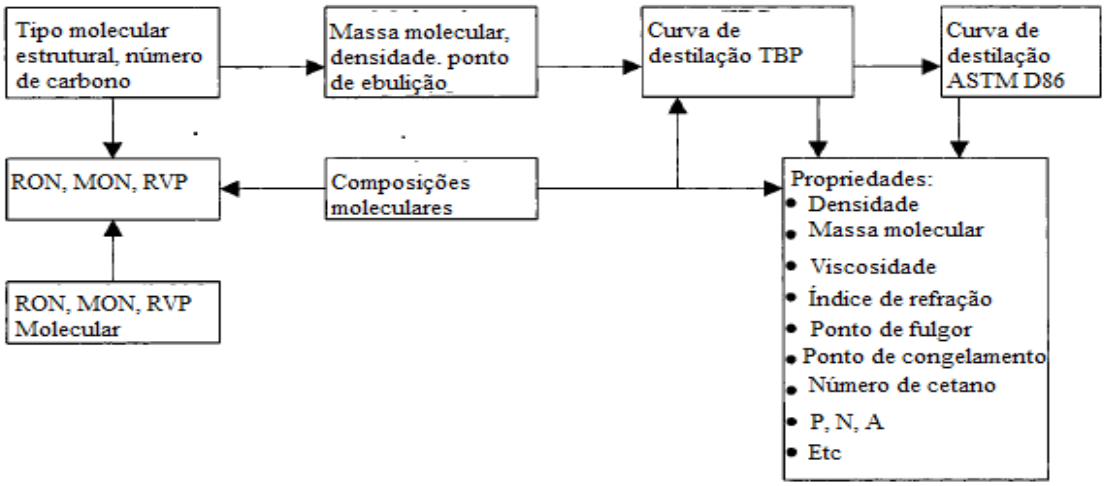


Figura 6 - Fluxograma para ilustrar cálculo das propriedades físico-químicas (Hu, 2002).

3.Materiais e Métodos

A metodologia descrita abaixo é divida em três partes diferentes. A primeira trata de uma breve descrição das condições utilizadas na espectrometria de massa para análise de amostras de asfaltenos e posteior obtenção das massas moleculares. A segunda é a descrição de séries homólogas para a geração do banco de dados, e a terceira é o funcionamento do sistema de identificação dos grupos moleculares.

3.1. *Espectrometria de massa*

As amostras de asfaltenos foram preparadas e sequencialmente analisadas pela técnica espectrométrica de MALDI TOF (*matrix- assisted laser desorption/ionization time-of-flight*). A influência da matriz foi avaliada, testes de repetitividade foram feitos, os espectros de massa foram obtidos gerando um banco de dados com esses espectros. Inicialmente foi realizada uma análise espectral para determinar a faixa de coleta dos espectros. A maioria dos picos significativos das amostras esteve entre 100 a 1200 Da sendo a faixa escolhida para as análises. O MALD TOF determina massas molares de macromoléculas, tais como asfaltenos, compostos de maior peso molecular entre os compostos do petróleo, que foram as amostras utilizadas para as análises através desta técnica (Pandoja, 2010).

As amostras de asfaltenos foram preparadas e analisadas por Pandoja, 2010. As massas moleculares foram cedidas como um resultado da parceria do trabalho em laboratório.

Uma vez concluídas as análises e obtidas as massas moleculares das amostras de asfaltenos, através da técnica acima citada, faz-se necessária a identificação ou classificação dos compostos através de suas massas. Deste modo, futuras determinações de propriedades destes compostos são necessárias para uma caracterização completa dessa fração.

3.2. Séries Homólogas dos Compostos

Para gerar moléculas baseado no conhecimento de componentes de petróleo, foram utilizadas as fórmulas gerais de componentes orgânicos

reconhecidamente presentes no petróleo. Tais fórmulas moleculares foram obtidas com base em referências bem estabelecidas na literatura (Hsu, 2003). Buscou-se congregar tais moléculas seguindo o conceito das séries homólogas. As séries homólogas são séries de compostos químicos orgânicos com propriedades químicas similares cujos componentes diferem entre si por uma massa molecular relativa constante (CH_2). Abaixo, na Tabela 2, estão alguns exemplos de séries de compostos e funções orgânicas presentes em petróleo, tais como, hidrocarbonetos simples, compostos nitrogenados, compostos sulfurados, compostos oxigenados, compostos aromáticos e compostos saturados de cadeias fechadas, entre outros, determinados para compor o banco de moléculas utilizado neste trabalho.

Tabela 2 - Séries de Compostos Presentes em Frações de Petróleo (Hsu, 2003)

Classe	Intervalos de Fórmulas de Tipos de compostos			Número de Séries Homólogas
Hidrocarbonetos	C_nH_{2n}	-	C_nH_{2n-52}	27
Compostos Sulfurados	$C_nH_{2n+2}S$	-	$CnH_{2n-42}S$	23
	$C_nH_{2n-12}S_2$	-	$C_nH_{2n-30}S_2$	11
Compostos Oxisulfurados	$C_nH_{2n-10}SO$	-	$C_nH_{2n-30}SO$	11
	$C_nH_{2n-10}SO_2$	-	$C_nH_{2n-18}SO_2$	5
	$C_nH_{2n-10}SO_3$	-	$C_nH_{2n-2}0SO_3$	6
Compostos Nitrogenados	$C_nH_{2n-3}N$	-	$C_nH_{2n-51}N$	25
	$C_nH_{2n-12}N_2$	-	$C_nH_{2n-26}N_2$	8
	$C_nH_{2n-5}N_3$	-	$C_nH_{2n-11}N_3$	4
Compostos Oxigenados	$C_nH_{2n-2}O$	-	$C_nH_{2n-50}O$	25
	$C_nH_{2n-6}O_2$	-	$C_nH_{2n-46}O_2$	19
	$C_nH_{2n-6}O_3$	-	$C_nH_{2n-32}O_3$	14
	$C_nH_{2n-10}O4$	-	$C_nH_{2n-20}O_4$	7
Compostos Nitrooxigenados	$C_nH_{2n-7}NO$	-	$C_nH_{2n-41}NO$	18

Como as frações de petróleo possuem substancialmente a mesma composição independente de sua origem (Thomas, 2004). É possível representar uma dada fração de petróleo, por séries de compostos, como descrito na Tabela 2.

A Tabela 3 apresenta exemplos de séries de funções orgânicas, mostrando algumas restrições de número de carbono. Sendo valido salientar que ao desenvolver séries de compostos é necessário obedecer a critérios de existência das moléculas, como número mínimo de carbono, para obter um banco de dados confiável.

Tabela 3 – Séries Homólogas de Compostos Orgânicos

Grupos Orgânicos	Fórmula Geral de Componentes Orgânicos	Número de Carbono
Alcanos ou Parafinas	C_nH_{2n+2}	-
Alcadienos	C_nH_{2n-2}	n>=2
Anéis benzênicos	C_nH_{2n-6}	n>=6
Olefinas	C_nH_{2n}	n>=2
Álcool ligações simples	$C_nH_{2n+2}O$	-
Álcool uma dupla ligação	$C_nH_{2n}O$	-
Ácidos Carboxílicos com Ligações simples, acíclicos	$C_nH_{2n}O_2$	-
Ácidos Carboxílicos com uma dupla ligação, acíclicos	$C_nH_{2n-2}O_2$	n>=3
Cetonas com Ligações simples, acíclicos	$C_nH_{2n}O$	n>=3
Éter com ligações simples, acíclico	$C_nH_{2n+2}O$	n>=2
Éter uma dupla ligação acíclico	$C_nH_{2n}O$	n>=3
Éster com ligações simples	$C_nH_{2n}O_2$	n>=2
Éster com uma ligação dupla	$C_nH_{2n-2}O_2$	n>=4
Amina ligações simples	$C_nH_{2n+3}N$	-
Amina com uma dupla ligação	$C_nH_{2n+1}N$	n>=2
Amidas ligações simples	$C_nH_{2n+1}ON$	n>=3
Mercaptanas	$C_nH_{2n+2}S$	-
Dissulfetos de Carbono	$C_nH_{2n+2}S_2$	n>=2
Benzotiofeno	$C_nH_{2n-10}S$	-
Di-Benzotiofenos	$C_nH_{2n-16}S$	-
Polissulfetos	$C_nH_{2n+2}S_3$	-

3.3. Desenvolvimento de sistema de busca

O sistema de busca, como ilustrado na Figura 7, foi gerado com o auxílio de algoritmos em linguagem VBA.

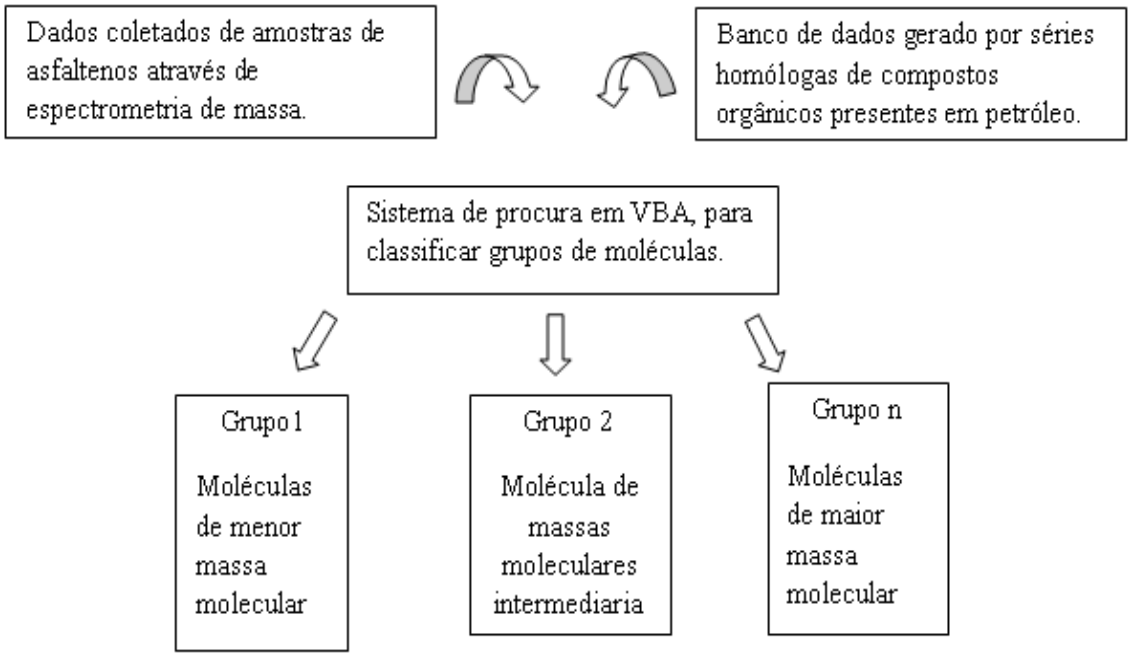
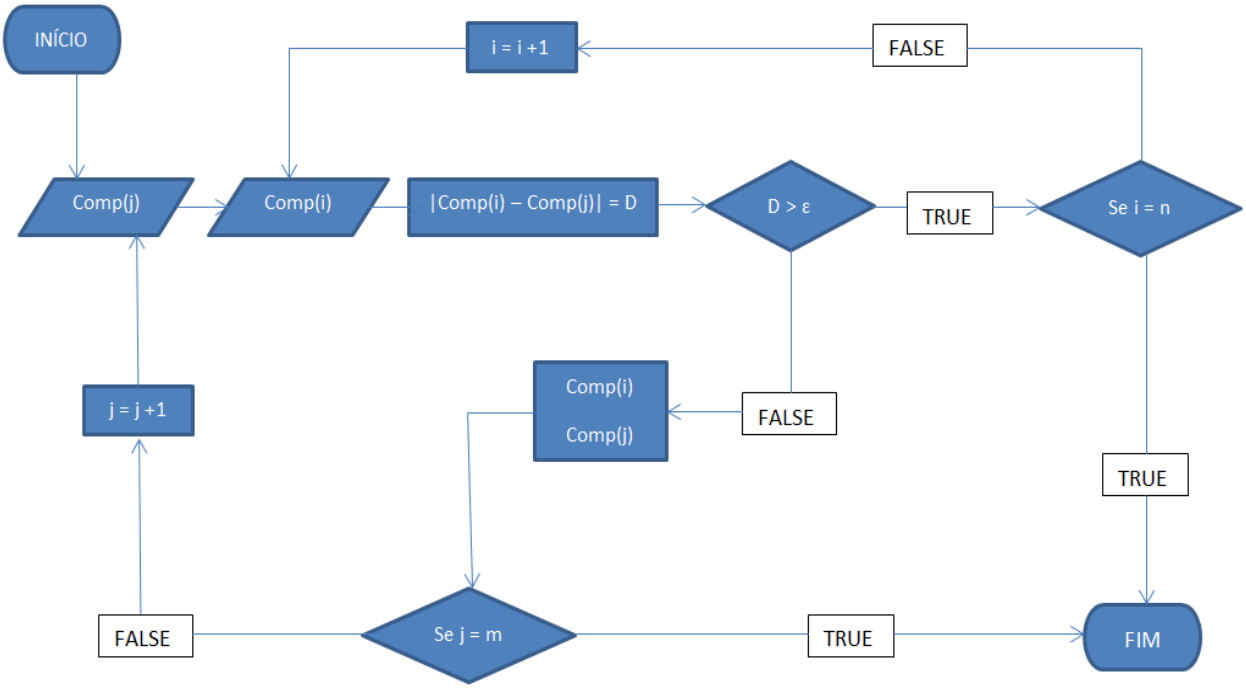


Figura 7 - Esquema ilustrativo para busca de grupos de moléculas.

O funcionamento do sistema de busca ocorre como descrito a seguir: i) ocorre a geração de séries homólogas seguida pela determinação de suas respectivas massas moleculares, conforme descrito anteriormente; ii) comparam-se as massas moleculares provenientes da espectrometria de massa com as massas das séries homólogas obtidas; iii) Cada massa molecular teórica (gerada por série homóloga) será comparada com as obtidas das análises experimentais e os grupos são formados quando a diferença entre as massas for menor ou igual a um limite pré-determinado, normalmente o limite de aferição do equipamento.

Para desenvolver o cálculo das massas moleculares das séries homólogas foi utilizado o Microsoft Excel, construindo-se o banco de dados e posteriormente armazenando os dados. A rotina computacional foi desenvolvida em linguagem VBA (*Visual Basic for Applications*), gerando as macros. As macros foram utilizadas para processar os dados das planilhas.

No Fluxograma 1 é apresentado o funcionamento da rotina em VBA para a identificação das moléculas de uma dada amostra.



Fluxograma 1 - Fluxograma da rotina em VBA.

A rotina compara os componentes j de uma dada amostra um a um com os componentes i do banco de dados, tendo i=1 até n e j=1 até m.

Se |comp(i)-comp(j)|<=épsilon, o componente comp(j) é classificado como pertence a um determinado grupo, deste modo, verifica-se o próximo componente em j que é comp(j+1). A ideia principal é que cada i percorra vários j.

Se |comp(i)-comp(j)|>épsilon, então busca-se o próximo componente em i, que é comp(i+1). E assim por diante até que se alcance o composto final m ou n.

O grupo será constituído por todos os elementos de j que atendem ao critério de diferença acima, e será representado pelo elemento de i. Do mesmo modo as propriedades físicas e químicas deste grupo, são atribuídas como sendo as propriedades do componente i, como exemplo a fórmula molecular. A

fórmula molecular do componente i é conhecida, visto que foi gerada juntamente com as massas moleculares, para compor o banco de dados.

O pseudocódigo, na Figura 8, tal qual o fluxograma, descreve em linhas gerais o código em VBA, utilizado para a separação dos grupos.

```
Start ' início do programa
Épsilon = 0.1, i=1, j=1, ' valores iniciais das variáveis.
Do i=1 to n
    Do j=1 to m
        k=1
        diferença = abs(comp_amostra(j) – comp_bancodados(i))
        if  diferença <=epsilon then
        p(k) = comp_amostra(j)
        k=k+1
         j=j+1
         else
         i=i+1
         end if
    next j
next i
end
```

Figura 8 – Pseudocódigo da rotina em VBA

4. Resultados e Discussão

Para a classificação dos componentes das amostras de asfaltenos foi gerado um banco de dados de massas moleculares das séries homólogas conforme demonstrado na metodologia acima. Análises das amostras de asfaltenos foram realizadas em espectrômetro de massa e, como consequência, obtidas suas massas moleculares com precisão satisfatória. Para sistematizar a classificação e separação de grupos via massa molecular foi utilizado um algoritmo de busca desenvolvido em linguagem VBA (*Visual Basic for Applications*). Este programa forma tais grupos com base em compostos que apresentam massas moleculares semelhantes.

As Tabelas 4,5,6 e 7 foram geradas como resultado da simulação da rotina feita em VBA para classificação de componentes moleculares de amostras em relação a um banco de dados gerado por séries homólogas. As tabelas descritas apresentam apenas uma pequena parte dos resultados da simulação devido ao elevado número de dados.

Na primeira coluna tem-se a massa molecular dos componentes da amostra que foram selecionados pelo critério de diferenças de massa.

A segunda coluna apresenta os componentes do banco de dados que classificam os componentes da amostra, ou seja, os componentes da amostra são classificados como pertencentes a um grupo de moléculas cuja massa molecular é de um componente do banco de dados. Em virtude dessa classificação obtém-se um erro porque as massas dos componentes comparados não são iguais. Esse erro é mostrado na quarta coluna.

Para cada peso molecular do banco de dados é associada uma fórmula molecular, como mostrado na terceira coluna. Tal fórmula molecular é a representante dos grupos gerados a partir desta classificação.

A Tabela 4 mostra alguns componentes classificados, como um resultado da simulação realizada, com as massas moleculares da amostra 1, com um critério de diferença de massa épsilon = 0,1. Verificando-se formação de grupos que é dado na segunda coluna com o nome agrupamento.

Tabela 4 - Resultados da classificação da amostra 1 comparada com o banco de dados.

Amostra 1 Identificada (kg/kmol)	Agrupamento (kg/kmol)	Fórmula Molecular	Erro
203,0525	203,1431	C12H17N3	0,0906
203,0855	203,1431	C12H17N3	0,0576
203,1067	203,1431	C12H17N3	0,0364
203,143	203,1431	C12H17N3	0,0001
203,1794	203,1431	C12H17N3	0,0363
204,0934	204,0067	C11H8S2	0,0867
205,0496	205,0894	C15H11N	0,0398
205,1011	205,0894	C15H11N	0,0117
206,0726	206,0038	C10H6SO3	0,0688
207,0805	207,1051	C15H13N	0,0246
207,1169	207,1051	C15H13N	0,0118
209,0597	209,0843	C14H11NO	0,0246
209,0961	209,0843	C14H11NO	0,0118
209,1199	209,0843	C14H11NO	0,0356
209,1324	209,0843	C14H11NO	0,0481
210,0675	210,0317	C13H6O3	0,0358
210,1039	210,0317	C13H6O3	0,0722
211,0059	211,1000	C14H13NO	0,0941
211,0576	211,1000	C14H13NO	0,0424
211,0753	211,1000	C14H13NO	0,0247
211,1118	211,1000	C14H13NO	0,0118
211,1481	211,1000	C14H13NO	0,0481
212,0832	212,0296	C13H8SO	0,0536
212,1196	212,0296	C13H8SO	0,0900
213,0546	213,1156	C14H15NO	0,0610
213,091	213,1156	C14H15NO	0,0246
213,1262	213,1156	C14H15NO	0,0106
213,1274	213,1156	C14H15NO	0,0118
213,1638	213,1156	C14H15NO	0,0482
213,1849	213,1156	C14H15NO	0,0693
215,0339	215,1313	C14H17NO	0,0974
215,0702	215,1313	C14H17NO	0,0611
215,0855	215,1313	C14H17NO	0,0458
215,1055	215,1313	C14H17NO	0,0258
215,1066	215,1313	C14H17NO	0,0247
215,143	215,1313	C14H17NO	0,0117
215,1794	215,1313	C14H17NO	0,0481
215,2005	215,1313	C14H17NO	0,0692
216,0781	216,0067	C12H8S2	0,0714
216,0933	216,0067	C12H8S2	0,0866

A Tabela 5 mostra alguns componentes classificados, como um resultado da simulação realizada, com as massas moleculares da amostra 2, com um critério de diferença de massa épsilon = 0,1. Verificando-se formação de grupos que é dado na segunda coluna com o nome agrupamento.

Tabela 5 - Resultados da classificação da amostra 2 comparada com o banco de dados.

Amostra 2 Identificada (kg/kmol)	Agrupamento (kg/kmol)	Fórmula Molecular	Erro
300,1145	300,0609	C20H12SO	0,0536
300,1509	300,0609	C20H12SO	0,0900
301,0859	301,1469	C21H19NO	0,0610
301,1223	301,1469	C21H19NO	0,0246
301,1461	301,1469	C21H19NO	0,0008
301,1587	301,1469	C21H19NO	0,0118
301,1825	301,1469	C21H19NO	0,0356
301,1951	301,1469	C21H19NO	0,0482
301,2162	301,1469	C21H19NO	0,0693
302,0938	302,0224	C19H10S2	0,0714
302,109	302,0224	C19H10S2	0,0866
302,1124	302,0224	C19H10S2	0,0900
302,1665	302,0765	C20H14SO	0,0900
303,1016	303,1051	C23H13N	0,0035
303,1202	303,1051	C23H13N	0,0151
303,138	303,1051	C23H13N	0,0329
303,1618	303,1051	C23H13N	0,0567
303,1744	303,1051	C23H13N	0,0693
303,1981	303,1051	C23H13N	0,0930
303,2141	303,1626	C21H21NO	0,0515
303,2318	303,1626	C21H21NO	0,0692
304,1246	304,0380	C19H12S2	0,0866
304,128	304,0380	C19H12S2	0,0900
304,1696	304,0922	C20H16SO	0,0774
304,1821	304,0922	C20H16SO	0,0899
305,1173	305,1207	C23H15N	0,0034
305,1199	305,1207	C23H15N	0,0008
305,1325	305,1207	C23H15N	0,0118
305,1536	305,1207	C23H15N	0,0329
305,1748	305,1207	C23H15N	0,0541
305,1775	305,1207	C23H15N	0,0568
305,19	305,1207	C23H15N	0,0693
305,2138	305,1207	C23H15N	0,0931
305,2475	305,1782	C21H23NO	0,0693
306,104	306,0537	C19H14S2	0,0503
306,1403	306,0537	C19H14S2	0,0866

306,1852	306,1045	C23H14O	0,0807
306,1978	306,1045	C23H14O	0,0933
307,1117	307,1364	C23H17N	0,0247
307,1328	307,1364	C23H17N	0,0036
307,1355	307,1364	C23H17N	0,0009

A Tabela 6 mostra alguns componentes classificados, como um resultado da simulação realizada, com as massas moleculares da amostra 3, com um critério de diferença de massa épsilon = 0,1. Verificando-se formação de grupos que é dado na segunda coluna com o nome agrupamento.

Tabela 6 - Resultados da classificação da amostra 3 comparada com o banco de dados.

Amostra 3 Identificada (kg/kmol)	Agrupamento (kg/kmol)	Fórmula Molecular	Erro
207,1169	207,1051	C15H13N	0,0118
209,1324	209,0843	C14H11NO	0,0481
211,1482	211,1000	C14H13NO	0,0482
213,1636	213,1156	C14H15NO	0,0480
215,0854	215,1313	C14H17NO	0,0459
215,1794	215,1313	C14H17NO	0,0481
217,1011	217,0894	C16H11N	0,0117
222,1277	222,0351	C11H10SO3	0,0926
223,1355	223,1000	C15H13NO	0,0355
223,1481	223,1000	C15H13NO	0,0481
225,1636	225,1156	C15H15NO	0,0480
227,1794	227,1313	C15H17NO	0,0481
229,101	229,0894	C17H11N	0,0116
231,1167	231,1051	C17H13N	0,0116
233,1324	233,1207	C17H15N	0,0117
235,1481	235,1000	C16H13NO	0,0481
237,1511	237,1156	C16H15NO	0,0355
237,1638	237,1156	C16H15NO	0,0482
239,0854	239,1313	C16H17NO	0,0459
239,1794	239,1313	C16H17NO	0,0481
243,1168	243,1051	C18H13N	0,0117
244,112	244,0194	C13H8SO3	0,0926
245,1199	245,1207	C18H15N	0,0008
245,1325	245,1207	C18H15N	0,0118
246,1277	246,0351	C13H10SO3	0,0926
247,148	247,1000	C17H13NO	0,0480
249,1512	249,1156	C17H15NO	0,0356
249,1637	249,1156	C17H15NO	0,0481
251,1667	251,1313	C17H17NO	0,0354
251,1794	251,1313	C17H17NO	0,0481

255,1167	255,1051	C19H13N	0,0116
256,1246	256,0380	C15H12S2	0,0866
257,1324	257,1207	C19H15N	0,0117
258,1278	258,0351	C14H10SO3	0,0927
259,1354	259,1000	C18H13NO	0,0354
259,1481	259,1000	C18H13NO	0,0481
260,1434	260,0473	C17H8O3	0,0961
261,1511	261,1156	C18H15NO	0,0355
261,1637	261,1156	C18H15NO	0,0481
262,0810	262,0452	C17H10SO	0,0358

A Tabela 7 mostra alguns componentes classificados, como um resultado da simulação realizada, com as massas moleculares da amostra 4, com um critério de diferença de massa épsilon = 0,1. Verificando-se formação de grupos que é dado na segunda coluna com o nome agrupamento.

Tabela 7 - Resultados da classificação da amostra 4 comparada com o banco de dados.

Amostra 4 Identificada (kg/kmol)	Agrupamento (kg/kmol)	Fórmula Molecular	Erro
207,1170	207,1051	C15H13N	0,0119
208,1121	208,0194	C10H8SO3	0,0927
209,1325	209,0843	C14H11NO	0,0482
213,1636	213,1156	C14H15NO	0,0480
215,0855	215,1313	C14H17NO	0,0458
223,1481	223,1000	C15H13NO	0,0481
227,1795	227,1313	C15H17NO	0,0482
229,1011	229,0894	C17H11N	0,0117
231,1168	231,1051	C17H13N	0,0117
233,1325	233,1207	C17H15N	0,0118
235,1482	235,1000	C16H13NO	0,0482
237,1512	237,1156	C16H15NO	0,0356
237,1639	237,1156	C16H15NO	0,0483
239,0855	239,1313	C16H17NO	0,0458
239,1795	239,1313	C16H17NO	0,0482
243,1169	243,1051	C18H13N	0,0118
244,112	244,0194	C13H8SO3	0,0926
245,1199	245,1207	C18H15N	0,0008
245,1326	245,1207	C18H15N	0,0119
246,1278	246,0351	C13H10SO3	0,0927
247,148	247,1000	C17H13NO	0,0480
249,1512	249,1156	C17H15NO	0,0356
249,1637	249,1156	C17H15NO	0,0481
251,1668	251,1313	C17H17NO	0,0355

251,1794	251,1313	C17H17NO	0,0481
252,0934	252,0067	C15H8S2	0,0867
255,1168	255,1051	C19H13N	0,0117
257,1324	257,1207	C19H15N	0,0117
258,1278	258,0351	C14H10SO3	0,0927
259,1355	259,1000	C18H13NO	0,0355
259,1481	259,1000	C18H13NO	0,0481
260,1196	260,0473	C17H8O3	0,0723
260,1435	260,0473	C17H8O3	0,0962
261,1511	261,1156	C18H15NO	0,0355
261,1637	261,1156	C18H15NO	0,0481
262,0811	262,0452	C17H10SO	0,0359
263,1668	263,1313	C18H17NO	0,0355
263,1794	263,1313	C18H17NO	0,0481
265,1012	265,1469	C18H19NO	0,0457
265,1825	265,1469	C18H19NO	0,0356

Na Tabela 8 tem-se um grupo representado separadamente. Esse grupo foi gerado como resultado da simulação feita com os dados de massa moleculares da Amostra 1, com épsilon=0,1. O grupo gerado é identificado como C12H17N3 e tem massa exata 203,1431 que é proveniente do banco de dados. Os componentes da amostra classificados por esse grupo serão representados pelo C12H17N3. Entretanto para cada determinação tem-se um erro atribuído como mostrado na tabela, para cada molécula.

Tabela 8 – Grupo do C12H17N3 para a Amostra 1.

Amostra1 (kg/kmol)	Referência do Banco de Dados	Erro
203,0161	203,1431 (C12H17N3)	0,1270
203,0525		0,0906
203,0855		0,0576
203,1067		0,0364
203,143		0,0001

A Figura 9 mostra o gráfico de erro que é gerado na determinação do grupo C12H17N3, observando-se que essa classificação apresenta erros variáveis para cada molécula.

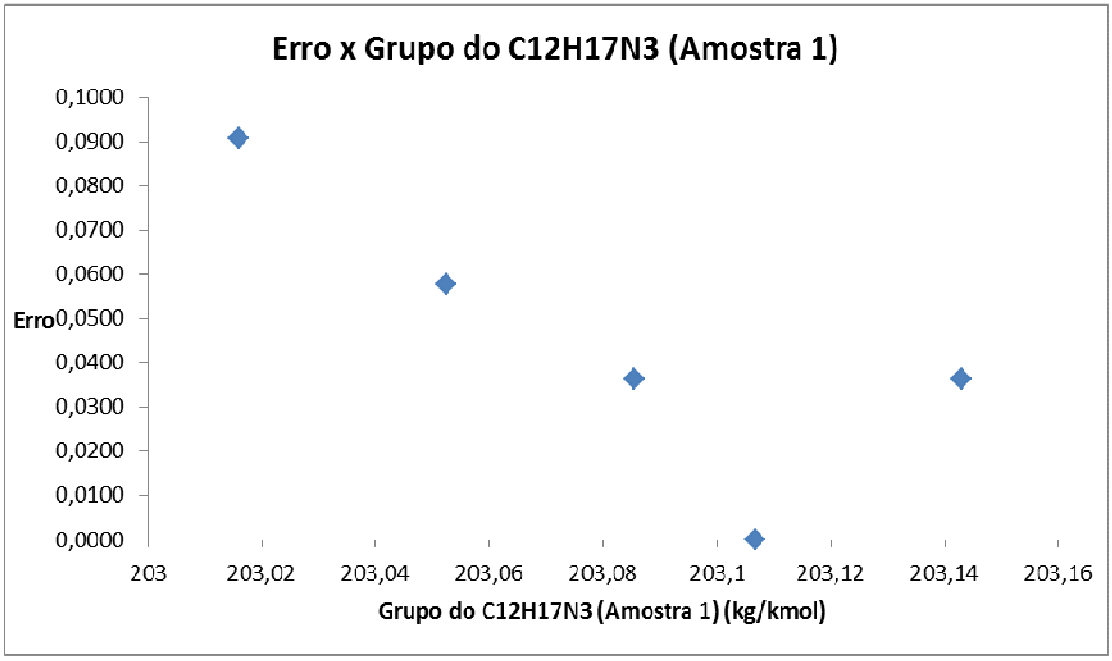


Figura 9 – Gráfico de erro para grupo C12H17N3 da Amostra1.

Na Tabela 9 tem-se um grupo representado separadamente. Esse grupo foi gerado como resultado da simulação feita com os dados de massa moleculares da amostra1, com épsilon=0,1. O grupo gerado é identificado como C14H17NO e tem massa exata 215,1313 que é proveniente do banco de dados. Os componentes da amostra classificados por esse grupo serão representados pelo C14H17NO. Entretanto para cada determinação tem-se um erro atribuído como mostrado na tabela para cada molécula diferente.

Tabela 9 – Grupo do C14H17NO para a Amostra 1.

Amostra1 (kg/kmol)	Referência do Banco de Dados	Erro
212,1196	215,1313 (C14H17NO)	0,0974
212,156		0,0611
213,0546		0,0458
213,091		0,0258
213,1262		0,0247
213,1274		0,0117
213,1638		0,0481
213,1849		0,0692

A Figura 10 mostra o gráfico de erro que é gerado na determinação do grupo C14H17NO, observando-se que essa classificação, apresenta erros variáveis para cada molécula.

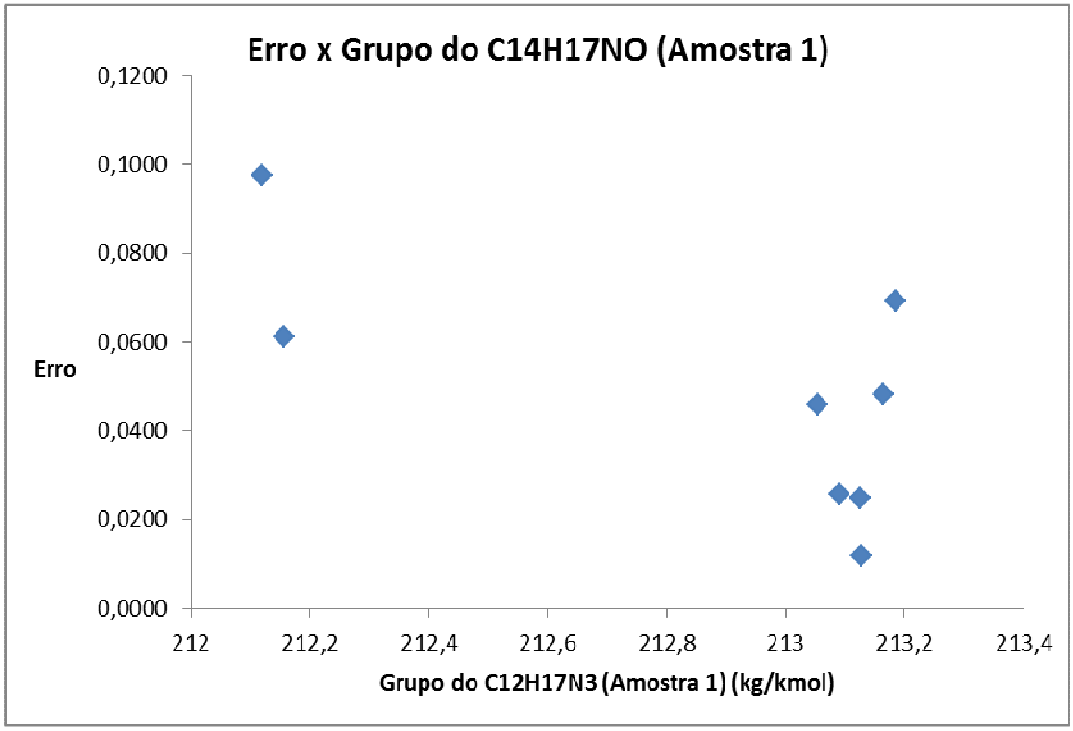


Figura 10 – Gráfico de erro para grupo C14H17NO da Amostra1.

Na Tabela 10 tem-se um grupo representado separadamente. Esse grupo foi gerado como resultado da simulação feita com os dados de massa moleculares da Amostra 2, com épsilon=0,1. O grupo gerado é identificado como C21H13N e tem massa exata 279,1051 que é proveniente do banco de dados. Os componentes da amostra classificados por esse grupo serão representados pelo C21H13N. Entretanto para cada determinação tem-se um erro atribuído como mostrado na tabela para cada molécula diferente.

Tabela 10 – Grupo do C21H13N para a Amostra 2.

Amostra2 (kg/kmol)	Referência do Banco de Dados	Erro
279,0652	279,1051(C21H13N)	0,0399
279,0933		0,0118
279,1016		0,0035
279,1168		0,0117
279,138		0,0329
279,1591		0,0540
279,1743		0,0692
279,1981		0,0930

A Figura 11 mostra o gráfico de erro que é gerado na determinação do grupo C21H13N, observando-se que essa classificação, apresenta erros variáveis para cada molécula.

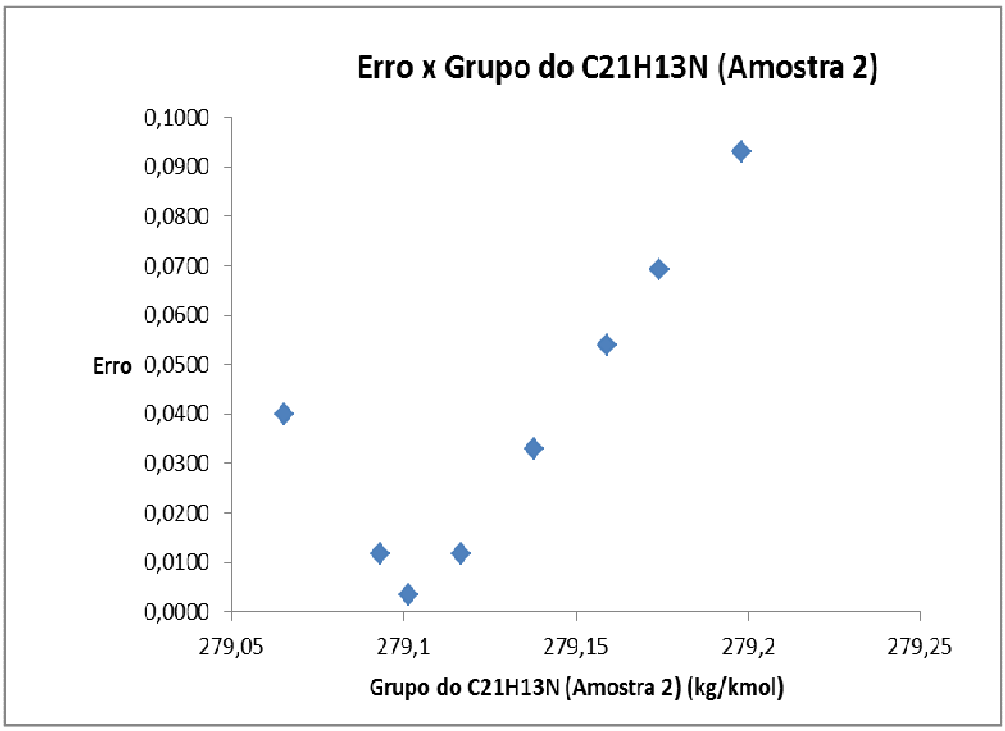


Figura 11 – Gráfico de erro para grupo C21H13N da Amostra2.

Na Tabela 11 tem-se um grupo representado separadamente. Esse grupo foi gerado como resultado da simulação feita com os dados de massa moleculares da Amostra 2, com épsilon=0,1. O grupo gerado é identificado como C20H19NO e tem massa exata 289,1469 que é proveniente do banco de dados. Os componentes da amostra classificados por esse grupo serão

representados pelo C20H19NO. Entretanto para cada determinação tem-se um erro atribuído como mostrado na tabela para cada molécula diferente.

Tabela 11 – Grupo do C20H19NO para a Amostra 2.

Amostra2 (kg/kmol)	Referência do Banco de Dados	Erro
289,0859	289,1469(C20H19NO)	0,0610
289,1012		0,0457
289,1046		0,0423
289,1223		0,0246
289,1587		0,0118
289,1825		0,0356
289,1951		0,0482
289,2162		0,0693

A Figura 12 mostra o gráfico de erro que é gerado na determinação do grupo C20H19NO, observando-se que essa classificação, apresenta erros variáveis para cada molécula.

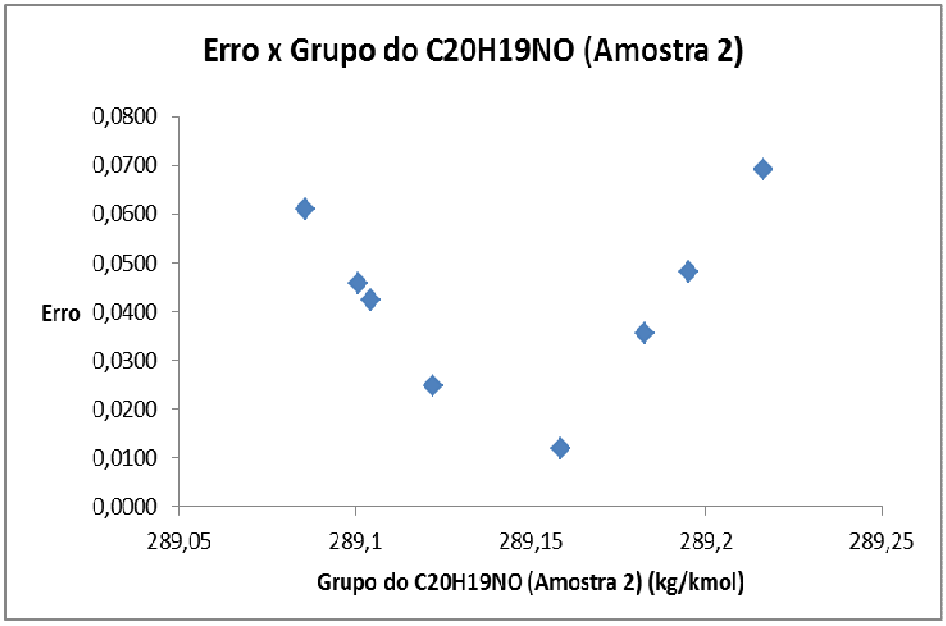


Figura 12 – Gráfico de erro para grupo C20H19NO da Amostra2.

Na Tabela 12 tem-se um grupo representado separadamente. Esse grupo foi gerado como resultado da simulação feita com os dados de massa moleculares da Amostra 2, com épsilon=0,1. O grupo gerado é identificado como C20H16S2 e tem massa exata 320,0693 que é proveniente do banco de

dados. Os componentes da amostra classificados por esse grupo serão representados pelo C20H16S2. Entretanto para cada determinação tem-se um erro atribuído como mostrado na tabela para cada molécula diferente.

Tabela 12 – Grupo do C20H16S2 para a Amostra 3.

Amostra3 (kg/kmol)	Referência do Banco de Dados	Erro
320,1435	320,0693(C20H16 S2)	0,0742
320,1554		0,0861
320,1561		0,0868
320,1594		0,0901

A Figura13 mostra o gráfico de erro que é gerado na determinação do grupo C20H16S2, observando-se que essa classificação, apresenta erros variáveis para cada molécula.

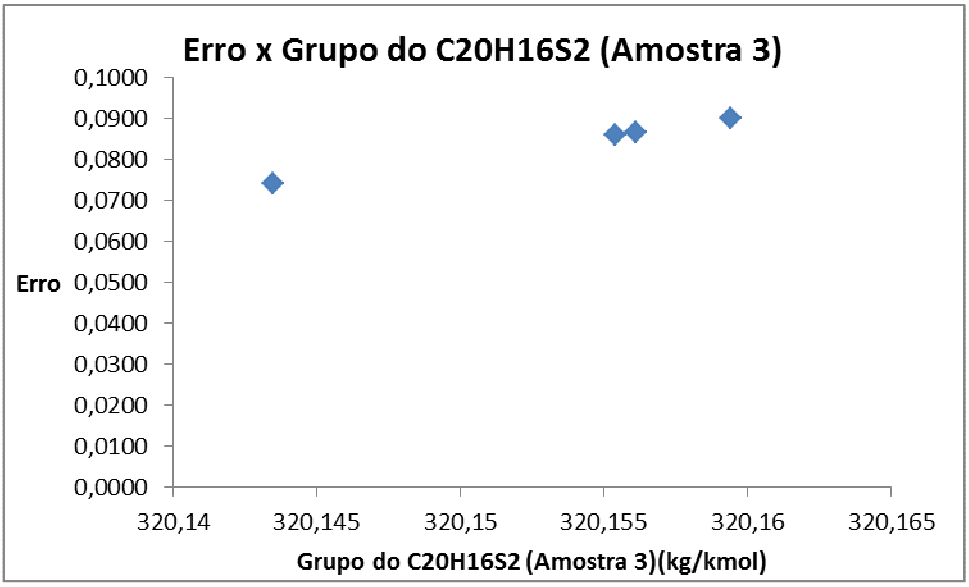


Figura 13 – Gráfico de erro do grupo C20H16S2 da Amostra 3.

Na Tabela 13 um tem-se um grupo representado separadamente. Esse grupo foi gerado como resultado da simulação feita com os dados de massa moleculares da Amostra 3, com épsilon=0,1. O grupo gerado é identificado como C18H19NO e tem massa exata 265,1469 que é proveniente do banco de dados. Os componentes da amostra classificados por esse grupo serão

representados pelo C18H19NO. Entretanto para cada determinação tem-se um erro atribuído como mostrado na tabela para cada molécula diferente.

Tabela 13 – Grupo do C18H19NO para a Amostra 3.

Amostra3 (kg/kmol)	Referência do Banco de Dados	Erro
265,1011		0,0458
265,1824	265,1469(C18H19NO)	0,0355
265,195		0,0481

A Figura14 mostra o gráfico de erro que é gerado na determinação do grupo C18H19NO, observando-se que essa classificação, apresenta erros variáveis para cada molécula.

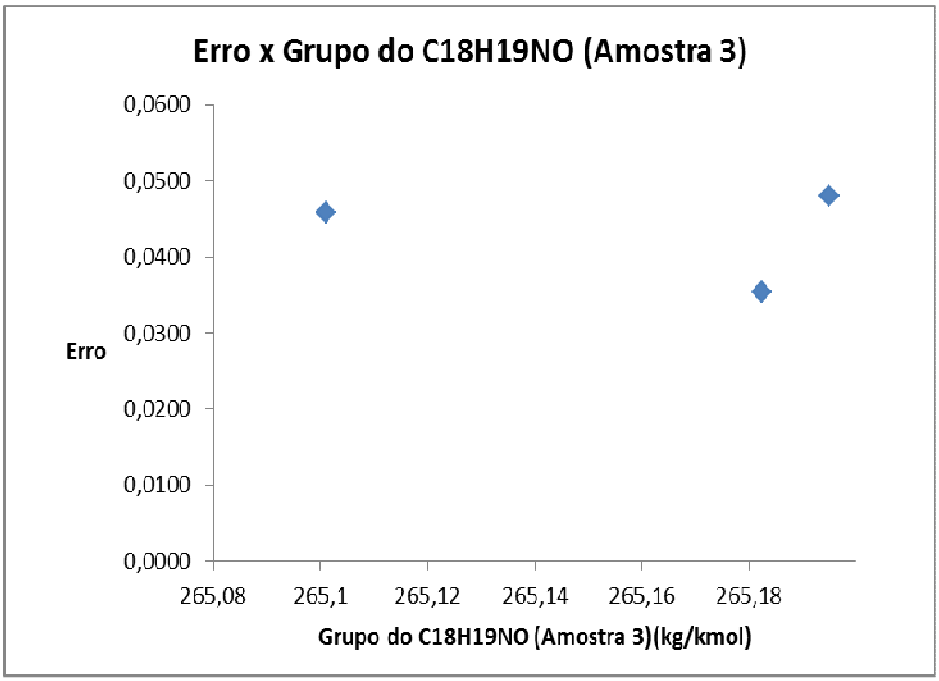


Figura 14 – Gráfico de erro para grupo C18H19NO da Amostra3.

Na Tabela 14 um tem-se um grupo representado separadamente. Esse grupo foi gerado como resultado da simulação feita com os dados de massa moleculares da Amostra 4, com épsilon=0,1. O grupo gerado é identificado como C23H17 e tem massa exata 307,1363 que é proveniente do banco de dados. Os componentes da amostra classificados por esse grupo serão representados pelo C23H17. Entretanto para cada determinação tem-se um erro atribuído como mostrado na tabela para cada molécula diferente.

Tabela 14 – Grupo do C23H17 para a Amostra 4.

Amostra4 (kg/kmol)	Referência do Banco de Dados	Erro
307,1355		0,0009
307,148	307,1363(C23H17)	0,0116
307,2296		0,0932

A Figura 15 mostra o gráfico de erro que é gerado na determinação do grupo C23H17, observando-se que essa classificação, apresenta erros variáveis para cada molécula.

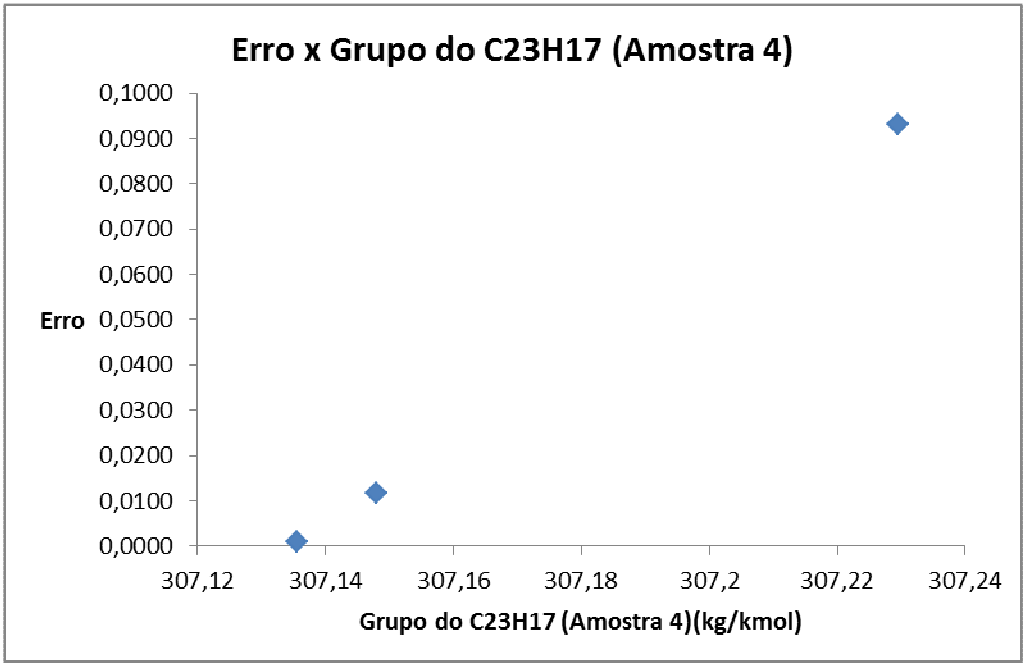


Figura 15 – Gráfico de erro para grupo C23H17 da Amostra 4.

Na Tabela 15 tem-se um grupo representado separadamente. Esse grupo foi gerado como resultado da simulação feita com os dados de massa moleculares da Amostra 4, com épsilon=0,1. O grupo gerado é identificado como C27H19NO e tem massa exata 373,1469 que é proveniente do banco de dados. Os componentes da amostra classificados por esse grupo serão representados pelo C27H19NO. Entretanto para cada determinação tem-se um erro atribuído como mostrado na tabela para cada molécula diferente.

Tabela 15 – Grupo do C27H19NO para a Amostra 4.

Amostra4 (kg/kmol)	Referência do Banco de Dados	Erro
373,1824		0,0355
373,195	373,1469(C27H19NO)	0,0481
373,2401		0,0932

A Figura 16 mostra o gráfico de erro que é gerado na determinação do grupo C27H19NO, observando-se que essa classificação, apresenta erros variáveis para cada molécula.

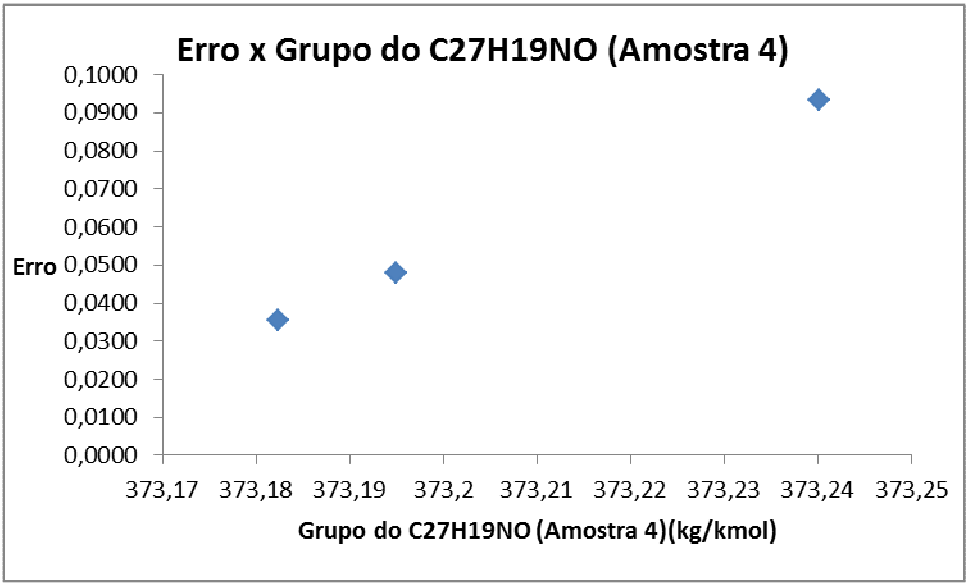


Figura 16 – Gráfico de erro para grupo C27H19NO da Amostra 4.

Para cada molécula classificada existe um erro associado à sua determinação, visto que elas são classificadas por uma diferença de massa. O critério para ser classificado é que essa diferença de massa entre a massa molecular de um constituinte da amostra e um constituinte do banco de dados seja menor ou igual a um limite pré-estabelecido. Tal limite é proveniente do conhecimento da incerteza do equipamento utilizado para as medidas de massa molecular das amostras e o erro é dado pela diferença entre a massa molecular do componente da amostra e do componente do banco de dados.

Na Amostra 1, a faixa dos pesos moleculares vai desde 203,0161 até 1471,5840 totalizando 5623 moléculas para serem comparadas com o mesmo intervalo de pesos moleculares do banco de dados. O banco de dados no intervalo de comparação vai de 203,1313 até 1471,4440, resultando em um

total de 16901 moléculas. Deste modo, têm-se uma relação de 16901 moléculas do banco de dados que podem ser comparadas com as 5623 moléculas da Amostra 1 no mesmo intervalo de massa. Na Amostra 1 havia 5623 moléculas no total, das quais apenas 2594 foram classificadas levando-se em conta um critério pré-estabelecido. Tal critério de classificação estabelece que a diferença de massas moleculares entre um componente da amostra e outro do banco de dados seja menor ou igual a um épsilon, que no caso da presente simulação foi igual a 0,1. Sendo assim, do total de moléculas presentes na Amostra 1, apenas 46,13% foram classificadas.

A Amostra 2, por sua vez, apresentou uma faixa de pesos moleculares de 203,0525 até 663,4537, totalizando 1870 moléculas para serem comparadas com o mesmo intervalo de pesos moleculares do banco de dados. O banco de dados no intervalo de comparação vai de 203,1313 até 663,6685, resultando em um total de 6409 moléculas. Deste modo, têm-se uma relação de 6409 moléculas do banco de dados que podem ser comparadas com as 1870 moléculas da Amostra 2 no mesmo intervalo de massa. Na Amostra 2 havia 1870 moléculas no total, das quais apenas 894 foram classificadas levando-se em conta um critério pré-estabelecido. Tal critério de classificação estabelece que a diferença de massas moleculares entre um componente da amostra e outro do banco de dados seja menor ou igual a um épsilon, que no caso da presente simulação foi igual a 0,1. Sendo assim, do total de moléculas presentes na Amostra 2, apenas 47,81% foram classificadas, conforme apresentado nas Tabelas 16 e 17.

Tabela 16 - Intervalos de Peso Molecular das Amostras.

Amostra	Intervalo de Peso Molecular da Amostra (kg/kmol)	Número de Moléculas da Amostra	Intervalo do Banco de Dados Utilizado para Cada Amostra (kg/kmol)
1	203,0161 - 1471,5840	5623	203,1313 - 1471,4440
2	203,0525 - 663,4537	1870	203,1313 -663,6685
3	207,1169 - 502,2658	1502	207,1051 - 607,5655
4	207,1117 - 601,4655	1911	207,1051 - 601,5589

A Amostra 3, por sua vez, apresentou uma faixa de pesos moleculares de 207,1169 até 502,2658, totalizando 1502 moléculas para serem comparadas com o mesmo intervalo de pesos moleculares do banco de dados.

O banco de dados no intervalo de comparação vai de 207,1051 até 607,5655, resultando em um total de 3561 moléculas. Deste modo, têm-se uma relação de 3561 moléculas do banco de dados que podem ser comparadas com as 1502 moléculas da Amostra 3 no mesmo intervalo de massa. Na Amostra 3 havia 1502 moléculas no total, das quais apenas 583 foram classificadas levando-se em conta um critério pré-estabelecido. Tal critério de classificação estabelece que a diferença de massas moleculares entre um componente da amostra e outro do banco de dados seja menor ou igual a um épsilon, que no caso da presente simulação foi igual a 0,1. Sendo assim, do total de moléculas presentes na Amostra 3, apenas 38,81% foram classificadas.

A Amostra 4, por sua vez, apresentou uma faixa de pesos moleculares de 207,1117 até 601,4655, totalizando 1911 moléculas para serem comparadas com o mesmo intervalo de pesos moleculares do banco de dados. O banco de dados no intervalo de comparação vai de 207,1051 até 601,5589, resultando em um total de 4994 moléculas. Deste modo, têm-se uma relação de 4994 moléculas do banco de dados que podem ser comparadas com as 1911 moléculas da Amostra 4 no mesmo intervalo de massa. Na Amostra 4 havia 1911 moléculas no total, das quais apenas 780 foram classificadas levando-se em conta um critério pré-estabelecido. Tal critério de classificação estabelece que a diferença de massas moleculares entre um componente da amostra e outro do banco de dados seja menor ou igual a um épsilon, que no caso da presente simulação foi igual a 0,1. Sendo assim, do total de moléculas presentes na Amostra 4, apenas 40,82% foram classificadas, conforme apresentado nas Tabelas 16 e 17.

Tabela 17 - Percentagens de Moléculas Classificadas das Amostras.

Amostra	Número de Moléculas da Amostra	Número de Moléculas Classificadas da Amostra	Número de Moléculas do Banco de Dados no Intervalo da Amostra	Percentagem de Moléculas Classificadas
1	5623	2594	16901	46,13%
2	1870	894	6409	47,81%
3	1502	583	3561	38.81%
4	1911	780	4994	40,82%

O percentual de moléculas classificadas foi semelhante para as quatro amostras analisadas. A Amostra 2 apresentou o maior percentual de moléculas classificadas com 47,81%, não muito distante do encontrado na Amostra 1, que foi de 46,13% de classificação.

A Amostra 1 apresentou um intervalo no banco de dados bem mais amplo para comparação das moléculas. Tal amplitude é de aproximadamente três vezes aquela encontrada na Amostra 2. Entretanto, não obteve o maior percentual de moléculas classificadas, mas teve o maior número de moléculas classificadas devido ao fato de ter o maior número de moléculas na amostra.

5. Conclusões

A partir da classificação utilizada neste trabalho com o auxílio de uma rotina computacional desenvolvida em linguagem VBA, foi possível concluir que, para um dado composto pesado (amostras de asfaltenos), as análises apresentaram para cada uma das quatro amostras, respectivamente, 46,13%, 47,81%, 38,81% e 40,82% das moléculas classificadas. Tais percentuais são bastante significativos devido ao fato de representarem compostos moleculares conhecidos tais como compostos nitrogenados, oxigenados, entre outros.

Também foi possível inferir que os erros (diferenças entre as massas moleculares de cada componente da amostra em relação ao banco de dados), para cada grupo se comportam de forma diferente entre si em cada uma das amostras não apresentando nenhum padrão específico. Apesar de os grupos individuais apresentarem, na maioria das vezes, erros com determinadas tendências, não foi possível determinar uma única tendência para cada amostra devido ao fato de que cada composto apresenta uma tendência diferente e também devido à existência de descontinuidades na representação dos grupos.

Por fim, a abordagem de classificação de compostos tendo por base o conceito de séries homólogas desenvolvido neste trabalho mostrou-se coerente, mas existe a necessidade de utilizar simulações com diferenças de massa cada vez menores visando à obtenção de uma única massa molecular para cada elemento de uma série homóloga, devido ao fato de que a cada massa molecular exata corresponde um único elemento de uma série homóloga existente;

A classificação por massa molecular apresenta-se, portanto, como um fator relevante na modelagem molecular, pois consegue separar compostos em determinados grupos nos quais se conhece a fórmula molecular dos compostos presentes

REFERÊNCIAS

ASTM, Annual Book of Standards, Section Five: Petroleum Prooducts, Lubricants and Fossil Fuels, ASTM International, West Conshohoken, PA, USA, 2002

Nigam, A., Klein, M.: A Mechanism-Oriented Lumping Strategy for Heavy Hydrocarbon Pyrolysis: Imposition of Quantitative Structure-Reactivity Relationships for Pure Components. I*nd. Eng. Chem. Res.* 1993,32, 1297-1303.

Quann, R. J., Jaffe, S.B.: Structure-Oriented Lumping: Describing the chemistry of Complex Hydrocarbon Mixtures. I*nd. Eng. Chem. Res.* 1992, 31, 2483-2497.

Hu, S., Shu, F.X.: A General Framework for Incorporating Molecular Modeling Into Overall Refinery Optimisation. *Applied Thermal Engineering*. 2001, 21, 1331-1348.

Hu, S., Towler, G., Zhu, F.X.: Combine Molecular Modeling with Optimization to Stretch Refinery Operation. *Ind. Eng. Chem*. 2002, 41, 825-841.

Gray, M. R.: Lumped Kinetics of Structural Groups; Hydrotreating of Heavy Distillate. *Ind. Eng. Chem. Res*. 1990, 29, 505-512.

Stangeland, B. E.: A Kinetic Model for the Prediction of Hydrocracker Yields. *Ind. Eng. Chem. Process Des. Dev*. 1974, 13, 72.
Parnas, R.S., Allen, D. T.: Compound Class Modeling of Hydropyrolysis. *Chem. Eng. Sci*. 1988, 43, 2845-2857.

Katz, D. L., Brown, G. G.: Vapor Pressure and Vaporization of Petroleum Fractions. *Ind. Eng. Chem*. 1993, 25, 1373-1384.

Watson, B. A., Klein, M. T., Harding, R. H.: Mechanistic Modeling of n-Heptane Cracking on HZSM-5. *Ind. Eng. Chem*. 1996, 35, 1506-1516.

Watson, B. A., Klein, M. T., Harding, R. H.: Mechanistic Modeling of a 1-Phenyloctane/*n*-Hexadecane Mixture on Rare Earth Y Zeolite. *Ind. Eng. Chem*. 1996, 36, 2954-2963.
Watson, B. A., Klein, M. T., Harding, R. H.: Catalytic Cracking of Alkylbenzenes: Modeling the Reaction Pathways and Mechanisms. *Appl. Catal. A: Gen.* 1997, 160, 13-39.

Watson, B. A., Klein, M. T., Harding, R. H.: Catalytic Cracking of Alkylcyclohexanes: Modeling the Reaction Pathways and Mechanisms. *Int. J. Chem. Kinet*. 1997, 29, 545-560.

Russel, C.L., Klein, M. T., Quann, R. J., Trewella, T.: Catalytic Hydrocracking Reaction Pathways, Kinetics, and Mechanisms of n-Alkylbenzenes. *Energy Fuels* 1994, 8, 1394-1400.

Broadbelt, L. J, Stark, S. M., Klein, M. T.: Computer Generated Pyrolysis Modeling: On-the-Fly Generation of Species, Reactions, and Rates. *Ind. Eng. Chem. Res*. 1994, 33, 790-799.

Walter, T. D., Klein, M. T.: A Mechanistic Model of the Pyrolysis Chemistry of 4-(1-Naphthylmethy) bibenzyl as a Probe of Hydrocarbon StructureReactivity Relationships. *Ind. Eng. Chem. Res*. 1995, 34, 4244-4253.

Fake, D. M., Nigam, A., Klein, M. T.: Mechanism-based Lumping of pyrolysis Reactions: Lumping by reactive intermediates. *Appl. Catal. A: Gen.* 1997, 160, 191-221.

Baltanas, M. A., Froment, G. F.: Computer Generation of Reaction Networks and Calculation of Product Distributions in the Hydroisomerization and Hydrocracking of Paraffins on Pt_Containing Bifunctional Catalysts. *Comput. Chem. Eng*. 1985, 9 , 71-81.

Baltanas, M. A., Van Raemdonck, K. K., Froment, G. F., Mohedus, S. R.: Fundamental Kinetic Modeling of Hydroisomerization and Hydrocracking on Noble-Metal-Loaded Faujasites. *Ind. Eng. Chem. Res*. 1989, 28, 899-910.

Clymans, P. J., Froment, G. F.: Computer-Generation of Reaction Paths and Rate Equations in the Thermal Cracking of Normal and Branched Paraffins. *Comput. Chem. Eng*. 1984, 8, 137-142.

Froment, F., Depauw, G. A., Vanrysselberghe, V.: Kinetic Modeling and Reactor Simulation in Hydrodesulfurization of Oil Fractions*. Ind. Eng. Chem. Res*. 1994, 33, 2975-2988.

Feng, W., Vynckier, E., Froment, G. F.: Single event kinetics of catalytic cracking. *Ind. Eng. Chem. Res*. 1993, 32, 2997-3005.

Svoboda, G. D. , Vynckier, E., Debrabandere, B., Froment, G. F.: Single-Event Rate Parameters for Paraffin Hydrocracking on a Pt/US-Y Zeolite. *Ind. Eng. Chem. Res*. 1995, 34, 3793-3800.

Quann, R. J., Jaffe, S. B.: Building useful models of complex reaction systems in petroleum refining. *Chem. Eng. Sci*. 1996, 51, 1615-1635.
Oka, M., Chang, H. C., Gavales, G. R.: Computer-assisted molecular structure construction for coal-derived compounds. *Fuel* 1977, 56, 3-8.

Hu, S., Towler, G., Zhu, X.X.: Combine molecular modeling with optimization to strech refinary operation. *Ind. Eng. Chem. Res.* 2002, 41, 825-841.

Hsu, Chang. Analytical Advances for Hydrocarbon Research, Kluwer Academic/ Plenum Publishers, New York, 2003

Pandoja, P. Caracterização de Petróleo por técnicas Espectroscópicas e sua Utilização na Avaliação e Desenvolvimento de Métodos de Inferência em Processo de Refino. Tese, São Paulo, DEQ-USP, 2010.

Farah, M.A., Caracterização do Petróleo e seus Produtos, Petróleo e combustíveis, 2ª versão, Rio de Janeiro, 2006.

Speight, J.G., The Chemistry and Technology of Petroleum, 3rd edition., Marcel Dekker, New York, 1998.

Denis, J., Briant, J,. e Hipeaux, J.C., Lubricant Properties Analysis and Testing, Editions Technip, Paris, 1997.

Riazi, M.R., Daubert, T.E., Molecular Weight of Heavy Fractions from Viscosity, Oil and Gas Journal, Vol 58, No52, 1987.

Riazi, M.R., Characterization and Properties of Petroleum Fractions, ASTM, 2005.

Kesler, M.G., Lee, B.I., Sandler, S.I., A Third Parameter for Use in generalized Thermodynamic Correlations, Industrial and Engineering Chemistry Fundamentals, Vol 18, 1979.

Goossens, A.G., Prediction of Molecular Weight of Petroleum Fractions, Industrial and Engineering Chemistry Research, Vol 35, 1996.

Jaffe, B.S., Freund, H., Olmstead, W., Extension of Structure-Oriented Lumping to Vacuum Residua, Industrial and Engineering Chemistry Research, Vol 44, 2005.

Thomas, J.E., Fundamentos de Engenharia de Petróleo, 2ª Edição, Editora Interciência, Rio de Janeiro, 2001.

Altgelt, K.H. e Bonduszinski, M.M., Composition and Analysis of Heavy Petroleum Fractions, Marcel Dekker, New York, 1994